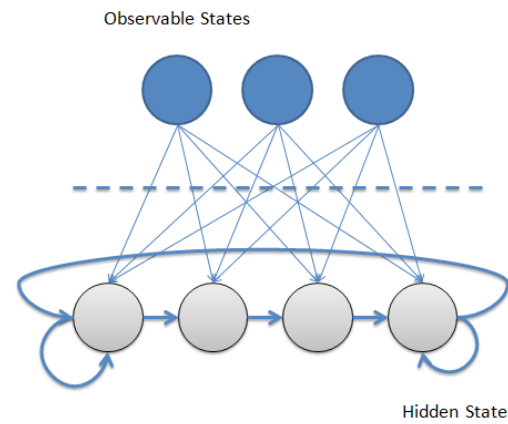




Master Thesis

Statistical analysis of network traffic in a Cloud Data Centre

2017-10-09

Context

With the increasing adoption of cloud infrastructure by different types of applications, network performance inside a data centre becomes a concern due to the time varying resource usage behaviour of the applications. In a cloud data centre, traffic sources are usually the virtual machines where the applications are being run. The performance of the network and these applications largely depends on how VMs are placed, since inappropriate placement of these VMs into a set of physical hosts can cause high load inside the network with a consequence of congestion, increased latency and performance degradation of the applications. Such consequence directly hampers QoS, cost and profit. In order to determine the network resource demands of the applications such as bandwidth as well as to predict how the network will perform in future for a certain period of time, it is essential to analyse the network traffic at first and then to find an appropriate traffic model that can represent the VMs' traffic pattern with a certain accuracy. Under H2020 CloudPerfect project, a tool called Allocation Optimiser has been designed which aims to assess the feasibility of placement of a new VM at each host with regard to their predicted availability of network resources over a specific time period, based on the statistical processing of monitoring data of network resource utilisation (in terms of data rates) over time, retrieved from VMs and hosts. The statistical models have been created from simple probability theory.

Scope of the Thesis

The main goal of this thesis is to find a method to enhance the probability distribution models by applying suitable traffic models. At first, the student should perform an in-depth study of different traffic models such as Bernoulli Process, Hidden Markov Model, Markov Modulated Fluid Models etc. and afterwards select the suitable ones that can be used to increase the accuracy level of the statistical models for better prediction, since accurate modeling of VMs' traffic can contribute to effective VM placement and capacity planning of the network. Finally the efficiency of the traffic models needs to be evaluated.

Requirements and Comments

1. Knowledge on fundamentals of cloud data centre and networks
2. Good knowledge in R/Python is a must.
3. Knowledge on statistical approaches

References

1. Chandrasekaran, Balakrishnan. "Survey of network traffic models." Washington University in St. Louis CSE 567 (2009).

If you are interested, please contact Stefan Wesner and Mitalee Sarker either by mail or directly in the office.

Mail: stefan.wesner@uni-ulm.de, **office:** Uni West, 43.2.210
mitalee.sarker@uni-ulm.de **office:** Uni West, 43.2.220

**Faculty of Engineering and
Computer Science**

Institute of Information
Resource Management