# DATA STORAGE TECHNOLOGIES & NETWORKS

## (CS C446, CS F446 & IS C446)

### LECTURE 22– STORAGE

# eRAID

- spin down partial or entire mirror group to standby state to save energy
  - read requests are served by data copies in the primary disks.
  - write requests to the standby disks are deferred in controller cache or active disks, and then flushed to the standby disks after they are spun up.
- Saves up to 32% energy

- Spin down disks to save energy
  - a balance needs to be made for providing enough active disks to meet performance based service quality.
- eRAID deploys a time- window based performance control scheme to do the tradeoffs between energy-saving and performance (response time and throughput) degradations.
- To predict the performance of storage system, first forecast the load features in next time window (e.g., request inter-arrival rate).
  - use the load parameters as input, feed them into a performance prediction model to make the prediction.

# EERAID – Energy Efficient RAID

- Takes advantage of redundant information
- EERAID 1 achieves up to 30% energy saving
- EERAID 5 achieves up to 11% energy saving
- No performance degradation
- Energy efficient, redundancy aware I/O request scheduling and RAID controller cache management
  - Assumption: Non volatile write back cache is available in RAID controller

- New policy: **Windows Round-Robin**
  - alternatively dispatch every N successive requests to the primary group and the mirror group
    - the idle period of disks of the other group (called idle group) is largely stretched

# Power and Redundancy-Aware Flush (PRF)

Replacement strategy

- when cache controller looks for victims to flush dirty cache lines
  - It selects a cache line that belongs to the group to which RAID controller is dispatching read requests.

# Remote Mirroring

- To protect against the failure of a disk subsystem
  - Example fire will destroy the data and data copies
  - Need to store data copy remotely
- Remote monitoring operation is handled by the 2 participating subsystems
  - Invisible to application servers and does not consume their resources
- Synchronous remote mirroring
  - First disk subsystem sends data to second disk subsystem before acknowledging write complete
  - Data is always up-to-date
  - Increases response time. Slight performance degradation
- Asynchronous remote mirroring
  - acknowledging a write command immediately
  - Preferable if the mirrored disk is in distant place

# Logical Unit Number [LUN]

- Physical drives or groups of RAID protected drives can be logically split in to volumes known as Logical Units
- Use of Logical Units improve disk utilization
- Logical Units are spread across all the physical disks that belong to that set
- Need to allocate only the required space in disk, not the entire disk
  - Host requirement is 100GB but HDD size is 1T
    - Without LUNs 1T allocated to host
    - With LUNs 100GB allocated to host

# Logical Unit Number [LUN]

- Each LU has unique ID called a LUN
- LUNs hide the organization and composition of the RAID set from hosts
- Conventional storage provisioning method of creating LUNs is referred to as thick LUNs
  - To distinguish from LUNs created by virtual provisioning method
- If LUNs are assigned to a non-virtualized host, a bus scan is required to identify the LUN
  - LUN appears as a raw disk to the OS
  - Formatted with a file system and FS is mounted