Register
Cache
RAM
HDD

Storage Architecture

Chapter 9

Storage Architecture

- ★ Storage Technology
  - ○ Protocol vs. Storage
- ★ Storage Virtualization
- ★ RAID
- ★ Network Storage
  - ○ Network Attached Storage (NAS)
    - ■ NFS
    - ■ CIFS (SMB)
  - ○ Storage Area Network (SAN)
    - ■ iSCSI

Storage Technology

★ Storage Technology
  ○ SSD (Solid State Drive) - (Flash) Memory
 ○ HDD (Hard Disk Drive) - Mechanical Disk

★ Interface Technology
  ○ Serial attached SCSI - Typically on server (SAS-3, 12.0 Gbit/s)
  ○ Serial ATA - (SATA 3.0, 6.0 Gbit/s)
  ○ PCI Express - (PCIe 3.0 ×4, 31.5 Gbit/s)
  ○ Fibre Channel - almost exclusively found on servers (128 Gbit/s)
  ○ USB - (10 Gbit/s)
  ○ Parallel ATA - (UDMA, 1064 Mbit/s)
  ○ (Parallel) SCSI - (40 Mbit/s- 2560 Mbit/s)

Storage Virtualization

★ Storage virtualization is "the process of presenting a logical view of the physical storage resources to" a host computer system, "treating all storage media (hard disk, optical disk, tape, etc.) (in the enterprise) as a single pool of storage.

★ Abstraction of storage -- one (or more) logical storage from several physical storage.

Software.

logical layer.

Physical layer
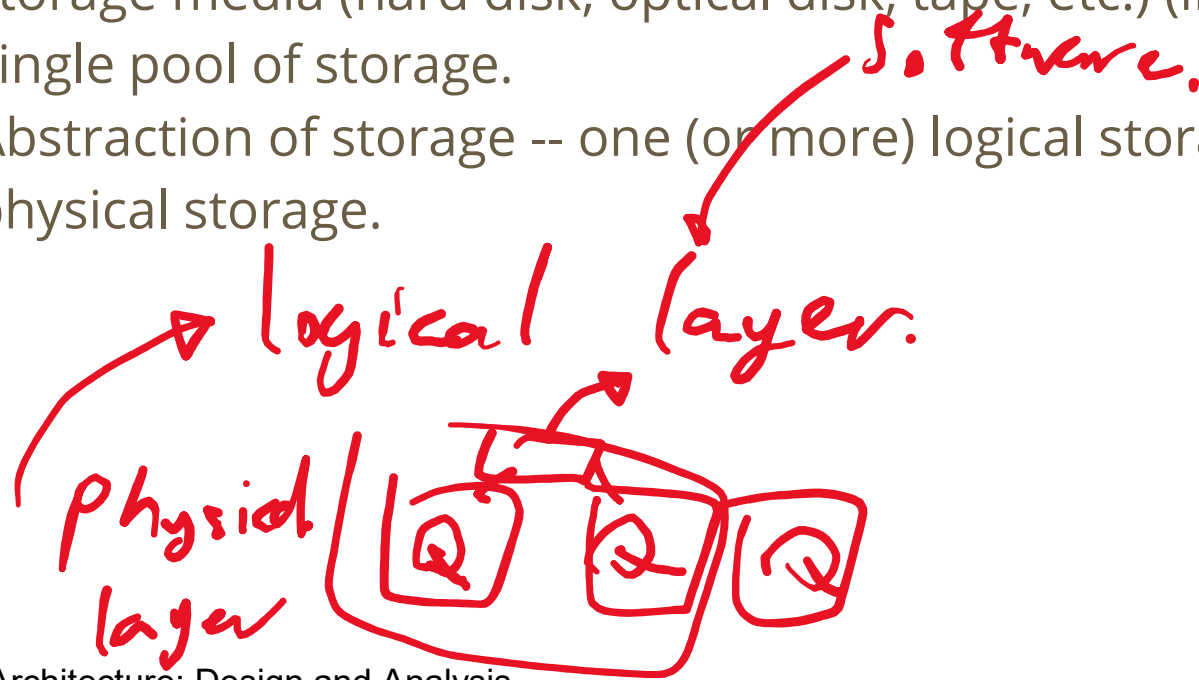
RAID

★ RAID (Redundant Array of Independent Disks) was originally (Redundant Array of Inexpensive Disks)
★ storage virtualization technology
★ "RAID" was invented by David Patterson, Garth A. Gibson, and Randy Katz at the University of California, Berkeley in 1987

David Patterson is the author of several Computer Architecture books. Randy Katz is the author of Contemporary Logic Design used in our Digital Computer Logic class.

RAID
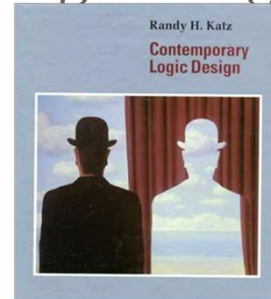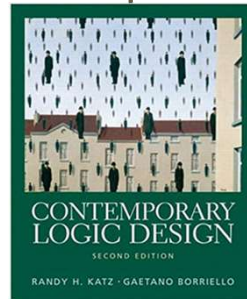
[handwritten: redundancy, performance]
[handwritten: (error correction w/o. backup).]

