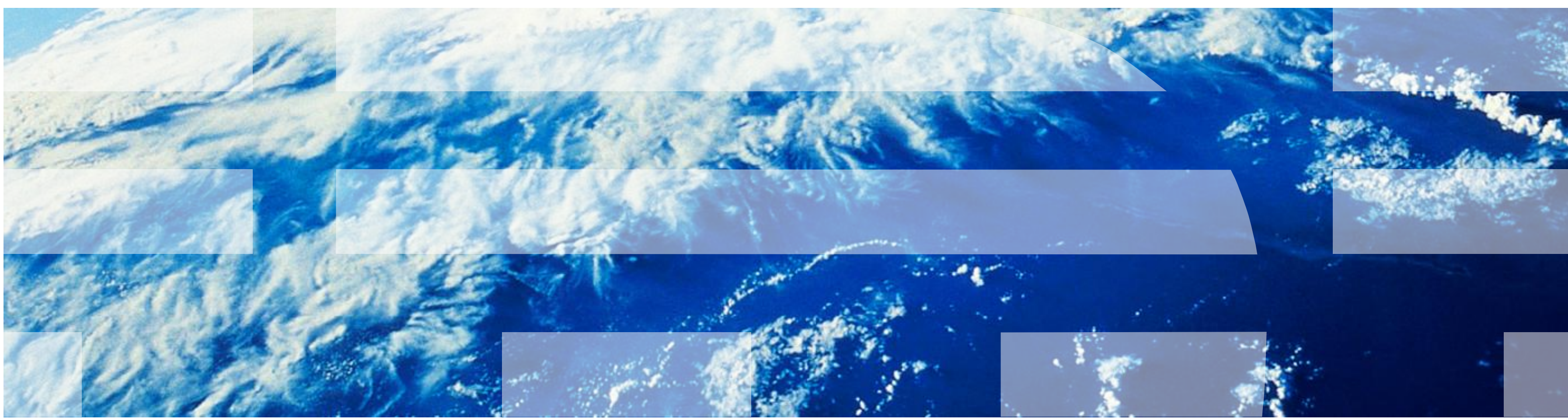
26th IEEE (MSST2010) Symposium on Massive Storage Systems and Technologies
May 2010

IBM

Leveraging Disk Drive Acoustic Modes for Power Management

Doron Chen, George Goldberg, Roger Kahn, Ronen Kat, Kalman Meth

IBM Haifa Research Lab



Motivation - Data Center Power Crisis

- Data Center demand for power continues to increase
  - Increased server density (and heat) leads to additional cooling
- Data Center energy bills are a significant factor due to
  - Increased energy consumption
  - Increased cost of energy
- Energy providers will impose limits on power supplied to Data Center
  - Need to make management decisions based on available power
  - Cannot provision power for worst case
  - Use power to capacity; maximize performance per watt
- Demand for storage continues to accelerate
  - Compliance, Content depots, Data Replication for Disaster Recovery
  - => Storage will require a greater percentage of data center energy consumption
- **Need *power savings* as well as *power management***

Power vs. Energy

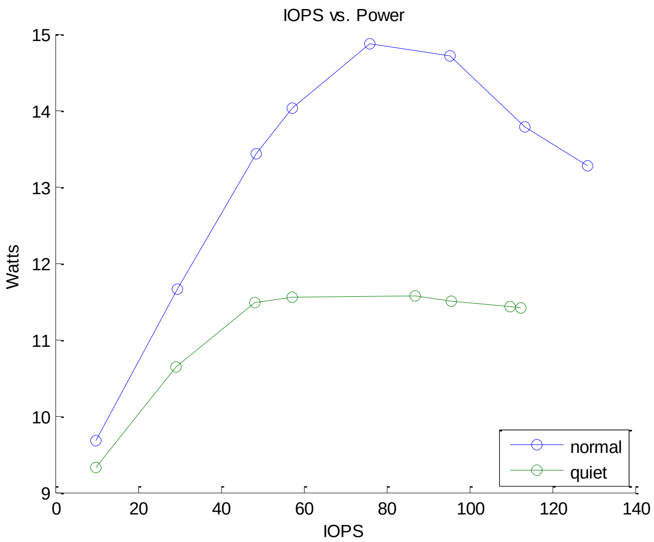
- Power is measured instantaneously
  - Power distribution to equipment
  - Rate (cost) may be based on peak power consumption
- Energy is the overall power consumption over time
  - Energy utility bills
  - Carbon emissions
  - Cooling

