I1 – High IOPS series

Benchmark

Index

I.   Introduction

Emergence of software for Machine Learning (ML) employed by corporations such as Google, Intel and Microsoft, independent research firms and educational institutions are frequently plagued by bottlenecks in their computing infrastructure. Complimenting ML and AI, the field of Big Data is rapidly expanding with equivalent demands at the risk of stagnation.

Regardless of client objective, enormous datasets must be parsed, analyzed and validated— in the fastest moving market to date, speed is certainly of the essence. Furthermore, database servers, a long time and ubiquitous element in the computing industry display growing reliance on these same technologies if they are to maintain pace with applications to which they are integral.

Early on, computational limitations were addressed via cooption of GPUs for parallel stream calculations; Hard Disk Drives, then Serial Solid-State Drives (SSDs) combined with advanced RAID configurations were implemented at high cost to meet exponentially increasing IO requirements. In the past few years, however, great strides have materialized within the storage market in the form of Non-Volatile Memory Express (NVMe) storage, a new form of solid-state device with enormous bandwidth and potential. Movement from Serial SSDs to Parallel NVMe storage now compliments GPUs, serving as great tools for their employing entities. Still, adoption of new technology is hindered due to real or perceived cost prohibition by would be users.

In this short evaluation, OVHcloud's new High IO cloud solutions are analyzed and scrutinized closely along with competing offerings from the Cloud Sphere's biggest players to see if they truly conquer current impediments of affordability while providing the necessary performance for continued industry growth.

II.   Virtual machine specifications and selection

VMs for this engagement focused on High Performance Storage VMs with local NVMe volumes. All VMs were deployed with a current release of Ubuntu 18.04 LTS from each respective provider. OVHcloud advertises the new High IO VMs as having PCI passthrough for blazing performance.

| Provider	Flavor	CPU	RAM (GB)	Storage (TB)	Device	Hourly price	Monthly price
AWS	i3.2xlarge	8	64	1.9	1	0.69	503.07
Azure	Standard_L8s_v2	8	64	1.9	1	0.74	540.20
Oracle	VM.DenseIO1.4	8	64	3.2	1	0.51	372.30
OVHcloud	I1-45	8	45	1.8	1	0.47	343.10
OVHcloud	I1-90	16	90	1.8	2	0.94	686.20
T-Systems	h2.3xlarge	16	128	3.2	1	2.42	1766.60

Monthly price
Lower is better
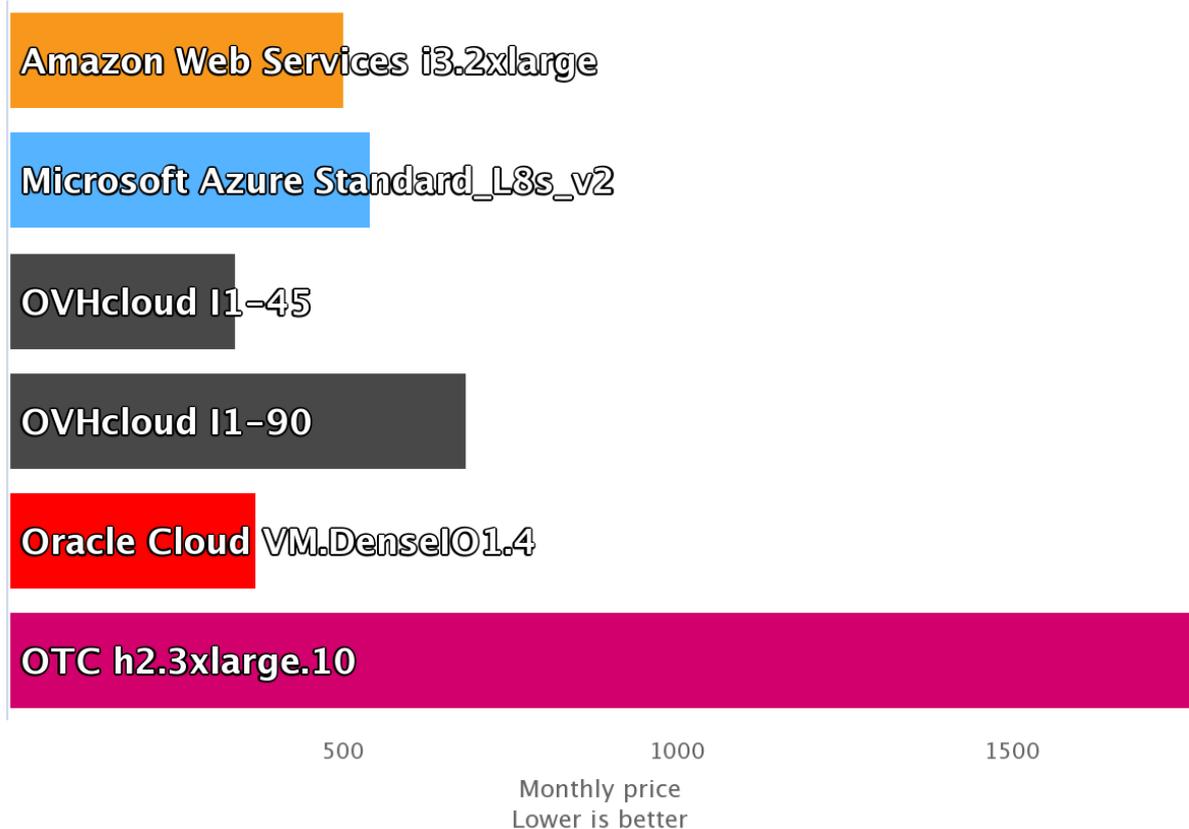
III.   Performance testing

