Chapter 6

Non–Intrusive Autonomic Approach with Self–Management Policies Applied to Legacy Infrastructures for Performance Improvements

Rémi Sharrock
LAAS-CNRS - University of Toulouse; UPS, INSA, INP, ISAE, France

Thierry Monteil
LAAS-CNRS - University of Toulouse; UPS, INSA, INP, ISAE, France

Patricia Stolf
IRIT and Université de Toulouse, France

Daniel Hagimont
IRIT and Université de Toulouse, France

Laurent Broto
IRIT and Université de Toulouse, France

ABSTRACT

The growing complexity of large IT facilities involves important time and effort costs to operate and maintain. Autonomic computing gives a new approach in designing distributed architectures that manage themselves in accordance with high-level objectives. The main issue is that existing architectures do not necessarily follow this new approach. The motivation is to implement a system that can interface heterogeneous components and platforms supplied by different vendors in a non-intrusive and generic manner. The goal is to increase the intelligence of the system by actively monitoring its state and autonomously taking corrective actions without the need to modify the managed system. In this paper, the authors focus on modeling software and hardware architectures as well as describing administration policies using a graphical language inspired from UML. The paper demonstrates that this language is powerful enough to describe complex scenarios and evaluates some self-management policies for performance improvement on a distributed computational jobs load balancer over a grid.

INTRODUCTION

Autonomic Computing Principles

Autonomic computing aims to provide methods and tools to answer the exponentially growing demand in IT (Information Technologies) infrastructures. These IT systems are getting increasingly complex while using a wide variety of technologies. Huebscher (2008) compares the actual situation to the one experienced in the 1920s in telephony: automatic branch exchanges finally supplanted trained human operators. Nowadays, large IT facilities involve important time and effort costs to operate and maintain hardware and software. Numerous new technologies are emerging and they consume considerable human resources in learning how to run, tweak, or configure. One of the challenges facing large companies that use such IT infrastructures is that of reducing their maintenance and operating costs (David, Schuff, & St. Louis, 2002) in order to increase their dependability (Sterritt & Bustard, 2003) and assurance levels to help them being more confident.

Here are some of the issues raised in this field of research:

- First of all, managing large scale infrastructures requires describing the global system in a synthetic way. This involves describing a deployment objective, a picture of what to deploy and how to deploy it. This picture represents what the system should look like upon deployment, the intended construction. Indeed, it is necessary to have an automatically orchestrated deployment process due to the huge number of machines (at least hundreds, and up to thousands) that are potentially involved. It has been shown that large deployments cannot be handled by humans, because this often leads to errors or inconsistencies (Flissi, Dubus, Dolet, & Merle, 2008).

- Once deployed, the system has to be configured and started. These multiple tasks are ordered (some parts of the system have to be configured or started before others), which also imply an automatic process. Indeed, Kon and Campbell (1999) argue that it is hard to create robust and efficient systems if the dynamic dependencies between components are not well understood. They found common issues were some parts of the system fail to accomplish their goals because unspecified dependencies are not properly resolved. Sometimes, failure of one part of the system could also lead to a general system failure.

- During runtime, human system operators tend to have slow reaction times and this can result in unavailability of critical services. For example, Gray (1986) clearly shows that human mistakes made during maintenance operations or reconfigurations are mainly responsible for failures in distributed systems. The need here is to introduce a rapid system repair or reconfiguration so that critical services are kept at an acceptable level (Flaviu, 1993). Moreover, new services (large audience services like social networks, video services) tend to follow heavy fluctuations in demands (Cheng, Dale, & Liu, 2008). Thus, these services experience scalability issues and need fast rescale mechanisms.

Raised Problematics

Exploring the autonomic computing principles lead us to the following questions:

- How to describe deployment objectives, repairs, reconfigurations or rescaling? This raises the need for easy to use, high-level policies to describe the management of distributed systems.

Related Content

Portfolio Optimization Using Evolutionary Algorithms

Lean Yu, Shouyang Wangand Kin Keung Lai (2008). *Reflexing Interfaces: The Complex Coevolution of Information Technology Ecosystems (pp. 235-245).*

www.irma-international.org/chapter/portfolio-optimization-using-evolutionary-algorithms/28381

Impact of Cross-Layer Adaptations of Mobile IP on IEEE 802.11 Networks on Video Streaming

P. De Cleynand C. Blondia (2012). *Technological Innovations in Adaptive and Dependable Systems: Advancing Models and Concepts (pp. 261-281).*

www.irma-international.org/chapter/impact-cross-layer-adaptations-mobile/63586

The Future Quantum Computer: Biotic Complexity

Hector Sabelliand Gerald H. Thomas (2008). *Reflexing Interfaces: The Complex Coevolution of Information Technology Ecosystems (pp. 138-172).*

www.irma-international.org/chapter/future-quantum-computer/28377

The Practical and Theoretical Trails in Hindsight

Lars Taxén (2010). *Using Activity Domain Theory for Managing Complex Systems (pp. 162-168).*

www.irma-international.org/chapter/practical-theoretical-trails-hindsight/39677

The User Requirement Survey and Analysis System of Knowledge Management for Laboratories in Universities

Jiaxun Wangand Hongtao Ren (2013). *International Journal of Knowledge and Systems Science (pp. 26-34).*

www.irma-international.org/article/user-requirement-survey-analysis-system/80419