

Sistem Operasi

Disk Scheduling, RAID

(Tanenbaum: Section 5.4)

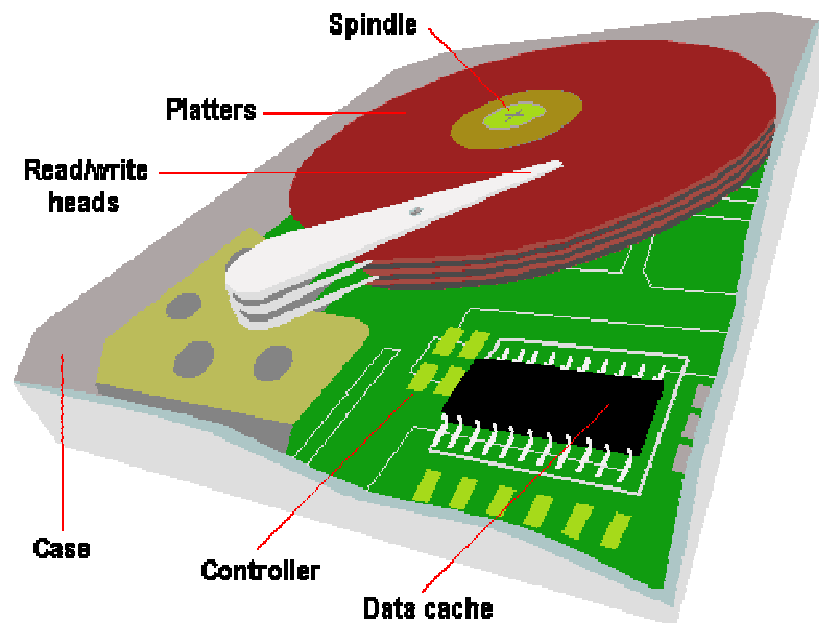
(Siberschatz: Section 14.1, 14.2, 14.5)

Outline

- Struktore Disk
- Penjadwalan Disk
- Struktur RAID

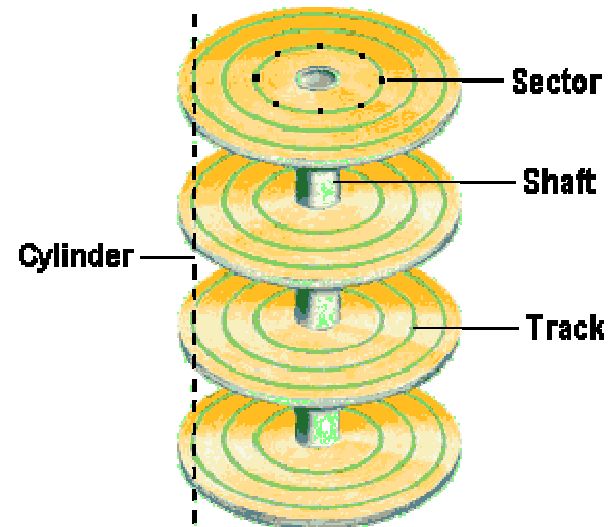
Hard Disk Drive (HDD)

- Controller ada yang di card (MFM, RLL, ESDI, SCSI) ada yang di disk itu sendiri (keluarga IDE/ATA)
- Diberi cache untuk mempercepat akses



HDD - Zoom

- Piringan
- 3000-10,000 RPM (floppy 360 RPM)
- Track, Silinder, Sektor
- Contoh: hdb:
 - CFS540A, 516MB w/64kB Cache, CHS=1050/16/63
 - 1050 silinder (trak), 16 head (8 piringan), 63 sektor per trak, 512 byte per sektor



Struktur Disk

- Disk drive diamati sebagai array besar 1-dimensi dari blok-blok logik, blok logik adalah unit transfer terkecil.
- array besar 1-dimensi dari blok-blok logik dipetakan ke sektor di disk secara sekuensial.
 - Sektor 0 adalah sektor pertama dari trak pertama di sisi silinder paling luar.
 - Pemetaan dilakukan berurut dari trak pertama ke sisa trak di silinder tersebut, dan kemudian ke seluruh sisa silinder dari paling luar ke paling dalam.

