Comparative Benchmark For Open Cloud Telekom

Open Telekom Cloud

2020

Index

I.Introduction

The IaaS segment of Cloud market is subject to a permanent race for efficiency where providers redouble their efforts to push the level of performance or the best match with their users' workloads. Virtual Machines being one of the main components of infrastructures, the constant pursuit of the best products is fast and we can observe that several offerings are released by providers each quarter.

Like the other hyperscalers, during 2020 Open Cloud Telekom designed new lines of VMs:

- C4: Dedicated CPU
- S3: General Purpose

These new products are equipped with custom Intel processors and are meant to perform better than the former series and the actual market competitors.

This document aims to benchmarks performance and price form Open Telekom Cloud's products against the Cloud Market leaders Amazon Web Services, Microsoft Azure and Google Cloud Platform.

An interactive visualization of this project is available at the URL below. It allows you to dynamically explore the data, display or export pieces of this study:
https://projector.cloud-mercato.com/projects/open-telekom-cloud-2020

II.Product specifications

For this benchmark, Virtual Machine from 2 to 16 CPU have been selected and for the sake of clarity, only 16-CPUs machine are present in the current document. The full inventory is available in the web version of the project.

The following exchange rates have been used:

- 1 USD = 0.8580 EUR
- 1 EUR = 1.1655 USD

1. Virtual Machines

Provider	Name	Series	CPU	RAM	Hourly
AWS	m5.4xlarge	General purpose 5	16	65536	$0.92
AWS	r5.4xlarge	Memory-optimized 5	16	131072	$1.22
Google	n1-highmem-16	N1 High Memory	16	106496	$1.04
Google	e2-standard-16	E2 Standard	16	65536	$0.69
Azure	Standard_D16s_v3	Standard Ds v3	16	65536	$0.96
Azure	Standard_E16s_v3	Standard Es v3	16	131072	$1.28
OTC	c3.4xlarge.4	Elastic compute 3	16	65536	$1.19
OTC	c4.4xlarge.4	Elastic compute 4	16	65536	$1.18
OTC	s2.4xlarge.4	General-purpose v2	16	65536	$0.79
OTC	s3.4xlarge.4	General-purpose v3	16	65536	$0.94

2. Block Storage

To apply a fair benchmark methodology, we size the volumes from Block Storage with a rule of thumb: **50GB per CPU**. This rule mimics a traditional scaling up where the more you have computing power (CPU), the more you may require storage. It allows to go through the different performance tiers applied to some storage classes.

Provider	Storage	Monthly Price
Amazon Web Services	General Purpose SSD	$0.11
Google Cloud	Persistent Disk SSD	$0.20
Microsoft Azure	Premium LRS	$0.20
Open Cloud Telekom	Ultra-High I/O	$0.10

