Why Speed is Not the Most Important Thing to Look For in a RAID System

SMPTE Technology Webcast Series Sponsored by:

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

SMPTE Technology Webcast Sponsors

• Thank you to our sponsor for their generous support:

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)
SMPTE Technology Webcasts

- Series of monthly 60- to 90-minute online, interactive webcasts covering a variety of technical topics
- Free professional development benefit for SMPTE members
- Sessions are recorded for member viewing convenience.

Views and opinions expressed during this SMPTE Webcast are those of the presenter(s) and do not necessarily reflect those of SMPTE or SMPTE Members.

This webcast is presented for informational purposes only. Any reference to specific companies, products or services does not represent promotion, recommendation, or endorsement by SMPTE.
Your Host

Joel E. Welch
Director of Education
SMPTE

Today’s Guest Speaker

Timothy Standing
VP Software Engineering - Mac
Other World Computing, Inc.
SoftRAID:

22 Years of Software RAID for Mac OS

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Understanding RAID Levels
Minimize the Number of Disks Which Fail
Worst Case Scenarios

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)
What is **RAID**?

**Redundant Array of Independent Disks**

---

Protecting Volumes from Disk Failure

**RAID Levels with no protection:**
- RAID 0 (stripe)

**RAID Levels which protect by data duplication:**
- RAID 1 (mirror)
- RAID 1+0

**RAID Levels which protect using parity:**
- RAID 4 and 5
- RAID 6 and 6+
RAID Levels with no protection:

How does RAID 0 (stripe) work?

RAID Levels Which Protect with Data Duplication:

How does RAID 1 (mirror) work?
RAID Levels Which Protect with Data Duplication (continued):

How does RAID 1+0 work?

4 pairs of mirrored disks in 2 enclosures: best performance and safety

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

RAID Levels Which Protect with Parity:

How do RAID 4 and 5 work?

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)
RAID Levels Which Protect with Parity (continued):

- 1 Parity Disk
- Data protected from failure of 1 disk
- Can detect data corruption
- Cannot fix data corruption

RAID 4 and 5

- 2 Parity Disks
- Data protected from failure of 2 disks
- Can detect data corruption
- Can fix data corruption from a 1 unreliable disk

RAID 6 and 6+

Hardware vs. Software RAID

**Hardware RAID:**
- Works on any OS
- Volume speed not dependent on CPU

**Software RAID:**
- Requires separate software for each OS
- Better disk testing and disk health monitoring
- Volume speed dependent on number CPU cores
- Able to combine speed of multiple ports
- More future proof
What about ZFS?

- Large memory footprint
- Volume meta-data held in memory for long periods of time
- Requires ECC RAM
- Great for NAS solutions
- Not suitable for volumes attached directly to a workstation

The Speed of Different RAID Levels:

RAID 0 vs RAID 1+0

<table>
<thead>
<tr>
<th></th>
<th>4 SATA HDDs</th>
<th>8 SATA HDDs</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAID 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAID 1+0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)
The Speed of Different RAID Levels:

HDDs vs SSDs

![Graph comparing RAID 0 and RAID 1+0 for 8 SATA HDDs and 8 SATA SSDs](image)

The Speed of Different RAID Levels:

RAID 0 on 8 M.2 Blades

![Graph comparing RAID 0 performance on 8 M.2 Blades](image)
Tim Standing, Firefighter

Tim Standing
Mendocino Volunteer Fire Department
2003 - 2006

Minimize the Number of Disks Which Fail

• Learn from the experience of others

• Use the best tools available
Learn from the experience of others

Hard Drive Failure Rates by Manufacturer

All drive sizes for a given manufacturer are combined

https://www.backblaze.com/blog/hard-drive-reliability-stats-q1-2016/

Learn from the experience of others

Hard Drive Failure Rates by Manufacturer and Model

Published each Quarter

https://www.backblaze.com/blog/hard-drive-stats-for-q2-2018/
Use the best tools available

Test your drive before first using it
Monitor your drive once it's in use

Disk Failure Rates

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)
Probable cause of Early Disk Failures

New

bad sector

1 Month
Probable cause of Early Disk Failures

<table>
<thead>
<tr>
<th>Time</th>
<th>Bad Sectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>New</td>
<td></td>
</tr>
<tr>
<td>1 Month</td>
<td></td>
</tr>
<tr>
<td>2 Months</td>
<td></td>
</tr>
<tr>
<td>6 Months</td>
<td></td>
</tr>
</tbody>
</table>

Disk reports errors when reading and writing from bad sector

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)
How can you test a disk before using it?