Penjadwalan Disk

- Waktu akses = waktu seek + jeda rotasi + waktu transfer
- Waktu seek adalah waktu yang diperlukan untuk memindahkan head ke silinder dimana terdapat sektor yang dituju.
- Jeda rotasi adalah waktu tunggu tambahan untuk disk berputar ke sektor yang dituju oleh head.
- Waktu transfer adalah waktu untuk mengirim bit-bit dari disk

Penjadwalan Disk

- Bandwidth disk adalah jumlah byte yang ditransfer dibagi dengan waktu akses
- Waktu seek mendominasi waktu akses
- Itu sebabnya blok yang lebih besar dikatakan mempunyai laju data lebih besar di sistem file.
- Minimalisasi waktu seek
- Waktu seek \approx jarak seek

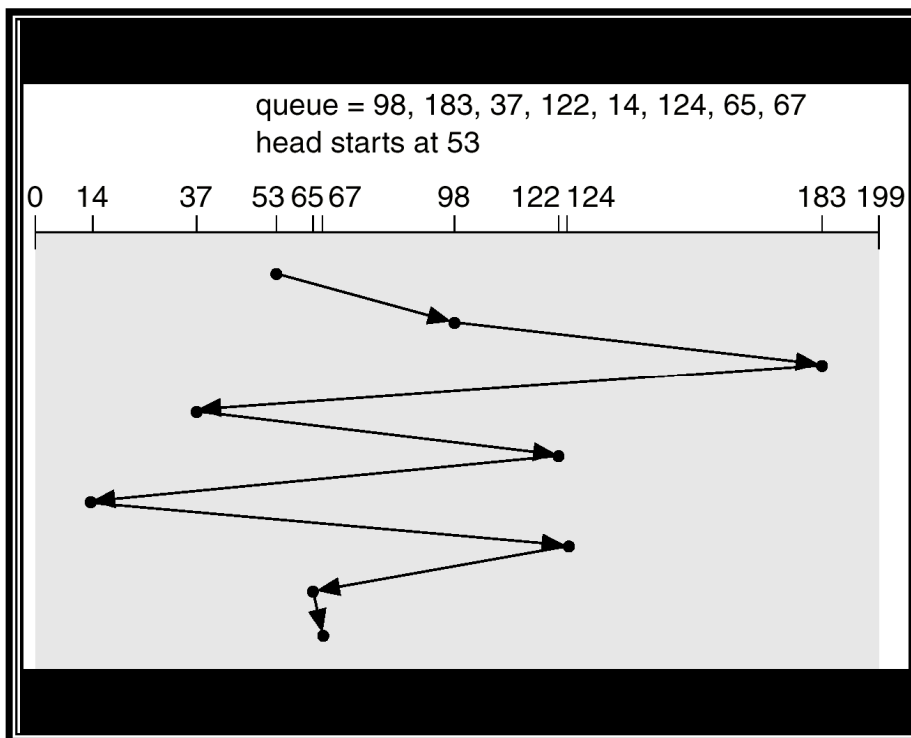
Penjadwalan Disk

- Terdapat beberapa algoritma untuk menjadwalkan layanan permintaan I/O disk.
- Kita gambarkan sebagai sebuah antrian permintaan silinder (0-199).

