#dcsarticle http://www.dcseurope.info/n/ngkv

SNIA™ Cloud Storage Initiative

SNIA™ Solid State Storage Initiative

www.snia-europe.org

A different approach to solid state storage in the cloud

By Mark Carlson, Vice Chair SNIA Technical Council, Toshiba.

WITH THE INCREASING use of the Cloud comes the desire to improve the performance of cloud hosted operations. In response, several cloud providers have started to offer solid state storage for cloud computing workloads. This article provides an overview of these offerings and compares the different approaches. Also discussed are the caveats that might be of concern to specific customers.

How is solid state being used?

Cloud Storage

The typical cloud storage offering is accessed over the Internet for use cases such as Sync and Share, Backup, and Media archive. As such, the latency of access to the data stored there is largely determined by the latency of the network connection to the offering. Solid state storage could provide faster response times to requests, but this improvement is negated by the overall latency of the network connection. Thus the cloud provider is hard pressed to pass on the additional expense of solid state to customers with a beneficial use case.

In terms of throughput, solid state can handle a larger number of simultaneous accesses such as in a media server use case. This can be justified in a Content Distribution Network use case as well and indeed several offerings are available that use solid state[1] storage. This is shown in Figure 1. While this usage may provide a competitive advantage to the provider, the user is not typically able to specify the type of storage used.

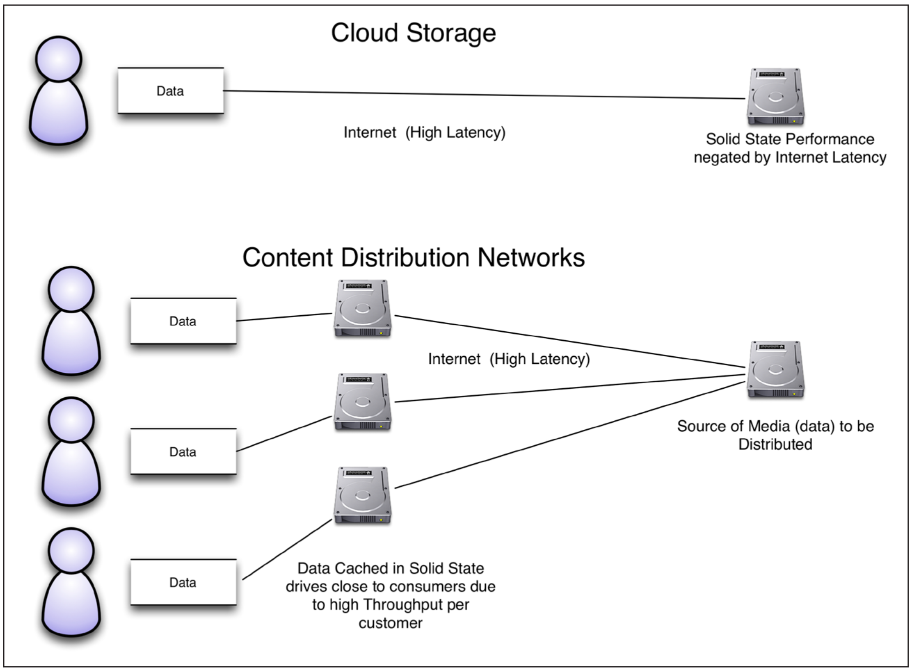

Figure 1: Use of Solid State in Cloud Storage

Cloud computing

Cloud Computing hosts your software on a machine or virtual machine (VM) in the provider's data center. These computing resources are in the same data center as the storage resources and are connected in some cases by low latency networking as shown in Figure 2. As a result Solid State storage can accelerate the performance of these compute tasks.

Several Infrastructure-as-a-Service (IaaS) cloud computing providers now offer solid state storage as part of their offering. In IaaS offerings, storage typically takes two forms: the ephemeral storage associated with a virtual machine instance (disappears when the VM disappears), and the block storage that persists beyond any one virtual machine's lifecycle. Either or both

Data system metadata item	Description
cdmi_latency	If this data system metadata item is present and set to a positive numeric string, it indicates that the client is requesting a desired maximum time to first byte, in milliseconds. This metadata is the desired latency (in milliseconds) to the first byte of data, as measured from the edge of the cloud and factoring out any propagation latency between the client and the cloud. For example, this metadata may be used to determine, in an interoperable way, from what type of storage medium the data may be served. When this data system metadata item is absent, or is present and is not set to a positive numeric string, this data system metadata item shall not be used.
cdmi_throughput	If this data system metadata item is present and set to a positive numeric string, it indicates that the client is requesting a desired maximum data rate on retrieve, in bytes per second. This metadata is the desired bandwidth to the data, as measured from the edge of the cloud and factoring out any bandwidth capability between the client and the cloud. This metadata is used to stage the data in locations where there is sufficient bandwidth to accommodate a maximum usage. When this data system metadata item is absent, or is present and is not set to a positive numeric string, this data system metadata item shall not be used.

