GIGAOM RESEARCH

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series

A Buyer's Lens Report by Anne Qingyang Liu

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series

02/06/2015

Table of Contents

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

2

1 Executive Summary

In Fall 2014, Microsoft Azure released its D-Series Infrastructure as a Service (IaaS), significantly improving performance over the standard A-Series.

Performance and pricing are both key considerations in evaluating public-cloud options. The notorious price wars between some of the largest providers have pressured the industry into cutting margins to stay competitive. Simultaneously, performance is quickly becoming a focus as providers introduce higher-performing offerings, such as Amazon EC2's C4 Family, Rackspace's Performance Servers, and Microsoft Azure's D-Series. This report examines Microsoft Azure's D-Series, and how it has improved from a performance perspective.

Key findings from this report include:

- The D-Series virtual machines (VMs) have higher vCPU, memory, and storage performance outputs than the A-Series VMs when examined across different machine sizes.

- A-Series VMs are priced lower, making them more suitable for applications like small websites and low-transaction databases.

- Applications that need higher-performance requirements (NoSQL clusters, high-transaction databases) will find a better fit with the D-Series or the more recently released G-Series.

Key Findings of vCPU, Memory, and Storage Performance

Performance	Key Findings		
vCPU	The D-Series' vCPUs outperformed the A-Series' vCPUs by 58%.	The vCPU performance of the D4 instances showed relatively large variation, primarily due to two abnormally low scores. This could be caused by a possible malfunction on one or more specific physical hosts.	The difference in physical hardware had no notable impact on the vCPU performance of the A-Series VMs.
Memory	The D-Series VMs displayed, on average, 65% more memory bandwidth than their A-Series counterparts.	The maximum memory bandwidth on the D-Series was 58% more than on the A-Series (17.2 GB/s vs. 10.9 GB/s).	Significant fluctuations on memory throughput were observed on the A2, A3, D2 and D4 instances. Variability ranged from 20% to 52% of the average across the five-day testing period.
Storage	The D-Series had an average of 6.3x more IOPS than the A-Series for local storage.	The D-Series write IOPS scaled significantly as the VM sizes became larger.	The A-Series showed much more storage performance variability than the D-Series. D-Series performance variability was under 5% for almost all measurements.

Source: Cloud Spectator

Thumbnail image courtesy of D3Damon/iStock.

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

3

2 Test Methodology

Cloud Spectator performed 10 iterations of vCPU, memory, and storage tests on several A-Series and D-Series VMs of equivalent size (1 vCPU, 2 vCPU, 4 vCPU and 8 vCPU, accordingly the A1/D1, A2/D2, A3/D3 and A4/D4 instances.) Tests were conducted twice per day for five consecutive days in order to capture variation in performance across different physical hosts. Machines were terminated and recreated between each test session to increase likelihood of provisioning on different physical nodes. Doing so provides better insight into variability that was caused by differences in physical hardware as well as potential "noisy neighbor" effects.

Performance variability (PV) was tested across different machines, and was calculated as the standard deviation expressed as a percentage of the average.

Performance Variability = (Standard Deviation / Average) * 100%

Cloud Spectator used the fio test to measure local storage performance by gauging the amount of IOPS realized over a two-minute interval with an 8KB block size. The block size was chosen based on its relevance to structured databases. According to official documentation, fio is an I/O tool meant to be used for both benchmark and stress/hardware verification with a specific focus on I/O-related tasks.

Key Considerations

When examining the results of these tests, the following should be kept in mind:

- The D-Series was introduced eight days prior to the beginning of Cloud Spectator's testing. The physical machines may not have been as saturated with users relative to the A-Series, which had been available for more than one year. Therefore, retesting the D-Series in a few months may result in different findings.

- The tests leveraged in this study were used to standardize and compare performance. They may not be reflective of variations in application performance.

- Cloud Spectator tested each machine size multiple times, terminating and recreating machines for each test iteration. This was done intentionally, to sample different physical hosts.

- Two of the test iterations on the D4 machines produced outlying data points that were significantly lower than other D4 data points. This skewed the average performance scores for this specific machine type.

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

4

3 Why Performance Matters

Cloud-infrastructure performance is a key consideration that should not be overlooked. Differences in performance outputs of VMs across IaaS providers and their variety of product offerings can greatly impact quality of service (QoS) as well as total cost of ownership (TCO). Selecting the wrong product or provider to house an application can result in unnecessary overspending and potential application performance problems.

