



- (51) **International Patent Classification:**
G06F 17/30 (2006.01)
- (21) **International Application Number:**
PCT/US2012/035762
- (22) **International Filing Date:**
30 April 2012 (30.04.2012)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
13/098,823 2 May 2011 (02.05.2011) US
- (71) **Applicant (for all designated States except US):** **AB INITIO TECHNOLOGY LLC** [US/US]; 201 Spring Street, Lexington, Massachusetts 02421 (US).
- (72) **Inventors; and**
- (75) **Inventors/Applicants (for US only):** **SCHECHTER, Ian** [US/US]; 94 Brook Road, Sharon, Massachusetts 02067 (US). **ALLIN, Glenn John** [US/US]; 132 Mt. Vernon Street, Arlington, MA 02476 (US).
- (74) **Agents:** **HENNESSEY, Gilbert H.** et al.; Fish & Richardson P.C., P.O. Box 1022, Minneapolis, Minnesota 55440-1022 (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, CL, 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, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, 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, RS, SE, SI, SK, SM, 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:** MANAGING DATA QUERIES

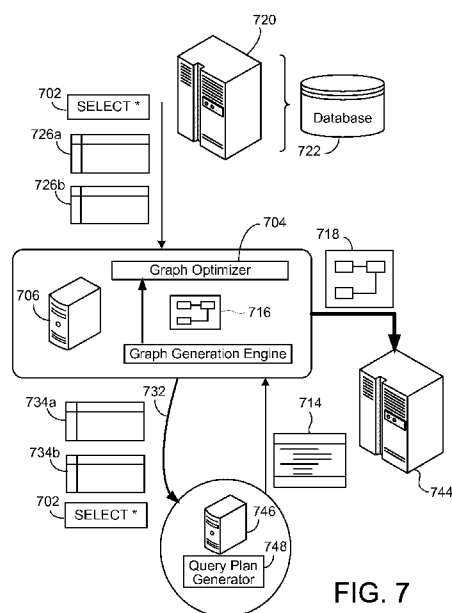


FIG. 7

(57) **Abstract:** In one aspect, in general, a method of generating a dataflow graph (716) representing a database query (702) includes receiving a query plan (714) from a plan generator (748), the query plan representing operations for executing a database query on at least one input (734a, 734b) representing a source of data (726a, 726b), producing a dataflow graph from the query plan, wherein the dataflow graph includes at least one node that represents at least one operation represented by the query plan, and includes at least one link that represents at least one dataflow associated with the query plan, and altering one or more components of the dataflow graph based on at least one characteristic of the at least one input representing the source of data.

MANAGING DATA QUERIES

BACKGROUND

This description relates to managing data queries.

Data can be stored in databases and arranged in various forms such as database
5 tables. A database table can contain a set of data having a common theme or purpose.
The arrangement of a database table can be defined by a database scheme, and multiple
database tables can have similar or identical arrangements. Further, the contents of a
database and its associated database tables can change over time as data is adjusted,
appended or deleted. Various techniques can be used to transfer data into and out of a
10 database and to manipulate the data in the database.

SUMMARY

In one aspect, in general, a method of generating a dataflow graph representing a
database query includes receiving a query plan from a plan generator, the query plan
representing operations for executing a database query on at least one input representing
15 a source of data, producing a dataflow graph from the query plan, wherein the dataflow
graph includes at least one node that represents at least one operation represented by the
query plan, and includes at least one link that represents at least one dataflow associated
with the query plan, and altering one or more components of the dataflow graph based on
at least one characteristic of the at least one input representing the source of data.

20 Aspects can include one or more of the following features. Altering one or more
components of the dataflow graph includes removing at least one component of the
dataflow graph. The component of the graph corresponds to an operation represented by
the query plan. The at least one operation is capable of being performed by executable
functionality associated with the source of data represented by the at least one input. A
25 characteristic of the at least one input includes executable functionality associated with
the source of data represented by the input. The executable functionality includes sorting
functionality. Altering one or more components of the dataflow graph includes merging a
first component representing a first operation applicable to a first input and a second
component representing a second operation applicable to a second input into a single

component, the single component representing a single operation equivalent to the first operation and the second operation, the single operation applicable to the first input and the second input. Altering one or more components of the dataflow graph includes merging a first input component representing a first source of data and a second input
5 component representing a second source of data into a single input component, the single input component including an operation to access data of the first source of data and the second source of data. Altering one or more components of the dataflow graph includes removing a first component representing a first operation applicable to a first input and applying a second component to the first input, the second component representing a
10 second operation corresponding to the same functionality as the first operation, and the second operation applicable to a second input. The method also includes identifying functionality associated with a database associated with a component representing the at least one source of data, and based on the identification, configuring the component to provide a database query to the database. The at least one input representing a dataset
15 includes at least one of a data file, a database table, output of a second dataflow graph, and a network socket. Output of the dataflow graph is assigned to at least one of a data file, a database table, a second dataflow graph, and a network socket. The database query includes an SQL query. The dataflow graph includes a component configured to receive output from the plan generator.

20 In another aspect, in general, a computer-readable medium stores a computer program for generating a dataflow graph representing a database query, and the computer program includes instructions for causing a computer to receive a query plan from a plan generator, the query plan representing operations for executing a database query on at least one input representing a source of data, produce a dataflow graph from the query
25 plan, wherein the dataflow graph includes at least one node that represents at least one operation represented by the query plan, and includes at least one link that represents at least one dataflow associated with the query plan, and alter one or more components of the dataflow graph based on at least one characteristic of the at least one input representing the source of data.

30 In another aspect, in general, a system for generating a dataflow graph representing a database query includes means for receiving a query plan from a plan

generator, the query plan representing operations for executing a database query on at least one input representing a source of data, means for producing a dataflow graph from the query plan, wherein the dataflow graph includes at least one node that represents at least one operation represented by the query plan, and includes at least one link that
5 represents at least one dataflow associated with the query plan, and means for altering one or more components of the dataflow graph based on at least one characteristic of the at least one input representing the source of data.

In another aspect, in general, a system for generating a dataflow graph representing a database query includes a computer system configured to receive a query
10 plan from a plan generator, the query plan representing operations for executing a database query on at least one input representing a source of data, produce a dataflow graph from the query plan, wherein the dataflow graph includes at least one node that represents at least one operation represented by the query plan, and includes at least one link that represents at least one dataflow associated with the query plan, and alter one or
15 more components of the dataflow graph based on at least one characteristic of the at least one input representing the source of data.

In another aspect, in general, a system for generating a dataflow graph representing a database query includes an interface in communication with a plan generator that provides a query plan, the query plan representing operations for executing
20 a database query on at least one input representing a source of data, and a processor configured to produce a dataflow graph from the query plan provided by the plan generator, wherein the dataflow graph includes at least one node that represents at least one operation represented by the query plan, and includes at least one link that represents at least one dataflow associated with the query plan, and alter one or more components of
25 the dataflow graph based on at least one characteristic of the at least one input representing the source of data.

DESCRIPTION OF DRAWINGS

FIG. 1 is a block diagram of a dataflow graph.

FIG. 2 illustrates a database query and a dataflow graph.

FIG. 3 is an overview of a database system and associated components.

FIG. 4 represents the execution of a dataflow graph.

FIG. 5 is a database table and a structural representation of a database table.

FIG. 6 illustrates a database query and a dataflow graph each containing an executable function.

FIG. 7 is an overview of a database system and associated components.

5 FIG. 8 illustrates a query plan and dataflow graphs.

FIG. 9 illustrates a database query user interface.

FIGS. 10-11 illustrate a query plan and dataflow graphs.

FIG. 12 is a flowchart for operations of a database system.

10

DESCRIPTION

1 Databases, Queries, and Graphs

A database management system handles data stored in one or more databases. To store such data, database storage can take one or more forms such as database tables, which can be collections of data organized into data structures such as rows and columns. In one construct, each row represents a record of data, and each column represents a field within each of the rows.

15 The information contained in a database can be accessed and processed (e.g., modified) using database queries. A database query is a set of instructions describing a subset of the database contents and actions to take upon the data in that subset. For example, some database systems perform database queries written in a dedicated database query language such as Structured Query Language (SQL). In these database systems, an SQL query is the primary instrument for manipulating the contents of the database.

In some implementations, database queries and other computations are associated with a database management system in a graphical representation. For example, data may be represented as passing through a collection of operations, referred to as a dataflow. In one arrangement, a dataflow may be provided through a directed graph, with components of the computation being associated with the vertices of the graph and dataflows between the components corresponding to links (arcs, edges) of the graph. A graph is a modular entity and may be connected to or combined with other modular graphs. Each graph can be made up of one or more other graphs, and a particular graph

25

30

can be a component in a larger graph. A graphical development environment (GDE) provides a user interface for specifying executable graphs and defining parameters for the graph components. A system that implements such graph-based computations is described in U.S. Patent 5,966,072, EXECUTING COMPUTATIONS EXPRESSED AS GRAPHS, which is incorporated herein by reference in its entirety.

Referring to FIG. 1, an example of a dataflow graph 101 (as represented in an exemplary user interface) includes an input component 102 providing a collection of data to be processed by the executable components 104a – 104j of the dataflow graph 101. In some examples, the input component 102 is a data source that can include data records associated with a database system or transactions associated with a transaction processing system. The data records may be stored in various forms such as a database table, for example. Each executable component 104a – 104j is associated with a portion of the computation defined by the overall dataflow graph 101. Work elements (e.g., individual data records from the data collection or database table) enter one or more input ports of a component, and output work elements (which are in some cases the input work elements, or processed versions of the input work elements) typically leave one or more output ports of the component. In dataflow graph 101, output work elements from components 104e, 104g, and 104j are stored in output data components 102a – 102c.

