CHALMERS
UNIVERSITY OF TECHNOLOGY

A Cloud Benchmark Suite Combining Micro and Application Benchmarks

Joel Scheuner, Philipp Leitner

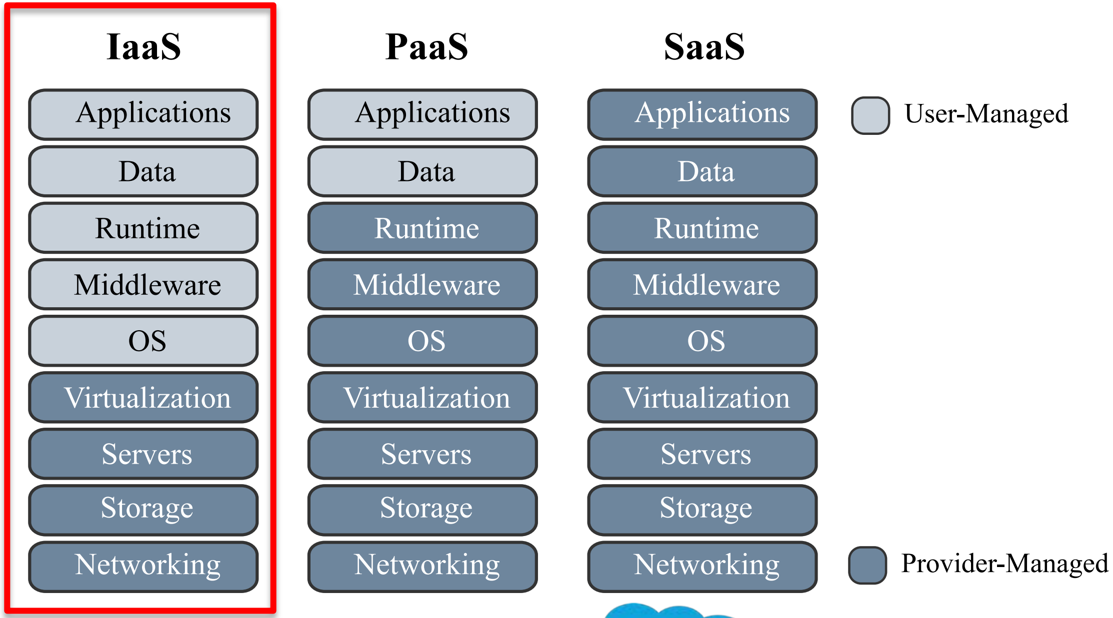
Joel Scheuner

scheuner@chalmers.se

joe4dev

@joe4dev

Context: Public Infrastructure-as-a-Service Clouds



amazon web services EC2 | HEROKU | salesforce

Infrastructure-as-a-Service (IaaS)
Platform-as-a-Service (PaaS)
Software-as-a-Service (SaaS)

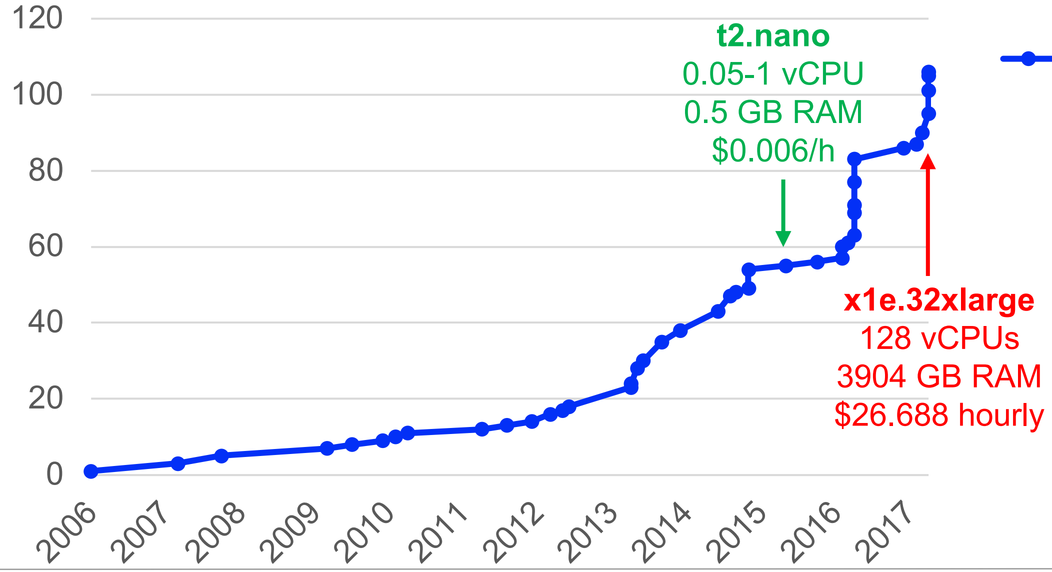
Motivation: Capacity Planning in IaaS Clouds

