Redundant Array of Independent Disks

Yashwant K. Malaiya
Redundant Array of Independent Disks (RAID)

• Enables greater levels of performance and/or reliability

• How? By concurrent use of two or more ‘hard disk drives’.

• How Exactly?

• Striping (of data), Mirroring, Error correction techniques
Hard Disks

• Rotate: one or more platters
• Efficient for blocks of data (sector 512 bytes)
• Inherently error prone
• CRC to check for errors
• Need a controller
• Can fail completely
Standard RAID levels

• RAID 0: striping
• RAID 1: mirroring
• RAID 2: bit-level striping, Hamming code for error correction (not used anymore)
• RAID 3: byte-level striping, parity (rare)
• RAID 4: block-level striping, parity
• RAID 5: block-level striping, distributed parity
• RAID 6: block-level striping, distributed double parity
RAID 0

- Data striped across n disks
- Read/write in parallel
- No redundancy.

\[ R_{sys} = \prod_{i=1}^{n} R_i \]

- Ex: 3 year disk reliability = 0.9 for 100% duty cycle, n = 14
- \( R_{sys} = (0.9)^{14} = 0.23 \)
RAID 1

- Disk 1 mirrors Disk 0
- Read/write in parallel
- One of them may be used as backup.

\[ R_{sys} = \prod_{i=1}^{n} [1 - (1 - R_i)^2] \]

- Ex: 3 year disk reliability = 0.9 for 100% duty cycle. \( n = 7 \) pairs
- \( R_{sys} = (2 \times 0.9 - (0.9)^2)^7 = 0.93 \)
RAID 2

- Used Hamming code check bits as redundancy
- Obsolete
**RAID 3**

- Byte level striping
- Dedicated parity disk
- If one fails, its data can be reconstructed using a spare

\[ R_{sys} = \sum_{j=n-1}^{n} \binom{n}{j} R_j^j (1 - R_i)^{n-j} \]

- Ex: 3 year disk reliability = 0.9 for 100% duty cycle. \( n = 13, j = 12, 13 \)
- \( R_{sys} = 0.62 \)
RAID 4

- Block level striping
- Dedicated parity disk
- If one fails, its data can be reconstructed using a spare

\[
R_{sys} = \sum_{j=n-1}^{n} \binom{n}{j} R_j^j (1 - R_i)^{n-j}
\]

- Ex: 3 year disk reliability = 0.9 for 100% duty cycle. \( n = 13, j = 12, 13 \)
- \( R_{sys} = 0.62 \)
RAID 5

- Distributed parity
- If one disk fails, its data can be reconstructed using a spare

\[ R_{sys} = \sum_{j=n-1}^{n} \binom{n}{j} R_j^j (1 - R_i)^{n-j} \]

- Ex: 3 year disk reliability = 0.9 for 100% duty cycle. \( n = 13, j = 12, 13 \)
- \( R_{sys} = 0.62 \)
RAID 6

- Distributed double parity
- If one disk fails, its data can be reconstructed using a spare
- Handles data loss during a rebuild

\[
R_{sys} = \sum_{j=n-2}^{n} \binom{n}{j} R_j^j (1 - R_i)^{n-j}
\]

- Ex: 3 year disk reliability = 0.9 for 100% duty cycle. \(n = 13, j = 11, 12, 13\)
- \(R_{sys} = 0.87\)
Nested RAID Levels

- RAID 01: mirror of stripes
- RAID 10: stripe of mirrors
- RAID 50: block-level striping of RAID 0 with the distributed parity of RAID 5 for individual subsets
- RAID 51: RAID5 duplicated
- RAID 60: block-level striping of RAID 0 with distributed double parity of RAID 6 for individual subsets.
RAID 10

- Stripe of mirrors: each disk in RAID0 is duplicated.

\[ R_{sys} = \prod_{i=1}^{ns} [1 - (1 - R_i)^2] \]

- Ex: 3 year disk reliability = 0.9 for 100% duty cycle. \( ns = 6 \) pairs,
- \( R_{sys} = 0.94 \)
RAID 01

- Mirror of stripes: Complete RAID0 is duplicated.

\[ R_{sys} = \left[ 1 - (1 - \prod_{i=1}^{ns} R_i)^2 \right] \]

- Ex: 3 year disk reliability = 0.9 for 100% duty cycle. ns = 6 for each of the two sets,
- \( R_{sys} = 0.78 \)
RAID 50
RAID 51
## RAIDS Comparison

<table>
<thead>
<tr>
<th>Level</th>
<th>Space efficiency</th>
<th>Fault tolerance</th>
<th>Read performance</th>
<th>Write performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>none</td>
<td>nx</td>
<td>nx</td>
</tr>
<tr>
<td>1</td>
<td>1/2</td>
<td>1 drive</td>
<td>2x</td>
<td>x</td>
</tr>
<tr>
<td>2</td>
<td>&lt;1</td>
<td>1</td>
<td>var</td>
<td>var</td>
</tr>
<tr>
<td>3</td>
<td>&lt;1</td>
<td>1</td>
<td>(n-1)x</td>
<td>(n-1)x</td>
</tr>
<tr>
<td>4</td>
<td>&lt;1</td>
<td>1</td>
<td>(n-1)x</td>
<td>(n-1)x</td>
</tr>
<tr>
<td>5</td>
<td>&lt;1</td>
<td>1</td>
<td>(n-1)x</td>
<td>(n-1)x</td>
</tr>
<tr>
<td>6</td>
<td>&lt;1</td>
<td>2</td>
<td>(n-2)x</td>
<td>(n-2)x</td>
</tr>
<tr>
<td>10</td>
<td>1/2</td>
<td>1/set</td>
<td>nx</td>
<td>(n/2)x</td>
</tr>
</tbody>
</table>