The input component 102 and the output data components 102a – 102c, for example, could each provide access to any one of a data file, a database table, output of a second dataflow graph, a network socket, or another source of data.

Some graph-based database systems are used to process database queries. For example, a database query can be applied to one or more database tables to extract an identified subset of the database table contents, for example, for processing in a dataflow. In some implementations, a graph-based database system accepts and executes database queries in the form of dataflow graphs. Other database systems may use other types of database queries.

In some cases, one or more database tables are moved to a graph-based database system from another kind of database system that uses SQL. The other database system may have many SQL queries already written that are incompatible with the graph-based

database system. In some implementations, the SQL queries can be converted to database queries compatible with the graph-based database system.

FIG. 2 shows an example of a database query 200 written in SQL. This example database query 200 is intended to operate on database tables 202, 204 managed by a graph-based database management system. The graph-based database management system can recognize the data in the database tables 202, 204 because the database tables are in a format native to the graph-based database management system. The database tables 202, 204 might contain data obtained from other database tables, for example, database tables originating from a database system that recognizes SQL queries and does not use dataflow graphs.

However, the graph-based database management system may not have built-in functionality for processing SQL queries, so a graph-based database query can be produced to emulate the SQL database query 200. The graph-based database query is recognizable by the graph-based database system. For example, the database query 200 can be converted 206 (using techniques described herein) from an SQL query into a dataflow graph 208. The dataflow graph 208 operates on the database tables 202, 204 by accepting them as input, and provides the execution results of the database query 200 as output.

2 Query Plans

Some database systems carry out database queries such as SQL queries based on a query plan (also sometimes called an explain plan). A query plan is a description of the database operations that may be performed if the database query is executed. The query plan may describe one possible arrangement of database operations, even if other arrangements of the operations or a different set of operations would accomplish the same result.

To provide such query plans, a database system may include a query plan generator (also sometimes called a query planner). For example, the query plan generator can produce a query plan when a database query is being executed, or the query plan generator can generate a query plan before any decision about executing the query is made.

In some arrangements, database operations may be executed in various orders while still providing equivalent outputs. As such, the query plan generator may have functionality that determines an optimal query plan. For example, an optimal query plan could be the query plan that describes the arrangement of database operations for
5 executing the database query in the least amount of time, or using the least amount of database resources such as data storage space, or otherwise accomplishing the database query within constraints that have been identified by the database system. The query plan generator's functionality for determining an optimal query plan may include functionality that scores or ranks many possible query plans, and may also include
10 functionality that rearranges possible query plans to an optimal or efficient configuration.

A single database query can be executed multiple times, and each execution could have a unique optimal query plan. For example, the data within a database table could change between two executions of a database query. In this example, the operations described in the query plan that was generated for the first execution of the database
15 query may need more or less execution time during the second execution of the database query than during the first execution. In this case, a different arrangement of operations may be better suited to the second execution of the database query, for example, a different arrangement of the same operations, or an arrangement of different operations. A query plan optimized for the second execution of the database query can be generated
20 for that execution, taking into account the momentary state of the database table.

3 System Overview

A query plan generator can be used in producing a graph-based database query that emulates another kind of database query such as an SQL query. FIG. 3 shows a database query management system 300 for preparing a database query 302 for execution
25 on a database management computer system 304. The database management computer system 304 shown includes a graph execution engine 306 that handles database operations implemented as dataflow graphs. The database query management system 300 also includes a graph generation computer system 308 having a graph generation engine 310 that can build a dataflow graph 312 from a description of operations to be
30 performed by the dataflow graph. For example, the description of operations could be a

query plan 314. The graph generation computer system 308 is configured with the ability to alter components of the dataflow graph based on one or more characteristics of an input representing a source of data, as described in more detail below in section 6 describing “federated queries.”

5 The database query management system 300 also includes a query planning computer system 316 that executes a query plan generator 318. The query plan generator 318 can be any query plan generator that produces a query plan from a database query, and need not be designed with any functionality related to dataflow graphs or graph generation. Further, the database query management system 300 also includes a database
10 computer system 320 having a database 322 in a data storage (e.g. a hard drive, optical disc, etc.) and containing one or more database tables 324a, 324b, 324c.

 Although separate computer systems are shown for the database management computer system 304, the graph generation computer system 308, the query planning computer system 316, and the database computer system 320, two or more of these
15 computer systems could be the same computer system, or components of the same computer system. All of the computer systems have at least one processor for executing their respective executable components and at least one data storage system. The computer systems can be connected to each other using a computer network such as a local area network (LAN), a wide-area network (WAN), a network such as the Internet,
20 or another kind of computer network.

 To demonstrate the production of a dataflow graph 312 from one or more database queries 302, a database query and a database table 326 are received and processed by the graph generation computer system 308 prior to an execution of the dataflow graph. The graph generation computer system 308 receives the database table
25 326 from the database computer system 320.

 The database table 326 can take any of several forms. For example, the database table 326 could be a relational database table, a partial database table, a flat file, or another kind of data file or collection of data files. In some examples, the database table 326 could be received in the form of information about the database table, e.g. metadata
30 about the database table, or a description of the database table.

In some implementations, the database table 326 could be identified by a data registry associated with the database computer system 320 or otherwise accessible to the graph generation computer system 308. The data registry could be in the form of lookup file catalog, for example, which may contain a data file location associated with the database table 326, and primary key and index information associated with the database table. The data registry could also provide information about the data formats for different types of database tables. Further, the data registry could also provide information about how a dataflow graph 312 can access the database table.

The graph generation computer system 308 also receives a database query 302 to be applied to the database table 326. For example, the database query could be an SQL query. The database query 302 could be received from any number of possible sources. For example, the database query 302 could be received from a user interface 328 where a user 330 has entered the database query. In some examples, the database query 302 is received from a data storage system, or the database query is received from a computer network such as the Internet, or the database query is generated based on another previously-received database query.

In some implementations, the graph generation computer system 308 provides (as represented by an arrow 332) a version of the database table 326 to the query planning computer system 316, produced from information about the database table 326. For example, the version of the database table 326 provided to the query planning computer system 316 could be a structural representation 334 of the database table that is smaller in size than the database table and thus requires fewer computational resources to process. The structural representation 334 of the database table 326 may contain information about the database table, but is absent some or all of the data of the database table. For example, the structural representation 334 of the database table 326 could contain a format that reflects the format of the database table, such as the columns, rows, or fields of the database table. The structural representation 334 of the database table 326 could also contain information about the data, such as data storage sizes of elements in the database table, or the data types of elements in the database table.

The graph generation computer system provides (as represented by the arrow 332) the database query 302 and the structural representation 334 of the database table 326 to

the query planning computer system 316. The query planning computer system 316 executes the query plan generator 318, which produces a query plan optimized for executing the database query 302 over the database table 326. The structural representation 334 of the database table 326 supplies the same information used by the query plan generator 318 as would be supplied by the database table 326 itself, for example, data sizes, data types, and other information about the data contained in the database table. In some implementations, the structural representation 334 of the database table 326 includes an index of data elements that is used to optimize data lookup and retrieval. The query plan generator 318 can use the index to calculate the speed of identifying and retrieving data elements from the indexed database table 326.

The graph generation computer system 308 receives a query plan 314 from the query planning computer system 316. The query plan 314 describes an arrangement of database operations that can be used to execute the database query 302 over the database table 326. For example, the operations in the query plan 314 can correspond to nodes of a dataflow graph 312. The query plan 314 can also include information about the data types used by the operations in the query plan. For example, the operations in the query plan could have parameters, and the data types of the parameters could be described in the query plan.

Once produced, the query plan 314 is provided to the graph generation computer system 308 for dataflow graph 312 production by the graph generation engine 310. The graph generation engine 310 outputs a dataflow graph 312 corresponding to the query plan 314. In some implementations, the dataflow graph 312 has nodes representing operations described in the query plan, and node links representing flows of data between the operations. Because a dataflow graph 312 may be generated for each instance of preparing a database query for execution, the graph generation engine 310 can generate a dataflow graph quickly enough to respond to real-time requests to execute a database query. In some implementations, the graph generation engine 310 can generate a dataflow graph from a query plan in less than one second.

The graph generation computer system 308 provides (represented by an arrow 336) the dataflow graph 312 generated by the graph generation engine 316 to the database management computer system 304. In some implementations, the graph

generation computer system 308 also prepares the database table 326 for use by the database management computer system and provides the prepared database table 338. For example, graph generation computer system 308 might convert the database table 326 from a format used by the graph generation computer system 308 to a format used by the database management computer system 304.

Once provided to the database management computer system 304, the dataflow graph 312 is prepared for execution. As shown in FIG. 4, the database management computer system 304 can execute operations of the dataflow graph 312 and use the database table 326 in order to produce results 402 of the database query. The database management computer system 304 provides the database table 326 to one or more nodes 404a, 404b, 404c of the dataflow graph 312 and executes the dataflow graph using the graph execution engine 306. The graph execution engine 306 performs the operations represented by the nodes 404a, 404b, 404c of the dataflow graph 312, which correspond to database operations for executing the underlying database query. Further, links 408a, 408b, 408c between the nodes represent flows of data between the database operations as the database table is processed. The dataflow graph 312 outputs the results 402 of the database query.

4 Structural Representation

FIG. 5 shows an example of a database table 500 and a structural representation 502 of the database table (containing none of the data from the database table). The database table 500 has columns 504a-504i that may contain similar or different types of data. The database table 500 also has rows 508a-508e each containing a field corresponding to each column. Each field of a row contains a data element 510a, 510b, 510c of the data type of the corresponding column (e.g. character string, integer, floating point number, etc.). Further, each row 508a-508e has an inherent data storage size 512a-512e. For example, the data storage size 512a-512e might be the amount of storage space used by the data elements of the row when the row resides in data storage such as memory.

