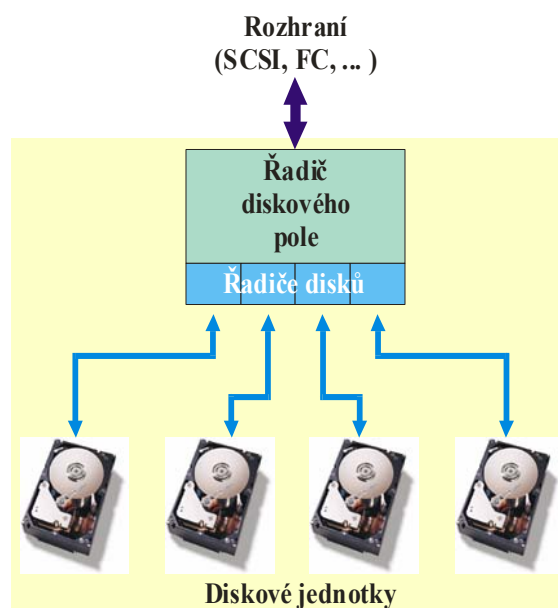


# Disková pole

1

## Princip

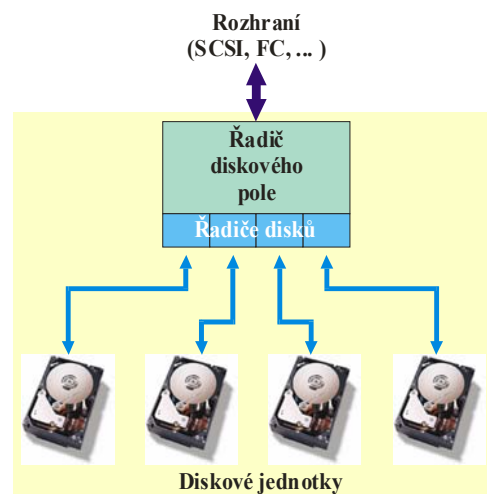
- Rozdělení celkové kapacity disků mezi několik diskových jednotek.
- Důvody:
  1. Zvýšení rychlosti.
    - Paralelní práce většího počtu diskových jednotek.
  2. Zvýšení bezpečnosti uložených dat (proti ztrátě).
    - Data jsou rozložena na více jednotkách.



2

## Princip

- Diskové jednotky bývají připojeny systémem pro **hot swap** – lze je vyměňovat za provozu.
- V systémech se zvýšenou bezpečností se po výměně postupně obnoví data na nové diskové jednotce.

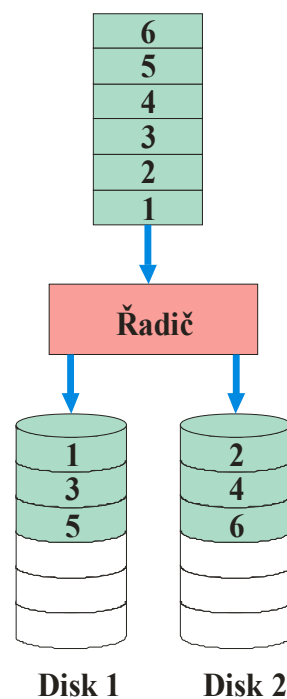


## Základní techniky

- Souhrn základních principů rozpracován na University of California in Berkeley.
- **RAID** – Redundant Array of Independent Disks.
- Má několik základních typů (RAID 0 – RAID 7).
- Lze použít i jinou klasifikaci.

## RAID 0

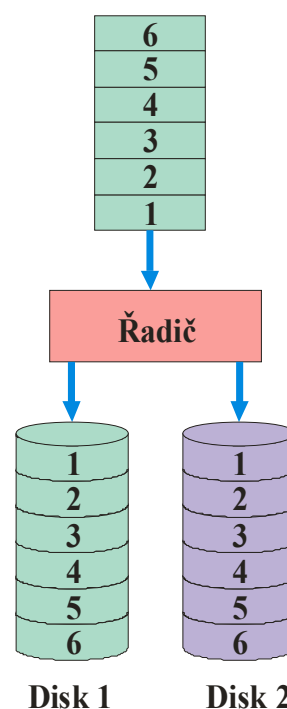
- Používá základní techniku **Striping**.
- Data jsou rozdělena na několik bloků (strip), které se střídavě zapisují na jednotlivé disky.
- Velikost stripu ~ 100 kB – 1 MB.
- Zvýšení rychlosti: ANO.
- Zvýšení bezpečnosti dat: NE.