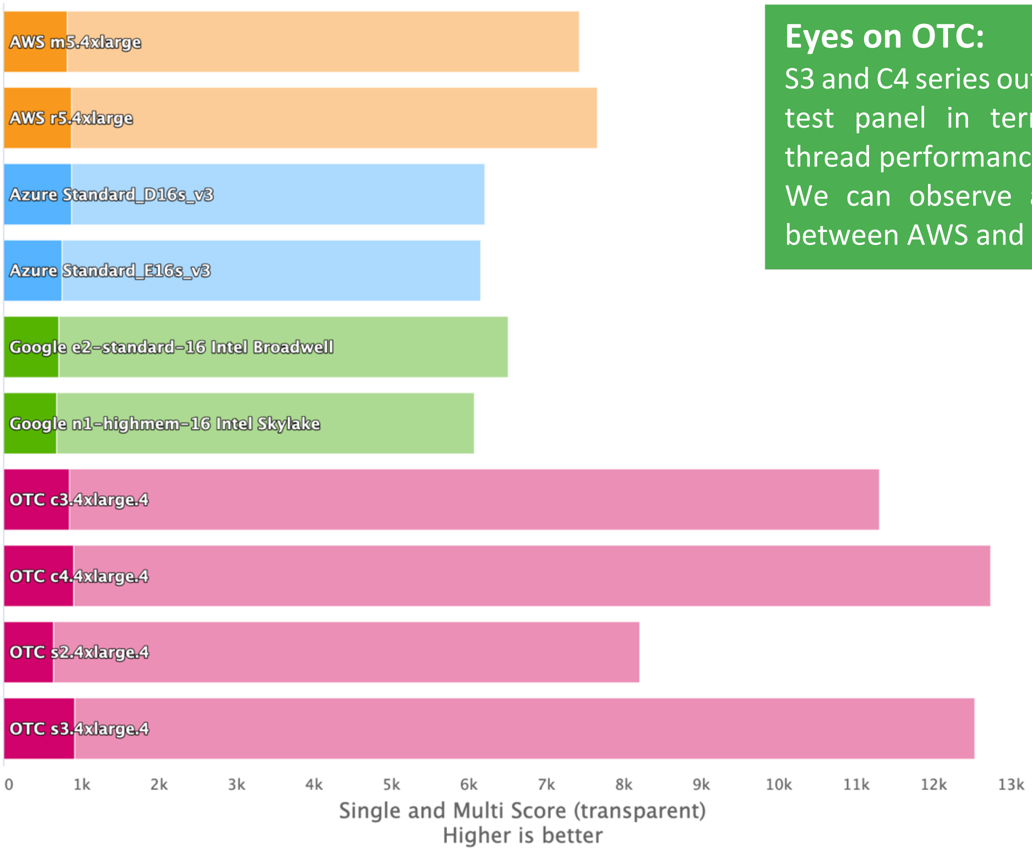
III.Performance testing

1. CPU Performance

Virtual Machines' performance is mainly qualified by computing power delivered by CPU model. This value derives from a lot of other characteristics bound to virtualization and CPU specifications such as frequency, built-in instructions and more. Facing a real-life workload, the simple MIPS (Millions of Instructions Per Seconds) or FLOPS (FLoating-point Operations Per Second) declared by CPU vendors aren't enough to declare if a machine performs better than another.

To collect a synthetic performance value, we use the benchmark suite Geekbench 5. This software runs workloads across Integer, Floating Point and Cryptography domains. The set of tests includes various kinds such as Compression, Machine Learning or Compute Vision and each test is attempted with single and multi-thread modes.

From the Geekbench 5 output, we'll focus on Single and Multi-Score which are respectively summary of a single CPU performance and the whole VM performance.

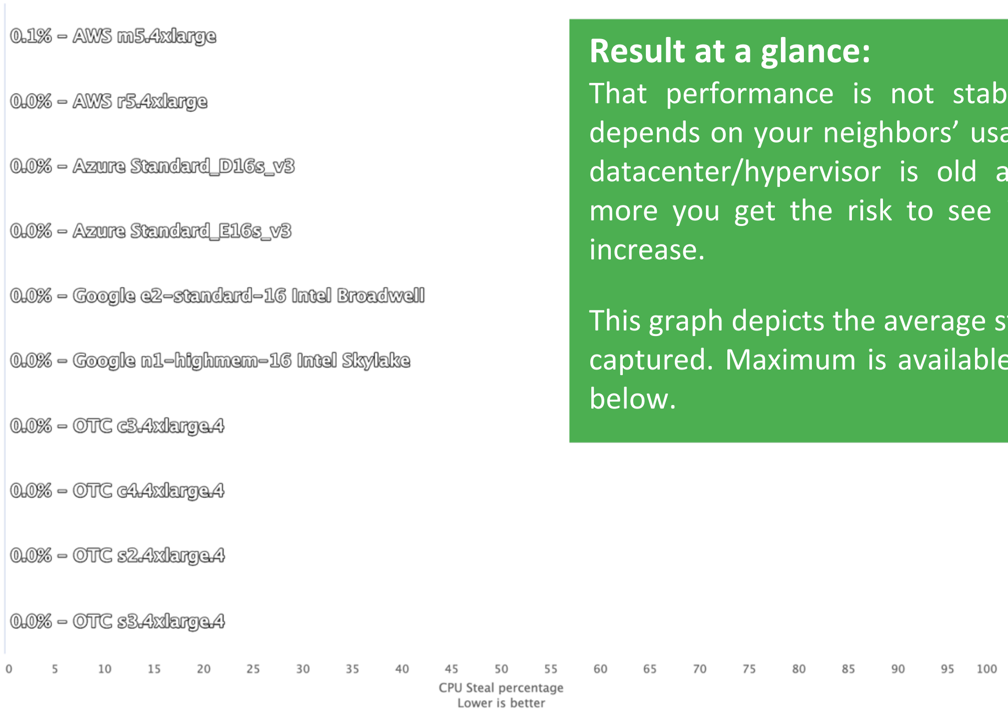


Eyes on OTC:
S3 and C4 series outperform the test panel in term of multi-thread performance.
We can observe a factor 1:2 between AWS and OTC

