UNIVERSITÄT BERN

ACM SIGCOMM Workshop on Distributed Cloud Computing (DCC 2014)

Modelling and Simulation of Concurrent Workload Processing in Cloud-Distributed Enterprise Information Systems

Florian Antonescu[1,2], Torsten Braun[2]

[1]SAP, [2]Universität Bern

braun@iam.unibe.ch, cds.unibe.ch

Motivation

- Cloud-distributed enterprise applications can be monitored and dynamically scaled considering SLAs.
- Goals
  - Understanding of
    - factors influencing performance of distributed cloud applications
    - dependencies between physical resource utilization and application performance metrics for SLA-driven scaling
  - Optimization of
    - cloud resource utilization / allocation
    - SLAs
- Large-scale, reproducible experiments are needed.
  → realistic simulations

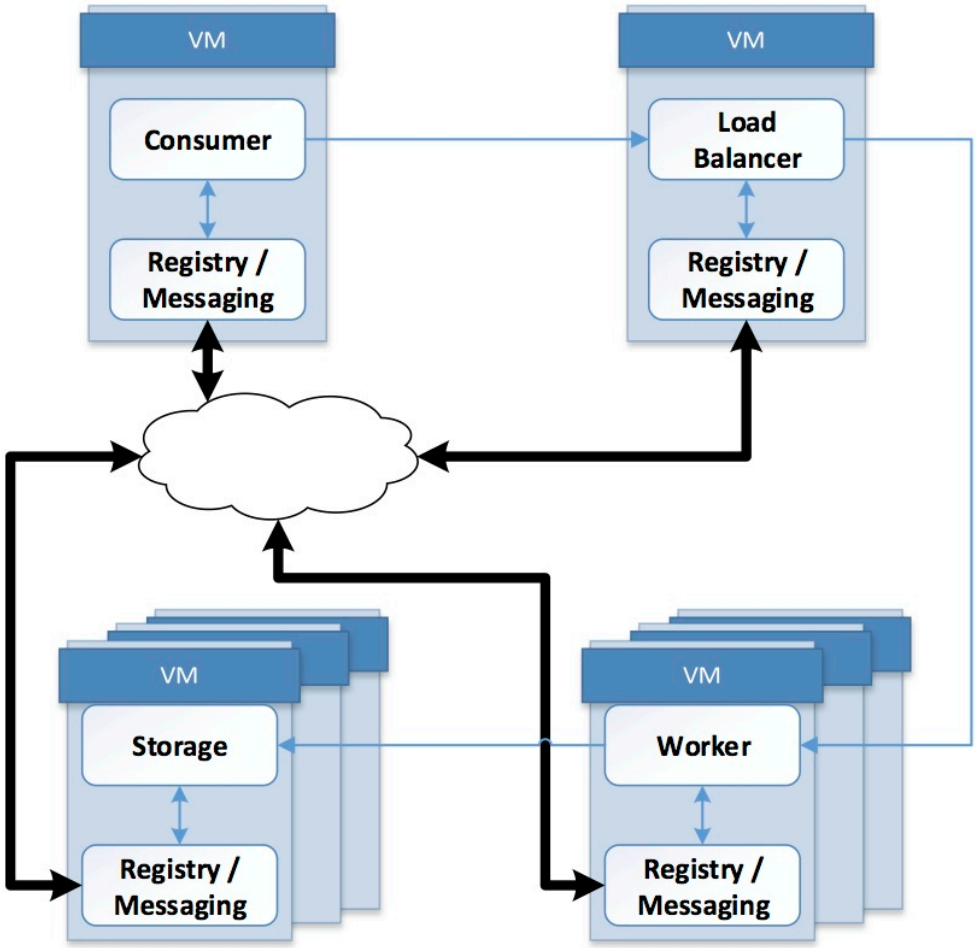
Methodology

1. Profiling of distributed Enterprise Information System (dEIS) application
2. Build dEIS simulation model using CloudSim
3. Validation of dEIS simulation model against real-world virtual machine experiments

CloudSim Simulator

- developed at University of Melbourne, www.cloudbus.org/cloudsim
- simulates
  - data center (power, inter-server network)
  - servers (CPU, memory, network, power)
  - VMs (CPU, memory, network)
  - cloudlets (CPU time, network I/O)

