

SPEC

SPECstorage™ Solution 2020_ai_image Result

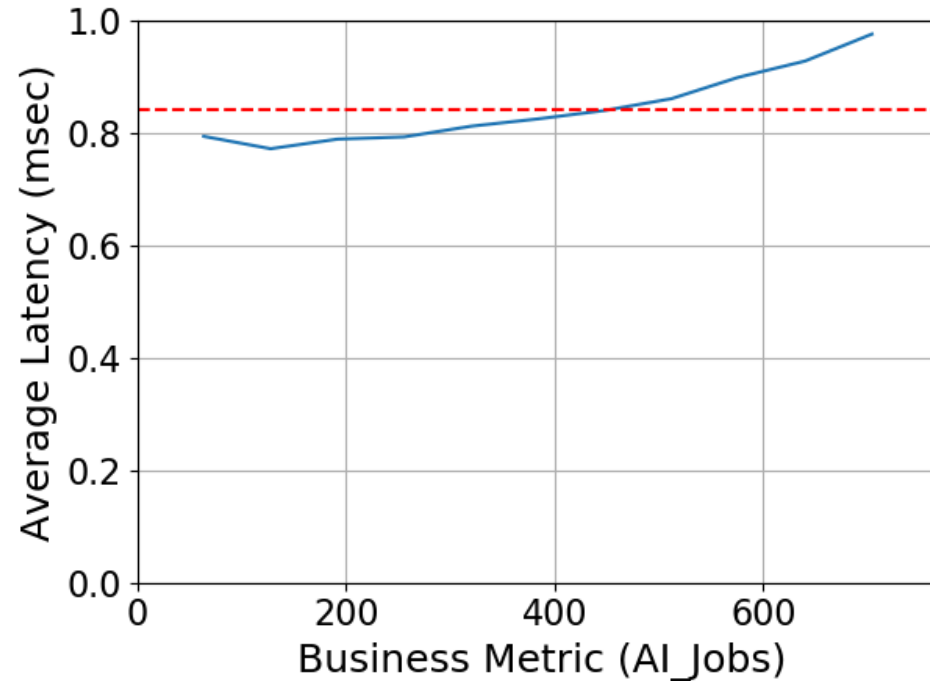
Copyright © 2016-2024 Standard Performance Evaluation Corporation

Qumulo, Inc.
Azure Native Qumulo - Public Cloud Reference

SPECstorage Solution 2020_ai_image = 704 AI_Jobs
Overall Response Time = 0.84 msec

Performance

Business Metric (AI_Jobs)	Average Latency (msec)	AI_Jobs Ops/Sec	AI_Jobs MB/Sec
64	0.794	27843	6261
128	0.772	55687	12519
192	0.789	83530	18778
256	0.793	111374	25039
320	0.812	139218	31294
384	0.825	167061	37552
448	0.840	194905	43811
512	0.861	222748	50071
576	0.899	250592	56332
640	0.928	278436	62591
704	0.976	306279	68849



Product and Test Information

Azure Native Qumulo - Public Cloud Reference	
Tested by	Qumulo, Inc.
Hardware Available	May 2024
Software Available	May 2024
Date Tested	May 11th, 2024
License Number	6738
Licensee Locations	Seattle, WA USA

None

Solution Under Test Bill of Materials

Item No	Qty	Type	Vendor	Model/Name	Description
1	33	Microsoft Azure Virtual Machine	Microsoft	Standard_E32d_v5	The Standard_E32d_v5 Azure Virtual Machine runs on the 3rd Generation Intel® Xeon® 8370C processor. Each VM features 32 vCPU cores, 256 GiB of memory, 8 NICs, and supports up to 16000 Mbps of network egress bandwidth. Network ingress is not metered in Azure, enabling exceptional throughput results of 68GB/s with an overall response time latency of 0.84ms. Thirty-two (32) VMs served as Spec netmist load generators, each with two dedicated ANQ mount points. One (1) VM was dedicated to the Prime Client role. Each VM utilized a single NIC with Advanced Networking enabled.

Configuration Diagrams

- [Azure Native Qumulo](#)

Component Software

Item No	Component	Type	Name and Version	Description
1	Azure Native Qumulo filesystem	File System	7.1.0	The Azure Native Qumulo (v2) File System is a cloud service provisioned directly from the Azure portal. This architecture differs from previous releases of Qumulo Core Filesystems. Customers specify the initial capacity and a customer subnet delegated to Qumulo.Storage/FileSystems. Qumulo uses vNet

				injection, resulting in IP addresses on customer networks. The storage service components are opaque to the customer. In ANQ, performance and capacity are completely disaggregated. The service offers non-disruptive and dynamic performance increases on-demand, currently requested through service requests. For this benchmark, an increase in aggregate performance capability up to 100GB/s of throughput and 500,000 total IOPS was requested.
2	Ubuntu	Operating System	22.04	The Ubuntu 22.04 operating system was deployed on all Standard_E32d_v5 VMs.

Hardware Configuration and Tuning - Virtual

Component Name		
Parameter Name	Value	Description
Accelerated Networking	Enabled	Microsoft Azure Accelerated Networking enables single root I/O virtualization (SR-IOV)

Hardware Configuration and Tuning Notes

N/A

Software Configuration and Tuning - Virtual

Component Name		
Parameter Name	Value	Description
mount options	64	Each NFSv3 client mounted (2) distinct IP addresss

Software Configuration and Tuning Notes

The single filesystem was attached via two (2) NFS version 3 mounts per client. The mount string used: mount -t nfs -o tcp,vers=3 10.0.0.4:/AI_IMAGE /mnt/AI_IMAGE-0 mount -t nfs -o tcp,vers=3 10.0.0.5:/AI_IMAGE /mnt/AI_IMAGE-1 The mount specifies that two (2) NFS version 3 exports were mounted on each client.