- ★ In the past (60s, 70s), high performance means expensive disk (e.g. battery-backed RAM disk).
- ★ RAID allows commodity hardware to meet the performance of expensive disk.
- ★ RAID level, provides a different balance among the key goals: reliability, availability, performance, and capacity.
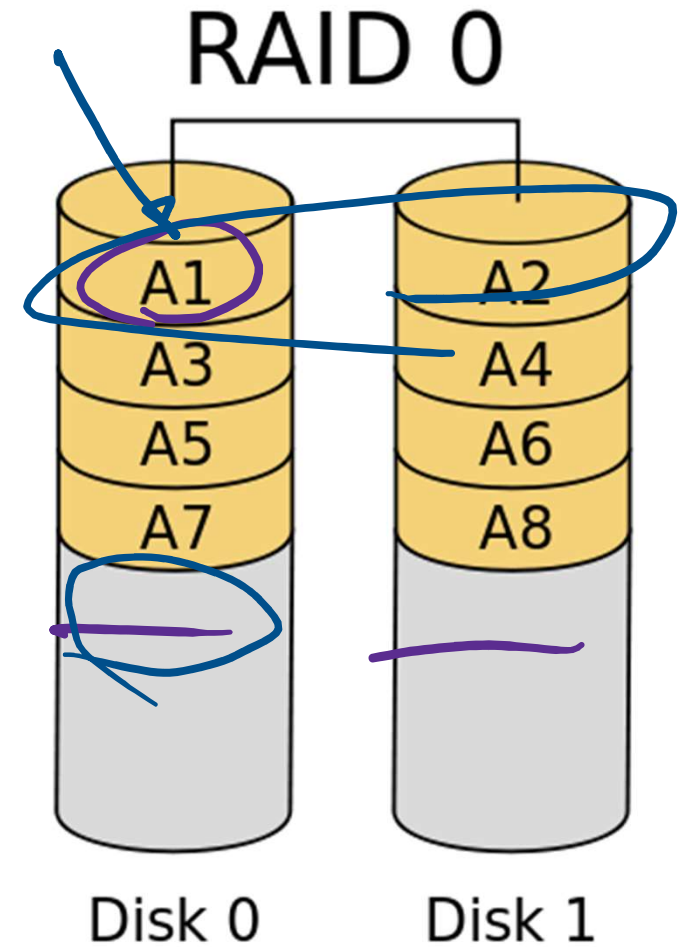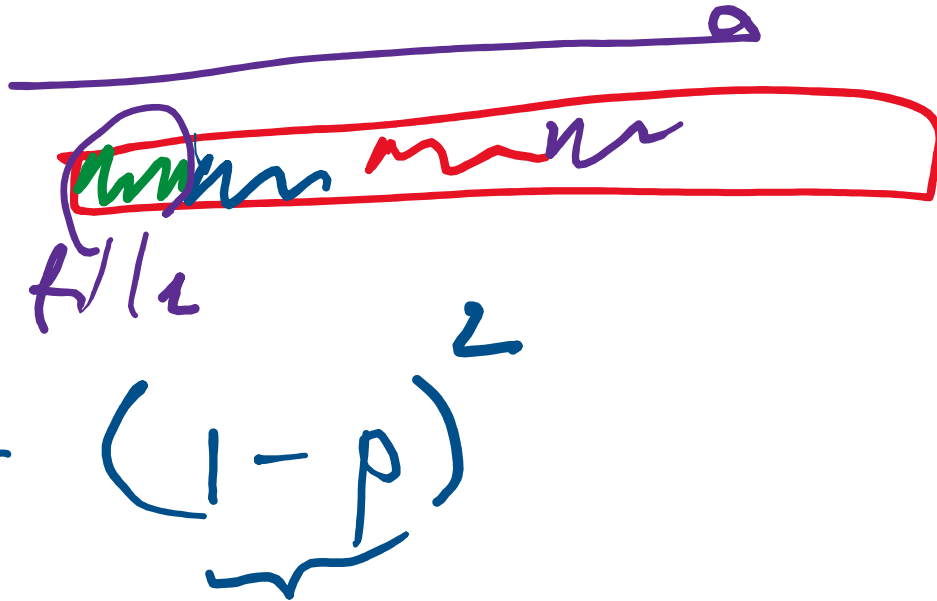
[handwritten: uptime]

RAID 0 (Stripe)

★ Good for high performance, but low reliability

$$1 - (1-p)^2$$

file

RAID 0

A1   A2
A3   A4
A5   A6
A7   A8

Disk 0   Disk 1

Pictures from wikipedia

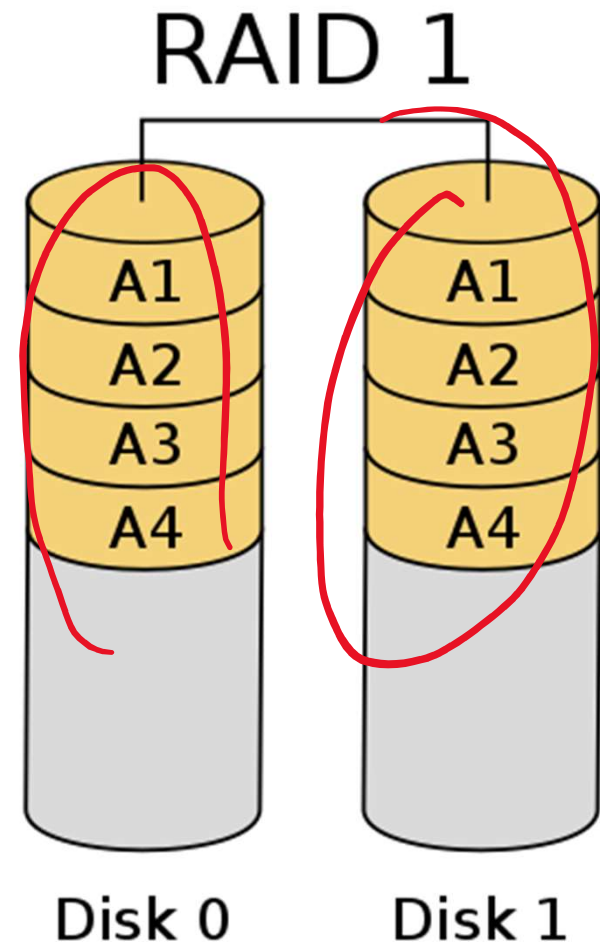Computer Architecture: Design and Analysis   Krerk Piromsopa, Ph.D. @ 2016

RAID 1 (Mirror)

★ Mirror
★ Useful when read and reliability is important
★ Tolerant to failure of a disk