98, 183, 37, 122, 14, 124, 65, 67

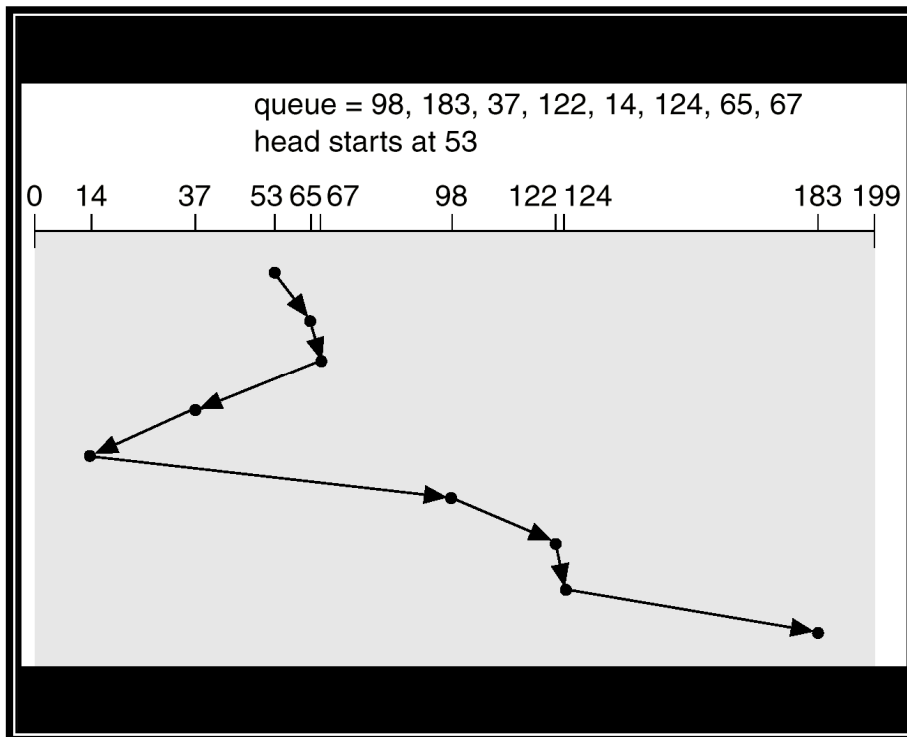
- Head diawali di 53
- Ukur total jarak pergerakan head

FCFS (First come first served)



Total pergerakan head:
640 silinder.

SSTF (shortest seek time first)

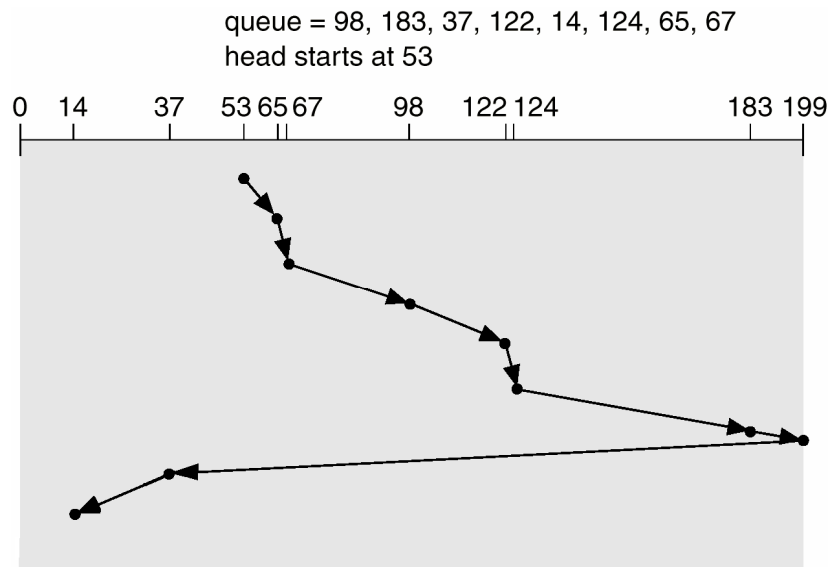


Pilih permintaan terdekat dari head

Total pergerakan head:
236 silinder.

Starvation!