Service SLA Notes

Azure Service SLA: <https://www.microsoft.com/licensing/docs/view/Service-Level-Agreements-SLA-for-Online-Services>

Storage and Filesystems

Item No	Description	Data Protection	Stable Storage	Qty
1	Azure Native Qumulo persists data entirely using Azure Blob Storage, which is a Microsoft-managed service providing cloud storage that is highly available, secure, durable, scalable, and redundant. Architecturally, Read cache is serviced from an in-memory L1 cache and a NVMe based L2 cache. The global read-cache is increased on-demand with a customer service request. Write transactions leverage high performance Premium SSD v2 disks, which act as a protected write-back cache for incoming writes, continuously flushing data to Azure Blob Storage.	Provided by Microsoft Azure Blob Storage	Azure Blob Storage	64

Number of Filesystems	1
Total Capacity	100 TiB
Filesystem Type	Azure Native Qumulo

Filesystem Creation Notes

Azure Native Qumulo is deployed directly from the Azure Portal, the Azure CLI, or using an Azure REST endpoint. Customers specify the capacity of storage and what subnet they want to deploy vNet-injected IP addresses in.

Storage and Filesystem Notes

Azure Native Qumulo is a cloud storage service with a different architecture than traditional storage, including the traditional Qumulo Core software. Data is persisted to Azure Blob, with caching occurring only at the compute layer. This disaggregates storage capacity/persistence from compute/performance into two completely separate and independently elastic layers. The Azure Native Qumulo filesystem persists data to Azure Blob using Locally Redundant Storage (LRS). Azure Blob LRS stores data synchronously across three physical locations within an Azure region. This architecture acts as an accelerator, executing parallelized reads that are prefetched from object storage and served directly from the filesystem's L1/L2 cache to the clients (over NFSv3 in this benchmark).

Transport Configuration - Virtual

Item No	Transport Type	Number of Ports Used	Notes
1	Ethernet	97	There were ninety-seven (97) vNics used in this benchmark, with thirty-two (32) used as the load generating Ubuntu NFS clients, sixty-four (64) used as vNet injected IP addresses from the ANQ storage service, and one (1) by the Prime Client. Each VM utilized a single NIC with Advanced Networking enabled.
2	ethernet	97	Azure networking allows up to 8 virtual NICs for the Standard_E32d_v5 with a combined aggregate EGRESS limit of 16,000 Gbps. In this benchmark only a single vNic was used, which still has the same EGRESS limit of 16,000 Gbps. There is not a INGRESS limit in Azure.

Transport Configuration Notes

The NFSv3 protocol was used with the mount options tcp and vers=3. There was no use of nconnect in this benchmark, as it does not result in a performance increase with this workload. IP addresses are distributed to clients using a round-robin DNS record, as shown in the architectural diagram. Each client in this benchmark used two of the ANQ vNet-injected IP addresses per NFSv3 client, resulting in two distinct NFSv3 mounts on each load-generating client.

Switches - Virtual

Item No	Switch Name	Switch Type	Total Port Count	Used Port Count	Notes
1	Azure	12.5 Gbps Ethernet with Accelerated Networking	64	64	Used by ANQ Storage Service
2	Azure	16.0 Gbps Ethernet with Accelerated Networking	33	33	Used by NFS Front-end Clients

Processing Elements - Virtual

Item No	Qty	Type	Location	Description	Processing Function
1	1	vStorage	Azure Region EastAsia	Azure Native Qumulo Cloud Storage Service	Qumulo File System, Network communication, Storage Functions

Processing Element Notes

None

Memory - Virtual

Description	Size in GiB	Number of Instances	Nonvolatile	Total GiB
Azure Standard_E32d_v5 VM memory	256	33	V	8448
Grand Total Memory Gibibytes				8448

Memory Notes

Each Standard_E32d_v5 VM has 256GiB of RAM

Stable Storage

Azure Blob Storage provides the stable storage function in the ANQ architecture. It is a cloud object storage solution from Microsoft Azure, designed to store large amounts of unstructured data and optimized for efficient storage and retrieval of large data objects. This solution offers virtually unlimited scalability, durability, and high availability to meet ANQ's stable storage needs. From a data protection perspective, Azure Blob Storage uses HMAC to ensure data integrity. When data is uploaded, a hash is computed and stored. Upon retrieval, the hash is recomputed and compared to verify data integrity. ANQ employs Azure Blob Storage's Locally Redundant Storage (LRS) offering, which keeps multiple copies of data within a single region.

Solution Under Test Configuration Notes

The solution under test was a standard Azure Native Qumulo cluster deployed natively in Azure. The cluster is elastic both in performance and capacity and can be requested to handle 100's of GB/s of throughput, millions of IOPS, and 100's of PB's of capacity. Performance and capacity are 100% disaggregated in ANQ.

Other Solution Notes

For this benchmark, the aggregate performance capability of 100GB/s and 500,000 IOPS was requested for the ANQ. Performance capability can be changed up or down by submitting a service ticket, at this time.

Dataflow

None

Other Notes

None

Other Report Notes

None

Generated on Sat Jun 8 03:12:36 2024 by SpecReport
Copyright © 2016-2024 Standard Performance Evaluation Corporation