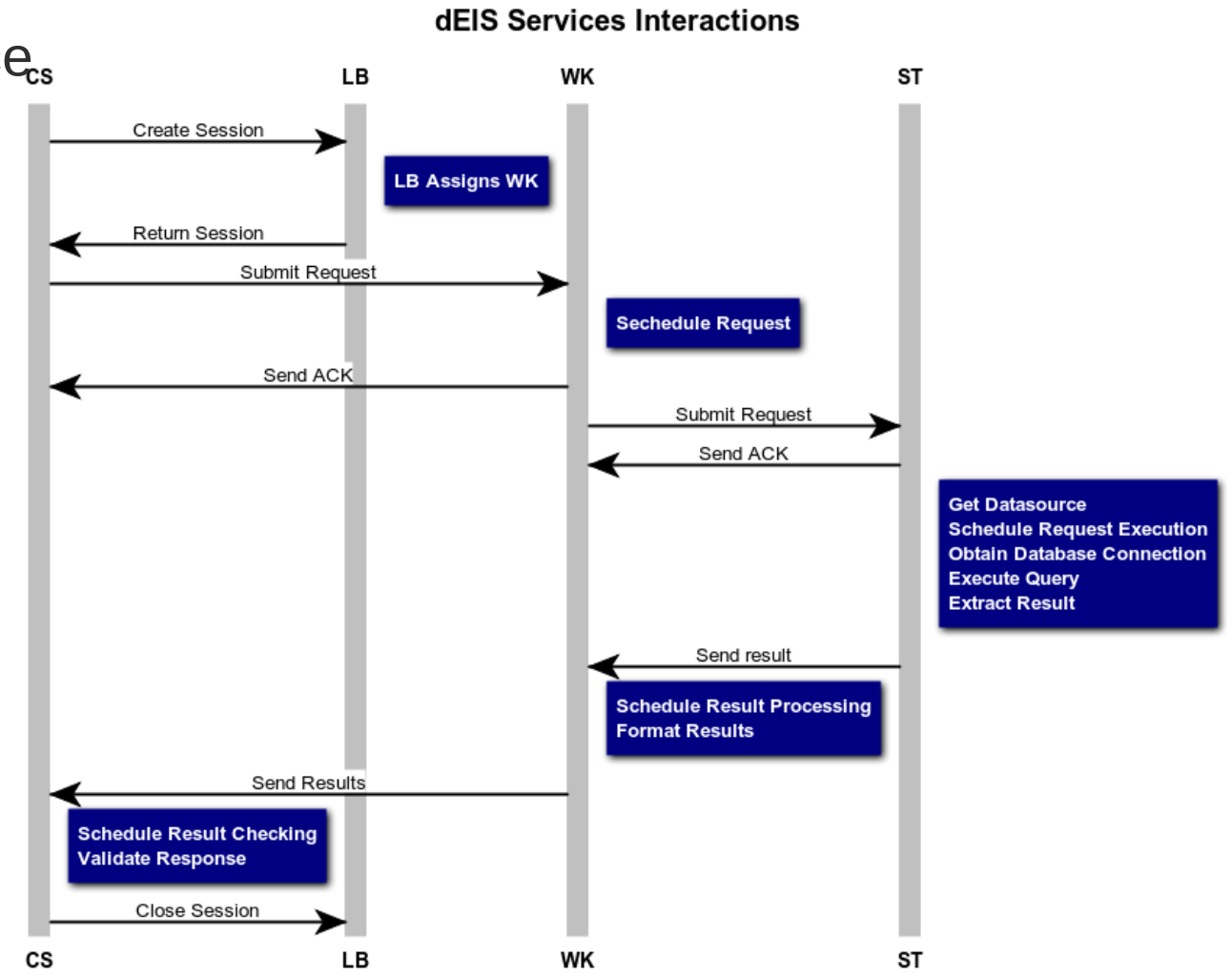
dEIS Topology



dEIS Service Interactions

In each VM / service measurement of

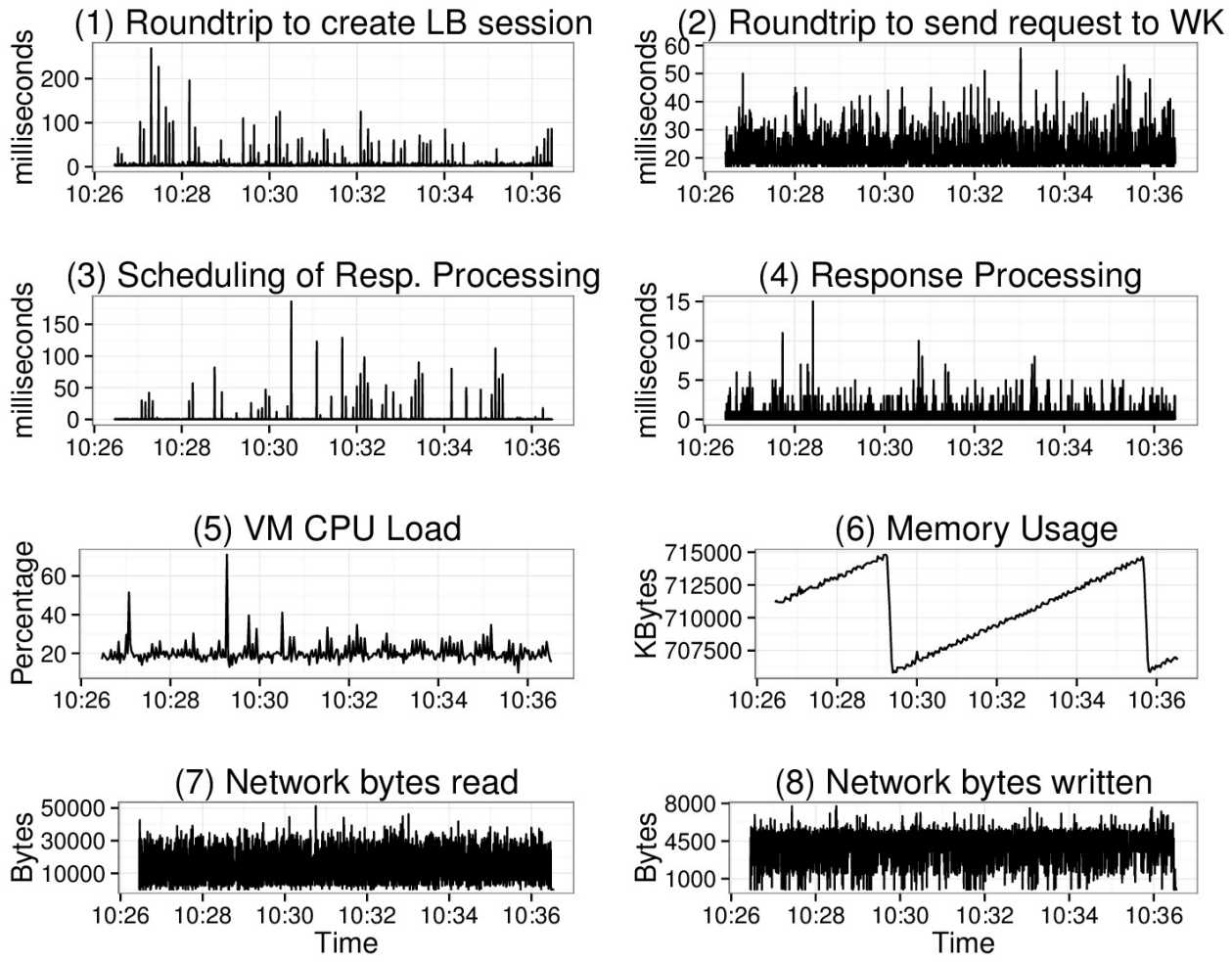
- service call durations
- memory
- network I/O