By examining the underlying hardware we can understand the theoretical maximum and sustained performance of each component that comprises the server. The table below lists the three virtual server components studied in this project.

Infrastructure Components Summary

| vCPU	Memory	Storage
The performance of all applications is highly dependent on the vCPU. The vCPU is responsible for the processing and orchestration of all applications.	While memory performance is not considered one of the key bottlenecks in performance for many applications, a subset of applications—particularly HPC and in-memory databases—is highly dependent on large sustained memory bandwidth.	Because most applications and all data reside on the storage, having fast storage performance is a key consideration for smooth application performance in many cases.

Source: Cloud Spectator

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

5

4 Performance Findings

vCPU Performance

Cloud Spectator used Geekbench 3 to test vCPU performance and memory bandwidth. Geekbench 3 is a modern, system-focused stress-testing tool that simulates real-world scenarios for servers, desktops, and mobile devices. Such use cases include but are not limited to encryption, decryption, compression, and solving equations.
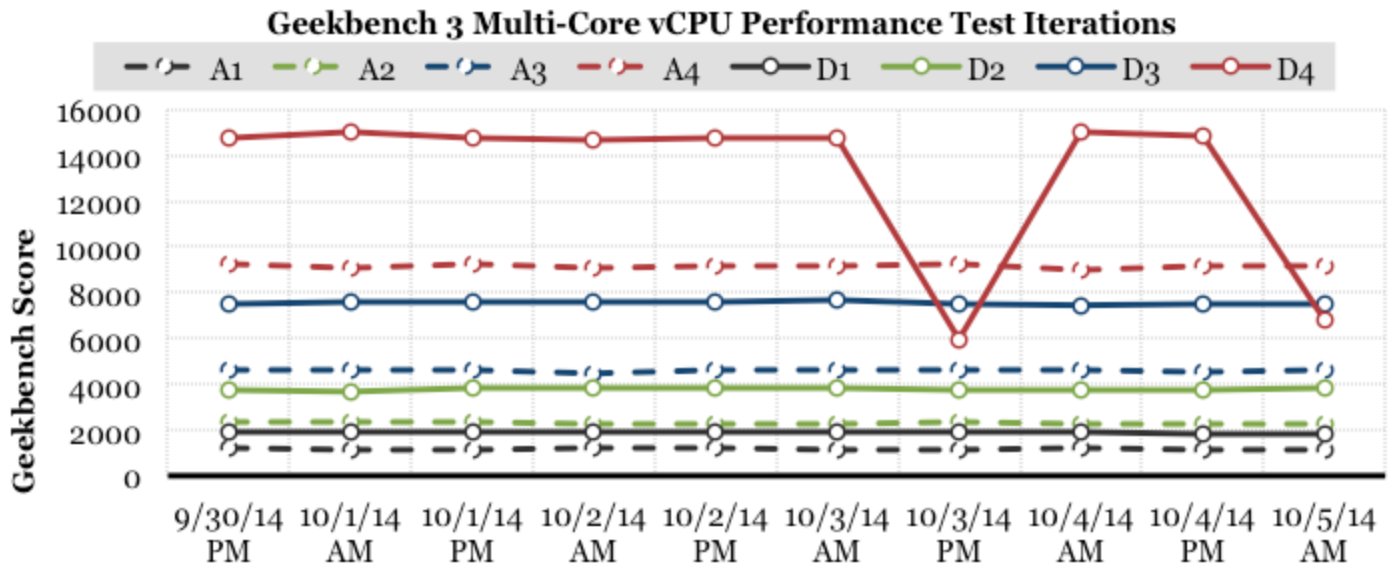
On average, Microsoft Azure's D-Series VMs outperformed their A-Series counterparts by 58 percent (see chart below). The difference between the A4 and D4 was only 43 percent due to the considerable performance variation within the D4 instance.

