CLOUD IAAS PERFORMANCE

AUGUST 2016

CLOUD SPECTATOR

Performance analysis of three public cloud providers: Amazon AWS, Microsoft Azure, and Schuberg Philis

TABLE OF CONTENTS

INTRODUCTION

In an effort to simulate an end-user experience regarding performance of virtual machines across various cloud providers, Cloud Spectator ran its iterative benchmark suite on virtual machines (VM) of each of the following providers: Amazon Web Services, Microsoft Azure and Schuberg Philis.

Cloud Spectator measured performance on 1vCPU, 2vCPU, and 4vCPU VMs for this study. For each size, two total VMs were provisioned, each one tested for a 4-hour period, every three days.

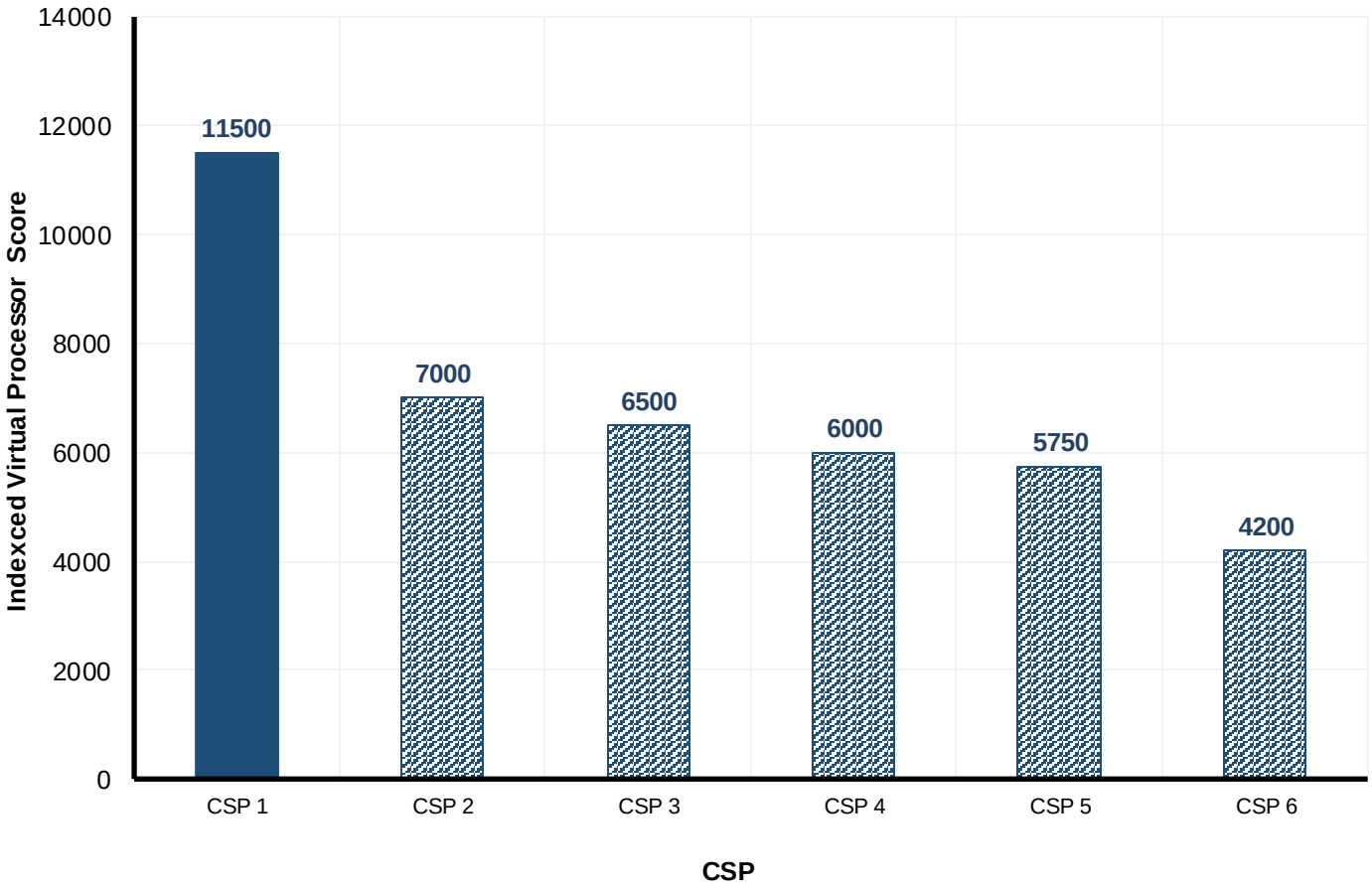
This study not only examined the performance of each vendor, but also tracked performance variability for each of the 1vCPU, 2vCPU, and 4vCPU machines over a period of 6-months. The methodology allowed Cloud Spectator to capture performance variability both over time on the same VM as well as across different VMs on multiple physical hosts. The providers, were observed with widely varying processor and memory bandwidth performance stability compared against each other. Some providers, such as Amazon, exhibited controlled periods of burst followed by throttled performance on network storage depending on the size of the storage volume. Both Amazon and Schuberg Philis exhibited variable performance across the storage volumes throughout the study. Internal network was tightly throttled on Amazon and Azure, compared to a wider range of values on Schuberg Philis, but are set at relatively lower throughput limits.

WHY PERFORMANCE MATTERS

Performance and pricing are both key considerations in the public cloud industry, together having a substantial impact on a company's annual operating costs. Cloud users may need fewer resources on better performing services, which can lower costs. Since many users only consider price and not price-performance, these users may be paying more because they require additional resources to achieve a desired level of performance. While some providers try to differentiate their offerings by cutting prices, others try to differentiate by focusing on improved performance and user experience.

Differences in performance outputs of VMs across IaaS providers can greatly impact quality of service as well as annual operating costs. The graph on the right illustrates an example of the average processor performance from a sample of six Cloud Service Providers (CSPs) as studied by Cloud Spectator. CSP 1 has a processor performance three times as high as CSP 6 (names removed), which gives CSP 1 a notable advantage in many processor-intensive workloads. CSPs 2-5 exhibit a closer resemblance in processor performance, but do not offer nearly as much processing power as CSP 1 does.

