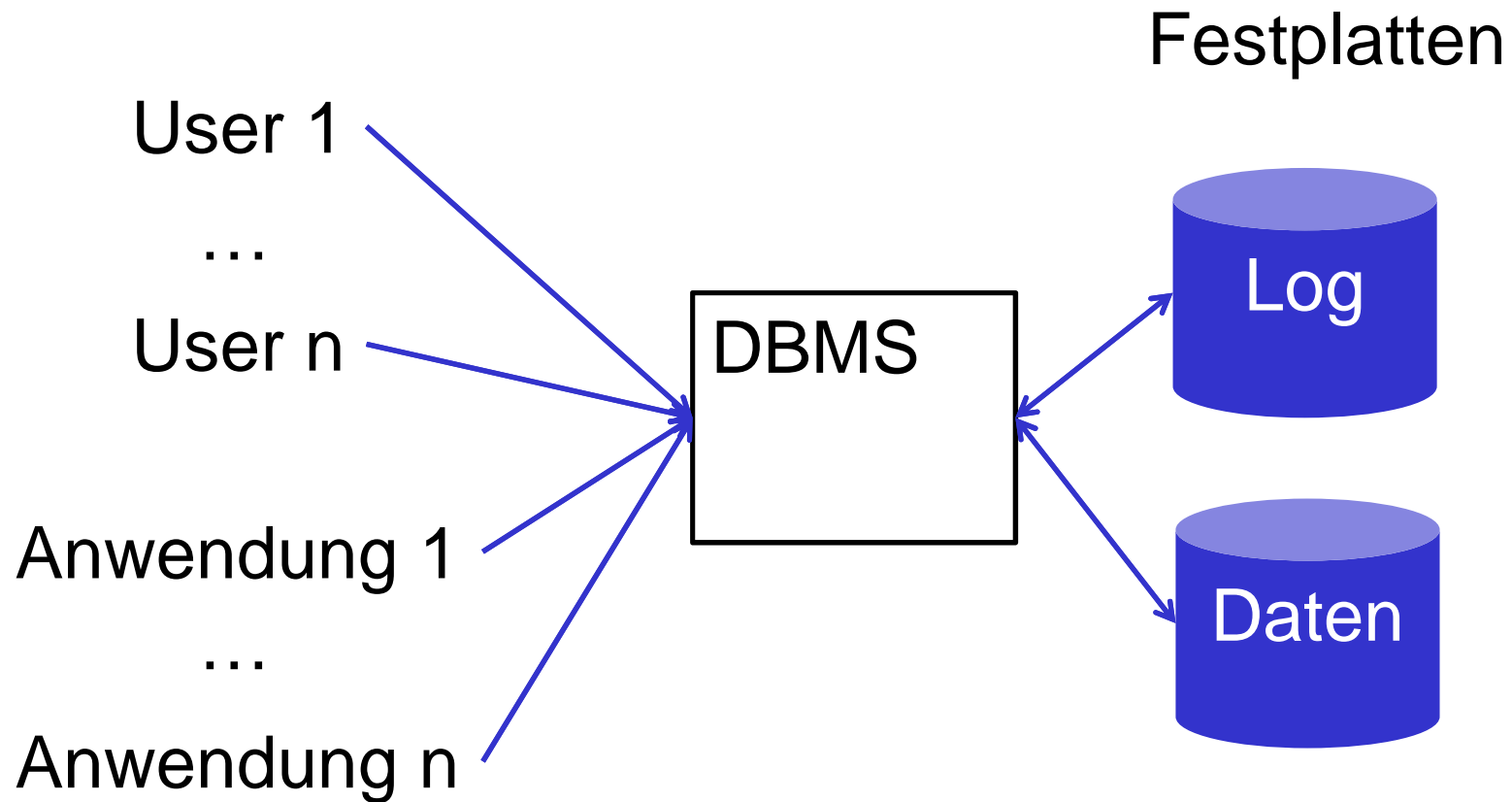


# Grundlagen der Datenbanksysteme 2 (M-DB2)

Dr. Karsten Tolle



# Probleme

---

1. Medienfehler → Festplatte
2. Zugriffszeiten beim Lesen und Schreiben
3. Wie speichert man z.B. 100 TB?

# Medienfehler - Festplattentypen

---

Welche Typen gibt es?

- IDE (ATA), **SATA**, SCSI, **SAS** (Serial Attached SCSI), Fibre Channel, ...