Geekbench 3 Multicore vCPU Performance Comparision
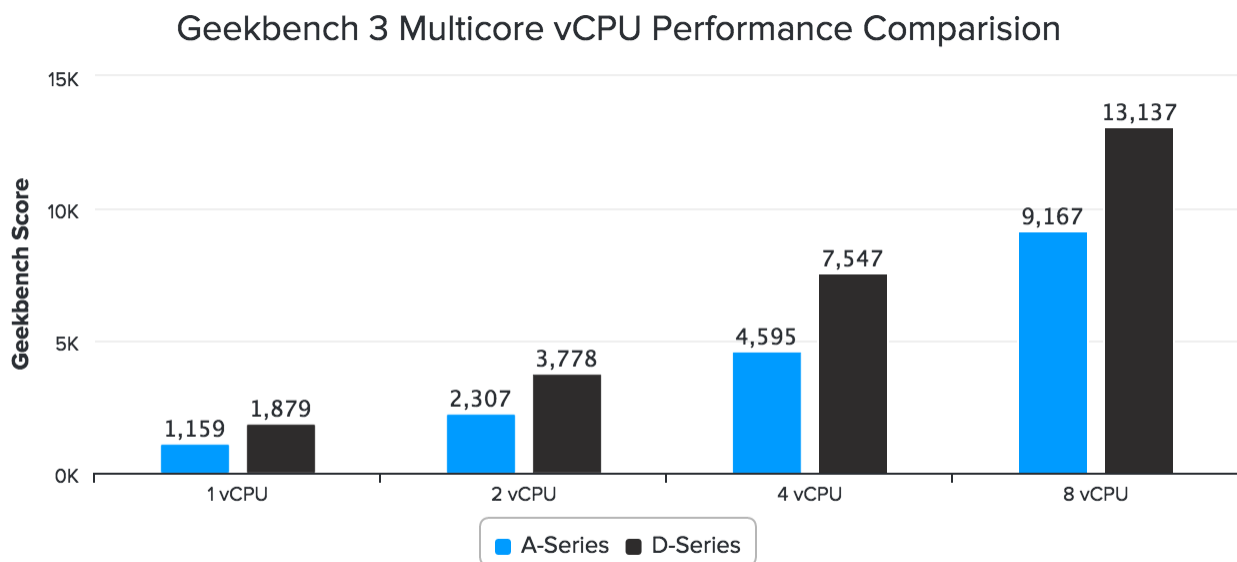
Cloud Spectator

The performance variability tests yielded consistently stable performances ranging between 1 percent and 4 percent. The PV for the D-Series D4 instance was significantly higher, at 27 percent, due to two extremely low vCPU performance scores observed on separate testing days (see chart below). The D-Series variation could be attributed to a new service that is not as saturated with VMs as the more established A-Series platform.

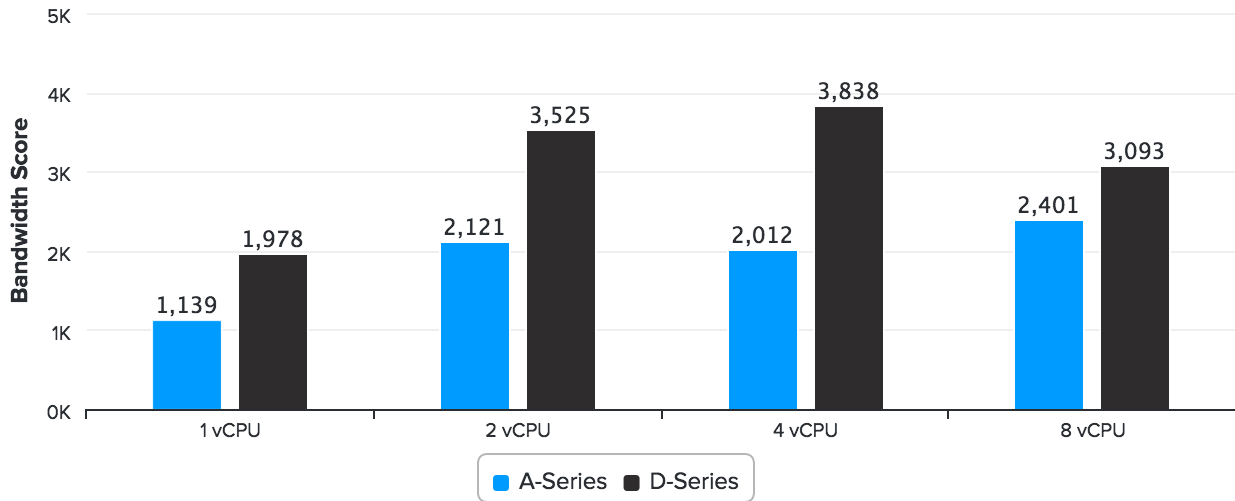**A-Series vs. D-Series vCPU Performance Variability**