Chart 1.1A: Virtual Processor Performance Across the IaaS Industry



EXECUTIVE SUMMARY

Schuberg Philis commissioned Cloud Spectator to gauge the performance of VMs on three different cloud providers': Amazon Web Services, Microsoft Azure, and Schuberg Philis. Raw VM performance was examined to evaluate the component attributes of processor, memory, disk, and internal network on each provider's VMs. The purpose of the study was to understand, from an end-user perspective, performance among cloud providers with similarly sized VMs.

FINDINGS

Table 1.1: Key Findings Observed in This Report

vCPU Performance Findings
Tested with Geekbench 3

Schuberg Philis VMs exhibited the highest maximum scores for processor performance on 1vCPU and 4vCPU VMs in the study.

Amazon VMs exhibited overall performance stability for the processor tests with coefficients of variation (CV) at 9% or lower.

Memory Performance Findings
Tested with Geekbench 3

Schuberg Philis VMs achieved the highest median score for memory bandwidth on 2vCPU and 4vCPU VMs.

Amazon VMs exhibited the highest performance stability for the memory bandwidth tests overall, with an average CV of 8%.

Block Storage Performance Findings
Tested with fio

Schuberg Philis VMs offered the highest median performance for VMs examined in the study.

Azure VMs exhibited the highest overall performance stability for the storage tests.

Internal Network Performance Findings
Tested with iperf

Schuberg Philis VMs achieved the highest internal network throughput, although the throughput fluctuates significantly - between 2-9Gbits.

Amazon exhibited highly stable internal network throughput with an average coefficient of variation (CV) of 8%.

METHODOLOGY

PROCESS

1. Test cycles were run for each VM on each provider for a total of 4 hours of testing per VM size, every 3-days, from February 2016 to August 2016. After each cycle, VMs were terminated before beginning another cycle of tests on newly provisioned machines.
2. Each VM provisioned with a Linux Ubuntu 14.04 operating system by default, available from all providers. For AWS, the HVM image was used.
3. Before each test period, and after provisioning the VMs, system updates and upgrades were conducted via apt-get.
4. The following dependencies were installed for testing:
    a. **Git**. Git was used to clone the test repository on the VM.
    b. **MySQL**. For automation purposes, mysql-server was installed to automate data uploads.
    c. **Pip**. Used to download the appropriate libraries for Python in order to run the testing. SQL Alchemy was downloaded to interact with MySQL and upload data.
    d. **Libmysqlclient-dev**. MySQL database's development files, which are necessary for the SQL Alchemy and MySQL interaction.
5. Each test cycled through in the following sequence: Geekbench 3 (process & memory), fio sequential operations, fio random operations, Iperf internal network throughput (for more information on testing, see Tests Used).
    a. For fio testing (to measure disk IOPS), sequential operations ran first. Files from the sequential tests were deleted, and fio recreated files before running random operations. Once random operations completed, the files were also deleted. Thus, before each disk IOPS test, the files associated with the tests were deleted and recreated.
6. Internal network testing was conducted by creating a clone of the existing VM within the same region/availability zone in a client and server architecture, connected through a private network. Iperf was installed on both the client and server. The server used "Iperf –s" to listen on port 5001 (default port) for a TCP connection made by the client. At the conclusion of the 4-hour cycle of testing, the cloned server was terminated alongside the client VM. Iperf ran bi-directionally for 60 seconds each cycle.
7. At the end of each test iteration, results were uploaded into Cloud Spectator's database through use of SQLAlchemy.

TESTS USED

PROCESSOR AND MEMORY

Geekbench 3

Geekbench 3, a licensable product created by Primate Labs, is a cross-platform processor benchmark that can measure single-core and multi-core performance by simulating real-world workloads. The Geekbench 3 test suite is comprised of 27 individual tasks/workloads: 13 integer workloads, 10 floating point workloads, and 4 memory-bandwidth tasks. While processor and memory bandwidth are both performance factors that contribute to the final score provided by Geekbench 3, the test suite weighs processing performance much more heavily than memory bandwidth. Also, memory bandwidth is not necessarily affected by the amount of memory available for the VM, so VMs with larger amounts of memory may not exhibit larger bandwidth. For more information on Geekbench 3 and to see its individual workloads, please see http://www.primatelabs.com/geekbench/.

DISK IOPS

Fio

Fio is an open source I/O generator that spawns a number of threads and processes to conduct a particular type of I/O action specified. For the purpose of this study, fio was used to measure disk IOPS by tracking direct I/O to the VM's network storage. 5 x 200 MB files were created for sequential operations testing, and 5 x 200 MB files were created for random operations testing. All operations were 100% read and 100% write, then aggregated for an overall score. Each test iteration used a 4 KB block size. Each test iteration lasted 60 seconds.

INTERNAL NETWORK

Iperf

Iperf is an open source tool used to measure TCP or UDP network bandwidth performance. Cloud Spectator used Iperf to measure the network throughput between VMs residing in the same region/availability zone. A screen session was created for Iperf as a server machine on each appropriate VM (see Process 6). Each Iperf test iteration lasted 60 seconds, and data was transferred in one direction, from the test VM to the Iperf server VM.

VM CONFIGURATIONS

VM configurations were matched to standardize by virtual processors. 1vCPUs, 2vCPUs, and 4vCPUs machines from each provider were used in the study that most closely matched in size. While some providers may offer the option for local storage, none was used and all disk testing was conducted on persistent SAN storage. The storage columns in Figures 2A through C are reflective of the SAN storage provisioned for each VM.