p^2

$2 \to 1$

$2:1$
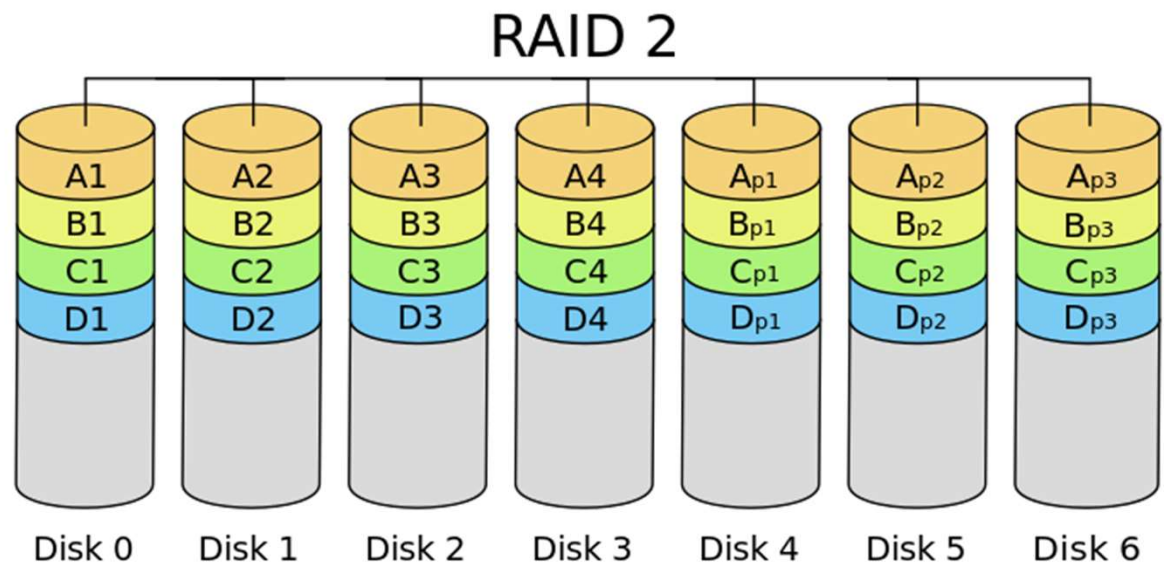
RAID 1

A1   A1
A2   A2
A3   A3
A4   A4

Disk 0   Disk 1

Pictures from wikipedia

RAID 2

★ Rarely used in practice
★ Bit-level parity
★ Use humming code for error correction.
★ Tolerant to failure of a disk

RAID 2

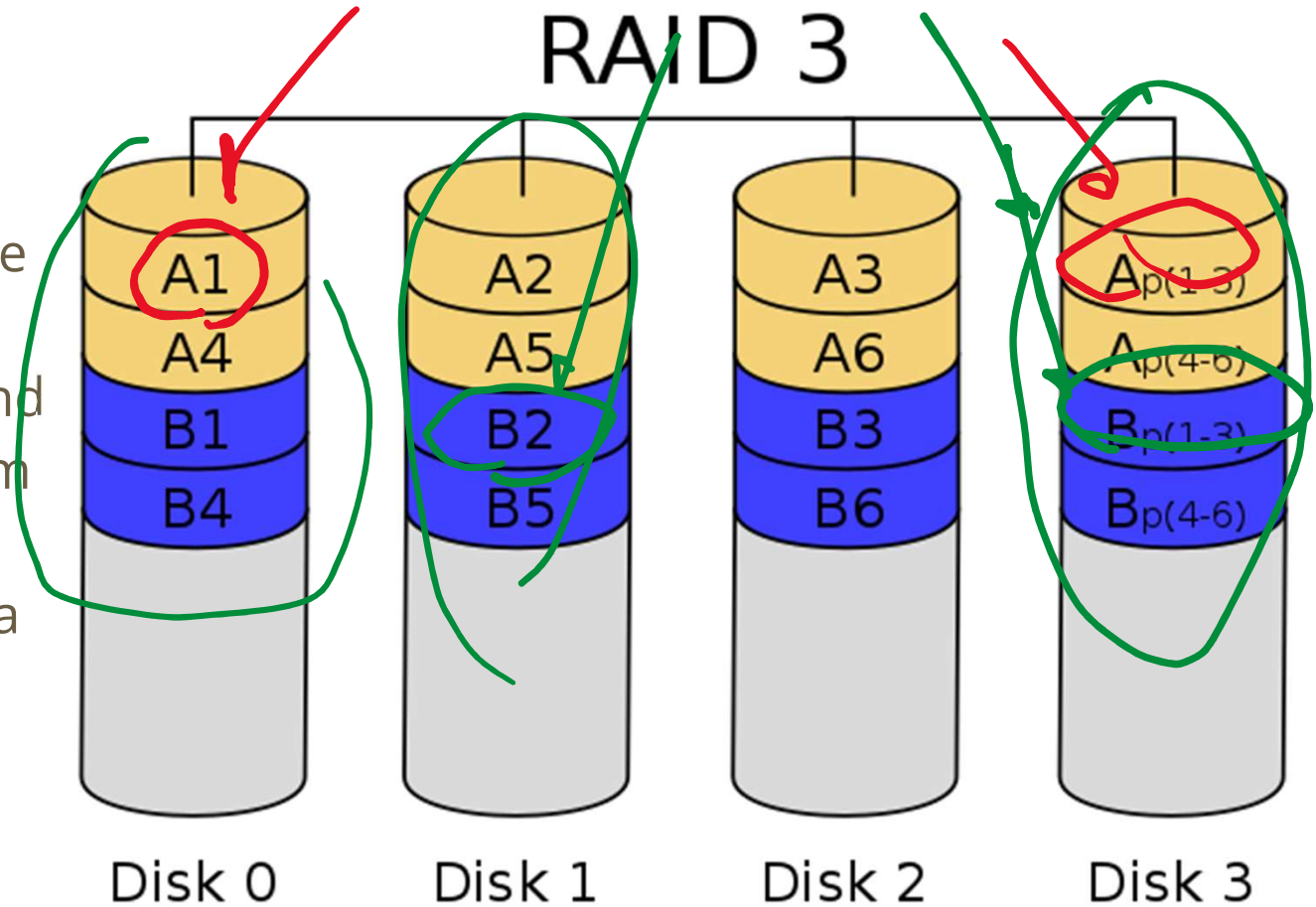| Disk 0	Disk 1	Disk 2	Disk 3	Disk 4	Disk 5	Disk 6
A1	A2	A3	A4	Ap1	Ap2	Ap3
B1	B2	B3	B4	Bp1	Bp2	Bp3
C1	C2	C3	C4	Cp1	Cp2	Cp3
D1	D2	D3	D4	Dp1	Dp2	Dp3

Pictures from wikipedia

RAID 3

- ★ Rarely used in practice
- ★ Byte-level parity
- ★ Good for long read and write. (Bad for random read and write)
- ★ Tolerant to failure of a disk

