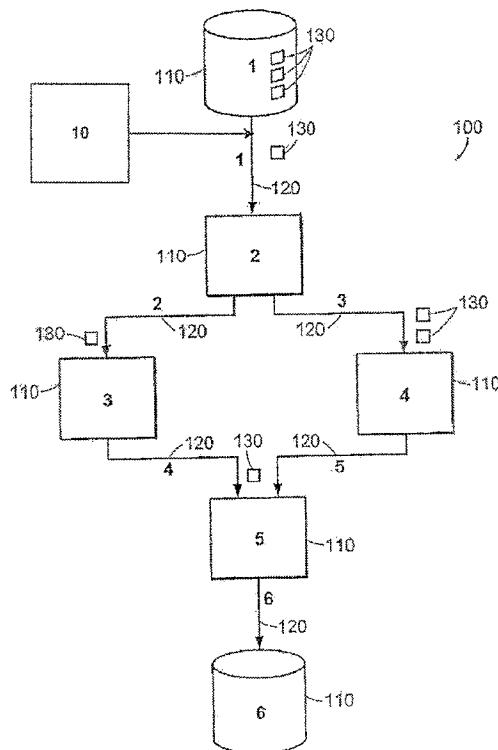




(22) Date de dépôt/Filing Date: 2008/07/25
(41) Mise à la disp. pub./Open to Public Insp.: 2009/01/29
(45) Date de délivrance/Issue Date: 2020/01/07
(62) Demande originale/Original Application: 2 697 306
(30) Priorité/Priority: 2007/07/26 (US60/952,075)

(51) Cl.Int./Int.Cl. G06F 9/44 (2018.01),
G06F 11/07 (2006.01)
(72) Inventeurs/Inventors:
STANFILL, CRAIG W., US;
WHOLEY, JOSEPH SKEFFINGTON, III, US
(73) Propriétaire/Owner:
AB INITIO TECHNOLOGY LLC, US
(74) Agent: SMART & BIGGAR LLP

(54) Titre : CALCUL A BASE DE GRAPHE TRANSACTIONNEL AVEC MANIPULATION D'ERREUR
(54) Title: TRANSACTIONAL GRAPH-BASED COMPUTATION WITH ERROR HANDLING



(57) Abrégé/Abstract:

Processing transactions using graph-based computations includes determining (305) that at least one of a plurality of graph elements of a computation graph of a set of one or more computation graphs includes a computation to be performed for a given transaction, associating the given transaction with an instance (310) of the computation graph that includes reusable computation elements associated with respective graph elements, and executing (320) the graph to perform the computation.

Abstract

Processing transactions using graph-based computations includes determining (305) that at least one of a plurality of graph elements of a computation graph of a set of one or more computation graphs includes a computation to be performed for a given transaction,

- 5 associating the given transaction with an instance (310) of the computation graph that includes reusable computation elements associated with respective graph elements, and executing (320) the graph to perform the computation.

Transactional Graph-Based Computation with Error Handling

[01] This is a divisional of Canadian Patent Application Serial No. 2,697,306 filed on July 25, 2008.

CROSS-REFERENCE TO RELATED APPLICATIONS

5 [01a] This application claims priority to U.S. Application Serial No. 60/952,075, filed on July 26, 2007.

Background

[02] This invention relates to execution of graph-based computations.

[03] Complex computations can often be expressed as a data flow through a directed graph, 10 with components of the computation being associated with the vertices of the graph and data flows between the components corresponding to links (arcs, edges) of the graph. A system that implements such graph-based computations is described in U.S. Patent 5,966,072, EXECUTING COMPUTATIONS EXPRESSED AS GRAPHS. One approach to executing a graph-based computation is to execute a number of processes, each associated with a different 15 vertex of the graph, and to establish communication paths between the processes according to the links of the graph. For example, the communication paths can use TCP/IP or UNIX domain sockets, or use shared memory to pass data between the processes.

Summary

[04] According to one aspect, there is provided a method for processing graph-based computations including: within a graph including vertices representing work element-handling graph components that process work elements according to links joining the vertices, providing at least one error-handling graph component within the computation graph, the at least one error-handling graph component being distinct from the work element-handling graph components and configured to provide error information to a process external 20 to the graph, and processing data, including, in response to a work element-handling graph component encountering an error while processing, redirecting processing to the error- 25

handling graph component, including directing at least some of the work elements to the error-handling graph component according to at least one link to a vertex representing the error-handling graph component.

[04a] According to another aspect, there is provided a system for processing graph-based computations, the system including: within a graph including vertices representing work element-handling graph components that process work elements according to links joining the vertices, means for providing at least one error-handling graph component within the computation graph, the at least one error-handling graph component being distinct from the work element-handling graph components and configured to provide error information to a process external to the graph, and means for processing data, including, in response to a work element-handling graph component encountering an error while processing, redirecting processing to the error-handling graph component, including directing at least some of the work elements to the error-handling graph component according to at least one link to a vertex representing the error-handling graph component.

[04b] According to still another aspect, there is provided a computer-readable medium storing a computer program for processing graph-based computations, the computer program including instructions for causing a computer system to: within a graph including vertices representing work element-handling graph components that process work elements according to links joining the vertices, provide at least one error-handling graph component within the computation graph, the at least one error-handling graph component being distinct from the work element-handling graph components and configured to provide error information to a process external to the graph, and process data, including, in response to a work element-handling graph component encountering an error while processing, redirecting processing to the error-handling graph component, including directing at least some of the work elements to the error-handling graph component according to at least one link to a vertex representing the error-handling graph component.

[04c] According to yet another aspect, there is provided a computing system for processing graph-based computations including: at least one processor configured to: within a computation graph including vertices representing work element-handling graph components

that process work elements according to links joining the vertices, provide at least one error-handling graph component within the computation graph, the at least one error-handling graph component distinct from the work element-handling graph components and configured to provide error information to a process external to the computation graph, process data,

5 including, in response to a work element-handling graph component encountering an error while processing, redirecting processing to the error-handling graph component, and direct at least some of the work elements to the error-handling graph component according to at least one link to a vertex representing the error-handling graph component.

[05] Aspects can include one or more of the following features.

10 [06] At least some instances of the graphs in the set of computation graphs share one or more of the computation elements.

83998618

5 [07] The computation elements include computations executed by at least one of an operating system process and a process thread.

[08] The graph elements include vertices of the computation graphs.

[09] Associating the transaction with an instance of the computation graph includes assigning a computation element corresponding to each graph element in the 10 computation graph to the instance of the computation graph before beginning executing the graph elements.

[10] Associating the transaction with an instance of the computation graph includes assigning a computation element corresponding to a graph element in the computation graph to the instance of the computation graph after executing another graph element 15 using a computation element already assigned to the instance.

[11] At least two of the graph elements use a common resource, and executing the graph to perform the computation includes assigning each of the graph elements using the common resource to a single computation element.

[12] The single computation element is already initiated when the graph elements 20 are assigned to the computation element.

[13] The common resource includes a database.

[14] The common resource includes a specific port.

[15] Processing the transaction includes receiving a request for the transaction.

[16] The method also includes determining that the same computation graph is 25 associated with a computation to be performed for a second transaction, associating the second transaction with a second instance of the computation graph, and executing the second instance of the graph to perform the computation for the second transaction.

[17] The computations for transactions performed using different instances of 30 computation graphs are performed in a time-interleaved manner.

[18] Multiple transactions are processed concurrently.

83998618

5 [19] Each transaction is associated with one or more work elements that are processed according to the corresponding computation graph.

[20] At least some transactions are each associated with one work element that is processed according to the corresponding computation graph.

10 [21] The method further includes forming multiple instances of at least some of the computation graphs.

[22] The method further includes identifying that an error has occurred in the performing of a computation for one of the transactions, and continuing the performing of a computation for another one of the transactions.

15 [23] The processing of a first transaction of the plurality of transactions starts at a first time, and the processing of a second transaction of the plurality of transactions starts at a second time later than the first time, the method further includes completing the performing of the computation for the second transaction before completing the performing of the computation for the first transaction.

20 [24] In another aspect, in general, a system for processing transactions using graph-based computations includes means for determining that at least one of a plurality of graph elements of a computation graph of a set of one or more computation graphs includes a computation to be performed for a transaction, means for associating the given transaction with an instance of the computation graph that includes reusable computation elements associated with respective graph elements, and means for executing the graph to perform the computation.

30 [25] In another aspect, in general, a computer-readable medium stores a computer program for processing transactions using graph-based computations. The computer program includes instructions for causing a computer system to: determine that at least one of a plurality of graph elements of a computation graph of a set of one or more computation graphs includes a computation to be performed for a given transaction, associate the given transaction with an instance of the computation graph that includes reusable computation elements associated with respective graph elements, and execute the graph to perform the computation.

83998618

5 [26] In another aspect, in general, a method for processing graph-based computations includes: within a graph including vertices representing graph components that process work elements according to links joining the vertices, providing at least one error-handling graph component configured to provide error information to a process external to the graph, and processing data, including, in
10 response to a graph component encountering an error while processing, redirecting processing to the error-handling graph component including directing at least some of the work elements to the error-handling component according to at least one link to a vertex representing the error-handling component.

[27] Aspects can include one or more of the following features.

15 [28] Redirecting processing to the error-handling graph component includes removing work elements from at least one input queue.