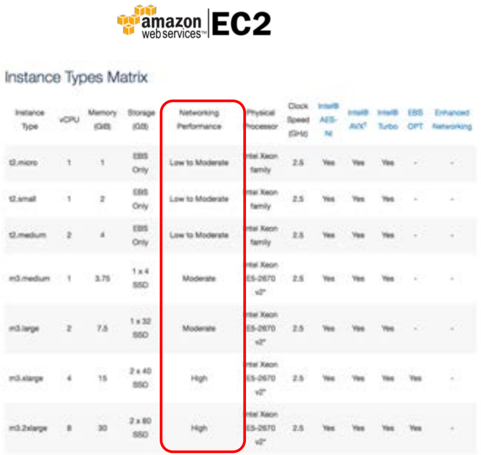
What cloud **provider** should I choose?

https://www.cloudorado.com

Motivation: Capacity Planning in IaaS Clouds

What cloud **service (i.e., instance type)** should I choose?





amazon web services EC2

Instance Types Matrix

Instance Type	vCPU	Memory (GiB)	Storage (GB)	Networking Performance	Physical Processor	Clock Speed (GHz)	Intel® AES-NI	Intel® AVX†	Intel® Turbo	EBS OPT	Enhanced Networking
t2.micro	1	1	EBS Only	Low to Moderate	ntel Xeon family	2.5	Yes	Yes	Yes	-	-
t2.small	1	2	EBS Only	Low to Moderate	ntel Xeon family	2.5	Yes	Yes	Yes	-	-
t2.medium	2	4	EBS Only	Low to Moderate	ntel Xeon family	2.5	Yes	Yes	Yes	-	-
m3.medium	1	3.75	1 x 4 SSD	Moderate	ntel Xeon E5-2670 v2*	2.5	Yes	Yes	Yes	-	-
m3.large	2	7.5	1 x 32 SSD	Moderate	ntel Xeon E5-2670 v2*	2.5	Yes	Yes	Yes	-	-
m3.xlarge	4	15	2 x 40 SSD	High	ntel Xeon E5-2670 v2*	2.5	Yes	Yes	Yes	Yes	-
m3.2xlarge	8	30	2 x 80 SSD	High	ntel Xeon E5-2670 v2*	2.5	Yes	Yes	Yes	Yes	-

Topic: Performance Benchmarking in the Cloud

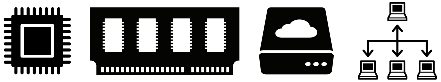
NETFLIX

"The instance type itself is a very major tunable parameter"

@brendangregg re:Invent'17
https://youtu.be/89fYOo1V2pA?t=5m4s

Background

Micro Benchmarks



CPU Memory I/O Network

Application Benchmarks

Overall performance
(e.g., response time)

Generic | Specific

Artificial | Real-World

Resource-specific | Resource-heterogeneous

Related Work

Micro Benchmarking / Application Kernels

Iosup et. al. Performance analysis of cloud computing services for many-tasks scientific computing.

Ostermann et. al. A performance analysis of EC2 cloud computing services for scientific computing.

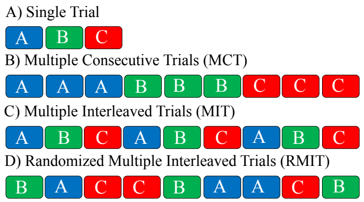
Application Benchmarking

Ferdman et. al. Clearing the clouds: a study of emerging scale-out workloads on modern hardware.

Cooper et. al. Benchmarking Cloud Serving Systems with YCSB.