Table 2A: 1vCPU Virtual Machine Configurations

Provider	Instance Name	vCPU	Memory (in GB)	Block Storage (in GB)
Amazon	m3.medium	1	3.75	100 (EBS SSD)
Schuberg Philis	x-tiny	1	2	100 (Gold)
Schuberg Philis	Small	1	4	100 (Gold)
Microsoft	A1 Standard	1	1.75	100 (Page Blob-LRS)
Microsoft	D1 v2	1	3.5	100 (Page Blob-LRS)

Table 2B: 2vCPU Virtual Machine Configurations

Provider	Instance Name	vCPU	Memory (in GB)	Block Storage (in GB)
Amazon	m4.large	2	8	250 (EBS SSD)
Schuberg Philis	Medium	2	8	250 (Gold)
Microsoft	D2 v2	2	7	250 (Page Blob-LRS)

Table 2C: 4vCPU Virtual Machine Configurations

Provider	Instance Name	vCPU	Memory (in GB)	Block Storage (in GB)
Amazon	m4.xlarge	4	16	500 (EBS SSD)
Schuberg Philis	Large	4	16	500 (Gold)
Microsoft	D3 v2	4	14	500 (Page Blob-LRS)

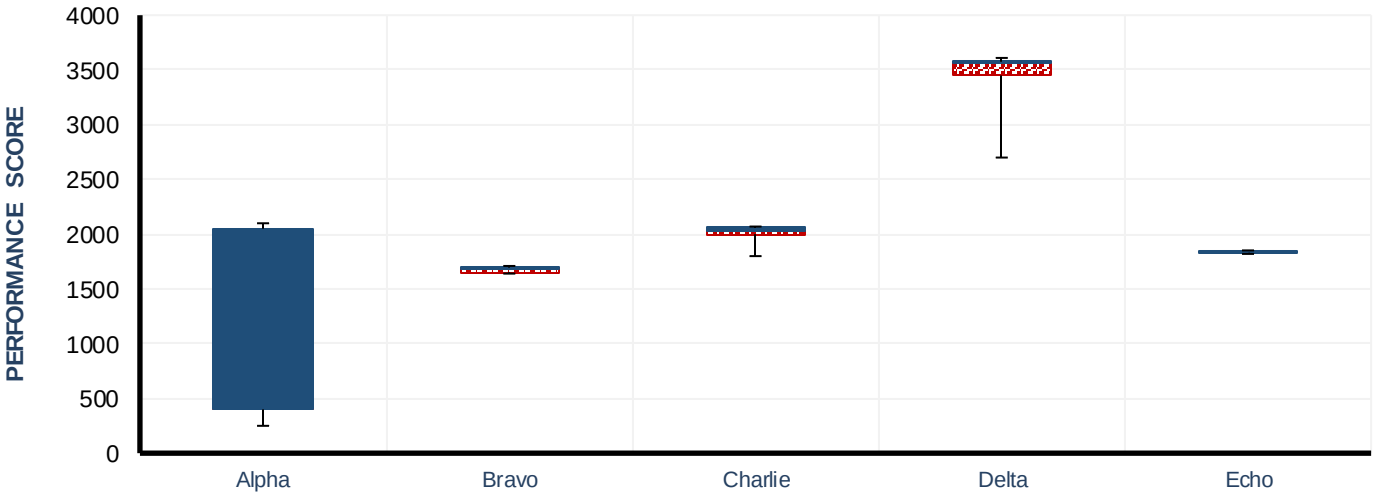
Table 2.1: Data Center Locations

Provider	Amazon	Microsoft	Schuberg Philis
Location	EU (Frankfurt)	West Europe	Schiphol

UNDERSTANDING THE CHARTS

VM performance is illustrated using percentile scores retrieved from all data points collected. 5th percentile and 95th percentile scores are often used instead of minimum and maximum scores in order to exclude potential outliers. The information has been integrated into percentile graphs and value tables designed to visualize performance variation captured while testing over time. An example of the performance percentile graph along with a corresponding value table is displayed in Chart 8A:

Chart 2: Sample Chart Guide



- **Maximum (MAX)**: The highest performance score(s) achieved on the VM over the course of the study.
- **95TH Percentile (95TH)**: 95% of all scores on the VM achieved this performance score or lower.
- **Median (MED)**: The number separating the higher and lower half of scores. If the median is closer to the 95TH percentile, then more high performance scores were observed than low performance scores (and vice versa).
- **5TH Percentile (5TH)**: 5% of all scores on the VM achieved this performance score or lower.
- **Minimum (MIN)**: The lowest performance score(s) achieved on the VM over the course of the study.

TEST CONSIDERATIONS

- Testing was conducted on specific VM types for each provider. Different VM configurations may yield different comparative results between the providers. Amazon and Azure offered fixed VM configurations, while Schuberg Philis offered independently customizable resource configurations.
- Users may experience different performance across different physical hosts. Factors such as user contention or malfunctions of the physical hardware can cause suboptimal performance. Cloud Spectator terminated and created new VMs for each test iteration to increase the likelihood of testing on different physical hosts.
- VMs selected were the base offerings across providers; greater performance may be obtained on certain providers by paying for additional features/services.

PROCESSOR PERFORMANCE

The processor tasks are segregated into two categories: integer and floating point tasks.

KEY OBSERVATIONS

The scores in Figures 3.1 A – F are indexed through a combination of floating point and integer performance on processors. Performance variation was elevated on the VMs which demonstrated higher performance, while variation was low on the lower performance VMs. Performance variation was determined by measurements of the coefficient of variation (CV), which is a percentage expressing the relationship between the average and standard deviation.

ADDITIONAL OBSERVATIONS

- Schuberg Philis VMs achieved the second highest performance across both the integer and floating point processor tests in the study; Azure achieved the highest performance across the VMs.
- Although AWS VMs' processor performance scored lower than Schuberg Philis and Azure, AWS demonstrated the lowest levels of performance variability throughout the test period, with CVs of 9% or lower on all VMs.

DETAILS

Chart 3.1A: Integer Performance (1vCPUs)

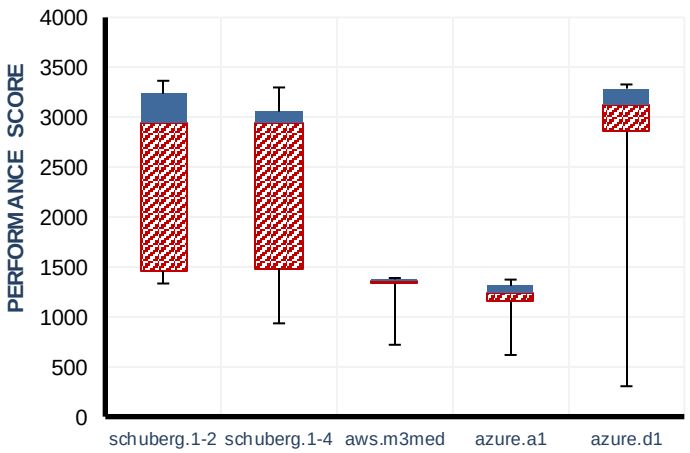


Table 3.1A: Integer Performance (1vCPUs)

