

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
9 August 2007 (09.08.2007)

PCT

(10) International Publication Number
WO 2007/089264 A2

(51) International Patent Classification: **Not classified**

(21) International Application Number:
PCT/US2006/022662

(22) International Filing Date: 9 June 2006 (09.06.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/171,154 29 June 2005 (29.06.2005) US

(71) Applicant (for all designated States except US): **THE BOEING COMPANY** [US/US]; 100 N. Riverside Plaza, Chicago, IL 60606 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **CRISPO, Richard, J.** [US/US]; 14406 Se 89th St., Newcastle, WA 98059 (US). **STANLEY, Edward, G.** [US/US]; 16101 Agatewood Road NE, Bainbridge Island, WA 98110 (US). **CHAMBERS, Patrick, J.** [US/US]; 27435 144th Avenue S.E., Kent, WA 98042 (US). **BARR, Dennis, N.** [US/US]; 430 SW 156th Street, #104, Burien, WA 98166 (US). **DIKE, Bruce, A.** [US/US]; 14800 245th Avenue S.E., Issaquah, WA 98027 (US).

(74) Agents: **RAD, Fariba** et al.; TGE BOEING COMPANY, 15460 Laguna Canyon Road, MC 1650-7006, Irvine, CA 92618 (US).

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

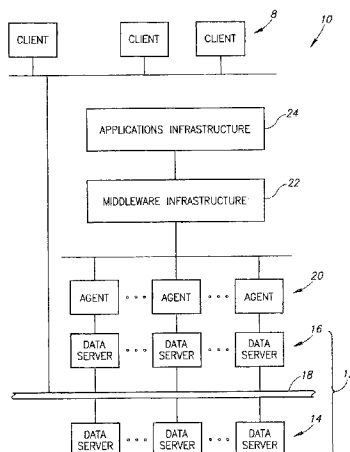
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, 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

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: DISTRIBUTED DATABASE SYSTEMS AND METHODS



(57) Abstract: Systems and methods for distributed database management are disclosed. In an embodiment, a method of accessing information in a distributed database system having one or more data servers includes establishing a plurality of lattice structures configured to store data elements having inter-related information. A data query initiated by a client is then formulated that includes selected data attributes. The data query is processed to generate a query product, and the plurality of lattice structures are searched to identify data elements corresponding to elements in the query product. Pertinent data files located within another data server are accessed that correspond to the data query and generating a result, and the result is transferred to the initiating client system.

WO 2007/089264 A2

DISTRIBUTED DATABASE SYSTEMS AND METHODS

GOVERNMENT LICENSE RIGHTS

This invention was made with Government support under U.S. Government contract DAAE07-03-9-F001 awarded by the U.S. Army and Defense Advanced
5 Research Projects Agency (DARPA). The Government has certain rights in this invention.

FIELD OF THE INVENTION

This invention relates generally to information systems and, more specifically, to systems and methods for data base management and integration.

BACKGROUND OF THE INVENTION

10

15

20

Databases are widely used to store and manage information. Briefly, and in general terms, a database is a collection of related information related to a particular subject that is organized in a useful manner that allows a database user to search for selected information, and to retrieve the information. In one known variation, which
is commonly known as a distributed database, the information storage is distributed throughout a plurality of storage sites that are coupled by a communications network. In many present distributed database systems, the information resides in disparate database systems that exist within distinct geographical and/or corporate boundaries, which presents a challenge to database designers and managers. Advantageously,
each database user should have the ability to search and access selected information within the disparate databases with the same performance and functionality.

25

Referring now to FIGURE 1, a distributed database system 10 according to the prior art is shown. The system 10 includes a plurality of data servers, or “federated” systems 12 that operably maintain information that is exclusively controlled by a
selected one of the data servers 12. For example, one of the data servers 12 may be controlled by a governmental agency, while another one of the data servers 12 may be controlled by a private organization. The data servers 12 may further include a secure

portion 14 and a public portion 16 that are separated by a firewall structure 18. In general, each of the disparate data servers 12 may require different authentication procedures, access control procedures, or require other protocols and/or procedures that are unique to a selected data server 12. Accordingly, agents 20 are positioned
5 between the data servers 12 and a middleware infrastructure 22 to allow information sharing between the plurality of data servers 12. The agents 20 may, for example, include adaptors, brokers or message queues as are understood by those skilled in the art. The middleware infrastructure 22 may manage the transaction between the various data servers 12, and may accelerate client requests by generally reducing the
10 overall number of resource-expensive connections that occur during an information exchange. The middleware infrastructure 22 may also communicate with an applications infrastructure 24 that provides applications access to the various data servers 12.

