J. Parallel Distrib. Comput. 74 (2014) 2918–2933

Contents lists available at ScienceDirect

J. Parallel Distrib. Comput.

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

A survey of Cloud monitoring tools: Taxonomy, capabilities and objectives

Kaniz Fatema[a,*], Vincent C. Emeakaroha[a], Philip D. Healy[a], John P. Morrison[a], Theo Lynn[b]

[a] Irish Centre for Cloud Computing & Commerce, University College Cork, Ireland
[b] Irish Centre for Cloud Computing & Commerce, Dublin City University, Ireland

HIGHLIGHTS

- Surveyed monitoring tools revealing common characteristics and distinctions.
- Identified practical capabilities of monitoring tools.
- Presented taxonomy of monitoring capabilities.
- Analysed strengths and weakness of monitoring tools based on taxonomy.
- Discussed challenges and identified future research trends in Cloud monitoring.

ARTICLE INFO

Article history:
Received 17 September 2013
Received in revised form
2 May 2014
Accepted 19 June 2014
Available online 5 July 2014

Keywords:
Cloud management
Monitoring tools
Cloud operational areas
Capabilities
Taxonomy
Survey

ABSTRACT

The efficient management of Cloud infrastructure and deployments is a topic that is currently attracting significant interest. Complex Cloud deployments can result in an intricate layered structure. Understanding the behaviour of these hierarchical systems and how to manage them optimally are challenging tasks that can be facilitated by pervasive monitoring. Monitoring tools and techniques have an important role to play in this area by gathering the information required to make informed decisions. A broad variety of monitoring tools are available, from general-purpose infrastructure monitoring tools that predate Cloud computing, to high-level application monitoring services that are themselves hosted in the Cloud. Surveying the capabilities of monitoring tools can identify the fitness of these tools in serving certain objectives. Monitoring tools are essential components to deal with various objectives of both Cloud providers and consumers in different Cloud operational areas. We have identified the practical capabilities that an ideal monitoring tool should possess to serve the objectives in these operational areas. Based on these identified capabilities, we present a taxonomy and analyse the monitoring tools to determine their strength and weaknesses. In conclusion, we present our reflections on the analysis, discuss challenges and identify future research trends in the area of Cloud monitoring.

1. Introduction

The emergence of Cloud Computing has ushered in a new era of Internet-based service provisioning opportunities. Cloud Computing is characterised by the provision of resources as general utilities that can be leased and released in an on-demand manner. Consequently, IT resources represent an operational rather than a capital expenditure. A broad variety of pricing models can be applied to Cloud resources, from simple fixed rental schemes to pay-as-you-go models. Monitoring techniques are indispensable in order to manage large-scale Cloud resources and enforce quality of service for consumers.

Given the multi-tenant nature of Cloud environments, efficient management in the face of quality of service and performance constraints can be a challenge. Monitoring tools have an important role to play in these areas by allowing informed decisions to be made regarding resource utilisation. Automated monitoring of physical and virtual IT resources allows for the identification and resolution of issues with availability, capacity, and other quality requirements. The benefits of automated monitoring have long been recognised, even in non-Cloud environments. The importance of monitoring has been widely addressed in the literature in

* Corresponding author.
E-mail address: k.fatema@cs.ucc.ie (K. Fatema).

http://dx.doi.org/10.1016/j.jpdc.2014.06.007
0743-7315/

various contexts, such as: system/network [13,35,71], distributed systems/Grid [93,4,95], application [47,10] and Cloud [1,33]. For Cloud environments, appropriate monitoring is crucial as usage-based billing and elastic scaling are impossible to implement in the absence of relevant metrics. Currently, a variety of Cloud monitoring tools is applied in an *ad-hoc* and non-systematic way, everywhere from low-level, general-purpose infrastructure monitoring to high-level application and service monitoring. The purpose of this paper is to comprehensively review these tools to assess whether they are adequate in satisfying the essential objectives for measuring intrinsic Cloud behaviours.

The focus of our work is to capture the evolutionary adaptation of monitoring tools' capabilities from general purpose to Cloud monitoring and to present a full capability analysis with respect to practical Cloud operational areas that would help Cloud providers and customers in making an informed choice of an appropriate monitoring tool. The monitoring platforms considered in this paper have been chosen based on literature reviews and perceived industrial acceptance.