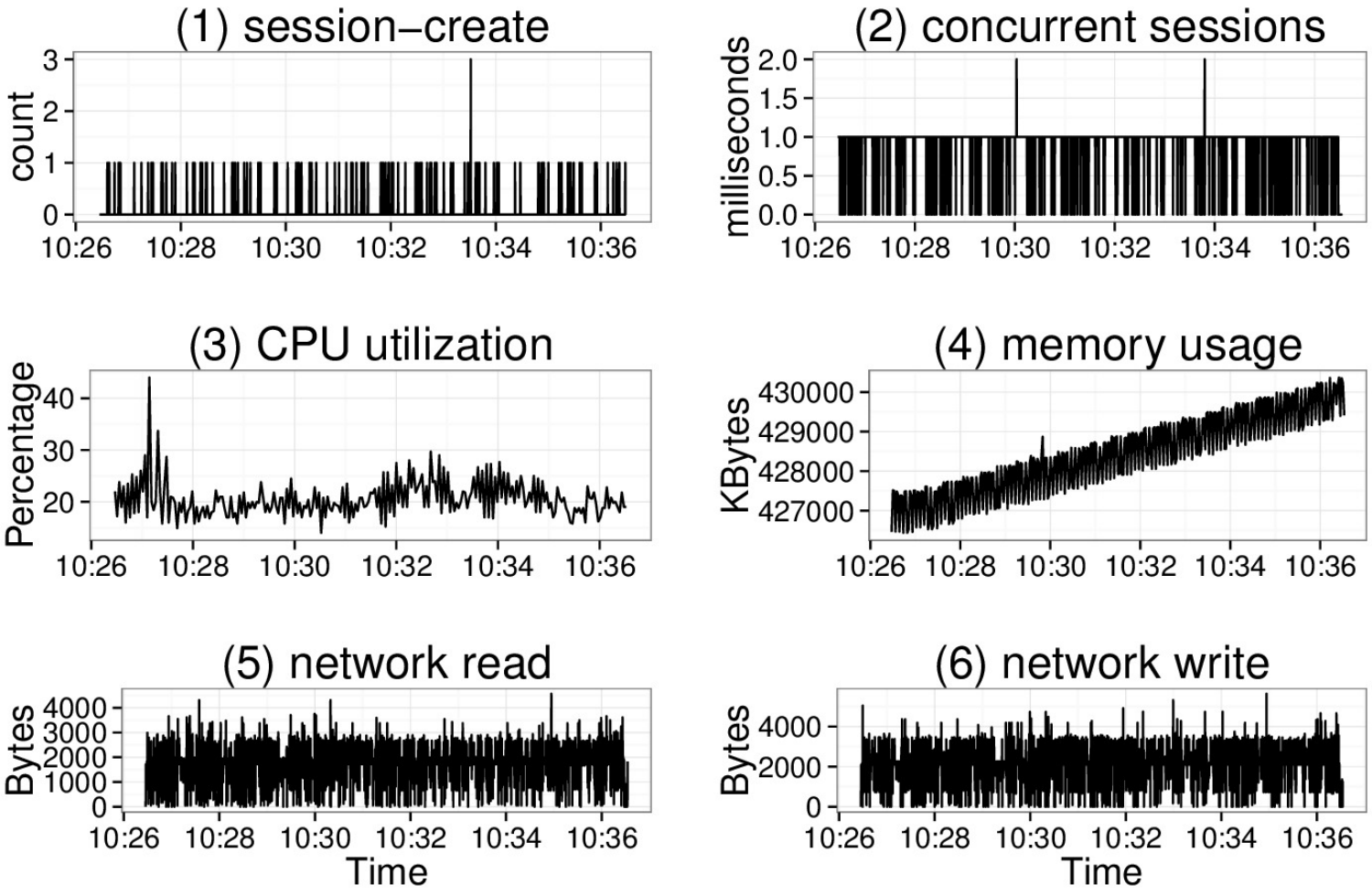
Physical Setup for Profiling

- 3 servers with
  - Intel Core 2 Duo CPU E8400 / 3 GHz
  - 4 GB memory
  - 1 Gbps network interface
- Consumer (CS) and Load Balancer (LB) on the same server, Worker (WK) and Storage (ST) on separate servers