One drawback associated with the foregoing distributed database system 10 is
15 that the middleware infrastructure 22 and/or the agents 20 are typically difficult and expensive to maintain. Moreover, changes or alterations to the middleware infrastructure 22 and/or the agents 20 may not be properly communicated to each of the data servers 12, so that transactions between various clients 8, or between a selected data servers 12 and the applications infrastructure 24, may occur without
20 notice.

Another drawback associated with the foregoing database system 10 involves data access in general. For example, if a selected one of the clients 8 initiates a search for desired information content that may reside on one of the data servers 12, the search must generally include searching each of the data servers 12 to find the desired
25 information. The foregoing procedure is necessarily time-consuming and may occupy significant resources of the system 10 before the desired information is located.

Still another drawback associated with the database system 10 of FIGURE 1 is that a user may be required to repeatedly enter authentication data as the access request moves throughout the system 10. For example, the user may be required to
30 enter a variety of passwords or authentication codes at various times during the access request.

Although desirable results have been achieved, new systems and methods for distributed database management that may have reduced cost and improved searching efficiency would have utility.

SUMMARY

5 The present invention is directed to systems and methods for distributed database management. Embodiments of systems and methods in accordance with the invention may have reduced cost and improved searching efficiency in comparison with the prior art. In one aspect, a method of accessing information in a distributed database system having one or more data servers includes establishing a plurality of
10 lattice structures configured to store data elements having inter-related information. A data query of a selected data server is then formulated that includes selected data attributes. The data query is processed to generate a query product, and the plurality of lattice structures are searched to identify data elements corresponding to elements in the query product. Pertinent data files located within another data server are
15 accessed that correspond to the data query and generating a result, and the result is transferred to a selected client.

BRIEF DESCRIPTION OF THE DRAWINGS

Preferred and alternate embodiments of the present invention are described in detail below with reference to the following drawings:

20 FIGURE 1 is a schematic block view of a distributed database system according to the prior art;

FIGURE 2 is a schematic block view of a distributed database management system according to an embodiment of the invention;

FIGURE 3 is a schematic block view of the distributed product system of
25 FIGURE 2; and

FIGURE 4 is a flowchart that describes a method of managing a distributed database, according to another embodiment of the invention.

DETAILED DESCRIPTION

The present invention relates to systems and methods for distributed data base management. Many specific details of certain embodiments of the invention are set forth in the following description and in FIGURES 2 through 4 to provide a thorough understanding of such embodiments. One skilled in the art, however, will understand that the present invention may have additional embodiments, or that the present invention may be practiced without one or more of the details described in the following description.

FIGURE 2 is a schematic block view of a distributed database management system 30 according to an embodiment of the invention. The distributed database management system 30 includes one or more data servers 12, as previously shown in connection with FIGURE 1. The data servers 12 are generally federated systems controlled and maintained by an agency or organization that maintains selected informational content. The data servers 12 may be further divided into subsystems that maintain informational content of different types, or into subsystems that maintain predetermined levels of security for information maintained on the subsystem. For example, the data servers 12 may be subdivided into a secure (or restricted) portion 14 and a publicly accessible portion 16 that are separated by one or more firewall structures 18 to maintain the security of information residing on the restricted portion 14. The data servers 12 generally include an input/output capability (not shown) that permits one or more users to communicate with the data servers 12. The input/output capability may include, for example, a computer terminal that is directly coupled to the data servers 12. Alternately, the input/output capability may also include a personal computing device that is coupled to the data servers 12 through a communications system, such as a local area network (LAN), a wide area network (WAN), or even through a Public Internet connection.

