A BPA-Based Approach for the Identification of SOA Candidate Software Services

Rana Yousef
Department of Computer Information Systems, KASIT
The University of Jordan
Amman, Jordan
rana_m1980@yahoo.com

Mohammed Odeh
Department of Computer Science, FET
The University of the West of England
Bristol, United Kingdom
Mohammed.odeh@uwe.ac.uk

Abstract—**The task of identifying candidate services is one of the main activities in developing service-oriented models. The current service identification approaches exhibit some limitations, where they are either too complex to be adopted, too simple to satisfy all Service Oriented Architecture (SOA) principles, or are theoretical approaches that are far from automation. In this paper we propose a novel service identification approach, which uses a simple and straightforward criterion for the service identification process based on the Riva business process architecture. The approach is demonstrated in this paper using the UWE's CEMS faculty administration case study. Evaluation revealed that this approach simplifies the process of identifying services while satisfying SOA principles. In addition, this newly introduced service identification approach adds a further dimension to the Riva BPA method to extend its role and application for the development of service oriented systems.**

Keywords-Service Identificaion; Service Oriented Architecture; Business Process Architecture; The Riva method.

I. INTRODUCTION

Service-Oriented Architecture (SOA) is becoming the mainstream for providing efficient and agile business solutions that can keep up with changes demanded by the business world. Service identification is one of the main activities in developing service-oriented applications. Errors made during this identification task can be propagated through to the detailed design, implementation and verification activities. Accordingly, the selection of an appropriate method for identifying services is an essential requirement to the development of service oriented architectures [8, 11].

Top-down service identification approaches are the most thoroughly investigated ones, where the business process is subdivided into sub-processes or decomposed into granular activities and tasks to represent services according to certain criteria. However, the complex relations between business activities that must be examined while analyzing the organization's business process models increase the overall complexity of such approaches. Business process architecture, on the other hand, provides a higher abstraction level that conforms to SOA. So, it would be more logical to identify services (i.e. SOA building blocks or service boundaries) from the process architecture rather than the process models. In this paper, we introduce a new service identification approach based on the Riva Business Process Architecture (BPA) in order to raise the abstraction level required to identify services to a level that conforms to SOA, and hence simplifying the identification process.

The rest of this paper is organized as follows; section 2 presents current work related to service identification. Section 3 briefly explains the Riva method for realizing an organization's business process architecture. Section 4 explains the new service identification approach demonstrated using the UWE's CEMS faculty administration case study. Section 5 evaluates the introduced service identification approach in terms of its satisfaction of the simplicity criterion as well as its conformance with SOA principles. Finally, section 6 concludes the paper.

II. RELATED WORK

The literature provides a lot of work in service identification approaches ranging from top-down to bottom-up. Service identification approaches make different use of process models; some deduce services from process models systematically as a meaningful representation of business processes, others restrict themselves to formulating general guidelines for identifying services. For example Klose et al [11] proposed a method for the identification of services from a business point of view based on process models. In their proposed method, functions are only implemented and provided as a service, if both business potential and technical feasibility have been verified. Their model for service identification was divided into the following three phases:

- The preparation phase: determine scope of the analysis and development framework, prepare existing process models and define stakeholders.
- The service analysis phase: conduct business analysis, identify visibility and takeover potentials and analyze IT feasibility based on SOA design principles.
- The service categorization phase: provide bottom-up definition of service-types and operations and define service-compositions.

Kim and Doh [10] suggest a formal approach to identify services at the right level of granularity from the business process model. Their approach uses the concept of graph clustering and provides a systematic approach by defining the cost metric as a measure of the interaction costs. To effectively extract service information from the business model, they take activities as the smallest units in service identification and cluster activities with high interaction cost into a task through hierarchical clustering algorithm, so as to reduce the coupling of remote tasks and to increase local task cohesion.

Kim et. al. [9] pointed out that business goals and business change factors should be analyzed because the ultimate aim of SOA is to achieve business goals and business agility in turbulent business environment. Accordingly, the authors proposed a service identification method based on goal-scenario modeling and a conceptual framework to elicit possible business changes. Traceability among business goals, business changes and identified services were also constructed in this approach.