Table 1: CDMI Data System Metadata from CDMI Table 119 in section 16.4

of these types can be based on solid state and the provider typically will pass on the additional cost of solid state in these offerings.

Standards
The SNIA has created a standard interface for managing data in the cloud, called the Cloud Data Management Interface (CDMI[2]), now an ISO standard. CDMI specifies the management of the requirements for data in a cloud environment. The data requirements are expressed by specifying Data System Metadata for individual data elements (such as Objects) or containers of data (such as a block device or file system). When used with cloud computing, CDMI can be used to specify requirements on block storage devices such as SSDs that are attached to a VM. The basic requirements that are defined include latency and throughput as shown in table 1.

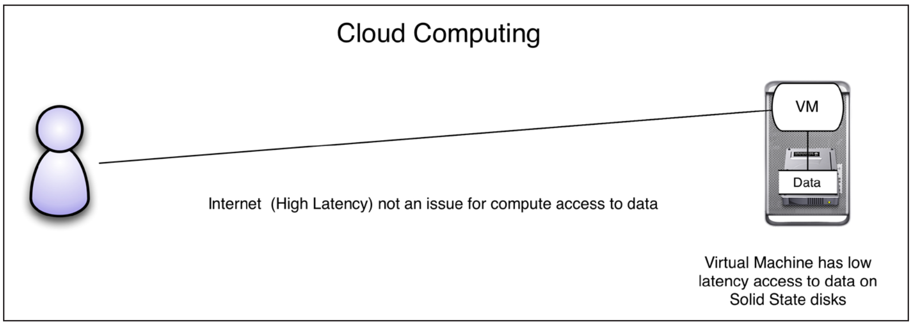

Figure 2: Use of Solid State in Cloud Computing

	Maximum IOPS / 100 GB (scales linearly up to 10 TB)	Maximum Sustained throughput / 100 GB	Maximum Sustained throughput / VM
Read	30 IOPS	12 MB/s	180 MB/s
Write	150 IOPS	9 MB/s	120 MB/s

	Expected IOPS / 100 GB	Expected throughput / 100 GB	Maximum Sustained throughput / VM
Read	3000 IOPS	48 MB/s	240 MB/s
Write	3000 IOPS	48 MB/s	240 MB/s

Table 3: For an SSD PD of 100GB they show the above

These parameters are 'hints' to the underlying infrastructure that may then be matched to different storage media types such as solid state, traditional hard drives, or even tape devices. The hints may also factor into what type of connection is used between the CPU and storage such as PCIe, NVMe, SAS, etc. CDMI also allows for the instrumentation of the actual provided values to be exposed through metadata as well. In CDMI table 120, cdmi_latency_provided contains the provided maximum time to first byte, and cdmi_throughput_provided contains the provided maximum data rate on retrieve. In a cloud environment, the billing would reflect the actual provided metrics.

Amazon and Google examples

Google Compute Engine Disks
Google, as part of their Google Compute Engine IaaS offering describes two types of storage:

- Standard Persistent Disk (PD) (storage backed by hard disk drives)
- SSD PD (storage backed by solid state drives)

Performance for Persistent Disk increases with volume size up to the per-VM maximums. While performance increases with volume size for both Standard and SSD PD, the performance increases much more quickly for SSD PD. Table 3 directly show the expected performance increase for VMs that need the additional IOPS and throughput. For pricing, Google lists Standard PDs priced at $0.04 per Gigabyte (GB) and SSD PDs priced at $0.17 per GB. A better way to think about this, however, is cost of the IOPs. Standard PD is approximately $0.133 per random read IOPS and $0.0266 per random write IOPS. SSD PD is $0.0057 per random read IOPS and $0.0057 per random write IOPS. So if the requirement for IOPS is key for the application, SSDs are clearly a win.

Google also offers an ephemeral solid state storage, called 'Local SSDs' which have the same lifecycle as the VM it is attached to. Local SSDs are 375 GB per device, and only four local SSD devices are allowed per virtual machine. Local SSDs cannot be used as root disks. Google offers two types of interfaces to Local SSD and lists the following performance as shown in table 4 below.