Storage testing methodology ensured all VMs performed a minimum of 100 iterations for read and write operations. Flexible I/O Tester (FIO) was employed for device benchmarks given the following parameters:

- **Block size**: 4K
- **Queue depth**: 32
- **Threads**: 1 per vCPU
- **Iteration time**: 60 seconds
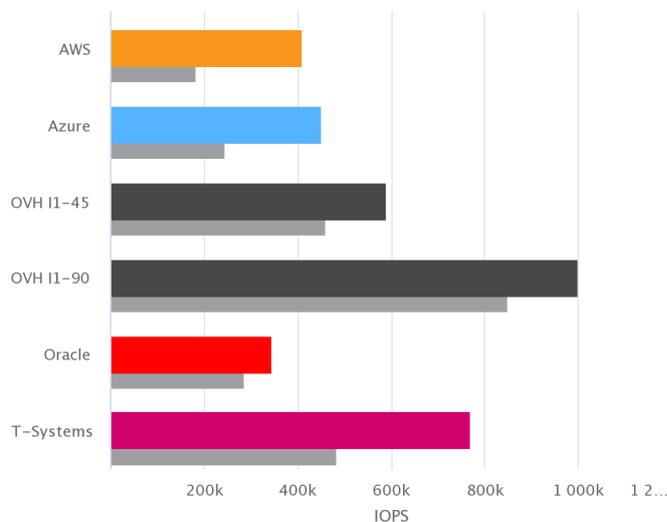
Eyes on OVHcloud NVMes: Read Performance

➢ Among VMS evaluated, OVHcloud's I1-90 configured with PCI passthrough and MDADM software striping provided strikingly high performance. The I1-45 and its single 1.8TB drive, delivered an average, just shy, of 600,000 read OIPs. The I1-90 striped array, indeed provided blazing speeds, the quickest observed by Cloud Mercato to date -- exceeding 1,000,000 read IOPs.

OVHcloud circumvents some of the above by eliminating a virtualization layer—supplying machines with PCI passthrough allowing for improved write operations, surpassing 450,000 IOPs per drive. All machines tested have NVMe conferred advantages, but the hardware is only as good as the method of implementation, perhaps an OVHcloud advantage in this case.

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. This value is calculated from the true mean of the benchmark divided by the (hourly price * 730 or the monthly price as given without discounts).
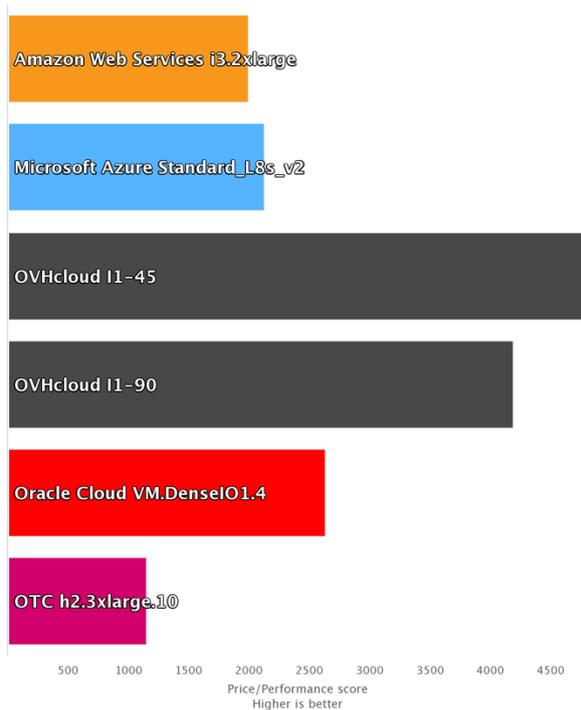
- Some machines in this contest were designed for High Performance computing, not just IO proficiency, as such those machines have much higher prices as depicted earlier. When it comes to "bang for the buck" of the test group, OVHcloud's I1-45, with sub $340/mo cost, clearly shows the effect of price relative to performance.
- In appearance, the I1-45 should be capable of handling heavy database operations, email servers, web applications or multiple roles simultaneously. This makes for a fit machine, suitable for most small businesses or essential component within an enterprise's infrastructure.
- The I1-90 is not far behind despite the higher cost. Equipped with double the hardware and delivering performance to scale. These points are particularly salient

V. Conclusion

Based on the study results, OVHcloud has indeed supplied a solution to the storage speed bottleneck—without sacrificing advantages of cloud scalability and flexibility nor incurring the penalties of purely virtualized hardware. The I1-45 and I1-90 evaluated, clearly harbor required storage capabilities for Machine Learning applications and those for Big Data enterprises, in addition to substantial platform improvements for traditional database operations.

Compared to competing offerings within the study, OVHcloud VM storage value displayed a combined average improvement of 66% and 78% for read and write operations respectively. Combined with a fair, competitive pricing model, these machines could revolutionize business for entities and individuals looking to move from other clouds or upgrading to this profoundly new storage technology without the massive overhead required only a few years ago.

VI.   Method and study notes

Synthetic Testing: Storage

Storage results were obtained using FIO (Flexible I/O tester) using 4KB blocks and threads corresponding to vCPU count. Several hundred 60-second random iterations were conducted to compensate for the high variability often seen when stressing storage volumes. Results were gathered and represented in IOPs (input/output operations per second).

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. The most notable challenge is the Noisy Neighbor Effect. Testing duplicate VMs mitigates most non-specific errors that could be attributed to a singular parent instance or storage volume. By minimizing possible sources of error, more accurate and precise performance samples can be collected during testing.

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.