Boerner and Goeken [2] pointed out that the lack of economic and governance aspects constitutes a problem in SI and leaves space for improvements. The authors propose a process-oriented method of service identification. This approach incorporates the business point of view, strategic and economic aspects as well as technical feasibility. By considering these aspects, it supports the governance of SOAs.

Alahmari and Zaluska [1] have pointed out that the wide range of current migration techniques for legacy systems in different implementations technologies do not address important aspects of service granularity, which affect service reusability, governance, maintainability and cohesion. In their paper, the authors proposed a novel framework for the effective identification of the key services in legacy code. The approach focuses on defining the right services based on standardized modeling languages (UML and BPMN). The framework provides effective guidelines for optimal service granularity for a wide range of possible service types.

Fareghzadeh [6] proposed a method for SI that combines different approaches and advantages and tries to avoid the disadvantages of each. This method is based on an in depth business process analysis coupled with use cases and existing assets analysis and goal service modeling.

Compared to the existing service identification approaches, our proposed approach considers an organization's business process architecture rather than its set of business process models to identify the service boundaries for the organization. This would raise the identification process to a higher abstraction level that conforms to the service oriented architecture and is accordingly expected to simplify the process.

III. THE RIVA METHOD

Ould [12, 13] proposed a methodological approach to derive process architectures from the essential entities of a business, which he later called the Riva method.

In order to identify an organization's process architecture in Riva, the following steps should be taken [12, 13]:

1) *Agree the boundary of the organisation*

2) *Brainstorm the organisations' subject matter to identify Essential Business Entities (EBEs)*

3) *Classify these EBEs that have a lifetime which is handled by, or are the responsibly of, members of the organisation as Units of Work (UOWs)*

4) *Draw a UOW diagram that depicts the dynamic relationships between UOWs.*

5) *Assume that for each UOW, there is:*

   a) *a case process (CP) that handles single instances of the UOW; and*

   b) *a case management process (CMP) for dealing with the flow of instances.*

6) *Transform the UOW diagram into a first-cut process architecture; then, use the provided heuristics to generate a second-cut process architecture.*

The Riva method was shown to be simple and easy to understand and apply [12]. The Riva-based architecture is derived from an understanding of what business the organization is in, rather than its current structure or culture. So, once the architecture is understood, it becomes apparent what is required from the IT systems supporting these processes.

IV. THE NEW SERVICE IDENTIFICAION APPROACH

From first observation, we noted that the second cut process architecture which is generated from applying the Riva method [12, 13] consists of a number of CPs and CMPs where some are related to other CPs and/or CMPs forming some set of clusters. We have called these RPA clusters. Hence, RPA clusters are either standalone CPs or consisting of a number of CPs and/or CMPs, which are related through the *Generate*, *Start* and *Request* relationships, and are not related to any other CP or CMP in other clusters. Figure 1 shows the 2nd cut PA diagram for the UWE's CEMS faculty administration case study. The RPA clusters are shown in Figure 1 bounded with circles.

From the figure we can identify 17 clusters, 13 of which are stand alone CPs. As can be seen, each cluster handles a set of some related functionality.

Clusters C1, C2, C3 and C4 are concerned with handling a student, a student withdrawal, a student request to transfer award, and a student problem, respectively. C5 cluster deals with extenuating circumstances. Clusters C6, C7, C8, C9, C10, C11, C12, C13 and C14 are concerned with handling an award handbook, an external examiner, a student loan company report, an induction week, a meeting, a faculty handbook, a referral day, an assessment offence and a quality inspection event, respectively. Cluster 15 is concerned with the administration of new awards, which in turn leads to the definition of any new modules required, the handling of any accreditation required, and the handling of validation events arising. C16 is concerned with the administration of a run of a module called for by the programme plan. Module runs require assignments to be defined, exams to be set, and submissions (including late submissions) to be marked. Cluster C17 is concerned with the administration of student appeals and examining board events [7].

