Direct-Attached Storage

Dr Peadar Grant

January 31, 2024

Contents

1 Disk service time

The disk service time, T_S is the time taken by a disk to complete an I/O request, composed of:

1. the seek time, T
2. average rotational latency, L
3. data transfer time, X

where:

$$T_S = T + L + X \tag{1}$$

1.1 Seek time

Seek time is the time required to position the head on the correct track. Obviously this isn't uniform, so seek time is given separately:

Full stroke: time taken to move from innermost to outermost track, T_F.

Track-to-track: time taken to move between adjacent tracks, T_T.

Average: time taken to move head from one random track to another.

We are normally concerned with the average seek time, defined as one-third of the full stroke:

$$\text{average seek time} = \text{full-stroke seek time}/3 \tag{2}$$
$$\text{full-stroke seek time} = \text{average seek time} * 3 \tag{3}$$

Typical average seek times would range from 5 ms to 20 ms.

Example 1: Seek time

Calculate the full-stroke seek time for a drive given an average seek time of 6ms.

$$\text{full-stroke seek time} = 6\,\text{ms} \times 3 \tag{4}$$
$$= 18\,\text{ms} \tag{5}$$

1.2 Rotational latency

The average rotational latency, L, is the time taken for the drive to revolve half a revolution:

$$L = 0.5/\text{revolutions per second} \tag{6}$$

This measure depends on drive speed, we must convert RPM to revolutions per second.

Example 2: Rotational latency

Determine the average rotational latency for a 5400-rpm drive:

```
5400 RPM = 5400 / 60 RPs
         = 90 RPs
       L = 0.5 / 90
         = 5.5 ms
```

1.3 Internal transfer time (X)

The data transfer time is how long it takes for one block of data (at a given size) to be transferred inside the drive.

$$X = \text{block size}/\text{internal transfer rate} \tag{7}$$

Example 3: Internal transfer time

Determine the transfer time given an internal transfer rate of 40MB/s and a block size of 32kB.

```
X = 32kB / 40 MB/s
  = ( 32 * 1024 ) / ( 40 * 1024 * 1024 )
  = 0.78 ms
```

1.4 IOPS

Storage performance is commonly quantified in Input/Output operations Per Second (or IOPS), which is the reciprocal of the disk service time, TS.

$$\text{IOPS} = 1.0/T_S \tag{8}$$

2 Native command queueing

- A hard disk receives multiple commands in quick succession. Each command will be delayed by seek time and rotational latency.
- SATA Native Command Queueing tries to optimise the overall latency by re-ordering the commands to reduce these latencies.
- All NCQ algorithms will optimise the seek time, but some will also optimise the rotational latency.

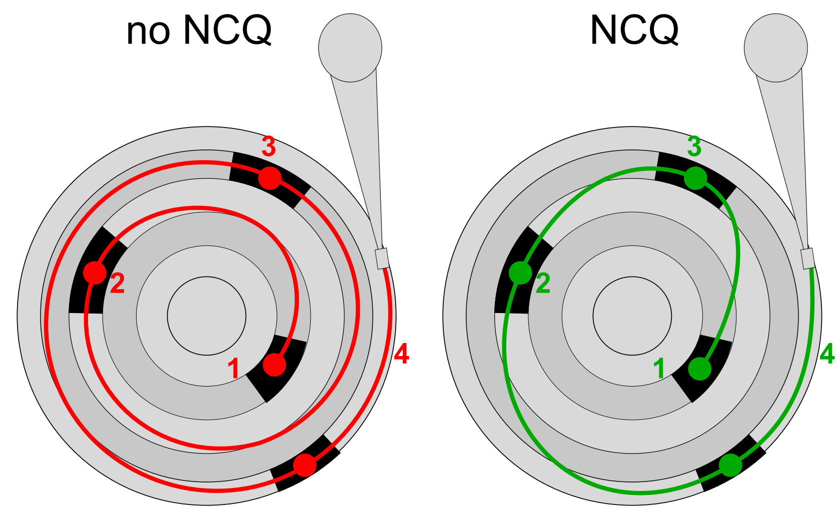

Figure 1: Native Command Queueing

3 Direct-Attached Storage

Direct Attached Storage is where a storage device connects 1:1 with host without a network:

Host is name given to computer system using storage. (Laptop PC, server or mainframe.)

- Host normally has an internal bus. (PC-based systems: PCI bus).

Storage device: magnetic disk, solid state drive, tape drive, multi-host storage appliance.

Host Bus Adapter (HBA) connects host to storage device. The HBA . . .

- . . . needs driver support from the host operating system, although many use one of a few generic drivers.
- . . . must match interface on the disk side AND the host's expansion bus (PCI, ISA, others).
- . . . nowadays often integrated on the host's motherboard.