5

## RAID 1

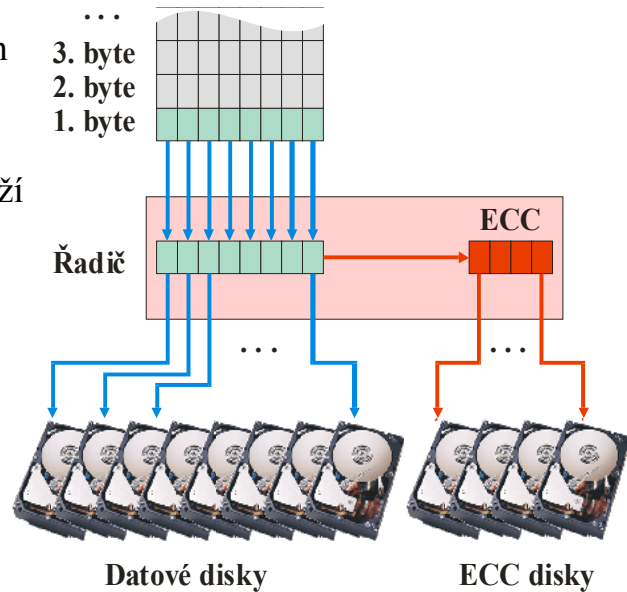
- Používá základní techniku **Mirroring** (zrcadlení dat).
- Alternativa: **Duplexing** – jsou zdvojené i řadiče.
- Data jsou pro zvýšení bezpečnosti zapisována 2× (na 2 diskové jednotky).
- Zvýšení rychlosti: Jen při čtení.
- Zvýšení bezpečnosti dat: ANO.
- Nízké využití diskových jednotek.



6

## RAID 2

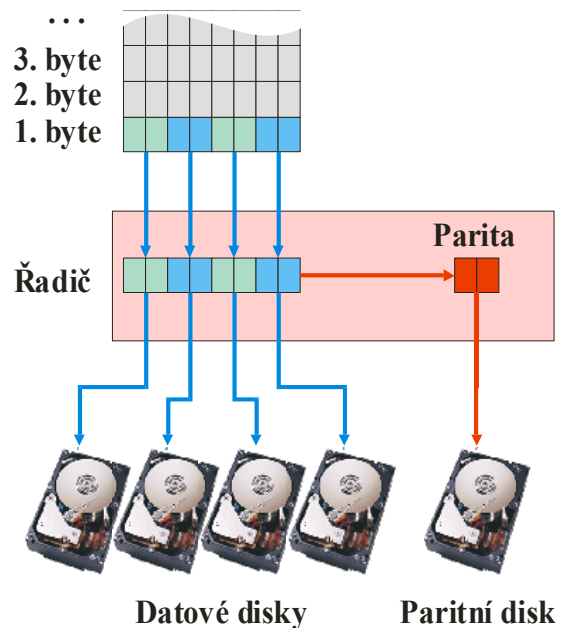
- Používá striping na úrovni jednotlivých bitů.
- Pro každé slovo (byte, word, ...) se vypočítá ECC (Hammingův kód) a uloží se na samostatný disk.
- Zvýšení rychlosti:
  - Pro dlouhé přenosy ANO,
  - Náhodný přístup: nepodporuje více požadavků současně, zápis pomalý.
- Zvýšení bezpečnosti: ANO.
- Složitější technická realizace, nízké využití kapacity (velká redundance).



