Journal of Network and Computer Applications 35 (2012) 97–110

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

A model-driven workflow fragmentation framework for collaborative workflow architectures and systems

Kwanghoon Kim *

Collaboration Technology Research Lab, Department of Computer Science, Kyonggi University, South Korea

ARTICLE INFO

Article history:
Received 3 September 2010
Received in revised form
11 January 2011
Accepted 29 March 2011
Available online 13 April 2011

Keywords:
Structured ICN workflow model
System-driven workflow fragmentation
Model-driven workflow fragmentation
Collaborative workflow architecture
Workflow resource allocation
Collaborative computing paradigm
Cloud computing environments

ABSTRACT

This paper focuses on a workflow distribution methodology for rationally deploying workflow models onto a distributed workflow system running on cloud computing environments, and we particularly lay a stress upon that those workflow systems operable on cloud computing environments are dubbed collaborative workflow systems, which are not only built upon the collaborative workflow architectures proposed in the paper, but pursuing the so-called collaborative computing paradigm characterized by focusing collaboration over cloud computing environments. The essential idea of the workflow distribution methodology is about how to fragment a workflow model and how to allocate its fragments to each of the architectural components configuring the underlying collaborative workflow architecture and system. As a reasonable solution to realize the essential idea, the paper proposes a model-driven workflow fragmentation framework, which provides a series of fragmentation algorithms that semantically fragmentate a workflow model by considering the semantic factors – performer, role, control-flow, data-flow, *etc.* – of the ICN-based workflow model as fragmentation criteria. The algorithms are classified into the vertical fragmentation approach, the horizontal fragmentation approach, and the hybrid approach of both. Conclusively, this paper conceives a possible set of collaborative workflow architectures embedding the collaborative computing paradigm, and describes the detailed formalism of the framework and about how the framework works on those collaborative workflow architectures and systems.

1. Introduction

A workflow management system (WfMS) is defined as a system that partially or fully automates the definition, creation, execution, and management of work procedures (workflows) through the use of software that is able to interpret the workflow definition, interact with workflow participants and invoke the use of IT tools, scripts and applications; and the majority of WfMS's infrastructures (Martin et al., 2008) for deploying and enacting workflow services, so far, has been pursuing the distributed computing paradigm, which is conceptually focusing on parallelization, distribution and sharing over distributed computing environments such as server–client, clustering, grid and P2P computing environments. However, as the cloud computing environment is recently hot-issued as an emerging new and powerful computing environment, it is increasingly needed to adopt the cloud computing environment as WfMS's infrastructure; consequently, it becomes a catalyst for the emerging concept of cloud workflow architectures and systems. More precisely speaking, in the distributed computing paradigm based on the cloud computing environment, the conceptual focus is not on parallelization, distribution and sharing but on collaboration (Liu et al., 2009), and so, the paradigm has been particularly dubbed *collaborative computing paradigm*,[1] which is becoming a conceptual basis for the so-called *collaborative workflow architectures and systems*[2] as well as a technological background of this paper.

In configuring a distributed workflow architecture, it might be very important to determine criteria of how to fabricate its architectural components and how to deploy and allocate workflow models and workflow resources to them, as well; those works of configuring, deploying and allocating workflow models are called a workflow distribution methodology. There are two styles of workflow distribution methodologies: system-driven and model-driven. Almost all of the traditional distributed workflow architectures aiming at operable on server–client, cluster, grid or P2P computing environments are apt to take the system-level functional aspects of workflow systems as the architectural criteria; their primary goal is to ensure the computing power for

* Tel.: +82 10 5231 9679; fax: +82 31 249 8949.
E-mail address: kwang@kgu.ac.kr
URL: http://ctrl.kyonggi.ac.kr

[1] This terminology has been defined by the co-editors of this special issue.
[2] The author has defined this terminology according to the conceptual definition of collaborative computing paradigm.

1084-8045/$ - see front matter
doi:10.1016/j.jnca.2011.03.029

enacting workflow models with a certain level of transparencies in parallelization, distribution and sharing of workflow resources, and so we would say that it is the style of system-driven workflow distribution methodologies (Tomaz et al., 2009), because configuring architectural components and fragmenting and deploying workflow models might be done model-independently. While on the other hand, in configuring a collaborative workflow architecture based upon cloud computing environments, the primary architectural theme is shifted into collaboration, not distribution as mentioned before, and so it is crucial to consider not only how to directly reflect the collaboration aspects on workflow models into the architectural collaborations, but also how to deploy and allocate workflow models and resources onto those collaborative architectural components. Consequently, those methodological works have to be done model-dependently, and so it is called the style of model-driven workflow distribution methodologies. Furthermore, in the style of model-driven workflow distribution methodologies, the most essential work is to devise a framework for fragmenting a workflow model and deploying its fragments onto the underlying collaborative architectural components. As the scope of this paper, we propose a model-driven workflow fragmentation framework, which provides a series of workflow fragment algorithms that fragmentates a workflow model into a group of workflow fragments and distributes them onto the architectural components of an underlying collaborative workflow architecture.