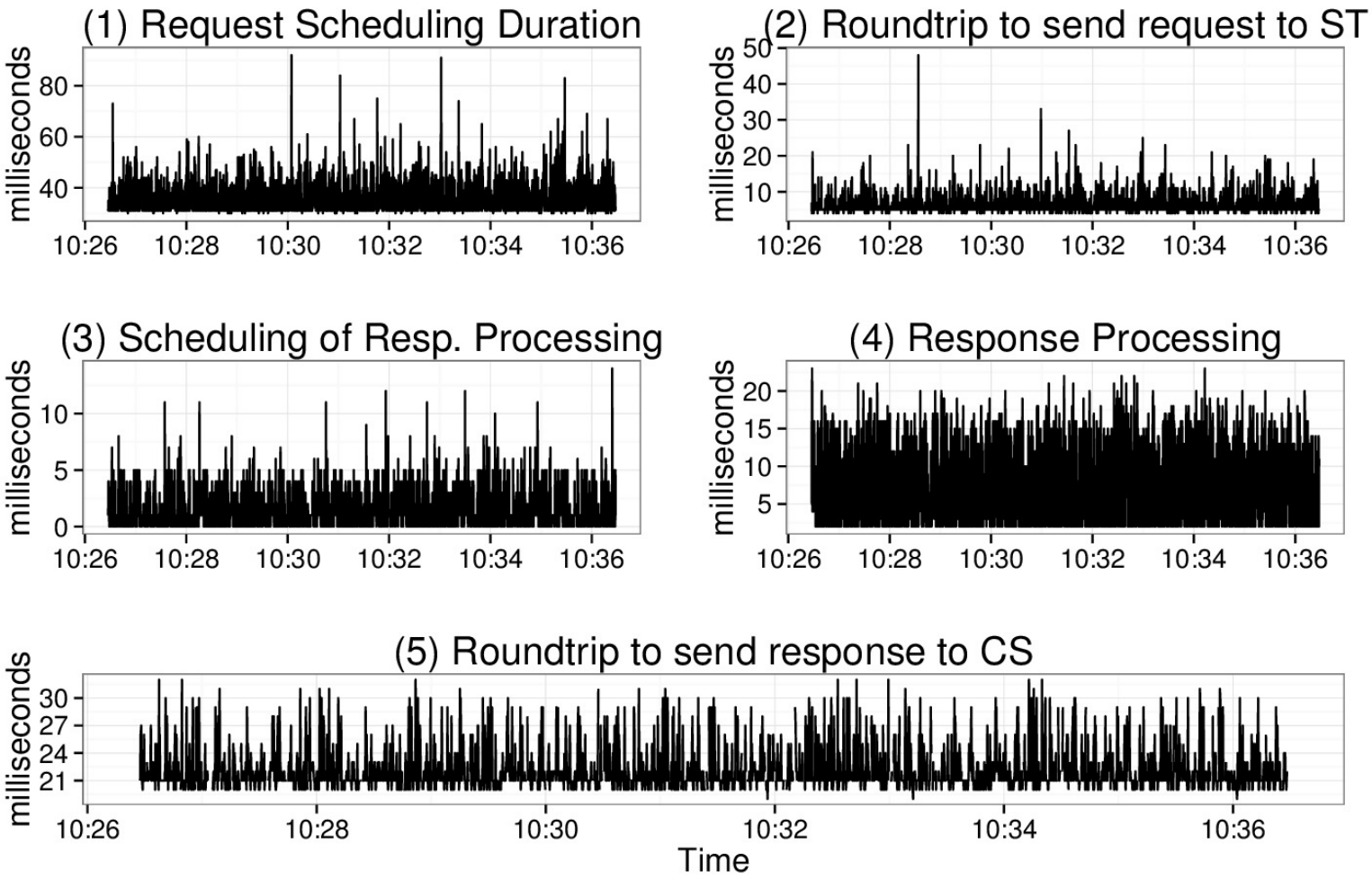
Consumer Service Performance



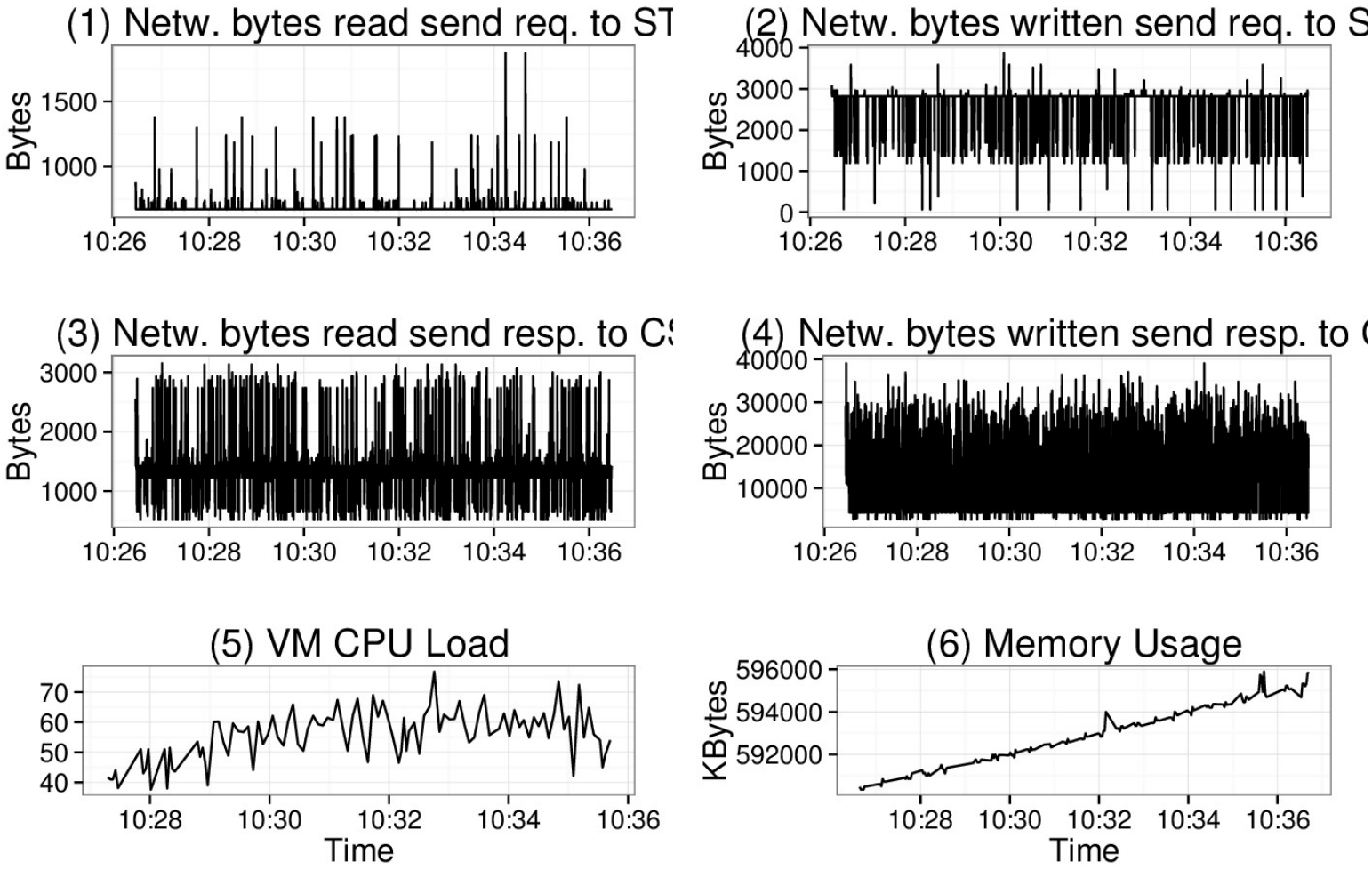
Load Balancer Service Performance



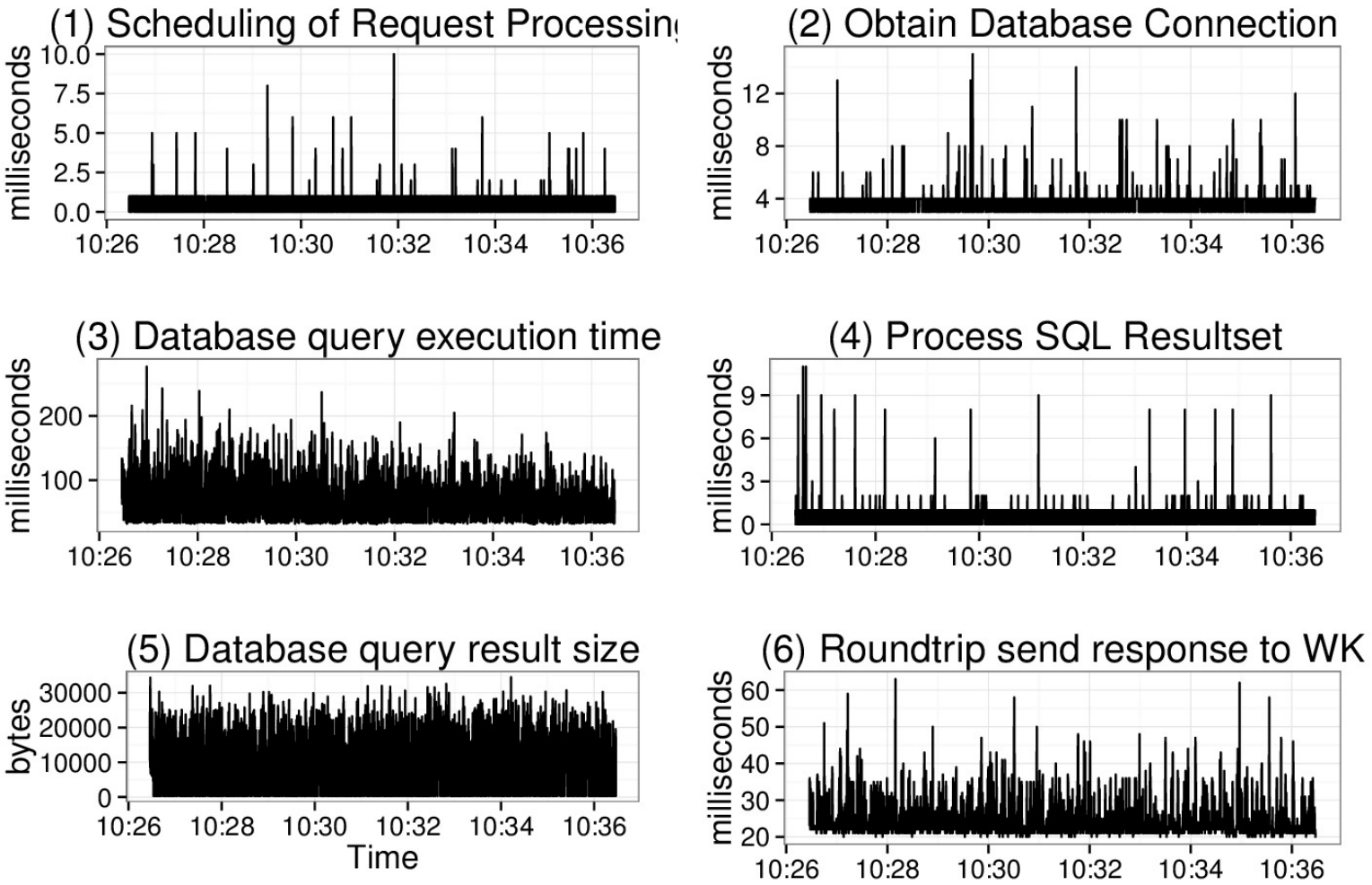
Worker Service Application Performance



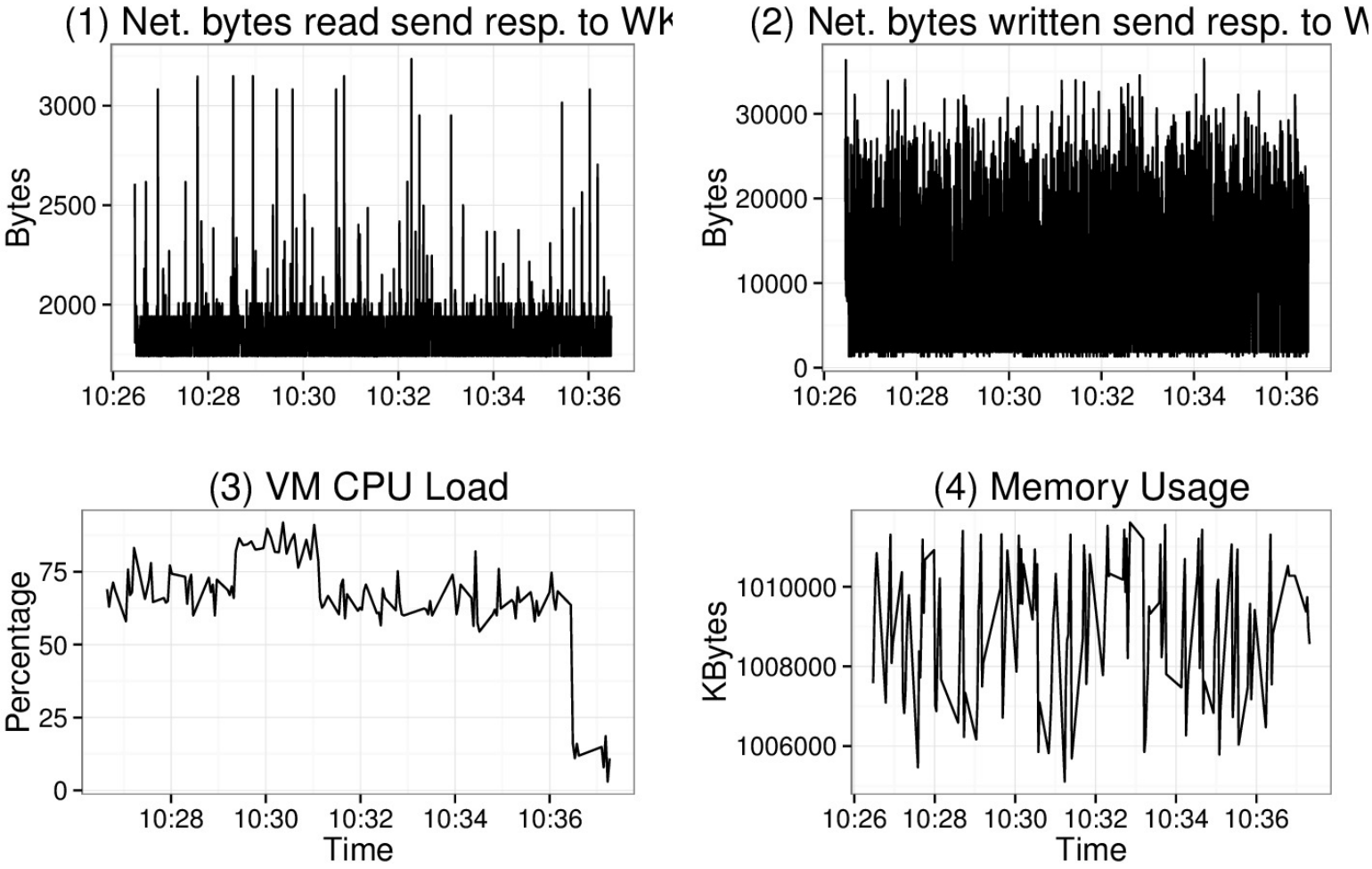
Worker Service VM Performance



Storage Service Application Performance



Storage Service VM Performance



Validation of Simulation Model

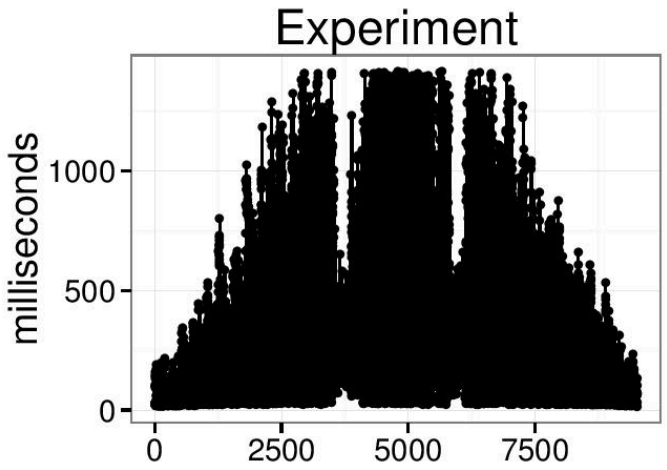
1. Perform varying load experiments using VMs on physical servers
2. Repeat experiment in CloudSim
3. Compare measured and simulated end-to-end response times