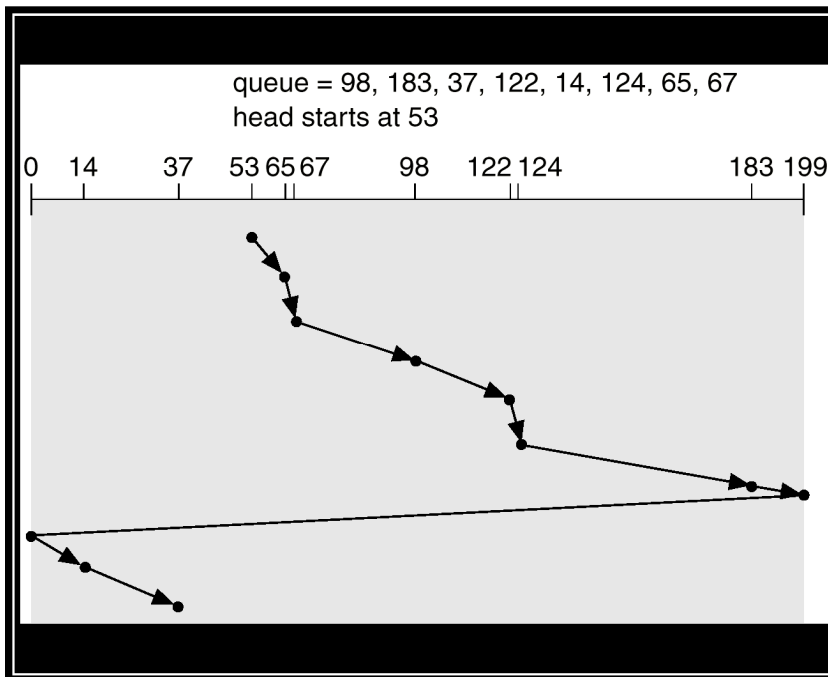
SCAN



Algoritma Elevator

Total pergerakan head:
331 silinder.

C-SCAN



Waktu tunggu lebih rata!

Scan satu arah dan
kembali seketika

Total: $183+199 = 382$

Pemilihan Algoritma

- SSTF adalah yang paling umum dan natural
- SCAN dan C-SCAN bekerja lebih baik untuk sistem berbeban tinggi pada disk.
- Performansi tergantung dari jumlah dan jenis permintaan.
- Permintaan layanan disk dapat dipengaruhi metoda alokasi file.
- Algoritma penjadwalan disk harus ditulis sebagai modul terpisah dari OS, memungkinkan penggantinya dengan algoritma lain jika diperlukan.
- Baik SSTF atau LOOK adalah pilihan yang bijak sebagai algoritma default.