[29] Redirecting processing to the error-handling graph component includes processing the work elements directed to the error-handling graph component.

20 [30] Processing the work elements directed to the error-handling graph component includes rolling back changes to a database made prior to the error.

[31] Processing the data includes, for graph components not included in handling the error, discarding work elements directed to those graph components.

[32] A sub-graph is provided, the sub-graph including an error-handling sub-graph component configured to provide an error code as an output of the sub-graph.

25 [33] If output provided by the sub-graph indicates that an error occurred in the sub graph, processing is redirected to the error-handling graph component.

[34] Redirecting processing to the error-handling graph component includes communicating, from the graph component that encountered the error, to the error-handling graph component, work elements that the graph component was processing
30 when the error occurred.

[35] The work elements are communicated according to the link to the vertex representing the error-handling component.

83998618

5 [36] Redirecting processing to the error-handling graph component includes communicating, from the graph component that encountered the error, to the error-handling graph component, reporting information about the error.

10 [37] The reporting information is communicated according to an implicit connection between the graph component that encountered the error and the error-handling component.

[38] The implicit connection is revealed as an explicit link between a vertex representing the graph component that encountered the error and a vertex representing the error-handling component in response to a user request.

15 [39] Providing the error-handling graph component includes providing a plurality of error-handling graph components, and redirecting processing to the error-handling graph component includes selecting an error-handling graph component based on output provided from the graph component that encountered the error.

[40] Processing the data also includes, if a graph component encounters an error while processing, outputting an identification of a work element that caused the error.

20 [41] Processing includes: enabling a first component of the graph; disabling the error-handling component; and for each component downstream of the first component other than the error-handling component, enabling the component if a component immediately upstream of the component is enabled.

25 [42] Redirecting processing to the error-handling graph component includes: stopping execution of each enabled graph component, disabling the component that encountered the error; enabling the error-handling component; disabling components downstream of the component that encountered the error that are not downstream of the error-handling component; and enabling components upstream of the error-handling component.

30 [43] Redirecting processing to the error-handling graph component includes, where the error occurred in a first component, if the error occurs under a first condition, directing process flow from the first component to a first error-handling component upstream of the first component, and if the error occurs under a second condition,

83998618

5 directing process flow from the first component to a second error-handling component downstream of the first component.

[44] The first condition is that a counter is below a limit.

[45] The second condition is that a counter is above a limit.

[46] Redirecting processing to the error-handling graph component also includes
10 enabling a set of graph components, the set having been determined prior to the error.

[47] In another aspect, in general, a system for processing graph-based computations includes, within a graph including vertices representing graph components that process work elements according to links joining the vertices, means for providing at least one error-handling graph component configured to provide error
15 information to a process external to the graph, and means for processing data, including, in response to a graph component encountering an error while processing, redirecting processing to the error-handling graph component including directing at least some of the work elements to the error-handling component according to at least one link to a vertex representing the error-handling component.

[48] In another aspect, in general, a computer-readable medium stores a computer program for processing graph-based computations. The computer program includes instructions for causing a computer system to: within a graph including vertices representing graph components that process work elements according to links joining the vertices, provide at least one error-handling graph component configured to
25 provide error information to a process external to the graph, and process data, including, in response to a graph component encountering an error while processing, redirecting processing to the error-handling graph component including directing at least some of the work elements to the error-handling component according to at least one link to a vertex representing the error-handling component.

[49] Other features and advantages of the invention are apparent from the following description, and from the claims.

Description of Drawings

[50] FIG. 1 is a diagram that illustrates an instance of graph-based computation.

83998618

5 [51] FIG. 2 is a logical block diagram of a system for processing work flows.

[52] FIG. 3A is a flowchart for processing each work flow.

[53] FIG. 3B is a flowchart for handling errors.

[54] FIGS. 4A, 4B, 5, and 6 are examples of error-handling graphs.

Description

10 1. Overview

[55] This application is related to United States patent application 10/268,509, Startup and Control of Graph-Based Computation, filed October 10, 2002, and 11/733,579, Transactional Graph-Based Computation, filed April 10, 2007, which is a continuation of application 10/268,509.

15 [56] The system described below implements a method for executing computations that are defined in terms of computation graphs. Referring to FIG. 1, an example of a computation graph 100 includes a number of vertices 110 that are joined by unidirectional links 120. In the example shown in FIG. 1, vertices 110 are numbered from 1 to 6, and links 120 are also numbered from 1 to 6. Computation graph 100 processes a work flow that is made up of a series of work elements 130, such as individual transactions that are processed according to a computation graph associated with a transaction processing system. A transaction may be composed of multiple work elements. Each vertex is associated with a portion of the computation defined by the overall computation graph. In this example, vertex 1 provides access to storage for an initial series of work elements 130 associated with one or more transactions, and passes that series on its output link 1. Processes that implement the computation associated with each of the vertices process the work elements 130 in turn, and typically produce a work element on one or more of the output links of that vertex.

20 [57] A process for a vertex is ready to run when at least one work element is queued at each of the vertex's inputs. As illustrated in FIG. 1, a work element 130 is in transit on link 1, a work element is queued ready for processing at vertex 3, and two work elements are queued for processing at vertex 4. Therefore, the processes for vertex 3 and vertex 4 are ready to run to process a queued work element. As

83998618

5 illustrated, vertex 5 has a work element queued on one of its inputs, link 4, but not on the other input, link 5. Therefore the process associated with vertex 5 is not ready to run.

[58] In some examples, a work flow may include work elements from multiple transactions (i.e., a first set of one or more work elements correspond to a first 10 transaction, a second set of one or more elements correspond to a second transaction, etc.). A transaction can include a set of work elements representing actions that are all to be processed as a set, such that if one action fails, none should be carried out. Multiple instances of a graph may be used to process multiple transactions, and 15 multiple instances of individual graph components (represented by vertices of a computation graph) may be created as needed by implementing computations of a graph component with a reusable computation element (e.g., an operating system process). By associating different transactions with different respective instances of graphs, multiple transactions can be processed concurrently. By enabling multiple computation elements to be assigned as needed to graph instances, efficient resource 20 sharing can be realized by having a computation element be used by one graph instance and reused by another graph instance, as described in more detail below.

[59] Referring to FIG. 2, a system 200 for executing computation graphs to process work flows comprising transactions includes stored graph data structures 210. These data structures include specifications of computation graphs that include 25 characteristics of the vertices and links of the graphs. Portions of these data structures may be accessed without loading an entire graph, for example, the specification of an individual graph component may be loaded in order to assign a work element to a newly-created instance of that graph component.

[60] A transaction subscription module 220 of the system receives control inputs 30 222 from a transaction subscribing graph component (e.g., a component providing commands without necessarily processing work elements, such as the component represented by vertex 10 FIG. 1) including commands to process particular work flows 232 using corresponding computation graphs, which are specified in the stored graph data structures 210. The transaction subscription module 220 keeps track of 35 graph computation processing resources 230 available for instantiating graph instances that are to be assigned to particular transactions. The transaction

83998618

5 subscription module 220 includes a scheduler that uses the specifications of the computation graphs to determine how to instantiate graph instances using the graph computation processing resources 230, which is generally made up of multiple processes (or pools of processes), where each process functions as a reusable computation element that instantiates a given graph component in a graph instance.

10 The processes that are executed to perform the computations of the components of a graph can make use of external data and processes 240, which include database engines, data storage, or other modules that are accessed during processing associated with vertices of the computation graphs. In some examples, a single process or set of processes capable of performing multiple different operations is bound to a given

15 instance of a graph to handle all the operations of that instance.

[61] In some examples, the scheduler of the transaction subscription module 220 uses a remote procedure call (RPC) process. When the scheduler receives a work element for a given transaction, it assigns the work element to the appropriate component of a graph instance associated with (i.e., assigned to) the transaction. The

20 process assigned to that graph instance executes the computation of that component. The data associated with the work element is written to a temporary space available for the graph instance and accessible by the process. The scheduler is notified that the transaction subscription module 220 is done with that component, and the scheduler then schedules any downstream graph components for execution. Eventually the

25 transaction will progress through the whole graph (as the graph is executed using the graph computation processing resources 230), and be output by way of an RPC publish process. This takes the data accumulated in the temporary space and commits it to the appropriate output channel, e.g., the database output 6 in FIG. 1. The RPC publish process can be multiplexed with the RPC subscribe process so that it can

30 access the socket on which the transaction was initially received.

[62] In general, different transactions may be processed concurrently, each being processed by a different instance of a graph. System 200, through the transaction subscription module 220, allocates resources for an instance of a computation graph for each transaction and, through the graph computation processing resources 230, controls their execution to process the work flows.

83998618

5 2. Graph data structures

[63] System 200 includes a number of features that provide rapid startup of graph computations as well as efficient sharing of limited resources.

[64] Before processing a transaction with an instance of a computation graph, the transaction subscription module 220 creates a runtime data structure for that graph 10 instance in a functionally shared memory. In one embodiment, a single shared memory segment is created in which all the runtime data structures for graph instances are created.