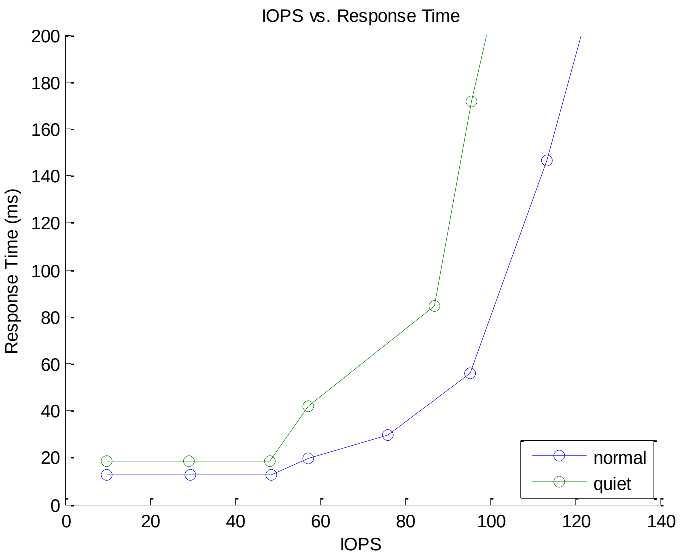
Power Capping and Energy Reduction

- Power capping
  - Allows a data center to provision power and cooling for reasonable scenario; not worst case
  - In case of temporary fluctuation in power and cooling:
    - Cap the power budget allowing the data center to continue operating with reduced performance
- Energy Reduction
  - Optimize energy consumption and performance using a utility function
  - Leads to energy optimized workloads

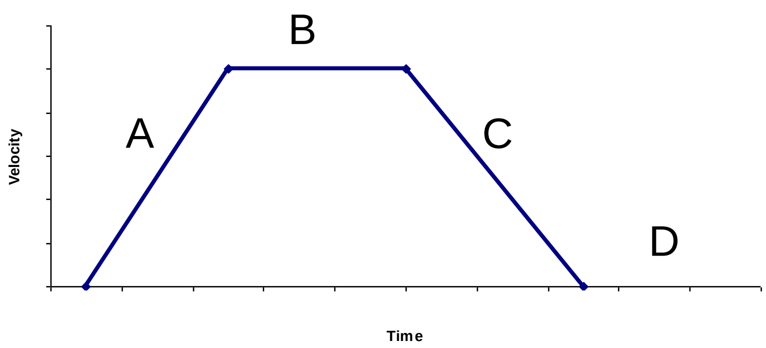
Exploitation of Acoustic Modes

- Goal: Scale energy consumption to fit performance workload
- Background:
  - ATA/ATAPI-6 standard support Automatic Acoustic Management (AAM) levels
    - A standardized way of setting limits on the disk acoustics (noise)
  - Current Hitachi and Western Digital drives support **normal** and **quiet** modes
  - Similar technologies are available on FC/SAS drives
  - **I/O operations run longer at reduced power (and noise)**