The data servers 12 also include various "federated" data files. In the present discussion, "federated" informational content refers to information that resides on, and is under the control of a single one of the data servers 12. The federated information content may include, for example, component part or product specifications, digital

representations of component parts and/or assemblies, text documents related to product assemblies, product and/or parts specifications, planning and/or management information and/or business-related documents, or other still other information.

The client data servers 12 are communicatively coupled to a distributed product system 32. The system 32 provides a centralized repository for selected lattice data structures that may, upon request, be accessed by clients 8. The system 32 also advantageously allows one of the data servers 12 to access information in one or more of the other data servers 12. Accordingly, the system 32 is operable to receive an information request from one or more of the clients 8 and to process the request. If the request is successfully processed by the system 32, the desired information that was the subject of the information request is returned to the requestor. If the system 32 fails to process the information request, the system 32 transmits a failure message to the client 8 that initiated the request.

With reference still to FIGURE 2, the distributed product system 32 includes a memory portion that is configured to store a plurality of lattice structures. The lattice structure is generally an array structure that couples selected data to a corresponding node coordinate value. Accordingly, node values in the lattice structure indicate the position of the selected data. In one particular embodiment of the invention, the distributed product system 32 is configured to store one-dimensional lattice structures. In other embodiments, the system 32 may be configured to store two, three, or "n" dimensional lattice structures. In either case, the lattice structures stored by the system 32 may include uniform, perimetric and curvilinear lattice structures. The distributed product system 32 is also communicatively coupled to an applications infrastructure 24 that permits the clients 8 to access commonly used applications of various types.

FIGURE 3 is a schematic block view of the distributed product system 32 of FIGURE 2 that will be used to describe the system 32 in greater detail. The system 32 includes an information broker 34 that is operable to receive information requests from the plurality of clients 8 and to process the request. In a particular embodiment, the informational request is communicated to the information broker 34 using the well-known extensible markup language common format (XML) that advantageously

allows the sharing of data across different systems, particularly systems connected through the Internet. The information request includes key attributes such as a name, a source for the requested information (if known), a data type, a product number, or other suitable information.

5 The information broker 34 creates a product that includes the foregoing attributes, and instantiates the product. In the present context, "instantiates" is understood to mean finding at least one instance of a desired word or term, or the particular usage of the word or term. The information broker 34 instantiates the product to a database 36 that retains a plurality of lattice structures. The lattice
10 structures include metadata or other similar content that corresponds to the attributes submitted in the information request. The information broker 34 may then generate other data structures that assist the user in navigating and viewing the desired information by selectively filtering the metadata present in the lattice structures. For example, the information broker 34 may generate one or more referential links that
15 permits the desired information to be viewed. In a specific embodiment, the referential links may be hyperlinks. In any event, the navigational information is returned to the requesting client 8. When the information broker 34 instantiates the product to the database 36, a data server 12 that maintains the federated data may also be notified that an inquiry that stems from a client 8 has occurred. Accordingly, the
20 system 12 that maintains the federated data may inform the requesting client 8 when a change to the federated data occurs at some future time.

Still referring to FIGURE 3, the information broker 34 further includes a data cabinet 38 in communication with each of the data servers 12. The cabinets 38 may be used to temporarily store information sent from the information broker 34 prior to
25 transferring the information to the requesting client 8. Alternately, the cabinets 38 may be used to sequester requested information that may be subject to security or proprietary restrictions. For example, a cabinet 38 that includes secure information may await the transfer of an appropriate security key prior to releasing the sequestered data.

30 FIGURE 4 is a flowchart that will be used to describe a method 40 of managing a distributed database, according to another embodiment of the invention.

At block 42, a data query is initiated at a selected one of the clients 8(as shown in FIGURE 2 and FIGURE 3). The data query may include a name, a source for the requested information, a data type, a product number, or other suitable information. The data query may then be converted into an appropriate format before it is
5 transferred to the information broker 34 (as shown in FIGURE 3). For example, the data query may be converted to the extensible markup language (XML), as discussed in detail above. At block 44, the formatted data query is transferred to the information broker 34, and a query product is generated. At block 46, lattice structures stored in the database 36 (also shown in FIGURE 3) are searched relative to the query product.
10 When the requested information is identified in the plurality of lattices stored in the database, one or many referential links may be requested by the information broker 34 allowing the information to be extended to other data servers and returned to the requesting client 8, as shown at block 48. At block 50, the one or more of the data servers 12 that retain the federated data is requested to retrieve and return the desired
15 data to the information broker 34 that will return the result set back to the client 8.