2. CPU Steal

Unlike VPS providers, Clouds are known to deliver a high-level of service by avoiding practice of CPU overcommitting. But by nature, VMs mandatorily share resources with other tenants or at least with the hypervisor. CPU sharing can be measured by collecting a Linux kernel counter called CPU Steal.

Expressed here in percentage, this number represents the amount of time that a task was not able to be done by CPU because of someone else usage. While we watch CPU steal counter, we load CPUs with Prime95. This software stresses all CPU up to 100% and allows us capture how much tick has been stolen during the load.



Result at a glance:
That performance is not stable and really depends on your neighbors' usage. More the datacenter/hypervisor is old and occupied, more you get the risk to see its Steal time increase.

This graph depicts the average stolen time we captured. Maximum is available on the table below.

		max	mean	std
Provider	Flavor			
Amazon Web Services	m5.4xlarge	0,1	0,1	0
	r5.4xlarge	0,1	0	0
Google Cloud	e2-standard-16	0,1	0	0
Microsoft Azure	Standard_D16s_v3	0	0	0
	Standard_E16s_v3	0	0	0
Open Telekom Cloud	c3.4xlarge.4	0	0	0
	c4.4xlarge.4	0	0	0
	s2.4xlarge.4	0	0	0
	s3.4xlarge.4	0	0	0

Result at a glance:
Cloud providers generally don't overcommit their hypervisor. A CPU stolen higher than 10% is extremely rare.