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

6

Geekbench 3 Multi-Core vCPU Performance Test Iterations

Source: Cloud Spectator

The A-Series machines were provisioned on two types of physical processors: the AMD OpteronTM Processor 4171 HE and the Intel Xeon CPU E5-2660. While the competing hardware did not lead to significant differences in overall vCPU performance, there were two exceptions noted. The AMD-based VMs exhibited a 6 percent performance advantage in integer-related compute tasks, while the Intel-based VMs exhibited a 3 percent performance advantage in floating point-related compute tasks.
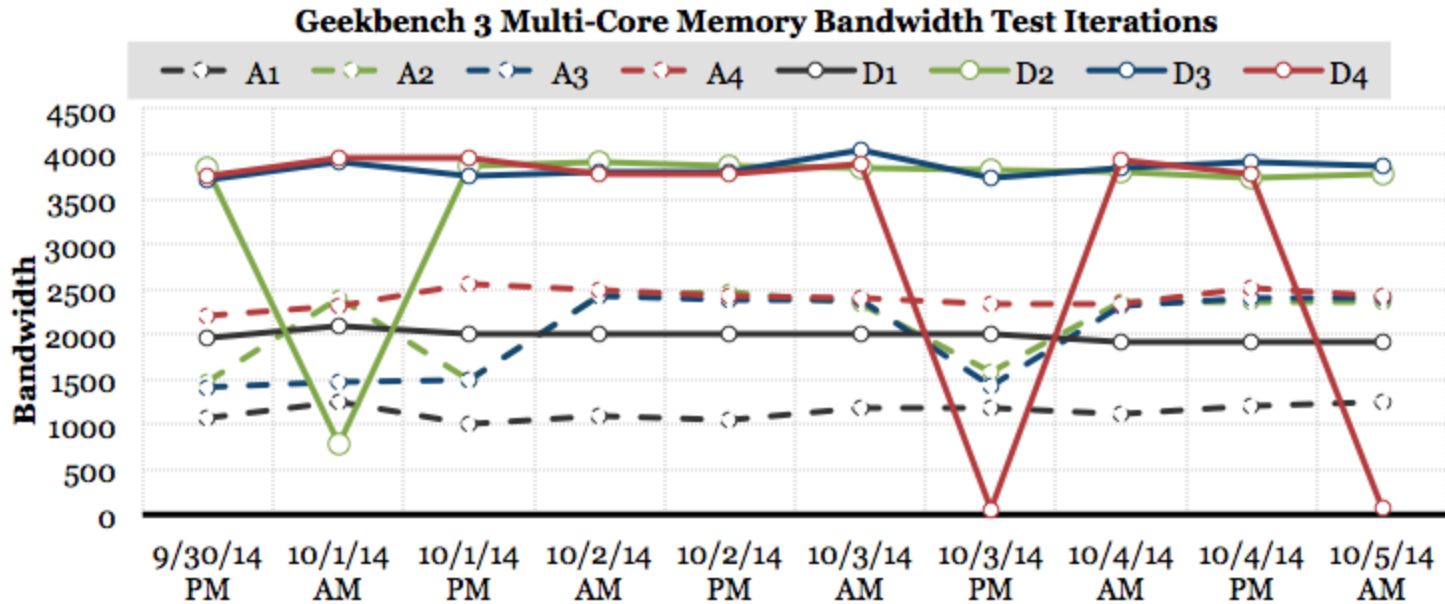
Memory Performance

Microsoft Azure D-Series machines had 65 percent higher memory bandwidth on average than their A-Series counterparts (see chart below). The performance differences between the A-Series and D-Series fluctuated across different machine sizes due to some abnormally low scores observed on certain instances for separate testing iterations, which skewed some averages. The highest memory bandwidth observed for the A-Series was 10.9 GB/s, whereas the highest recorded bandwidth for the D-Series was 58 percent higher, at 17.2 GB/s.