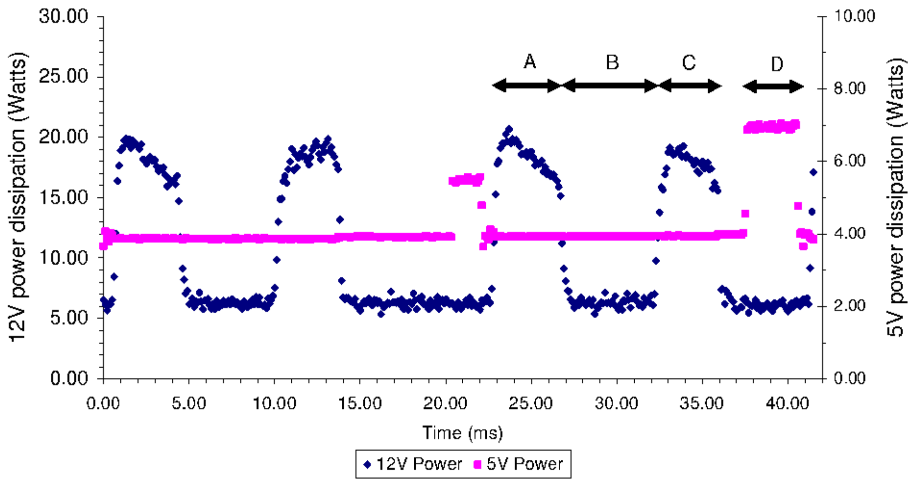




Comparing Acoustic Modes

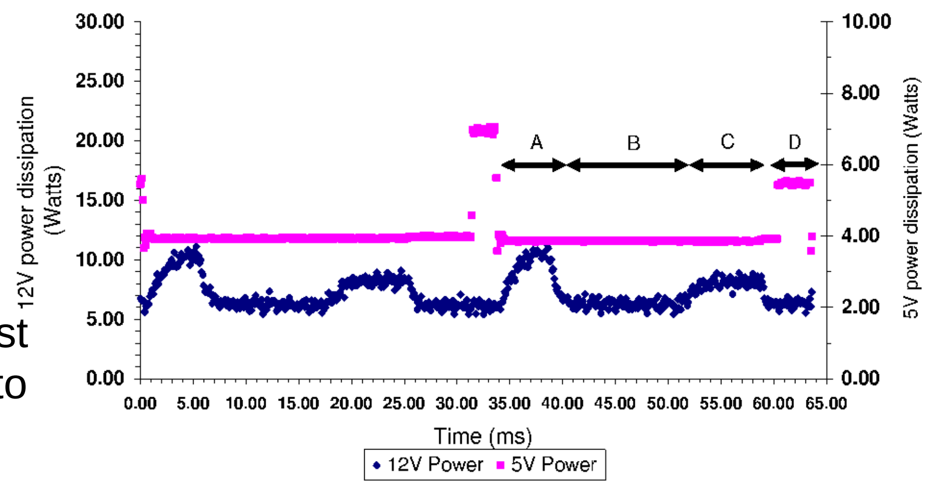


A – Acceleration
B – Coast
C – Deceleration
D – Data transfer



Normal Mode