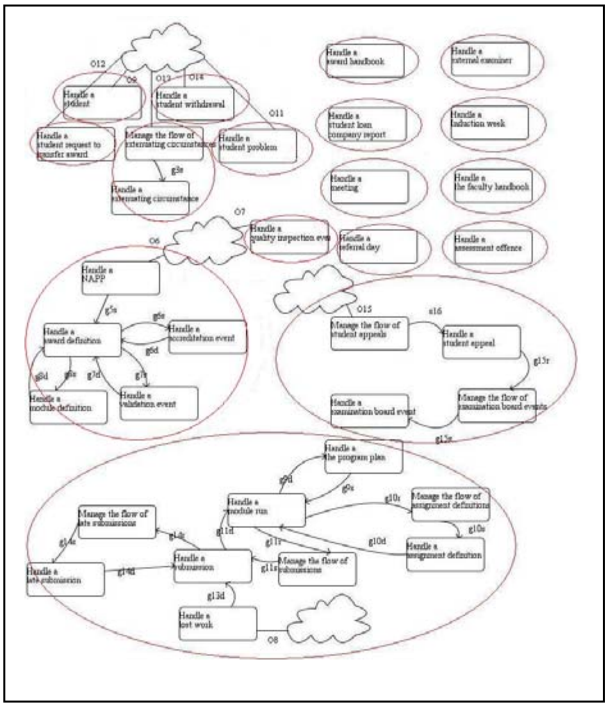

Figure 1. RPA clusters for the UWE's CEMS faculty administration

In this paper we hypothesize that:

RPA clusters that can be identified from the 2nd cut business process architecture diagram using the Riva method can be considered as candidate services suitable for building a SOA-based system.

As we have mentioned previously, our aim is to simplify the service identification process by starting with the business process architecture of an organization so that the effort already utilized in identifying the architecture can be deployed to identify services capable of building a SOA-based system. Hence, in the following section, we analyze RPA clusters identified using the CEMS faculty administration case study to assess the satisfaction of the simplicity criterion for the identified services and their conformance with SOA principles.

V. EVALUATING THE NEW SERVICE IDENTIFICAION APPROACH

A. RPA Clusters to Satisfy Simplicity

Considering RPA clusters as candidate services is in line with our goal to simplify the SI approach, where it was shown in the above example that RPA clusters can be easily identified from the 2nd cut PA identified using the Riva method. Riva is a simple method which is easy to apply and use (except for the first step in identifying the set of EBEs). It is concerned with modeling the organizational behavior in a way that is revealing and communicative, where it exactly shows what processes there are and how they interact [12].

Accordingly, we can conclude that the service identification approach using the RPA clusters criterion inherits the Riva method simplicity.

This also should be a normal consequence of using BPA rather than BPMs to identify services, where both BPA and SOA raise the level of abstraction to a higher level that does not require considering detailed information which results in higher complexity.

B. RPA Clusters to Satisfy Service Definitions and Principles

By conducting a thorough analysis of the RPA clusters generated from the UWE's CEMS faculty administration case study, which we hypothesize are equivalent to candidate services, we provide a discussion in order to assess the extent to which service definitions and principles [4, 5] are satisfied by the suggested identified services. This is summarized in Table 1.

An entity service is a service whose functional boundary and context is based on one or more related business entities [3]. This makes it highly reusable, where a single entity service can be leveraged to automate multiple parent business processes.

We can deduce that RPA clusters conform to the above definition, where according to Riva definition; each CP and CMP in the business process architecture diagram handles an instance of a unit of work or manages the flow of instances of the unit of work, respectively, where each unit of work was derived from an EBE. Hence, each CP's or CMP's functional boundary is based upon one business entity. Accordingly, RPA clusters which are stand alone CPs or sets of CPs and CMPs related through start, require and deliver relations, have functional boundaries based on one or more related business entities.