	Amazon	SP		Microsoft	
VM Size	**m3.med**	**1c2m**	**1c4m**	**A1**	**D1**
Minimum	720	1334	935	618	305
5th	1331	1460	1486	1150	2852
Median	1359	2936	2936	1244	3117
95th	1375	3235	3056	1309	3287
Maximum	1389	3365	3298	1373	3328

Chart 3.1B: Integer Performance (2vCPU VMs)

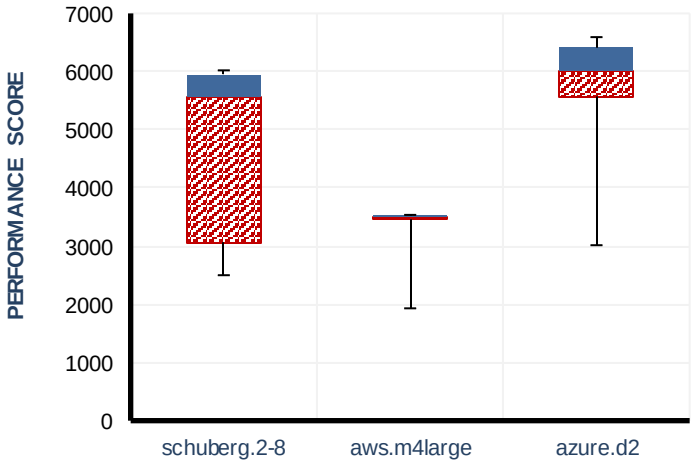


Table 3.1B: Integer Performance (2vCPUs)

	Amazon	SP	Microsoft
VM Size	**m4.large**	**2c8m**	**D2**
Minimum	1931	2500	3016
5th	3461	3067	5573
Median	3503	5554	6003
95th	3524	5956	6409
Maximum	3537	6021	6592

Chart 3.1C: Integer Performance (4vCPU VMs)

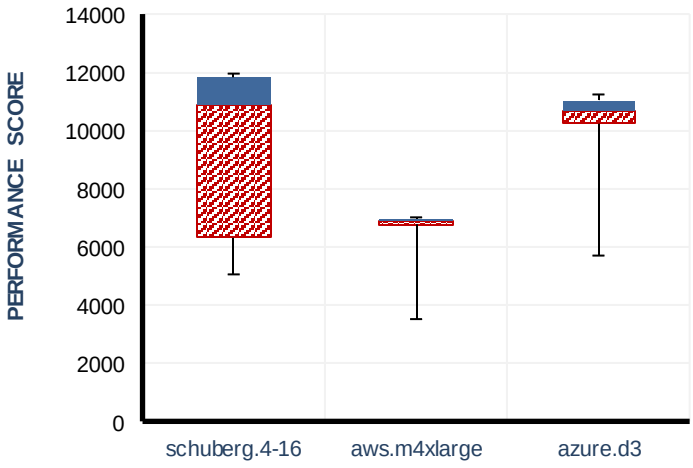


Table 3.1C: Integer Performance (4vCPUs)

	Amazon	SP	Microsoft
VM Size	**m4.xlarge**	**4c16m**	**D3**
Minimum	3515	5052	5698
5th	6775	6305	10231
Median	6913	10907	10660
95th	6965	11827	11044
Maximum	7012	11953	11236

Chart 3.1D: Floating Point Performance (1vCPU VMs)

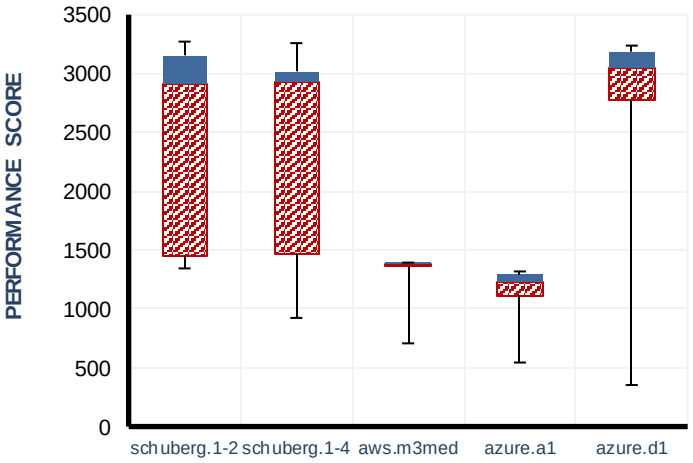


Table 3.1D: Floating Point Performance (1vCPUs)

	Amazon	SP		MSFT	
VM Size	**m3.med**	**1c2m**	**1c4m**	**A1**	**D1**
Minimum	707	1344	923	544	353
5th	1362	1450	1466	1116	2769
Median	1382	2911	2934	1229	3045
95th	1389	3155	3019	1290	3186
Maximum	1393	3271	3258	1318	3237

Chart 3.1E: Floating Point Performance (2vCPU VMs)

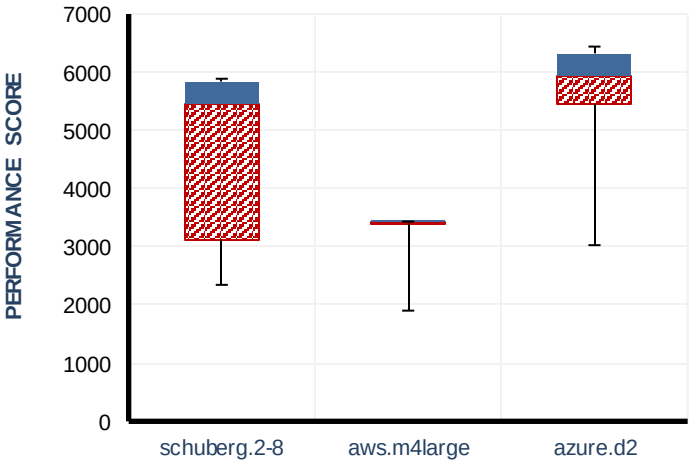


Table 3.1E: Floating Point Performance (2vCPUs)

	Amazon	SP	Microsoft
VM Size	**m4.large**	**2c8m**	**D2**
Minimum	1902	2345	3027
5th	3400	3119	5451
Median	3414	5451	5930
95th	3424	5839	6324
Maximum	3433	5887	6440