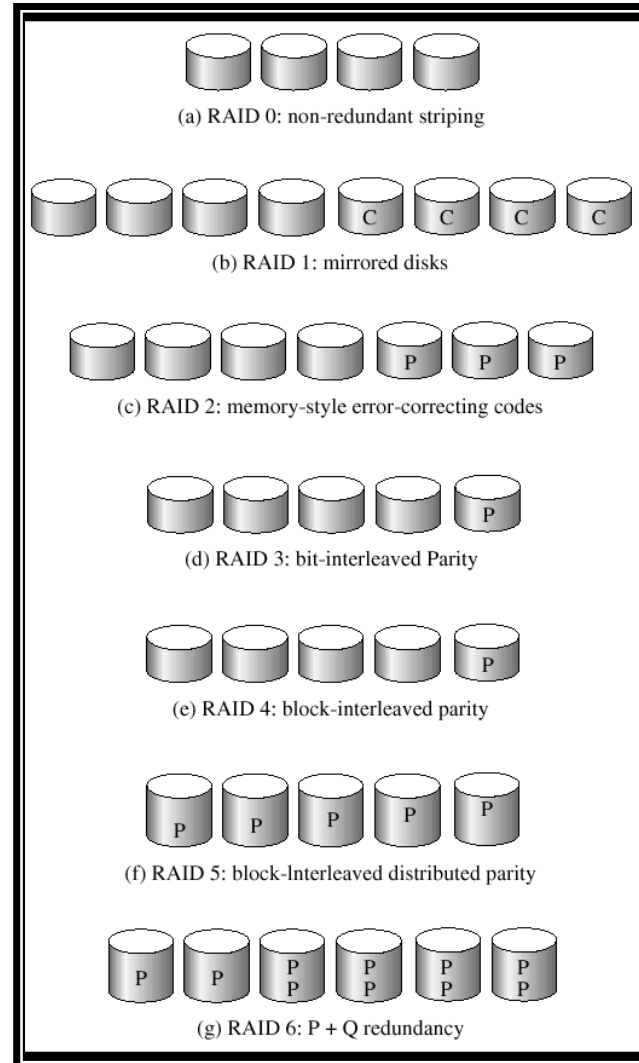
Disk Scheduling

Name	Description	Remarks
Selection according to requestor		
RSS	Random scheduling	For analysis and simulation
FIFO	First in first out	Fairest of them all
PRI	Priority by process	Control outside of disk queue management
LIFO	Last in first out	Maximize locality and resource utilization
Selection according to requested item		
SSTF	Shortest service time first	High utilization, small queues
SCAN	Back and forth over disk	Better service distribution
C-SCAN	One way with fast return	Lower service variability
N-step-SCAN	SCAN of N records at a time	Service guarantee
FSCAN	N-step-SCAN with $N =$ queue size at beginning of SCAN cycle	Load sensitive

Struktur RAID

- **RAID** – Redundant Array of Inexpensive Disk
- Sekarang – **Redundant Array of Independent Disk**
- Menyediakan **kehandalan** melalui **redundansi**.
- Misal . Mirroring
- Meningkatkan **performansi** melalui **parallelisme**.
- Misal. Data Stripping
- RAID dibentuk dalam 7 tingkat berbeda.

Tingkat RAID



RAID Tingkat 0 ~ 3



(a) RAID 0: non-redundant striping



(b) RAID 1: mirrored disks



(c) RAID 2: memory-style error-correcting codes



(d) RAID 3: bit-interleaved Parity

RAID Tingkat 4 ~ 6



(e) RAID 4: block-interleaved parity



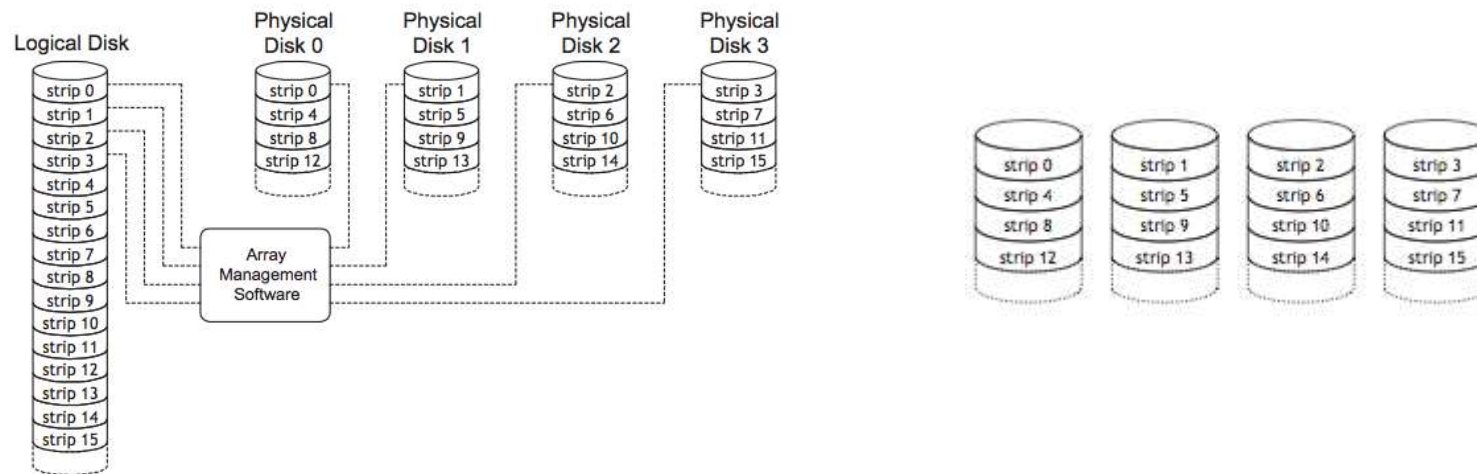
(f) RAID 5: block-Interleaved distributed parity