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

7

Geekbench 3 Multicore Memory Bandwidth Comparison

Cloud Spectator

The A2, A3, D2 and D4 instances had relatively large fluctuations in memory bandwidth. The Performance Variability for A2, A3, D2 and D4 were 20%, 24%, 27% and 52% respectively, while the average Performance Variability for the other instances was approximately 4% (see chart below). Different types of hardware had an impact on the memory bandwidth of the A-Series. VMs running on an AMD Opteron Processor 4171 HE showed 60% higher memory bandwidth than VMs running on the Intel Xeon CPU E5-2660. Both the D2 and D4 instances had sharp decreases in memory bandwidth on separate testing days, which may be attributed to malfunctions on the physical host.
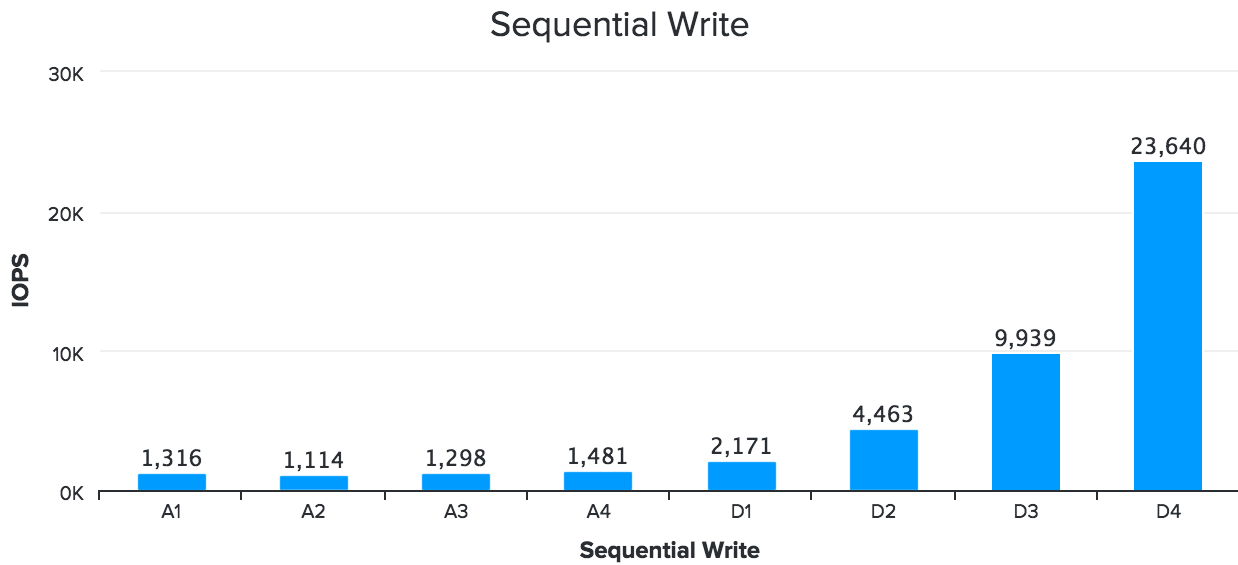
Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

8

Geekbench 3 Multi-Core Memory Bandwidth Test Iterations
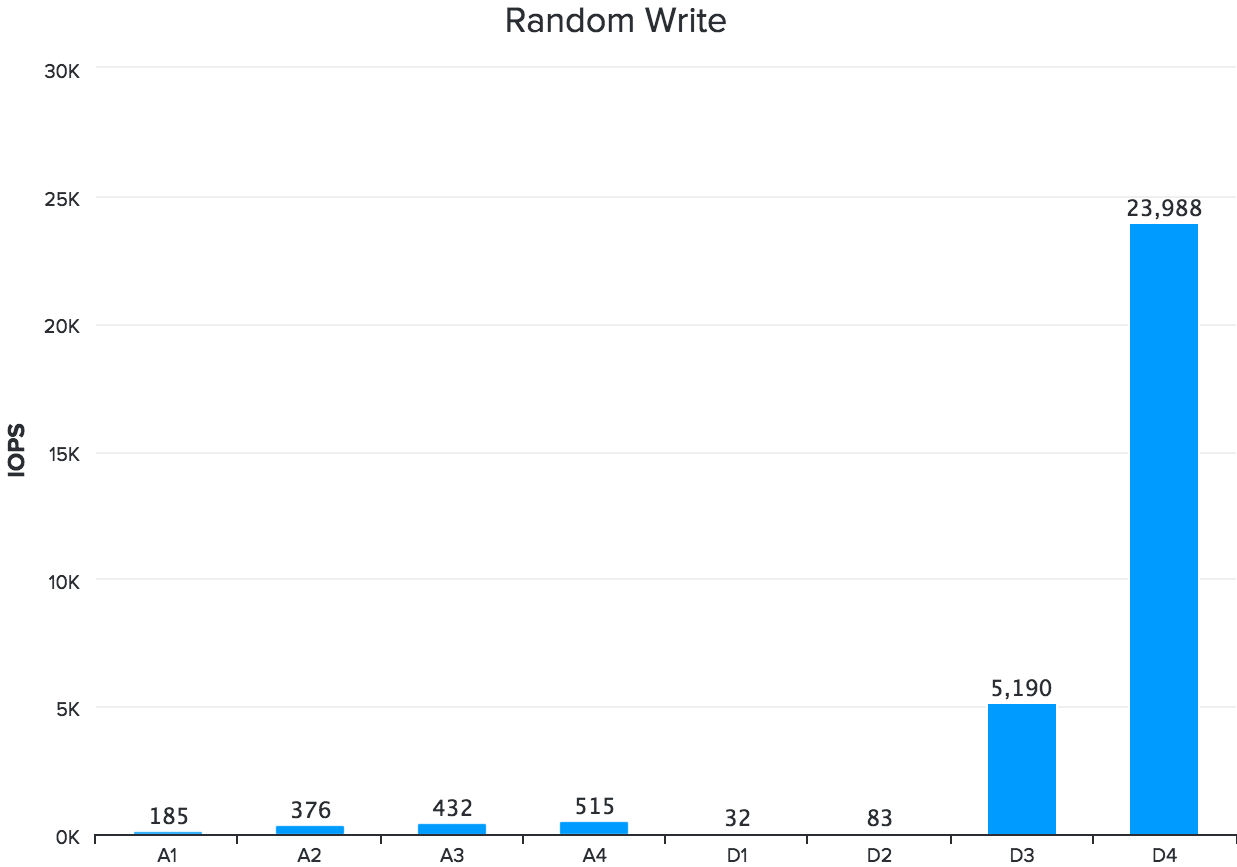
Storage Performance