Chart 3.1F: Floating Point Performance (4vCPU VMs)

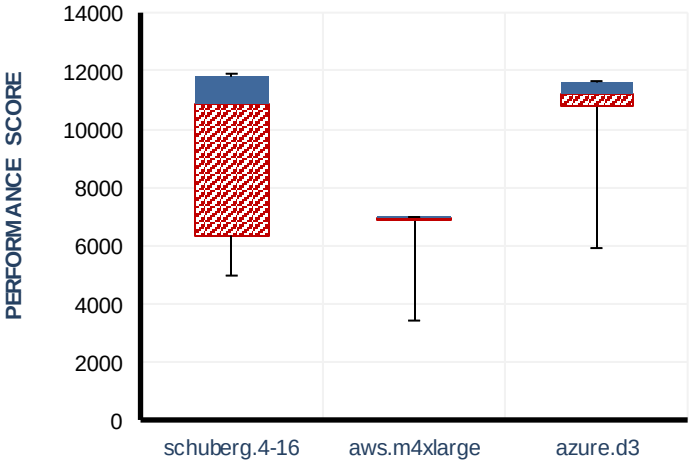


Table 3.1F: Floating Point Performance (4vCPUs)

	Amazon	SP	Microsoft
VM Size	**m4.xlarge**	**4c16m**	**D3**
Minimum	3416	4962	5908
5th	6839	6341	10758
Median	6950	10861	11234
95th	6966	11795	11601
Maximum	6976	11901	11647

MEMORY PERFORMANCE

Because memory bandwidth is generally not affected by the amount of provisioned RAM, VMs with larger amounts of RAM do not necessarily have higher scores.

KEY OBSERVATIONS

The scores in Figures 3.2 A – C are indexed STREAM test results on the memory bandwidth of the RAM. Schuberg Philis' VMs achieved the highest memory performance for the 2vCPU and 4vCPU VMs; Azure was highest for the 1vCPU VMs.

ADDITIONAL OBSERVATIONS

- Azure's performance varied the least across the three VM sizes, with an average CV of 8%. Schuberg Philis' CVs averaged 16% while the CVs on Amazon's VMs averaged 10%.
- Amazon's memory bandwidth scales up by 183% from the 1vCPU to the 2vCPU VM, and 73% up to the 4vCPU VM. Schuberg Philis scales up by 91% from 1vCPU to 2vCPU VM, and 87% from 2vCPU to 4vCPU. Microsoft scales 20% from 1vCPU to 2vCPU among the D-series VMs, but stops scaling from there.

DETAILS

Chart 3.2A: Memory Performance (1vCPU VMs)

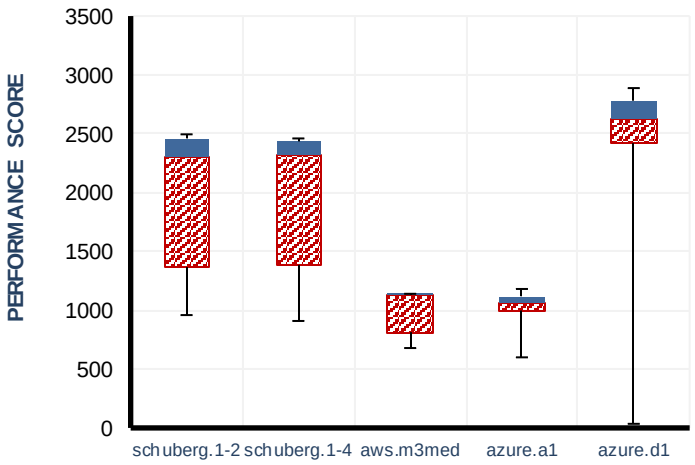


Table 3.2A: Memory Performance (1vCPUs)

	Amazon	SP		MSFT	
VM Size	**m3.med**	**1c2m**	**1c4m**	**A1**	**D1**
Minimum	679	959	909	600	35
5th	814	1368	1385	998	2420
Median	1126	2304	2311	1061	2629
95th	1135	2460	2435	1119	2781
Maximum	1140	2495	2460	1181	2888

Chart 3.2B: Memory Performance (2vCPU VMs)

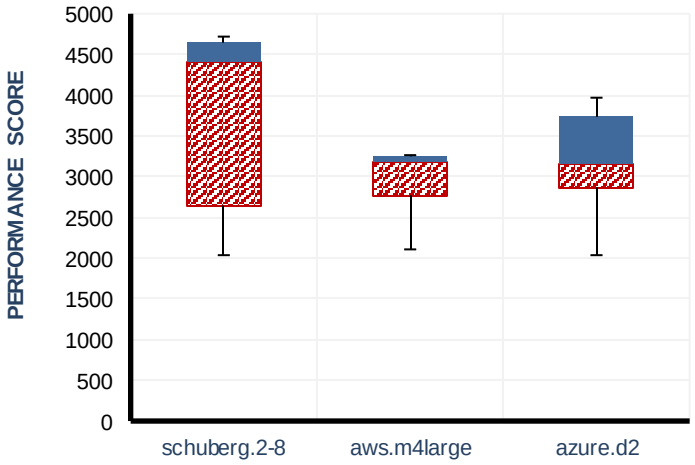


Table 3.2B: Memory Performance (2vCPUs)

	Amazon	SP	Microsoft
VM Size	**m4.large**	**2c8m**	**D2**
Minimum	2107	2036	2036
5th	2772	2647	2871
Median	3190	4418	3145
95th	3250	4644	3732
Maximum	3264	4719	3969

Chart 3.2C: Memory Performance (4vCPU VMs)

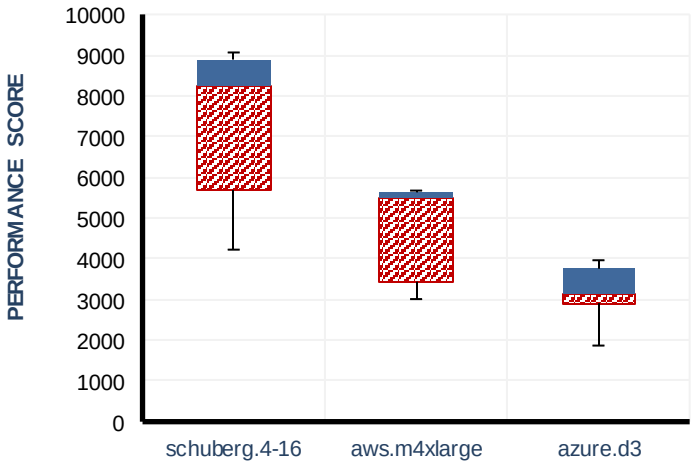


Table 3.2C: Memory Performance (4vCPUs)

