Design of an Autonomous Framework for Efficient Large Scale Management of Next Generation Web Service Mashups

Anna Hristoskova, Filip De Turck

Department of Information Technology, Ghent University - IBBT,
Gaston Crommenlaan 8 bus 201, B-9050 Ghent, Belgium,
{anna.hristoskova,filip.deturck}@intec.ugent.be

Abstract. The objective of this PhD research is the study and development of an autonomous framework focusing on the dynamic composition and deployment of service mashups. By enriching the available building blocks with semantic descriptions, automatic composition of new service mashups is achieved through the use of planning algorithms. The composed mashups will be automatically deployed on the available resources making optimal use of bandwidth, storage and computing power of the network and server elements. This system will also be extended with dynamic recovery from external errors, such as network failure or services delivering erroneous results. The last step will encompass the automatic extraction of new knowledge from request patterns, behavior and feedback of users for the optimization of future requests.
A number of relevant case studies will be defined in the domain of eHealth and Web 3.0, since the use of workflows for efficient data processing in these areas is very important.

1 Introduction

A major trend in the field of telecommunication research is the need for scalable processing of data. The main causes are the deployment of sensor networks and the growing importance of context-aware applications. Sensor networks consist of large quantities of sensors, placed in a building at regular distances on the walls, for monitoring temperature, the presence of persons or goods. Context-aware applications cover software components interacting with each other, so that the application behavior reacts on the current context, the user's preferences and the state of the network to which the user is connected.
An important aspect of the scalable processing of data is the optimal structuring of the available information. In recent years the Semantic Web [1, 2] is utilized to overcome this issue offering more access not only to content but also to Web services. The use of Web resources based on keywords and the provided service content is feasible [3]. This allows for easy sharing of reusable knowledge. In addition, ontologies not only facilitate the communication between people and machines, but also between machines. Since the information is presented in a

formal way, a computer can reason about it and new knowledge can be inferred. In this way the wide range of components used in sensor networks and context-aware applications can be dynamically and more efficiently combined through semantic reasoning, accomplishing new service mashups.
A service mashup is a new service that combines functionality or content from existing sources. These sources can be Web services, software components capable of being accessed via standard network protocols such as but not limited to SOAP over HTTP. The mashup construction is realized due to a workflow description: definition of how the basic building blocks interact. Today there are a number of popular standards and implementations, such as BPEL [4], BPMN, WFF, which define workflows. However they still exhibit a number of shortcomings: no automatic or dynamic deployment support, limited reliability guarantees, etc. The main goal of this PhD research is to study in depth these topics and design a framework focusing on the following aspects:

- Reasoning algorithms for automatic composition of service mashups satisfying quality of service constraints and requirements.
- Deployment strategies of the service mashups on the available resources making optimal use of bandwidth, storage and computing power of the network and server elements, extended with dynamic recovery from external errors.
- Optimization of future request through the automatic extraction of new knowledge from request patterns, behavior and feedback of users.

2 Autonomous Semantic Composition

Initially, our research will focus on the development of reasoning algorithms for the automatic composition of new service mashups [5]. For this purpose the Semantic Web will be studied that enables users and software agents to automatically discover, compose, invoke, and monitor Web resources offering services, subject to specified constraints. Semantic languages such as OWL-S, WSMO, SWSO, provide a semantic description of Web services specifying service inputs, outputs, preconditions and effects, and non-functional properties. Thanks to their semantic description, Web services can be compared and matched accomplishing a mashup composition, as shown in figure 1, realizing user defined goals.

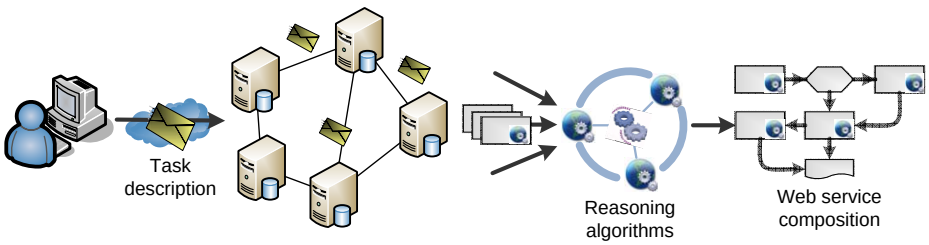

Fig. 1. Automatic composition of Web services

As there is no need for user intervention, this process can be fully automated. A number of reasoning algorithms are already developed that take into account certain quality constraints (execution time, cost) of the available building blocks used in the Intensive Care Unit of Ghent University Hospital. Semantically enriched medical services are automatically combined accomplishing a medical decision support system monitoring the patient's condition and registering critical changes in his state. This approach is still being refined with more comprehensive HTN planning techniques.