[65] The process or processes bound to a transaction are associated with the 15 vertices of the graph and each of these processes maps the shared memory segment into its address space. The processes may be associated with vertices when graph instances are created for individual transactions or they may not be associated with vertices until instances of individual graph components are created or executed. The 20 processes read and write work elements from and to the runtime data structures for the graph instances during processing of the transaction. That is, data for the transactions that flow through the graph are passed from component to component, and from process to process if more than one process is bound to the transaction, through these 25 runtime data structures in the shared memory segment. By containing the data for a given transaction in a memory space accessible to each component of the graph and executing each component with a consistent process or set of processes, state can be shared between the components. Among other advantages, this allows all the database 30 operations associated with executing the computations for a transaction to be committed at once, after it is confirmed that the transaction executed successfully.

3. Process pools

[66] As introduced above, graph computation processing resources 230 for 30 executing the components of a graph instance can be implemented using process pools managed and allocated by the scheduler. For each of a number of different types of computation, a pool of processes is created prior to beginning processing of work flows of transactions using graph components requiring that type of computation. When a transaction is assigned to a graph instance, if computation of a particular type 35 will be needed to perform the computation for a given component of the graph

83998618

5 instance, the scheduler allocates a member of the process pool for use by the graph instance and with the given component. The member of the process pool remains associated with that graph instance for the duration of processing of the transaction, and may be re-used for other components within that graph instance that require the same type of computation. The process may be released back to the pool once no
10 work elements remain upstream of the last component in the graph instance for that transaction that needs that type of computation. There may be many different pools of processes, each associated with a corresponding type of computation. Processes in a pool may be used for components in the same or different graph instances, including for a given type of component in different graph instances, and for multiple different components in one graph instance, for example.
15

[67] In some implementations, each process in a process pool is a separate process (e.g., a UNIX process) that is invoked by the transaction subscription module 220, which manages the process pools. The module 220 maintains a separate work queue for each process pool. Each entry in a work queue identifies a specific vertex of a
20 graph instance for which the process is to perform computation.

[68] Some processes reserve or consume fixed resources. An example of such a process is one that makes a connection to a database, such as an Oracle® database. Since resources are consumed with forming and maintaining each database connection, it is desirable to limit the number of such processes that are active. If a
25 graph includes multiple components that access a database, it may be desirable for all the database operations for a given transaction to take place in a single database process. To accommodate this, a set of processes may be established that each maintain a connection to the database and are each capable of performing the database functions that a given graph instance may require. When a graph instance is assigned to a given transaction, one process from the set is assigned to that graph instance for
30 the entire transaction, as described above, and all of the database components are multiplexed to that process. When a vertex requires a process for accessing the database to process a work element of the transaction, the assigned process (which has already established its connection with the database) is associated with that vertex. In this way, the overhead of the initialization steps of that process that would have been
35 required to connect to that database is avoided, and all database actions for a given

83998618

5 transaction are handled by the same process. Other types of processes can be handled in the same way.

[69] System 200 supports different approaches to configuring processes for vertices, which differ in when the vertices are associated with processes and when the computation for the vertices is initiated. In one type of configuration, a process is not 10 associated with a vertex until all the data at all its input work elements are completely available. If a work element is large, it may take some time for the entire work element to be computed by the upstream vertex and to be available. This type of configuration avoids blocking the process waiting for input to become available, so that it can be used by other vertices in that graph instance.

15 [70] Another type of configuration uses a streaming mode. A process is associated with a vertex and initiated when at least the start of each input is available. The remainder of each of its inputs becomes available while the process executes. If that input becomes available sufficiently quickly, the process does not block waiting for input. However, if the inputs do not become available, the process may block.

20 4. Computation control

[71] FIG. 3A is a flowchart for a process 300 for processing each transaction using 25 respective graph instances. When the transaction subscription module 220 (FIG. 2) receives a request to process a transaction, it first determines which computation graph (and corresponding type) is appropriate to process the transaction (step 305). For example, the scheduler determines that a certain computation graph is appropriate (e.g., includes appropriate components) to perform a computation for the transaction. The transaction itself may specify this, or the transaction subscription module 220 may include or have access to data that associates particular transaction types with 30 particular computation graphs. The transaction subscription module 220 then creates a graph instance (if necessary) of a computation graph of the type needed to process that transaction (step 310) and associates the transaction with it. As part of this process, the transaction subscription module 220 allocates a portion of the shared memory segment for the runtime data structure for the graph instance, and copies a graph template for that type of computation graph into runtime data structure, thereby 35 initializing the runtime data structure. Examples of the use of graph templates are

83998618

5 described in more detail in U.S. Patent No. 7,167,850.

In some examples, graph instances have already been created and at this stage one is merely assigned to the present transaction. The transaction subscription module 220 then executes the graph instance (step 320), under control of the scheduler, as described in more detail below. The graph instance includes 10 computation elements (e.g., processes) associated with (assigned to) respective components that are reusable. When the transaction's entire work flow has been processed, the transaction subscription module 220 commits results of the execution of the graph (e.g., commits changes to an output database) and optionally releases the assigned resources and computation elements and deletes the runtime data structure 15 for the graph instance, thus permitting that portion of the shared memory segment to be reused for other graph instances (step 330).

5. Alternatives

[72] As noted above, it is possible to pre-create graph pools of already instantiated instances of computation graphs in anticipation of there being transactions that will 20 require them. When a transaction is received and needs a graph instance, if one is available from a graph pool, it is assigned from the pool rather than having to be created. In this way, the startup cost for a transaction is further reduced. When the computation for the transaction is completed, the graph is reset by restoring variables to their initial values prior to having been assigned to the transaction and freeing any 25 dynamically-assigned memory. After the graph instance is reset it is returned to the pool.

[73] In some examples, the number of graph instances in a graph pool can be allowed to grow as needed. For instance, there might be a minimum number of instances of each graph, and more may be created as needed.

30 [74] In the description above, processes may be assigned to vertices in the graph in an on-demand manner where they are not associated with a vertex until after all the inputs to that vertex are available, though they are bound to the particular graph instance and transaction. Another approach is to associate the processes to the vertices 35 when the transaction is associated with the graph instance and to maintain the association until the transaction's entire work flow has been processed.

83998618

5 6. Applications

[75] One application of computation graphs of the type described above is for processing financial transactions in a banking application. In general, different types of transactions require different types of computation graphs. A typical computation graph is associated with some combination of a type of customer transaction and 10 “backend” services that are needed to process the transaction. For example, transactions can be ATM requests, bank teller inputs, and business-to-business transactions between computers or web servers. Different customers might have different backend systems, particularly when banks consolidate and customers are combined from different original banks. Their accounts may be maintained on very 15 different backend systems even though they are all customers of the acquiring bank. Therefore, different vertices in a graph may be used to process different transactions. Different services may be associated with vertices in the graph. For example, some of the vertices may be associated with functions such as updating a balance, depositing money in an account, or performing an account hold so funds are held in an account. 20 In some implementations, on-the-fly assignment of processes to vertices avoids the overhead of having processes for unused vertices remain idle.

[76] An advantage of allocating graph instances on a per-transaction basis is that it allows parallelization of data streams that otherwise would have to be processed serially. Graph instances assigned to different transactions may finish in a different 25 order than they started, for example, if the first transaction was more complicated than the second. This may allow the second graph instance to be released and available to process a third transaction when a serialized system would still be processing the first transaction.

7. Error Handling

30 [77] An advantage of allocating graph instances on a per-transaction basis is that failures due to errors in executing a graph instance can be contained to that transaction, and do not compromise the concurrent processing of other graph instances. By delaying committing the results of the computation graph until the entire transaction is completed, the data can be “rolled-back”, in the event of an error,

83998618

5 to the state that it was in before the system began to process the transaction. Errors can be handled in several ways.

[78] In some examples, an “error handling” component is included in a graph. The error handling component is a special case in that it does not have to execute for the graph to complete. In the event that the component at any vertex generates an error, 10 instead of causing the whole computation to abort, execution of the graph is redirected to the error handling component. An explicit relationship between a given component and an error handling component (including a work flow from an output port of a component to an input port of the error handling component) is referred to as an exception flow. The scheduler removes work elements that were part of the failed 15 computation from the graph instance and the error handling component provides an output which the graph can use to provide an error message as output to the process that called it. The error handling component may receive data input other than through an exception flow, depending on the implementation.

[79] FIG. 3B shows a flowchart for an exemplary procedure 350 for executing a 20 graph and handling an error that occurs in the graph. The scheduler processes a work flow of work elements (step 360) in the graph components according to the links. When the scheduler recognizes (step 370) that an error has occurred in a graph component, the scheduler redirects processing to the error-handling component. One 25 aspect of this redirection is directing (step 380) work elements to the error-handling component according to any exception flows to that error-handling component. As described in more detail below, processing of the exception flow enables the error-handling component to provide (step 390) error information to a process external to the graph that represents the state of the graph processing before the graph began processing the transaction in which the error occurred.

30 [80] For any component in a graph, there is a designated error handling component. This may be a component that directly receives an exception flow output or other error data output from another graph component, or it may be defined as the designated error handling component for a set of components regardless of whether it receives an exception flow. In some examples, exception flow is handled as shown in 35 FIGs. 4A-B. In this example, the graph is designed for carrying out a transactional computation and subscribe 902 and publish 904 components are shown, but the same

83998618

5 technique can be used in graphs for non-transactional work flows. In FIG. 4A, the scheduler has activated a graph 900. Beginning with the first component, subscribe 902, each component in any non-exception path downstream is marked as “enabled.” An exception path is a path that only receives a flow of work elements or other error data in the case of an exception (e.g., an exception flow leading to an error handling 10 component as described above). This is referred to as enablement propagation. A given component that is downstream of others is enabled if any of its inputs is connected to an upstream component that is enabled. That is, replicate 906, reformat 908, call web service 910, rollup 912, fuse 914, and publish 904 are all enabled, but the error handler 916 and the two components rollback 918 and error log 920 15 downstream of it, which do not receive non-exception input flows from any enabled components, remain “disabled.”