1. Write to every sector on a disk
2. Read each sector back
3. Verify that the data read is what was written out
4. Random access testing to stress disk thermally

Certifying a disk with SoftRAID

Certifying a disk in SoftRAID: writing a random pattern to every sector on a disk, then reading it back to verify the pattern
Certifying a disk with terminal commands

Step 1:
Write a random pattern to every sector on the disk

```bash
> Last login: Fri Sep 30 16:36:10
> #!/bin/bash
> status=0
> index=0
> echo "creating 1 GB test file"
> rm /tmp/random_file
> head -c 1000000000 /dev/urandom > /tmp/random_file
> status=$?
> echo "test file created"
> > while [ $status == 0 ]
> > do
> >   path="/Volumes/""$1""/testfile "$index
> >   echo "copying test file to $path"
> >   cp /tmp/random_file "$path"
> >   status=$?
> >   let index=index+1
> > done
> > echo "volume $1 full"
> >
> © 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)
```

Certifying a disk with terminal commands

Step 2:
Read the pattern back and verify it

```bash
> Last login: Fri Sep 30 16:36:10
> #!/bin/bash
> status=0
> index=0
> path="/Volumes/""$1""/testfile "$index
> > while [ -f "$path" ]
> > do
> >   echo "verifying file at $path"
> >   cmp /tmp/random_file "$path"
> >   status=$?
> >   let index=index+1
> >   path="/Volumes/""$1""/testfile "$index
> > done
> >
> © 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)
```
Disk Failure Rates After Certifying

Failure Rate vs. Time (years)

Uncertified disk failure rates

Certified disk failure rates

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)
Use the best tools available

✓ Test your drive before first using it
Monitor your drive once it's in use

SMART for Disks
Self Monitoring, Analysis and Reporting Technology

SMART Test:
• Pass/Fail test which a disk can perform on itself to see if it is operating correctly
• Performed when requested by attached computer
• Limited use for disk health monitoring

SMART Attributes:
• List of operating parameters which the disk reports back to the computer
• Examples, current operating temperature, number of power on hours, etc.
• Out of over 80 parameters, only 3 are useful for disk health monitoring
Study by Google of 100,000 disk drives over 8 months

2007

- Checked SMART attributes every 24 hours for every disk in their server pods
- Noticed correlation between certain SMART Attributes and future disk failure

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Meanings</th>
<th>Acceptable Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Number of reallocated sectors</td>
<td>0</td>
</tr>
<tr>
<td>197</td>
<td>Unreliable sector count</td>
<td>0</td>
</tr>
<tr>
<td>198</td>
<td>Uncorrectable error count</td>
<td>0</td>
</tr>
</tbody>
</table>


Backblaze uses SMART to predict disk failure

- Checked SMART attributes every 24 hours for every disk in their server pods
- Noticed correlation between certain SMART Attributes and future disk failure

[https://backblaze.com/blog/hard-drive-smart-stats](https://backblaze.com/blog/hard-drive-smart-stats)
Disk which is predicted to fail by SoftRAID

Note: This disk passes the SMART test!

Predicting disk failure on Mac OS

Software which uses SMART to predict disk failure: Drive DX, SMARTReporter, SoftRAID, SoftRAID Lite
Disk Failure Rates
Predicted and Unpredicted

Unpredicted Disk Failure Rates

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)
Disk Failures Avoided

- Test your drive before first using it
- Monitor your drive once it's in use

SSD Failure Rates
Flash Memory Usage in an SSD

<table>
<thead>
<tr>
<th>available storage</th>
<th>extra</th>
</tr>
</thead>
<tbody>
<tr>
<td>960 GB</td>
<td>70 GB</td>
</tr>
</tbody>
</table>

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)
Flash Memory Usage in an SSD

- Available storage: 960 GB
  - Extra: 70 GB
- Available storage: 960 GB
  - Extra: 35 GB
- Available storage: 960 GB
  - Extra: 0 GB
Predicting Failure of SSDs

- SSDs fail catastrophically and without warning
- All SSDs which we have seen fail have a Media Wearout Indicator greater than 80%
- No reliable way of predicting failure

Worst Case Scenarios

Events that happen to 1 in 1,000 or 1 in 10,000 users

- We don't think they can happen to us
- When they do, they can destroy our project or even our business
Case 1: Cirina Catania

Founder of The Catania Group

- Documentary Filmmaker
- Featured photographer with National Geographic
- 8 years as a senior executive with MGM and United Artists
Case 1: Cirina Catania

Cirina's post-catastrophe desktop:

all disks used for RAID volumes had been initialized and re-labeled: **My Book**

both hardware and software RAID volumes were destroyed

Case 2: Ryan Francis

Editor and Co-founder of NakNed Design

- Commercial and Music Videos

- Clients including: LandRover, Chick-Fil-A, Mirabella Makeup and TR Golf
Case 2: Ryan Francis

What we can learn from Worst Case Scenarios

Remember these important rules!

- Have 2 offsite backups of all critical data
- Alternate which backup you use
- Never connect all backup disks to a computer at the same time
- When creating volumes or certifying disks, always use a computer which has no critical data
Run a Fire Drill for your Data

Scenarios to think about:
• Imagine that your computer and all your disks disappear
• Imagine that your office burns to the ground
• Imagine that your computers are infected with ransomeware

Questions to ask yourself:
• Do you have all your important work stored off-site?
• How long will it take you to restore to a new computer and disks?
• What happens if you lose your computer and all your disks during a backup?

(Verbal) Questions?

Timothy Standing
VP Software Engineering - Mac
Other World Computing, Inc.

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)
SMPTE Technology Webcast Sponsors

• Thank you to our sponsor for their generous support:

AJA VIDEO SYSTEMS

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)