7

## RAID 3

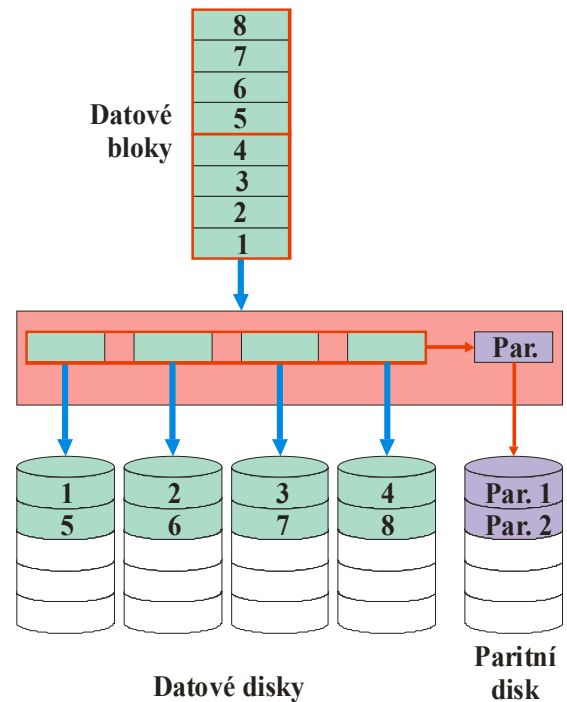
- Používá striping na úrovni bitů.
- Každé slovo doplněno paritními bity, které se zapisují na samostatný disk.
- Zvýšení rychlosti:
  - Pro dlouhé přenosy ANO,
  - Náhodný přístup: nepodporuje více požadavků současně, zápis pomalý.
- Zvýšení bezpečnosti: ANO (toleruje výpadek 1 disku).
- Složitější technická realizace.



8

## RAID 4

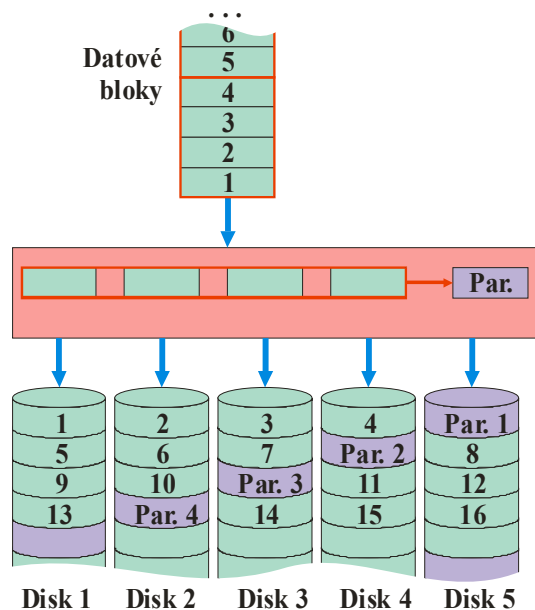
- Podobný jako RAID 3, ale používá striping na úrovni větších bloků.
- Paritní stripy se zapisují na vyhrazený paritní disk.
- Vlastnosti podobné jako RAID 3, ale lepší obsluha vícenásobných požadavků na čtení.



9

## RAID 5

- Podobný jako RAID 4.
- Paritní stripy rozloženy na všechny disky.
- Lepší obsluha vícenásobných požadavků na zápis než u RAID 4.