[81] If an error occurs, the scheduler halts execution of the erring component, allows any other components that are already executing to finish, and propagates any relevant data (e.g., exception flow output of the completed components, or “error 20 reporting output” of the erring component) to the error handling component. For example, if the call web service component 910 triggers an error, the exception flow from replicate component 906 and error reporting output from a reject port 921 of the call web service component 910 are input to the error handling component 916 at inputs 922, 924, respectively. Error reporting output ports (shown as ports on the 25 bottom of some of the components in the graph 900) can be used to provide information about any errors that have occurred including, for example, information characterizing what error(s) occurred, where the error(s) occurred, and any rejected work elements associated with the error(s).

[82] In this example, there are three error reporting output ports for the replicate 30 component 906. The reject port 921 provides work elements that may have caused the error or are in some way related to the error. The error port 923 provides an error messages describing relevant information about the error. The log port 925 can optionally provide information logging that the error occurred. The log port 925 can also provide log information about events during the normal course of execution even 35 if no errors occur. In this example, the reject port 921 is explicitly shown as connected for those components (e.g., the call web service component 910) that may need to use

83998618

5 the port. However, the error port 923 and log port 925 are not explicitly shown as connected, but have implicit connections to the error handling component 916. For example, the ports can be connected by a developer and then hidden using an interface control. In some implementations, the system can automatically determine implicit connections to a default error handling component, which may then be
 10 overridden by the developer. For large and/or complicated graphs, this “implicit wiring” for one or more types of error reporting ports improves visual comprehension of a graph by a developer, which is one of the benefits of graph-based programming. In some implementations, visual cues can be provided to indicate that a port is implicitly connected to a port of another component (e.g., an icon or a shaded or
 15 colored port). Some or all of the hidden implicit work flow connections can also be revealed as explicit links in response to a user request (e.g., clicking a button or hovering over a port).

[83] The exception flow output from the replicate component 906 may have already been queued at the input 922, if the replicate had finished operation before the
 20 error occurred. The scheduler then enables the error handling component (916 in this example), disables the erring component (910 in this example), and performs enablement propagation from the error handling component (enabling 918, 904, 920 in this example). Any component downstream of the disabled erring component is also disabled as long as that component does not receive a flow from an enabled
 25 component downstream of the error handling component (disabling 912 and 914 in this example). Finally, any remaining component that provides a flow to an enabled component is enabled (enabling 906 and 902 in this example).

[84] Thus, the result of this procedure is shown by the indication of “<enabled>” and “<disabled>” components in FIG. 4B. Connecting the publish component 904 back into the flow after the error handler 916 allows the transaction to be completed, albeit with an error message for its output. Data that had already propagated to the now-disabled components, e.g., output from the reformat component 908, is discarded.

[85] As noted above, data may flow to the error handling component as part of an
 35 exception flow or as part of an error reporting output of another component. Data that is available before the error occurs, for example, output data from the replicate

83998618

5 module 906 in FIG. 4B, is held in escrow in an input queue for the error handler 916 until it is needed, if it ever is. If the graph completes without error, the error handler 916 is never activated and the data is discarded. If an error does occur, the error handler 916 uses whatever input data it has received to formulate a response. In some examples, as in FIG. 4B, a rollback component 918 is used. The input data from the
 10 replicate component 906 tells the error handler 916 what the state of things was before the graph began processing the transaction. The error handler 916 outputs this to the rollback component 918 which uses it to restore any data that was modified by the other components to its state prior to the execution of the transaction. Execution flow then goes to both the error log 920 which logs the error, and to the publish component
 15 904 so that the error can be reported and appropriately handled by whatever higher-level process delivered it to the graph 900. An exception flow from any component to the error handler 916 may also include data. If there is input to the error handler 916 other than the original data from the replicate component 906, such as error output from the call web service component 910 or an exception flow from any other
 20 component (not shown), this may be used to formulate a more-detailed error message at either the error log or the publish component.

[86] In some examples, as shown in FIG. 5, a graph includes vertices that are implemented as sub-graphs, e.g. sub-graph 950, each of which may have its own error handling component 952. Thus, there can be a hierarchy of sub-graphs with a top
 25 level graph, having vertices that are sub-graphs at a lower “graph level,” and so on. If an error occurs in any component 954, 956, 958, 960, 962 of the sub-graph 950, the processing flow is routed to the error handling component 952 which provides error reporting output on a sub-graph error reporting port 974. The scope of the error handling component 952 is the sub-graph 950. The error handling component may
 30 have inputs 966, 968 that receive output from either an exception flow from another graph element (e.g., element 954) or error output 959 of another graph element (e.g., element 958) which may themselves be nested sub-graphs. In some examples, if an error handling component has multiple inputs, only the input that most recently received data is used. If all the components of the sub-graph 950 complete their
 35 operations successfully, output (a work flow) is delivered to the normal sub-graph output port 970 and processing flow beyond the sub-graph 950 continues normally. If an error occurs, it can be handled and reported on an error flow output 972, or an error

83998618

5 reporting output 974. In other examples, errors can also be reported on the standard output 970.

[87] If a sub-graph does not have error handling, its errors flow upwards in the hierarchy of sub-graphs of which it is a part until they reach a graph level that does have error handling, at which point that level's error-handling component is activated.

10 [88] The data escrowed at the input of the error handling component may be a subset of a work flow, it may be all the data associated with a transaction, or it could be an entire data flow. If the error-handling component has an error output port, it will output the record that caused the error or other error information based on the escrowed data or the input received from the component that had the error. If it does 15 not have such a port, it may simply output the offending record as normal output on its output port.

[89] If a sub-graph does not have error handling, errors in its components flow upwards in the hierarchy of sub-graphs of which it is a part until they reach a graph level that does have error handling, at which point that level's error-handling 20 component receives appropriate input and generates an appropriate error output.

[90] Error handling can allow cyclic graph arrangements that would ordinarily be avoided in graph-based computation processing. For example, as shown in FIG. 6, in a graph 1100 an error output 1116 from a computation component 1112 downstream from an error handler 1104 returns flow back to that same error handler 1104. The 25 error handler 1104 also receives input from a subscribe component 1102 and provides output to a roll-back component 1106 as in FIG. 4A. The roll-back 1106 returns the data to the state it was in before the failed computation was attempted, based on the data input to the error handler 1104 by the subscribe component 1102. A counter component 1108 may receive the flow from the roll-back component 1106 and 30 increment its value before returning flow to the gather component 1110. The computation component 1112 can use the value input from the counter component 1108 in several different ways. It can consult the value before carrying out its computation, for example, to see if it should change something about its operation. It can also consult the counter after an error, to see if some threshold number of attempts 35 has been made. If the threshold has been exceeded, instead of returning the error

83998618

5 output to the error handler 1104 through the output 1116 again, it directs its output to
a second error output 1118 that leads to a second error handler 1120. If a counter
component is not used, some other technique can be used to break the cycle and
assure that the graph eventually completes.

10 [91] To assure that a cyclic graph is well-defined, the set of elements that will be
enabled on error is determined in advance based on the topology of the graph, rather
than being done as-needed as described above.

15 [92] In some examples, other rules are used to assure that error handling works
correctly. For example, in some implementations, error handling can only be triggered
on one exception port of one component within a graph (any simultaneous errors may
be ignored). If a graph component or sub-graph is linked to an error handling
component, it must use that component on any error. If a graph component or sub-
graph is not linked to an error handling component, errors must be handled by the
generic error handler for the present scope. Each graph component is typically
associated with exactly one error handler. These rules may be modified or combined
20 depending on the requirements of the system. They can be useful where tight control
of the process for each transaction is needed.

25 [93] In some examples, when an error occurs, the operating system determines
which error-handling component is associated with the component that experienced
the error, and then determines which input flow, if any, to that error-handling
component should be used. If there are multiple inputs, the one that most recently had
data written to it is used.

30 [94] Error handling may be active, as just described, where components or sub-
graphs handle their own errors and produce error codes that can be used by other
components to diagnose or work around the error, or it can be passive. In a passive
system, a graph that encounters an error simply fails, and allows the operating system
to provide error handling, for example by providing a stack dump to a debugging
process.

35 [95] Each component of a graph is implicitly connected to a scheduler, which
doesn't need a specific invitation from a graph to intervene and handle errors. The
scheduler can remove data related to an error from a graph instance and, in some

83998618

5 examples, does not need to know the nature of the error. In some cases, the scheduler
may return resources assigned to a graph to their respective pools in stages, allowing
the graph to complete processing work elements that were not affected by the error.