The structural representation 502 of the database table 500 (produced by the graph generation computer system 308 as shown in FIG. 3) has columns 514a-514i identical to

the database table, including the same data types as the original columns 504a-504i. The structural representation 502 of the database table 500 also has rows 518a-518e corresponding to the rows 508a-508e of the database table. However, the rows 518a-518e do not contain the data elements 510a, 510b, 510c from the database table 500.

5 Each row 518a-518e is associated with a data storage size value 520a-520e. While the data storage size 512a-512e of a row in the database table 500 could be calculated from the data storage sizes of the individual data elements 510a, 510b, 510c, the data storage size value 520a-520e of each row 518a-518e can itself be a data element stored alongside each row in the structural representation 502 of the database table. The total data storage
10 size of the structural representation 502 of the database table 500 may be a small percentage of the total data storage size of the database table, because the fields 522a, 522b, 522c of the structural representation of the database table are absent the data elements 510a, 510b, 510c from the fields of the database table. In some implementations, the structural representation 502 of the database table 500 may retain
15 some data elements from the database table, for example, the structural representation of the database table may retain data elements corresponding to key values 524, 526 or other data elements used in the structure, organization, or format of the database table. In some implementations, the structural representation 502 of the database table 500 may contain an index or another data structure that provides information used in data retrieval. In
20 some implementations, the structural representation 502 of the database table 500 may contain other statistics about the data contained in the database table.

5 Function Calls

As shown in FIG. 6, in some implementations, a custom graph function 602 can be embedded in a database query 604. For example, the custom graph function 602
25 might represent an operation that a dataflow graph 606 can carry out during the execution of the dataflow graph. The custom graph function 602 might not have any functionality within the context of the database query 604, and is placed in the database query to be passed to the generated dataflow graph 606. For example, the custom graph function 602 might be added to a database query 604 to prepare it for conversion to a dataflow graph
30 but replacing or augmenting part of the existing database query. Further, the query plan

generator might not have information about the functionality of the custom graph function 602. The custom graph function 602 may be a function that has no equivalent function in the language of the database query 604. In some implementations, the query plan generator may be aware of the input data type or output data type for the custom graph function 602. When the dataflow graph 606 is generated 608, the custom graph function 602 could act as a node 610 of the dataflow graph.

In some implementations, the custom graph function 602 may be a function used for accessing data in a special or customized format, or the custom graph function may be a function used to execute a regular expression or pattern matching expression, or the custom graph function may be a function implemented by a user, for example, a user of the database management computer system.

6 Federated Queries

In some implementations, a query plan generator (e.g., query plan generator 318 shown in FIG. 3) can operate on a database query that is constructed to operate on input data acquired from multiple data sources and/or multiple types of data sources. For example, the database query could be constructed to operate on multiple database tables, and the database tables may belong to databases of different types and having different internal functionality. Also, the database query could be constructed to operate on data accessible in other data sources such as multiple data files, multiple network sockets, multiple outputs of other dataflow graphs, or any combination of these types of data sources or other types of data sources that can be accessed using a database query. When the graph generation engine generates a dataflow graph from a query plan, the resulting dataflow graph can be optimized for accessing each input representing a data source based on characteristics of the respective data sources, depending on the number and respective types of the data sources.

FIG. 7 shows a graph generation computer system 706 that includes a graph generation engine 908 and a graph optimizer 704. The graph generation computer system 706 receives one or more database tables 726a, 726b. For example, the graph generation computer system 706 may receive the database tables 726a, 726b from one or more database computer systems, e.g., a database computer system 720 having a database 722

in a data storage system (e.g., a hard drive, optical disc, etc.). The graph generation computer system 706 also receives a database query 702 to be applied to the database tables 726a, 726b. For example, the database query 702 could be received from a user interface, or, the database query 702 could be received from a data storage system, or the database query could be received from a computer network such as the Internet, or the database query could be generated based on another previously-received database query.

The graph generation computer system 706 provides (as represented by an arrow 932) versions 734a, 734b of the database tables 726a, 726b to a query planning computer system 746. In some implementations, the versions 734a, 734b of the database tables 726a, 726b are the database tables 726a, 726b themselves. In some examples, the versions 734a, 734b of the database tables 726a, 726b are structural representations of the database tables 726a, 726b. For example, the versions 734a, 734b of the database tables 726a, 726b could be structural representations formatted in the same manner as the structural representation 502 of a database table 500 shown in FIG. 5.

The graph generation computer system 706 provides (as represented by the arrow 932) the database query 702 and versions 734a, 734b of the database tables 726a, 726b to the query planning computer system 746. The query planning computer system 746 executes a query plan generator 748, which produces a query plan 714 for executing the database query 702 over the database tables 726a, 726b. In some implementations, if the versions 734a, 734b of the database tables 726a, 726b are structural representations of the database tables 726a, 726b, the structural representations supply the same information used by the query plan generator 748 as would be supplied by the database tables 726a, 726b themselves.

The graph generation computer system 706 receives the query plan 714 from the query planning computer system 746. Like the query plan 314 shown in FIG. 3, this query plan 714 describes an arrangement of database operations that can be used to execute the database query 702 over the database tables 726a, 726b. The graph generation computer system 706 executes the graph generation engine 908 using as input the query plan 714. The graph generation engine 908 generates a dataflow graph 716 based on the contents of the query plan 714. In some implementations, the dataflow graph

716 has nodes representing operations described in the query plan 714, and node links representing flows of data between the operations.

The dataflow graph 716 can be converted by a graph optimizer 704 to an optimized dataflow graph 718. In some implementations, the dataflow graph 716 can be executed by a computer system (for example, a database management computer system 744) to carry out operations corresponding to operations defined by the query plan 714. Some of the operations carried out by the dataflow graph 716 may be redundant and can be removed or consolidated with other operations. For example, the dataflow graph 716 may include a group of components that can be merged into a single component that performs the same operations as would the group of components. The graph optimizer 704 analyzes the dataflow graph 716 and performs an optimization to alter the dataflow graph, for example, to remove redundant components, merge components, and otherwise reduce the number of components in the dataflow graph 716 to generate the optimized dataflow graph 718. After the optimized dataflow graph 718 is generated, the optimized dataflow graph 718 can be provided to the database management computer system 744 for execution.

In some implementations, the dataflow graph 716 and/or the optimized dataflow graph 718 contains a component that is configured to provide output (e.g., a query plan 714) of the query plan generator 748 to other components of the dataflow graph 716. For example, the dataflow graph 716 or the optimized dataflow graph 718 may receive a database query 702 during execution (“on the fly”) and so a component of the dataflow graph 716 or the optimized dataflow graph 718 can, also during execution, request a query plan 714 corresponding to the database query 702 from the query plan generator 748. The component providing output of the query plan generator 748 can then provide the query plan 714 to other components of the dataflow graph 718.

In some implementations, the graph optimizer 704 may increase the number of components in the dataflow graph 716 when generating the optimized dataflow graph 718. For example, the dataflow graph 716 may operate faster or more efficiently when executed if a complex component is replaced by a group of more efficient components.

In some implementations, the graph generation computer system 706 includes a query plan optimizer that operates on the query plan 714, and the query plan 714 is optimized before being converted to a dataflow graph.

FIG. 8 shows an example of a dataflow graph 802 generated from a query plan 806 and an optimized data flow graph 804. The query plan 806 represents the following SQL query:

```
select count(*) as num_transactions,  
       sum(transaction_amt) as total,  
       account_id  
from federated_transactions  
group by account_id  
order by account_id
```

This SQL query selects records from a database table called federated_transactions and groups and orders them. The query plan 806 describes three operations that can be performed to execute the query. The operations of the query plan 806 are a data selection operation 810, a sort operation 812, and a data summarization operation called a group aggregate operation 814. These operations can be translated to a dataflow graph 802 having components that approximately correspond to the operations. The dataflow graph 802 has a data selection component 820, a sort component 822, a data summarization component called a rollup component 824, as well as an output component 826.

The data selection component 820 represents a source of data. For example, the data selection component 820 could be a database table, output of a second dataflow graph, a network socket, or another source of data.

The output component 826 represents a data destination. For example, the output component 826 could be any one of a data file, a database table, an input to a second dataflow graph, a network socket, or another destination for output data.

The data selection component 820 may represent a source of data that has executable functionality. For example, the data selection component 820 may represent a database table in a database that can execute database queries and returns data based on an executed database query such as the SQL query shown above. The dataflow graph 802 may have access to the functionality associated with the database storing the database

table. For example, the dataflow graph 802 may have a component that can send a request to the database to perform operations associated with the executable functionality.

In some implementations, the SQL query operates on data belonging to the database. For example, the federated_transactions database table represented by the data selection component 820 may belong to a database that can perform the same functions as performed by the sort component 822 and the rollup component 824. An optimized dataflow graph 804 can be generated from the query plan 806 that takes advantage of this built-in functionality. The optimized dataflow graph 804 only has two components, a data selection component 830 that includes the rollup functionality, and an output component 832.

The optimized dataflow graph 804 may have fewer components than the dataflow graph 802 generated directly from the query plan 806, and the optimized dataflow graph 804 may also operate more efficiently. For example, in the example shown, the dataflow graph 802 selects 796,096 records from the federated_transactions database table, which are reduced in number to 3,754 records once the rollup operation is complete. The optimized dataflow graph 804 need not take the step of selecting and processing the full number of 796,096 records, however, because it acquires the already sorted and summarized 3,754 records.

