Distributed Storage Networks and Computer Forensics

4 Volume Manager and RAID

Christian Schindelhauer

University of Freiburg
Technical Faculty
Computer Networks and Telematics
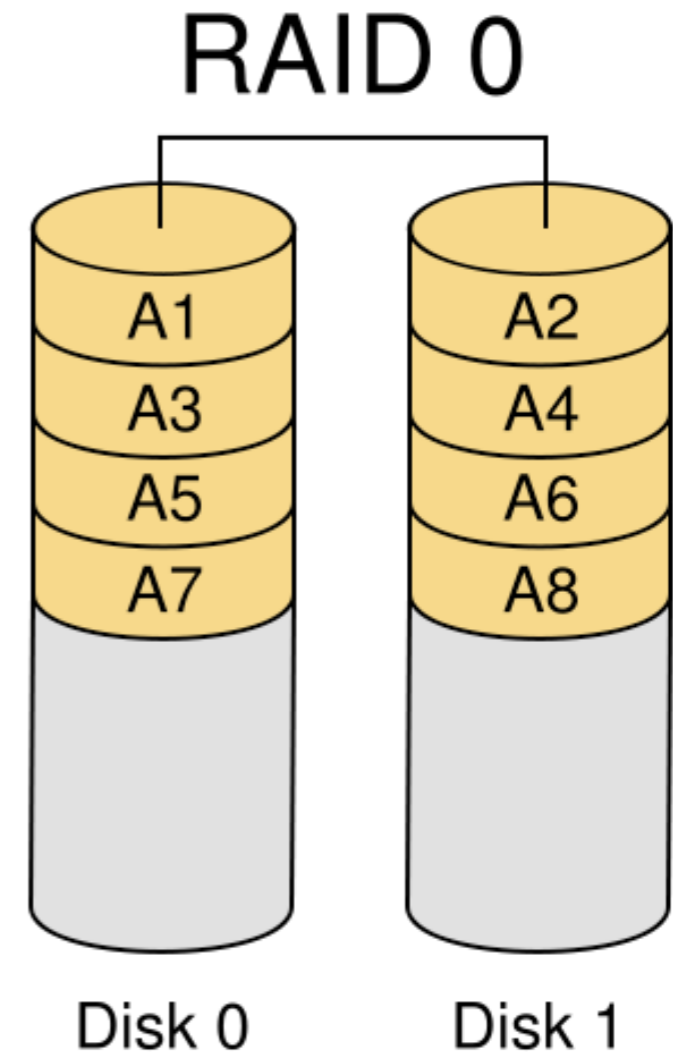Winter Semester 2011/12

CoNe
Freiburg

IIF
INSTITUT FÜR
INFORMATIK
FREIBURG

RAID

‣ **Redundant Array of Independent Disks**

- Patterson, Gibson, Katz, „A Case for Redundant Array of Inexpensive Disks", 1987

‣ **Motivation**

- Redundancy

  - error correction and fault tolerance

- Performance (transfer rates)

- Large logical volumes

- Exchange of hard disks, increase of storage during operation
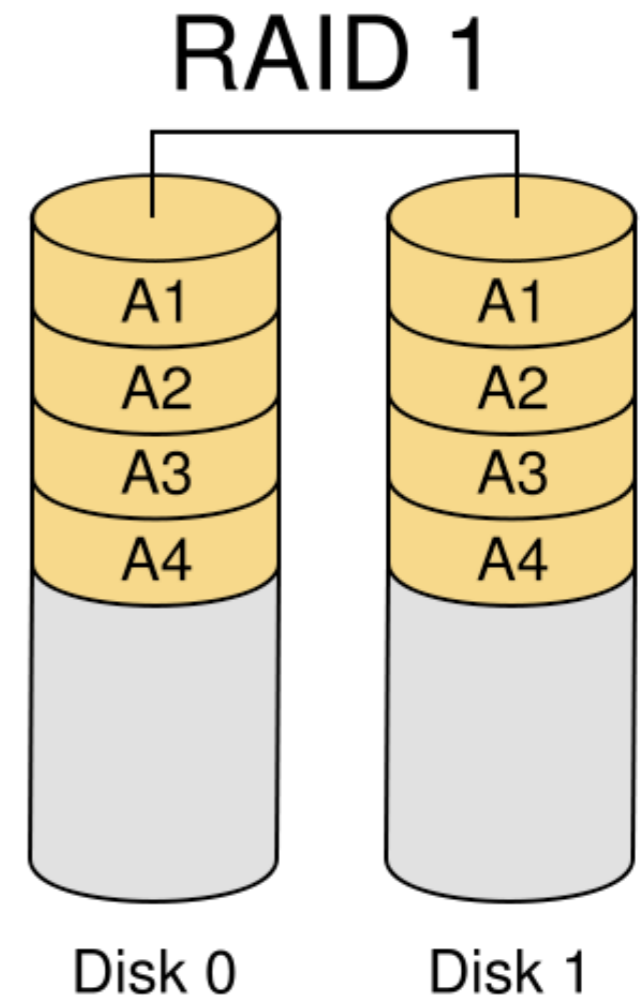
- Cost reduction by use of inexpensive hard disks