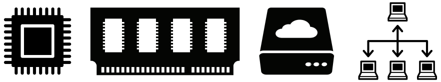
Repeatability of Cloud Experiments

Abedi and Brecht. Conducting Repeatable Experiments in Highly Variable Cloud Computing Environments. @ICPE’17



Problem: Isolation, Reproducibility of Execution

Micro Benchmarks



CPU Memory I/O Network

Generic

Artificial

Resource-specific

Application Benchmarks

Overall performance
(e.g., response time)

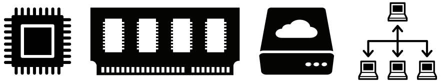
Specific

Real-World

Resource-heterogeneous

Question: **How** can we systematically **combine and execute** micro and application benchmarks?

Micro Benchmarks



CPU | Memory | I/O | Network

Generic

Artificial

Resource-specific

Application Benchmarks

Overall performance
(e.g., response time)

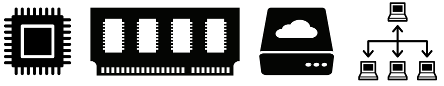
Specific

Real-World

Resource-heterogeneous

Idea

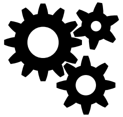
Micro Benchmarks



Generic

Artificial

Resource-specific



Systematically Execute Together

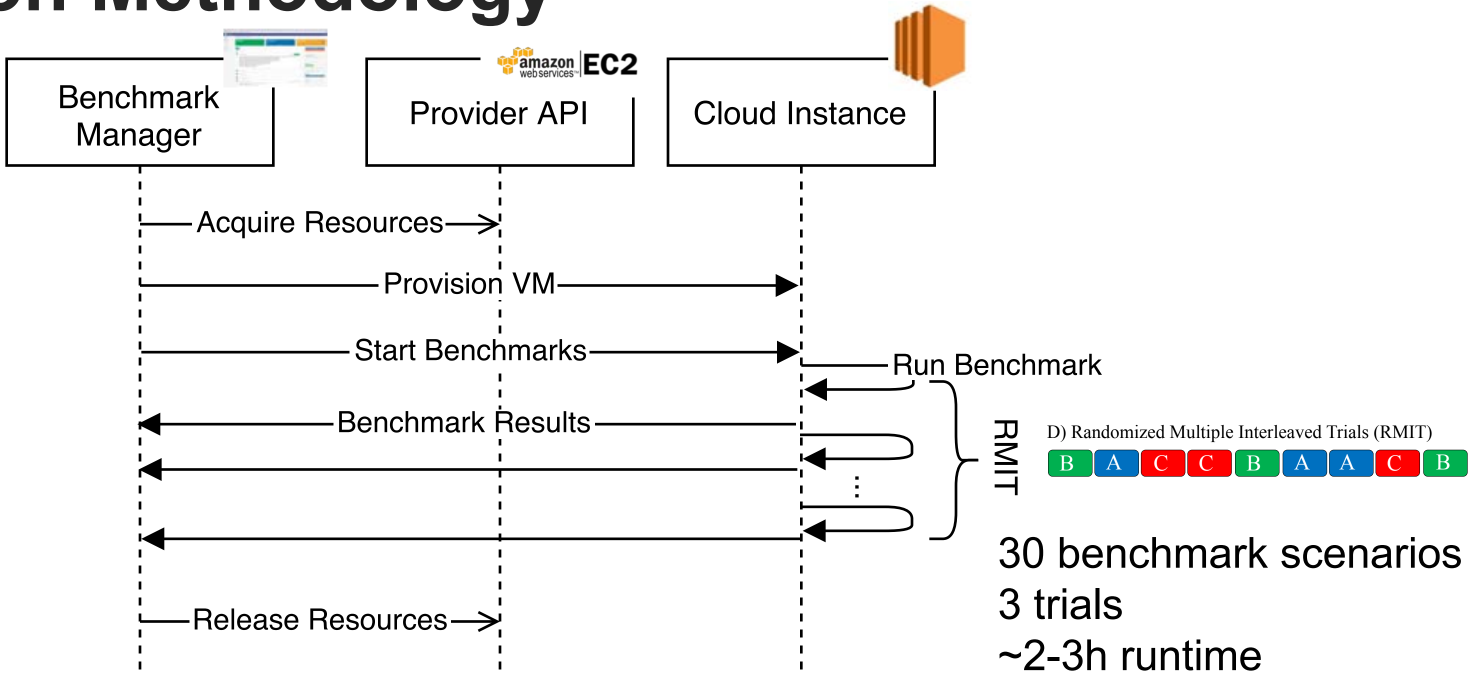
Application Benchmarks

Specific

Real-World

Resource-heterogeneous

Execution Methodology

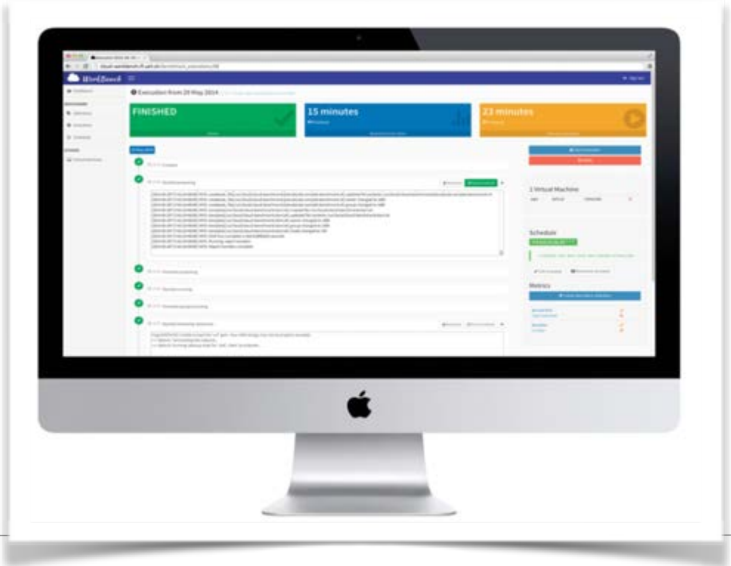


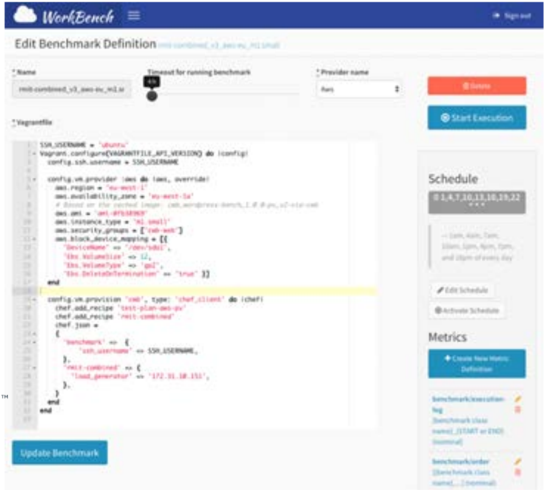
30 benchmark scenarios
3 trials
~2-3h runtime

Benchmark Manager

Cloud WorkBench (CWB)

Tool for scheduling cloud experiments



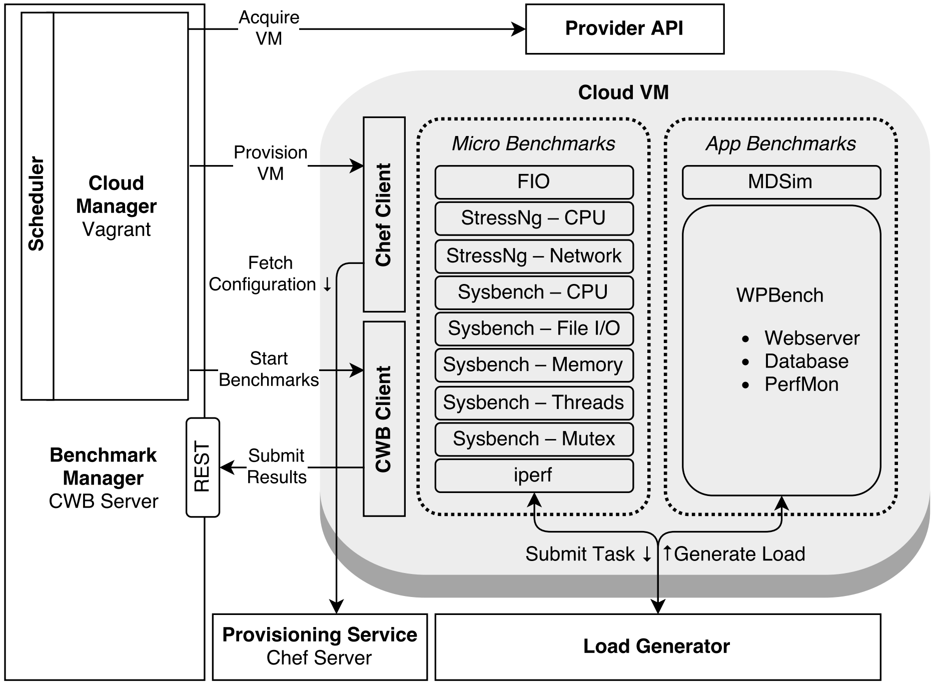


sealuzh/cloud-workbench

CloudCom 2014 "Cloud Work Bench – Infrastructure-as-Code Based Cloud Benchmarking"
Scheuner, Leitner, Cito, and Gall