Übertragungsrate: SATA III und SAS (bis 6 Gbit/s)

Bem.: SAS hat Dualport, dort also bis 12Gbit/s möglich

- HDD, SSHD (Hybrid), SSD

z.B. SSHD:

Hybrid-Festplatte mit 1 TB HDD und 8 GB SSD

6,35 cm (2,5") Festplatte mit 5.400 rpm, 64 MB

Cache und SATA III-Anschluss

**SATA: Serial Advanced Technology Attachment**

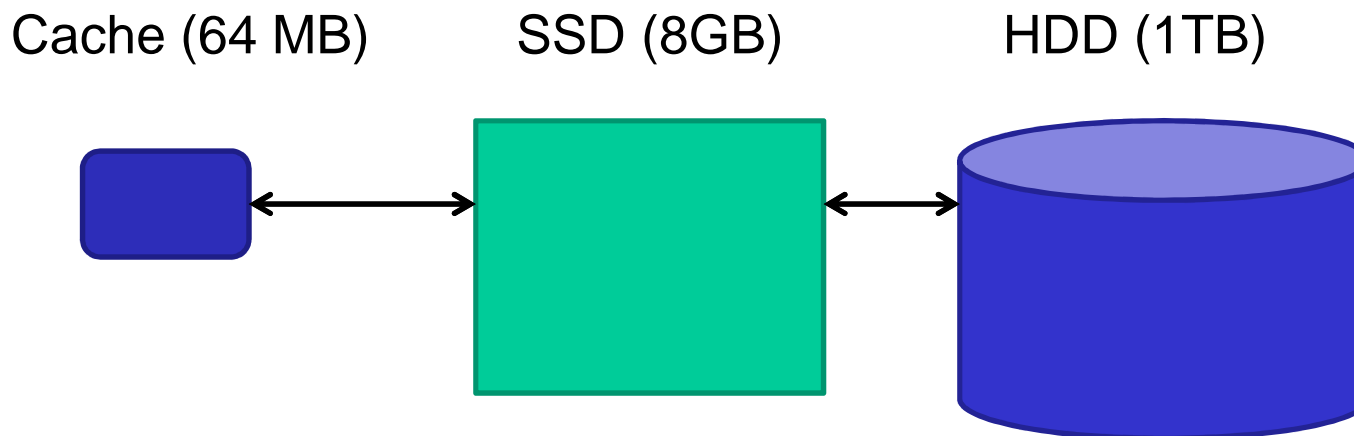
**SCSI: Small Computer System Interface**

# Medienfehler - Festplattentypen

## Aufbau?

z.B. SSHD:

- Hybrid-Festplatte mit 1 TB HDD und 8 GB SSD
- 6,35 cm (2,5") Festplatte mit 5.400 rpm, 64 MB Cache und SATA III-Anschluss



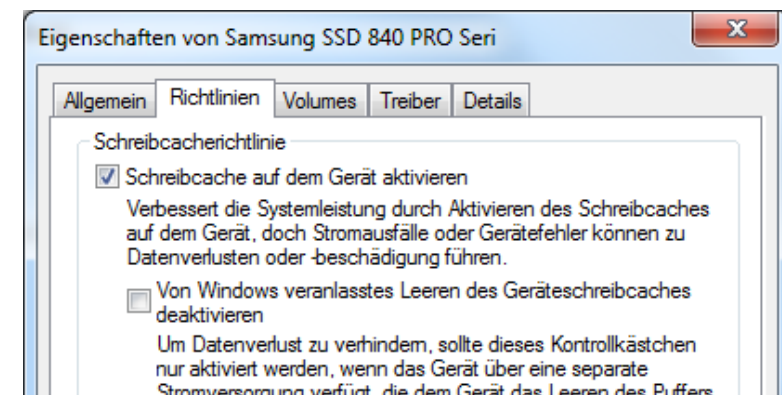
Was ist eigentlich mit dem Cache, wenn der Strom ausfällt?

