HITACHI
Inspire the Next

Hard Drive for Low Power
Energy Efficiency in Disk Storage

Sridhar Chatradhi
June 2009

Hitachi Global Storage Technologies

5 Decades of Spinning Disks

HITACHI
Inspire the Next

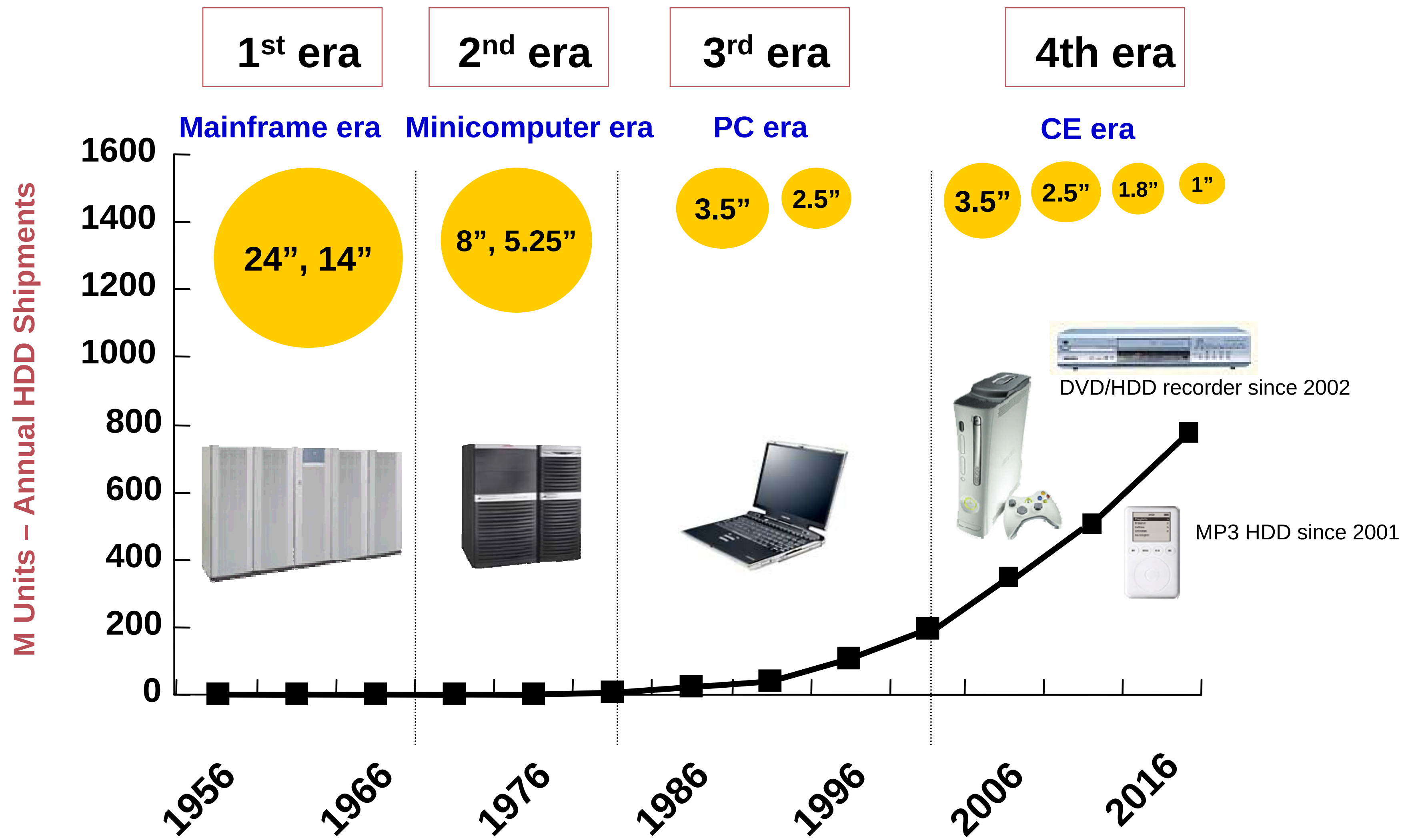
1956

RAMAC - first HDD
- 5 MegaBytes
- Fifty 24” disks
- 1200 RPM
- 2 kbits/sq.in.
- 150 kbit/s
- 800 microinch head to disk spacing
- Power – 52 KVA

