Telematics and Informatics 33 (2016) 793–807

Contents lists available at ScienceDirect

Telematics and Informatics

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

Real-time business activity monitoring and analysis of process performance on big-data domains

Alejandro Vera-Baquero [a,*], Ricardo Colomo-Palacios [b], Owen Molloy [c]

[a] Universidad Carlos III de Madrid, Department of Computer Science and Engineering, Leganés, Madrid, Spain
[b] Østfold University College, Department of Computer Science, B R A Veien 4, 1783 Halden, Norway
[c] National University of Ireland, Galway, Department of Information Technology, Galway, Ireland

ARTICLE INFO

Article history:
Received 26 August 2015
Received in revised form 21 November 2015
Accepted 10 December 2015
Available online 12 December 2015

Keywords:
Big data
Cloud computing
Business activity monitoring
Business Process Improvement

ABSTRACT

Real-time access to business performance information is critical for corporations to run a competitive business and respond to a continuously changing business environment with ever-higher levels of competition. The timely analysis and monitoring of business processes are essential to identify non-compliant situations and react immediately to those inconsistencies in order to respond quickly to competitors. In this regard, the integration of business intelligence (BI) systems with Process Aware Information Systems (PAIS) can become a key tool for business users in decision making. However, current BI systems are not suitable for optimising and improving end-to-end processes since these are normally business domain specific and are not sufficiently process-aware to support the needs of process improvement type activities. In addition, highly transactional business environments may produce vast amounts of event data that cannot be efficiently managed by the use of traditional storage systems which are not designed to manage vast amounts of event data. We introduce a cloud-based architecture that leverages big-data technology to support performance analysis on any business domain, in a timely manner and regardless of the underlying concerns of the operational systems. Likewise, we demonstrate the ability of the solution to provide real-time business activity monitoring on big-data environments with low hardware costs.

1. Introduction

ICT-based tools in general and real-time measurement and data analysis of the performance of operational activities is essential for companies to remain competitive (Keller and von der Gracht, 2014). The monitoring of business process execution allows business users to detect error rates and non-compliant business situations, such as supply chain issues. This action must be performed on-time in order to react quickly to those situations. In a well running process it is expected that arrival (demand) and throughput rates should be in balance. Processes or activities which do not have the capacity to work to this arrival rate will cause delays and bottlenecks, thereby starving proceeding activities of input. This may result in increase delays and a loss of profit due to a waste of valuable resources that are underutilised, and consequently, a loss of customer satisfaction and loyalty.

* Corresponding author.
E-mail addresses: averabaq@gmail.com (A. Vera-Baquero), ricardo.colomo-palacios@hiof.no (R. Colomo-Palacios), owen.molloy@nuigalway.ie (O. Molloy).

http://dx.doi.org/10.1016/j.tele.2015.12.005
0736-5853/

A successful analysis of business processes is essential for organisations to gain competitiveness (Delgado et al., 2014). Moreover, process improvement based on analysis is seen as a way to lead organisations to effectiveness (Herranz et al., 2014; Ruiz-Rube et al., 2015). Process models are often inadequately understood and optimised within organisations, and consequently processes under-performed. This leads to long response times, unbalanced utilisation of resources, low service levels, and so on, thereby causing high costs and dramatic loss of profits to corporations (van der Aalst et al., 2010). In this regard, the use of advanced analytical techniques would help analysts to continuously improve their processes, thus meeting their business goals.

The combination of BI and business activity monitoring (BAM) technologies may provide mechanisms to infer knowledge about business performance, but these are not sufficient for answering most of the demanding questions of today's business users. There currently exists an increasing demand for more advanced analytics such as root cause analysis of performance issues, predictive analysis and the ability to perform "what-if" type simulations. These features are powerful assets for analysts, expanding their knowledge beyond the limits of what current platforms typically offer. Furthermore, these platforms are normally business domain specific and have not been sufficiently process-aware to support the needs of process improvement type activities, especially on large and complex supply chains, where it entails integrating, monitoring and analysing a vast amount of dispersed, unstructured event logs produced on a variety of heterogeneous environments, in a timely manner.

In general, the monitoring and analysis of operational data aims to be fact-based and therefore empirically evaluated with real data which leads to trustworthy analysis results. However, this is complex to achieve as there exists a noticeable disconnect between idealised business processes and their actual event-data. Current BI platforms by their own do not fill this gap as they are focused on local decision making rather than end-to-end processes. As a consequence, their outputs tend to be unreliable since they are based on idealised models of reality rather than on observed facts (van der Aalst, 2012). In order to allow business users to gain visibility on their business processes, the execution outcomes must be gathered from operational sources, unified and correlated across organisational boundaries (van der Aalst, 2012).

The latest advances in technology have made it possible for organisations to co-operate with each other, necessitating the integration of diverse business information systems across large and complex supply chains in several domains (Stantchev et al., 2014; Colomo-Palacios et al., 2014). In these scenarios, isolated optimisation within individual organisations is insufficient to optimise and improve end-to-end processes. This leads to the management of complex operational processes, where web services technology and cloud computing have become widely used, producing cross-functional event logs that are beyond company (and increasingly software) boundaries. This has promoted an incredible growth in corporate event data that needs to be merged for analysis (van der Aalst, 2012). Moreover, enterprises' business data are usually handled by heterogeneous systems which run on different technological platforms, and even use incompatible standards. In addition, the continuous execution of distributed business processes (BP) may produce vast amounts of event data that cannot be efficiently managed by the use of traditional storage systems which are not adequate to manage event data of the order of hundreds of millions of linked records (Vera-Baquero et al., 2013). Therefore, innovative methods and techniques are needed to put real expert systems technology in the hands of business users. Nowadays, there exist emerging technologies such as big-data and cloud-computing that can be leveraged to drive the generation of business process analytics (BPA) solutions with the capabilities to produce outcomes on a timely basis. Notwithstanding, the successful implementation of a fully distributed BPA solution involves significant challenges that are not easy to address:

(1) First, BI-like platforms must be re-engineered to support business process analytics, and these are typically business domain-specific.
(2) Processes and enterprise events are intrinsically related to each other but these need to be correlated across organisational boundaries, and this is challenging.
(3) Measuring and improving overall business performance is especially hard to achieve on highly distributed environments whose business processes are part of complex supply chains. In turn, these processes are typically executed under a variety of heterogeneous systems, which makes them even harder to measure.
(4) Continuous execution of distributed business processes may produce a vast amount of event data that cannot be efficiently managed by means of traditional storage systems, which are not adequate to manage event data in the order of hundreds of millions of linked records.
(5) Existing centralised approaches such as described in Vera-Baquero and Molloy (2012), cannot provide real-time analytics on complex business cases that produce large amounts of event data. These systems are not suitable to deal with such volumes of information since they neither include sharding mechanisms nor provide big-data support. Furthermore, these approaches may entail a significant latency from the time the event occurs on source to the time the event is recorded in central repositories. This pitfall is intensified on very large and complex supply chains which normally involve a high number of business units and a greater number of operational systems.
(6) Dealing with highly distributed supply chains demands some collaborative analysis where individual stakeholders are geographically separate and need a platform to perform BPM in a collaborative fashion, rather than depending on a single centralised process owner to monitor and manage performance at individual supply chain nodes. This is especially complex to accomplish using centralised approaches.