## Was ist eigentlich mit dem Cache, wenn der Strom ausfällt?

### Schreibcache ist weg!

- Was dort geschrieben wurde, wurde per Handshake als geschrieben an den Scheduler gemeldet! → es kann also zu Problemen kommen!
- Lösungen:
  - Batterie Backup Unit (teilweise speziell für HDDs)
  - Schreibcache deaktivieren → schlechtere Performance

Windows: Gerätemanager → HDD →



# Medienfehler - Fehlertypen

---

- Head-Crash – Aufsetzen des Lesekopfes  
→ Kratzer oder Abreißen des Kopfes
- Überhitzungen, Magnetfelder, langes nicht verwenden („sticky disk“), ...
- UBE (unrecoverable bit error) – maximale Anzahl von Fehlern pro Leseroperation, z.B.  $10^{-14}$  (max. 1 Fehler pro  $10^{14}$  gelesene Bits) – wird vom Hersteller angegeben – normal sind: Consumer:  $10^{-14}$  Server: bis  $10^{-16}$   
*Siehe auch: unrecoverable read errors (URE) oder latent sector errors (LRE)*

# Probleme lösen

---

1. Medienfehler → Festplatte  
Lösungsansatz: Redundante Speicherung
2. Zugriffszeiten beim Lesen und Schreiben  
Lösungsansatz: mehrere Platten gleichzeitig nutzen
3. Wie speichert man z.B. 100 TB?  
Lösungsansatz: mehrere Platten zu einer „virtuellen“ Platte zusammenlegen

Womit: RAID - redundant array of independent disks



# RAID Umsetzungen

- Software-RAID

- RAID Controller



z.B. DC-324E RAID BLI

4-Kanal SATA II RAID 5 PCI-Express Controller

- Unterstützt RAID 1, 0, 10, 5, JBOD und Backup
- Vier unabhängige Serial ATA II Kanäle
- Zwei Kanäle wahlweise als eSATA-Port nutzbar
- 300 MByte/s Datentransferrate

- Hardware-RAID



Disk-Array (auch DAS, SAN oder NAS)

# RAID

---

- (früher) redundant array of inexpensive disks
- (heute) redundant array of independent disks

In Büchern findet man oft:

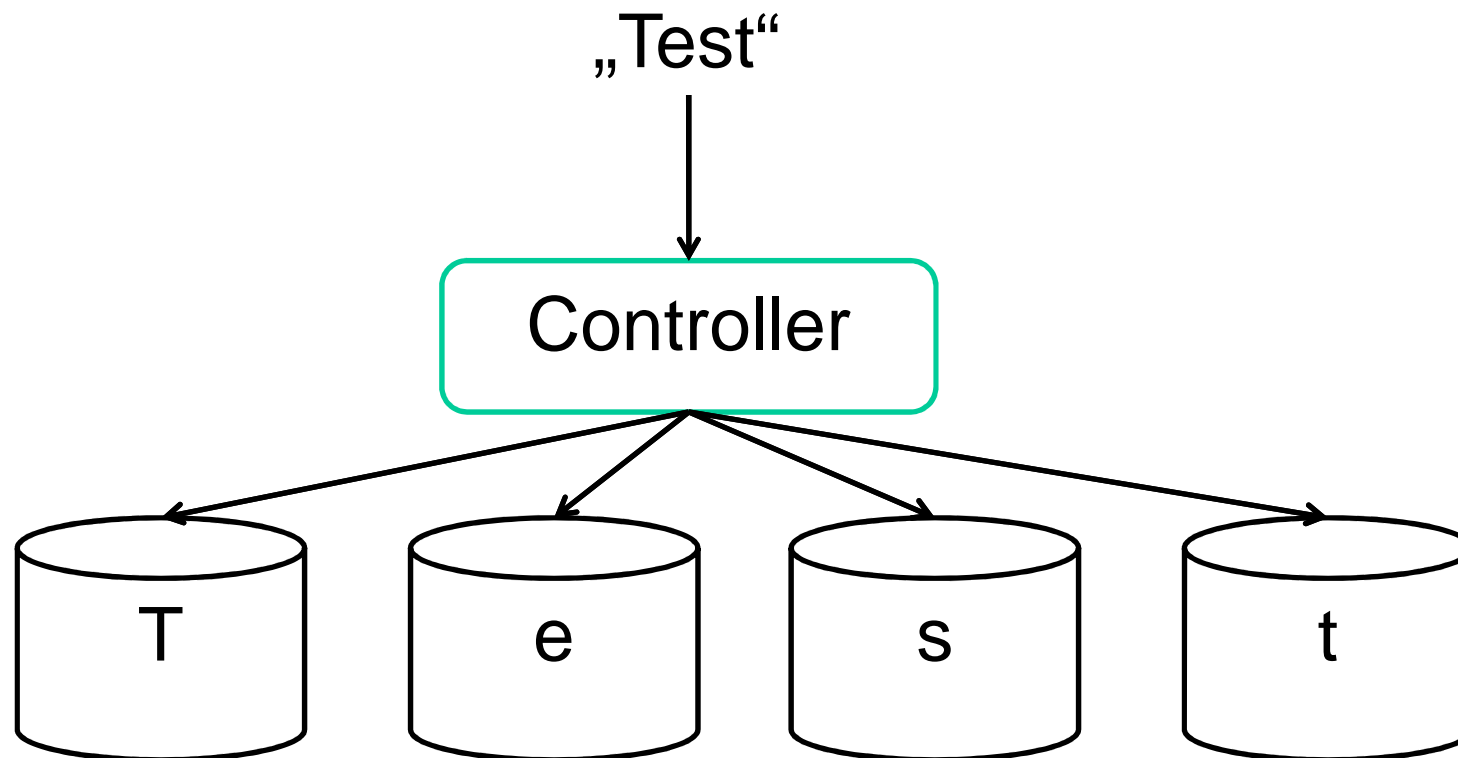
RAID-Levels 0 – 6

(Kemper: RAID-Level 0-5 ;

Elmasri, Navathe: RAID-Level 0-6)

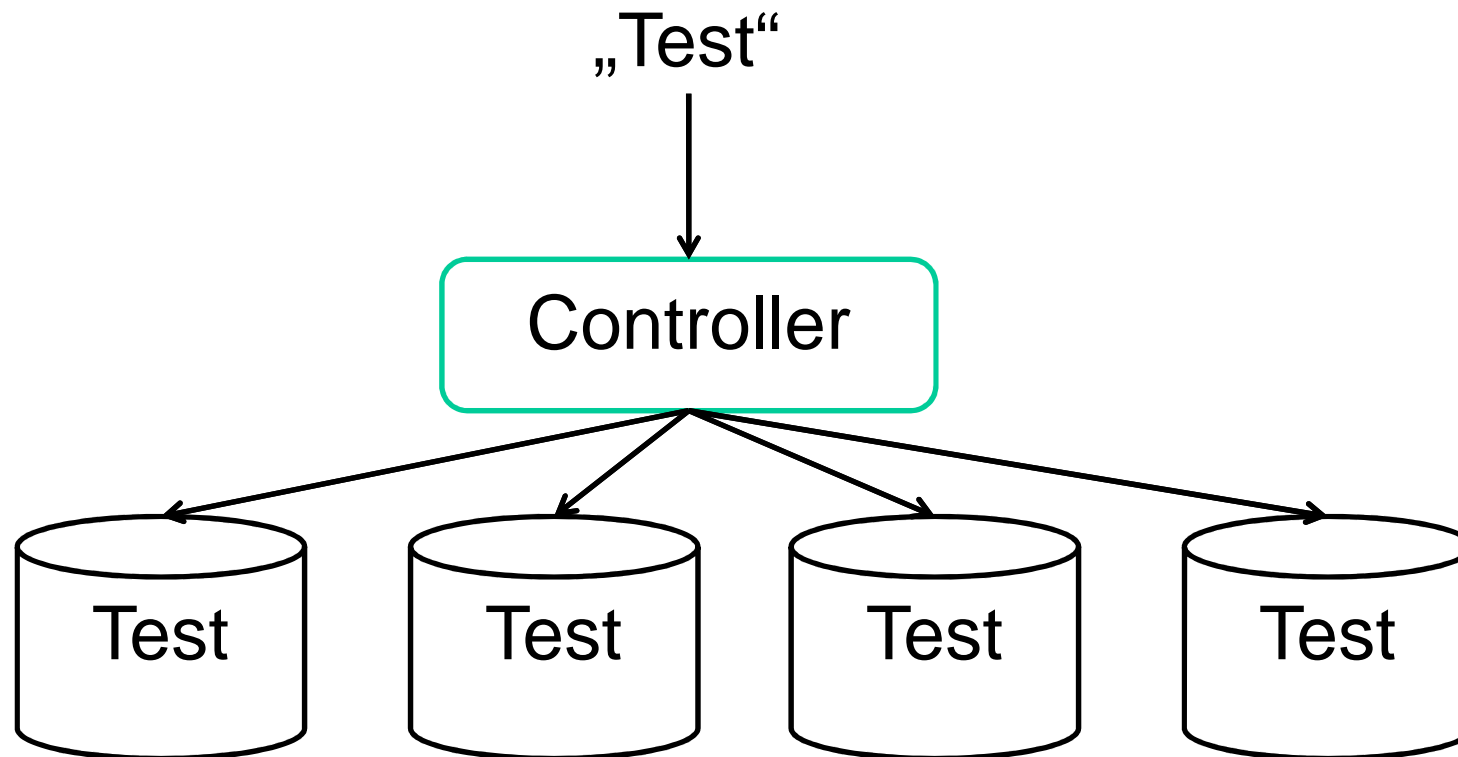
# RAID-Level 0 - Striping

Eigentlich gar kein „R“!