8. Implementation

[96] The invention may be implemented in hardware or software, or a combination
10 of both (e.g., programmable logic arrays). Unless otherwise specified, the algorithms
described are not inherently related to any particular computer or other apparatus. In
particular, various general purpose machines may be used with programs written in
accordance with the teachings herein, or it may be more convenient to construct more
15 specialized apparatus (e.g., integrated circuits) to perform particular functions. Thus,
the invention may be implemented in one or more computer programs executing on
one or more programmed or programmable computer systems (which may be of
various architectures such as distributed, client/server, or grid) each comprising at
least one processor, at least one data storage system (including volatile and non-
20 volatile memory and/or storage elements), at least one input device or port, and at
least one output device or port. Program code is applied to input data to perform the
functions described herein and generate output information. The output information is
applied to one or more output devices, in known fashion.

[97] Each such program may be implemented in any desired computer language
(including machine, assembly, or high level procedural, logical, or object oriented
25 programming languages) to communicate with a computer system. In any case, the
language may be a compiled or interpreted language.

[98] 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
30 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
35 to perform the functions described herein.

83998618

5 [99] 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. Other embodiments are within the scope of the following claims.

CLAIMS:

1. A method for processing graph-based computations including:
within a graph including vertices representing work element-handling graph components that process work elements according to links joining the vertices,
5 providing at least one error-handling graph component within the computation graph, the at least one error-handling graph component being distinct from the work element-handling graph components and configured to provide error information to a process external to the graph, and
processing data, including, in response to a work element-handling graph
10 component encountering an error while processing, redirecting processing to the error-handling graph component, including directing at least some of the work elements to the error-handling graph component according to at least one link to a vertex representing the error-handling graph component.
2. The method of claim 1 in which redirecting processing to the error-handling
15 graph component includes removing work elements from at least one input queue.
3. The method of claim 1 in which redirecting processing to the error-handling graph component includes processing the work elements directed to the error-handling graph component.
4. The method of claim 3 in which processing the work elements directed to the
20 error-handling graph component includes rolling back changes to a database made prior to the error.
5. The method of claim 3 in which processing the data includes, for graph components not included in handling the error, discarding work elements directed to those graph components.
- 25 6. The method of claim 1 also including providing a sub-graph, the sub-graph including an error-handling sub-graph component configured to provide an error code as an output of the sub-graph.

7. The method of claim 6 in which if output provided by the sub-graph indicates that an error occurred in the sub graph, redirecting processing to the error-handling graph component.
8. The method of claim 1 in which redirecting processing to the error-handling graph component includes communicating, from the graph component that encountered the error, to the error-handling graph component, work elements that the graph component was processing when the error occurred.
9. The method of claim 8 in which the work elements are communicated according to the link to the vertex representing the error-handling component.
10. The method of claim 8 in which redirecting processing to the error-handling graph component includes communicating, from the graph component that encountered the error, to the error-handling graph component, reporting information about the error.
11. The method of claim 10 in which the reporting information is communicated according to an implicit connection between the graph component that encountered the error and the error-handling component.
12. The method of claim 11 also including revealing the implicit connection as an explicit link between a vertex representing the graph component that encountered the error and a vertex representing the error-handling component in response to a user request.
13. The method of claim 1 in which providing the error-handling graph component includes providing a plurality of error-handling graph components, and redirecting processing to the error-handling graph component includes selecting an error-handling graph component based on output provided from the graph component that encountered the error.
14. The method of claim 1 in which processing the data also includes, if a graph component encounters an error while processing, outputting an identification of a work element that caused the error.
15. The method of claim 1 in which processing includes:

enabling a first component of the graph;
disabling the error-handling component; and
for each component downstream of the first component other than the error-handling component, enabling the component if a component immediately upstream of the
5 component is enabled.

16. The method of claim 15 in which redirecting processing to the error-handling graph component includes:

10 stopping execution of each enabled graph component;
disabling the component that encountered the error;
enabling the error-handling component;
disabling components downstream of the component that encountered the error
that are not downstream of the error-handling component; and
enabling components upstream of the error-handling component.

17. The method of claim 1 in which redirecting processing to the error-handling
15 graph component includes, where the error occurred in a first component,

if the error occurs under a first condition, directing process flow from the first component to a first error-handling component upstream of the first component,
and if the error occurs under a second condition, directing process flow from the first component to a second error-handling component downstream of the first component.

20 18. The method of claim 17 in which the first condition is that a counter is below a limit.

19. The method of claim 17 in which the second condition is that a counter is above a limit.

20. The method of claim 17 in which redirecting processing to the error-handling
25 graph component also includes enabling a set of graph components, the set having been determined prior to the error.

21. A system for processing graph-based computations, the system including:

within a graph including vertices representing work element-handling graph components that process work elements according to links joining the vertices,

means for providing at least one error-handling graph component within the computation graph, the at least one error-handling graph component being distinct from the

5 work element-handling graph components and configured to provide error information to a process external to the graph, and

means for processing data, including, in response to a work element-handling graph component encountering an error while processing, redirecting processing to the error-handling graph component, including directing at least some of the work elements to the
10 error-handling graph component according to at least one link to a vertex representing the error-handling graph component.

22. A computer-readable medium storing a computer program for processing graph-based computations, the computer program including instructions for causing a computer system to:

15 within a graph including vertices representing work element-handling graph components that process work elements according to links joining the vertices,

provide at least one error-handling graph component within the computation graph, the at least one error-handling graph component being distinct from the work element-handling graph components and configured to provide error information to a process external
20 to the graph, and

process data, including, in response to a work element-handling graph component encountering an error while processing, redirecting processing to the error-handling graph component, including directing at least some of the work elements to the error-handling graph component according to at least one link to a vertex representing the
25 error-handling graph component.

23. A computing system for processing graph-based computations including:
at least one processor configured to:
within a computation graph including vertices representing work element-handling graph components that process work elements according to links joining the vertices,

provide at least one error-handling graph component within the computation graph, the at least one error-handling graph component distinct from the work element-handling graph components and configured to provide error information to a process external to the computation graph,

5 process data, including, in response to a work element-handling graph component encountering an error while processing, redirecting processing to the error-handling graph component, and

direct at least some of the work elements to the error-handling graph component according to at least one link to a vertex representing the error-handling graph
10 component.

24. The computing system of claim 23 in which redirecting processing to the error-handling graph component includes removing work elements from at least one input queue.

25. The computing system of claim 23 in which redirecting processing to the error-handling graph component includes processing the work elements directed to the error-handling graph component.
15

26. The computing system of claim 25 in which processing the work elements directed to the error-handling graph component includes rolling back changes to a database made prior to the error.

27. The computing system of claim 25 in which processing the data includes, for
20 graph components not included in handling the error, discarding work elements directed to those graph components.

28. The computing system of claim 23 wherein the processor is configured to provide a sub-graph, the sub-graph including an error-handling sub-graph component configured to provide an error code as an output of the sub-graph.

25 29. The computing system of claim 28 in which if output provided by the sub-graph indicates that an error occurred in the sub graph, redirecting processing to the error-handling graph component.

30. The computing system of claim 23 in which redirecting processing to the error-handling graph component includes communicating, from the graph component that encountered the error, to the error-handling graph component, work elements that the graph component was processing when the error occurred.
- 5 31. The computing system of claim 30 in which the work elements are communicated according to the link to the vertex representing the error-handling component.
32. The computing system of claim 30 in which redirecting processing to the error-handling graph component includes communicating, from the graph component that encountered the error, to the error-handling graph component, reporting information about the
10 error.
33. The computing system of claim 32 in which the reporting information is communicated according to an implicit connection between the graph component that encountered the error and the error-handling component.
- 15 34. The computing system of claim 33 wherein the processor is configured to reveal the implicit connection as an explicit link between a vertex representing the graph component that encountered the error and a vertex representing the error-handling component in response to a user request.
- 20 35. The computing system of claim 23 in which providing the error-handling graph component includes providing a plurality of error-handling graph components, and redirecting processing to the error-handling graph component includes selecting an error-handling graph component based on output provided from the graph component that encountered the error.
36. The computing system of claim 23 in which processing the data also includes, if a graph component encounters an error while processing, outputting an identification of a work element that caused the error.
- 25 37. The computing system of claim 23 in which processing includes:
enabling a first component of the computation graph;
disabling the error-handling component; and

for each component downstream of the first component other than the error-handling component, enabling the component if a component immediately upstream of the component is enabled.

38. The computing system of claim 37 in which redirecting processing to the error-handling graph component includes:

stopping execution of each enabled graph component,
disabling the component that encountered the error;
enabling the error-handling component;
disabling components downstream of the component that encountered the error
10 that are not downstream of the error-handling component; and
enabling components upstream of the error-handling component.

39. The computing system of claim 23 in which redirecting processing to the error-handling graph component includes, where the error occurred in a first component,
if the error occurs under a first condition, directing process flow from the first
15 component to a first error-handling component upstream of the first component,
and if the error occurs under a second condition, directing process flow from
the first component to a second error-handling component downstream of the first component.

40. The computing system of claim 39 in which the first condition is that a counter
is below a limit.

20 41. The computing system of claim 39 in which the second condition is that a
counter is above a limit.

42. The computing system of claim 39 in which redirecting processing to the error-handling graph component also includes enabling a set of graph components, the set having
been determined prior to the error.

25 43. The computer-readable storage medium of claim 22 in which redirecting
processing to the error-handling graph component includes removing work elements from at
least one input queue.

44. The computer-readable storage medium of claim 22 in which redirecting processing to the error-handling graph component includes processing the work elements directed to the error-handling graph component.

45. The computer-readable storage medium of claim 44 in which processing the work elements directed to the error-handling graph component includes rolling back changes to a database made prior to the error.