Distributed Storage Networks
and Computer Forensics
Winter 2011/12

2

Computer Networks and Telematics
University of Freiburg
Christian Schindelhauer

Raid 0

‣ **Striped set without parity**

  • Data is broken into fragments

  • Fragments are distributed to the disks

‣ **Improves transfer rates**

‣ **No error correction or redundancy**

‣ **Greater disk of data loss**

  • compared to one disk

‣ **Capacity fully available**

RAID 0

A1	A2
A3	A4
A5	A6
A7	A8

Disk 0   Disk 1

http://en.wikipedia.org/wiki/RAID

Distributed Storage Networks
and Computer Forensics
Winter 2011/12

3

Computer Networks and Telematics
University of Freiburg
Christian Schindelhauer

Raid 1

‣ **Mirrored set without parity**
- Fragments are stored on all disks

‣ **Performance**
- if multi-threaded operating system allows split seeks then
- faster read performance
- write performance slightly reduced

‣ **Error correction or redundancy**
- all but one hard disks can fail without any data damage

‣ **Capacity reduced by factor 2**

RAID 1

Disk 0	Disk 1
A1	A1
A2	A2
A3	A3
A4	A4

http://en.wikipedia.org/wiki/RAID

Distributed Storage Networks
and Computer Forensics
Winter 2011/12

4

Computer Networks and Telematics
University of Freiburg
Christian Schindelhauer

RAID 2

- ▸ **Hamming Code Parity**

- ▸ **Disks are synchronized and striped in very small stripes**

- ▸ **Hamming codes error correction is calculated across corresponding bits on disks and stored on multiple parity disks**

- ▸ **not in use**

Distributed Storage Networks
and Computer Forensics
Winter 2011/12

5

Computer Networks and Telematics
University of Freiburg
Christian Schindelhauer

Raid 3

‣ **Striped set with dedicated parity (byte level parity)**

  • Fragments are distributed on all but one disks

  • One dedicated disk stores a parity of corresponding fragments of the other disks
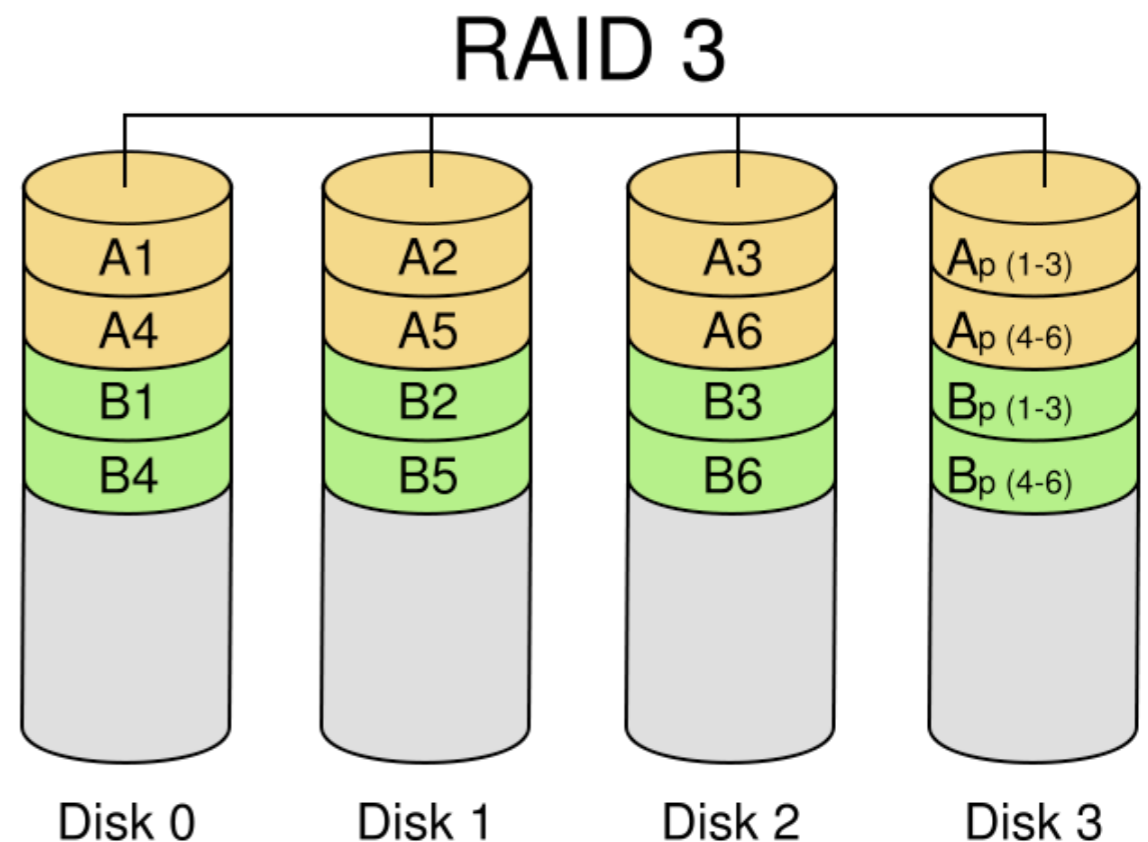
‣ **Performance**