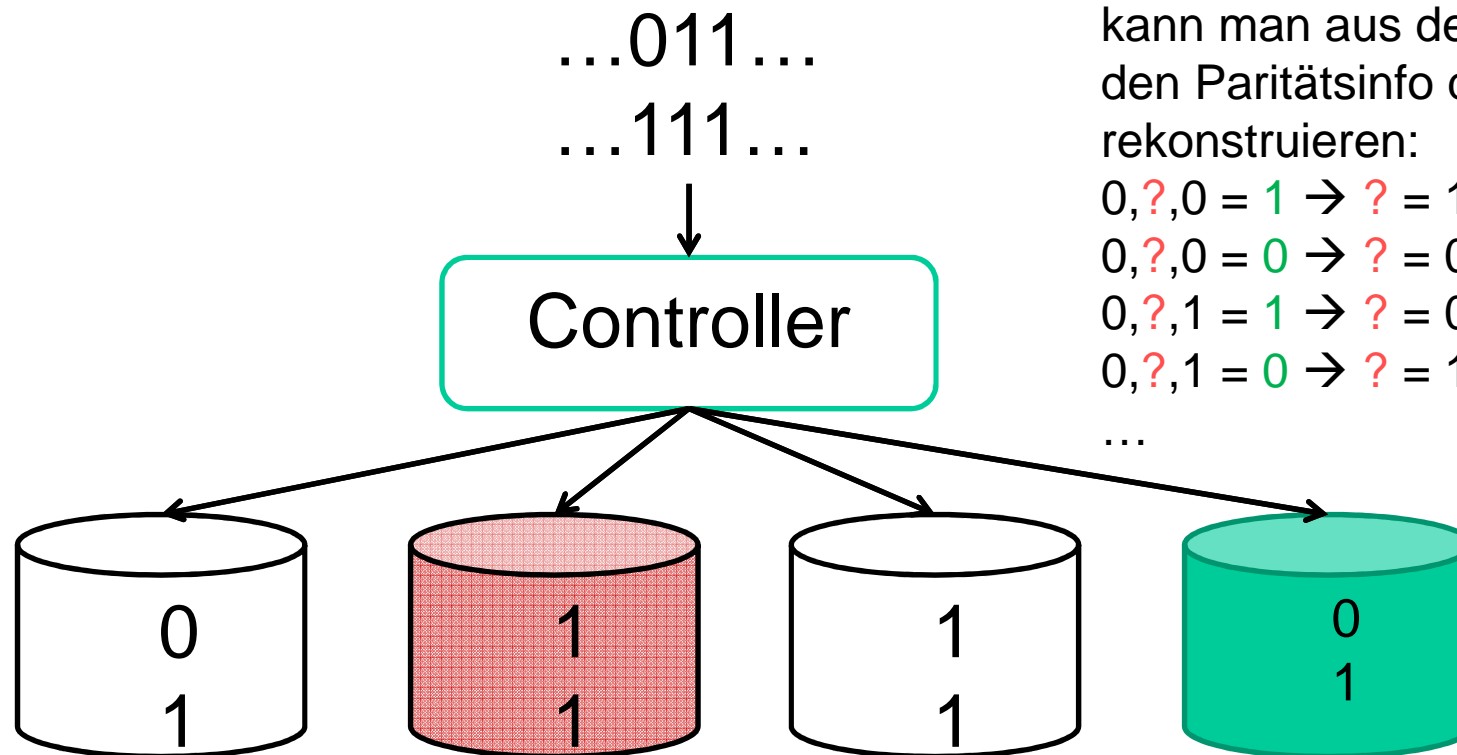
RAID 0 mit 4 Platten

# RAID-Level 1 - Mirroring



RAID 1 mit 4 Platten

# RAID-Level 3 – Prinzip Paritätsinfo



RAID 3 mit 4 Platten

Fällt die rote Platte aus,  
kann man aus den anderen und  
den Paritätsinfo die Daten  
rekonstruieren:

$$0,?,0 = 1 \rightarrow ? = 1,$$

$$0,?,0 = 0 \rightarrow ? = 0,$$

$$0,?,1 = 1 \rightarrow ? = 0,$$

$$0,?,1 = 0 \rightarrow ? = 1,$$

...