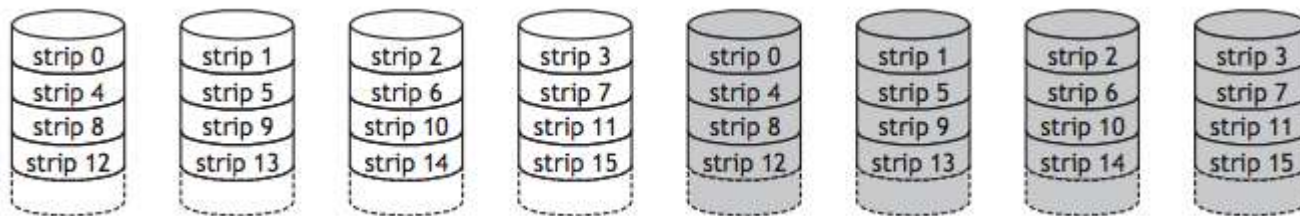
(g) RAID 6: P + Q redundancy

RAID 0 (non-redundant)

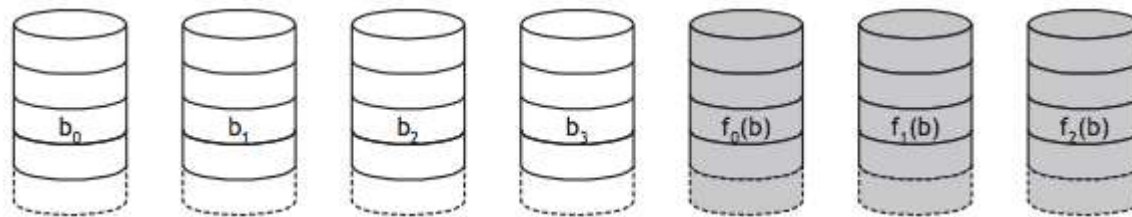
Data Mapping for a RAID Level 0 Array



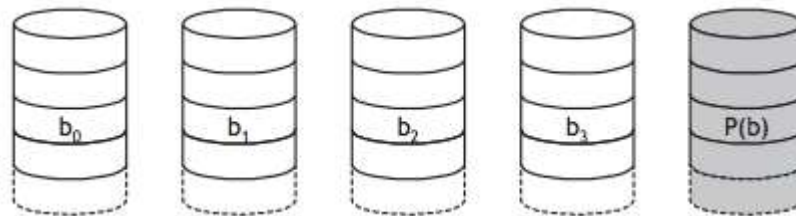
RAID 1 (mirrored)



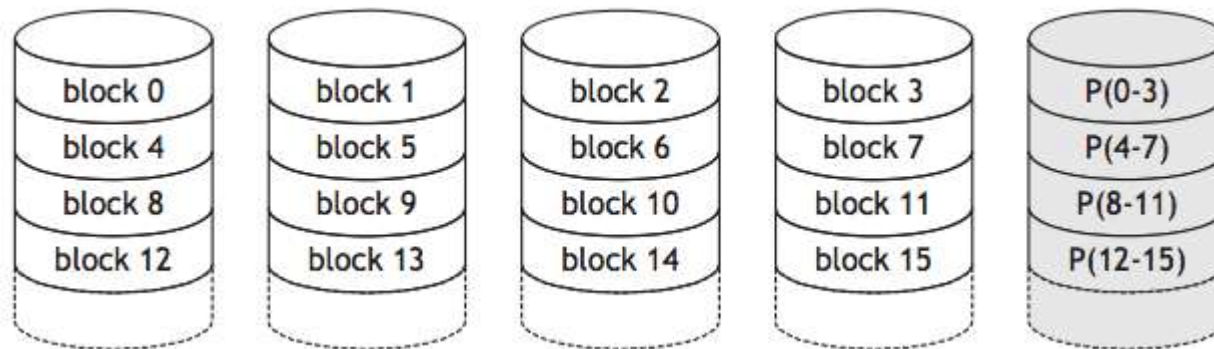
RAID 2 (redundancy through Hamming code)