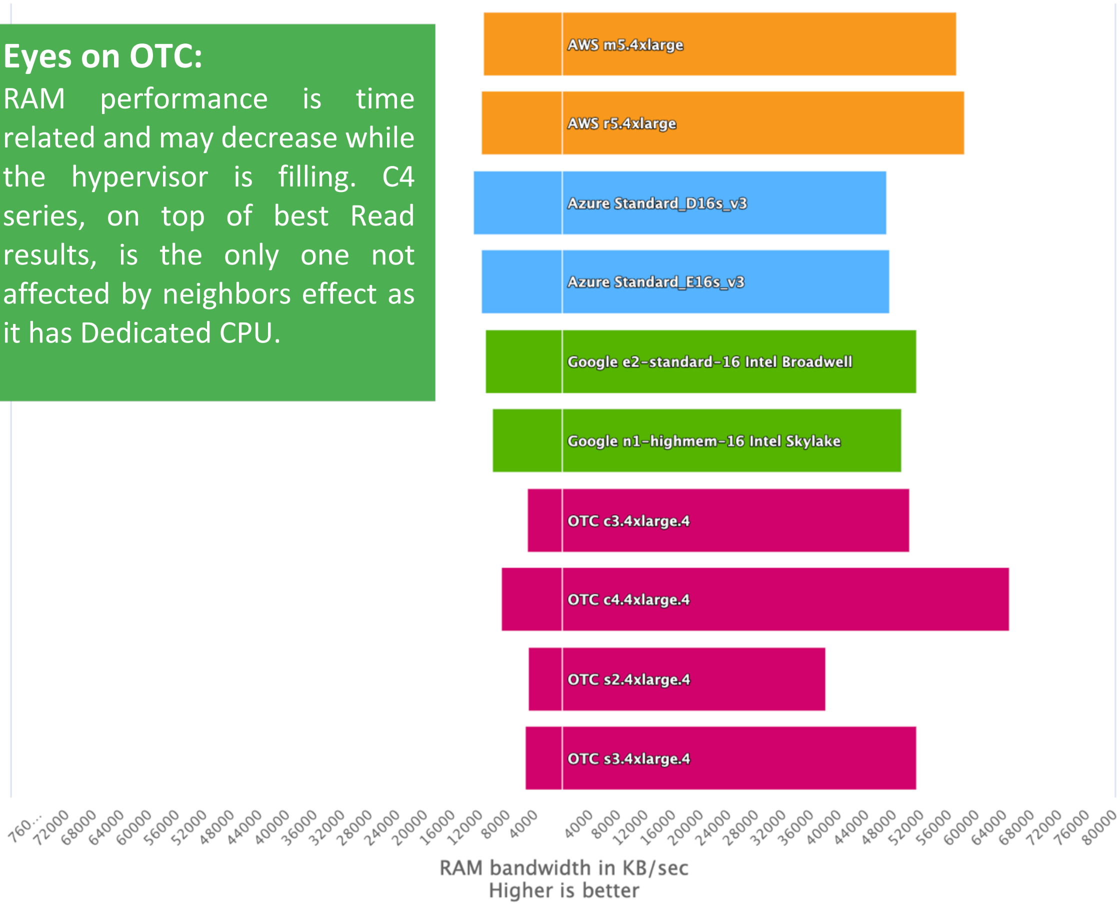
3. RAM Performance

The volatile memory is the fastest one present on a system, this is the main one used by software, cache, buffer and more. RAM access testing has several interests:

- It gives available performance for memory-intensive workloads such as in-memory databases.
- The RAM buses being shared among CPU cores, like CPU steal, it's an indicator of noisy neighbors.

To evaluate RAM performance, we test the bandwidth between CPU and memory with the opensource software Sysbench. Its test scenarios are based on random access to memory with 1KB blocks.

Eyes on OTC:
RAM performance is time related and may decrease while the hypervisor is filling. C4 series, on top of best Read results, is the only one not affected by neighbors effect as it has Dedicated CPU.



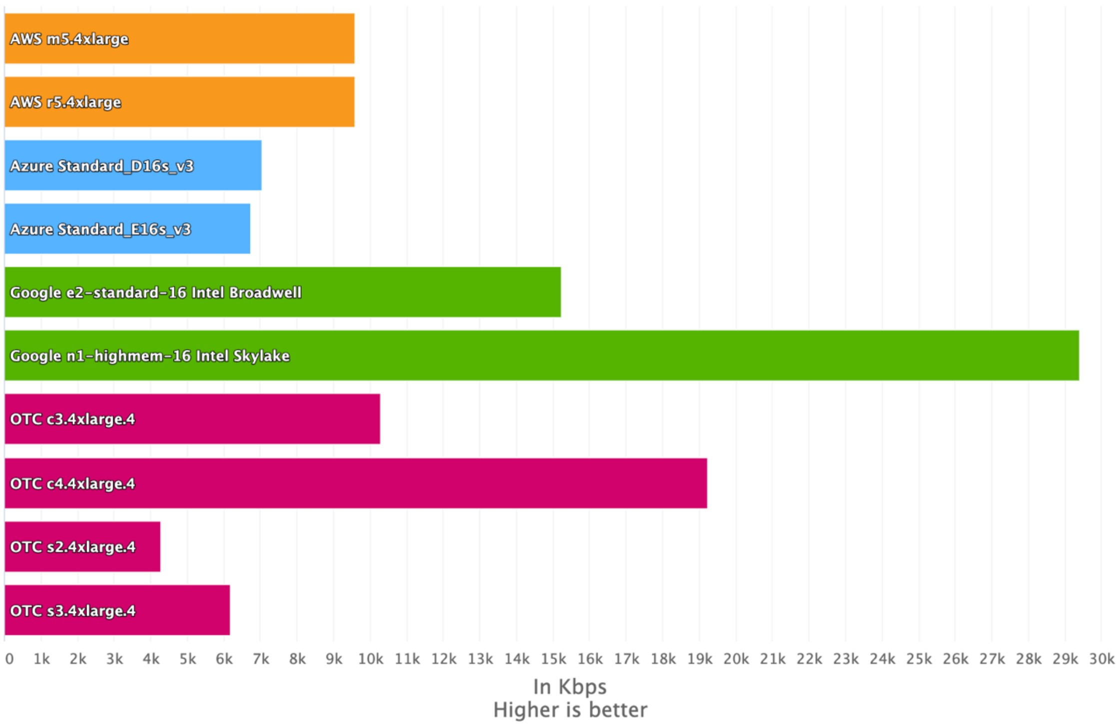
Read and write access modes have been used for this testing. The Read mode depicts the capacity to intensively retrieve cached data such as in-memory database. By nature, Write mode is slower as it suffers from latency occurred by data storing.