  • improved read performance

  • write performance reduced by bottleneck parity disk

‣ **Error correction or redundancy**

  • one hard disks can fail without any data damage

‣ **Capacity reduced by 1/n**

RAID 3

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

http://en.wikipedia.org/wiki/RAID

Distributed Storage Networks
and Computer Forensics
Winter 2011/12

6

Computer Networks and Telematics
University of Freiburg
Christian Schindelhauer

Raid 4

‣ **Striped set with dedicated parity (block level parity)**

- Fragments are distributed on all but one disks
- One dedicated disk stores a parity of corresponding blocks of the other disks on I/O level
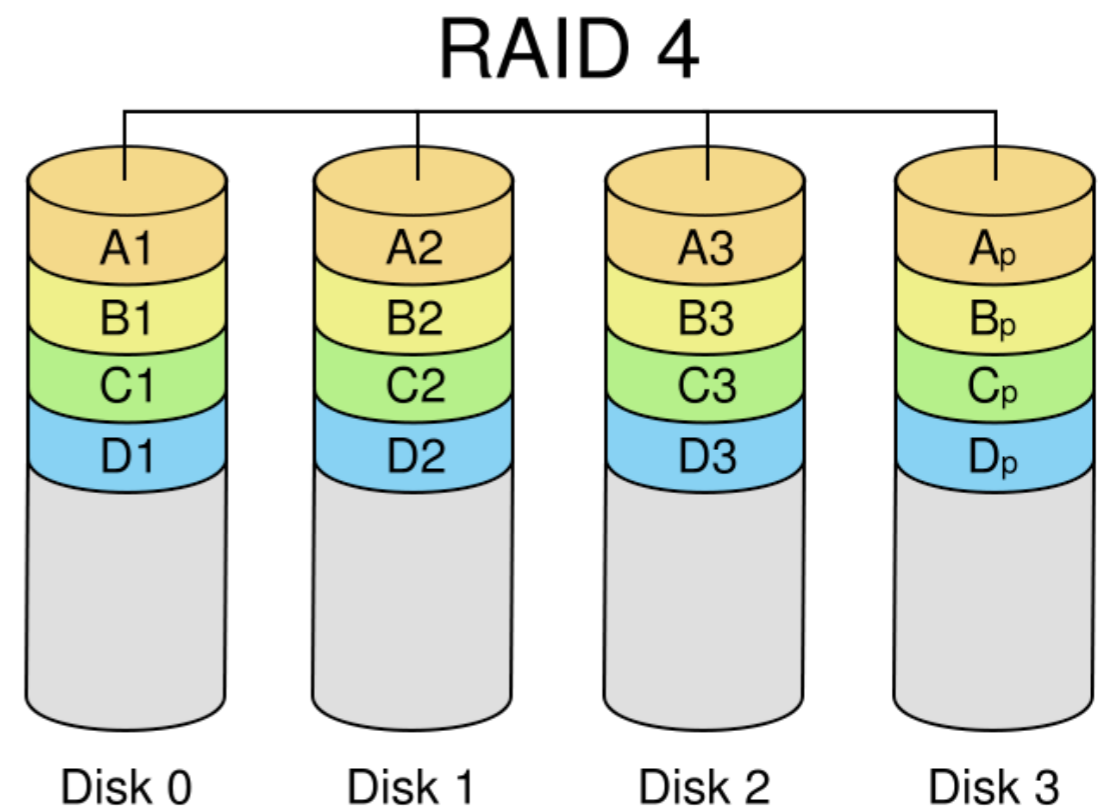
‣ **Performance**