RAID 3

Disk 0: A1, A4, B1, B4
Disk 1: A2, A5, B2, B5
Disk 2: A3, A6, B3, B6
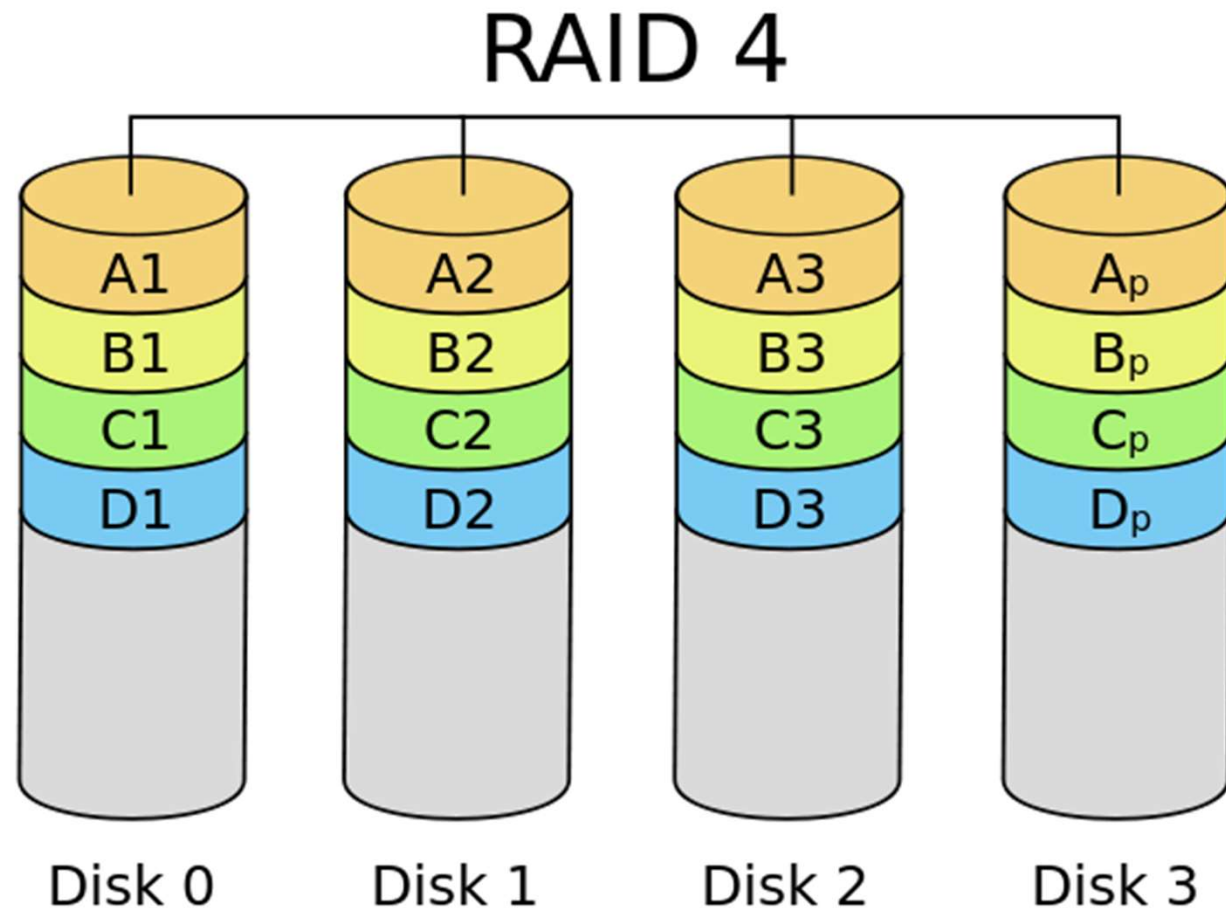Disk 3: $A_{p(1-3)}$, $A_{p(4-6)}$, $B_{p(1-3)}$, $B_{p(4-6)}$

Pictures from wikipedia

RAID 4

- ★ Rarely used in practice
- ★ Block-level parity to a dedicate disk
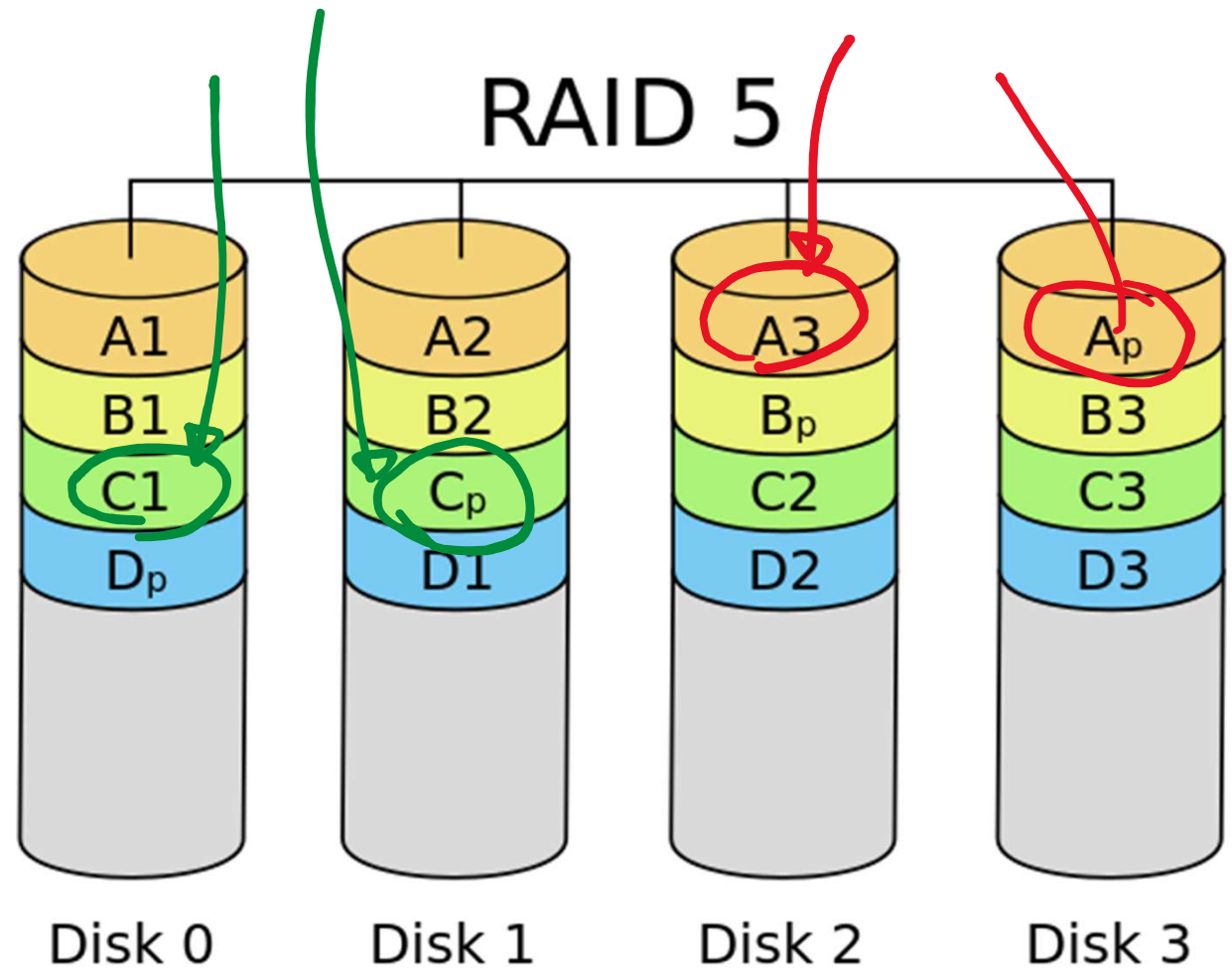- ★ Good read, slow write (parity disk is a bottleneck.)