4. Network Performance

Cloud Providers generally have a high-performance internal network throttled in consumer usage to guarantee a certain level of service for all tenants. The maximum performance is completely virtual and is defined by vendor in the virtual machine's network specification.

Despite the theoretical maximum bandwidth is announced by providers, only performance testing can assert what they claim. We used iPerf3 to evaluate the maximum available bandwidth set by providers with the following methodology:

- 2 identical VMs
- Same region and availability zone
- A number of thread equal to CPU to generate the max. throughput



Eyes on Open Telekom Cloud:
C4 with 20Gbps offers the best performance of this benchmark. Network performance are progressive relative to VM size, while OTC may not be the best fit in smallest sizes, 16 CPUs and more products offer the highest bandwidth.

5. Block Storage Performance

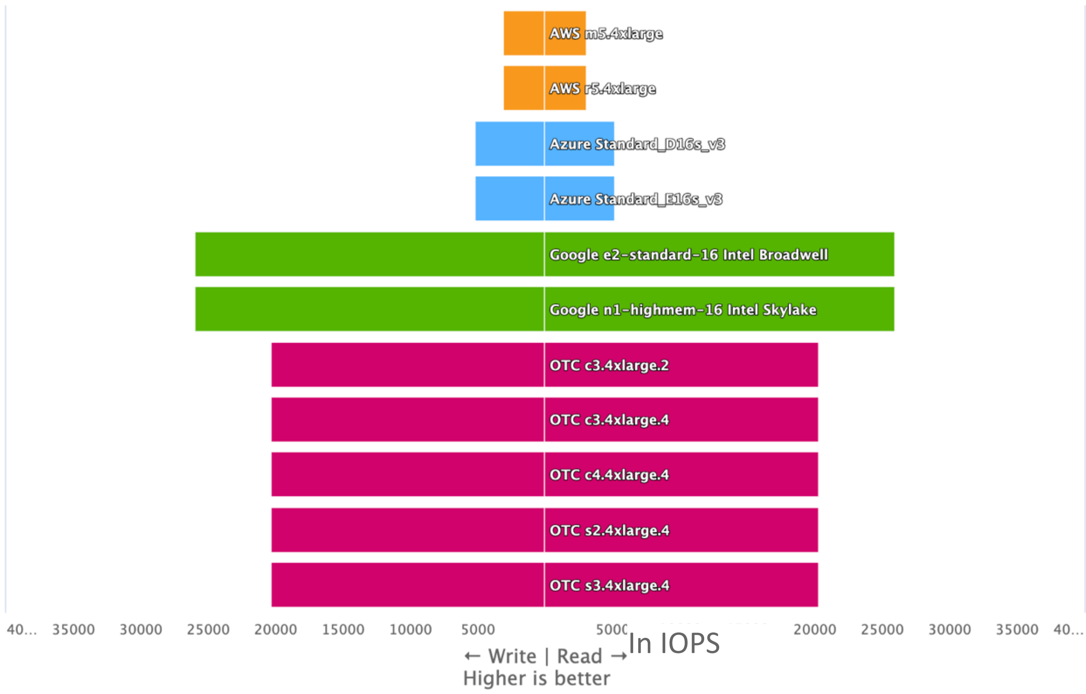
By its design, block storage is the natural Cloud answer to persistent storage. This kind of memory offers the inherent flexibility bound to Cloud Computing while providing a high level of availability and durability. On top of that it brings managed features such as snapshots/backups.

Block storage performance is mainly qualified by its maximum Input/Ouput Per Seconds (IOPS). This value helps to determine the maximum number of transactions that the virtual device is able to provide and at this game, no storage system is equal to another. Like for network bandwidth, storage is a shared resource, where vendors prefer to throttle usage rather than let users with erratic performance.

Each provider has defined its own rules about IOPS and briefly here's the performance factors:

- **Storage class:** Define the base rules to establish performance
- **Volume size:** The maximum performance can be driven by the size of volume.
- **Total volume size:** At Azure, the performance depends of size tiers calculated on the total ownership. The performance being shared across all volumes.

