

The Research of Workflow Engine States

Shi Yan-rong

Electronic Information Department
Xi'an Railway Vocational & Technical Institute
Xi'an, China
E-mail: shishiyr@126.com

Hou Tao

College of Electrical and Information Engineering
Shanxi University of Science and Technology
Xi'an, China
E-mail: houtao@sust.edu.cn

Abstract—Based on the analysis of the workflow patterns, this paper designs the active state of the workflow, and puts forward the idea of the logical design of the engine driven. It can conveniently define the information model of the workflow based on relational database systems, it can effectively reduce the difficulty on the technical development of the workflow engine and take full advantage of existing relational database system to realize the information system that has the nature of the workflow and reduce the cost of the implementation.

Keywords-Workflow engine; Status; conditions of state changes; Engine-driven logic

I. INTRODUCTION

Workflow technology is used in the telecommunications industry, the field of software engineering, manufacturing, financial services, banking, scientific experiments, the field of health care, the shipping industry and office automation, etc. Many products (such as ERP, CRM, call centers, etc.) have formed the function of workflow management^[1]. The definition of the workflow given by the Workflow Management Coalition: workflow a type of business process that can fully or partially performed automatically, it can transfer and execute among the different executors based on a series of process rules, documents, information or tasks. In addition, many scholars have also tried to give some description to the definition of the workflow, but as everyone has the different understanding of the workflow, it is difficult to have a uniform definition^{[2][3][4][5]}.

The workflow engine is a set of data models and software that is used to provide support when the Workflow Management System is in definition and provide Interpretation and implementation services when it is running, it is the core part of the workflow system in a certain sense. But if we develop the Workflow Reference Model that given by the Workflow Management Coalition systematically, the workflow products are too complex, a lot of small and medium-sized enterprises are difficult to accept for its higher costs and longer cycles. Therefore, it is a good solution for small and medium-sized enterprises to research and develop a lightweight workflow engine to implement workflow technology; it can also apply more technology and automation technology in the business of small and medium-sized enterprises to improve the efficiency of their operations. In this way, they can compete with the larger companies in the same horizontal level and not be deprived of the right to use the advanced technology

to enhance their operational capacities.

This paper takes a relational database as the driven database of workflow engine, redesigned modeling elements of workflow, and mainly focuses on the node status and conditions of the state changes. It separates the process logic from the system and put the work directly onto the user's desktop by the workflow engine to achieve automation of process control. As the engine driven state-based, so it can define the workflow data model based on a traditional relational database; All data models of workflow engine are expressed by the relational structure; A variety of program logic of the control workflow engine operation is also realized by procedure package, trigger and other mechanisms that provided by the conventional relational database management system to achieve. At the same time, Concurrency Control in the affairs is also realized by the mechanism provided by the database system. Application developers can embed it into the application system to achieve the information systems with the nature of the work flow.

II. THE DESIGN OF WORKFLOW MODELING ELEMENTS

Use the design of 10 workflow modeling elements to represent the corresponding activities.

- Start: the first activity of the business process.
- judgment: The judgment should in accordance with the set conditions, but not for a specific part of the business, and choose the successor route based on the judgment result.
- Branches: not for specific business sectors, this event will also derive a number of follow-up activities.
- Or branch: not for specific business sectors, according to the rules, this activity will select a branch from several branches to push back the process, and the other branches will be terminated and abandoned.
- Polymerization: not for specific business sectors, the task flowing through here will be synchronized polymerization. Before the flow of the follow-up activities, we could only wait until all the branches have executed there.
- Or polymerization: not for a specific part of the business, the task flowing through here will be or aggregation synchronization. As long as there have a branch to meet the specified conditions it will flow to the follow-up activities.

- Interactive processing: The Interactive processing corresponds to the actual links of the business, and it corresponds to the actual application logic in the foreground. It requires the staff to take part in it to complete the activity, and in all types of activities only interactive processing activities requires the actual participation of staff.
- Automatic processing: automatic activities also correspond to the practical aspects of the business, but the actual application logic is in the background, automatically call the appropriate body of the execution of the application logic to complete by the workflow engine invokes the execution of the application logic automatically to complete the activity.
- Sub-processes: When the main flow run to this activity location, it will activate a sub-process.
- End: termination of the activity indicates the end of the corresponding business process, not for a specific business part. In the graphical system of the modeling tools, these activities are represented by different icons, it is also called node, and the nodes are connected by directed arcs.
- Arc: a directed arc connects two nodes, and indicates the direction of the control flow and data flow. Some directed arcs have conditional expressions, they can calculate the value of the expression when the workflow runs, and the corresponding changes will occur when the result is true.

Using the ten types of nodes defined above and the combination and nested of the directed arcs, we can express the business processes that we needed. In the process, the real business processing is actually the interactive process nodes (manual) and automatic processing nodes (automatic), and other nodes are just used to form a control structure, it is not targeted at specific business sectors, so it is also called the control node (except the start and end nodes).