The foregoing methods and systems may advantageously avoid several of the shortcomings present in the prior art. For example, the requirement for extensive data searching that extends into each of the data servers is avoided, since the search extends only to the lattice structures that include various nodes having metadata
20 positioned at the nodes. The metadata thus serve as “placeholders” for the actual information. The foregoing methods and systems also advantageously avoid large, cumbersome middleware infrastructures generally required in distributed database structures. In particular, the requirement for agents, or “adaptor programs” is eliminated.

25 Still other advantages are readily apparent in the foregoing methods and systems. Since the data servers are relatively loosely coupled to the distributed product system, each of the data servers retains the ability to alter internal processes without affecting the operation of other data servers that operate within the distributed product system, thus enhancing partnering agility. General infrastructure benefits are
30 also apparent, since process advantages occur in a collaborative and controlled process. For example, cycle times are shortened, since a search of each data servers is

avoided. Moreover, since changes to metadata in the lattice structures are communicated to all users having a common interest in the information, the need for communicating revision information and data is advantageously avoided. Still further, the foregoing methods and systems advantageously avoid the need for a user to repeatedly enter identifying information, such as passwords and the like, as the user accesses different data servers.

While preferred and alternate embodiments of the invention have been illustrated and described, as noted above, many changes can be made without departing from the spirit and scope of the invention. Accordingly, the scope of the invention is not limited by the disclosure of these preferred and alternate embodiments. Instead, the invention should be determined entirely by reference to the claims that follow.

What is claimed is:

1. A method of accessing information in a distributed database system having one or more data servers, comprising:

establishing a plurality of lattice structures configured to store data elements

5 having inter-related information;

formulating a data query in a selected data server that includes selected data attributes;

processing the data query to generate a query product;

10 searching the plurality of lattice structures to identify data elements corresponding to elements in the query product;

accessing pertinent metadata and data files located within another data server that corresponds to the data query and generating a result; and

transferring the result to the selected client.

15 2. The method of claim 1, wherein establishing a plurality of lattice structures further comprises storing the plurality of lattice structures in a database.

3. The method of claim 1, wherein establishing a plurality of lattice structures further includes forming lattice structures having at least one of a uniform, perimetric and curvilinear lattice structure.

20 4. The method of claim 1, wherein formulating a data query further comprises identifying at least one of a name, a source, a data type and a product number.

5. The method of claim 1, wherein processing the data query further comprises formatting the data query in an extensible markup language (XML) format.

25 6. The method of claim 1, wherein searching the plurality of lattice structures further comprises filtering the lattice structures to identify a desired data element.

7. The method of claim 1, wherein accessing pertinent data files located within another data server further comprises generating a plurality of referential links that permit navigation to the data files.

8. The method of claim 1, wherein transferring the result to the selected client further comprises sequestering at least a portion of the result in a cabinet associated with one of the selected data servers and the another data server.

9. A method of administering information in a distributed database system, comprising:

requesting a data element;

10 selecting a data element in a lattice structure and initiating a message to a federated data repository;

executing the request directly in the federated data file; and

retrieving the data.

10. The method of claim 9, wherein requesting a data element further comprises generating a product that includes attributes of the data request.

11. The method of claim 9, wherein selecting a data element in a lattice structure further comprises filtering the lattice structure to identify the desired data element.

12. The method of claim 9, wherein executing the request directly in the federated data file further comprises generating a referential link that permits navigation to the data file.

13. The method of claim 12, wherein generating a referential link further comprises generating a hyperlink.

14. The method of claim 9, wherein retrieving the data further comprises transferring the data to a selected data server.