We use Flexible I/O Tester (FIO) to evaluate IOPS with the following settings:



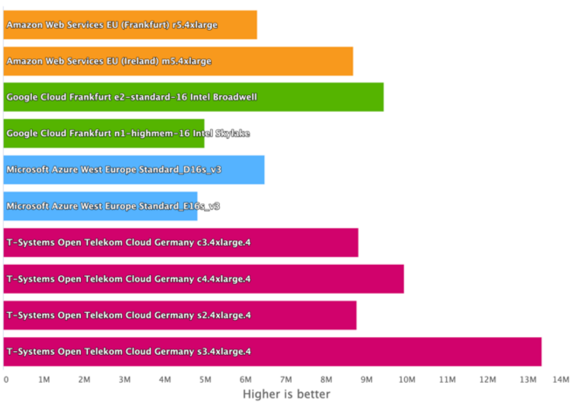
Eyes on results:
This benchmark highlights 2 groups of classes: Low and high IOPS storage. The first group is throttled around 3,000. IOPS, the second caps at 26,000 IOPS.

IV.Price/Performance value

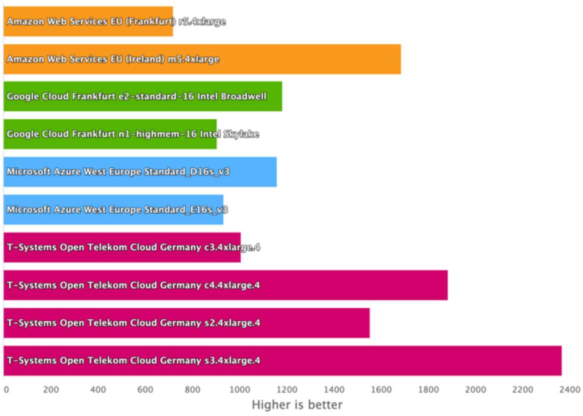
Price/Performance is a ratio of raw performance to price for a Cloud Provider's service or feature subset. Thus, price- performance offers a universal metric for comparing service value.

1. Compute

As different billing options exist across providers, we compare the hourly and yearly consumptions. The performance value used for the charts below is the Geekbench 5 multi-score.