Interface defines electrical and communication characteristics between the HBA and the storage device.

- Common interfaces ATA/IDE/PATA, SATA, SCSI and SAS.

4 Interface types

ATA/IDE/PATA

Serial ATA (SATA)

Small Computer Systems Interface (SCSI)

Serially Attached SCSI (SAS)

4.1 Operating system

- Host operating system will see so-called block devices representing each individual disk attached to the HBA.
- Block devices can then be partitioned or otherwise spatially managed by the host operating system.
- Operating system normally inserts another layer of abstraction — the filesystem.

5 Multi-device DAS

5.1 Port multipliers

Port multipliers can allow a single DAS port to connect to more than one device:

- HBA must support port multiplier usage
- Port multiplier is transparent to individual disks
- Bandwidth is shared

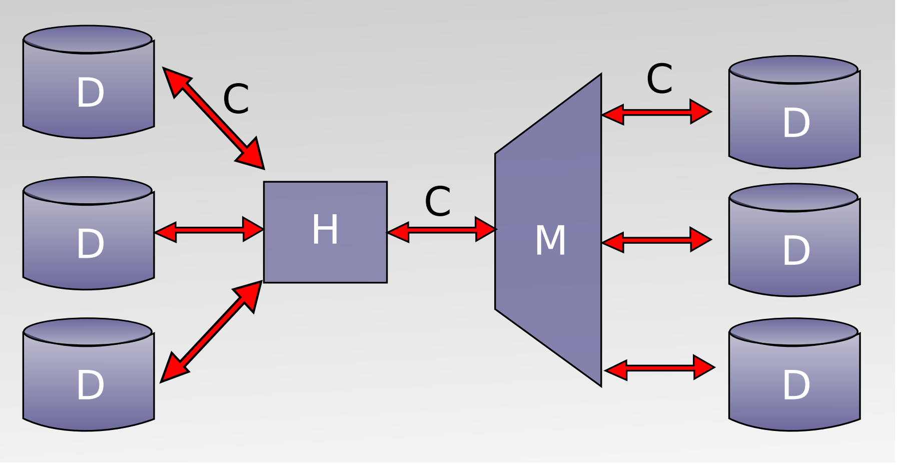

Figure 2: SATA Multipler

5.2 Storage appliances

A storage appliance with multiple disks can connect these to 1 or more hosts in DAS-type configuration (usually SAS).

6 Controller utilisation

The utilisation, U, of a disk controller ranges from 0 to 1:

$$0 <= U <= 1 \tag{9}$$

Although we commonly talk about utilisation as a percentage, we must convert it to a decimal fraction when working out calculations that involve utilisation. For example, 30% utilisation would mean that `U=0.3`.

6.1 Average response time

The average response time, R, depends on the disk service time, taking into account the controller utilisation:

$$R = Ts/(1 - U) \tag{10}$$

Example 4: Response time under load

The disk service time of a disk under no load is 7.8ms. What would the response time expected under 70% load be?

$$R = 7.8\,\text{ms}/(1 - 0.7) \tag{11}$$

$$= 26\,\text{ms} \tag{12}$$

6.2 IOPS vs utilisation

A drive will be advertised as capable of doing a certain number of IOPS. However, if we want to keep response times within a certain limit, we may limit the number of IOPS to a certain number.

$$\text{IOPS limit} = \text{U desired} \times \text{IOPS advertised} \tag{13}$$

Example 5: IOPS under load

A drive is advertised as being capable of 180 IOPS. Determine the maximum permissible IOPS if load is limited to 70%.

```
IOPS @ 70%  load = 180 * 0.7
                 = 126 IOPS
```

6.3 Multiple disks

Sometimes one disk cannot meet the application requirements on its own. For a given application, the number of disks required, DR, will be determined by two other quantities:

- Number of disks required to meet the capacity, DC.
- Number of disks required to meet the application IOPS requirement, DI, at a given utilisation U.

The higher of these two quantities determine DR:

```
DR = max ( DC, DI)
```

Note that this requirement does not specify at what level the disks are combined.

6.4 Meeting IOPS requirement

A common situation is where a given application seeks to guarantee a minimum number of IOPS with a given controller utilisation time. This can be done by:

1. Use the controller utilisation to determine the number of available IOPS.
2. Divide the required IOPs among the number of available IOPS to get the minmum number of disks.

Example 6: Disk required for IOPS at utilisation

given application has a requirement of 3600 IOPS. The disks available have a maximum of 180 IOPS. Determine how many disks are required assuming a maximum utilisation of 80%.

```
available IOPS = 180 * 0.8
               = 144
disks required = ceil( 3600 / 144 )
               = 25
```