Lecture 2: Storage Devices

I/O problems

- How should we store the multimedia file?
- How should we layout the file in the storage system?
- What are the issues that affect the storage and transmission?
- How can we take care of timing?
- How should we consider fault tolerance?

RAID Technology

- Redundant Array of Inexpensive Disk (RAID).
- view a set of disk drives as one logical drive.
- data is distributed across the set of drives
- Proposed by IBM Almaden, UC Berkeley gave the taxonomy.

RAID 0: Disk Striping

Based on spreading of data across multiple drives connected to a single disk controller. The data being written to disk is broken into segments. If segment size is 512 bytes and block size is 4 KB. First segment goes to the first disk, second segment to the second disk and so on. Performance gain, parallel read/write.

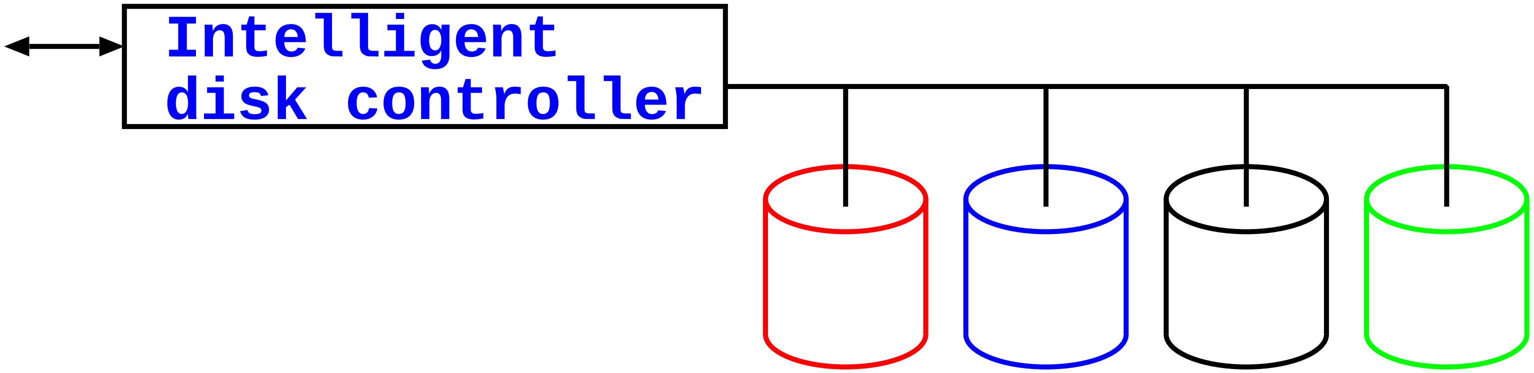

Figure 1: Hardware view of RAID

RAID 1: Disk Mirroring

RAID 1 deals with fault tolerance. Each copy of the file is written on two separate drives. Write takes twice as long while read operations can be optimized. The overhead for this case is the major storage overhead.

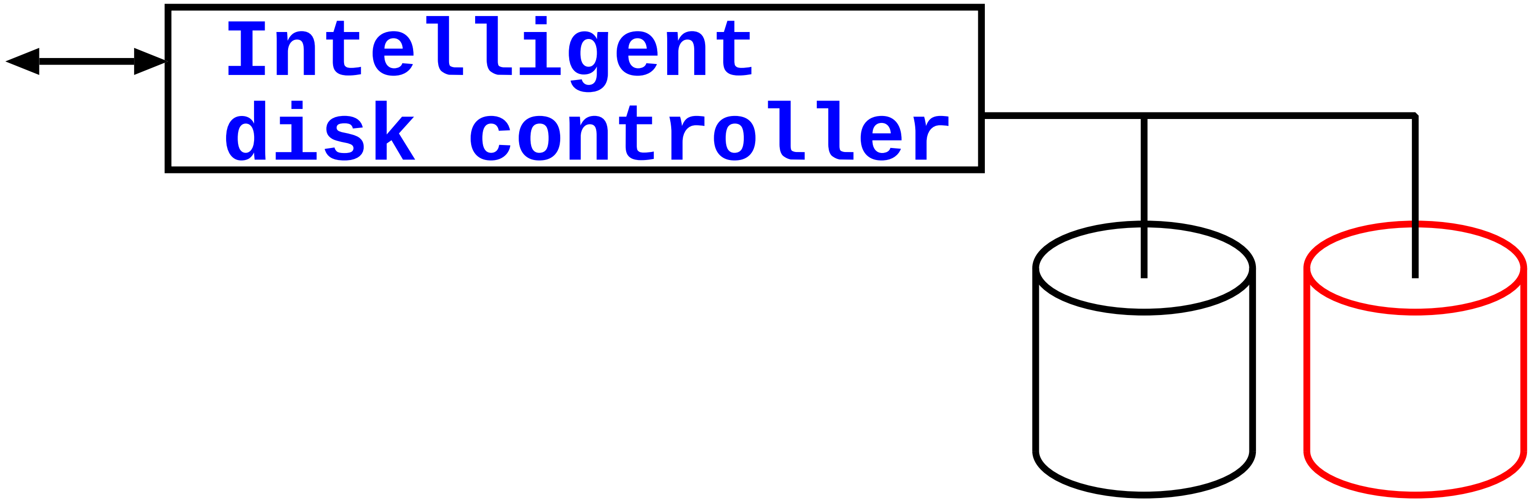

Figure 2: Disk Mirroring

RAID 2: Bit interleaving

Data is written one bit at a time and is interleaved across multiple drives. In RAID 2, drive spindles are synchronized as a single I/O operation accesses all drives. This scheme has a large transfer rate (sum of transfer rate of all disks). This scheme is poor for small read/write.