15. A distributed database management system, comprising:

a plurality of data servers having federated data sources and having input/output capabilities that are configured to permit communications between a selected data server and a user;

5 a distributed product system in communication with the plurality of client systems, the distributed product system further comprising:

a database that is configured to store a plurality of lattice structures that include data elements corresponding to information stored in the federated data sources; and

10 a information broker coupled to the plurality of data servers and the database, the information broker being operable to accept a data query from a client, search the database according to attributes included in the data query and to extract data from at least one data servers and transfer the extracted data to the initiating client.

16. The system of claim 15, wherein the plurality of data servers further comprises a firewall structure that separates a private portion of a selected client
15 system from a public portion.

17. The system of claim 15, wherein the database is configured to store lattice structures having metadata located at nodes of the lattice structures.

18. The system of claim 15, wherein one or more of the client systems are coupled to a data cabinet operable to store information associated with the data query.

20 19. The system of claim 15, further comprising an applications infrastructure coupled to the system.

20. The system of claim 17, wherein the information broker further comprises a digital filter configured to filter the lattice structures to identify relevant metadata.

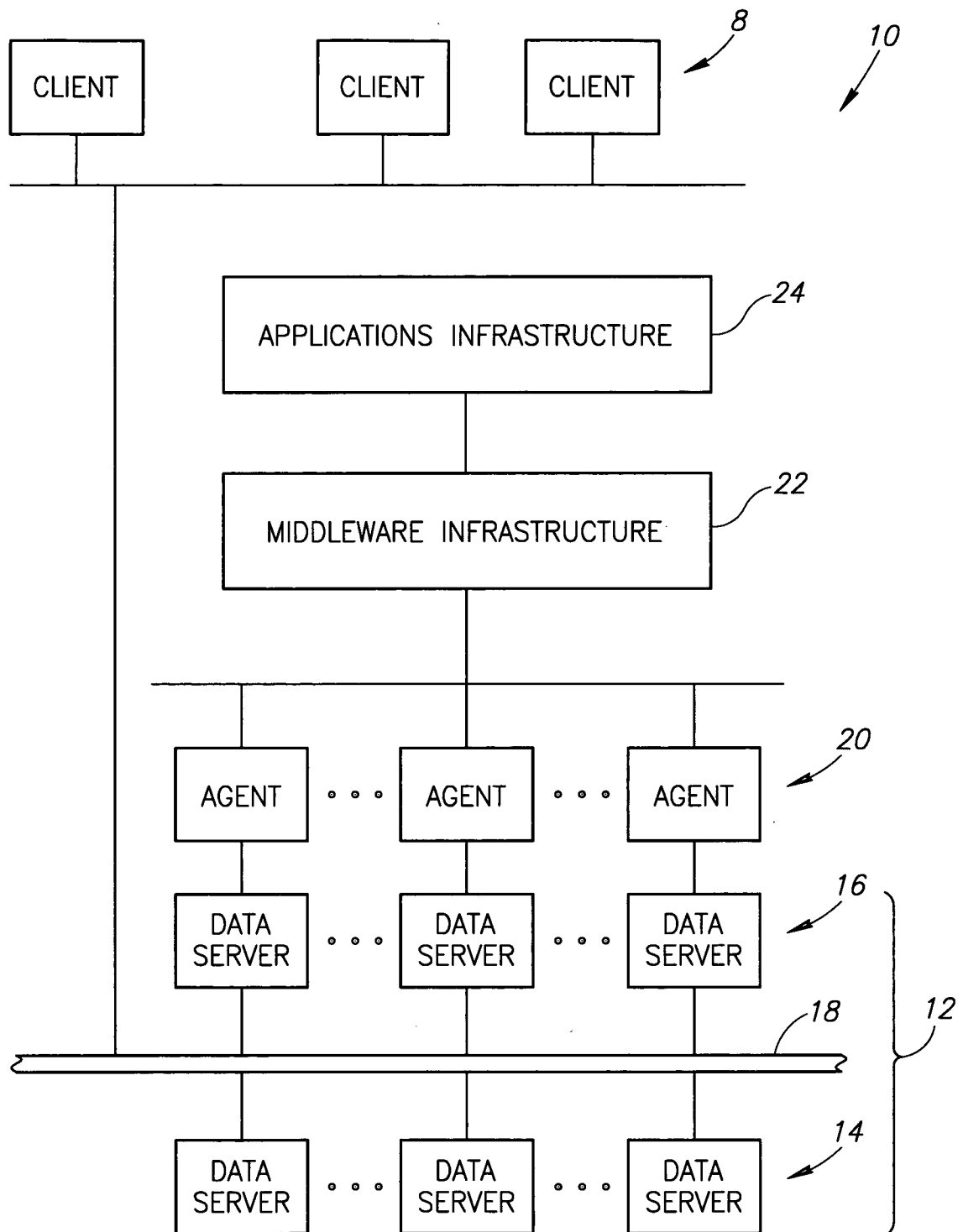


FIG.1
(PRIOR ART)