The start node of the main process has only one output arc, and the end node has only one input arc; the nodes of Interactive processing (manual) and automatic processing (automatic) have one input arc and one output arc; But the control node has multiple input arcs or multiple output arcs; predecessor and successor nodes of control nodes is either the interactive processing nodes or the automatically processing nodes.

III. NODE STATE TRANSITIONS

As it is driven by the changes of the node states, it requires designing the node states and its changes combines with the workflow patterns. According to requirement, we set several node statuses as follows, and also give the conditions of the state transition.

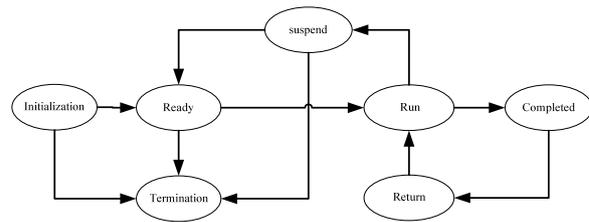


Fig1 Node state transition

A. Node states

1) Initialization: the node has been generated but has not been activated (e.g. start condition of the activity is not satisfied).

2) Ready: the node has been activated, it satisfies the running conditions, as long as the business process reaches the node it can start running immediately.

3) Run: business processes is processing at the node.

4) Completed: the node has dealt with a workflow. According to the additional information of the node, the management system of the workflow inquiries and activate the next node which is meet the start conditions.

5) Suspend: The node is temporarily suspended because of the unusual circumstances or error, the activity is in a quiescent state, and the node will not participate in the business process operation.

6) Return: the successor node requests re or supplementary treatment.

7) Termination: It stops midway by the system.

B. The conditions of the node state changes

1) Initialization → Ready: After the node activation conditions are met, the engine sends the tasking notice to the to-do people, and the node state converts to the ready state. And at the moment the to-do people has not received activated processing tasks yet.

2) Ready → Run: The to-do people receive the processing tasks and processes, the node state converts to the running state.

3) Run → Completed: the node has finished processing, after the processing result is submitted to the system, the node status converts to complete.

4) Run → Suspend: In the running process, when the node is unable to run temporarily for some reason, it will enter the suspend state and the activities becomes static temporarily.

5) Suspend → Ready: When the problem that caused suspend solved, the node re-equipped operating conditions, and the state converted to ready state.

6) Completed → Return: the predecessor node is already completed, but the successor node finds the information conveyed by the predecessor does not comply with the requirements of the current node, thus the successor node requires to fallback processing, the current node status converts to returned state.

7) Return → Run: When the back processing request of the successor node is accepted by the current node transitor,

it will have re or supplementary processing and the current node status converted from the return state to the running state.

8) Termination: As the follow-up process does no longer need the running this node, the system forces the node status converted to terminate the state.

IV. THE DESIGN CONCEPT OF THE ENGINE-DRIVEN LOGIC

When the workflow model is instantiated, all the status of the nodes are "initialized" status, Converting the start node status to "Ready" state to activate the running of the entire process, the nodes that follow the start node will continue to be activated and the node status converts to the "ready" status by the scheduling of the workflow, and its status converts to "Run" when the node starts processing business. After the business processing, the node will return the results of the implementation to system, and the node status converted to "complete", then in the light of the result, the engine dispatches the successor node.

According to the description above, we can see the running process of the workflow model instance as a process of the change of each node status. Taking advantage of the stored procedures, triggers and other mechanisms that provided by the relational database management system to capture the changes of the node state and drive the entire process of running combined with the corresponding processing logic. In this way, the workflow engine that designed can seamlessly embed into the traditional application development environment to reduce the pressure on the system server and client. Make full use of the management function of Unified, centralized and independent for the data by the database management system, it can play a powerful query search capabilities and digs out the real role of the database management system, but not just as a data warehouse.

A. Driven Algorithm of Nodes Tasking

When a node status changes to "Ready" state, the workflow engine will complete the operation of the prescribed operation of the current node automatically in accordance with the appropriate rules. The control node will change the status of the correlation output of the arc by the different types of nodes automatically and the pre-set conditions and provides the basis for the choosing of the subsequent flow path. For the node requiring human intervention, the engine will send do notices to the specified operator and request artificial participate in the execution of the current node. When it is completed, the node status converted to "Complete".

The processing algorithm is described as follows:

When the node status becomes "Ready" state, it collects useful information. If the node is:

1) Start: Output arc is set to "True"; Node status is set to "Completed".

2) Judgment: Seek the value of the judgment expression; According to the results, the output arc that meets the

requirements is set to "True"; and the node status is set to "Completed".

3) Branches: All output arc are set to "True"; And node status are set to "completed".

4) Or branch: According to the rules, select an output arc and sets to true; and the node status is set to "completed".

