Technical Open Challenges on Benchmarking Workflow Management Systems

Marigianna Skouradaki, Dieter H. Roller, Frank Leymann

Institute of Architecture and Application Systems
University of Stuttgart
Germany
{skouradaki, dieter.h.roller, leymann}@iaas.uni-stuttgart.de

Vincenzo Ferme, Cesare Pautasso
Faculty of Informatics
University of Lugano
Switzerland
{firstname.lastname@usi.ch}

Abstract: The goal of the BenchFlow project is to design the first benchmark for assessing and comparing the performance of BPMN 2.0 Workflow Management Systems (WfMSs). WfMSs have become the platform to build composite service-oriented applications, whose performance depends on two factors: the performance of the workflow system itself and the performance of the composed services (which could lie outside of the control of the workflow). Our main goal is to present to the community the state of our work, and the open challenges of a complex industry-relevant benchmark.

1 Introduction

This paper is an introduction to the Benchflow project[1]. The goal of this project is to design the first benchmark for assessing and comparing the performance of BPMN 2.0[2] Workflow Management Systems (WfMSs). WfMSs have become the platform to build composite service-oriented applications (SOA), whose performance depends on two factors: the performance of the workflow system itself and the performance of the composed services (which could lie outside of the control of the workflow). Therefore the development of such a benchmark, reveals a number of challenges that were not present in benchmark of other types of systems, such as database management systems, or programming language compilers.

In this work we present a list of challenges that are recognized during the design of such a benchmark. We present the current state of our work, as well as the design decisions taken so far, and also discuss the open challenges.

[1] `http://www.iaas.uni-stuttgart.de/forschung/projects/benchflow.php`
[2] `http://www.bpmn.org/`

Proc. SOSP 2014, Nov. 26–28, 2014, Stuttgart, Germany

2 Description of the Workflow Management System environment

The purpose of this section is to provide to the reader a high-level understanding of the WfMSs. For that we first need to clarify the concept of "workflow", which is basically, the focus point of a WfMS. According to [Spe99] a "workflow is the computerized facilitation or automation of a business process, in whole or part". In other words, a workflow is the automation of a series of business activities that are needed for achieving a goal. In this respect a Workflow Management System (WfMS) is a system that supports the reengineering and automation of business and information processes. Its main characteristics are the definition of workflows, and the provision of fast re(design) and re(implementation) as the business needs change [GHS95]. All the components of the WfMS, are compliant with the semantics of a *metamodel* that defines concepts such as the structure of a process and the operations that can be performed on a process model instance [LR00].

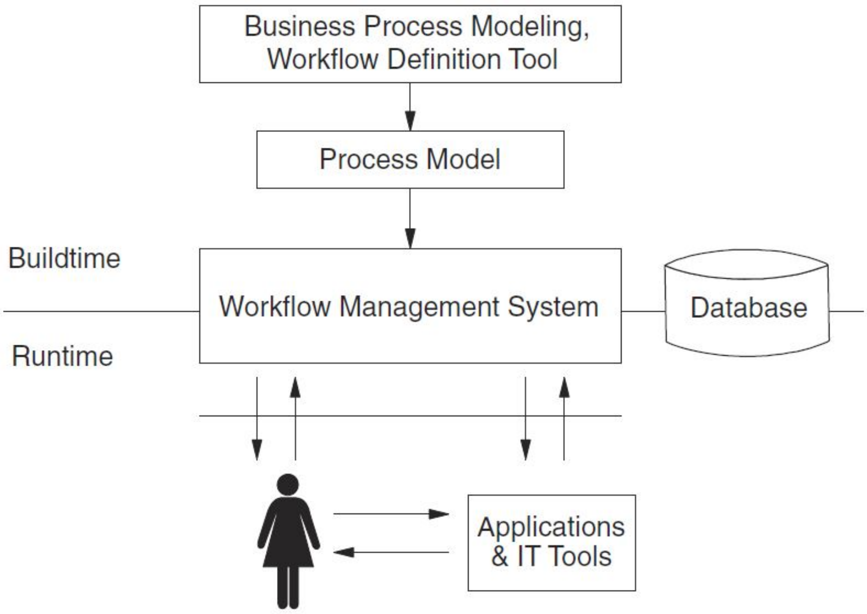

Figure 1: Major Components of a Workflow Management System [LR00]

In Figure 2 we can see the main design and run-time components of a WfMS:

Buildtime provides means to the user to define the constructs such as process models, organizational structures, and information technology aspects (e.g. scripts, programs).
Workflow Definition Tool is the tool with which the users define the workflow according to a workflow language that is compliant with the WfMS.
Process Model is an intermediate representation [Ley10] of the previously defined workflow, in a format that the WfMS can process. Process Models can be expressed through

different definition languages. For many years there was no standard definition language, and each company chose to implement a customized form of the stardard definition languages on their engines. WS-BPEL [Org07] became an industry-accepted standard for executing processes over WSDL-based Web services. Later, it was followed by the creation and industrial acceptance of the BPMN 2.0 standard [JE11]. This work focuses on benchmarking WfMSs that support BPMN 2.0.
Runtime is the part that performs the actual execution of the processes, featuring operations, such as creating, navigating and controlling processes.
Database of the WfMS keeps and handles all the information that are essential for the build and runtime. Thus, in the database we find stored the process models, and also the state of their instances.
User and IT Tools the user (either the owner of a workflow or a participant in it) interacts with the WfMS to report about the progress of his or her tasks, which may be performed directly or through applications and other IT tools.

Currently, WfMSs can be seen as one of the key middleware components of Service Oriented Architectures (SOA). As human users are less and less part of the workflow execution, whereby many different, possibly geographically distributed, service components interact to provide the corresponding functionality, under the control of the WfMS [ZDGH05, GKW+12].

3 Open Technical Challenges

Given the peculiar characteristics of WfMSs, in the following we identify and briefly discuss some of the technical challenges in benchmarking WfMSs.

Automate the generation of realistic workload for different use case scenarios In order to keep our measurements accurate we need to create realistic workload scenarios. As workflow-based applications are currently present in various types of application domains[LR97], it is challenging to select a sufficiently large subset of domains and synthesize a domain-independent workload. In order to address this challenge we have started designing a workload generator for creating process models according to pre-defined criteria. The workload generator needs to take into consideration a set of variables that are related to the correct operation of the workflows, e.g. data-flow, key performance indicators (KPI), actors, and interactions with external entities (e.g. Web Services, databases, external scripts).

The setup of the benchmark environment During the setup of the benchmark environment the challenge is to eliminate the interference of non-WfMS resources as for example database systems, and Web Services used by a process. In order to do that, we design the benchmark environment to be distributed on different physical machines that are connected through the same local network. To do this we need to ensure a flexible deployment mechanism.

System Internal Load Optimization The request load to WfMS is different for the various day times. In order to handle this situation the WfMS shifts work to the daytime where the load is lower. The challenge in this case is to take into account these optimizations in the measurements of the throughput.

Benchmark Long Running Processes The lifetime of a process can span from some milliseconds to hours, weeks, or even some years. The processes that run for more than some hours are called "long-running" processes. An example of a long running process might be a process that sends a message to an external partner and needs to wait for weeks or even years for the reply. The long-running processes have two main characteristics:

1. the storage of the instances in database increases in size and
2. caching cannot be used from either the WfMS or the DBMS

The challenge in this case is to find a method to benchmark this type of processes without having to wait for years for the completion of the benchmark.

Performance Impact of Workflow Language Features BPMN 2.0 provides a large set of constructs, that express iteration, parallelism, exception handling, interactions with external entities etc. According to a research among the Websites and release notes of the currently available engines, we have observed that the documentation on the coverage of BPMN 2.0 constructs is usually inadequate. There is also the hypothesis that WfMSs avoid implementing all of the constructs for performance reasons. It is also important to consider which of these constructs are actually used in real-world processes [MR08]. The challenge in this case is to: a) run compliance tests to define which constructs are actually supported by the available WfMSs, and b) decide which constructs need to go into the workload mix.

4 Project Status

The initial phase of the project focused on analyzing a large collection of process models, for collecting data to proceed to workload characterization and generation. The set up of the benchmark environment is also planned at this phase.

Currently we have collected a large collection of 8363 real-life process models that are expressed in various modelling languages. Collecting process models can be a big challenge as companies are not willing to share them to protect their corporate assets. The ways by which we countenanced the process model sharing was to sign confidentiality agreements, and to develop tools for obfuscating the models [SRPL14], and thus protecting the company's intellectual property. Our collection contains: 1% WS-BPEL, 4% EPC, 7% YAWL, 24% Petri Net, and 64% BPMN & BPMN 2.0 Models, where 2/3 of them are BPMN 2.0. These data justify our decision to focus on the engines that support the BPMN 2.0 standard, as it seems to have become widely accepted in industry and academia.

However, the focus on the BPMN 2.0 standard was not the only decision that was taken with respect to the data derived from the analysis of the process models collection. We have currently ran statistical analysis on more than hundred different metrics, observing process models characteristics (e.g. control flow gateways fan-in, fan-out, cyclomatic complexity), and the frequency of appearance of the BPMN 2.0 elements in the collected process models. This led to a better understanding of which BPMN 2.0 features are present in the real world, and thus would need to be supported by the engines participating in the benchmark. We also arrived at the definition of representative process model mockups, generated based on the results of a cluster analysis over structural complexity, and language construct usage metrics.