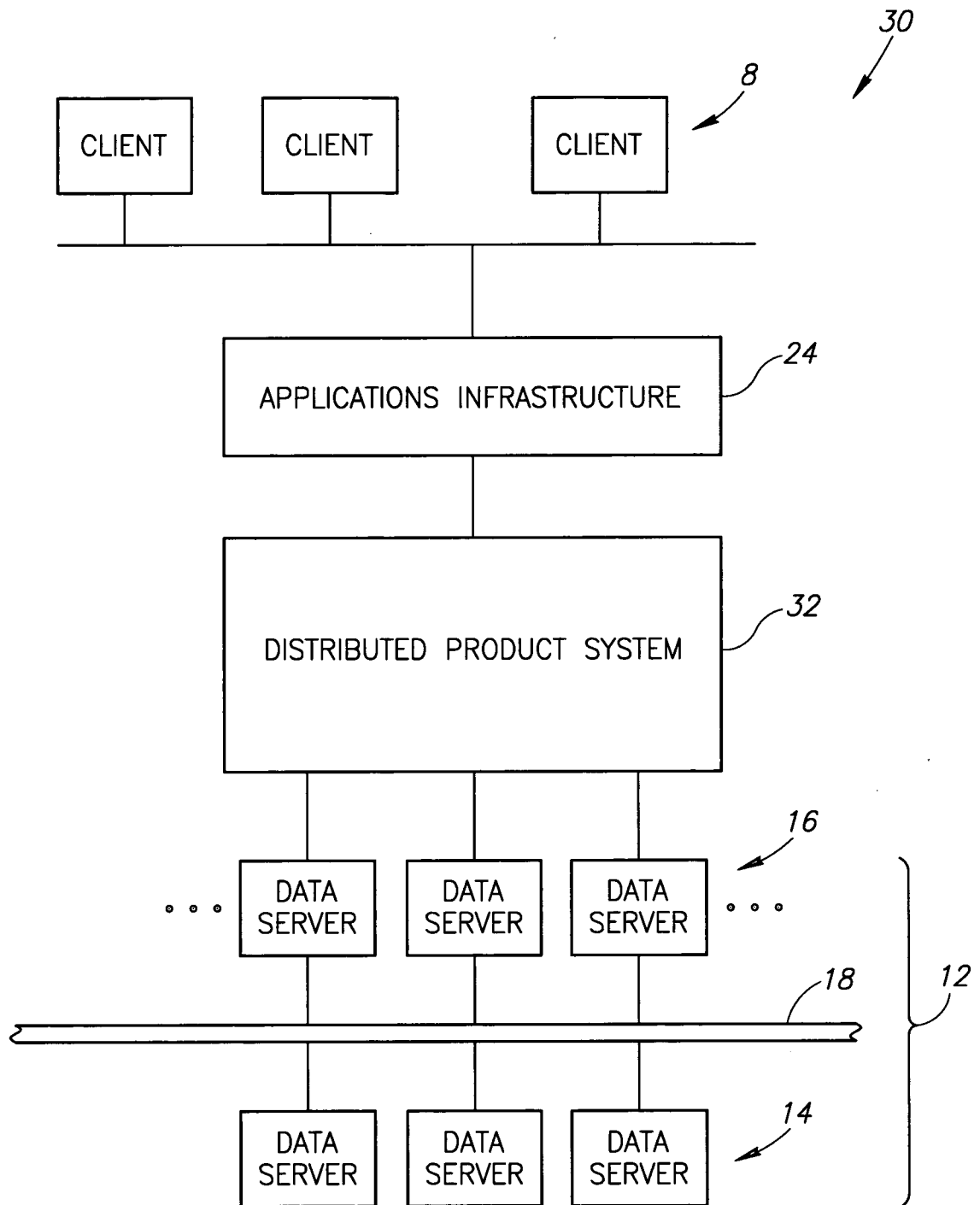


FIG.2