In some implementations, the optimization can be performed by translating a portion of the query plan 806 back into a query appropriate for the data source having the functionality being utilized by the optimized dataflow graph 804. FIG. 9 shows an example of a user interface 940 associated with the data selection component 830 of the optimized dataflow graph 804. The user interface 940 allows the data selection component 830 to be configured with a SQL query 942. When the dataflow graph 804 associated with the data selection component 830 is executed, the SQL query 942 is delivered to the source of data represented by the data selection component 830 and executed by database functionality associated with the data selection component 830. For example, if the data selection component 830 is associated with a database 722 stored by a database computer system 720 (FIG. 7), the SQL query 942 can be delivered to the database computer system 720 for execution. In this way, the data selection component 830 can use the SQL query 942 to access functionality associated with the database 722.

Dataflow graphs representing query plans having multiple sources of data as inputs can be optimized based on characteristics of the multiple inputs. FIG. 10 shows a query plan 1010 undergoing transformation to an optimized dataflow graph 1030. The query plan 1010 represents a federated query and has two database tables 1012, 1014 as input. The query plan can be converted 1016 to a dataflow graph 1020. For example, the graph generation engine 908 shown in FIG. 7 can perform the conversion.

The dataflow graph 1020 includes a data selection component 1022 representing one of the database tables 1012 and a data component 1026 representing the other database table 1014. The data selection component 1022 outputs data to a rollup component 1024 and the data component 1026 outputs data to a second rollup component 1028. The data selection component 1022 has internal functionality that can be used to perform the same operations as the rollup component 1024 performs within the dataflow graph 1020. For example, the data selection component 1022 may represent a database system that has executable functionality and can perform its own operations on the data that it provides. For example, the data selection component 1022 could represent a database table accessible on a database such as the database 722 shown in FIG. 7.

The data component 1026 does not have functionality that can be used to perform the same operations as the rollup component 1028 performs within the dataflow graph 1020. For example, the data component 1026 may be an entity such as a data file that does not include executable functionality.

The dataflow graph 1020 can be converted 1018 (e.g., by a graph optimizer such as the graph optimizer 704 shown in FIG. 7) to an optimized dataflow graph 1030 in which some of the components have been removed based on the internal functionality of the data selection component 1022 and the data component 1026. Because the data selection component 1022 includes rollup functionality, the data selection component 1022 and its associated rollup component 1024 can be merged into a combined data selection component 1032 incorporating the rollup operations otherwise performed by the rollup component 1024. The combined data selection component 1032 can be provided a database query instructing the combined data selection component 1032 to perform the rollup operations before providing data to the optimized dataflow graph 1030. For example, the combined data selection component 1032 can be provided a database query

such as the SQL query 942 shown in FIG. 8. The rollup component 1028 associated with the data component 1026 remains in the optimized dataflow graph 1030 because the data component 1026 does not have internal rollup functionality that is otherwise duplicated by the rollup component 1028.

5 FIG. 11 shows a query plan 1110 undergoing transformation to an optimized dataflow graph 1130. The query plan 1110 represents a federated query and has two database tables 1112, 1114 as input. The query plan can be converted 1116 to a dataflow graph 1120. For example, the graph generation engine 908 shown in FIG. 7 can perform the conversion.

10 The dataflow graph 1120 contains components corresponding to operations of the query plan 1110. This dataflow graph 1120 includes data source components 1122, 1124 representing the database tables 1112, 1114 of the query plan 1110, and also includes a join component 1126 that operates on the data provided by the data source components 1122, 1124 to combine the data of the database tables 1112, 1114 into a single source of
15 data.

 The database tables 1112, 1114 may belong to the same database, e.g., a database 722 shown in FIG. 7. The database 722 may have functionality that can provide the data of the two database tables 1112, 1114 as a single source of data, e.g., a single database table that is the result of a join operation. A graph optimizer, e.g., the graph optimizer
20 704 shown in FIG. 7, can identify the functionality of the database 722 associated with the database tables 1112, 1114 and determine whether the dataflow graph 1120 can be simplified by reducing the number of components in the dataflow graph 1120. The graph optimizer 704 can determine that the database 722 can perform a join operation on the database tables 1112, 1114 and generate an optimized dataflow graph 1130 containing a
25 merged data source component 1132 representing data that is provided by the database 722 when the database 722 performs the join operation upon the two database tables 1112, 1114. The optimized dataflow graph 1130 does not contain the join component 1126 because the functionality represented by the join component 1126 is instead performed by the database 722.

30 FIG. 12 is a flowchart 1200 showing exemplary operations of a graph generation computer system 706 (shown in FIG. 7). In step 1202, the graph generation computer

system receives a query plan, for example, a query plan produced by a query plan generator of a query planning computer system. The query plan represents operations for executing a database query on at least one input representing a source of data. In step 1204, the graph generation computer system uses a graph generation engine to produce a dataflow graph from the query plan based on the operations described in the query plan. The resulting dataflow graph includes at least one node that represents at least one operation represented by the query plan, and includes at least one link that represents at least one dataflow associated with the query plan. In step 1206, the graph generation computer system alters components of the dataflow graph based on a characteristic of the input representing the source of data. The components may be altered to optimize the dataflow graph, for example, to reduce the number of components in the dataflow graph.

The database query managing approach described above can be implemented using software for execution on a computer. For instance, the software forms procedures in one or more computer programs that execute on one or more programmed or programmable computer systems (which may be of various architectures such as distributed, client/server, or grid) each including at least one processor, at least one data storage system (including volatile and non-volatile memory and/or storage elements), at least one input device or port, and at least one output device or port. The software may form one or more modules of a larger program, for example, that provides other services related to the design and configuration of computation graphs. The nodes and elements of the graph can be implemented as data structures stored in a computer readable medium or other organized data conforming to a data model stored in a data repository.

The software may be provided on a storage medium, such as a CD-ROM, readable by a general or special purpose programmable computer or delivered (encoded in a propagated signal) over a communication medium of a network to the computer where it is executed. All of the functions may be performed on a special purpose computer, or using special-purpose hardware, such as coprocessors. The software may be implemented in a distributed manner in which different parts of the computation specified by the software are performed by different computers. Each such computer program is preferably stored on or downloaded to a storage media or device (e.g., solid state memory or media, or magnetic or optical media) readable by a general or special

purpose programmable computer, for configuring and operating the computer when the storage media or device is read by the computer system to perform the procedures described herein. The inventive system may also be considered to be implemented as a computer-readable storage medium, configured with a computer program, where the storage medium so configured causes a computer system to operate in a specific and predefined manner to perform the functions described herein.

A number of embodiments of the invention have been described. Nevertheless, it will be understood that various modifications may be made without departing from the spirit and scope of the invention. For example, some of the steps described above may be order independent, and thus can be performed in an order different from that described.

It is to be understood that the foregoing description is intended to illustrate and not to limit the scope of the invention, which is defined by the scope of the appended claims. For example, a number of the function steps described above may be performed in a different order without substantially affecting overall processing. Other embodiments are within the scope of the following claims.

What is claimed is:

1. A method of generating a dataflow graph representing a database query, the method including:
receiving a query plan from a plan generator, the query plan representing operations for executing a database query on at least one input representing a source of data;
producing a dataflow graph from the query plan, wherein the dataflow graph includes at least one node that represents at least one operation represented by the query plan, and includes at least one link that represents at least one dataflow associated with the query plan; and
altering one or more components of the dataflow graph based on at least one characteristic of the at least one input representing the source of data.
2. The method of claim 1, wherein altering one or more components of the dataflow graph includes removing at least one component of the dataflow graph.
3. The method of claim 2, wherein the component of the graph corresponds to an operation represented by the query plan.
4. The method of claim 2, wherein the at least one operation is capable of being performed by executable functionality associated with the source of data represented by the at least one input.
5. The method of claim 1, wherein a characteristic of the at least one input includes executable functionality associated with the source of data represented by the input.
6. The method of claim 5, wherein the executable functionality includes sorting functionality.

7. The method of claim 1, wherein altering one or more components of the dataflow graph includes merging a first component representing a first operation applicable to a first input and a second component representing a second operation applicable to a second input into a single component, the single component representing a single operation equivalent to the first operation and the second operation, the single operation applicable to the first input and the second input.
8. The method of claim 1, wherein altering one or more components of the dataflow graph includes merging a first input component representing a first source of data and a second input component representing a second source of data into a single input component, the single input component including an operation to access data of the first source of data and the second source of data.
9. The method of claim 1, wherein altering one or more components of the dataflow graph includes removing a first component representing a first operation applicable to a first input and applying a second component to the first input, the second component representing a second operation corresponding to the same functionality as the first operation, and the second operation applicable to a second input.
10. The method of claim 1, further including identifying functionality associated with a database associated with a component representing the at least one source of data, and based on the identification, configuring the component to provide a database query to the database.
11. The method of claim 1, wherein the at least one input representing a dataset includes at least one of a data file, a database table, output of a second dataflow graph, and a network socket.