Paritätsinfo:

$$\text{XOR}(011) = 0$$

$$\text{XOR}(111) = 1$$

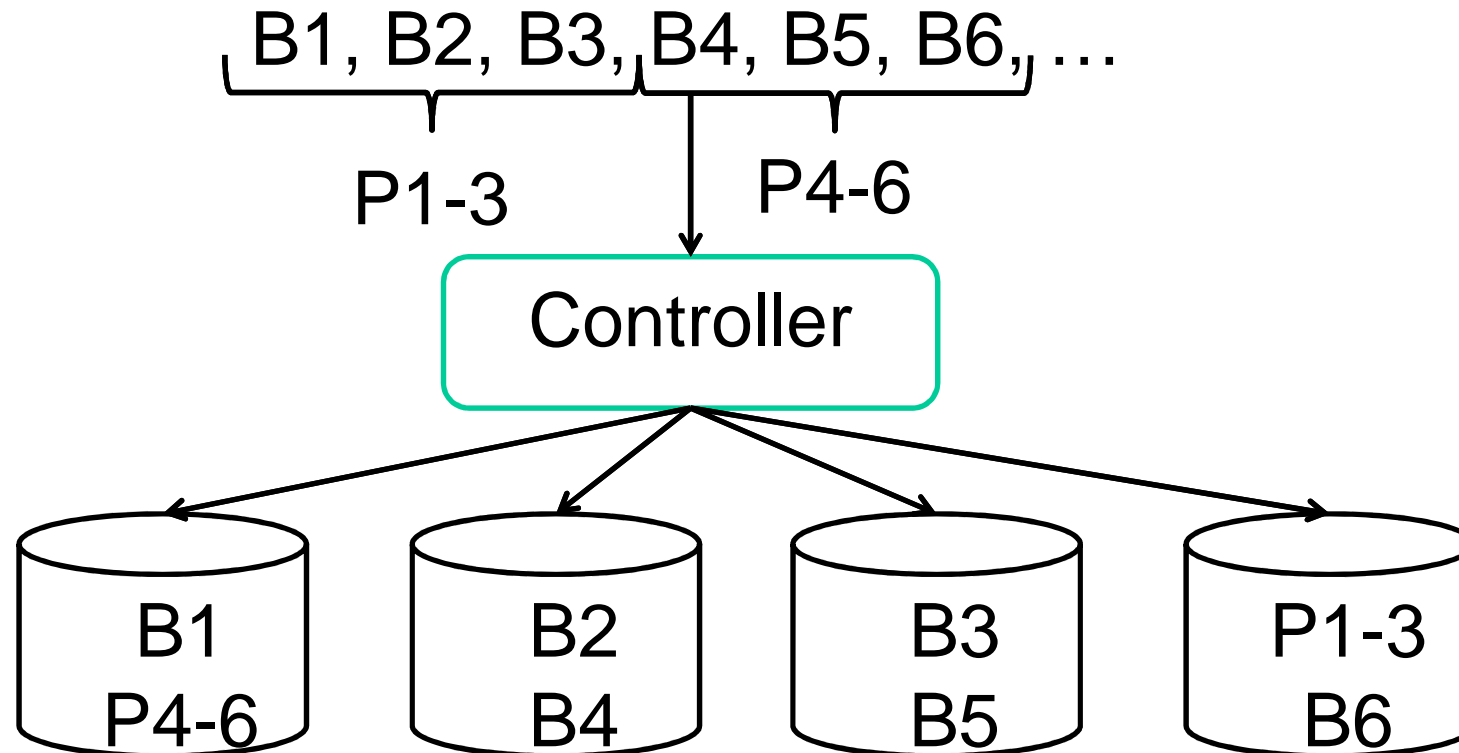
## RAID-Level 2 und 3, 4 (veraltet)

---

- RAID-Level 2: Bit-Level Striping mit Hamming-Code (veraltet)
- RAID-Level 3: Bit-Level Striping mit Paritätsinfo auf extra Platte (veraltet)
- RAID-Level 4: Block-Level Striping mit Paritätsinfo auf extra Platte (veraltet)

## RAID-Level 5

- Block-Level Striping mit verteilter Paritätsinfo



RAID 5 mit 4 Platten

Vermeidet Paritätsplatte als Flaschenhals beim Block-Striping!