Hourly pricing

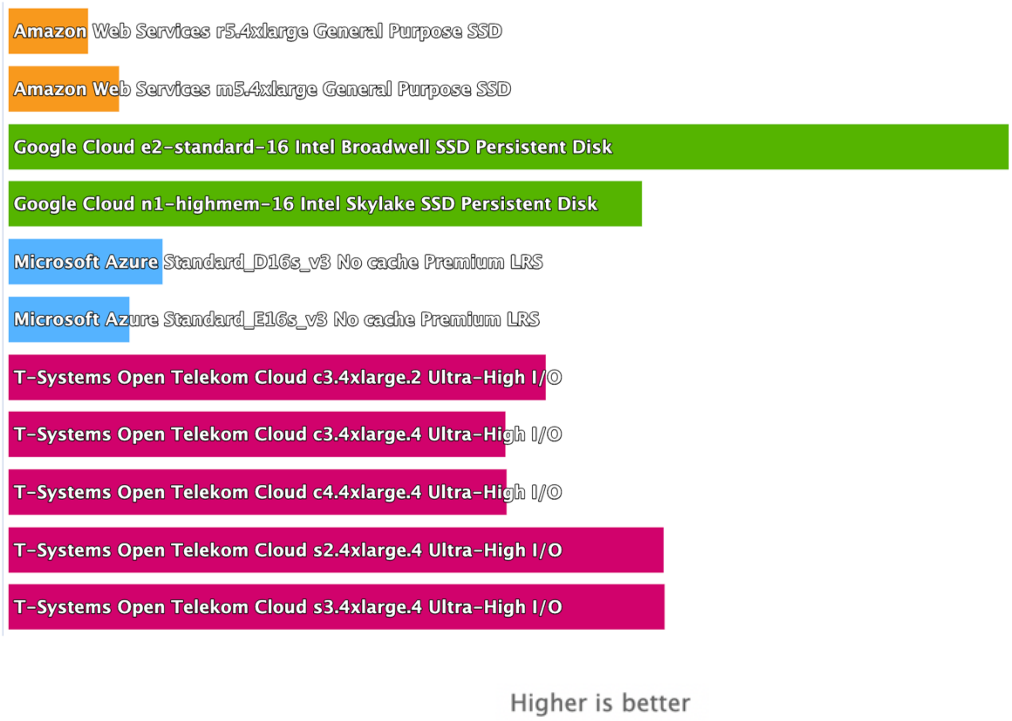


1-year pricing

Eyes on billing options:
The long-term option accentuates the price/performance ratio and the efficiency of OTC products. While T-Systems' hourly pricing is higher, the yearly pricing offers a better ratio.

2. Block Storage

From the IOPS testing, we create an apple-to-apple comparison gathering the VM and volume hourly prices. Read and write access performance are also taken in account to compute the price/performance value.



Eyes on results:
Block Storage's efficiency follows the Compute's one, except for the High-IOPS classes at Google, Oracle and T-Systems. Like for raw performance analysis, 2 groups are drawn by this study: Low and high IOPS.

V.Test methodology

1. CPU Performance

Geekbench 5 provides a command line tool with a simple usage:

geekbench5 --no-upload --export-json

This command will launch close to 60 micro-benchmarks and summarize results in JSON format. Geekbench outputs each performance with its score but we focus on the aggregation: Single and Multi-Score

2. CPU Steal

We capture CPU steal by reading Linux Kernel's counters and stressing CPU at 100% with Prime95. The CPU steal is also available via top, htop or atop command line tools.

3. RAM Performance

Sysbench is an opensource benchmark tool available on GitHub. The memory testing is ran by the following commands:

- sysbench --threads=<cpu_number> --time=30 memory --memory-oper=read run
- sysbench --threads=<cpu_number> --time=30 memory --memory-oper=write run

This C program stresses RAM by reading or writing small block of 1KB. The outputted values could not represent the maximum bandwidth but the capacity to CPU and RAM to communicate.

4. Network Performance

Iperf 3 requires to be launched as client and as server, here's the command line used:

- iperf3 --client <ip> --interval 30 --parallel <cpu_number> --time 30 --format M –json
- iperf3 --server --version4 --interval 30

These commands will fill the bandwidth with the goal is to capture the maximum TCP throughput between 2 hosts.

5. Block Storage Performance

Storage results were obtained using FIO (Flexible I/O tester) using 4KB blocks and threads corresponding to vCPU count.

- fio --numjobs=<cpu_number> --bs=4k --rw=randread --ioengine=libaio --iodepth=32 --direct=1 --invalidate=1 --end_fsync=1 --time_based --runtime=30 --timeout=30 --filename=/dev/sdX --group_reporting --output-format=json --name=fio
- fio --numjobs=<cpu_number> --bs=4k --rw=randwrite --ioengine=libaio --iodepth=32 --direct=1 --invalidate=1 --end_fsync=1 --time_based --runtime=30 --timeout=30 --filename=/dev/sdX --group_reporting --output-format=json --name=fio

VI.Study notes

Test Design Considerations

Testing was conducted on specific VM types for each provider. Provider VM configurations may yield different results based on underlying infrastructure, virtualization technology, settings (e.g. shared resources), and other technology factors. Furthermore, issues such as user contention or physical hardware malfunctions can also cause suboptimal performance. Cloud Mercato therefore provisioned multiple VMs with the same configuration to better sample the underlying hardware and enabling technology, as well as to improve testing accuracy and limit the effects of underlying environmental variables.

The VMs selected for this engagement were generally-available specified offerings from the various providers. While better performance can often be attained from providers when additional features or support services are purchased, the selected VMs used in Cloud Mercato's testing do not leverage such value-added services. This helps provide data and test results that are indicative of real-world customer choices and ensures the most direct comparisons possible.

Error Minimizing Considerations

Duplicate VMs were deployed during testing to minimize sources of error prevalent in a Cloud hosting environment.

VII.About Cloud Mercato

Cloud Mercato is a neutral research and consulting firm dedicated to the study of the Cloud Market. Our goal is to bring transparency to the Cloud Market by the study and analysis of the different products and services.

We proactively benchmark the industry and share our analysis through our Cloud Transparency Platform.