	Amazon	SP	Microsoft
VM Size	**m4.xlarge**	**4c16m**	**D3**
Minimum	3010	4224	1869
5th	3436	5690	2910
Median	5516	8250	3153
95th	5631	8891	3756
Maximum	5674	9065	3964

BLOCK DISK IOPS

Disk IOPS was tested with direct I/O, so results are not reflective of cached performance, which may sustain higher IOPS on each provider.
For Random and Sequential Operations, Reads and Writes were run independently and combined for the overall Random/Sequential scores.

KEY OBSERVATIONS

Providers offer unique approaches to SAN disk, from hardware components (SSD vs. traditional magnetic drives) to performance throttling for differently tiered products. Even within a performance tier, throttling may be set to a range of values, rather than a set value that a customer can always expect.

ADDITIONAL OBSERVATIONS

- Schuberg Philis VMs demonstrated the highest median and maximum performance than all other providers tested, with maximums exceeding 50,000 IOPS for sequential operations.
- Overall the three providers express significant performance variability across the period of testing. Examining Random Read/Write operations, Amazon's variability drops as the storage volume increases; exhibiting a CV of 20% on the 1vCPU/100GB VM and dropping down to 6% on the 4vCPU/500GB VM (which is due to AWS's burst functionality). Azure's CVs ranged from 10% to 12%, while Schuberg Philis displayed noticeable variability with CVs ranging up to 38%.
- Azure storage performance is determined by the amount of storage utilized and not the volume size allocated, thus lower performance thresholds were observed.

DETAILS

Chart 3.3A: Sequential IOPS
(1vCPUs - 100GB)

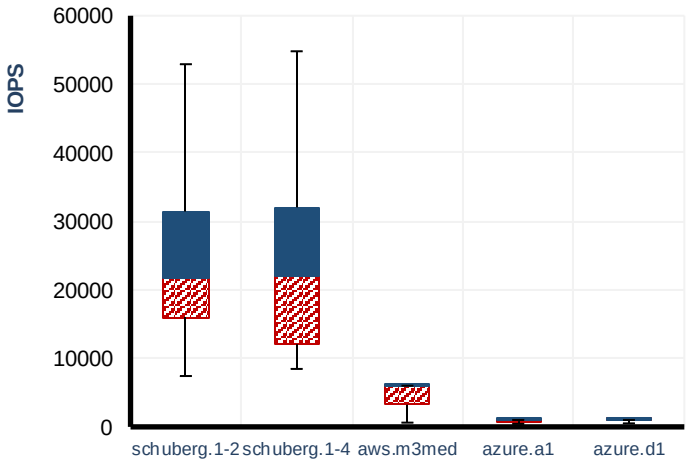


Table 3.3A: Sequential IOPS
(1vCPUs)

	Amazon		SP	Microsoft	
VM Size	**m3.med**	**1c2m**	**1c4m**	**A1**	**D1**
Minimum	618	7421	8456	438	510
5th	3224	16033	12074	839	910
Median	6004	21797	21949	1007	1007
95th	6008	31431	32035	1009	1009
Maximum	6008	52923	54812	1010	1009

Chart 3.3B: Sequential IOPS
(2vCPUs - 250GB)

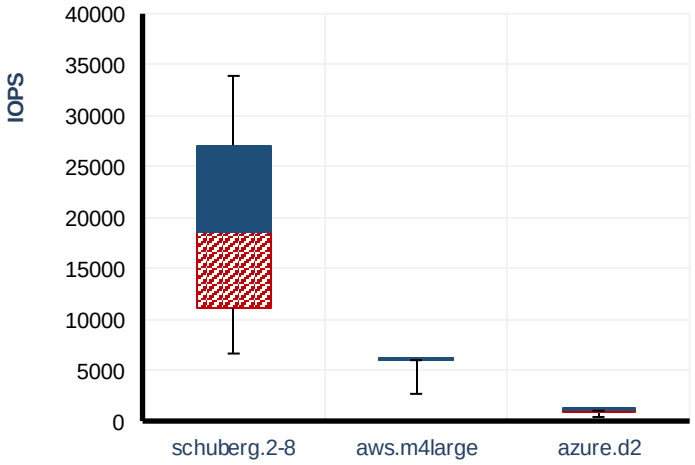


Table 3.3B: Sequential IOPS
(2vCPUs)

	Amazon	SP	Microsoft
VM Size	**m4.large**	**2c8m**	**D2**
Minimum	2668	6627	396
5th	5915	11081	795
Median	5966	18581	1006
95th	5973	27100	1008
Maximum	5975	33895	1010

Chart 3.3C: Sequential IOPS
(4vCPUs - 500GB)

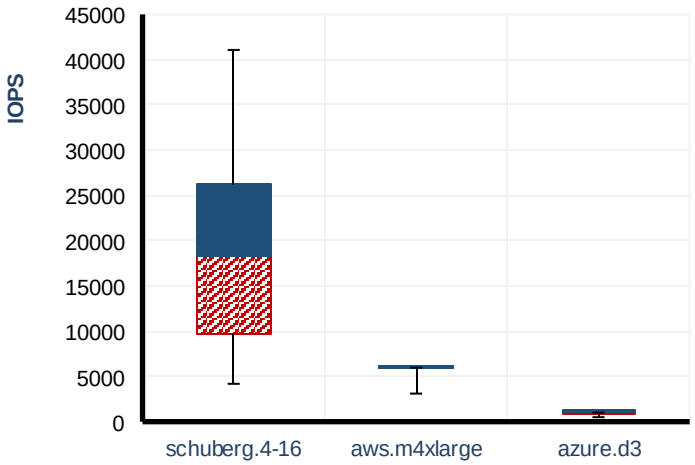


Table 3.3C: Sequential IOPS
(4vCPUs)