A Stand alone CP does not have any relation with other CPs or CMPs, i.e. it doesn't require, start or deliver to any other CP or CMP. Accordingly, it has low dependability on other processes. The same applies to the other type of RPA clusters; the set of CPs and CMPs that are related together, where we are grouping CPs and CMPs that request, start and/or deliver to each other. Thus, this means that members of each cluster depend on each other, but not on other clusters, i.e. clusters are loosely coupled and hence satisfy SOA loose coupling principle.

We can detect the direct dependencies between the identified cluster members in the example, and the lower dependencies between the clusters. Cluster C15 of the UWE's CEMS faculty administration case study, for example, is concerned with the administration of new awards, which in turn leads to the definition of any new modules required, the handling of any accreditation required, and the handling of validation events arising. All of these business processes depend on each other but not on processes from other clusters such as those concerned with handling a student or external examiner, etc.

As each RPA cluster corresponds to a set of related processes presented with a high abstraction level, they act as black boxes abstracting underlying functionalities that are considered as the associated service capabilities. This also hides interaction details leaving CPs and CMPs stateless and hence satisfying the SOA principle of Statelessness.

TABLE I. MAPPING THE CHARACTERISTICS OF RPA CLUSTERS TO SERVICE DEFINITIONS AND PRINCIPLES

Characteristics of RPA Clusters, According to Riva Definition	**Service Principles and/or Definitions**	**Mapped RPA cluster Characteristics to Service Definitions and/or Principles**
Each CP in an RPA cluster handles an instance of a unit of work, and Each CMP in an RPA cluster manages the flow of instances of a unit of work, where units of work are initially EBEs with lifetimes handled by members of an organization.	Entity Service definition	The functional boundary of each RPA cluster is based upon one or more business entities.
The first type of RPA clusters are stand-alone CPs, where they do not have require, start or deliver relations with other CPs or CMPs.	Principle of Loose Coupling	Stand alone CPs of RPA clusters have low dependability on other clusters.
The second type of RPA clusters is a set of CPs and CMPs that are related together through request, start and/or deliver relations.	Principle of Loose Coupling	RPA clusters that group CPs and CMPs have low dependability on other clusters.
Each CP and CMP corresponds to a process which is comprised of a set of functionalities.	Principle of Abstracting Underlying Logic	RPA clusters act as black boxes, where they abstract the underlying functionalities that are considered service capabilities.
RPA clusters are concerned with one or more related entities, where granularity level is finer than a BPA or a BPM and is coarser than tasks, and is also coarser or equal to a CP in granularity.	Principle of Reusability/ Principle of Composability	RPA clusters are highly reusable and are composable. The granularity level is not too coarse-grained nor too fine-grained.
CPs and CMPs are related through require, start and deliver relations (i.e. relations between CPs and CMPs are request/response relation, not conversational). The conversational relations between roles are included within each CP or CMP.	Principle of Statelessness	RPA clusters minimize the amount of state information they manage.

Reusability is an important principle of services to be “SOA-able”, and it is related to other service principles, such as loose coupling, composability and statelessness. From both the entity service definition and Riva definitions, we can infer that RPA clusters are reusable. In addition, we note that RPA clusters have granularity levels that are not too fine-grained nor too coarse-grained. If we can represent the process architecture and its consecutive components into a hierarchy as in Figure 2 to indicate granularity, we can observe the RPA clusters’ position in this hierarchy.

