Design and Evaluation of an Autonomous Load Balancing System for Mobile Data Stream Processing Based On a Data Centric Publish Subscribe Approach

Rafael Oliveira Vasconcelos, Department of Informatics, Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil
Markus Endler, Department of Informatics, Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil
Berto de Tácio Pereira Gomes, Federal University of Maranhão (UFMA), São Luís, Brazil
Francisco José da Silva e Silva, Department of Informatics, Federal University of Maranhão (UFMA), São Luís, Brazil

ABSTRACT

Several new applications of mobile computing environments, such as Intelligent Transportation Systems, Fleet Management and Logistics, and integrated Industrial Process Automation share the requirement of remote monitoring and high performance processing of huge data streams produced by large sets of mobile nodes. Two key requirements for the deployment and operation of such mobile infrastructures are the handling of large and variable numbers of wireless connections to the monitored mobile nodes regardless of their current use or locations, and to automatically adapt to variations in the volume of the mobile data streams. This article describes the design, implementation, and evaluation of an autonomic mechanism for load balancing of mobile data streams. The autonomic capability has been incorporated into a scalable middleware system based on a Data Centric Publish Subscribe approach using the OMG Data Distribution Service (DDS) standard and aimed at real-time and adaptive handling of mobile connectivity and data stream processing for great sets of mobile nodes. A significant amount of evaluation experiments of the proposed infrastructure is presented, reinforcing its viability and the benefits arising from the use of an autonomic approach to handle the requirements of high variability and scalability.

Keywords: Autonomic Computing, Data Stream Processing, DDS, Load Balancing, Mobile Communication Middleware

DOI: 10.4018/ijaras.2014070101
INTRODUCTION

A large number of applications require continuous and timely processing of high-volume of data originated from many distributed sources to obtain real-time notifications/alerts from complex queries over the steady flow of data items (Stonebraker, Çetintemel, & Zdonik, 2005) (Cugola & Margara, 2012). This has led to a new computing model called data stream processing (Stonebraker et al., 2005). This model is focused at sustained and timely analysis, aggregation and transformation of large volume of data streams that are continuously updated (Cugola & Margara, 2012). The need to process data streams comes from several application areas such as network monitoring, traffic engineering systems, intelligent routing of cars in metropolitan areas, sensor networks, telecommunication systems, financial applications and meteorology. Although some sorts of data processing may be performed locally at the mobile nodes (i.e., simple sensor data encoding or interpretation), most other application-relevant information about the monitored mobile system as a whole that requires execution of complex analysis and correlation functions over these mobile data streams, in real-time, has to be done by dedicated machines in a cluster or cloud, which increases the need for load balancing solutions (Calsavara & Lima Jr., 2011) (Randles, Lamb, & Taleb-Bendiab, 2010) (Zhang, Cheng, & Boutaba, 2010). Moreover, these systems must also support timely and reliable communication with the monitored mobile nodes, in order to send instructions, share alternative routes, make inquiries or disseminate alerts, either to nodes individually, or to groups of nodes (Karnouskos & Colombo, 2011).

However, today’s mobile communication and data stream processing systems lack important features that are necessary in order to support the large amounts of data flows envisioned by the massive and ubiquitous dissemination of sensors and mobile devices in industry and city-scale applications. In particular, the deployment and 24x7 operation of such mobile data stream processing and communication infra-structures pose two intrinsic technical challenges: they must be capable of (a) handling huge and variable numbers of connections to the monitored mobile nodes regardless of their current locations, and (b) automatically adapt to variations in the volume of the mobile data streams, either because of sudden increase in the set of nodes (e.g. a popular event happening in a region), intensified message exchange among the mobiles, or because new sensed data is required for a more precise analysis of a regional problem, such as road defect or an potential accident. In order to address these challenges we have developed a scalable middleware system that supports efficient and adaptive handling of mobile connectivity and data stream processing for thousands of mobile nodes. In this paper, we specifically explain the autonomic load distribution mechanisms implemented in the middleware, and discuss their potential benefits.

The remainder of this paper is organized as follows: Section Implementation Technologies: Overview provides an overview of the three technologies used in the implementation of our autonomous load balancing system, which are: the Data Distribution Service for Real-Time Systems standard, used for implementing the communication infrastructure of the autonomous load balancing system; the Scalable Data Distribution Layer (SDDL), used as the middleware for mobile communications; and the MAPE-SDDL Framework, an extension for the SDDL, which adds autonomic management services for the SDDL. Section Load Balancing of Mobile Data Streams delves into the proposed Data Processing Slice Load Balancing approach for mobile data streams and how it was implemented, while Section Evaluation describes the detailed evaluation of the implemented system using a prototype application. Section Related Works reviews related work on load balancing for Publish/Subscribe systems, including DDS (OMG, 2012a), while Section Discussion discusses the advantages of using an autonomic approach for load balancing and the benefits of the load balancing solution proposed by this work.
Software Agents for Human Interaction in Social Networks
www.igi-global.com/chapter/software-agents-human-interaction-social/46289?camid=4v1a

Comparative Study of Evolutionary Computing Methods for Parameter Estimation of Power Quality Signals
www.igi-global.com/chapter/comparative-study-evolutionary-computing-methods/66817?camid=4v1a