46. The computer-readable storage medium of claim 44 in which processing the data includes, for graph components not included in handling the error, discarding work elements directed to those graph components.

10 47. The computer-readable storage medium of claim 22 wherein the computer program includes instructions for causing the computer system to provide a sub-graph, the sub-graph including an error-handling sub-graph component configured to provide an error code as an output of the sub-graph.

15 48. The computer-readable storage medium of claim 47 in which if output provided by the sub-graph indicates that an error occurred in the sub graph, redirecting processing to the error-handling graph component.

49. The computer-readable storage medium of claim 22 in which redirecting processing to the error-handling graph component includes communicating, from the graph component that encountered the error, to the error-handling graph component, work elements 20 that the graph component was processing when the error occurred.

50. The computer-readable storage medium of claim 49 in which the work elements are communicated according to the link to the vertex representing the error-handling component.

51. The computer-readable storage medium of claim 49 in which redirecting 25 processing to the error-handling graph component includes communicating, from the graph component that encountered the error, to the error-handling graph component, reporting information about the error.

52. The computer-readable storage medium of claim 51 in which the reporting information is communicated according to an implicit connection between the graph component that encountered the error and the error-handling component.

53. The computer-readable storage medium of claim 52 wherein the computer program includes instructions for causing the computer system to reveal the implicit connection as an explicit link between a vertex representing the graph component that encountered the error and a vertex representing the error-handling component in response to a user request.

54. The computer-readable storage medium of claim 22 in which providing the error-handling graph component includes providing a plurality of error-handling graph components, and redirecting processing to the error-handling graph component includes selecting an error-handling graph component based on output provided from the graph component that encountered the error.

55. The computer-readable storage medium of claim 22 in which processing the data also includes, if a graph component encounters an error while processing, outputting an identification of a work element that caused the error.

56. The computer-readable storage medium of claim 22 in which processing includes:

20 enabling a first component of the computation graph;
disabling the error-handling component; and
for each component downstream of the first component other than the error-handling component, enabling the component if a component immediately upstream of the component is enabled.

57. The computer-readable storage medium of claim 56 in which redirecting processing to the error-handling graph component includes:

stopping execution of each enabled graph component,
disabling the component that encountered the error;
enabling the error-handling component;

disabling components downstream of the component that encountered the error that are not downstream of the error-handling component; and
enabling components upstream of the error-handling component.

58. The computer-readable storage medium of claim 22 in which redirecting processing to the error-handling graph component includes, where the error occurred in a first component,
if the error occurs under a first condition, directing process flow from the first component to a first error-handling component upstream of the first component,
and if the error occurs under a second condition, directing process flow from
10 the first component to a second error-handling component downstream of the first component.

59. The computer-readable storage medium of claim 58 in which the first condition is that a counter is below a limit.

60. The computer-readable storage medium of claim 58 in which the second condition is that a counter is above a limit.

15 61. The computer-readable storage medium of claim 58 in which redirecting processing to the error-handling graph component also includes enabling a set of graph components, the set having been determined prior to the error.

62. The method of claim 1 including committing the performed computation to a database after multiple computations for a given transaction are performed.

