Soft Computing
Job Performance Evaluation Using Back-propagation Network

Bc. Petr Stehlík <xstehl14@stud.fit.vutbr.cz>

1 Introduction

Every user of a supercomputer need to know whether their submitted job finished successfully and performed well. So far these tedious tasks are usually performed manually using only the output of their program and over-simplified metrics such as job run time and utilizied resources.

The aim of this project is to create a back-propagation neural network which can determine whether a job run was successful or in some way suspicious of unwanted behaviour such as poor performance or execution failure.

The network is supplied with fine-granular metric data acquired via Examon framework[2] which was run on Galileo supercomputer located in CINECA, Bologna, Italy.

The structure of this document is as follows: in section 2 the theoretical background needed for this project is presented together with the description of Examon framework. In the following section 3 the data supplied to the network are described as well as the final format of the data. Afterwards in section 4 the implementation of the network and its structure are laid out and in the last two sections 5 and 6 the achieved results, summary and further work are discussed.

2 Theoretical Background

2.1 Backpropagation Network

Backpropagation networks are multi-layer feed-forward networks with supervised learning. There is no interconnection between neurons in the same layer but layers are fully connected to each other in order to be able to do forward and backward propagation.

Each neuron disposes of a weight for each input initialized to a random value in range $< 0, 1 >$, activation function and transfer function. The input neurons

2.1.1 Feedforward Propagation

2.1.2 Backward Propagation

2.2 Examon

Examon framework is used for exascale monitoring of supercomputing facilities. It is built on top of MQTT protocol[4] which allows measured metrics to be send to a central broker where received data are processed and stored in KairosDB[1] database utilizing Cassandra[3] cluster.

KairosDB is used for storing metric data in time-series format whereas Cassandra, serving as a backend for KairosDB is also used for storing job-related data. More on data semantics is described in 3.

3 Data

As previously stated in 2.2 data are gather via Examon framework and stored in Cassandra database. We can split the data into two categories. Job data and metric data.

The job data come from

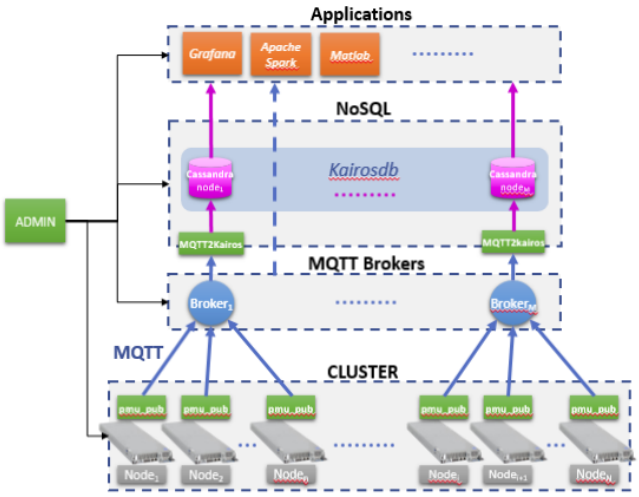

Figure 1: Examon framework architecture

4 Implementation

5 Achieved Results

6 Summary

References

[1] KairosDB. `https://kairosdb.github.io`. Accessed: 2017-11-25.

[2] F. Beneventi, A. Bartolini, C. Cavazzoni, and L. Benini. Continuous learning of hpc infrastructure models using big data analytics and in-memory processing tools. In *Design, Automation Test in Europe Conference Exhibition (DATE), 2017*, pages 1038–1043, March 2017.

[3] A. Lakshman and P. Malik. Cassandra: a decentralized structured storage system. *ACM SIGOPS Operating Systems Review*, 44(2):35–40, 2010.

[4] D. Locke. Mq telemetry transport (mqtt) v3. 1 protocol specification. *IBM developerWorks Technical Library*, 2010.