RAID 4

Pictures from wikipedia

RAID 5

- ★ Block-level parity
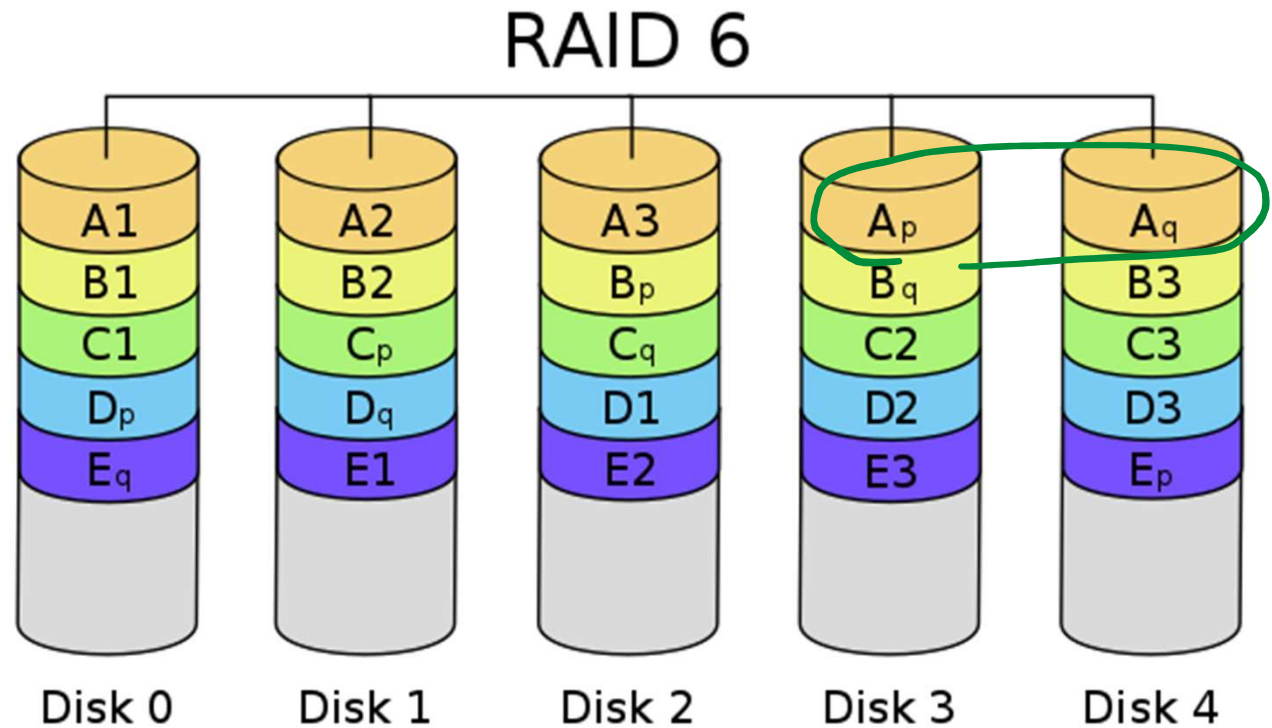- ★ Balance between performance and redundancy
- ★ Tolerant to failure of a disk

RAID 5

Disk 0   Disk 1   Disk 2   Disk 3

Pictures from wikipedia

RAID 6

★ Two block-level parities
★ Tolerant up to failure of 2 disks

RAID 6

Disk 0: A1, B1, C1, D_p, E_q
Disk 1: A2, B2, C_p, D_q, E1
Disk 2: A3, B_p, C_q, D1, E2
Disk 3: A_p, B_q, C2, D2, E3
Disk 4: A_q, B3, C3, D3, E_p

Pictures from wikipedia

Pop Quiz

What is the minimum number of disks required for RAID 0,1, and 5?

2 2 3

Nested RAID

Hot swapable (handwritten)