12. The method of claim 1, wherein output of the dataflow graph is assigned to at least one of a data file, a database table, a second dataflow graph, and a network socket.
13. The method of claim 1, wherein the database query includes an SQL query.
14. The method of claim 1, wherein the dataflow graph includes a component configured to receive output from the plan generator.
15. A computer-readable medium storing a computer program for generating a dataflow graph representing a database query, the computer program including instructions for causing a computer to:
 - receive a query plan from a plan generator, the query plan representing operations for executing a database query on at least one input representing a source of data;
 - produce a dataflow graph from the query plan, wherein the dataflow graph includes at least one node that represents at least one operation represented by the query plan, and includes at least one link that represents at least one dataflow associated with the query plan; and
 - alter one or more components of the dataflow graph based on at least one characteristic of the at least one input representing the source of data.
16. A system for generating a dataflow graph representing a database query, the system including:
 - means for receiving a query plan from a plan generator, the query plan representing operations for executing a database query on at least one input representing a source of data;
 - means for producing a dataflow graph from the query plan, wherein the dataflow graph includes at least one node that represents at least one operation

- represented by the query plan, and includes at least one link that represents at least one dataflow associated with the query plan; and
- means for altering one or more components of the dataflow graph based on at least one characteristic of the at least one input representing the source of data.
17. A system for generating a dataflow graph representing a database query, the system including:
- a computer system configured to receive a query plan from a plan generator, the query plan representing operations for executing a database query on at least one input representing a source of data,
 - produce a dataflow graph from the query plan, wherein the dataflow graph includes at least one node that represents at least one operation represented by the query plan, and includes at least one link that represents at least one dataflow associated with the query plan, and
 - alter one or more components of the dataflow graph based on at least one characteristic of the at least one input representing the source of data.
18. A system for generating a dataflow graph representing a database query, the system including:
- an interface in communication with a plan generator that provides a query plan, the query plan representing operations for executing a database query on at least one input representing a source of data; and
 - a processor configured to produce a dataflow graph from the query plan provided by the plan generator, wherein the dataflow graph includes at least one node that represents at least one operation represented by the query plan, and includes at least one link that represents at least one dataflow associated with the query plan, and

alter one or more components of the dataflow graph based on at least one characteristic of the at least one input representing the source of data.

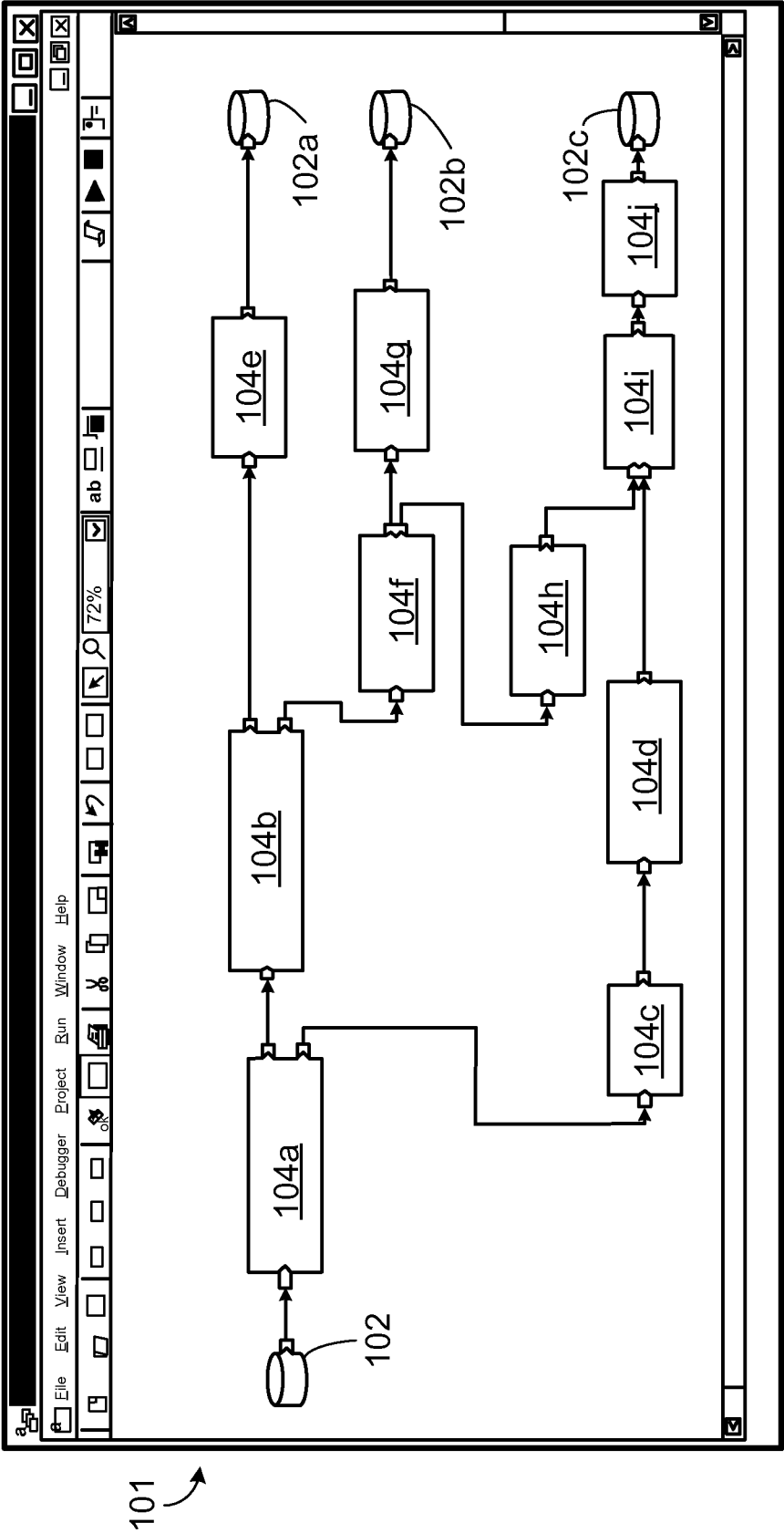
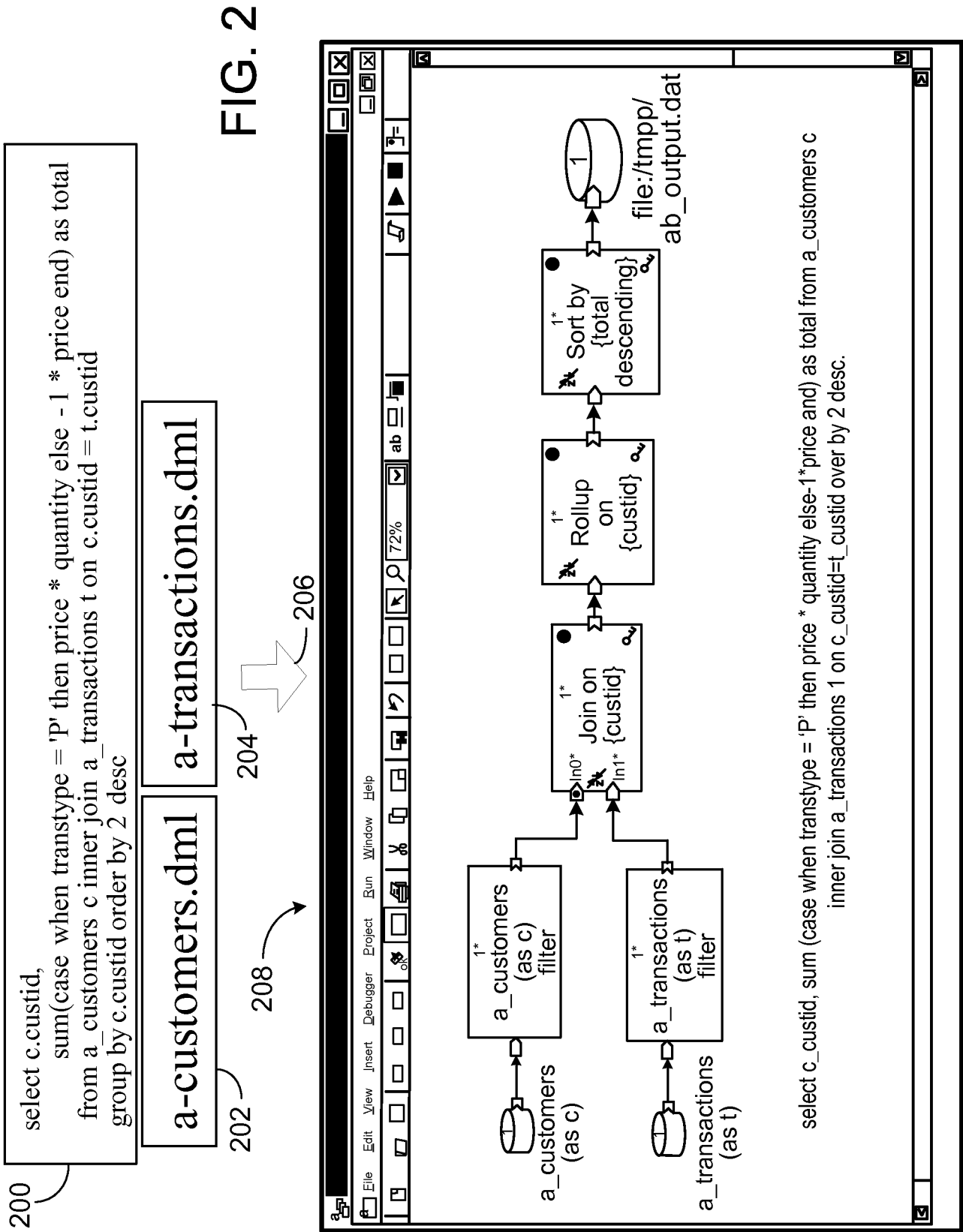