RAID 3: Parallel Disk Array

Introduce concept of parity and interleave data at a bit level across several drives. There is a single drive *dedicated* to store the parity information generated for disks in the array group. When a disk fault is detected, the parity disk can be used to reconstruct the affected information.

RAID 4: Sector interleaving

Similar to RAID 3 but striping is done at a sector or block level across several disks. The system can handle multiple read/write to multiple disks but the parity disk can be the bottleneck, unless it is a large write across multiple disks. The storage overhead is $\frac{1}{n}$ where n is the number of disks.

RAID 5: Block Interleaving

Similar to RIAD 4 but the parity block is distributed across multiple drives. Therefore, the workload of writing parity is distributed across all disks.

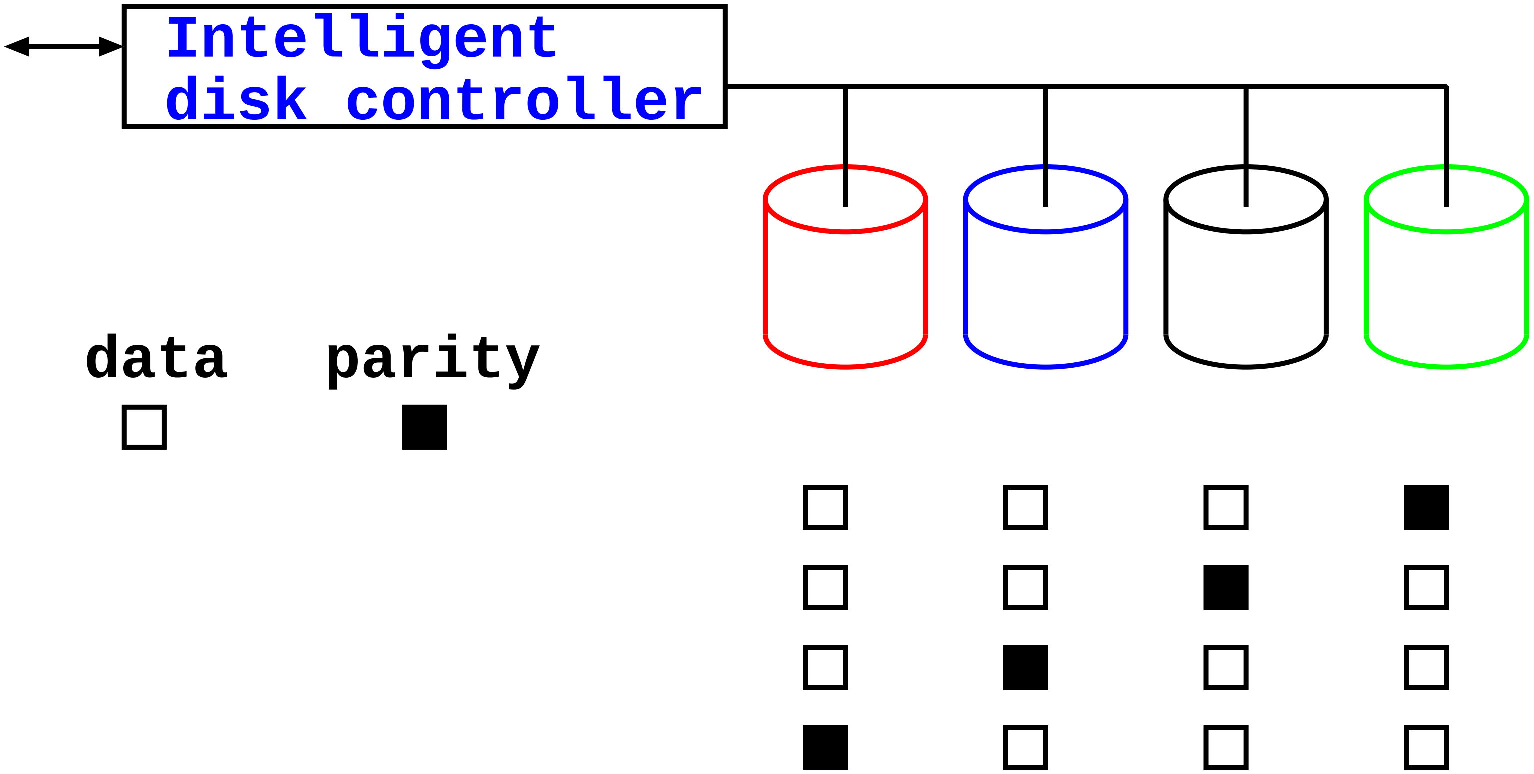

Figure 3: RAID 5

RAID 6 Consider two dimension disk arrays with addition fault tolerance. Parity is generated for rows and for columns of disks in the array group.

RAID 7 embedded real-time OS into disk array architecture. Improved on data recovery.

Issues, how groups should be designed? How to reconstruct data on a failed disk so as to reduce the mean time to failure?