10

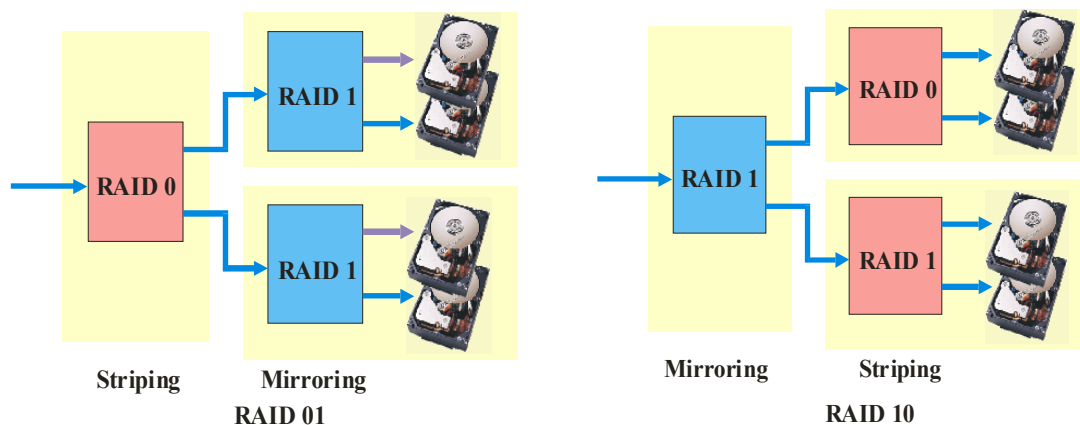
## RAID 6 a RAID 7

- RAID 6 je odvozen od RAID 5. Používá dvojnásobnou paritu  $\Rightarrow$  zvýšení bezpečnosti proti RAID 5.
- RAID 7 – není standardní. Je to rozšíření RAID 3 a 4 podle Storage Computer Corporation. Používá velké CACHE pro obsluhu více požadavků.

11

## Kombinace různých systémů RAID

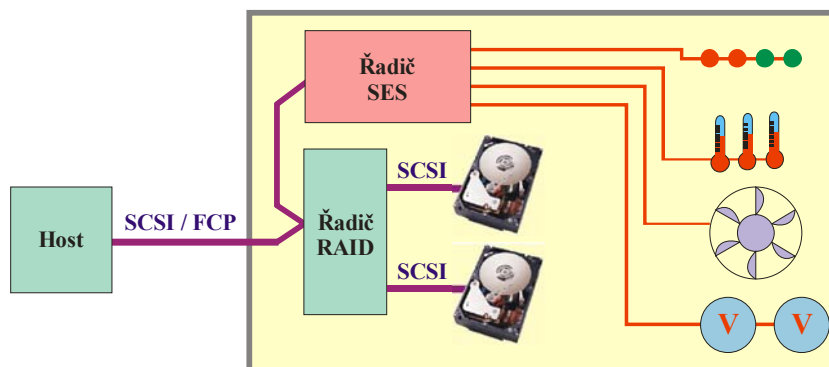
- RAID 01 = RAID 0 + RAID 1
- RAID 10 = RAID 1 + RAID 0
- Podobně např. RAID 03, RAID 30 (= RAID 53).



12

## SCSI Enclosure Services (1)

- **SES (SCSI Enclosure Services)** definuje způsob komunikace s příslušenstvím ve skříní disku / diskového pole.
- Skříň se chová jako zařízení s vlastním LUN nebo SCSI číslem.

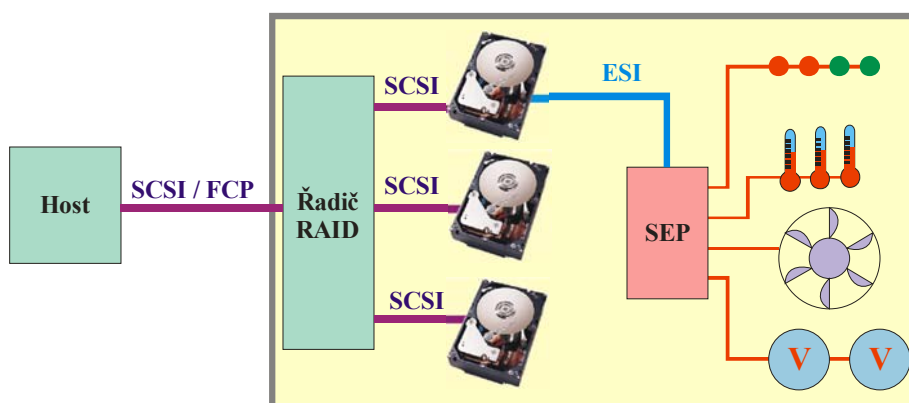


Řadič SES má vlastní SCSI číslo.