	Amazon	SP	Microsoft
VM Size	**m4.xlarge**	**4c16m**	**D3**
Minimum	3105	4184	496
5th	5918	9741	838
Median	5965	18240	1003
95th	5971	26208	1008
Maximum	5974	41074	1009

Chart 3.3D: Random IOPS
(1vCPUs - 100GB)

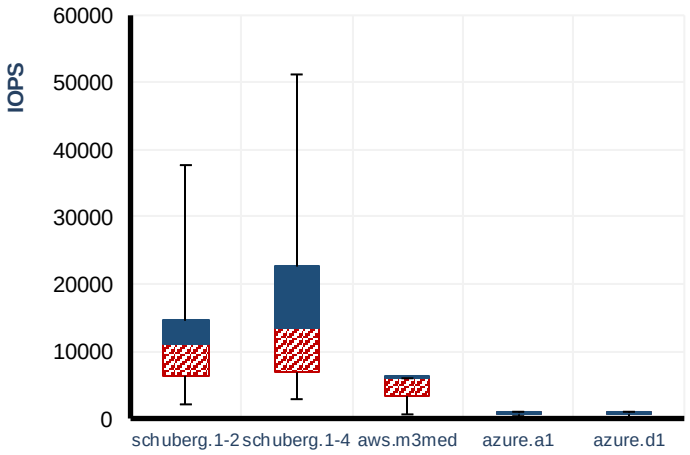


Table 3.3D: Random IOPS
(1vCPUs)

	Amazon	SP		Microsoft	
VM Size	**m3.med**	**1c2m**	**1c4m**	**A1**	**D1**
Minimum	614	2100	2877	339	182
5th	3327	6341	6866	618	620
Median	5986	11028	13528	742	741
95th	5987	14546	22612	811	816
Maximum	5987	37707	51200	997	997

Chart 3.3E: Random IOPS
(2vCPUs - 250GB)

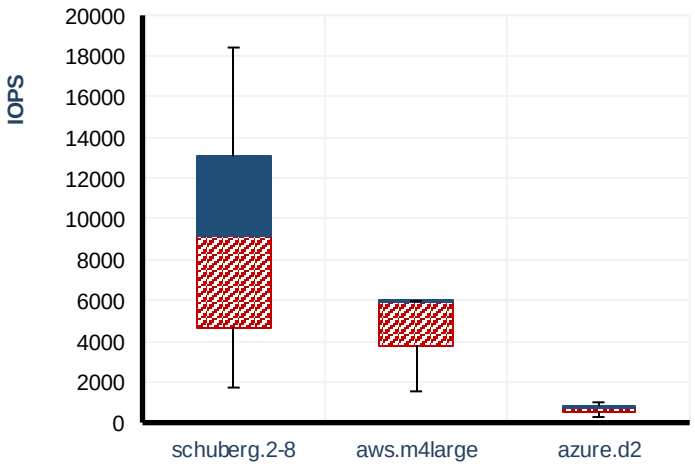


Table 3.3E: Random IOPS
(2vCPUs)

	Amazon	SP	Microsoft
VM Size	**m4.large**	**2c8m**	**D2**
Minimum	1535	1721	274
5th	3799	4675	542
Median	5958	9182	743
95th	5960	13071	817
Maximum	5962	18418	997

Chart 3.3F: Random IOPS
(4vCPUs - 500GB)

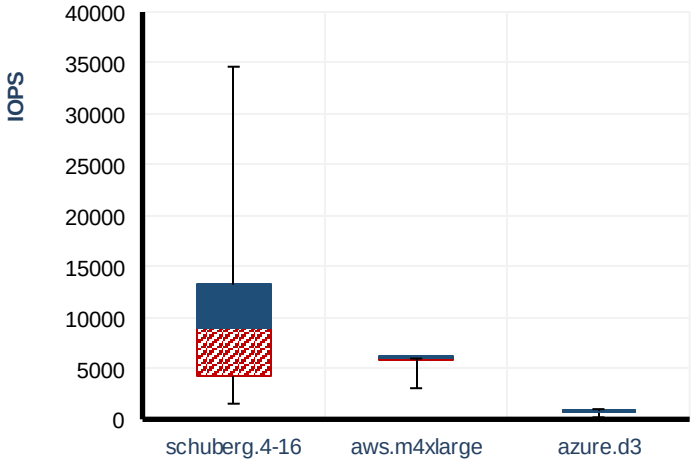


Table 3.3F: Random IOPS
(4vCPUs)

	Amazon	SP	Microsoft
VM Size	**m4.xlarge**	**4c16m**	**D3**
Minimum	3052	1532	191
5th	5908	4249	629
Median	5958	8984	744
95th	5960	13218	812
Maximum	5961	34625	998

INTERNAL NETWORK THROUGHPUT

Internal network was examined by running TCP connections between two servers within the same data center/region/zone.

KEY OBSERVATIONS

With regards to internal network throughput, two out of three providers exhibited high performance variability with CVs 10% or higher. Azure VMs expressed high variability that ranged up to 27% (A1 VM). Schuberg Philis VMs' performance variability was similar, exhibiting CVs ranging up to 29% (2vCPU, 8GB RAM VM). AWS's throughput was relatively stable, exhibiting CVs that ranged no higher than 9%. When examining the median value of the same VM size, Schuberg Philis achieves much higher throughput (above 5Gbits) than the other VMs. Despite the fluctuations in performance, Schuberg Philis' VMs still achieved higher throughput than the other provider VMs did.

ADDITIONAL OBSERVATIONS

- Schuberg Philis provides a wide range of throughput roughly 2-9 Gbit/s regardless of VM size.
- Amazon's VMs deliver a specific network throughput depending on the size and family of VM. As the size of the VMs increase, the network throughput increases as well. The throughput starts at 297 Mbit/s for the 1vCPU m3 VM, rises to ~445 Mbit/s for the 2vCPU m4 VM, and settles at ~754 Mbit/s for the 4vCPU m4 VM.
- Azure VMs also appear to deliver a specific throughput depending on the size of the VM. However Azure offers a general range rather than a set throughput target.