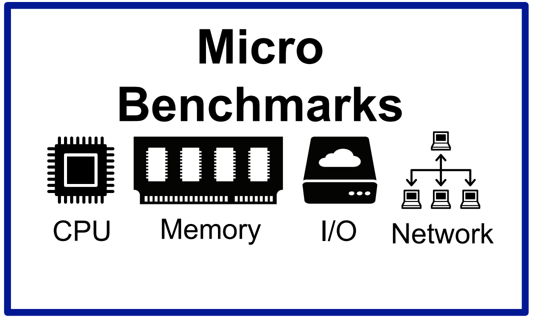
Demo@WWW 2015
Scheuner, Cito, Leitner, and Gall

Architecture Overview



Micro Benchmarks

Broad resource coverage and specific resource testing



CPU

- sysbench/cpu-single-thread
- sysbench/cpu-multi-thread
- stressng/cpu-callfunc
- stressng/cpu-double
- stressng/cpu-euler
- stressng/cpu-ftt
- stressng/cpu-fibonacci
- stressng/cpu-int64
- stressng/cpu-loop
- stressng/cpu-matrixprod

Memory

- sysbench/memory-4k-block-size
- sysbench/memory-1m-block-size

I/O

- [file I/O] sysbench/fileio-1m-seq-write
- [file I/O] sysbench/fileio-4k-rand-read
- [disk I/O] fio/4k-seq-write
- [disk I/O] fio/8k-rand-read

Network

- iperf/single-thread-bandwidth
- iperf/multi-thread-bandwidth
- stressng/network-epoll
- stressng/network-icmp
- stressng/network-sockfd
- stressng/network-udp

Software (OS)

- sysbench/mutex
- sysbench/thread-lock-1
- sysbench/thread-lock-128

Micro Benchmarks: Examples

Micro Benchmarks

CPU | Memory | I/O | Network

I/O — File I/O: 4k random read

1) Prepare

2) Run

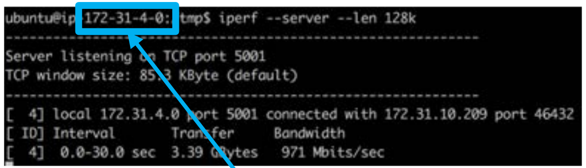
3) Extract Result

3.5793 MiB/sec

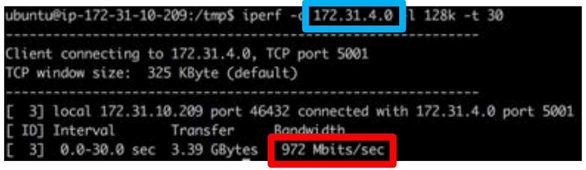
4) Cleanup

Network — Bandwidth

Server



Client



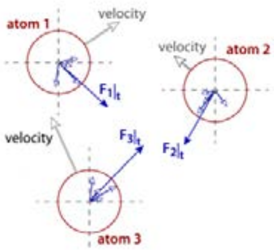
Result
972 Mbits/sec

Application Benchmarks

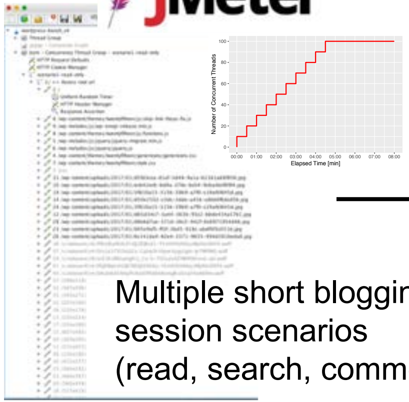
Application Benchmarks

Overall performance
(e.g., response time)

Molecular Dynamics
Simulation (MDSim)



APACHE JMeter™ WordPress Benchmark (WPBench)



Multiple short blogging
session scenarios
(read, search, comment)