★ Combined levels of RAID. see RAID 01, RAID 50, RAID 10

Pictures from https://www.lacie.com/as/en/manuals/lrm/raid/

Computer Architecture: Design and Analysis   Krerk Piromsopa, Ph.D. @ 2016
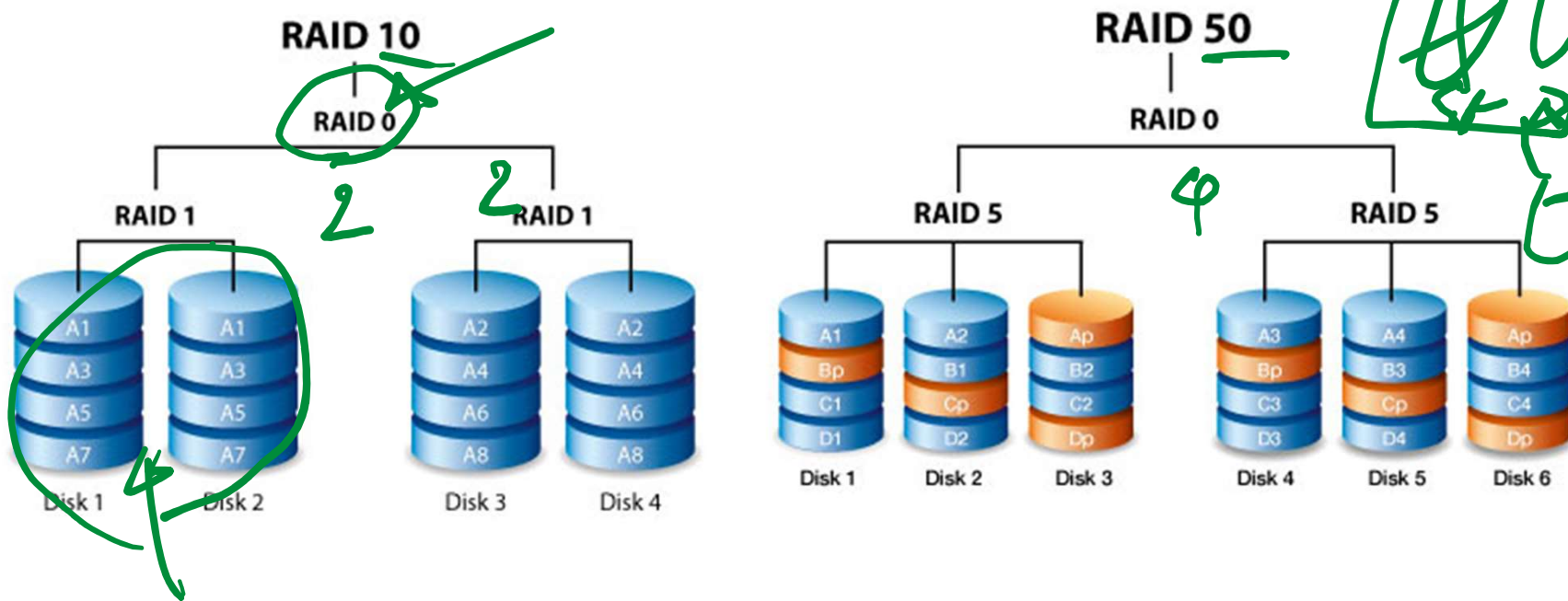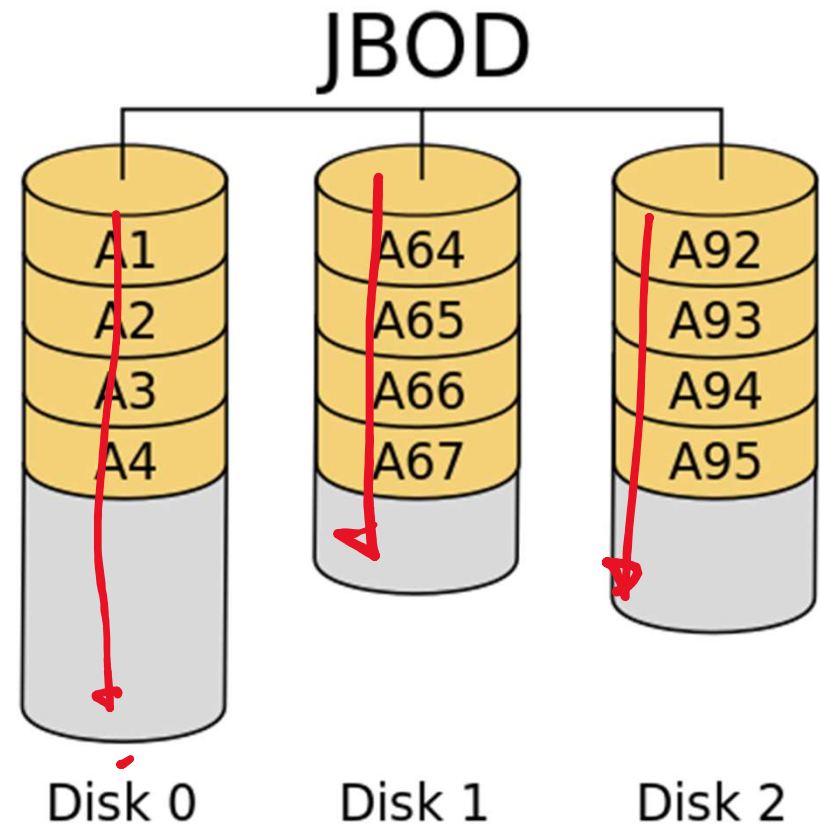
Other non-RAID

★ JBOD (just a bunch of disks/drives)
  ○ Span
  ○ Allows disk with different size to be combined.
★ Redundant?

JBOD

Disk 0	Disk 1	Disk 2
A1	A64	A92
A2	A65	A93
A3	A66	A94
A4	A67	A95

Network Storage

★ Block vs. File
★ SAN vs. NAS

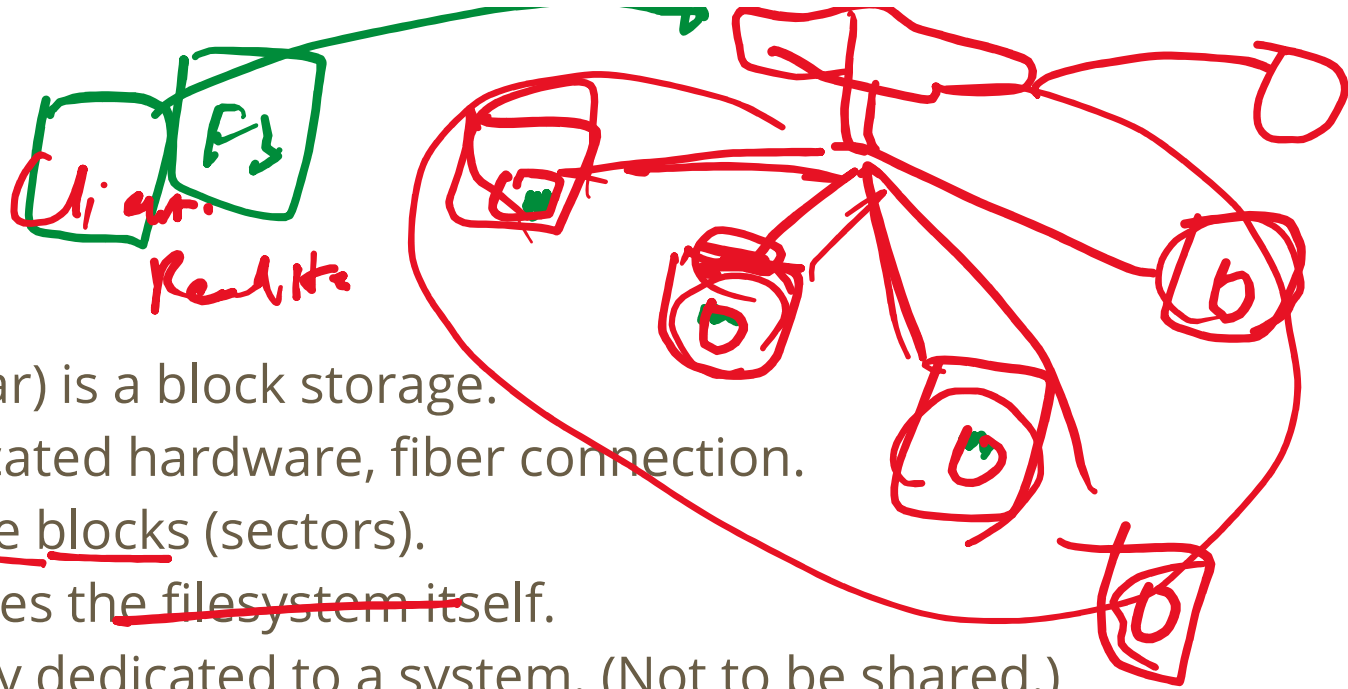★ There exists object (e.g. document) storage, which is not mentioned here.
★ Similar concepts on cloud services
  ○ Amazon Elastic Block Store vs. Amazon S3
  ○ Google Persistent Disk vs. Google Cloud Filestore