13

## SCSI Enclosure Services (2)

- **SEP (Service Enclosure Processor)** je připojen k některé diskové jednotce jako samostatné logické zařízení (má vlastní LUN).
- **SEP** je s diskovou jednotkou propojen rozhraním **ESI (Enclosure Services Interface)**.
  - Pro vývody ESI se na diskové jednotce používá konektor pro nastavení SCSI čísla.



14

# SCSI příkazy pro SES

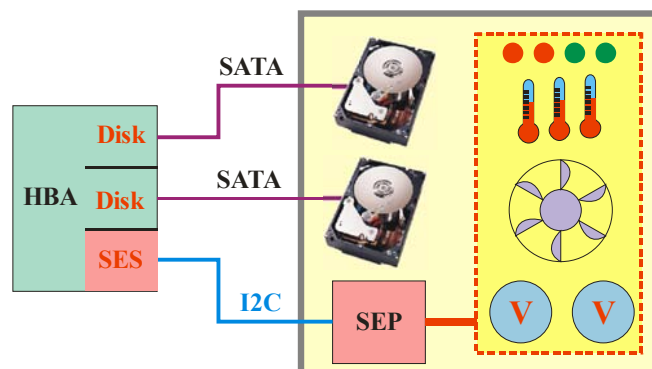
- Pro řízení SES jsou definovány speciální příkazy SCSI:
  - Send Diagnostic.
  - Receive Diagnostic Results.
- Příkazy zapisují / čtou datové bloky s diagnostickými daty.

Command name	Operation code	Type	Subclause
CHANGE DEFINITION	40h	O	8.2.1
COMPARE	39h	O	8.2.2
COPY	18h	O	8.2.3
COPY AND VERIFY	3Ah	O	8.2.4
FORMAT UNIT	04h	M	9.2.1
INQUIRY	12h	M	8.2.5
LOCK-UNLOCK CACHE	36h	O	9.2.2
LOG SELECT	4Ch	O	8.2.6
LOG SENSE	4Dh	O	8.2.7
MODE SELECT (6)	15h	O	8.2.8
MODE SELECT (10)	55h	O	8.2.9
MODE SENSE (6)	1Ah	O	8.2.10
MODE SENSE (10)	5Ah	O	8.2.11
PRE-FETCH	34h	O	9.2.3
PREVENT-ALLOW MEDIUM REMOVAL	1Eh	O	9.2.4
READ (6)	08h	M	9.2.5
READ (10)	28h	M	9.2.6
READ BUFFER	3Ch	O	8.2.12
READ CAPACITY	25h	M	9.2.7
READ DEFECT DATA	37h	O	9.2.8
READ LONG	3Eh	O	9.2.9
REASSIGN BLOCKS	07h	O	9.2.10
RECEIVE DIAGNOSTIC RESULTS	1Ch	O	8.2.13
RELEASE	17h	M	9.2.11
REQUEST SENSE	03h	M	8.2.14
RESERVE	16h	M	9.2.12
REZERO UNIT	01h	O	9.2.13
SEARCH DATA EQUAL	31h	O	9.2.14.1
SEARCH DATA HIGH	30h	O	9.2.14.2
SEARCH DATA LOW	32h	O	9.2.14.3
SEEK (6)	0Bh	O	9.2.15
SEEK (10)	2Bh	O	9.2.15
SEND DIAGNOSTIC	1Dh	M	8.2.15
SET LIMITS	33h	O	9.2.16
START STOP UNIT	1Bh	O	9.2.17
SYNCHRONIZE CACHE	35h	O	9.2.18
TEST UNIT READY	00h	M	8.2.16
VERIFY	2Fh	O	9.2.19
WRITE (6)	0Ah	O	9.2.20
WRITE (10)	2Ah	O	9.2.21
WRITE AND VERIFY	2Eh	O	9.2.22
WRITE BUFFER	3Bh	O	8.2.17
WRITE LONG	3Fh	O	9.2.23
WRITE SAME	41h	O	9.2.24

15

# Podpora SES v SATA II (1)

- SATA II má ve specifikaci podporu pro SES.
- SES se ovládá jako samostatná jednotka přímo v HBA (Host Bus Adapter).