Performance Data Set

amazon web services EC2

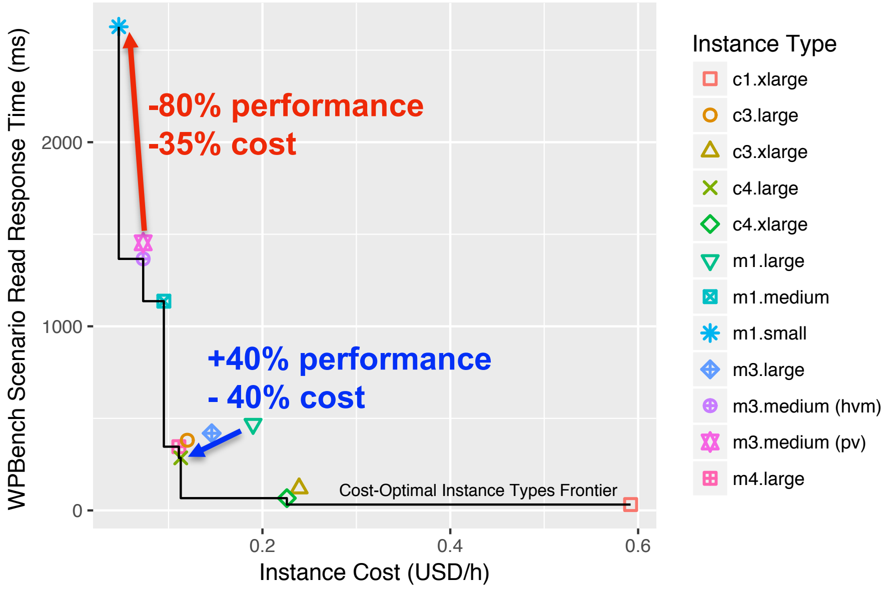
Instance Type	vCPU	ECU *	RAM [GiB]	Virtualization	Network Performance	
m1.small	1	1	1.7	PV	Low	eu + us
m1.medium	1	2	3.75	PV	Moderate	
m3.medium	1	3	3.75	PV /HVM	Moderate	
m1.large	2	4	7.5	PV	Moderate	eu + us
m3.large	2	6.5	7.5	HVM	Moderate	
m4.large	2	6.5	8.0	HVM	Moderate	eu
c3.large	2	7	3.75	HVM	Moderate	
c4.large	2	8	3.75	HVM	Moderate	
c3.xlarge	4	14	7.5	HVM	Moderate	
c4.xlarge	4	16	7.5	HVM	High	
c1.xlarge	8	20	7	PV	High	

* ECU := Elastic Compute Unit (i.e., Amazon's metric for CPU performance)

- >240 Virtual Machines (VMs) à 3 Iterations → ~750 VM hours
- >60'000 Measurements (258 per instance)

WPBench Response Time Cost Frontier

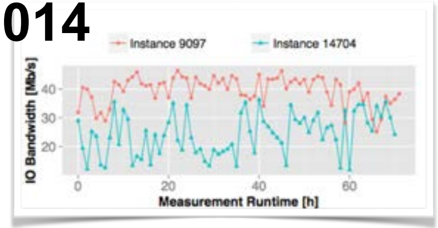
Cost/Performance is a trade-off but there exist unfavorable instance types



Intra-Cloud Network Bandwidth over Time

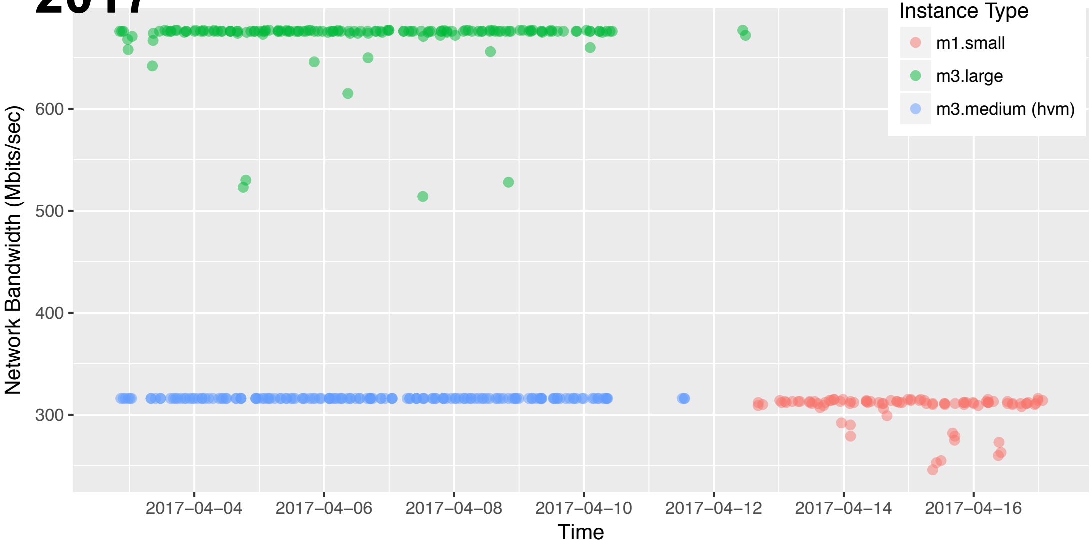
Almost perfect stability in comparison to previous results

