RAID

Chapter 38

Previously in CS212…

- We talked about Hard disk drives
- We looked at some ways of determining performance of a single disk
- We learned that large contiguous writes are the best choice for I/O performance
- Individual drives are limited

Redundant Array of Inexpensive Disks (RAID)

- Special purpose hardware complete with its own memory, processor, and firmware dedicated to the task of managing a collection of hard disks
- Better performance
  - We can perform parallel I/O operations across multiple disks
- More capacity
  - We can use the disks together to increase our overall capacity, but treat it like one disk within the OS
- Reliability
  - Spreading the data over multiple disks can help avoid data loss (especially with some form of redundancy)

Measures to Evaluate RAID

- Capacity
  - While we have multiple disks, some raid configurations do not give us access to the complete storage space
- Reliability
  - How many disk faults/failures can we tolerate before we cannot recover?
- Performance
  - With more disks comes the opportunity for parallelism
  - However, some configurations have more overhead as well

RAID Level 0: Striping

- Data block are spread across all the disks in **stripes**
  - Not really a RAID level as there is **NO REDUNDANCY**

- Multiple data blocks written consecutively before switching disks is referred to as a **chunk**
  - Smaller chunks means greater intra-file parallelism, but requires more positioning time
  - Larger chunks means less intra-file parallelism, needs less positing time, but multiple requests (inter-file) can yield high throughput

- We have access to the complete storage space of all disks, and receive performance nearly equivalent to a single disk

RAID Level 1: Mirroring

- Each time we write a block to disk, we copy that content to another disk
  - Content is both mirrored and striped across the disk
- Can read from either copy of the data, but writes require both copies be updated
  - Suffers worst case seek and rotational delays
- Available capacity is halved (everything is duplicated)
- Can tolerate one disk failure, and possibly more (if you are lucky)
- With the duplication and striping, we get roughly half of the bandwidth for sequential reading
  - Random reads fair better with the duplication allowing for more distribution among the disks

RAID Level 4

- Data striping with one disk reserved for storing the parity of the stripe
- Can withstand the loss of any ONE block from the stripe
- Parity is calculated using **XOR**
  - Even number of 1s in the bits, parity is 0
  - Odd number of 1s in the bits, parity is 1
  - This means that for any stripe there is an even number of 1s (including the parity bit)
- We lose one whole storage disk to the parity data
  - We can tolerate ONLY ONE disk of failure
- Performance is limited by the need to update the parity disk
  - With the subtractive method, each write is two reads and two writes
    - Read Old Block Data, Read Parity Data, Write new data, Write new parity data
  - Small writes are the largest bottle neck

RAID Level 5

- Performance is like RAID 4 as is its failure tolerance
- Instead of using one disk for parity bits, it rotates the parity blocks across all disks
- The big difference is that with parity blocks across all the drives, requests can be parallelized and are not bottlenecked by a single parity drive (like with RAID 4)
- That big difference is why RAID 5 mostly supplemented RAID 4

So, which to use?

- RAID 0
  - + Best performance and capacity
  - - no fault protection

- RAID 1
  - + Best random I/O performance and fault protection
  - - Half the storage of what you paid for

- RAID 5
  - + Best balance of capacity and reliability
  - - Poor performance with small write workloads

Next Time

- We investigate file systems and directories in the OS

Have a great break!