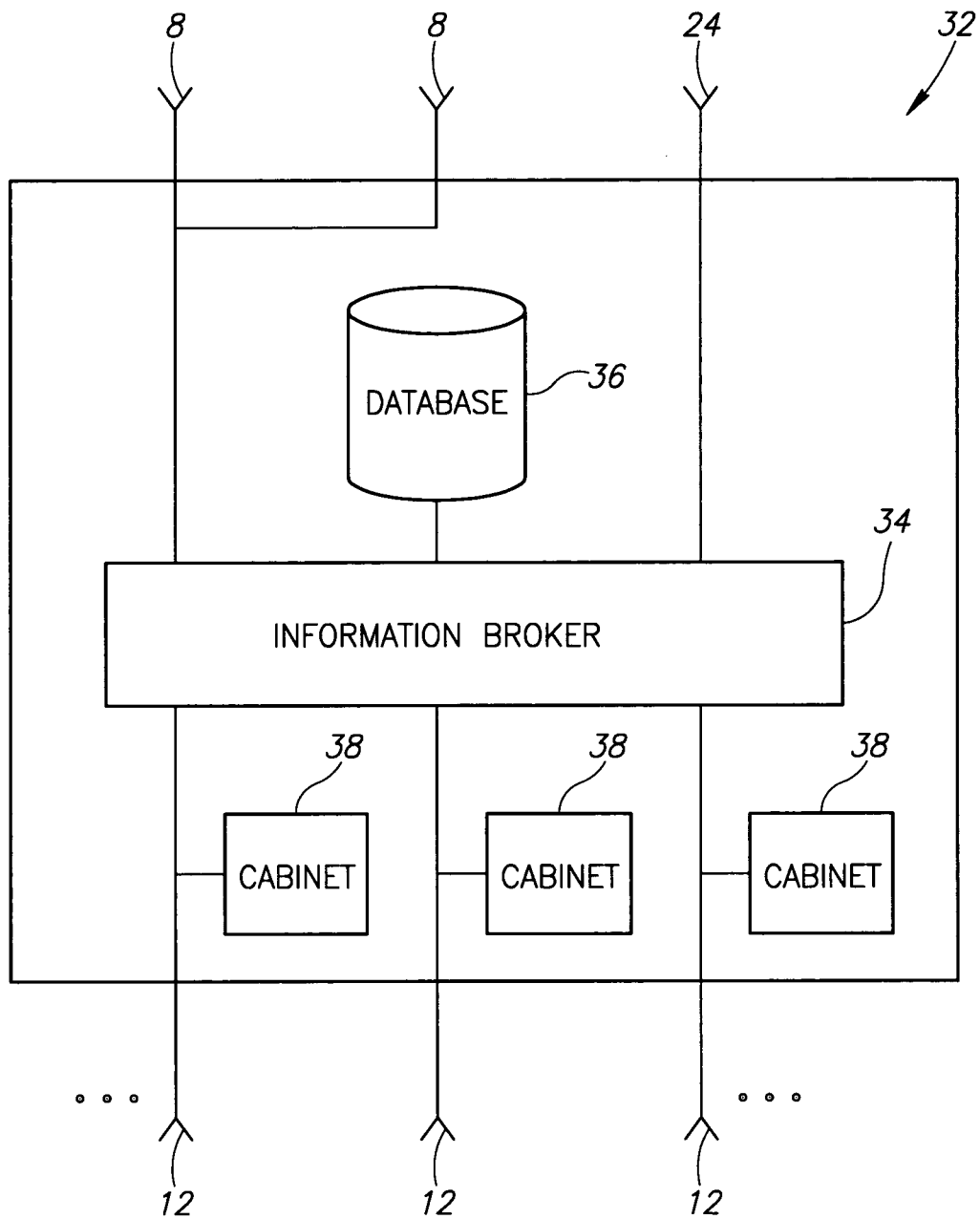


FIG.3