2014



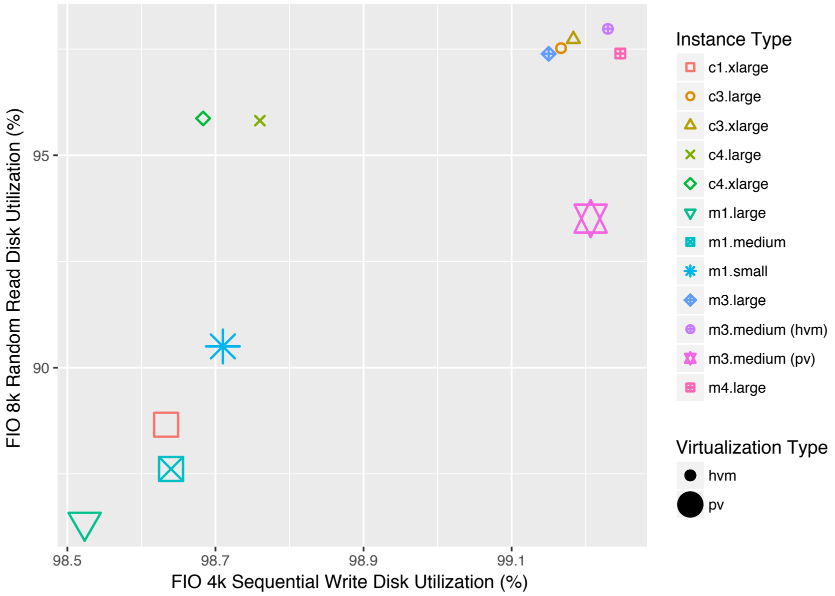
P. Leitner, J. Cito. Patterns in the Chaos - A Study of Performance Variation and Predictability in Public IaaS Clouds. TOIT 2016

2017



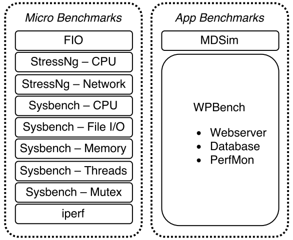
Disk Utilization during I/O Benchmark

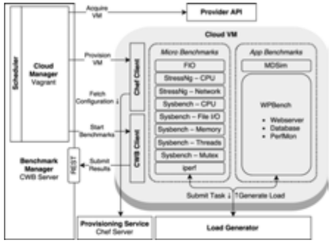
The newer virtualization type hvm is more I/O efficient than pv

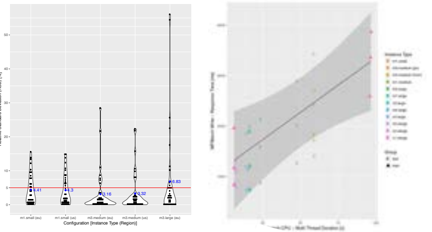


Future Work

Benchmark Design → Benchmark Execution → Data Pre-Processing → Data Analysis







Accepted

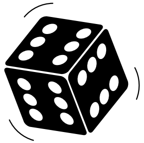
QUDOS@ICPE 2018 "A Cloud Benchmark Suite Combining Micro and Applications Benchmarks"
Scheuner and Leitner

Under Submission

"Estimating Cloud Application Performance Based on Micro Benchmark Profiling"
Scheuner and Leitner

Conclusions

Selecting an optimal instance type can save up to 40% costs while increasing up to 40% performance



Support trend towards more predictable performance (AWS EC2)

The newer virtualization type (hvm) improves I/O utilization rates up to 10% (vs pv)

scheuner@chalmers.se