The statistical analysis that we did on the process models gave us information about the model characteristics at a "micro" level. We also wanted to define complete structures that are repeatedly found in the process models. By this way we can combine different structures with each other and have an even more realistic workload. The problem of discovering the reoccurring structures is reduced to "Frequent Pattern Discovery", and it is NP-complete [Ben02]. To address this challenges we have implemented an extension of the VF-2 algorithm which is said to be one of the most efficient algorithms in graph isomorphism [PCFSV04]. The exported structures must now be characterized according to benchmark-related criteria, and combined with respect to the metrics found in the aforementioned statistical analysis. The idea is to give all the acquired information as an input to a workload generator, and create realistic workload, that is compliant for different benchmarking scenarios.

The choice BPMN 2.0 engines to be included in the benchmark is also a decision that needs to be taken at this phase of the project. The current state is that there are more than 20 engines available implementing the BPMN 2.0 Standard. We are currently collecting the available engines, and relevant information in a Wikipedia page[3]. Unfortunately, the lack of documentation in terms of the engine's compliance to the BPMN 2.0 standard, is currently delaying our decision. In order to find out this information we are contacting the vendors asking for information, and conducting compliance tests towards the engines. The aforementioned statistical analysis also plays its roles at this point, as we need to choose among these engines which implement at least the most frequently used elements of the BPMN 2.0 Standard.

In terms of the benchmark set-up we developed and deployed the first prototype of a WFMS on different physical machines connected through the same local network. The challenges that we needed to address are currently solved using Docker[4] that offers a flexible deployment mechanism with a minimal impact on the performance measurement. Docker guarantees a good level of isolation and a quick start up, and enables repeatability of the tests because the initial conditions are kept inside the Docker containers and are exactly the same for each execution of the benchmark.We are currently using Faban[5] to develop the first prototypes of benchmark drivers, we are running the first experiments using the three models we derived from the statistical analysis of a large model collection, and we are analyzing the execution log to compute the first KPIs such as throughput.

[3]http://en.wikipedia.org/wiki/List_of_BPMN_2.0_engines
[4]https://www.docker.com
[5]http://www.faban.org

5 Related Work

In the topic of benchmarking the performance of WfMSs, only some work is available (e.g., [GMW00, BBD10, DPZ11]). In the most recent approaches, the SOABench project [BBD10] can be seen as an initial step to provide a performance assessment and comparison framework for SOA middleware systems.

SPEC has also introduced a group that focuses on benchmarking SOA infrastructures [Sta10]. However information on their progress of research is still not available on the website. Vendors of proprietary systems are also executing internal benchmarks (e.g. [SAP11], [Inc11], [IC07]) in order to evaluate their work and inform the prospective customers. However, there is not any standard method followed in these becnhmark approaches. As easily concluded, and also emphasized by the literature, the need to introduce a benchmark standard for that addresses the industry state-of-art needs is now imperative [KKL06, WLR+09, RvdAH07, LMJ10]. Our work intends to the creation of such a standard benchmark, that differs from the related work, in terms of: a) the number of WfMSs to be compared, b) the complexity and diversity of the workload mix, c) the number of the executed performance tests, and d) the number of performance metrics that will be taken into consideration, and their aggregation into a meaningful number.

6 Discussion & Conclusion

In this paper we have presented the open challenges for the creation of a benchmark for WfMSs that we have been addressing in the first phase of the BenchFlow project as well as some solutions for addressing them. We have seen how the synthesis of the workload mix needs to be created through a workload generator, how we plan to setup the benchmark environment and isolate it from internal interferences, and how the compliance tests and statistics on BPMN 2.0 support are important before deciding the workload mix. We have also raised issues such us long-running processes benchmarking, and the consideration of WfMS internal load optimization into our benchmark.

By presenting this set of open research challenges in this position paper, we aim to present the SOSP community with a complex industry-relevant benchmarking challenge and discuss how to deal with the aforementioned open challenges.

References

[BBD10] D. Bianculli, W. Binder, and M.L. Drago. SOABench: Performance Evaluation of Service-Oriented Middleware Made Easy. In *Proc. of the 19th International World Wide Web Conference (WWW 2010)*, Raleigh, NC, 2010.

[Ben02] E. Bengoetxea. *Inexact Graph Matching Using Estimation of Distribution Algorithms*. PhD thesis, Ecole Nationale Supérieure des Télécommunications, Paris, France, Dec 2002.

[DPZ11] Florian Daniel, Giuseppe Pozzi, and Ye Zhang. Workflow Engine Performance Evaluation by a Black-Box Approach. In *Proc. of the International Conference on Informatics Engineering & Information Science*, ICIEIS 2011, pages 189–203. Springer, November 2011.