- improved read performance
- write performance reduced by bottleneck parity disk

‣ **Error correction or redundancy**
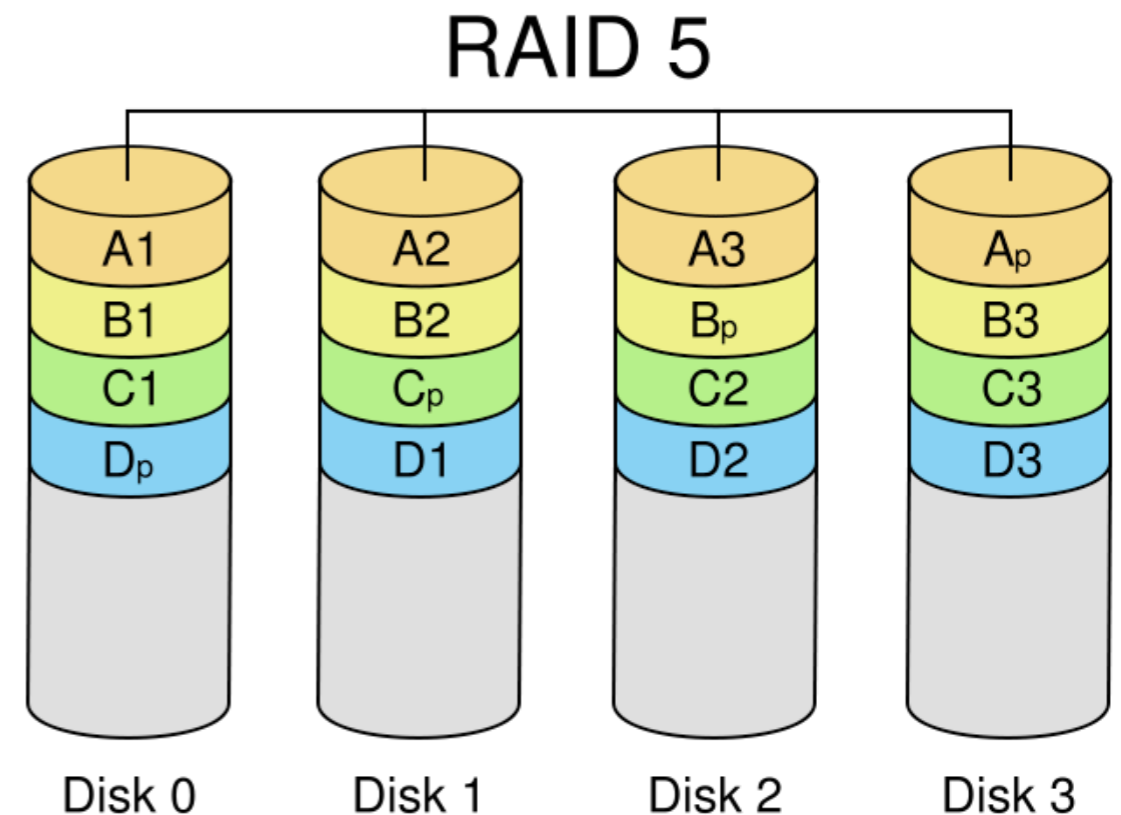
- one hard disks can fail without any data damage