Summarily, this paper focuses on the model-driven workflow distribution methodology consisting of the architectural configuration step configuring model-driven collaborative workflow architectures, and the model fragmentation step fragmenting a workflow model by one of the model-driven workflow fragmentation techniques and deploying its fragmented models to the corresponding architecture. The framework is based upon the information control net (ICN) methodology (Ellis, 1979) to represent workflow models. The ICN-based workflow modeling methodology was originally developed to describe and analyze information flows by capturing several entities within work procedures, such as activities, roles, actors, activity precedences, applications, and relevant-data/repositories. Note that it has been used within actual as well as hypothetical automated offices (1) to yield a comprehensive description of activities, (2) to test the underlying office description for certain flaws and inconsistencies, (3) to quantify certain aspects of office information flow, and (4) to suggest possible office restructuring permutations. For the sake of algorithmic simplicity, it is assumed that the ICN-based workflow model used in the paper preserves the structural properties – proper nesting and matched pairing properties (Park and Kim, 2008) – as a structured workflow model (Liu and Kumar, 2005). Once a structured workflow model is defined by using the ICN methodology, it is fragmented and deployed over its corresponding collaborative workflow architecture according to the proposed framework.

In the workflow literature, there has not been existing any types of the model-driven workflow fragmentation methods, as yet. So,the remainders of this paper describe the details of the architectural styles to be devised as the model-driven collaborative workflow architectures, and exemplify that the model-driven workflow fragmentation approaches are able to handle all of the possible fragmentation cases through the concepts of vertical and horizontal fragmentation methods. That is, the next sections present the meta-model of the structured workflow model with graphical and formal notations, formalize the model-driven workflow fragmentation framework that gives a detailed description about how to fragment a workflow model—vertical fragmentation *vs*. horizontal fragmentation, and illustrate those approaches and their detailed descriptions as well as related algorithms with some examples. Finally, the paper finalizes with an example through an operable architecture with the role-based fragmentation algorithm in order to show how the fragmentation approaches are feasibly applicable in realizing a certain style of collaborative workflow architecture and system.

2. ICN-based structured workflow model

In this paper, basically it is assumed that the information control net methodology (Ellis, 1979) is used to represent structured workflow models (Liu and Kumar, 2005) to be deployed on the collaborative workflow architectures proposed in the paper. So, this section describes the details of the structured workflow modeling concept as a theoretical basis of the model-driven workflow distribution methodology defined in this paper.

2.1. Workflow meta-model

In describing a structured workflow model, the literature has been using the basic workflow terminology—workflow procedure, activity, job, workcase, role, actor/group, and invoked application including web services. These terms become the primitive entity types to compose workflow models, and also they have appropriate relationships with each other as shown in Fig. 1. These entity types and their relationships play an important roles as the fundamental criteria for the model-driven collaborative workflow architectures and their fragmentation

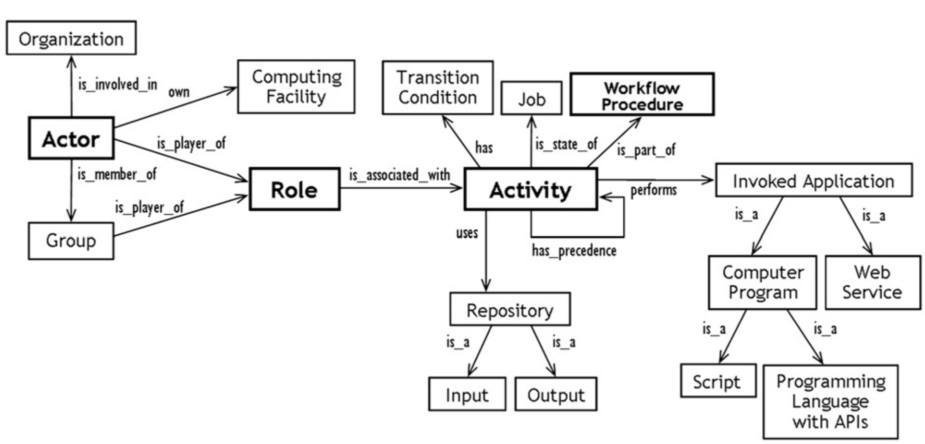

Fig. 1. The workflow meta-model.