Deskstar 7K1000B
- 1 TeraBytes
- 3.5” form factor
- 7200 RPM
- 244 Gbits/sq.in.
- SATA 3 Gb/s
- Nanometers in head to disk spacing
- Power – 5.2 W (Idle)

2008

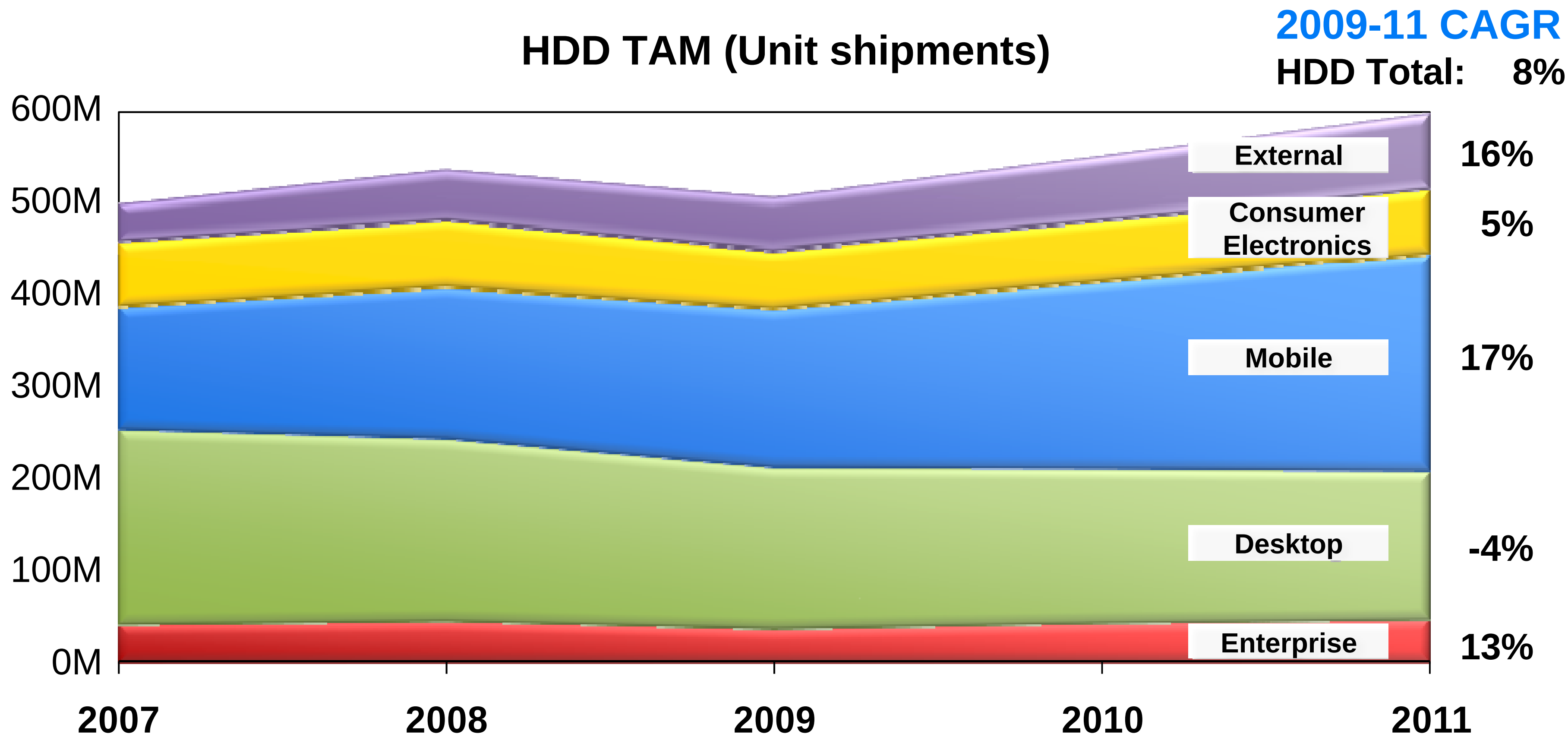
Fourth Era of HDD: Consumer Electronics



CE: Consumer Electronics
Source : Hitachi GST

HDD Market Outlook by Segment

HITACHI
Inspire the Next



Source: Hitachi GST estimates, May 2009

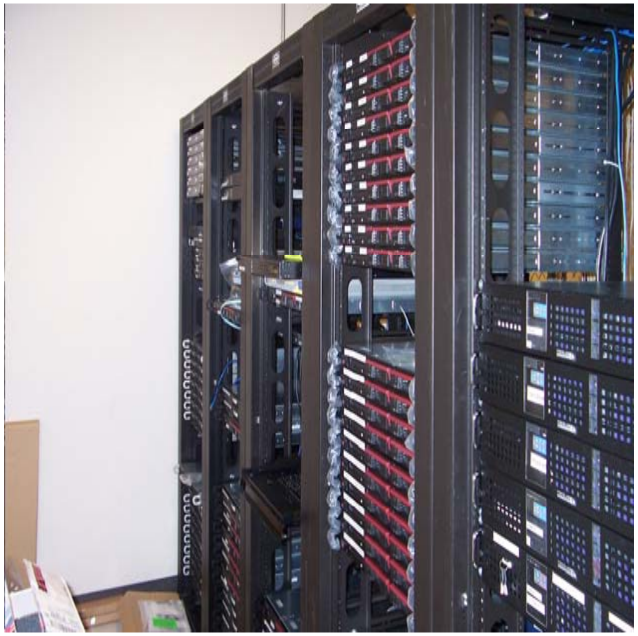
- Large and growing market; 2009 - 2011 overall growth rate is 8%