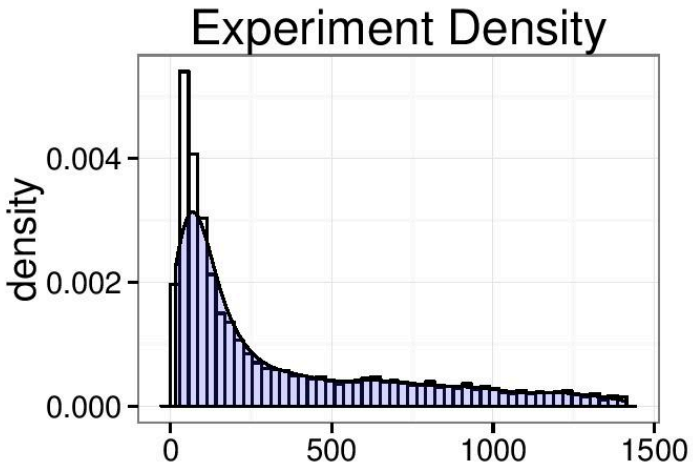
Modelling and Simulation

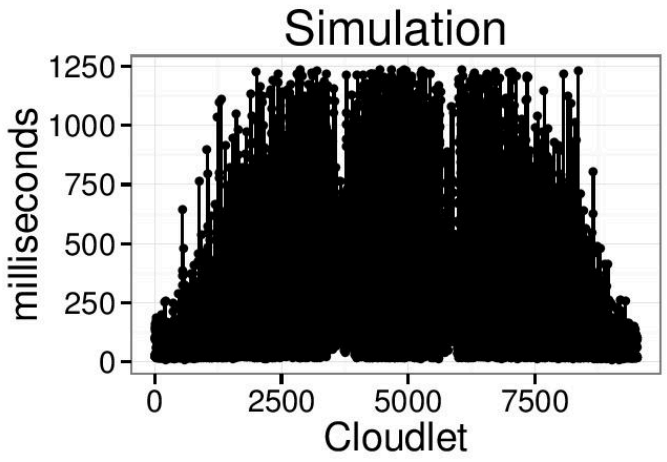
- Performance profile: values for
  - round-trip times of service calls between two services
  - local operation of a service
  - CPU, memory, and network utilization of the VMs running the services
- Dictionary of performance profiles for different load values (1 – 20)
- Transformation of times into numbers of instructions for the simulator
  - $instr = \frac{time\ [ms]}{1000} \cdot CPU_{MIPS}$; $instr^* = \frac{instr}{cl_{new} - concurrency_{VM}} \cdot (1 - CPU_{OS})$

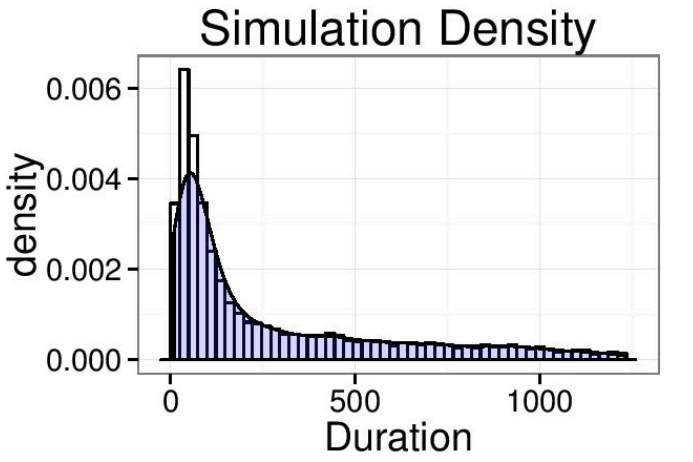
- Simulation scenario: list of pairs:
  (simulation time, expected number of concurrent CS requests;
- Lookup of parameters in performance profile (considering load) for running cloudlets in CloudSim