‣ **Hardly in use**

http://en.wikipedia.org/wiki/RAID

Distributed Storage Networks
and Computer Forensics
Winter 2011/12

7

Computer Networks and Telematics
University of Freiburg
Christian Schindelhauer

Raid 5

‣ **Striped set with distributed parity (interleave parity)**

  • Fragments are distributed on all but one disks

  • Parity blocks are distributed over all disks

‣ **Performance**

  • improved read performance

  • improved write performance

‣ **Error correction or redundancy**

  • one hard disks can fail without any data damage

‣ **Capacity reduced by 1/n**

RAID 5

Disk 0	Disk 1	Disk 2	Disk 3
A1	A2	A3	A_p
B1	B2	B_p	B3
C1	C_p	C2	C3
D_p	D1	D2	D3

http://en.wikipedia.org/wiki/RAID

Distributed Storage Networks
and Computer Forensics
Winter 2011/12

8

Computer Networks and Telematics
University of Freiburg
Christian Schindelhauer

Raid 6

‣ **Striped set with dual distributed parity**

- Fragments are distributed on all but two disks
- Parity blocks are distributed over two of the disks
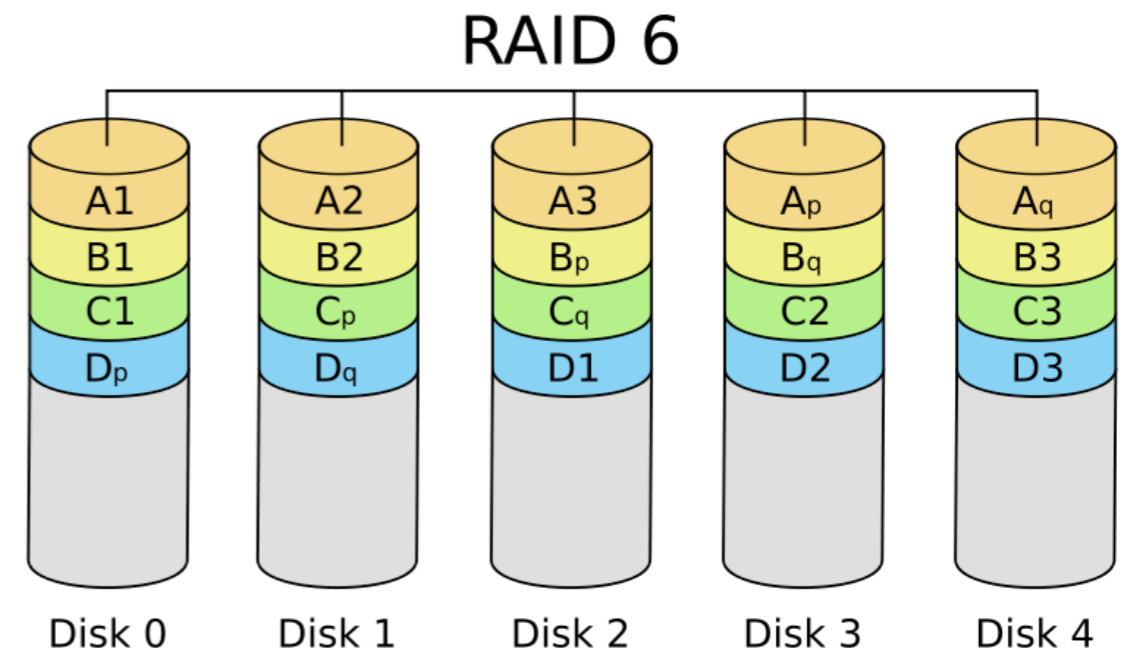 - one uses XOR other alternative method