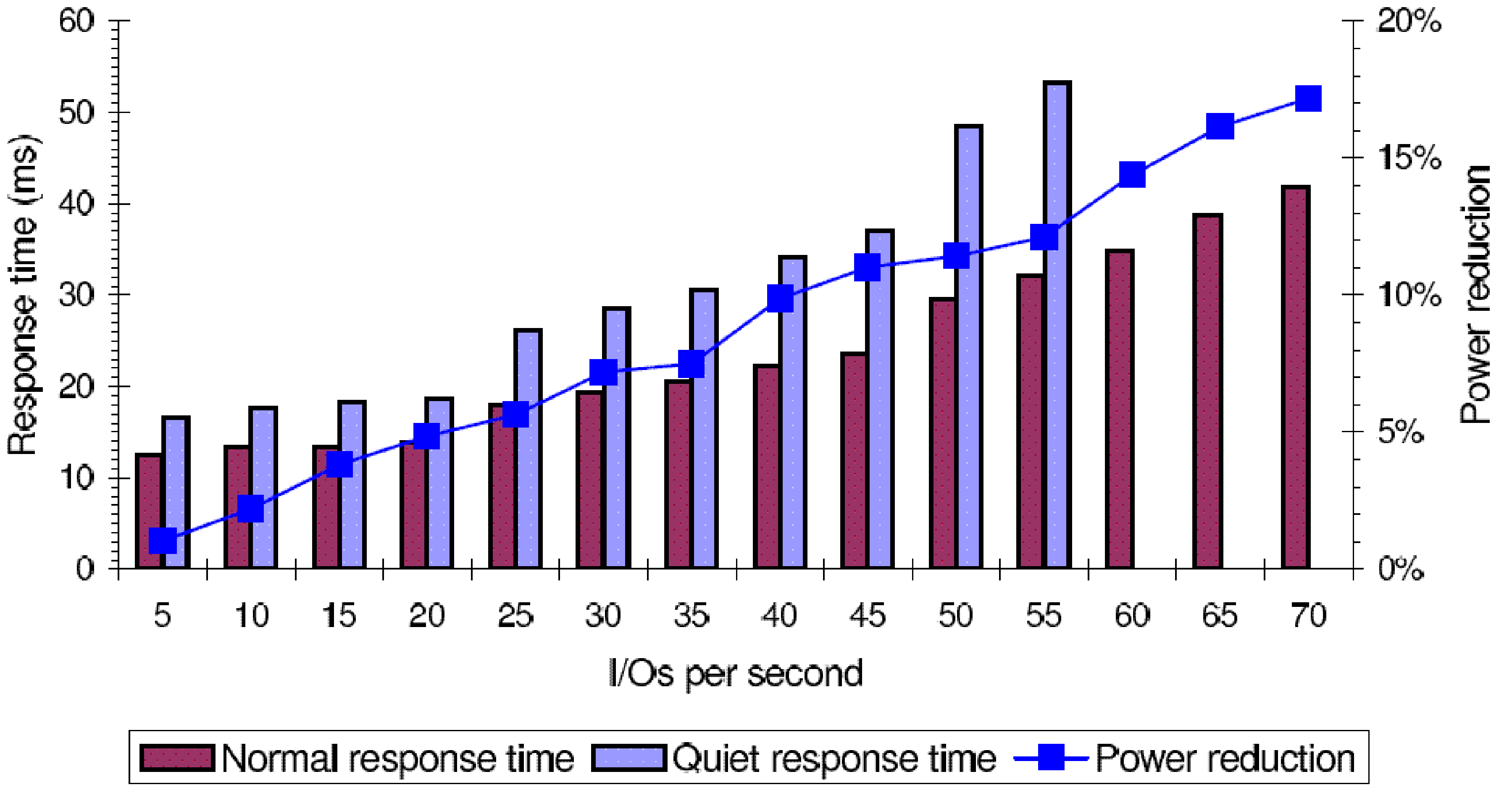
- Quiet mode uses:
  - Less power for acceleration and deceleration
  - Equal power for data transfer and coast
  - More energy per individual seek, due to longer coast