The D-Series machines outperformed the A-Series machines in all four local storage subtests: sequential read/write and random read/write. The D-Series' storage performance scaled considerably with the VM size for sequential and random write tasks while the A-series scaled slightly for the random write task. A possibility for this pattern may be attributed to I/O thresholds set by the virtualization platform. No significant scaling pattern was observed for read tasks (see charts below).

The D-Series VMs on average had 6.3X greater local sequential write IOPS than their A-Series counterparts:

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

9

Sequential Write

Cloud Spectator

The D-Series VMs on average had 13.7x greater local random write IOPS than their A-Series counterparts:

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

10

Random Write

Cloud Spectator

The D-Series VMs on average had 2.7x greater local sequential read IOPS than their A-Series counterparts:

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)
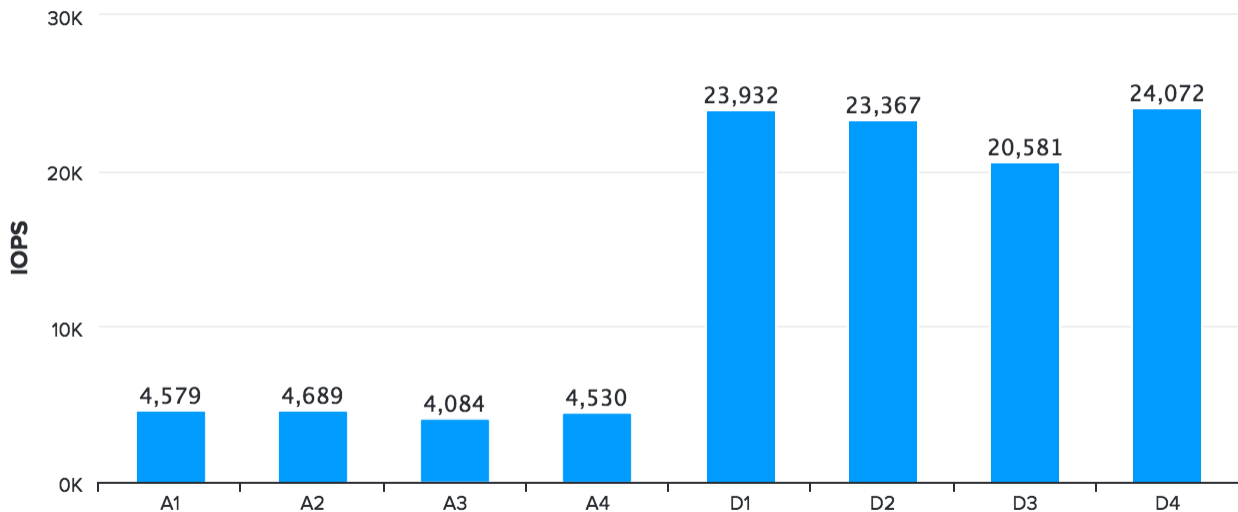
11