(Handwritten annotation on right: three stacked boxes labeled "OS", "FS", "Logical disk")

SAN

★ Storage Area Network
★ SAN (via iSCSI or similar) is a block storage.
★ Usually required dedicated hardware, fiber connection.
★ (Operating) System see blocks (sectors).
★ (Client) System manages the filesystem itself.
★ Block storage is usually dedicated to a system. (Not to be shared.)

★ There exists similar technology such as iATA and AoE.
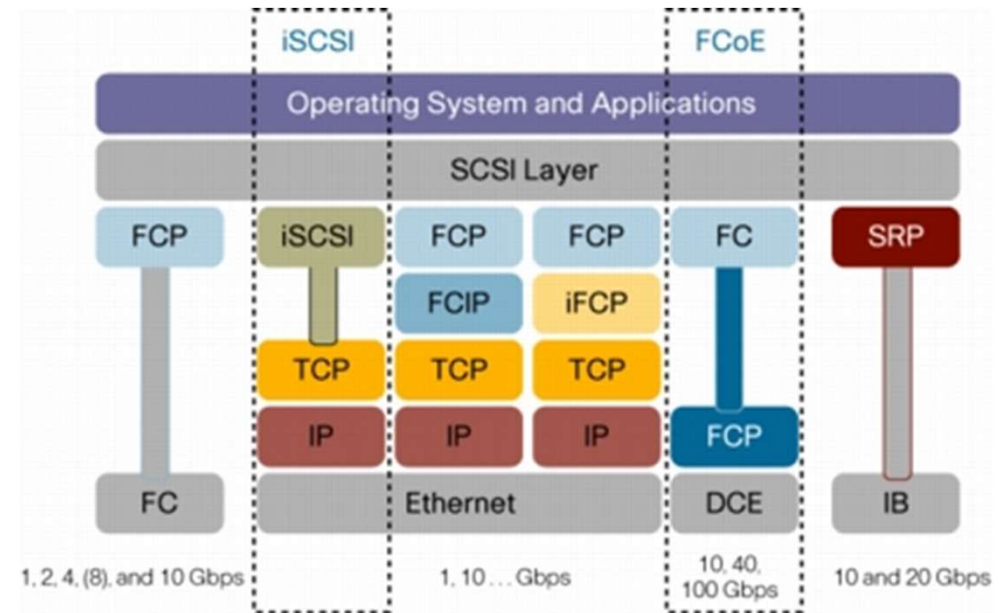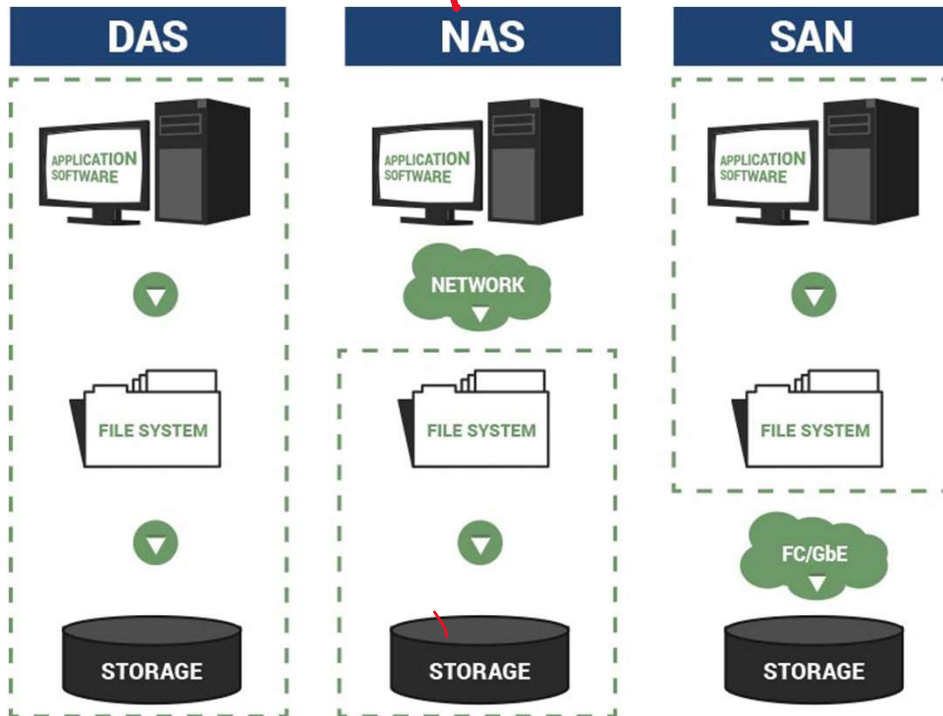However, it is not widely used.

NAS

- ★ Network Attached Storage
- ★ NAS is a file storage.
  Think Network Drive. (e.g. Windows map drive)
- ★ Can easily be shared among clients.
- ★ Host (Server) manages file system (and permissions).
- ★ Several protocols for sharing.
  - ○ NFS
  - ○ CIFS (aka. SMB)
  - ○ AFP (Apple File Protocol, deprecated.)
  - ○ FTP, SFTP
  - ○ etc.