[GHS95] D. Georgakopoulos, M. Hornick, and A. Sheth. An overview of workflow management: From process modeling to workflow automation infrastructure. *Distributed and parallel Databases*, 3(2):119–153, 1995.

[GKW+12] Spyridon Gogouvitis, Kleopatra Konstanteli, Stefan Waldschmidt, George Kousiouris, Gregory Katsaros, Andreas Menychtas, Dimosthenis Kyriazis, and Theodora Varvarigou. Workflow management for soft real-time interactive applications in virtualized environments. *Future Generation Computer Systems*, 28(1):193 – 209, 2012.

[GMW00] Michael Gillmann, Ralf Mindermann, and Gerhard Weikum. Benchmarking and Configuration of Workflow Management Systems. In *Proc. of the 7th International Conference on Cooperative Information Systems*, CoopIS 2000, pages 186–197, 2000.

[IC07] Intel and Cape Clear. BPEL Scalability and Performance Testing. White paper, 2007.

[Inc11] Active Endpoints Inc. Assessing ActiveVOS Performance, 2011.

[JE11] Diane Jordan and John Evdemon. Business Process Model And Notation (BPMN) Version 2.0. Object Management Group, Inc, January 2011.

[KKL06] Rania Khalaf, Alexander Keller, and Frank Leymann. Business processes for Web Services: Principles and applications. *IBM Systems Journal*, 45(2):425–446, 2006.

[Ley10] Frank Leymann. BPEL vs. BPMN 2.0: Should You Care? In Jan Mendling, Matthias Weidlich, and Mathias Weske, editors, *Business Process Modeling Notation*, volume 67 of *Lecture Notes in Business Information Processing*, pages 8–13. Springer Berlin Heidelberg, 2010.

[LMJ10] Guoli Li, Vinod Muthusamy, and Hans-Arno Jacobsen. A distributed service-oriented architecture for business process execution. *ACM Trans. Web*, 4(1):2:1–2:33, January 2010.

[LR97] Frank Leymann and Dieter Roller. Workflow-Based Applications. *IBM Systems Journal*, 36(1):102–123, 1997.

[LR00] Frank Leymann and Dieter Roller. *Production Workflow: Concepts and Techniques*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 2000.

[MR08] M. Muehlen and J. Recker. How much language is enough? Theoretical and practical use of the business process modeling notation. In *Advanced Information Systems Engineering*, pages 465–479. Springer, 2008.

[Org07] Organization for the Advancement of Structured Information Standards (OASIS). *Web Services Business Process Execution Language (WS-BPEL) Version 2.0*, April 2007.

[PCFSV04] Luigi P. Cordella, Pasquale Foggia, Carlo Sansone, and Mario Vento. A (Sub)Graph Isomorphism Algorithm for Matching Large Graphs. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(10):1367–1372, October 2004.

[RvdAH07] N. Russell, W.M.P. van der Aalst, and A.H.M. Hofstede. All That Glitters Is Not Gold: Selecting the Right Tool for Your BPM Needs. *CUTTER IT JOURNAL*, 20(11):31–38, November 2007.

[SAP11] IBM SAP. SAP NetWeaver Business Process Management Performance, Scalability, and Stability Proof of Concept. Technical report, IBM SAP International Competence Center (ISICC)Walldorf, Germany SAP NetWeaver BPM Product Management, Walldorf, Germany SAP Labs, Berlin, Germany SAP Active Global Support, Walldorf, Germany, 2011.

[Spe99] Workflow Management Coalition Specification. *Workflow Management Coalition, Terminology & Glossary (Document No. WFMC-TC-1011).* Workflow Management Coalition Specification, February 1999.

[SRPL14] Marigianna Skouradaki, Dieter Roller, Cesare Pautasso, and Frank Leymann. BPELanon: Anonymizing BPEL Processes. In *Proc. of ZEUS'14*, pages 09–15, 2014.

[Sta10] Standard Performance Evaluation Corporation. SPEC SOA Subcommittee, February 2010. `http://www.spec.org/soa/`.

[WLR+09] B. Wetzstein, P. Leitner, F. Rosenberg, I. Brandic, S. Dustdar, and F. Leymann. Monitoring and Analyzing Influential Factors of Business Process Performance. In *Proc. of the IEEE International Conference on Enterprise Distributed Object Computing Conference*, EDOC 2009, pages 141 –150, September 2009.

[ZDGH05] Olaf Zimmermann, Vadim Doubrovski, Jonas Grundler, and Kerard Hogg. Service-oriented Architecture and Business Process Choreography in an Order Management Scenario: Rationale, Concepts, Lessons Learned. In *Companion to the 20th Annual ACM SIGPLAN Conference on Object-oriented Programming, Systems, Languages, and Applications*, OOPSLA '05, pages 301–312, New York, NY, USA, 2005. ACM.