The main contributions of this paper can be summarised as follows: (i) it surveys the range of monitoring tools currently in use to gain their technical insights, (ii) it identifies the desired capabilities of monitoring tools to serve different Cloud operational areas from both providers' and consumers' perspectives, (iii) it then presents a taxonomy of the identified capabilities, (iv) it analyses the available tools based on the identified capabilities and unveils the capabilities that are under-represented, Cloud operational areas that are currently strongly supported by those monitoring tools and the areas that need further development and, (v) it discusses future research challenges and trends in Cloud monitoring and management.

Our paper flows as follows: First, we assign the tools into categories, Cloud specific and non-Cloud specific. After studying the tools and extracting technical capabilities, we summarise these in Tables 1 and 2. We then focus on Cloud-specific monitoring capabilities and derive a taxonomy, which we present in Section 4. We then re-examine all of the tools again in light of the taxonomy in Section 5 in order to identify their strengths and weakness when applied to particular operational areas.

The remainder of this paper is organised as follows: Section 2 provides information on various computing environments, ranging from single machine to various distributed systems, and the usage of monitoring in those environments. In Section 3, available monitoring tools are identified and described. Section 4 describes taxonomy of desirable monitoring capabilities that forms the basis for the analysis of the identified monitoring tools. In Section 5, we analyse the identified monitoring tools. Section 6 presents analysis of the related work. Section 7 discusses the identified challenges while Section 8 concludes the paper and focuses on the future research trends in Cloud monitoring.

2. Background

Monitoring tools have long been used for tracking resource utilisation and the performance of systems and networks. These tools have traditionally been administrated by a single administrator in a single domain. Subsequently, the development of distributed systems like clusters forced the evolution of monitoring tools to meet the demand of these new environments. As more sophisticated distributed environments emerged, including Grids and Clouds, monitoring tools had to be developed to capture their salient characteristics. In the case of Grid these included computations that span multiple administrative domains and in the case of Clouds on-demand multi-tenant services. According to National Institute of Standards and Technology (NIST) definition [58]: *"Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction".*

Cloud computing has shifted computation from local machines to services accessed via the network. Services in Cloud are typically offered via three different service models which can be viewed as a layered model of services: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS).

IaaS providers, such as Amazon EC2, offer virtual resources, such as machine instances, storage, and networking devices, as a service to the end user and/or the layers above, enabling self-service for virtualized resources. A virtual machine manager, or hypervisor, is required in order to make the physical resources available to virtual machine instances. Through the application of clustering techniques, multiple physical servers can be aggregated into a resource pool from which storage, CPU, memory and networking resources can be dynamically allocated to a set of VMs. Clustering of the resource pool ensures high availability in the presence of hardware failures.

PaaS providers utilise the virtual machines' environment (e.g. operating systems and tools) to provide a scalable abstractions of the underlying resources onto which applications can be deployed. For example, Google AppEngine provides developers with scalable application back end functionality for dynamic web serving, automatic scaling and load balancing [86]. The resulting abstraction is, from the developer's point of view, completely divorced from the underlying infrastructure. Similar offerings exist for Java EE (*e.g.*, Red Hat OpenShift [61]) and Ruby on Rails (*e.g.*, Engine Yard [73]), among a host of others. Some Cloud service providers, such as Microsoft Azure, offer tools and features as part of their offerings that simplify the development of PaaS services [56]. Open-source tools such as Nimbus [62] and Cloudify [32] are available that simplify PaaS integration in a provider-agnostic fashion.

SaaS (Software-as a-Service) providers offer software as a service which may hide the service implementation, thus the SaaS customer is not necessarily aware of the underlying platforms, infrastructure, or hardware.

Given the rich architecture of Cloud, effective monitoring requires an appropriate suite of tools capable of monitoring in the IaaS, PaaS and SaaS layers.

3. Review of existing monitoring systems

In the light of the above discussion, we examined the available monitoring tools and divided them into two broad categories: (i) general-purpose monitoring tools; and (ii) Cloud-specific monitoring tools. We examine each of these categories in turn in subsequent subsections to gain an understanding of their common characteristics and functionalities. The technical details, as well as its reported limitations and usage experiences are summarised in Tables 1 and 2. These tabular presentations give readers the opportunity to gain technical insights into the tools. Furthermore, the presented information also helps to identify the features of the tools, which are later used for capability based analysis and the development of a taxonomy.

3.1. General-purpose monitoring tools

Before the advent of Cloud Computing, a number of tools were already available for the purpose of monitoring diverse IT infrastructure resources such as networks and compute nodes. Some specialise in particular domains such as HPC clusters and Grids. Many of these tools continue to be developed, and could be adopted in Clouds for monitoring at various abstraction levels. In