- Growth driven by mobile, external and enterprise drives

Power Consumption in Data Centers

HITACHI
Inspire the Next

"What matters most to the computer designers at Google is not speed, but power - low power, because data centers can consume as much electricity as a city" – Eric Schmidt CEO of Google

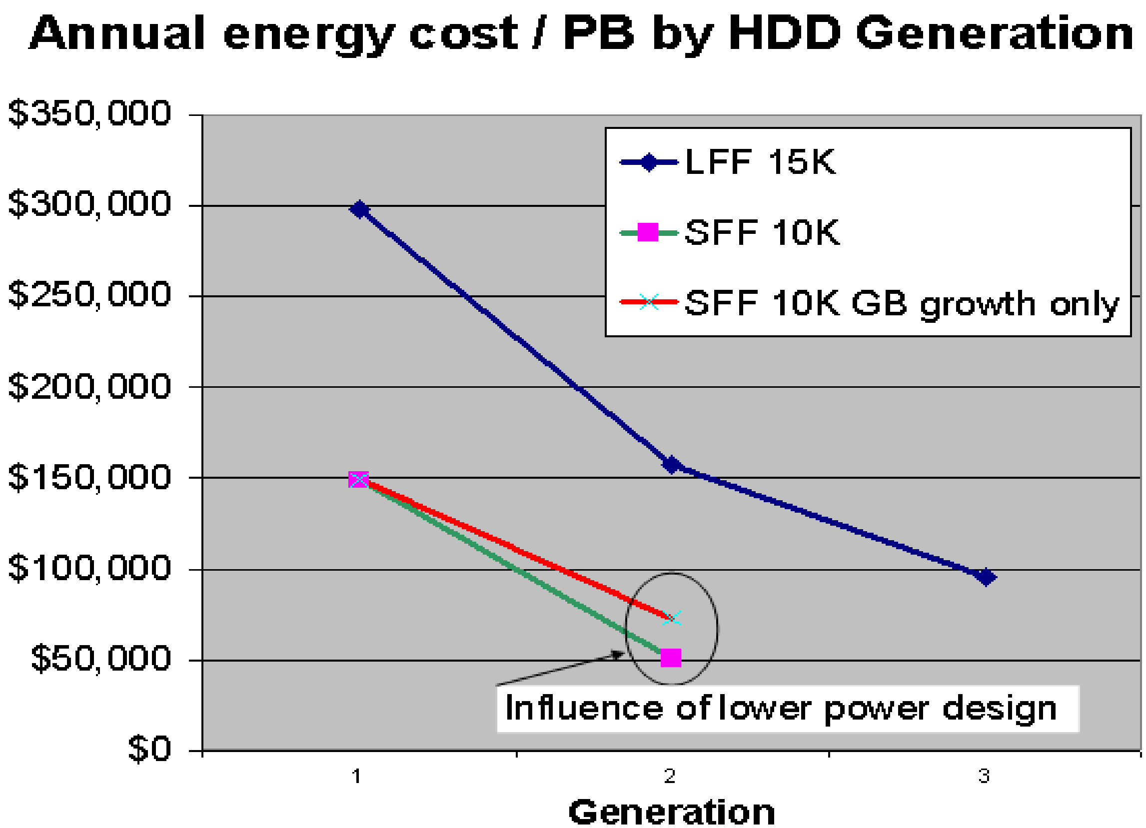
Storage is the 3rd largest power consumer in a data center behind Servers and cooling – consumes about 27% power