FIG. 1



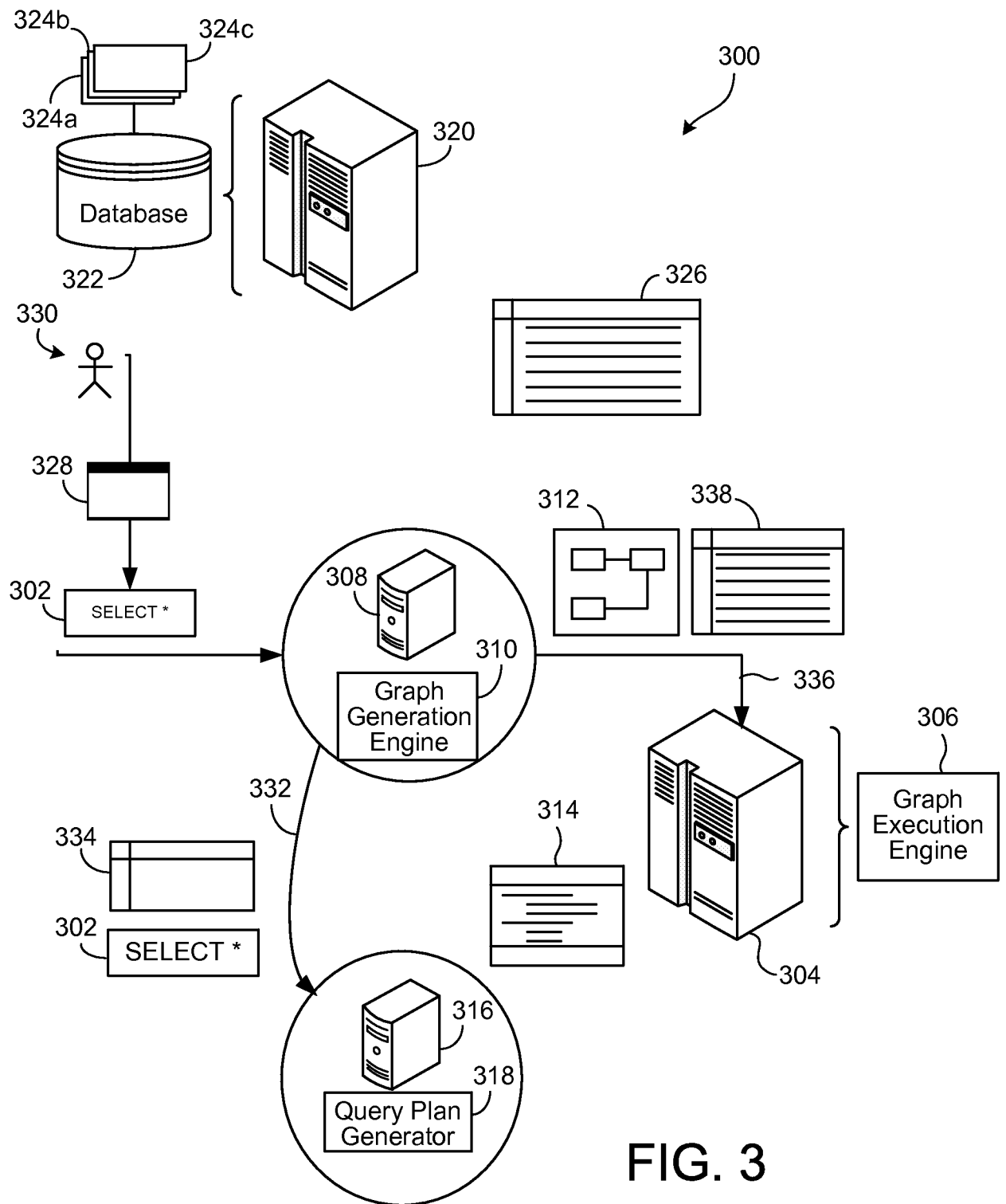
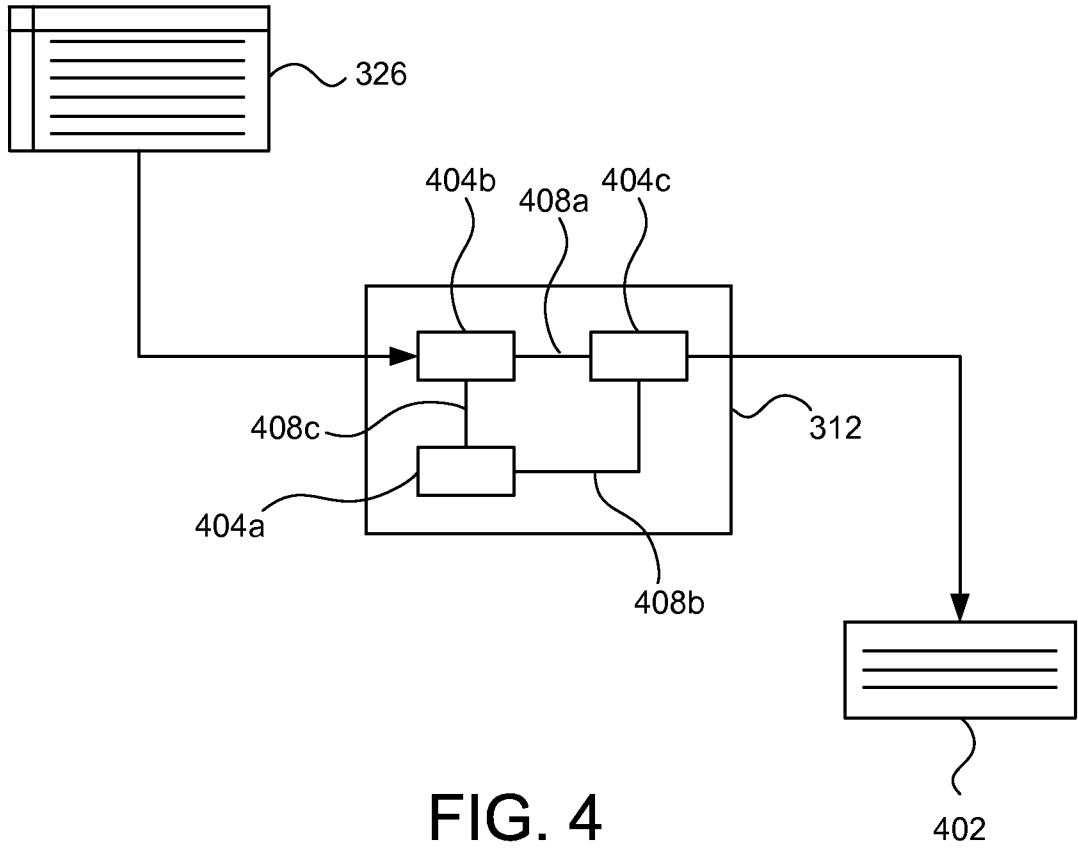