‣ **Performance**

- improved read performance
- improved write performance

‣ **Error correction or redundancy**

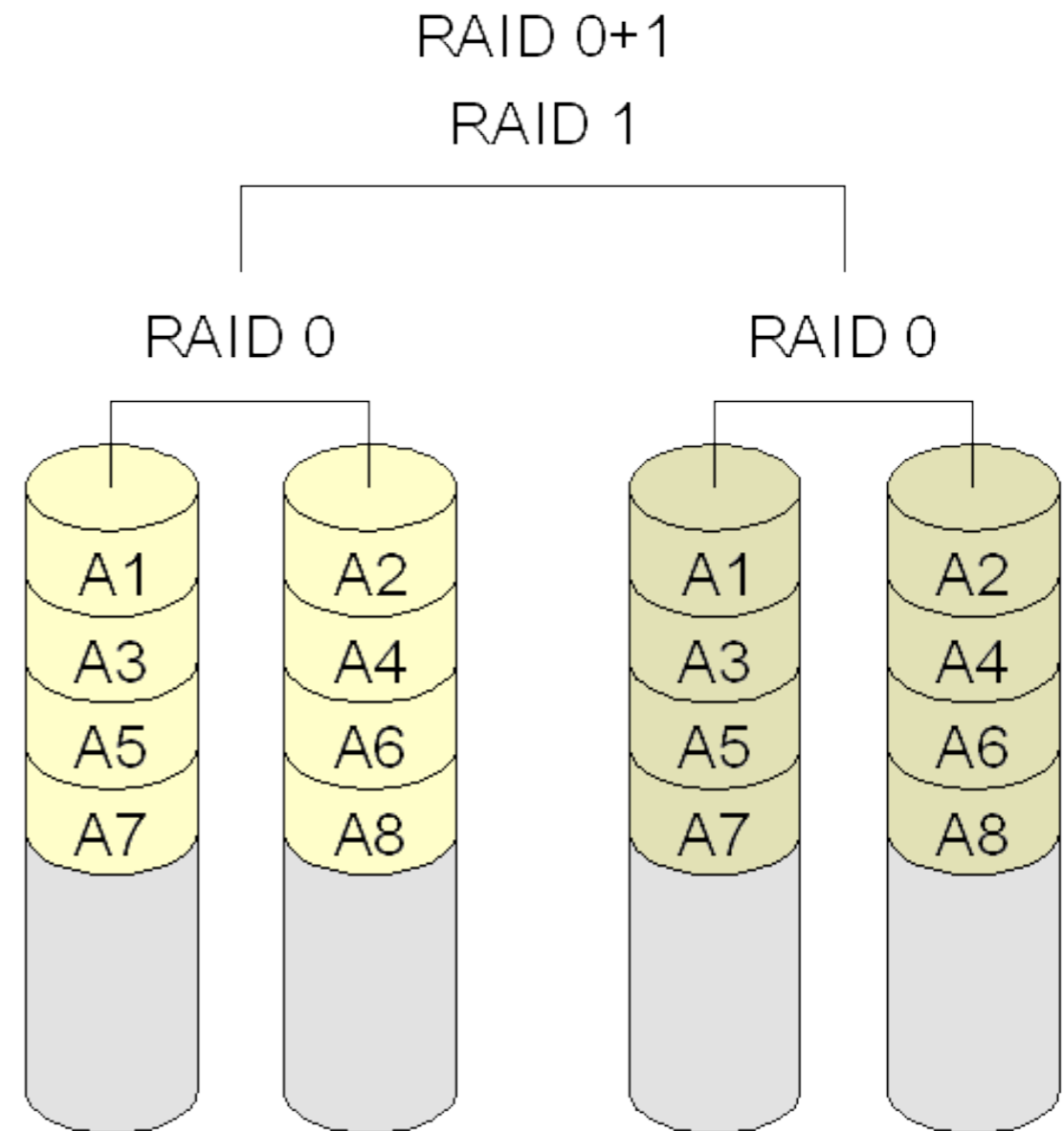- two hard disks can fail without any data damage

‣ **Capacity reduced by 2/n**

RAID 6

Disk 0	Disk 1	Disk 2	Disk 3	Disk 4
A1	A2	A3	Ap	Aq
B1	B2	Bp	Bq	B3
C1	Cp	Cq	C2	C3
Dp	Dq	D1	D2	D3

http://en.wikipedia.org/wiki/RAID

Distributed Storage Networks
and Computer Forensics
Winter 2011/12

9

Computer Networks and Telematics
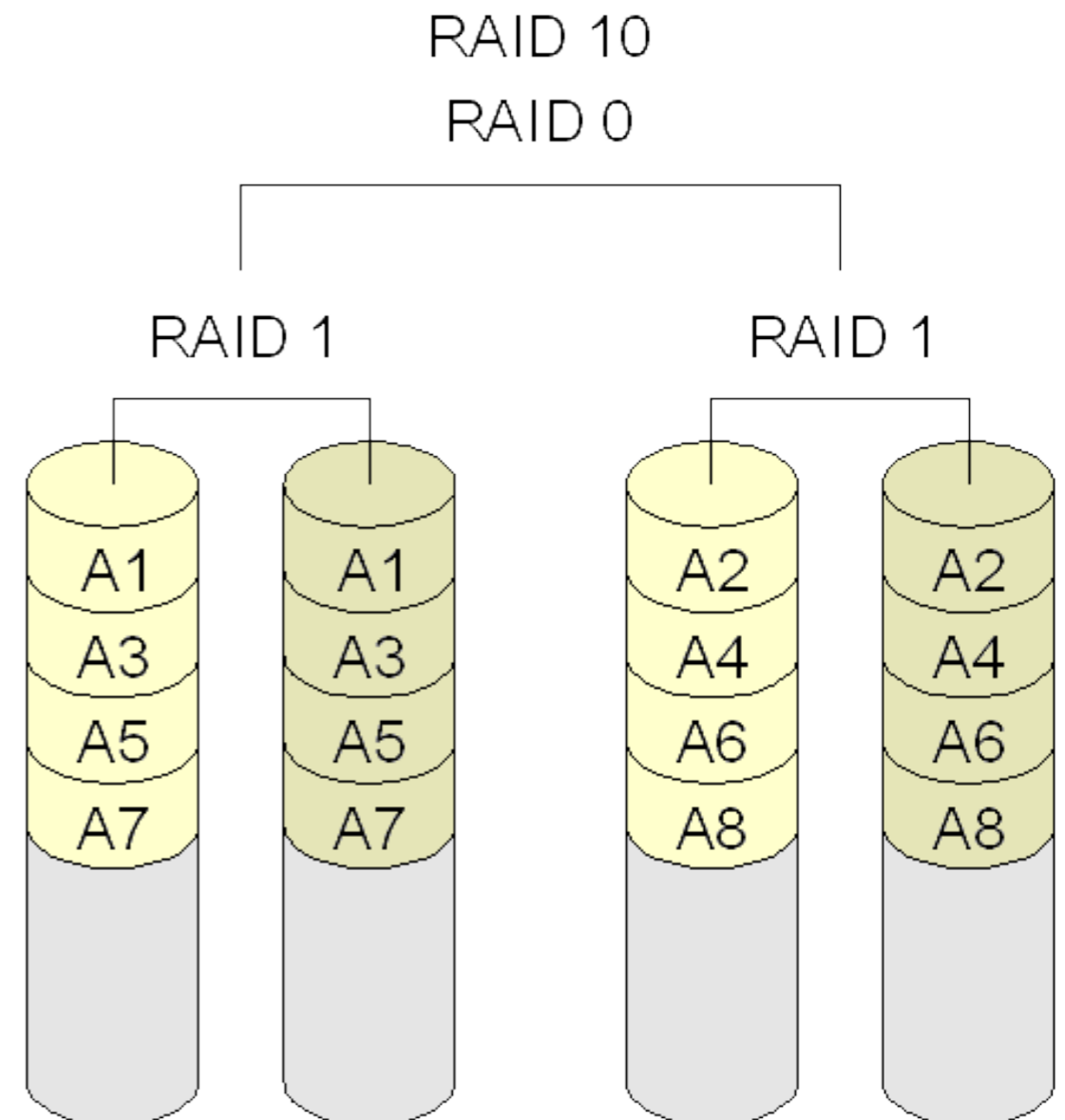University of Freiburg
Christian Schindelhauer