Measured and Simulated Worker Service Execution

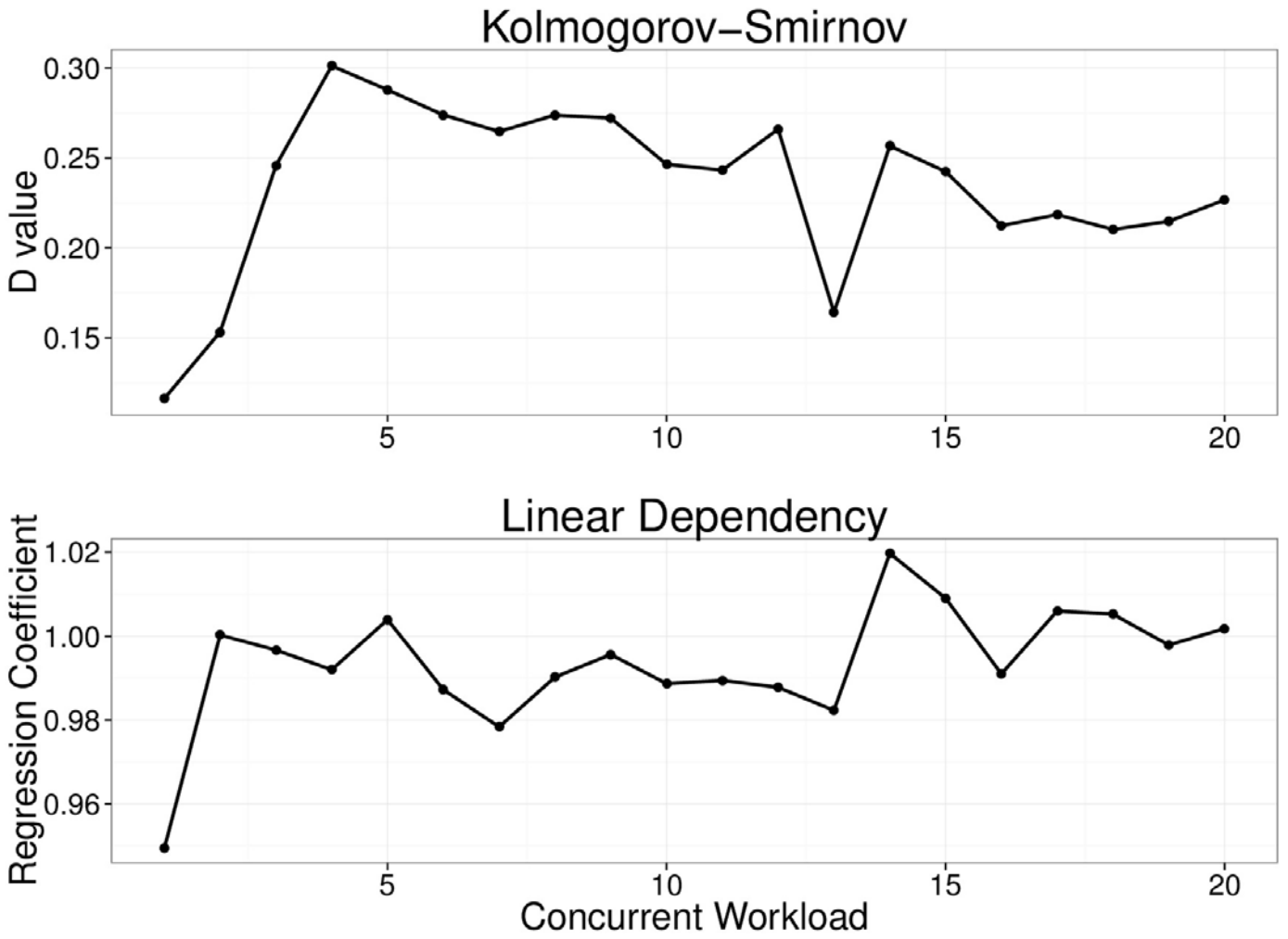




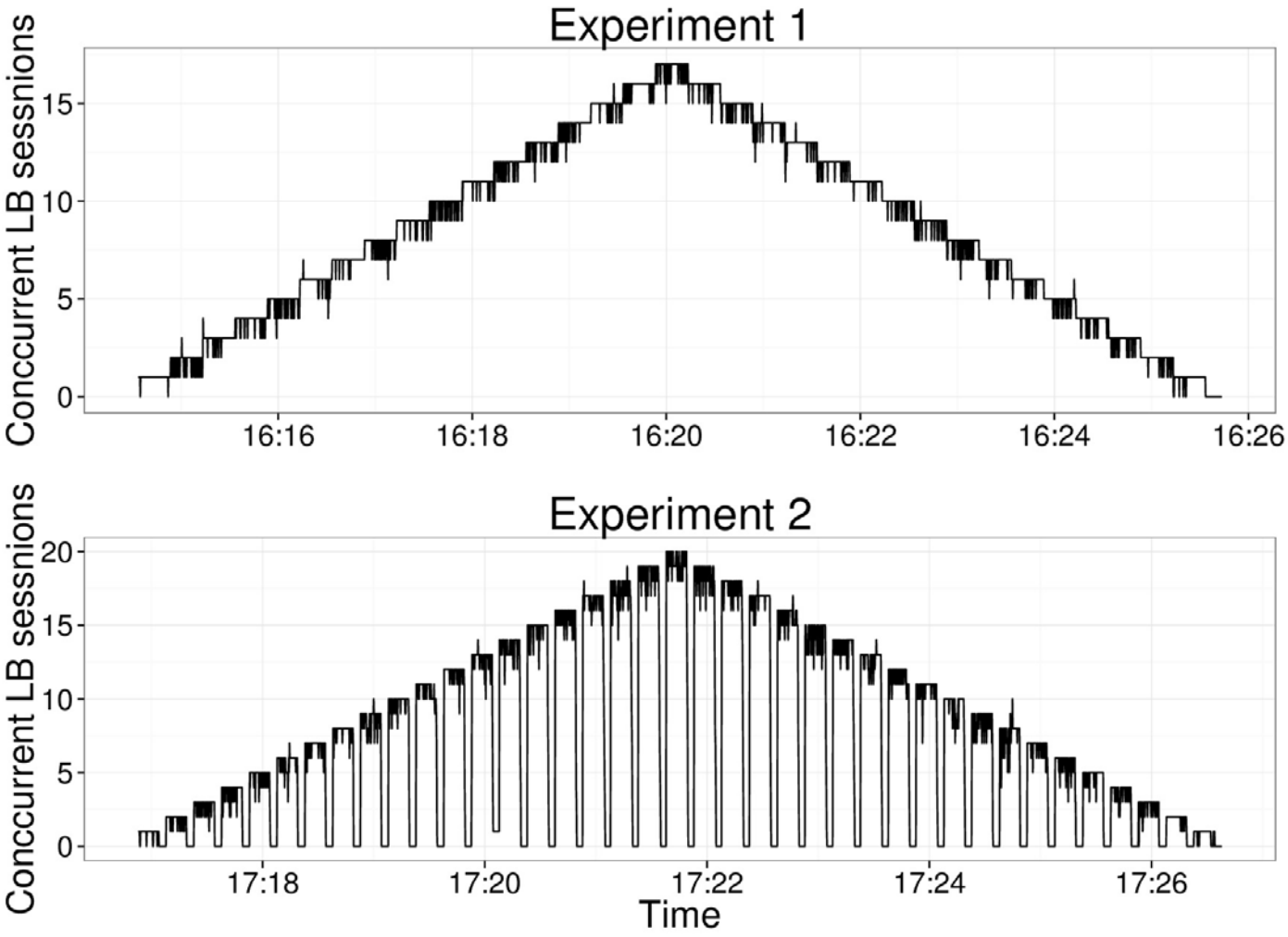




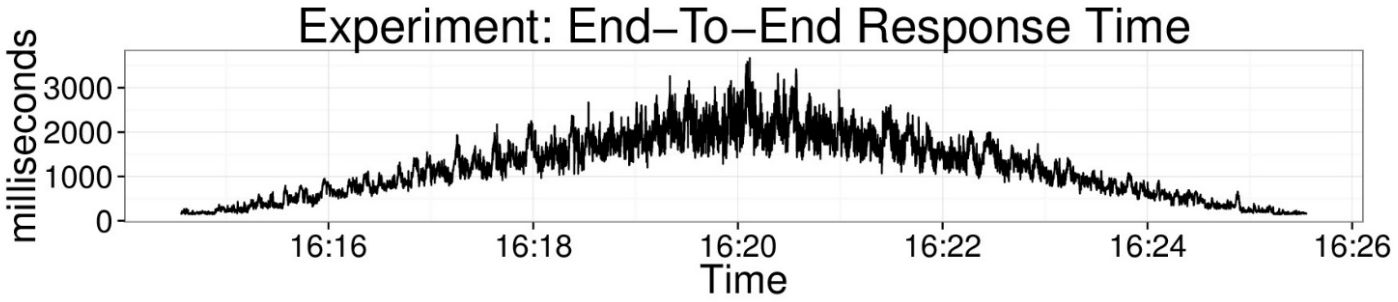
Comparison of Measured/Simulated End-to-End Response Times