FIG. 3



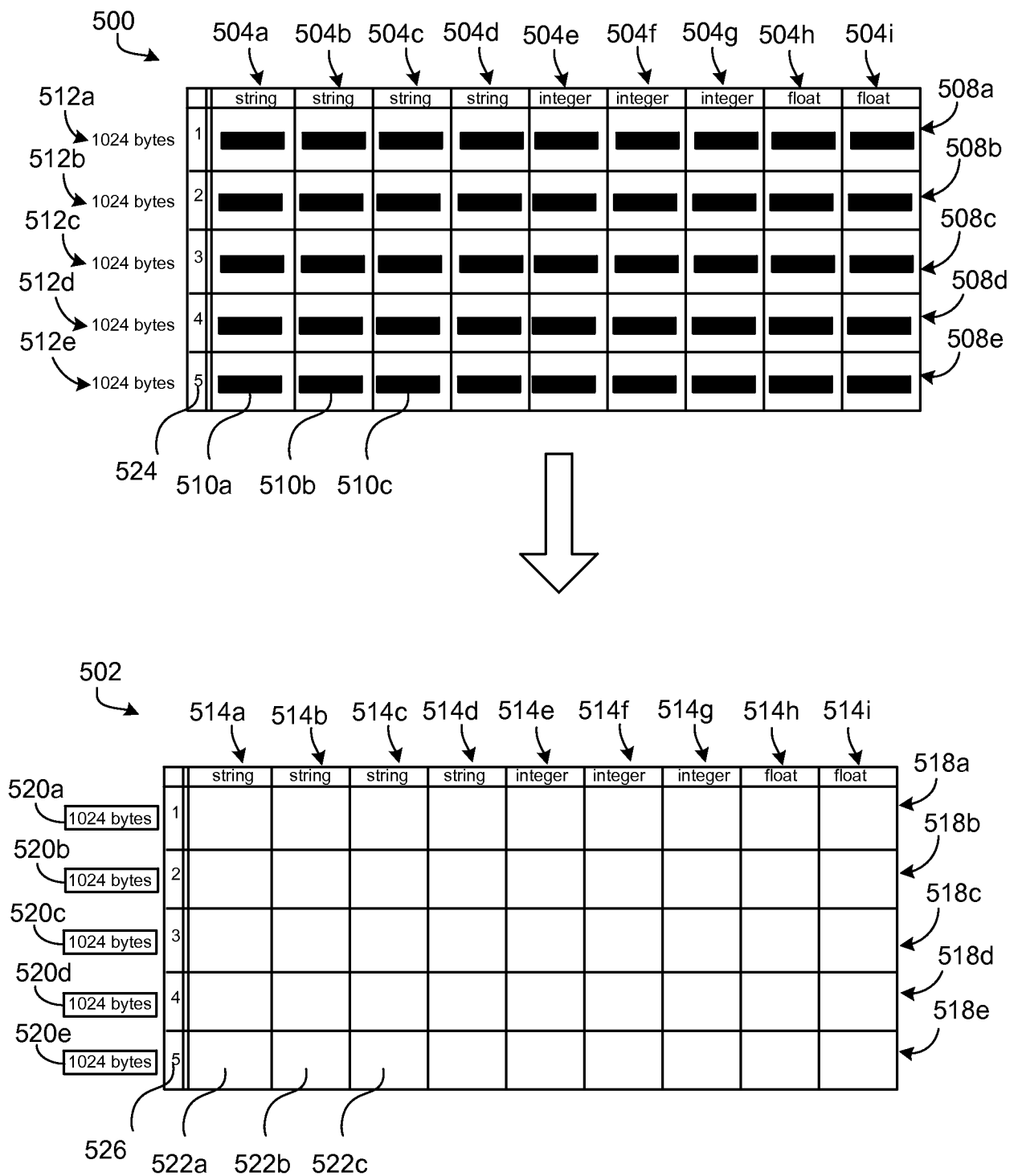


FIG. 5

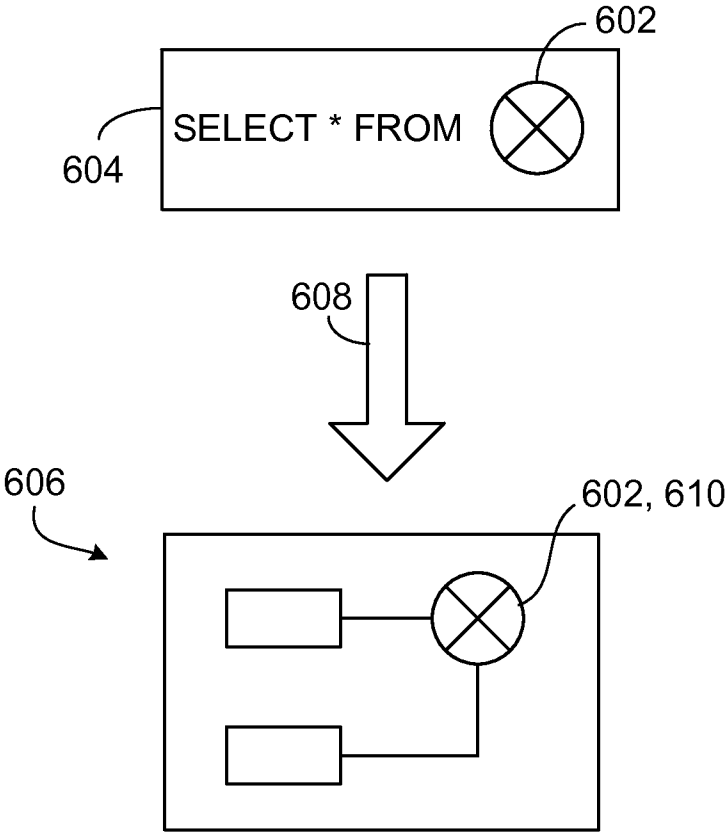


FIG. 6

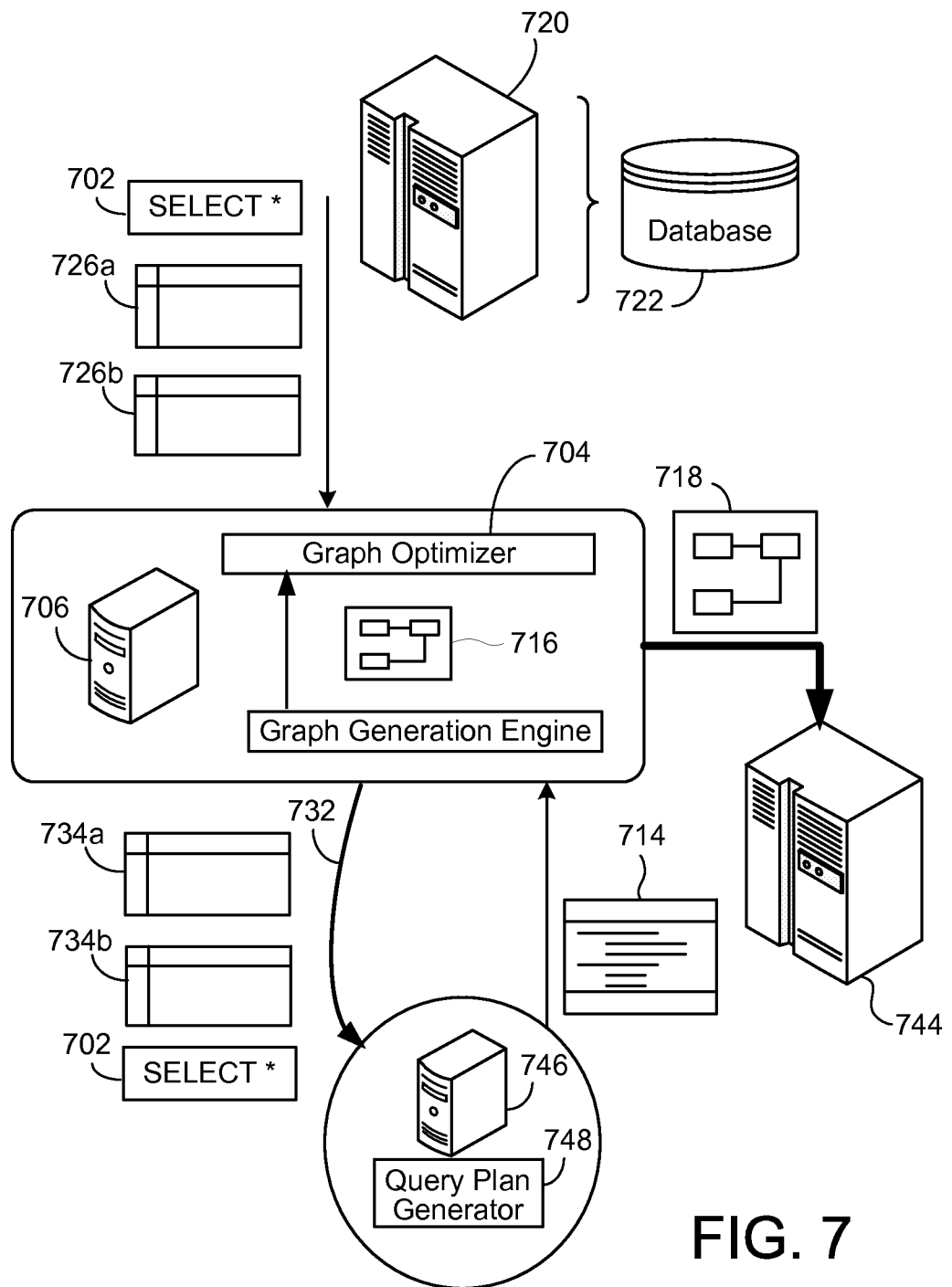


FIG. 7

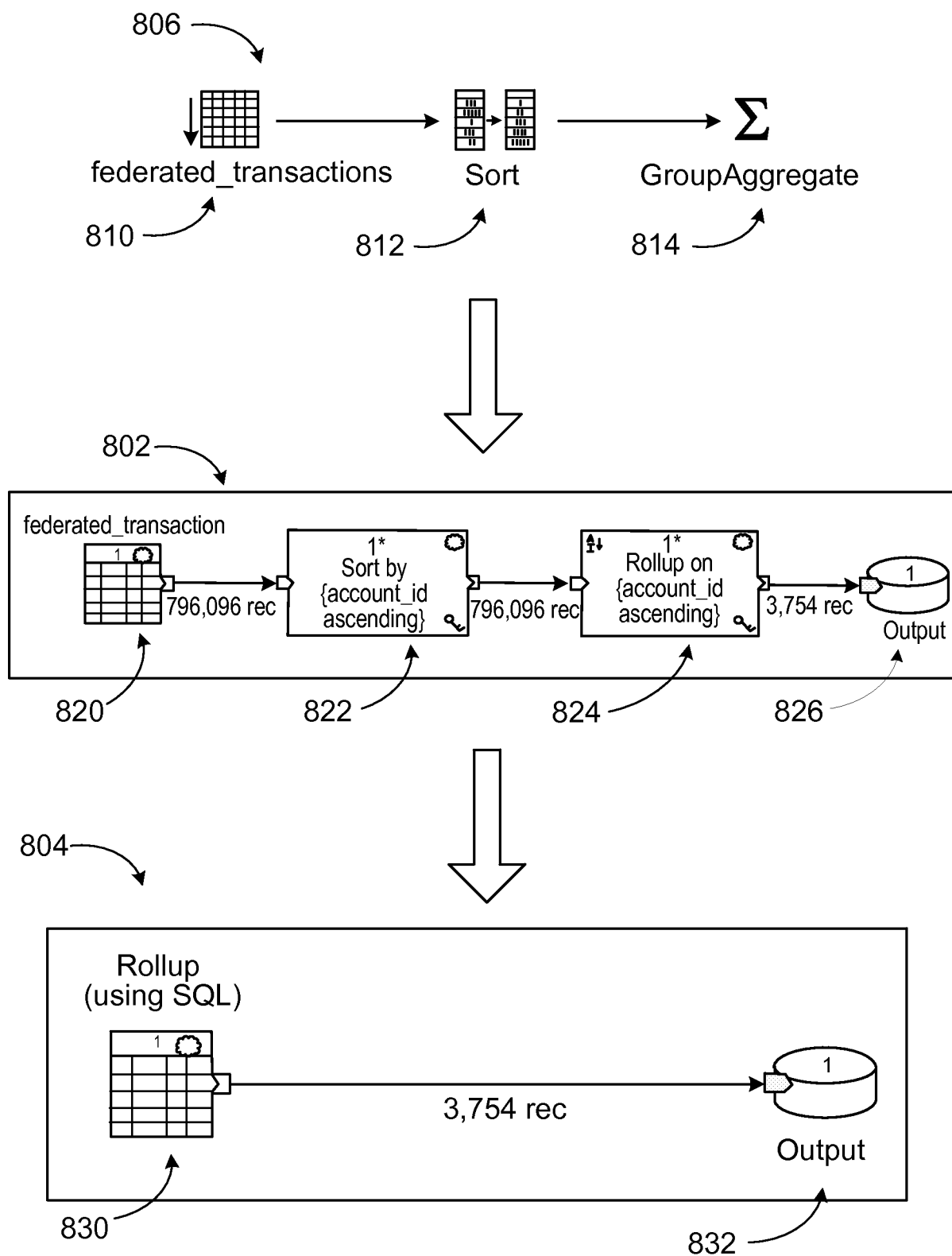


FIG. 8

940

Table 'Rollup(using SQL)' Properties

Description ☐ Data ☐ Parameters ☐ Layout ☐ Ports ☐ Condition

Configuration file _____

/usr/share/oracle-linux12.dbc
DBMS type: oracle

Config File ▶

Data location _____

Name: _____

Select statement:

```
select count (*) as num_transactions,  
sum (federated_transaction.transaction_amt) as total,  
federated_transactions.account_id as account_id  
from OPS&ISCHECHT.FEDERATED_TRANSACTIONS  
GROUP BY federated_transactions.account_id  
ORDER BY federated_transactions.account_id ASC
```

942

Edit...