RAID 0+1

‣ **Combination of RAID 1 over multiple RAID 0**

‣ **Performance**

  • improved because of parallel write and read

‣ **Redundancy**

  • can deal with any single hard disk failure

  • can deal up to two hard disk failure

‣ **Capacity reduced by factor 2**

RAID 0+1

RAID 1

RAID 0   RAID 0

A1	A2		A1	A2
A3	A4		A3	A4
A5	A6		A5	A6
A7	A8		A7	A8

http://en.wikipedia.org/wiki/RAID

Distributed Storage Networks
and Computer Forensics
Winter 2011/12

10

Computer Networks and Telematics
University of Freiburg
Christian Schindelhauer

RAID 10

- **Combination of RAID 0 over multiple RAID 1**
- **Performance**
  - improved because of parallel write and read
- **Redundancy**
  - can deal with any single hard disk failure
  - can deal up to two hard disk failure
- **Capacity reduced by factor 2**

RAID 10
RAID 0

RAID 1   RAID 1

A1	A1	A2	A2
A3	A3	A4	A4
A5	A5	A6	A6
A7	A7	A8	A8

http://en.wikipedia.org/wiki/RAID

Distributed Storage Networks
and Computer Forensics
Winter 2011/12

11

Computer Networks and Telematics
University of Freiburg
Christian Schindelhauer

More RAIDs

‣ **More:**

- RAIDn, RAID 00, RAID 03, RAID 05, RAID 1.5, RAID 55, RAID-Z, ...

‣ **Hot Swapping**

- allows exchange of hard disks during operation

‣ **Hot Spare Disk**

- unused reserve disk which can be activated if a hard disk fails

‣ **Drive Clone**

- Preparation of a hard disk for future exchange indicated by S.M.A.R.T

Distributed Storage Networks
and Computer Forensics
Winter 2011/12

12

Computer Networks and Telematics
University of Freiburg
Christian Schindelhauer

Volume Manager

‣ **Volume manager**

- aggregates physical hard disks into virtual hard disks
- breaks down hard disks into smaller hard disks
- Does not provide files system, but enables it

