

## Lecture 15

*Lecturer: Emery Berger**Scribe: Bryan McCaffery*

## 15.1 Raid

- RAID stands for Redundant Array of Independent Disks.
- The idea is to store data on multiple disks to increase performance and redundancy.
- Normally disks are limited to by writing bandwidth, by increasing the number of disks and striping the data you can get around a n way speed up for n disks.
  - Striping is the process of writing parts of the data to multiple different disks.
- As the number of disks increase the probability of any one disk failing also exponentially increases.
  - If the probability of any one disk failing is 1 percent,
    - \* The probability 10 disks do not fail is approximately 90 percent
    - \* The probability 100 disks do not fail is around 36.6 percent
    - \* Probability any fails  $p(\text{any}) = 1 - (1 - (1 - 0.01))^n$
- Max Reliability but minimum parallelism(slowest) is one disk
- Max Parallelism(fastest) but minimum reliability is N disks
- Levels of Raid
  - RAID 1 - Mirroring of Data
  - RAID 2 - Hamming Code
  - RAID 3 - Byte level striping, plus level II functionality
  - RAID 4 - Block level interleaving/striping
  - RAID 5 - Block level interleaving/striping and check information
- Recovery From Failures
  - Raid has issues recovering from failures
  - If a disk totally dies, all information has to be retrieved from replicated disks
  - There is a high chance of failure as millions of blocks could have to be rewritten
  - In many cases, back ups are kept in tape as it is easier to just recover from a backup than replicate data from disks

## 15.2 Robustness

- SSD's
  - SSD's have wear. The crystals can only reshape themselves so many times until they can no longer reshape themselves
  - They keep track of this and do not put things in the bad blocks of memory
- N Variant Systems
  - Have three different machines running program that should produce the same output
  - If two of the machines get same output, and third does not use the value that two machines got
  - Used by NASA
  - These machines if on the same spacecraft, can not be truly independent as they are subject to same environmental conditions
- Failures are not normally Independent of Each Other
  - Paper by Google discovered unusual ways things were not independent
  - Things as simple as buying all hard drives at the same time from same factory can lead to data centers having multiple failures around the same number of cycles
  - Contractors hired to right similar programs independent of each other commonly find similar solutions
- Failure Oblivious Computing
  - In the case of a bad address/seg fault put value in map with bad address pointing to value
  - In the case of read from a bad address, retrieve value from map