Sequential Read
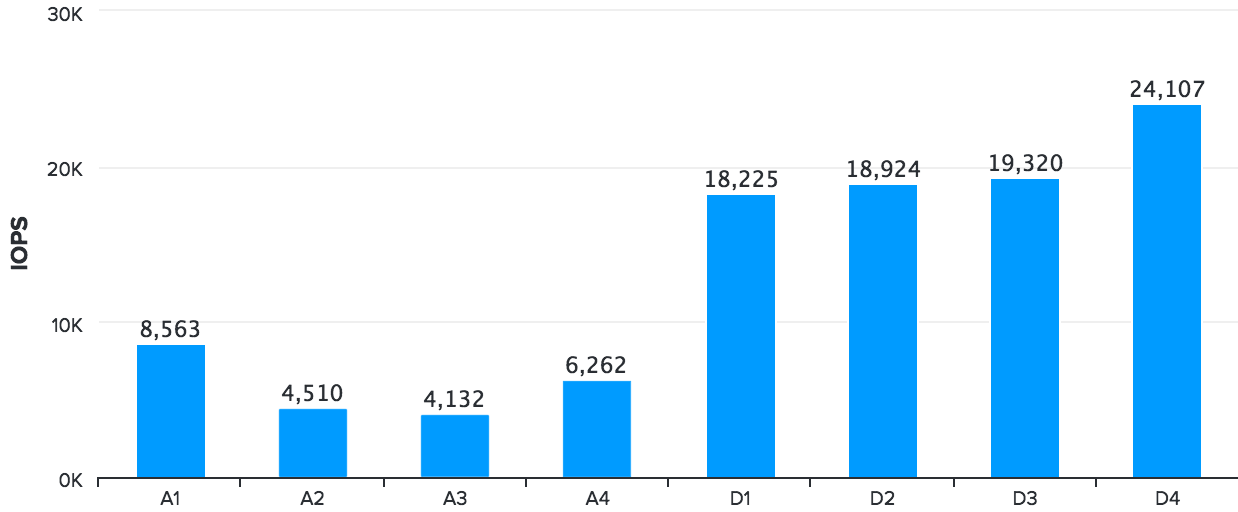
Cloud Spectator

The D-Series VMs on average had 4.1x greater local random read IOPS than their A-Series counterparts:

Random Read

Cloud Spectator

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

12

As mentioned in previous sections, the A-Series VMs were provisioned on different physical processors during the tests: the AMD Opteron Processor 4171 HE and the Intel Xeon CPU E5-2660. This variable affected the storage performance variability among the A-Series instances over the test iterations. Sequential operations were higher on the AMD OpteronTM processor-based VMs (on average 23.9x higher for reads and 1.3x higher for writes), while random operations were higher on the Intel Xeon processor-based VMs (on average 31 percent higher for reads and 108 percent higher for writes).

Compared to the A-Series, whose storage PV ranged from 8 percent to 213 percent, the D-Series showed significantly less variability. The D-Series' storage PV mainly ranged from 0 percent to 4 percent, with only two exceptions in the random write tests (see chart below). The stability of the D-Series environments, however, may be attributed to a lower number of users, thereby limiting user contention.

Performance Variability (%)

Task	A1	A2	A3	A4	D1	D2	D3	D4
Sequential Write	45	42	40	33	2	2	3	2
Random Write	22	20	22	24	4	16	37	1
Sequential Read	114	213	194	203	2	1	2	0
Random Read	22	13	8	15	2	2	3	0

Source: Cloud Spectator

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

13

5 Key Takeaways

The combination of more-consistent performance and more-powerful hardware (in the form of modern Intel processors and SSD-backed local storage) in the D-Series provides a powerful IaaS option for Azure users looking for a higher-performance option to run applications. According to the findings, Cloud Spectator recommends that the A-Series VMs are more economical for certain applications (e.g., small websites, low-transaction databases) because of the lower comparative cost, but applications with high performance requirements (e.g., NoSQL cluster environments, high-transaction databases) should look to D-Series environments.