(Hint: use \${} for parameter reference)

EME dataset _____

☐ Associate table with EME dataset:

Edit... Default

Export...

OK Cancel Apply Help

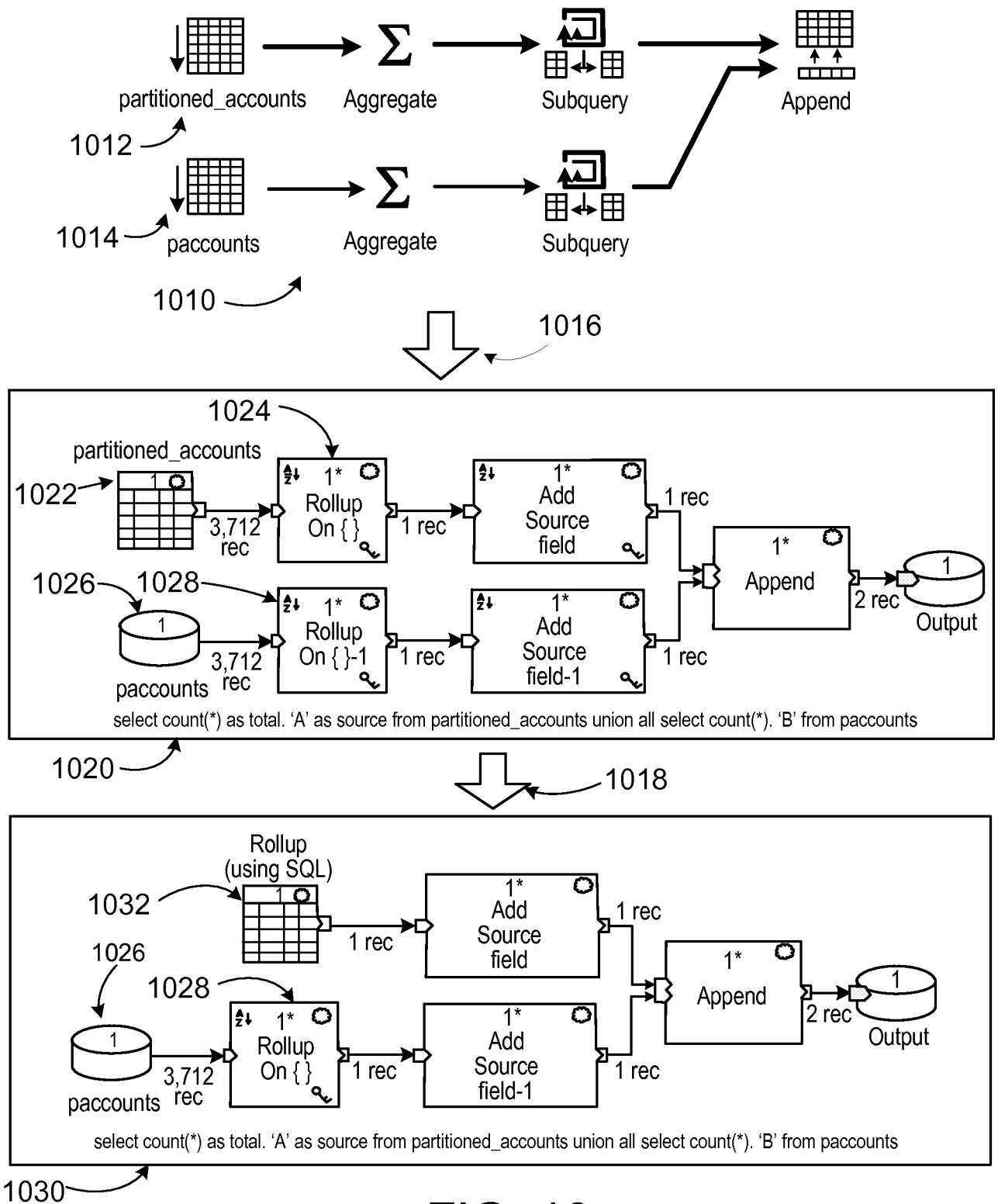


FIG. 10

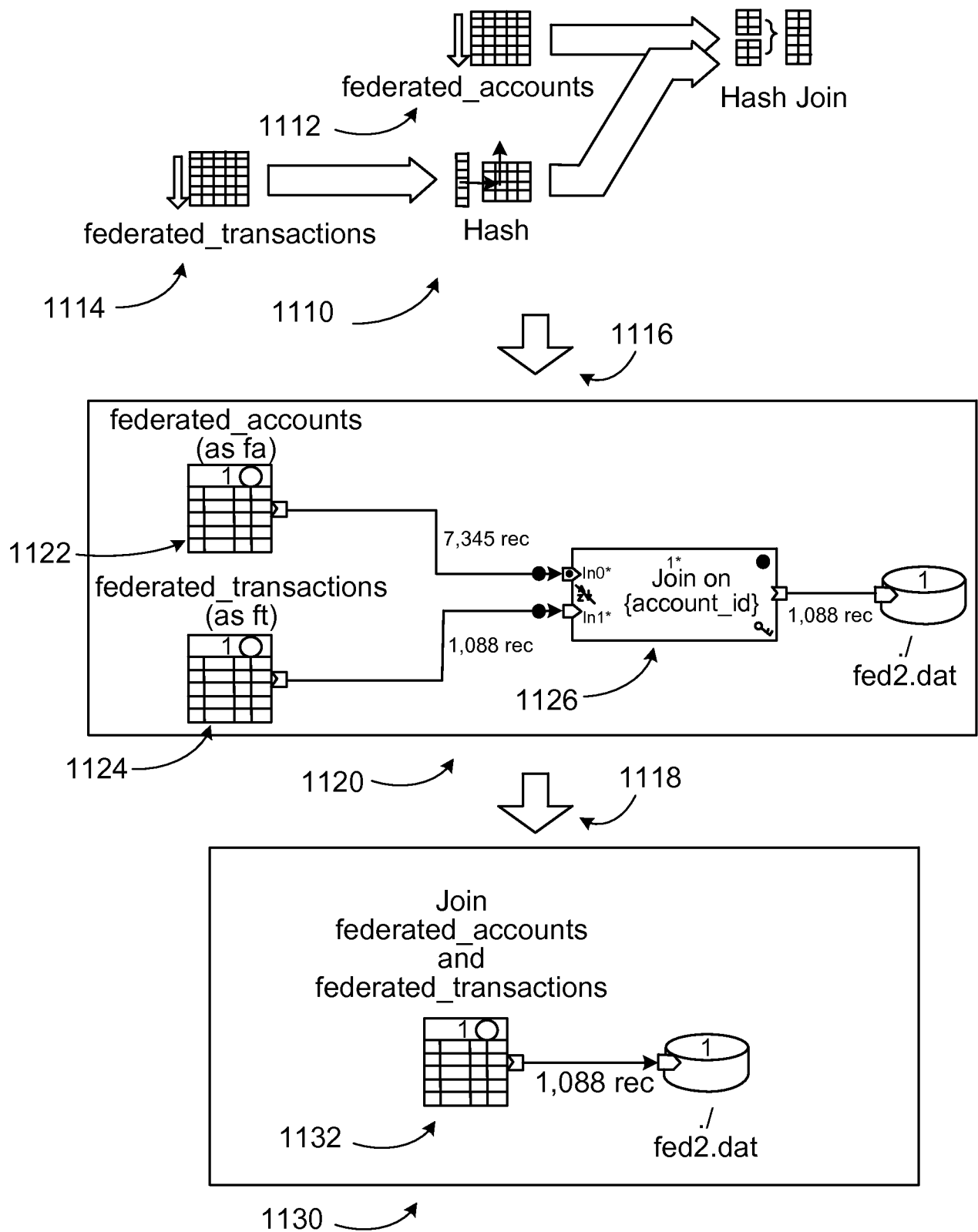
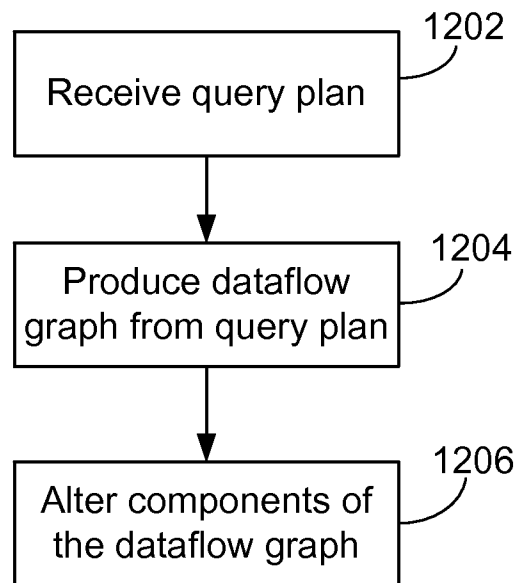


FIG. 11

**FIG. 12**