‣ **Can provide**

- resizing of volume groups by adding new physical volumes
- resizing of logical volumes
- snapshots
- mirroring or striping, e.g. like RAID1
- movement of logical volumes
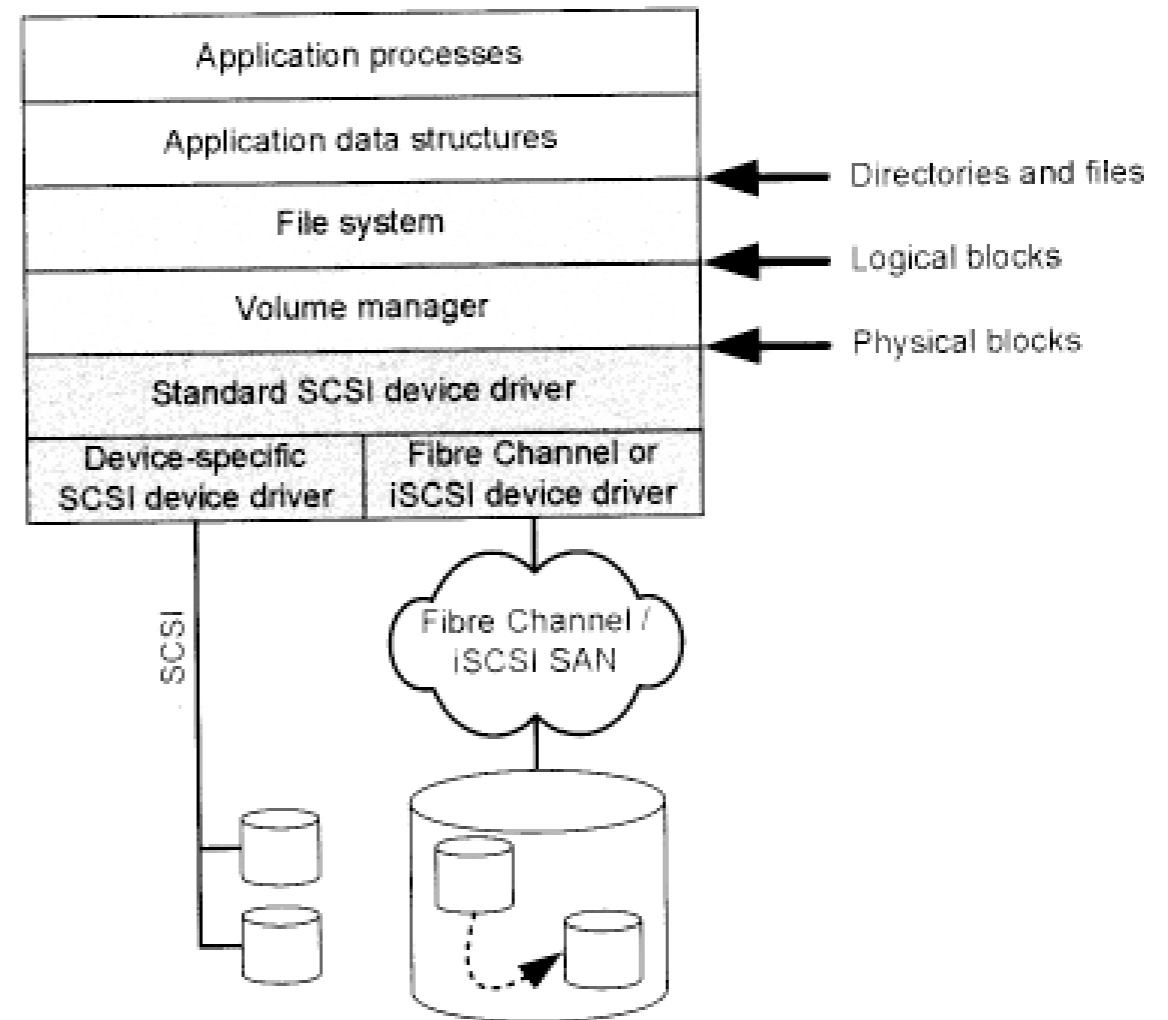
From: Storage Networks Explained, Basics and Application of Fibre Channel SAN, NAS, iSCSI and InfiniBand, Troppens, Erkens, Müller, Wiley

Overview of Terms

‣ **Physical volume (PV)**
- hard disks, RAID devices, SAN

‣ **Physical extents (PE)**
- Some volume managers splite PVs into same-sized physical extents

‣ **Logical extent (LE)**
- physical extents may have copies of the same information
- are addressed as logical extent

‣ **Volume group (VG)**

- logical extents are grouped together into a volume group

‣ **Logical volume (LV)**
- are a concatenation of volume groups
- a raw block devices
- where a file system can be created upon

Distributed Storage Networks
and Computer Forensics
Winter 2011/12

14

Computer Networks and Telematics
University of Freiburg
Christian Schindelhauer

Distributed Storage Networks and Computer Forensics

4 Volume Manager and RAID

Christian Schindelhauer

University of Freiburg
Technical Faculty
Computer Networks and Telematics
Winter Semester 2011/12

CoNe
Freiburg

IIF
INSTITUT FÜR
INFORMATIK
FREIBURG