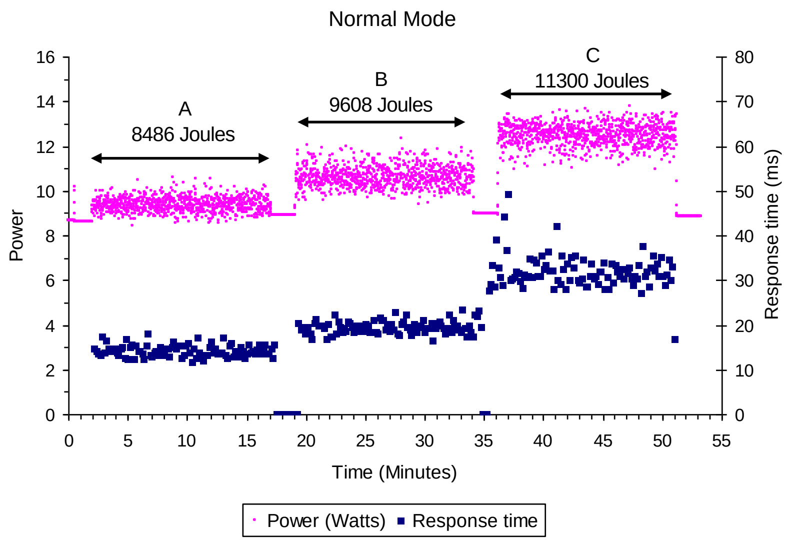
Quiet Mode

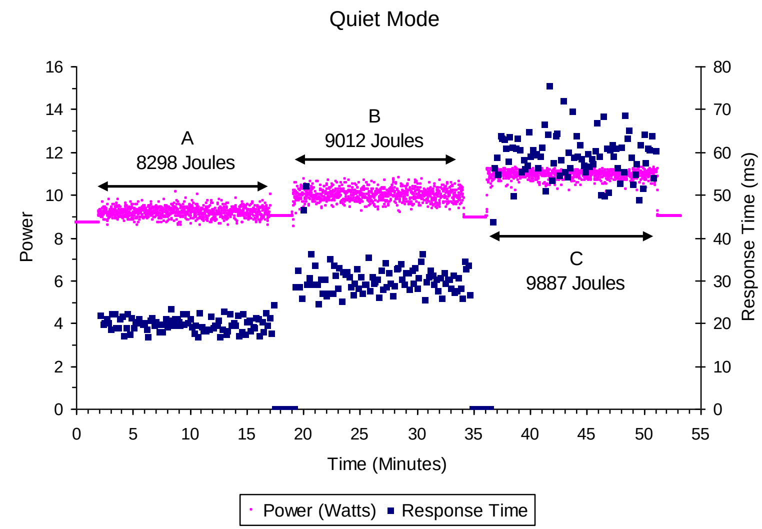
Note: The graphs describe long range seek operations.

Response Time and Power for SPC-1 like Workload



SPC-1 like Workload: Energy Reduction



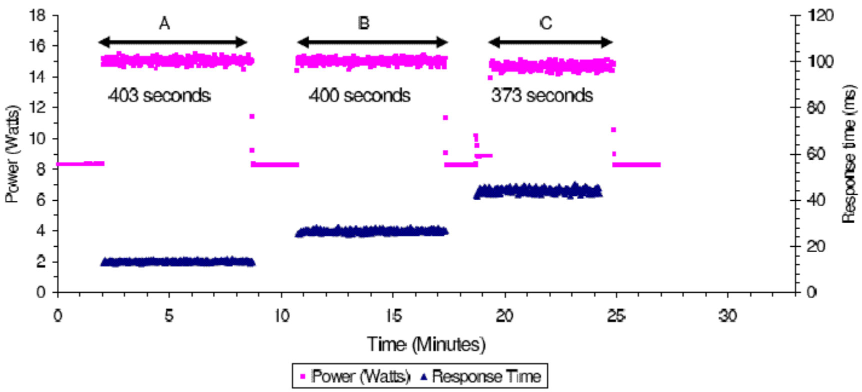


Energy Saving For various SPC-1 like Workloads

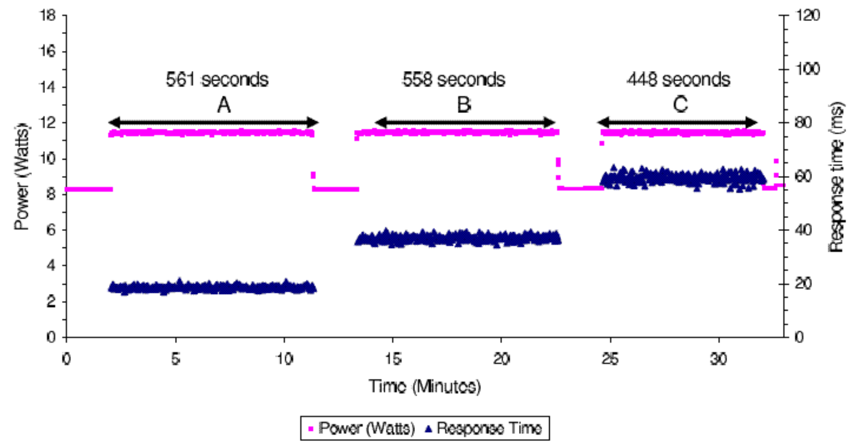
SPC-1 I/O rate	Normal mode	Quiet mode	Energy difference
10	8486 J	8298 J	−2.2%
25	9608 J	9012 J	−6.2%
50	11300 J	9887 J	−12.5%

- Exchange “wasted” disk idle time with a longer and slower seek

Trace Workload¥: Power Reduction, Energy Waste



(a) Normal mode.



(b) Quiet mode.