## RAID-Level 6

---

- Block-Level Striping mit doppelten Paritätsinfos
  - Fast wie RAID-Level 5, werden zwei verschiedene Paritätsinfos gebildet und gespeichert. Hierdurch können bis zu zwei zwei Platten gleichzeitig ausfallen.



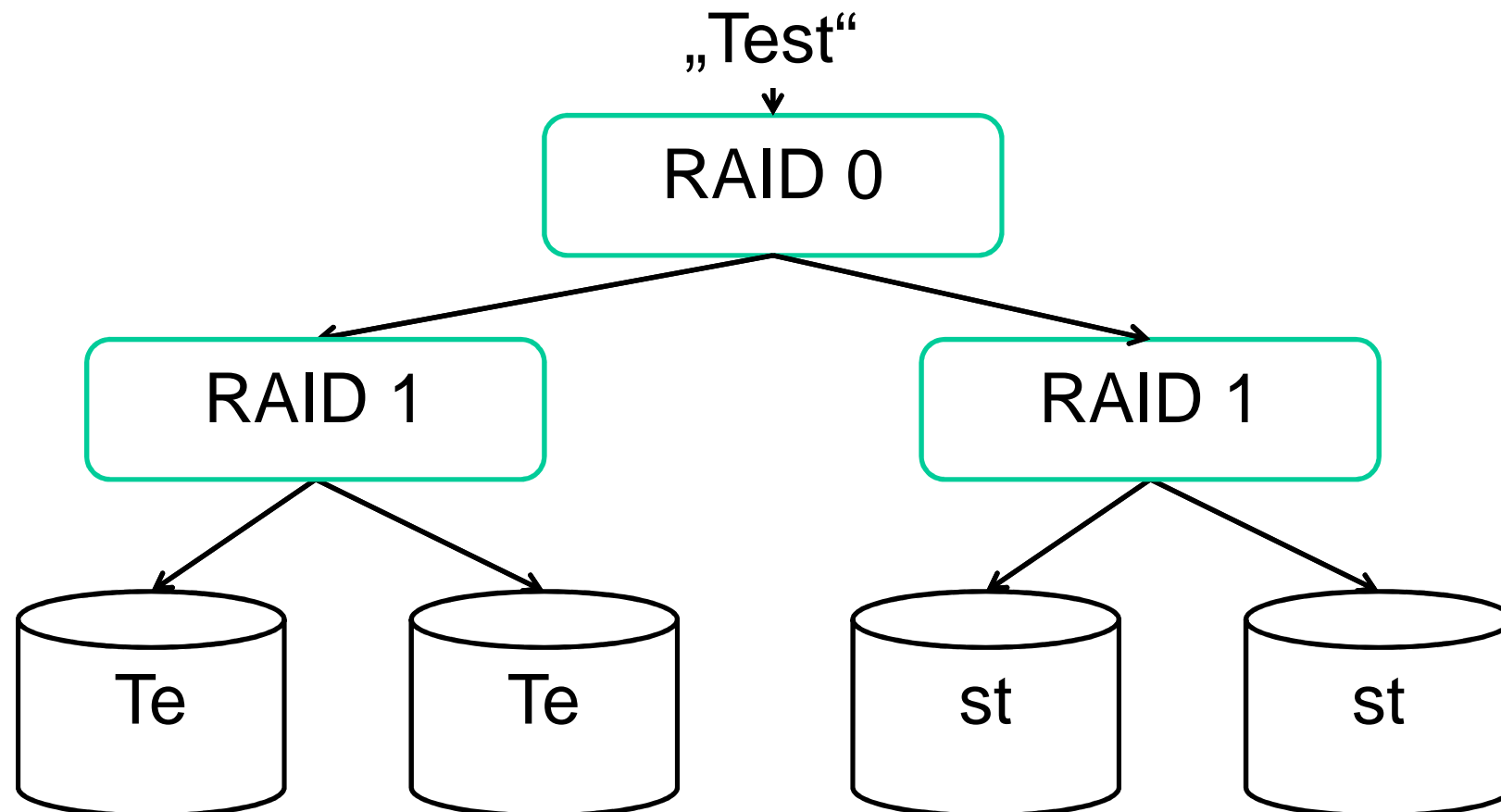
# RAID-Kombinationen

---

- Oft findet man heute Kombinationen der RAID-Level auf zwei Ebenen
  - RAID  $xy \rightarrow$  RAID  $x$  Verbände über ein RAID  $y$  verbunden
  - ... es gibt noch viele weitere spezielle Kombinationen und Entwicklungen für spezielle Zwecke