The performance improvements of the D-Series are representative of the general trends within the IaaS industry. Other large players such as Amazon, Rackspace, and Google have also introduced updated offerings that are focused on improved performance. VM performance is becoming a very important consideration not only because users need their applications to be fast, but also because of the potential cost savings realized when buying less infrastructure on better-performing services. As cloud-infrastructure performance continues to improve and prices continue to drop, the IaaS industry as a whole is consistently delivering more value to its users.

Further Study

This report provides an introductory analysis into the new Azure D-Series machines to create a foundation for further study. It should only be used as a starting point to address further questions.

Additional studies should consider:

- Internal network performance testing, which is a key component for cluster environments.

- Application performance, which may experience different bottlenecks for different requirements. For further interest in specific application performance, users should target application-focused benchmark tests.

- Comparing Microsoft Azure's D-Series with recent offerings from other providers to achieve a better understanding of Microsoft Azure's position in the market.

- Testing Microsoft's persistent SAN storage to gain a more thorough understanding of storage performance.

- A price-performance value analysis to understand whether the D-Series offers more price-performance value per unit of cost as compared to the A-Series.

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

14

- An evaluation of Microsoft Azure's G-Series.

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

15

6 Appendix

Test Descriptions

Tests	Geekbench 3	fio
Description	Tests 27 common computational tasks (e.g. compression, encryption). Index scores aggregate test results from individual subtests. Index indicates overall memory and CPU performance.	I/O tool used both for benchmark and stress/hardware verification, with a specific focus on I/O related tasks. Gauges the amount of IOPS performed over a 2-minute interval. The results includes sequential and random read and write scores.
Testing Notes	The 64-bit processor benchmarks were used.	Fio used a queue depth of 32 for 1 vCPU and 2 vCPU VMs, and an queue depth of 128 for 4 vCPU and 8 vCPU VMs. File size of 1 GB for all disk tests with 8KB block size.

Server Setups

Providers	Microsoft Azure A-Series	Microsoft Azure D-Series
1 core VM	A1	D1
Virtual Cores	1	1
Memory (in GB)	1.75	3.5
2 cores VM	A2	D2
Virtual Cores	2	2
Memory (in GB)	3.5	7
4 cores VM	A3	D3
Virtual Cores	4	4
Memory (in GB)	7	14
8 cores VM	A4	D4
Virtual Cores	8	8
Memory (in GB)	14	28
Operating System	Ubuntu 14.04	Ubuntu 14.04
Data Center Location	U.S. South Central	U.S. South Central

Notes on Provisioning

- All virtual machines were provisioned in U.S. South Central.

- After each test iteration, the virtual machines were deleted. Every test iteration started with a newly provisioned machine. This was done to enable the possibility of switching physical machines throughout the process of testing, which would increase the relevancy to any potential Azure users and the effectiveness of the sample size.

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

16

- Once a new VM was created for each new iteration of testing, "apt-get update" and "apt-get upgrade –yes" commands were issued, followed by a reboot of the system before installing any tests.

- A cloud service is provisioned with each new instance unless a specific cloud service is assigned. By default, a limit of 20 cloud services can be active at once. Cloud services do not get deleted along with the VM, so cloud services had to be manually deleted.

- The only additional installation was libaio, because Cloud Spectator originally planned on testing asynchronous disk I/O. In the end, Cloud Spectator conducted synchronous testing. The libaio package should not affect the results.

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

17

7 About Cloud Spectator and Anne Liu

Cloud Spectator is a cloud analyst agency focused on cloud Infrastructure-as-a-Service (IaaS) performance. The company actively monitors 20+ 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 businesses 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.

Anne Liu is an analyst at Cloud Spectator, where she specializes in researching the emerging cloud-based IaaS market. Her primary area of focus is in the public cloud market. Within that sector, Liu publishes information and trends on performance and pricing with a standardized methodology for comparison. This is done across a number of public cloud providers to help guide enterprises in selecting the right provider and environments for their cloud strategy.

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

18

8 About Gigaom Research

Gigaom Research gives you insider access to expert industry insights on emerging markets. Focused on delivering highly relevant and timely research to the people who need it most, our analysis, reports, and original research come from the most respected voices in the industry. Whether you're beginning to learn about a new market or are an industry insider, Gigaom Research addresses the need for relevant, illuminating insights into the industry's most dynamic markets.

Visit us at: research.gigaom.com.

Generational Performance Comparison: Microsoft Azure's A-Series and D-Series
This PDF prepared for: Danny Gee (Dgee@cloudspectator.com)

19