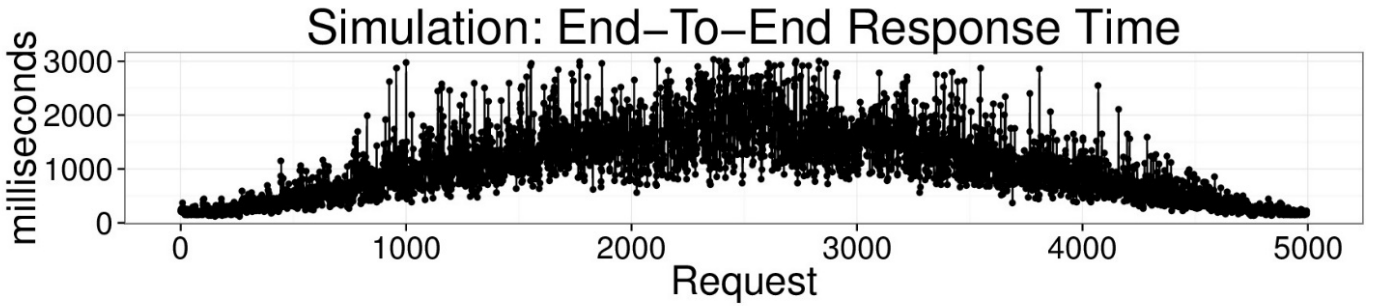


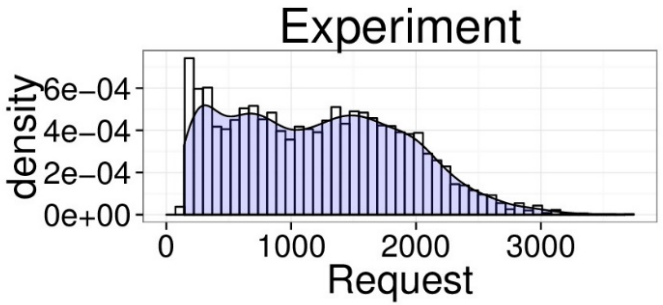
Experiments: Concurrent Load Balancer Sessions

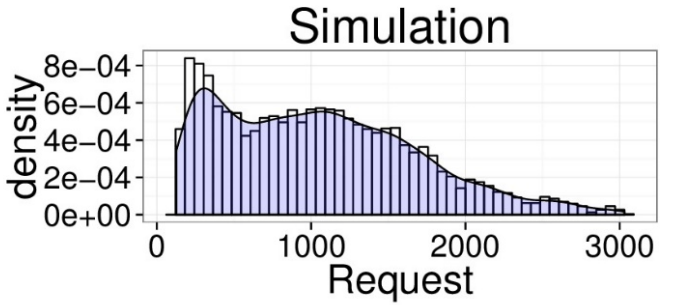


End-to-End Response Time (Experiment 1, 1 WK & 1 ST VM)

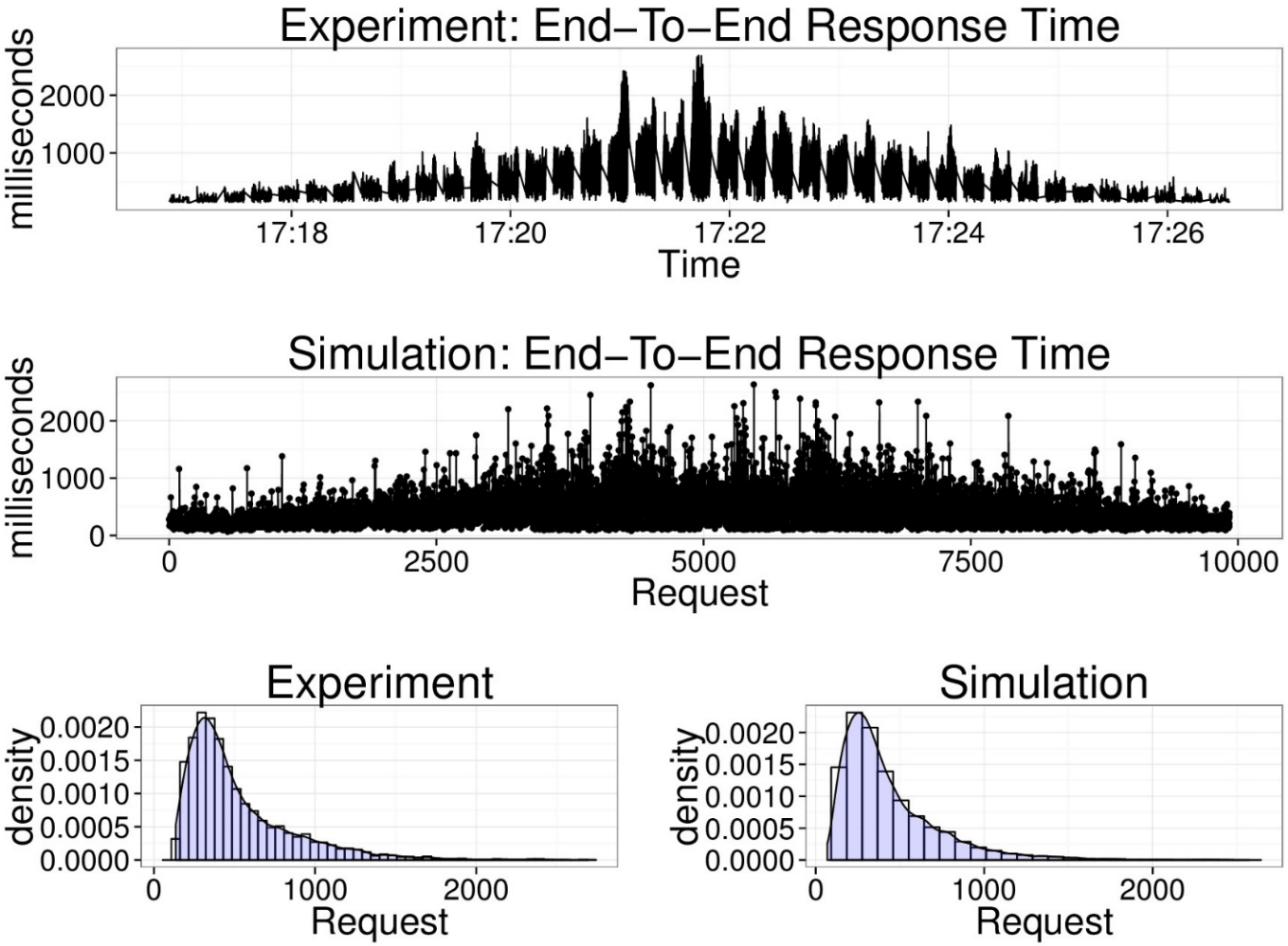








End-to-End Response Time (Experiment 2, 2 WK & 2 ST VMs)



Conclusions

- Summary
  - Implementation of a dEIS simulation model in CloudSim
  - Extension of CloudSim to support accurate concurrent simulations
- Future Work
  - Larger evaluation scenarios and improvement of modelling
  - Evaluation of different SLA scaling policies in larger scenarios

Thanks for Your Attention !

- braun@iam.unibe.ch
- cds.unibe.ch