Data Storage Infrastructure – Part 3

This lesson continues the discussion on data storage infrastructure.

We'll cover the following

Hard disk drive (HDD) and solid-state drive (SDD)

A *hard disk drive* is an electromechanical storage device that has a rotating disk coated with magnetic material. On the other hand, a *solid-state drive* uses an _ Integrated Circuit (IC)_ with flash memory to store data. It doesn't have a physical spinning disk. This makes *SSDs* more resistant to physical shocks. Also, the use of flash memory cuts down the latency significantly.

When it comes to archival storage *HDDs* are the preferred choice since semiconductor-based *solid-state drives* tend to lose data over time if left for long periods without power. *HDDs* are also more affordable in comparison to *SSDs*. A large amount of data can be stored with a lot less expense with *HDDs*, as opposed to when using *SSDs*. *SSDs* are typically used to implement a distributed cache layer where low latency is vital.

Returning to persistent disks, they are measured by a metric known as *IOPS*, which is *input-output operations per second*. So, if we need high *IOPS*, we would pick an *SSD* because they provide latencies in as low as single-digit milliseconds.

If we need to process a large amount of workload data having sequential read-write operations, we would pick a standard *HDD* persistent disk.

If you are wondering what a sequential read-write operation is, read about the difference between sequential and random write on StackOverflow.

Transient disks

Transient disks are *SSDs* that are attached to the compute, and they store data of a

specific instance as long as it is not terminated. If the compute instance is

decommissioned, all the data is wiped off the transient SSD disk. Also, the data on a transient disk is non-redundant.

A persistent disk may be deployed in isolation decoupled from *VMs*, but a local transient *SSD* is generally co-located with the *VMs*.

Why do we need transient disks?

Transient SSD disks provide lower latency than persistent *SSDs*. So, if you need to perform a very high *IOPS* operation with low latency, a transient *SSD* would be a good option.

In the next lesson, let's discuss the different types of data storage. But, first, a quiz.