Amazon elastic compute

Amazon offers solid state storage in a similar manner to that of Google. The Elastic Block Storage (EBS) General Purpose Volumes are allocated by the GB with no additional charge for I/O. EBS Provisioned IOPS (SSD) Volumes are provisioned with a specific IOPS value that is charged for alongside the charge for the storage space. Amazon also has EBS Magnetic Volumes that do not use solid state.

Amazon's ephemeral storage (called Instance Store) attached to a VM can also be based on solid state. Here is a table with the various instance types that include solid state instance store volumes as shown in Table 5 below.

These are just some of the offerings available as examples. Note that some of the solid state devices also support the TRIM command.

Caveats

When using solid state storage in cloud computing, your data writes are not usually overwriting the data that was in the same storage location. SSDs have a Flash Translation Layer (FTL) that balances writes across flash memory locations since the media wears out over time. Thus, your data may still be part of the drive even though it has been returned to the general pool. To get around this, some providers encrypt your data at 'rest' using their own encryption keys. Some providers (i.e. Google) will even let you provide your own keys for this purpose. If this is a concern, you should look for these specific features in your choice of cloud provider.

Future

Some of the latest solid state advances are appearing for systems developers, but have not made their way into cloud offerings as yet. Persistent Memory, for example, is accessed at or near DRAM speeds and has address and data lines for an interface. There are operating system and hypervisor changes needed to take advantage of these new technologies, but it is only a matter of time before support is there and cloud computing providers start offering virtual machines with provisioned persistent memory. While applications will see increased performance as a result, optimal performance comes with some pain caused by the new model's changes to libraries and language compilers in some instances.

The SNIA has specified the programming model for Non-Volatile Memory (Persistent Memory) in the NVM Programming Model[3]. The NVM Programming Model was developed to address the ongoing proliferation of new non-volatile memory (NVM) functionality and new NVM technologies. An extensible NVM Programming Model is necessary to enable an industry wide community of NVM producers and consumers to move forward together through a number of significant storage and memory system architecture changes.

This model defines recommended behavior between various user space and operating system (OS) kernel components supporting NVM. This model does not describe a specific API. Instead, the intent is to enable common NVM behavior to be exposed by multiple operating system specific interfaces. After establishing context, the model describes several operational modes of NVM access. Each mode is described in terms of use cases, actions and attributes that inform user and kernel space components of functionality that is provided by a given compliant implementation.

For more information about the work SNIA is doing around solid state www.snia.org/forums/sssi, and for more information about the work SNIA is doing around Cloud visit www.snia.org/forums/csi

Type	Maximum sustained read IOPS/device* IOPS / 100 GB	First write IOPS/device*
SCSI	100,000	70,000
NVMe	170,000	90,000

Table 4: Google offers 2 types of interfaces to Local SSD and lists the performances above

Instance Type	Instance Store Volumes
c3.large c3.8xlargec3.2xlarge	2 x 16 GB SSD (32 GB) 2 x 320 GB SSD (640 GB)2 x 80 GB SSD (160 GB)
cr1.8xlargecc2.8xlargec3.4xlarge	2 x 120 GB SSD (240 GB)4 x 840 GB (3360 GB) 2 x 160 GB SSD (320 GB)
i2.8xlargei2.2xlargehs1.8xlarged2.8xlarge	8 x 800 GB SSD (6400 GB), supports TRIM*2 x 800 GB SSD (1600 GB), supports TRIM*24 x 2000 GB (48 TB) 24 x 2000 GB (48 TB)
m1.smalli2.4xlargei2.xlargeg2.2xlarge	1 x 160 GB 4 x 800 GB SSD (3200 GB), supports TRIM*1 x 800 GB SSD, supports TRIM*1 x 60 GB SSD
r3.8xlarge	2 X 320 GB SSD (640 GB), supports TRIM*

Table 5: This table show various Amazon instance types that include solid state instance store volumes.

References

1. Fastly Blog Post: 'How our Solid State Drives Result in Cost Savings for Customers' https://www.fastly.com/blog/how-our-solid-state-drives-result-in-cost-savings-for-customers

2. Cloud Data Management Interface: ISO/IEC 17826:2012 - http://www.iso.org/iso/catalogue_detail.htm?csnumber=60617

3. NVM Programming Model v1.1 http://www.snia.org/sites/default/files/NVMProgrammingModel_v1.1.pdf