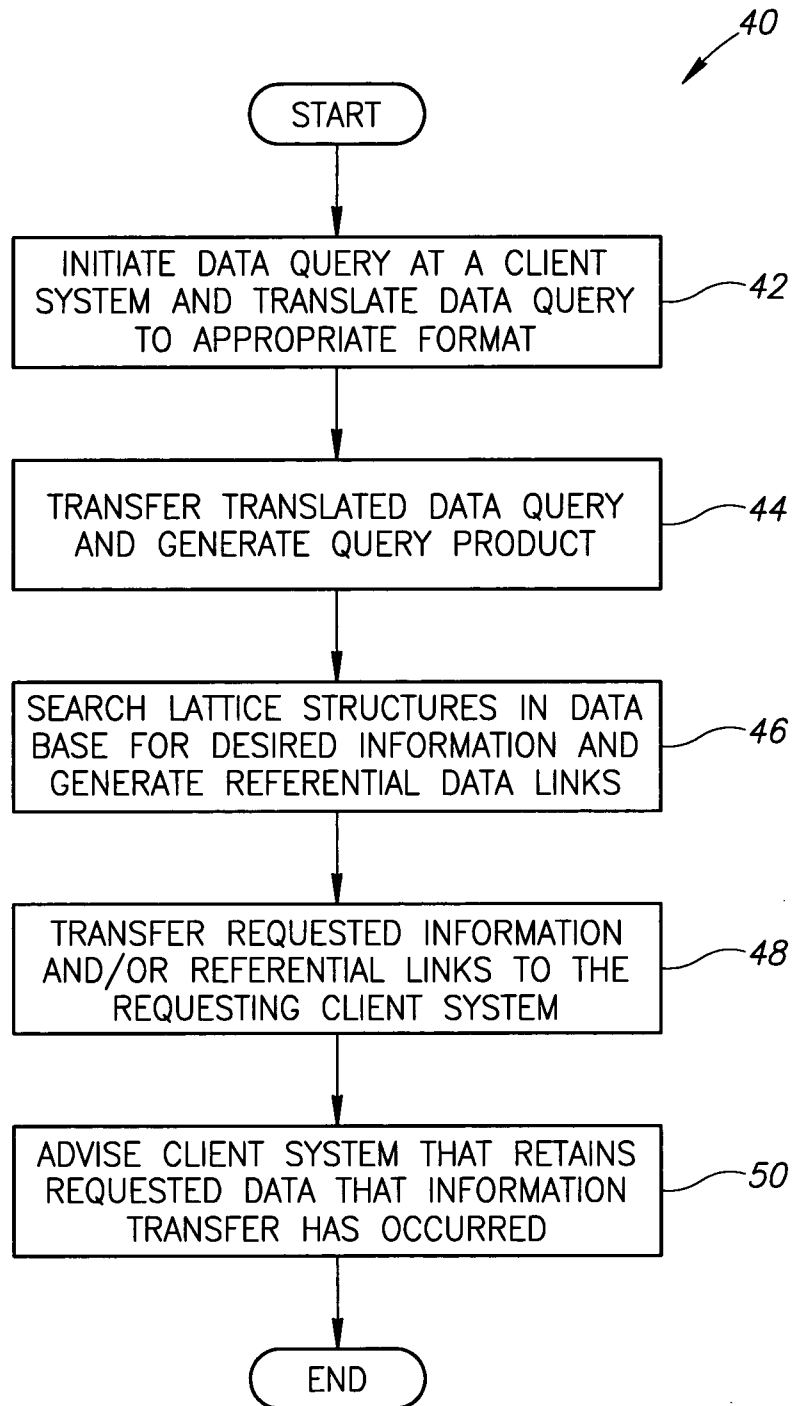


FIG.4