Number of concurrent requests	Normal mode	Quiet mode	Energy difference
1	6063 J	6419 J	+5.9%
2	6001 J	6390 J	+6.5%
4	5234 J	5122 J	−2.2%

- In C we have queuing optimizations that reduce overall energy cost

¥ The trace workload is a 4K random read

Summary

- Quiet mode always helps with **power capping**

- Quiet mode helps with energy reduction in **some cases**
  - Good for workloads with constant IOPS
  - Good for multi threaded (I/O independent) applications
  - Not beneficial for single threaded applications
  - Incurs a response time delay

- Quiet mode has no effect at idle or during sequential access

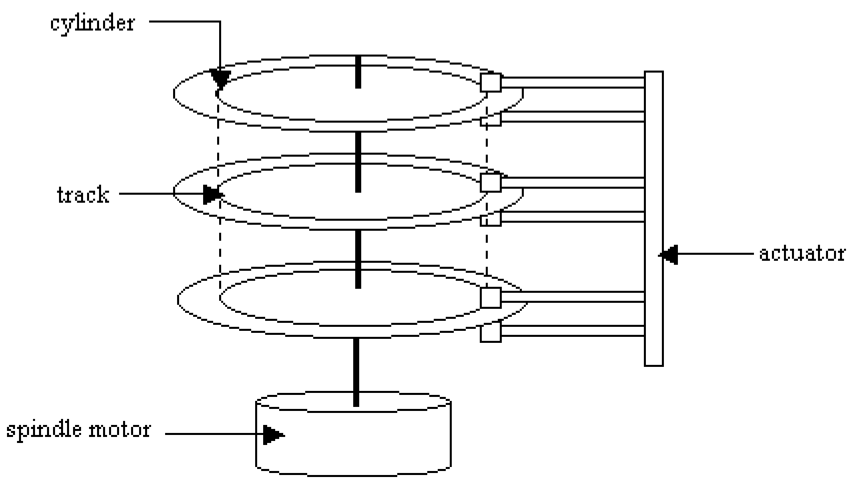
Questions?

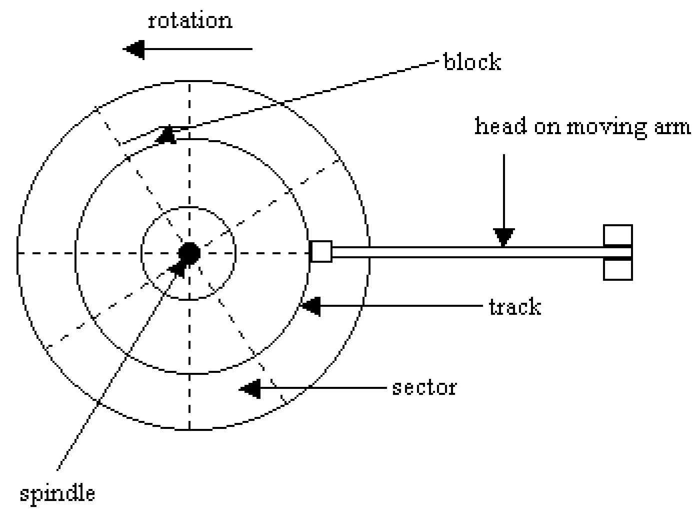
Backup Slides

Related Work – energy savings

- DRPM – Dynamic Rotations per Minute
- Spin-up and Spin-down
  - MAID – Massive Array of Idle Disks
  - Write offloading
  - …..
- Flash / SSD

Introduction to Power Modeling for Storage





Fixed: Spindle power, electronics

Dynamic: Seek activity, data transfer

$$\mathrm{Power(disk)} = \mathrm{Fixed}_{\mathrm{power}} + \mathrm{Dynamic}_{\mathrm{power}}$$

$$\mathrm{Fixed}_{\mathrm{power}} = \mathrm{Power(spindle)} + \mathrm{Power(electronics)}$$

$$Dynamic_{power} = \sum_{\mathrm{I/O}} [Power(\mathrm{seek}) + Power(\mathrm{datasize})]$$

12V – Mechanical

5V – Electronics