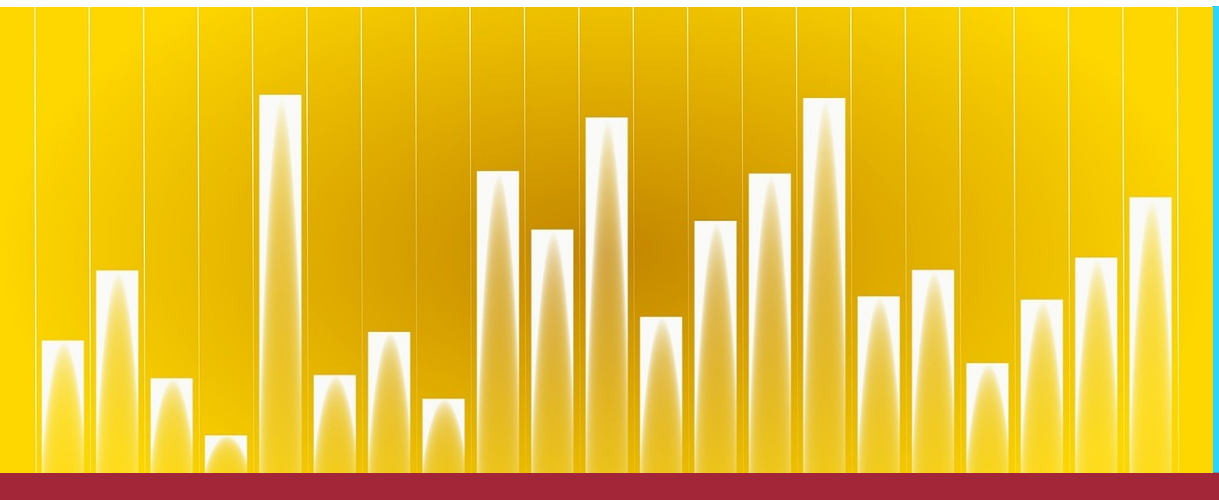
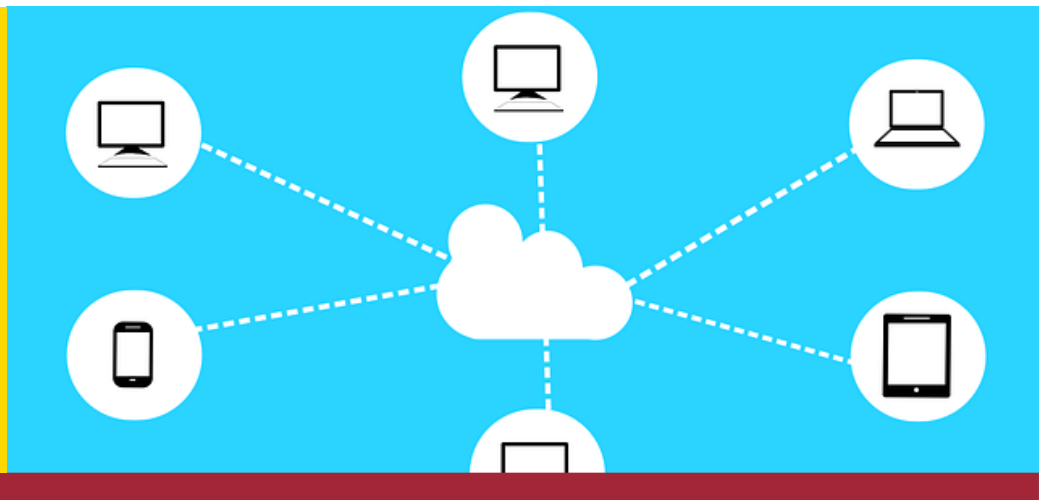
ulm university universität ulm





Master Thesis

2018-10-08

Classification of Virtual Machines from network traffic analysis

Context

Over the last years, the adoption of Cloud Infrastructure has not only increased in numbers but more and more resource demanding, and business critical applications are moved from a dedicated infrastructure towards shared Cloud based solutions. To cope with the increasing number of cloud applications, data centres are expanding at a high rate by adding up hundreds of thousands of servers and other necessary equipments.

A cloud provider wants to have a better resource management inside the infrastructure for guaranteed QoS and profit. When an experienced Cloud system administrator plans a manual deployment for a Cloud application, a couple of properties of the application are investigated beforehand. If a user is deploying without assistance from an operator, a set of base images is selected that indicates at least the application type. However, the application behaviour might be highly time variant, user dependent or use case driven (e.g. problem size).

Existing solutions do not consider a placement decision for a VM, the relationship between the new VM and other already deployed virtual machines. It is necessary to understand the level of communication between VMs, as it could be an important parameter to let VMs with a high cohesion to be put close to each other. An inappropriate placement of such VMs might lead to a scenario where they are placed at opposite ends of a data centre infrastructure, which will in turn stimulate high network load inside the data centre as well as latency. As a result, it will directly impact the job completion time of the applications running inside those VMs which in turn degrade their performance. The challenge that needs to be addressed is to find a way to classify the communication behaviour of the VMs for their better placement inside the physical hosts.

Scope of the Thesis

The first step is to find different correlations of VMs from monitored network traffic of a real cloud infrastructure. Those can be internal, communicative VMs. The later task is to model their communication pattern such as inter-arrival time, burstiness, direction of traffic e.g. one to many, all-to-all in order to determine their resource consumption. The student needs to come out with an approach to determine the type of VM communication such as closely coupled, strongly coupled, not coupled at all as well as different phases of the VM's life cycle such as initial phase, run-time phase, termination phase etc. Also possible pattern can be obtained from the offered VM flavors by the Cloud Providers. Finally the approach should be described in a statistical way/ mathematical way and should be validated in a simulation environment using different scenarios.

Requirements and Comments

1. Knowledge on fundamentals of cloud data centre and networks
2. Skills in a functional or object oriented programming language
3. Knowledge on statistical approaches

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