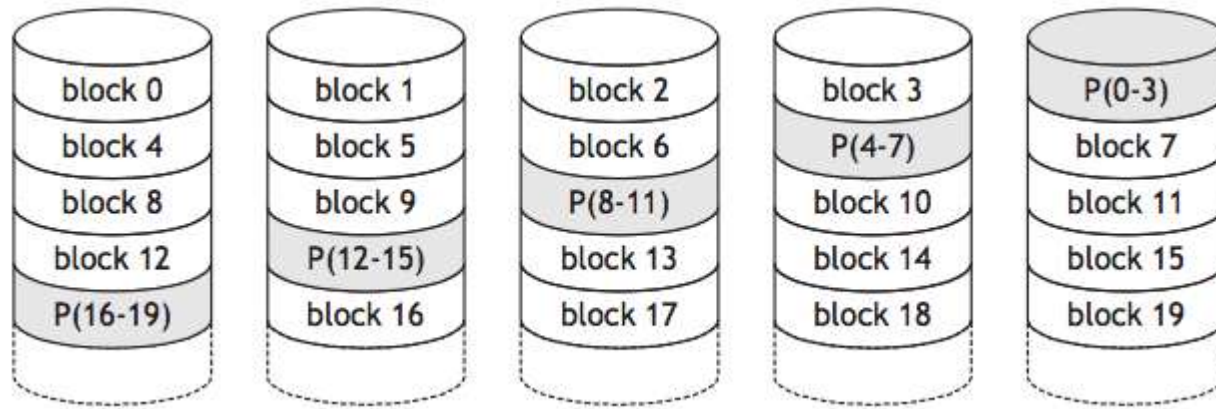
RAID 3 (bit-interleaved parity)



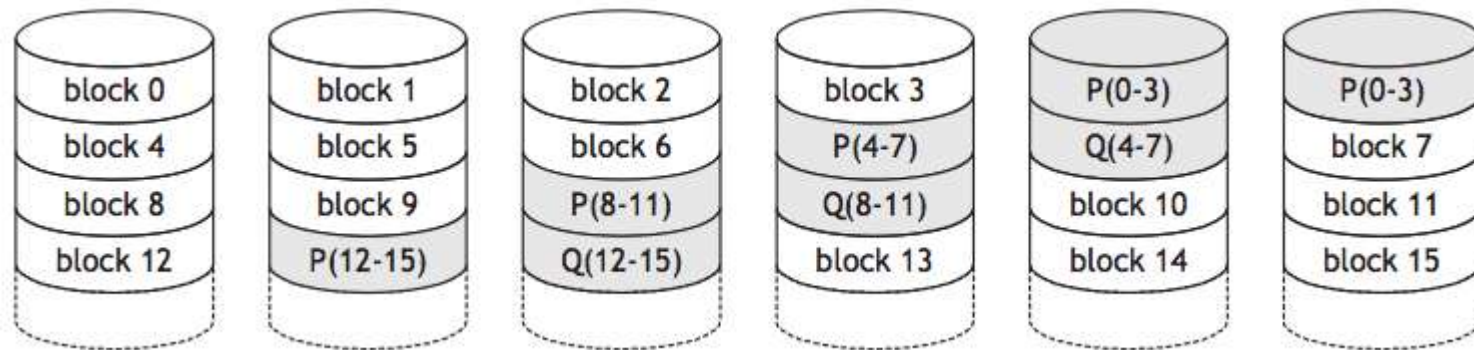
RAID 4 (block level parity)



RAID 5 (block-level distributed parity)



RAID 6 (dual redundancy)