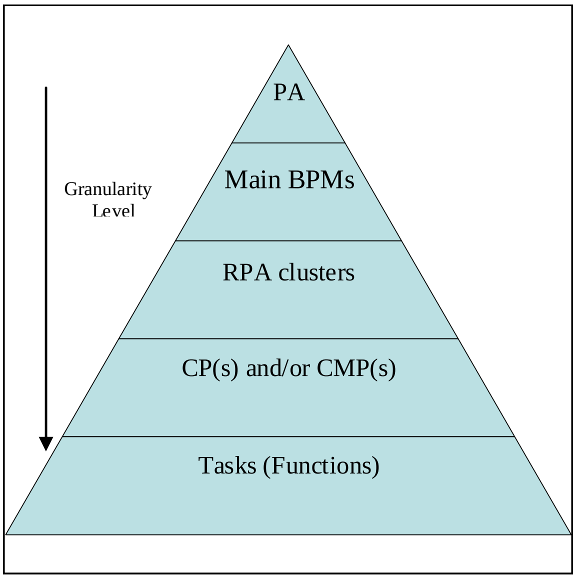

Figure 2. RPA Clusters Position in the PA Granularity Hierarchy

From the above figure we observe that the RPA clusters middles the hierarchy, they are not as coarse-grained as a PA, nor too fine-grained like CPs, CMPs or tasks. So, they provide a suitable granularity level that goes in the middle, satisfying both reusability and the possibility to hide interaction details.

VI. COCLUSION

In this paper, we have proposed a novel approach for identifying services based on the newly defined concept of Riva RPA clusters. These clusters can be easily extracted form a Riva 2^{nd} cut PA diagram, where they are the set of standalone CP (CP with no Start, Deliver or Request relationships) and the set of CPs and CMPs related together using the same relationships.

We have demonstrated using the UWE’s CEMS faculty administration case study, how to identify RPA clusters, which we hypothesized as excellent candidate services. Also, we were able to explain, after analyzing the identified RPA clusters, that the proposed SI approach is simple and can identify services conforming to service definitions and principles, where we were able to map characteristics of RPA clusters which are based on Riva definitions to the related SOA principles.

In Conclusion, this newly developed SI approach adds a further dimension to the Riva BPA method to extend its role and application for the development of service oriented systems, and is also a form of a stronger resemblance between business process activities and software services. In addition, the simplicity of the Riva Method (as being systematic in identifying and modeling the BPA) is reflected into the service identification process. This is in particular reflected in the benefit of considering the business process architecture rather than its associated BPMs to identify services, as BPA raises the level of abstraction to a higher level that does not entail considering detailed information to identify service boundaries that can result in higher complexities.

REFERENCES

[1] S. Alahmari and E. Zaluska, “Optimal Granularity for Service-Oriented Systems”. The 3rd Saudi International Conference (SIC09). University of Surrey, Guildford, UK, 2009.

[2] R. Boerner and M. Goeken, “Process-Oriented Service Identification, A Method for Business-Driven Service Modeling”, 2009.

[3] Cambridge Technology Enterprises. (2010) *SOA Terms* [online]. Available: www.ctepl.com/soaterms.shtml.

[4] T. Erl, SOA: Principles of Service Design. Prentice Hall, 2007.

[5] T. Erl, T. “SOA Principles: An Introduction to the Service-Oriented Paradigm” [online]. Available: http://www.soaprinciples.com/.

[6] N. Fareghzadeh, “Service Identification Approach to SOA Development”. World Academy of Science, Engineering and Technology, 45, 2008, pp. 258-266.

[7] S. Green and M. Ould, “The Primacy of Process Architecture”. 5th Workshop on Business Process Modelling, Development, and Support, in conjunction with the 16th Conference on Advanced Information Systems Engineering (CAiSE04), Riga, 2004.

[8] S. Inaganti and G. Behara, “Service Identification: BPM and SOA Handshake”. BPTrends, 2007, pp. 1-12.

[9] S. Kim, M. Kim., S. Park, “Service Identification Using Goal and Scenario in Service Oriented Architecture”. 15th Asia-Pacific Software Engineering Conference, 2008, pp.419-426.

[10] Y. Kim and K. Doh. “Formal Identification of Right-Grained Services for Service-Oriented Modeling”. Web Information Systems Engineering - WISE 2009. Springer Berlin, Heidelberg, 2009, pp. 261-273.

[11] K. Klose, R. Knackstedt and D. Beverungen, “Identification of Services - A Stakeholder-based Approach to SOA Development and its Application in the Area of Production Planning”. ECIS'07, 2007, pp. 1802-1814.

[12] M. Ould, Business Process Management: A rigorous Approach, BCS, UK, 2005.

[13] M. Ould, Business Processes – modelling and analysis for re-engineering and improvement, John Wiley and Sons, Chichester, 1995