Comparisons

Network attached Storage
Storage area network

Picture from
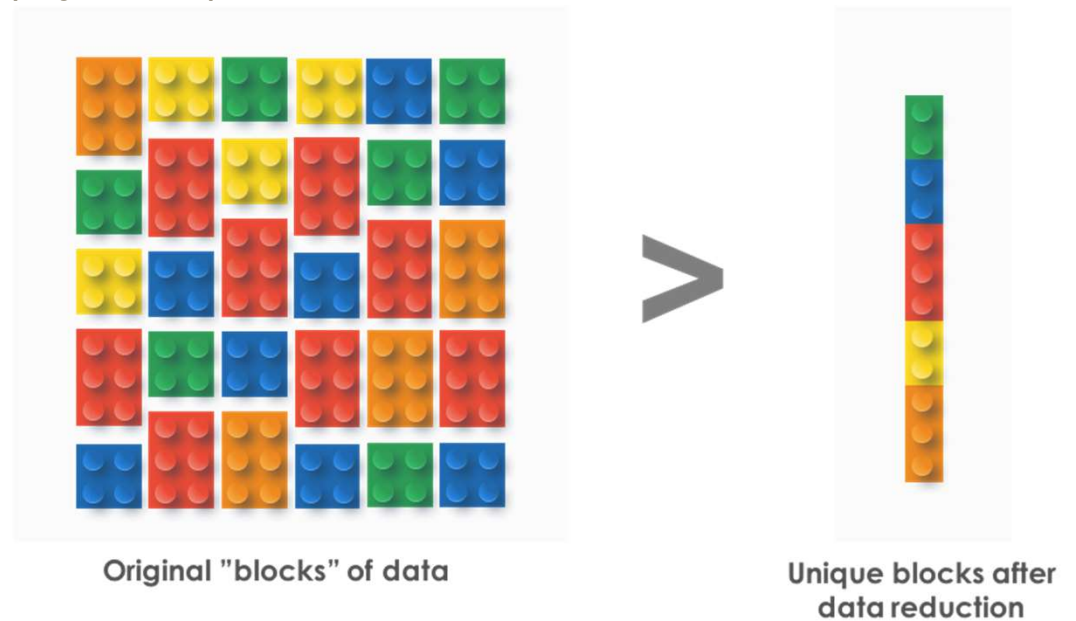https://sites.google.com/site/andreanguyencis4397/module-1/nas-and-san-introduction
https://www.cisco.com/c/en/us/products/collateral/switches/nexus-5000-series-switches/white_paper_c11-495142.html

Other Technologies

★ Deduplication
  ○ Several systems usually have similar files/blocks.
  ○ Can we share files/blocks to save (physical) spaces?

Original "blocks" of data

Unique blocks after data reduction

Picture from https://kaminario.com/company/blog/data-dedupe-all-flash-arrays/

Real-World Consideration

★ Performance
★ Price
★ Block vs. File (vs. Object)
★ Sharing
★ Reliability
★ Redundancy
★ Deduplication

Summary

★ Several Storage Technology
★ RAID allows independent disks to be used for large storage system.
★ SAN provides block access for clients.
★ NAS provides file access for clients

Exercises

Cloud Storage

★ Nowaday, cloud storage is available at reasonably low price. Please name type of storage (e.g. block vs. file vs. object) provided by each technology.
  ○ Amazon S3
  ○ Amazon EBS
  ○ Amazon Dynamo
  ○ Google Cloud Storage Bucket
  ○ Google Drive
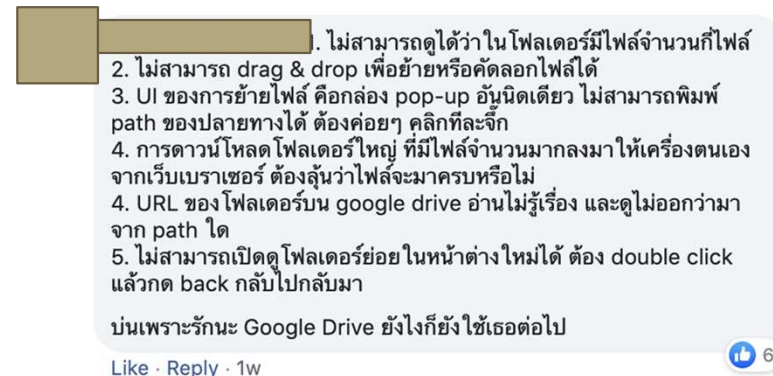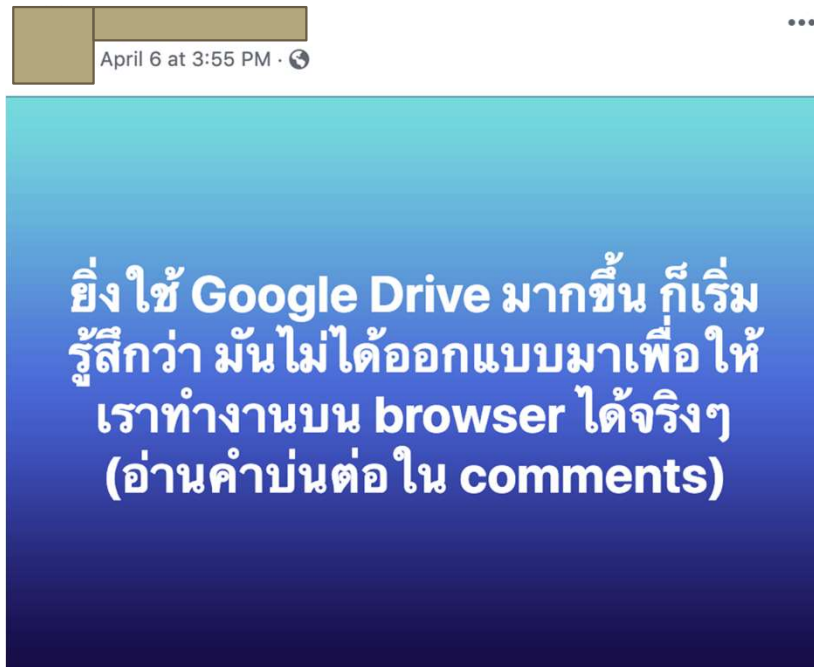 ○ One Drive
  ○ Dropbox
  ○ Apple iCloud

Raid Performance

★ Why is RAID 5 faster when comparing to RAID 4?
  ○ Hint. Parity

Can you explain why the user experience for Google Drive is like this?

Computer Architecture: Design and Analysis   Krerk Piromsopa, Ph.D. @ 2016

End of Chapter 9