5) Polymerization: If there is an input arc set as "False", the status is reset to "initialize"; if all input arcs are "True": the output arc is set to "True"; node status is set to "Completed".

6) Or polymerization: If the input arc is set as "False", its precursor node status should set to "Terminate"; output arc is set to "True" and node status is set to "Complete".

7) Interactive processing: Send the do notices to the to-do people. (The node state converts to "Run" when the to-do people receives the task; the node status converts to "Finish" when the processing is completed)

8) Automatic processing: It invokes the specified executable of the application logic to process automatically. (When the processing is completed, the node status is set to "completed".)

9) Sub-processes: If the rule is the wait mode of the main process: the node status is set to "Suspend"; The "Start" node of the successor is set to "Ready". If the rule is the operating mode of the main flow : all the output arcs are set to "True"; and all subsequent nodes status are set to "Ready".

10) End: If there are output arcs, then the output arcs are set to "True"; node status are set to "Completed".

B. Algorithm of Forward Control-driven

When a node status changes to "Completed" status, the workflow engine activates designated successor node automatically according to the result of setting of the output arc.

The processing algorithm is described as follows:

When the node state converts to "Complete", it collects useful information; if the node is:

1) Start: Check the successor node, its status is set to "Ready".

2) Judgment: Find successor nodes whose output arcs are set as "True", their status are set to "Ready"; and output arc reverts to "False".

3) Branch: Find all the successor nodes, their status are set to "Ready".

4) Or branch: Find successor nodes whose output arcs are set as "True", their status are set to "Ready"; And find the successor node whose output arcs are set as "False", their status are set to "Terminated".

5) Polymerization: Find the successor nodes, their status are set to "Ready".

6) Or polymerization: Find successor nodes whose output arcs are set as "True", their status are set to "Ready".

7) Interactive processing: The output arc is set to true; check successor node, its status is set to "Ready".

8) Automatic processing: The output arc is set to true; check successor node, its status is set to "Ready".

9)Sub-processes: Find successor nodes that are not set as "Start", their status are set to " Ready "; all output arcs are set to "True".

10)End: If there are output arcs: If the successor node status is set as "Suspend", the successor node status is set to "Completed"; otherwise, the successor node status is set to "Ready".

V. THE HANDING OF THE CONCURRENCY ISSUES ON WORKFLOW ENGINE

As the multiple clients' access shared resources simultaneously in the concurrency runtime, it is inevitable to occurred "dirty data" if the concurrency control logic is not good designed. The present researches on the concurrency control are Altruistic agreement, the method of NT / PV model-based as well as the processing method of the semantic-based, but these control methods are not ideal in the practical application.^{[6] [7]}

In this paper, the realization of the workflow engine status and its change mechanism depend on the internal operation mechanism of the database management system, therefore, the processing of the multitasking concurrent issues can take advantage of concurrency control management of the database management system itself to handle, and it can be relatively easier to solve the conflict appeared in the parallel processing flow.

VI. SUMMARIES

With the continuous development of Web service technology, more and more Web Service based workflow solutions have been proposed. But in the process of implementation and application, it is difficult to integrate the legacy systems effectively. Therefore, the companies have to abandon the original information systems and result in a tremendous waste of resources. The workflow engine state and its change mechanism researched in this paper can provide support on the workflow mechanism for the development of the lightweight workflow engine; it can construct a variety of application system with the workflow features. As the engine is driven by the node status, relational structure can be used to express the information model in the workflow engine, in this way, it reduces the technical difficulty of the development and workload from the technical point, so it can be very convenient for the enterprises to use the existing relational database system.

REFERENCE:

- [1]Irene VandeerFeesten,Hajo A. Reijers. A Human-Oriented Tuning of Workflow Management System[J]. Lecture Notes in Computer Science, 2005, 3649:80-95.
- [2] Yuan Qing-ke,Zhao Ru-jia. The Study of Theory and Method for Design Task Scheduling in Concurrent Engineering Environment[J]. CHINESE JOURNAL OF COMPUTERS 2000, 23(4). Vo1.23 No4:2000.
- [3]A. Manolescu,E. Johnson. A Micro-Workflow Component for Federated

Workflow. 2000, <http://micro-workflow.com/research>

[4]Workflow Management Coalition. Terminology and Glossary. Wfmc-TC- 1011, 1996, <http://www.wfmc.org>

[5]D.F.Ferguson,Web Services Architecture:Direction and Position Available at <http://www.w3.org/2001/03/WSWS-popa/paper44>. April 2001.

[6]Ren Zhi,Han Wei-hong,Jia Yan. The Research of Transaction Management Technology in Loosely—Coupled Distributed Environments[J].computer science.2003,30(10):336-339

[7] Yi Xin-ying,Song Bao-yan,Yu Ge.Workflow Concurrency Control Mechanism Based on Semantics[J].computer science.2003,30(10):182-186