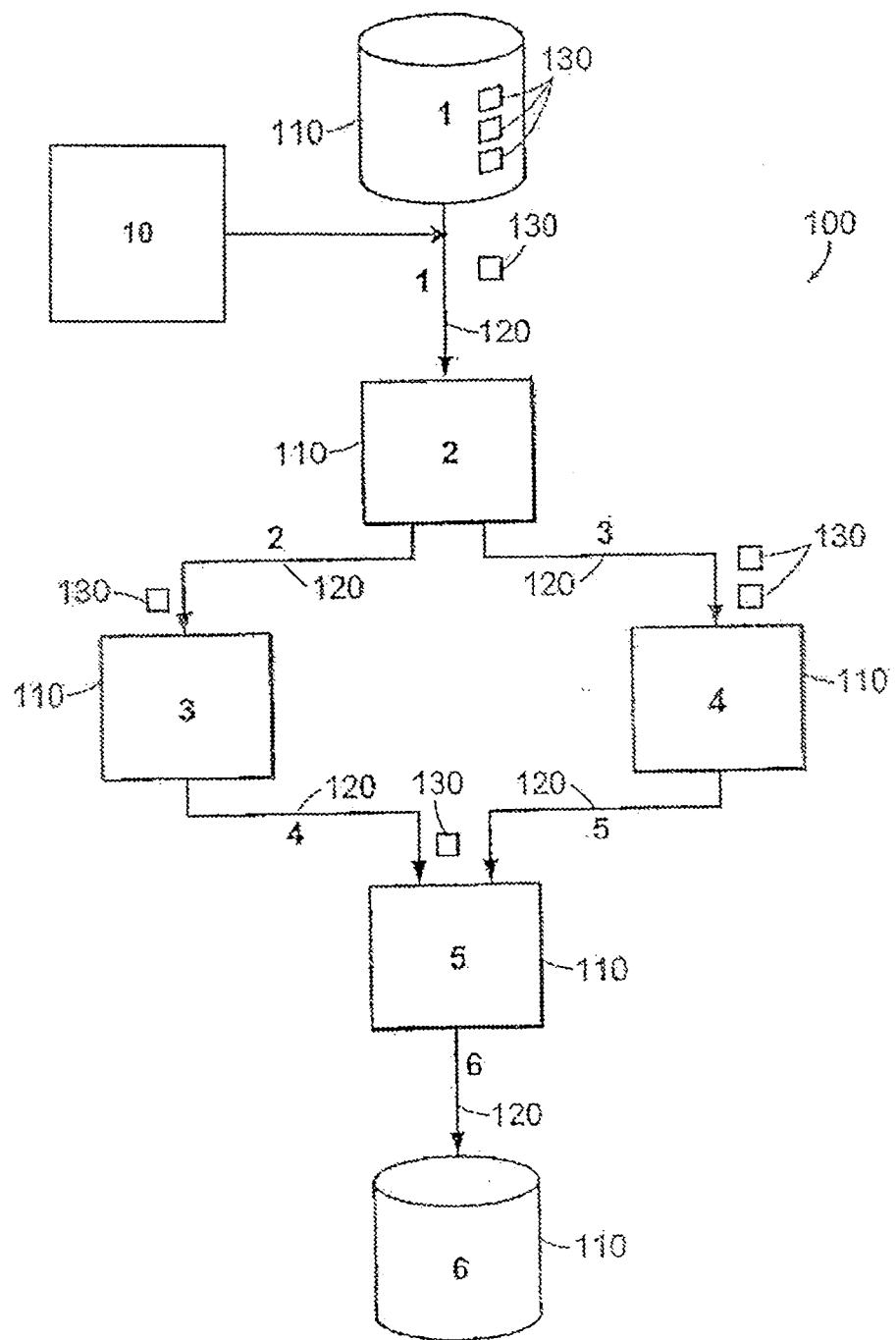


Fig. 1

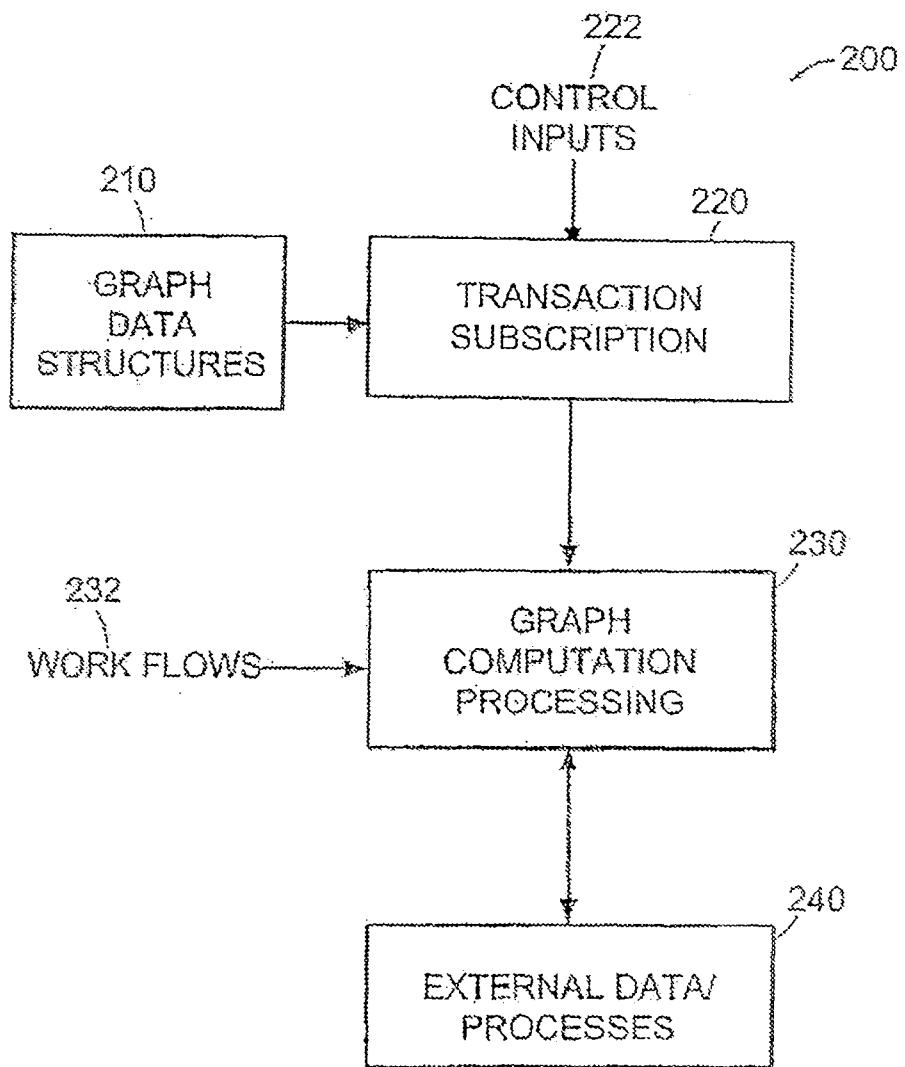


Fig. 2

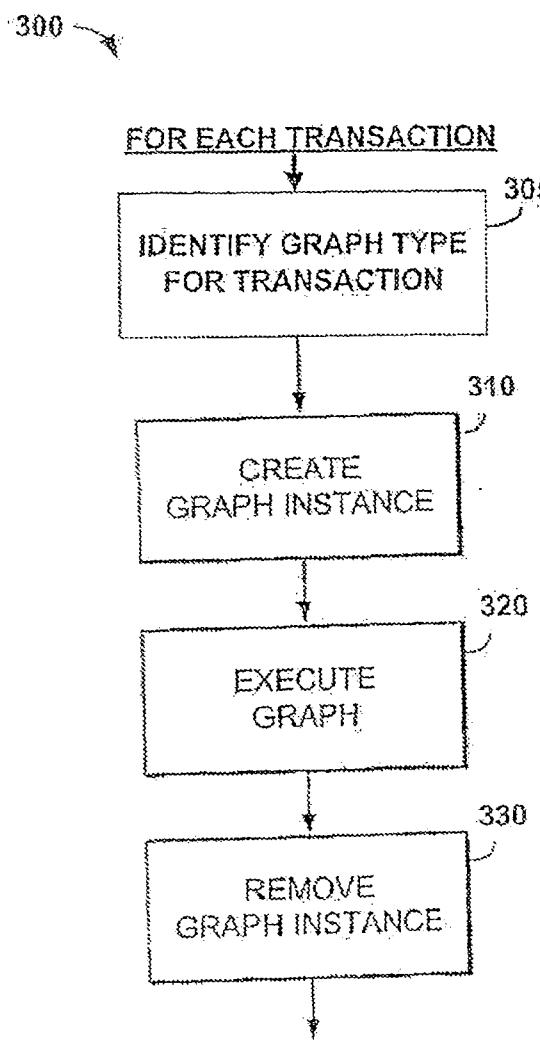


Fig. 3A

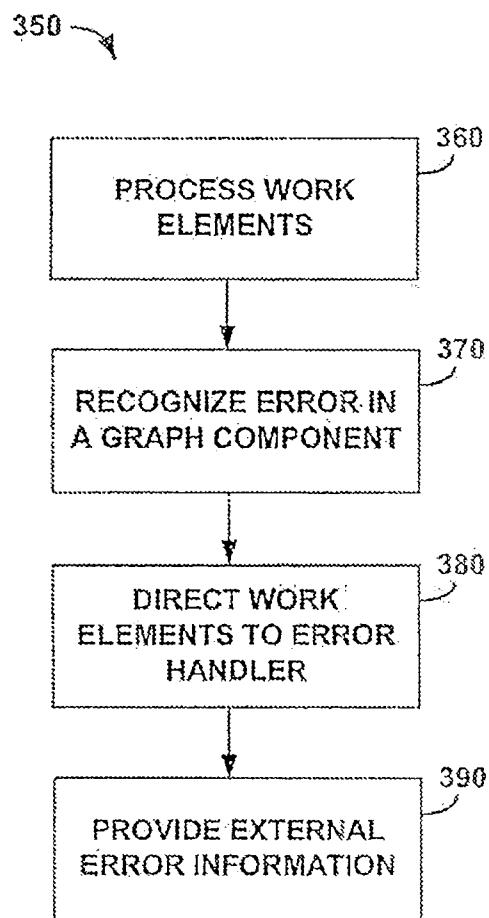
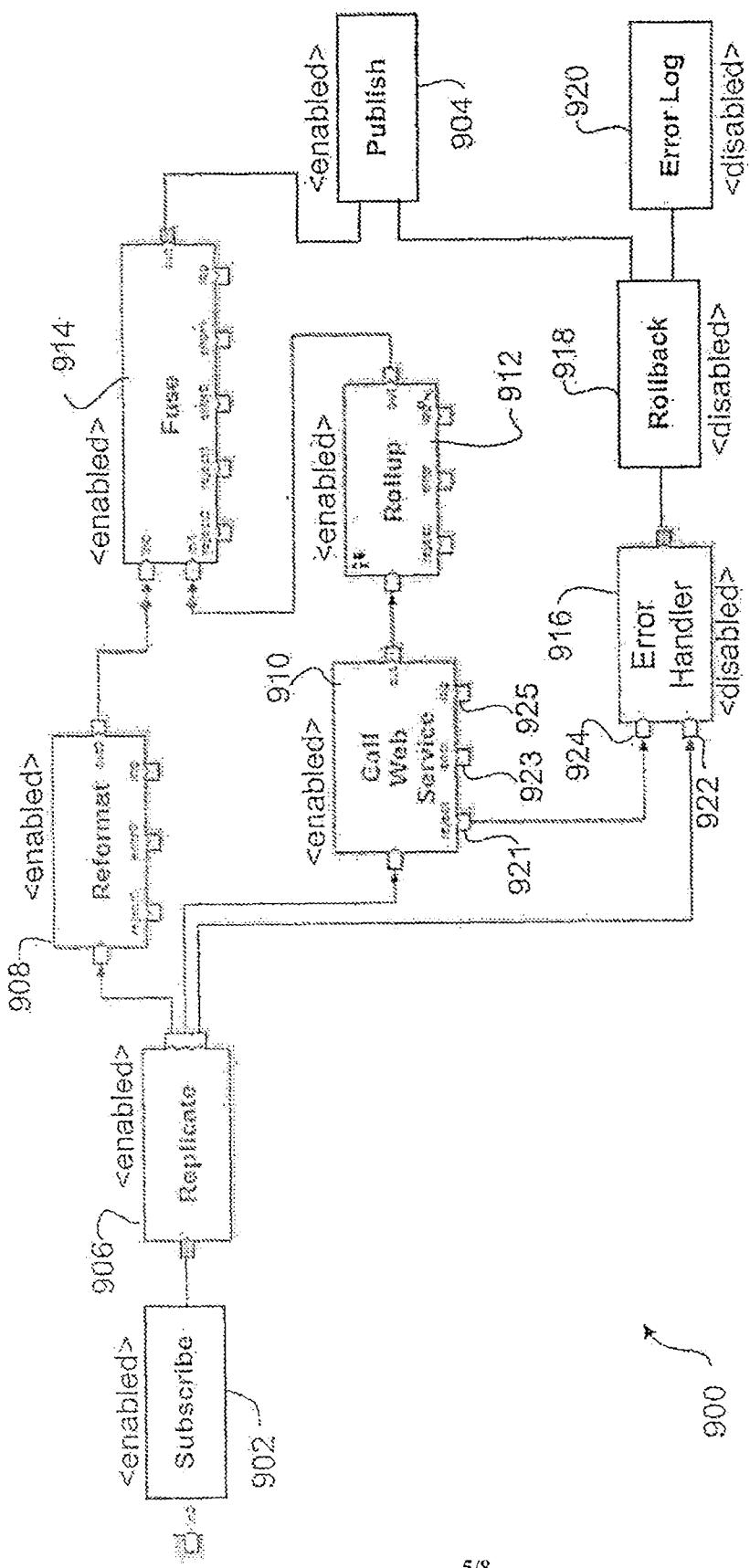
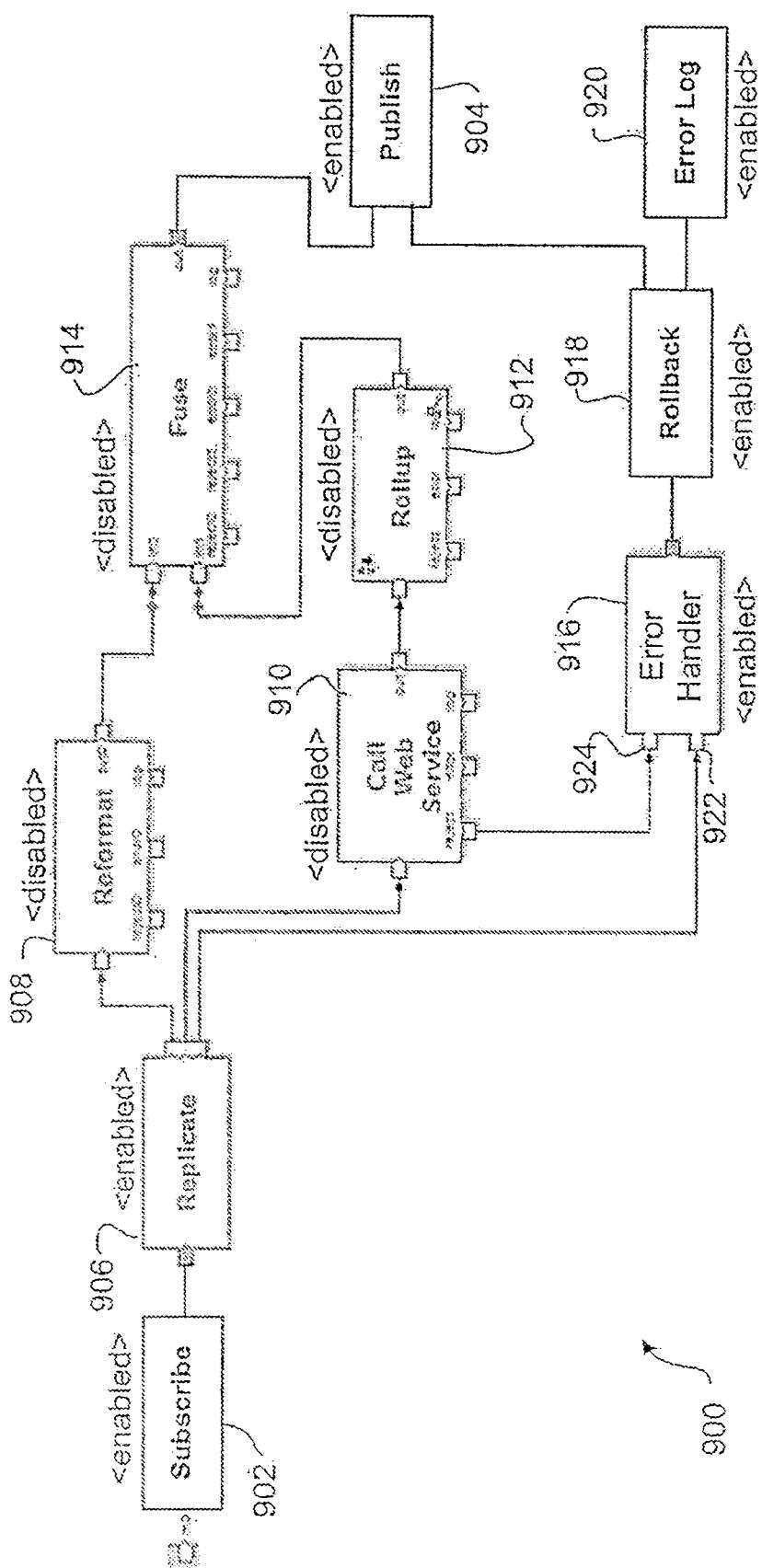