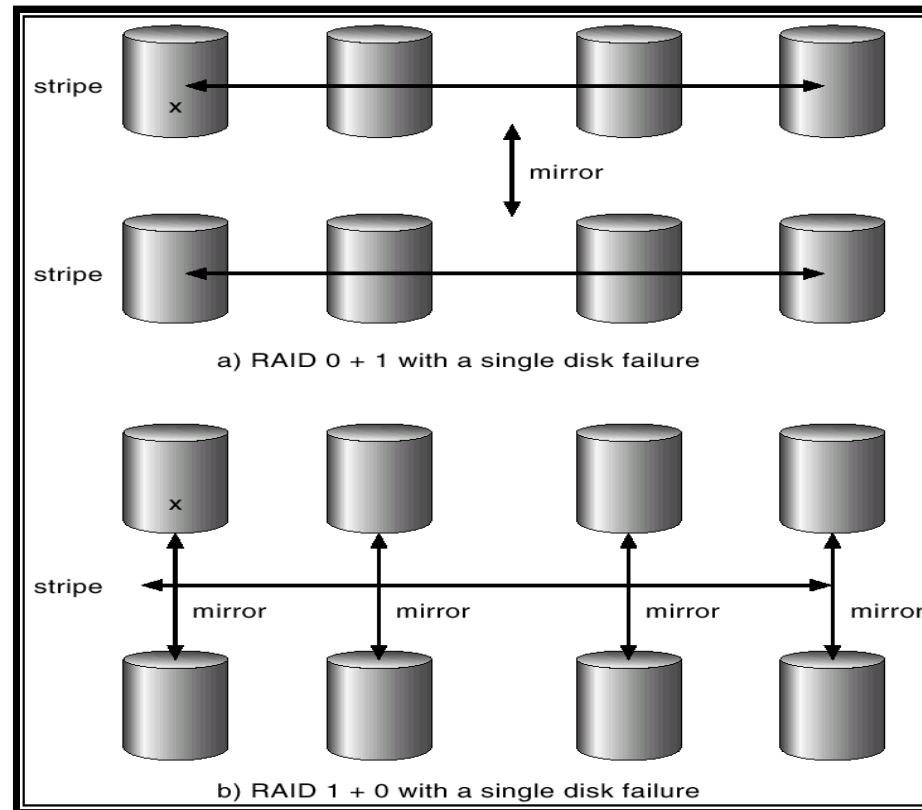


Perbandingan RAID

Category	Level	Description	Disks required	Data availability	Large I/O data transfer capacity	Small I/O request rate
Striping	0	Nonredundant	N	Lower than single disk	Very high	Very high for both read and write
Mirroring	1	Mirrored	$2N$	Higher than RAID 2, 3, 4, or 5; lower than RAID 6	Higher than single disk for read; similar to single disk for write	Up to twice that of a single disk for read; similar to single disk for write
Parallel access	2	Redundant via Hamming code	$N + m$	Much higher than single disk; comparable to RAID 3, 4, or 5	Highest of all listed alternatives	Approximately twice that of a single disk
	3	Bit-interleaved parity	$N + 1$	Much higher than single disk; comparable to RAID 2, 4, or 5	Highest of all listed alternatives	Approximately twice that of a single disk
Independent access	4	Block-interleaved parity	$N + 1$	Much higher than single disk; comparable to RAID 2, 3, or 5	Similar to RAID 0 for read; significantly lower than single disk for write	Similar to RAID 0 for read; significantly lower than single disk for write
	5	Block-interleaved distributed parity	$N + 1$	Much higher than single disk; comparable to RAID 2, 3, or 4	Similar to RAID 0 for read; lower than single disk for write	Similar to RAID 0 for read; generally lower than single disk for write
	6	Block-interleaved dual distributed parity	$N + 2$	Highest of all listed alternatives	Similar to RAID 0 for read; lower than RAID 5 for write	Similar to RAID 0 for read; significantly lower than RAID 5 for write

N = number of data disks; m proportional to $\log N$

RAID (0 + 1) dan (1 + 0)



Review

- Apa ketiga bagian dari waktu akses harddisk? Siapa yang mendominasi?
- Apa yang menjadi masalah dari SSTF (shortest seek time first)?
- Apa dua kegunaan utama dari RAID?