

Failure Panel

NSF HEC FSIO Workshop

Narasimha Reddy
Texas A&M University

Current Trends

- SSDs gaining increased acceptance and deployment
- SSD sizes in 80GB and above
- SSD costs about x8-x10 more than disk
- PCM and other NV memories becoming available
 - Displacing SSDs or achieving mass market status requires considerable investment

Likely Outcomes

- SSDs and other NV memories will find place in the hierarchy
- Working sets will fit in SSDs for most applications
- During the application run, most I/O can be done to SSDs
- May need occasional swaps of data from SSDs to disks

Likely Outcome scenario 1

- Disks can sleep most of the time
 - Failures due to spin up and spin down
 - Automatic scrubbing during normal access absent
 - On average, longer scrubbing time will imply higher latent media errors
 - What should be the scrubbing frequency, parity protection levels with reduced activity at disk?

Likely Outcome scenario 1

- SSDs active most of the time
 - SSDs have higher media error rates than disks
 - Lifetime of cells and quality of data
 - Recent work is showing scrubbing like activity will improve lifetimes
- Media errors and scrubbing will become more dominant in the future

Likely outcome Scenario 2

- We will use SSDs and disks in a collaborative way instead of a hierarchy
 - Asymmetric read/write advantages
 - Asymmetric small/large I/O advantages
- Recent work shows possibly better organizations than strict hierarchy
 - Equalize access times
 - Improved throughput/latency by simultaneous use of all the devices (Wu & Reddy MASCOTS 2010), even without performance asymmetry

Likely outcome Scenario 2

- All the devices active all the time
- May not be ideal for power/energy
- May be required for performance
- Worry about SSD failures in addition to disks
 - SSD arrays?
 - Aggravate SSD write problems
 - Write parity on to disk (FAST 2010)
 - Mixed failure models of SSDs and disks

Likely Outcome Scenario 2

- Independent and dependent failures of different devices
- Small I/Os on SSDs and large I/Os on disks
 - Should we protect different devices differently?
- Mostly read data on SSDs, writes on disks
 - Again what should be levels of protection?

Likely outcome scenario 3

- Disk becomes the new tape
- What can we learn from 30 years ago?
 - When disk and tape systems were prevalent
- Dusty tapes could not be trusted
 - Multiple copies on tapes
- Multiple copies on disks?
 - Disk bandwidth could become a problem
 - Google doesn't seem to have this problem
 - Exhaust disk capacity quicker
- Provide on-disk parity across multiple blocks?