Fig. 3B



40



四
四
四

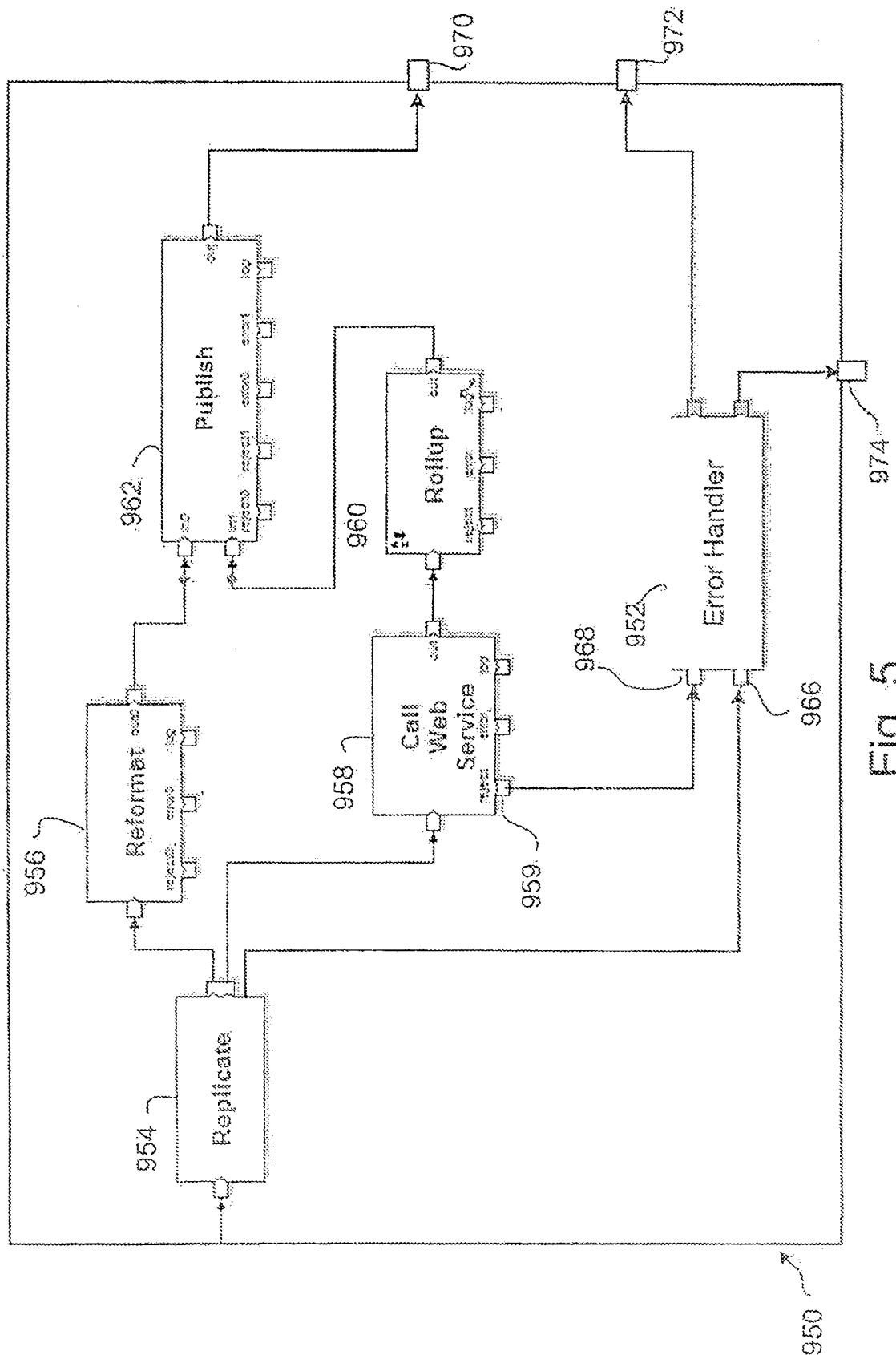
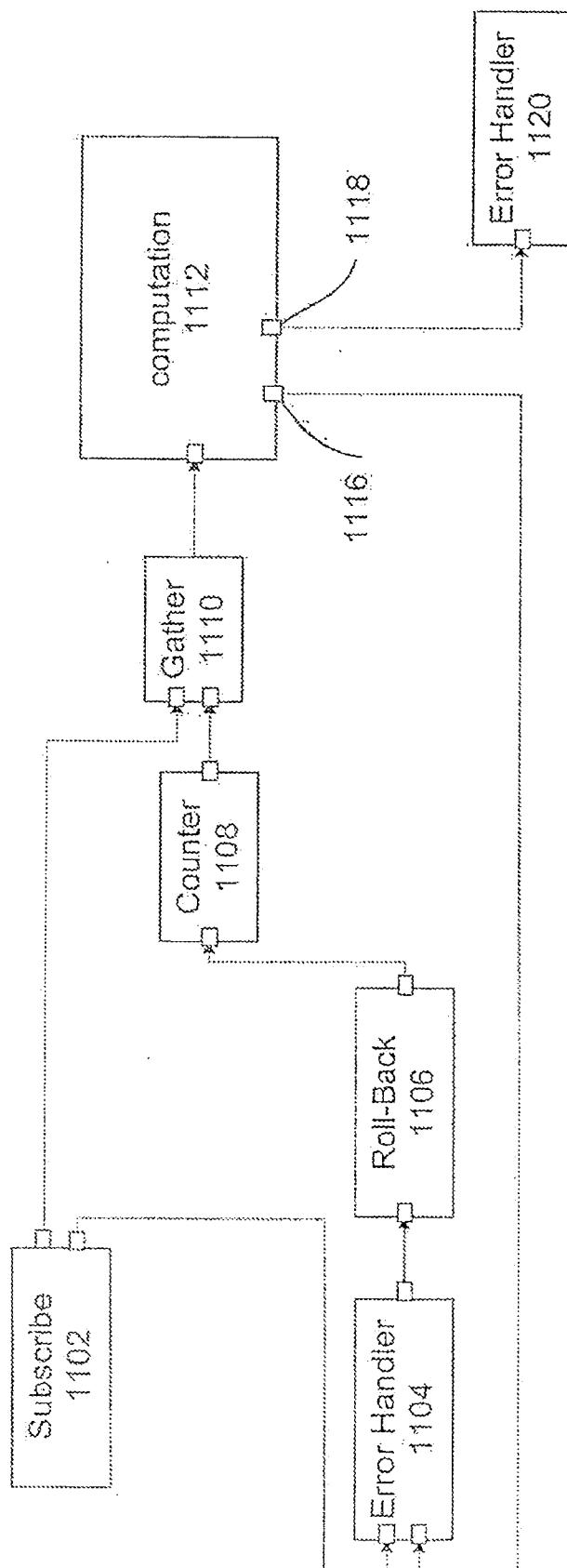


Fig. 5

974



卷六