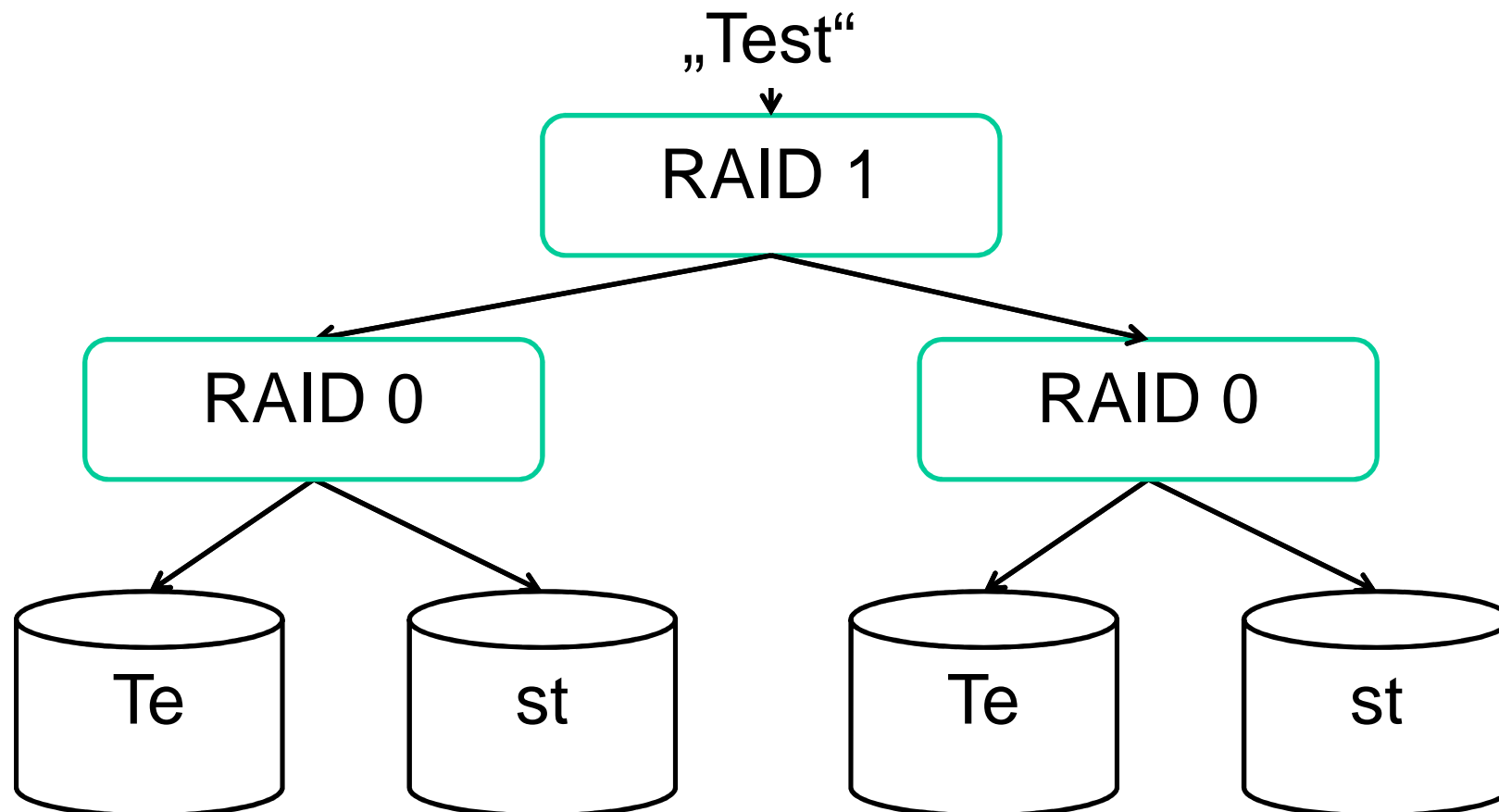
# RAID 10

- RAID 1 Verbünde über ein RAID 0 verbunden



# RAID 01

- RAID 0 Verbünde über ein RAID 1 verbunden



## Folgendes ist bei RAID zu beachten ...

---

- Plattengrößen sind oft nicht egal. → eventuell Platz lassen, sonst kann es Probleme beim Ersetzen der Platten geben.
- Ohne funktionierenden RAID-Controller sind die Daten eventuell nicht lesbar.