Chart 3.4A: Internal Network Throughput (1vCPUs)

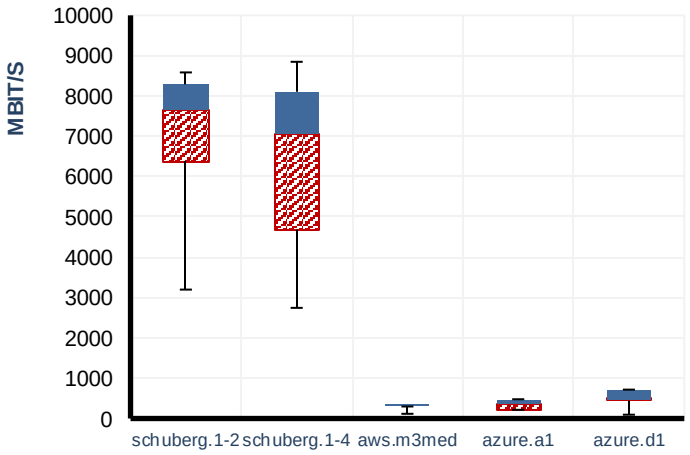


Table 3.4A: Internal Network Throughput (1vCPUs)

	Amazon	SP		Microsoft	
VM Size	**m3.med**	**1c2m**	**1c4m**	**A1**	**D2**
Minimum	116	3197	2744	215	94
5th	297	6387	4686	229	446
Median	297	7640	7054	356	474
95th	298	8289	8105	476	711
Maximum	298	8584	8847	477	716

Chart 3.4B: Internal Network Throughput (2vCPUs)

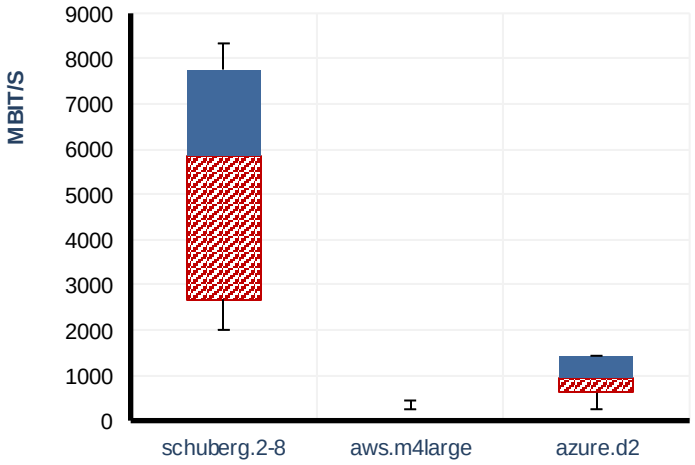


Table 3.4B: Internal Network Throughput (2vCPUs)

	Amazon	SP	Microsoft
VM Size	**m4.large**	**2c8m**	**D2**
Minimum	251	2005	252
5th	445	2681	613
Median	445	5837	935
95th	445	7759	1430
Maximum	445	8335	1433

Chart 3.4C: Internal Network Throughput (4vCPUs)

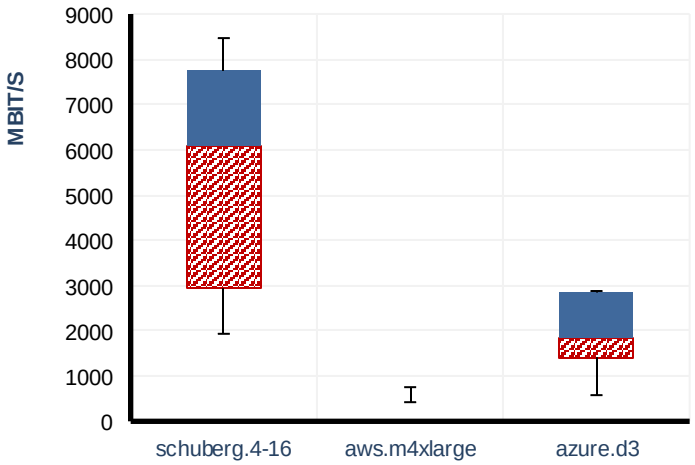


Table 3.4C: Internal Network Throughput (4vCPUs)

	Amazon	SP	Microsoft
VM Size	**m4.xlarge**	**4c16m**	**D3**
Minimum	422	1933	578
5th	754	2941	1418
Median	754	6068	1852
95th	754	7742	2845
Maximum	754	8467	2879

CONCLUSION

Selecting the right provider and virtual machines requires a thorough and accurate performance comparison. The study conducted for this report offers a general understanding of performance strengths and weaknesses across each included vendor's VMs, and should be considered a source of information to help guide readers in their own testing and analyses. The processor, memory bandwidth, network storage, and internal network were all examined and results clearly show that no single provider can offer everything to fit everyone's needs. Therefore, in order to accurately select the right provider for a business or an application, performance analysis is crucial.

Results from this study show that, Schuberg Philis VMs displayed strong overall performance. Schuberg Philis VMs demonstrated high performance for processing, memory bandwidth, storage and internal network. Its network-attached storage produced the largest amount of maximum IOPS observed in the study for read/write operations. Internal network throughput levels exceeded those of AWS and Azure by magnitudes of 3x and more.

Performance in the industry cannot be assumed to be equal or even similar, as illustrated in this report. When it comes to processor and memory bandwidth performance, tiered providers such as Amazon and Microsoft offer varying performance depending on the family/series of the VM, despite having equivalent amounts of vCPUs and similar amounts of memory.

While this study was conducted in the manner of understanding a typical end user experience, it should not be assumed to be accurate for all use cases. Stress testing was conducted to better understand fluctuation and theoretically sustained performance, and should be seen as a general indication of overall provider performance.

ABOUT CLOUD SPECTATOR

Cloud Spectator is a cloud analyst agency focused on cloud Infrastructure-as-a-Service (IaaS) performance. The company actively monitors several of the largest IaaS providers in the world, comparing VM performance (i.e., CPU, RAM, disk, internal network, and workloads) and pricing to achieve transparency in the cloud market. The company helps cloud providers understand their market position and helps business make intelligent decisions in selecting cloud providers and lowering total cost of ownership. The firm was founded in early 2011 and is located in Boston, MA.

For questions about this report, to request a custom report, or if you have general inquiries about our products and services, please contact Cloud Spectator at +1 (617) 300-0711 or contact@cloudspectator.com.