Big Data on Performance Logs –
A Collaborative Monitoring Cloud for ERP Systems

H. Müller$^1$ and K. Turowski$^1$

$^1$Very Large Business Applications Lab, Otto von Guericke University, Magdeburg, Germany

Abstract - Although outsourcing is a viable instrument to save operational costs, the majority of ERP systems is still operated in-house due to privacy, security and dependency concerns coupled with ERP’s exceptional significance for business continuity. In this abstract paper, we propose a research artefact that is planned to become an alternative option to classical in-house or off-promises operation models and enables fully controlled in-house operation with cloud-supported performance analyses. Therefore, we started to analyze 230 million performance log entries of about 8,700 standard SAP ERP systems and evaluate its suitability for a value creating Big Data scenario. Integrating performance data and hardware information of ERP systems enables cross-system and cross-customer analyses and, potentially, to deliver additional knowledge to ERP operating IT departments through a cloud service.

Keywords: ERP; Performance; Big Data; Cloud.

1 Introduction

Outsourcing of information technology has attracted significant interest from both research communities and industries [1]. Firms benefit from economies of scale by combining their operational costs for hardware, software and staff through external service providers. Particularly in cases of planned greenfield projects or major architecture changes, outsourcing is a viable instrument to save investment costs by spreading them over the entire contract period. Besides, information management has transformed from a technical perspective of “plan-build-run” to a business perspective of “source-make-deliver”. Driven by market orientation, product orientation and product lifecycle management [2], modern business departments consume IT products delivered by both internal and external providers, while quality attributes are ensured by means of service level agreements. While firms make use of these effects for various business applications, outsourcing, on the other hand, involves risks that still dominate decisions for certain core applications. Bryson and Sullivan discussed reasons for and against outsourcing which include concerns about security, privacy and application service provider (ASP) dependency [1], [3]. Olsen states that the biggest risks of outsourcing are downtime and loss of operational data. Therefore, companies tend to outsource applications which are not business critical in the effect that business continuity can still be ensured. According to the senior director of IT applications at Informatica Corp., the human resources (HR) module of their business suite could be outsourced, because the business will still continue to run if HR goes down [4]. Olsen argues that companies view ERP as too mission-critical to yield control. The CIO at Federal-Modul Corp., Mike Gaynor, clearly states a lack of control and trust when it comes to ERP outsourcing. He illustrates the dominating attitude of IT executives by a metaphor advising to never hand off the brains of operation, although companies might farm out for extra arms and legs [4]. One of the most widely used ERP system is developed by the German SAP AG. A recent study among German SAP customers from October 2014 [5] states that 75% of the surveyed companies do not have any cloud strategy regarding their SAP application landscape, whereby 62% of these evaluate SAP clouds as ‘not relevant’. For an average of 56% of the companies across all surveyed industries, any software-as-a-service (SaaS) offerings for SAP are neither used nor planned or discussed. Only 6% of the surveyed companies are discussing infrastructure-as-a-service (IaaS) alternatives to their SAP in-house operation and 72% of those companies who make use of SAP IaaS are extending their existing SAP solution [5]. Thomas and Chirania discuss on whether or not to outsource ERP and explain existing restraints by the lack of implementing unique business processes [6].

Based on [7], Olsen summarizes seven options of ERP operation from in-house till ASP and states that each specific organization might generate variants of these that suit their particular needs [1]. Our research project seeks to develop and evaluate an additional option, which includes in-house operation with cloud-supported performance analyses. Therefore, additional knowledge derived from performance data of various ERP systems is going to be extracted by means of Big Data techniques.

2 State of the Art

For ERP systems that can be distributed across multiple application servers and serve hundreds of users simultaneously, performance monitoring becomes a complex task and response times need to be measured and analyzed in various dimensions. Therefore, IT departments make use of consultancy services, e.g. offered by hardware partners. Those services include ERP system performance monitoring during
productive operation and subsequent analyses of log records [8], they are requested and delivered whenever necessity is assumed. Generated performance reports serve as decision support for system administrators regarding sizing requirements, load balancing strategies, job scheduling and other points of action which affect the system performance. Within SAP systems, log entries are called statistical records and created automatically after each dialog step, performed by any user. These records are used to assess, e.g., system performance on business transaction level. When adding information about utilized hardware components, statistical records can be used for benchmarking, too. Cloud-based benchmark tools for hardware components already exist in other domains where performance matters, e.g., the gaming community [9], [10]. For SAP ERP systems, standard application benchmarks exist, e.g., the Sales & Distribution (SD) benchmark, which simulates a defined amount of simultaneously working users and measures response times within a given load interval [11]. In the following, we provide an overview of an SAP ERP performance cloud, which we are going to evaluate with respect to its ability to provide dedicated transaction performance analyses for each customer, performance comparisons across systems, hardware and customers, as well as benchmarking capabilities.

3 Proposed Research Artefact

Taking into account the existing concerns and demands regarding the operation of core business functionality, we developed an operational option that supports in-house operation but cloud based performance analyses of ERP systems. In that manner, only monitored performance data will need to be shared with the service provider while master and transaction data as well as operational control remains at operator site. Therefore, customers benefit from tools, data models and analysis techniques implemented once by the service provider and utilized for multiple systems and customers. In addition, integration of monitored data from different systems enables cross-system and cross-customer analytics. Therefore, customers are given the opportunity to assess their ERP system landscape by comparing own system performance with, e.g., mean values derived from empirical distribution functions across various systems on similar hardware. Hence, ERP performance logs contain information that can be extracted on a global scale leading into a Big Data use case, which we are going to investigate and implement as appropriate. During our research, we will focus on both operator’s and provider’s perspective and their different objectives.

4 Research Design and Outlook

We planned our research based on the information systems research framework provided by [12]. Therefore, we design an option of collaborative ERP in-house operation as a research artefact in multiple iterations, considering stated business needs from the environment and contributing results from planned analytics to the research knowledge base. Together with a major infrastructure partner of SAP customers, we started to clean and preprocess more than 230 million statistical records of about 8,700 SAP systems. An entity relationship (ER) model for a common database schema was developed and performance data has been imported into one in-memory database, which integrates a web server that can be used for providing the customer’s user interface. Using statistics, we started preliminary work on identified use cases for response time predictions and hardware comparisons on business transaction level. In the current project phase, we collect further use cases of data analyses and seek feedback from both scientific community and industry. During the whole research, we plan to focus on both data analytics and the technical infrastructure required for the proposed research artefact.

5 References