3 Resource Efficient Service Deployment

Figure 2 illustrates the autonomous deployment of abstract workflows on a set of available resources [6]. This part of the research requires the linking of abstract service definitions to concrete service instances, adaptation in the communications layer, selection of the appropriate workflow engines and possible creation of additional workflow code. Optimization algorithms will be developed to ensure translation that makes the best use of the available network infrastructure.
These algorithms will be further extended to support dynamic deployment of the workflows [7]. This is necessary to avoid down-time but it may be also crucial to modify the deployment so as to anticipate changes in the network situation [8]. Triggers can be a change in the network topology, or an alteration in the nature of the workflow distribution.
As a last step, the design of a framework will be studied for the development of large-scale context-aware applications. We will develop procedures for automatically extracting new knowledge from request patterns, behavior and feedback of users. This knowledge will be utilized for the optimization of the composition and deployment process for future requests.
We will closely look into a Web 3.0 application addressing the enrichment of business components and services (such as sensors) with semantics, reasoning, autonomous behavior (agent technology), and distributed data management. We will proceed with the selection and composition of these building blocks into an application and providing assurance about the performance of the composite services on the available resources. The most suitable composition and deployment techniques for this domain will be examined.

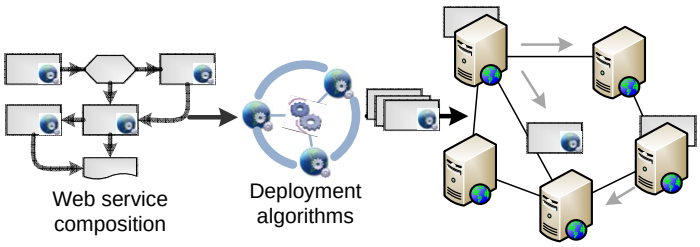

Fig. 2. Deployment of a service mashup

4 Summary

This PhD research focuses on the study of an autonomous framework for dynamic composition and deployment of the building blocks of service mashups. Based on semantic descriptions of Web services and resources, reasoning algorithms are developed for automatically composing new service mashups realizing defined goals. In addition, deployment strategies are designed for the distributed deployment and execution of the service mashups on the available resources. This system is optimized for the dynamic response to changing context such as failure or overload of network elements or Web services. Furthermore, techniques will be studied to take into account trends in user and resource behavior, in order to optimally design context-aware service mashups. Important application areas such as eHealth and Web 3.0 will be implemented to verify this framework.

Acknowledgement

A. Hristoskova would like to thank the Special Research Fund of Ghent University (BOF) for financial support through her PhD grant. F. De Turck is funded by a postdoctoral grant of the Fund for Scientific Research, Flanders (FWO-V).

References

1. L. Berners, J. Hendler, O. Lassila, "The Semantic Web: A new form of web content that is meaningful to computers will unleash a revolution of new possibilities", Journal of the Scientific American, 284(5):34.43, 2001.
2. H. Stuckenschmidt, M. Klein, "Structure-based partitioning of large concept hierarchies", In Proceedings of the 3th International Semantic Web Conference (ISWC2004), 2004.
3. J. Keeney, D. Lewis, D. O'Sullivan, "Ontological semantics for distributing contextual knowledge in highly distributed autonomic systems", Journal of Network and Systems Management, 15(1):pp. 75-86, 2007.
4. D. Karastoyanova, A. Houspanossian, M. Cilia, F. Leymann, A. Buchmann, "Extending BPEL for run time adaptability", In EDOC Enterprise Computing Conference, 2005 Ninth IEEE International, pp. 15.26, 19-23 Sept. 2005.
5. T. Klie, F. Gebhard , S. Fischer, "Towards automatic composition of network management web services", Integrated Network Management, pp. 769-772, 2007.
6. C. Wang, Z. Zou, Q. Ma, R. Fang, H. Wang, "A comprehensive semantic-based resource allocation framework for workflow management systems", IEEE Network Operations and Management Symposium (NOMS 2008), pp. 831-834, 2008.
7. M. Lee, H. Yoon, H. Shin, "Supporting dynamic workflows in a ubiquitous environment", In Multimedia and Ubiquitous Engineering, 2007. MUE .07. International Conference on, pp. 272.277, April 2007.
8. M. Gajewski, M. Momotko, H. Meyer, H. Schuschel, M. Weske, "Dynamic failure recovery of generated workflows", In Proceedings of the 16th International Workshop on Database and Expert Systems Applications (DEXA.05), 2005.