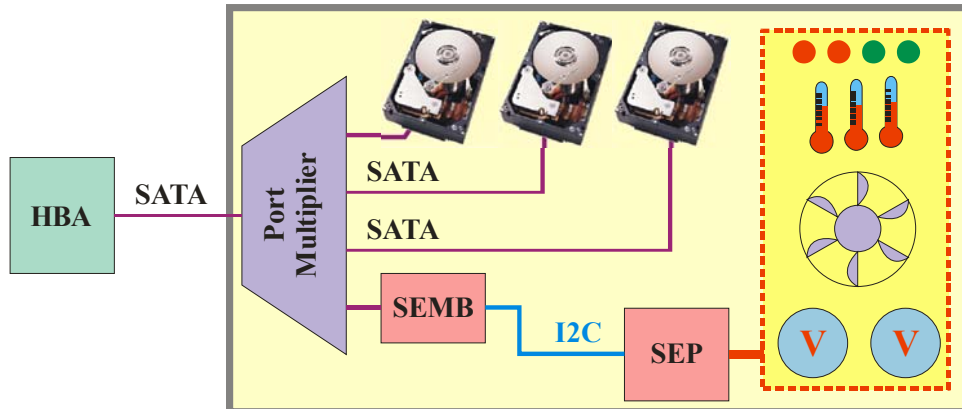


16



## Podpora SES v SATA II (2)

- Složitější případ: SEP je připojen jako disková jednotka na port Multiplieru.
  - SEMB (SATA Enclosure Management Bridge) předává diagnostická data do SEP.
  - SEMB a SEP mohou být fyzicky součástí Multiplieru.



17

## Příkazy SES v SATA II

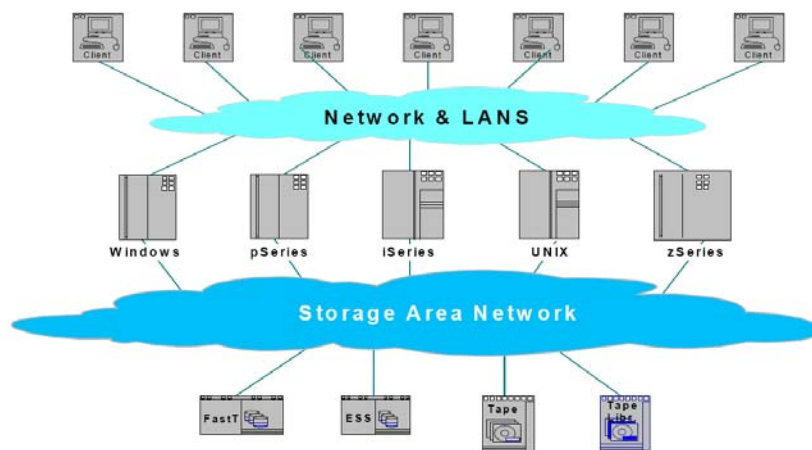
- Pro řízení SES rozeznává jednotka příkaz SEP\_ATTN (SEP Attention).
  - V **CMD\_TYPE** je rozlišení příkazů SEP nebo SAF-TE (SCSI Accessed Fault-Tolerant Enclosure).
  - V **SEP\_CMD** je kód příkazu – viz SES.

Register	7	6	5	4	3	2	1	0
Features				SEP_CMD				
Features (exp)				Reserved				
Sector Count				LEN				
Sector Count (exp)				Reserved				
Sector Number				CMD_TYPE				
Sector Number (exp)				Reserved				
Cylinder Low				Reserved				
Cylinder Low (exp)				Reserved				
Cylinder High				Reserved				
Cylinder High (exp)				Reserved				
Device/Head	Reserved		0		Reserved			
Command	SEP_ATTN (67h)							

18

## Diskové sítě

- SAN (Storage Area Network)
  - Postavené obvykle na bázi FC.
  - Propojuje diskové jednotky, pole atd.
  - Přístup na nižší úrovni než v LAN.



19

## Reference

- ❖ <http://www.t10.org/drafts.htm>  
(draft specifikací SES).
- ❖ <http://www.snia.org>  
(diskové sítě).

20