



ELSEVIER

Contents lists available at ScienceDirect

Data in Brief

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



Data Article

Data on performance prediction for cloud service selection



Abdullah Mohammed Al-Faifi*, Biao Song,
Mohammad Mehedi Hassan, Atif Alamri, Abdu Gumaei

College of Computer and Information Science, King Saud University, Riyadh, Saudi Arabia

ARTICLE INFO

Article history:

Received 10 March 2018

Accepted 24 August 2018

Available online 31 August 2018

Keywords:

Performance metrics

Workload parameters

Cloud computing

ABSTRACT

This paper contains data on Performance Prediction for Cloud Service Selection. To measure the performance metrics of any system you need to analyze the features that affect these performance, these features are called "workload parameters". The data described here is collected from the KSA Ministry of Finance that contains 28,147 instances from 13 cloud nodes. It was recorded during the period from March 1, 2016, to February 20, 2017, in continuous time slots. In this article we selected 9 workload parameters: Number of Jobs in a Minute, Number of Jobs in 5 min, Number of Jobs in 15 min, Memory Capacity, Disk Capacity, Number of CPU Cores, CPU Speed per Core, Average Receive for Network Bandwidth in Kbps and Average Transmit for Network Bandwidth in Kbps. Moreover, we selected 3 performance metrics: Memory utilization, CPU utilization and response time in milliseconds. This data article is related to the research article titled "An Automated Performance Prediction Model for Cloud Service Selection from Smart Data" (Al-Faifi et al., 2018) [1].

© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

DOI of original article: <https://doi.org/10.1016/j.future.2018.03.015>

* Corresponding author.

E-mail addresses: abdullah_ah50@yahoo.com (A.M. Al-Faifi), bsong@ksu.edu.sa (B. Song), mmhassan@ksu.edu.sa (M.M. Hassan), atif@ksu.edu.sa (A. Alamri), abdugumaei@gmail.com (A. Gumaei).

<https://doi.org/10.1016/j.dib.2018.08.108>

2352-3409/© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Specifications Table

Subject area	Computer Science
More specific subject area	Performance prediction, cloud computing
Type of data	Tables
How data was acquired	Data was collected from the KSA Ministry of Finance that contains 28,147 instances from 13 cloud nodes. It was recorded during the period from March 1, 2016, to February 20, 2017, in continuous time slots. It is collected using manage engine (application manager) and solar winds (virtualization manager software).
Data format	Raw data with class labels
Experimental factors	A set of attributes include the number of Jobs in a Minute, number of Jobs in 5 min, a number of Jobs in 15 min, memory capacity, disk capacity, number of CPU cores, CPU speed per core, average receive for network bandwidth in Kbps, and average transmit for network bandwidth in Kbps. A set of predictors are memory utilization, CPU utilization and response time.
Experimental features	The experiment aims to build two prediction models. The first model is used to learn from labeled workload attributes and predict memory utilization, CPU utilization, and response time. The data set used in this model contains 28,147 instances. A random subset of 2450 instances is utilized as a testing set. The second model is used to learn from CPU utilization and response time of one model type node as a benchmark and predict CPU utilization and response time of another model type node.
Data source location	Ministry of Finance, Riyadh, Saudi Arabia
Data accessibility	Data is available with this article

Value of the data

- This dataset is important to the field of performance prediction and cloud computing as it provides a log of workload parameters as well as performance metrics.
- The data could be used as a benchmark for performance prediction.
- Analysis of the data can provide direction towards enhancing the performance of the systems and helping in identifying the resources required before in migrating to cloud service.

1. Data

The supplementary dataset contains 28,147 instances from 13 cloud nodes. These data was recorded during the period from March 1, 2016, to February 20, 2017, in continuous time slots. These data contains nine workload parameters include: Number of Jobs in a Minute, Number of Jobs in 5 min, Number of Jobs in 15 min, Memory Capacity, Disk Capacity, Number of CPU Cores, CPU Speed per Core, Average Receive for Network Bandwidth in Kbps and Average Transmit for Network Bandwidth in Kbps. Other than that, three performance metrics were selected include: Memory utilization, CPU utilization and response time in milliseconds. Details of the testing scenario can be found in section 4 in [1].

The supplementary files contains all 28,147 instances of both workloads and performance metrics.

Download English Version:

<https://daneshyari.com/en/article/11000409>

Download Persian Version:

<https://daneshyari.com/article/11000409>

[Daneshyari.com](https://daneshyari.com)