- **According to recent EPA report data centers and servers consumed 61 billion KWH of electricity in 2006 and it is likely to double by 2011**
- **According to IDC report in 2008, total cost to power and cool a drive is 48 watts**
  - 12 watts for running HDD
  - 12 watts for storage shelf (HBAs, fans, power supply)
  - 24 watts to cool the HDDs and storage shelf

Power Requirements Driving New Thinking in Storage

HITACHI
Inspire the Next



Data for Enterprise SAS drives

Assumptions:
Power usage effectiveness: 3
Power supply efficiency: 85%
Electricity cost ($/KWhr): $0.0951

Notes:

1. Annual energy costs calculated at the data center level. Energy cost of CPU, HBA, NIC, etc. not considered.
2. Power Usage Effectiveness = Total Facility Power / Computer Equipment Power
3. Assume national average commercial energy costs from http://www.eia.doe.gov/cneaf/electricity/epm/table5_6_a.html
4. PB = Peta bytes, LFF=Large form factor (3.5”) SFF= Small form factor (2.5”)

“Power” as a Key Parameter in Hard Drive Design

HITACHI
Inspire the Next

Economics

- Lower power saves money on critical system components
  - Power supplies
  - Cooling (fans, clearances)
  - Cost of powering the HDD
- Lower power increases reliability
  - Longer HDD life means less frequent replacements of HDDs in data centers

Environmental and Regulatory

- Low power increasingly a part of bids from government and private accounts
- Energy use impact on environment
- ENERGY STAR®, other programs emphasizing product efficiency
- Efficiency standards from initiatives like Green Grid, Climate Savers Computing

Storage Market Push to Lower Power

- **Enterprise storage**
  - Pursuit of lower RPM HDDs in large storage systems
  - Advanced power management features in SAS drives (similar to SATA)
  - Large form factor (3.5”) to small form (2.5”) factor shift
- **CE (Set-top Box, DVR)**
  - Low RPM drives in Set-top Boxes
    - Other factors include: heat dissipation, acoustics and reliability
- **Desktop PC**
  - ENERGY STAR specification for Desktop PC driving lower power HDD
  - Movement to 2.5-inch for compact systems and acoustics
- **Notebook computing**
  - Battery life continues to drive power reductions

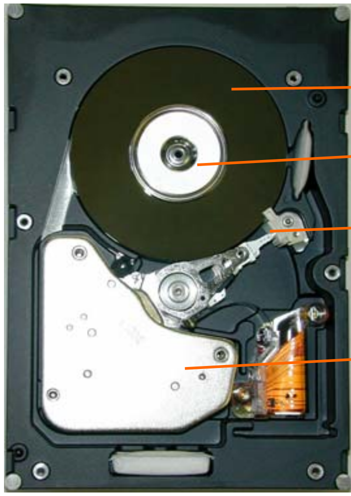
HDD Specs’ Influence on Power

- **Lower RPM - Lower power**
  - Reducing RPM from 7200 to 5400 can reduce power by about 2-3 W
- **Diameter of Platters: Smaller diameter – Lower power**
  - 2.5” drives consume about 3-4 Watts less power than 3.5” drives
- **Number of platters – Fewer platters – Lower power**
- **Higher performance requires higher power**
  - Higher sequential performance requires higher disk data rate
  - Faster seeks

- **Trade off between capacity / performance requirements and power based on application**

Disk Drive Components

HITACHI
Inspire the Next

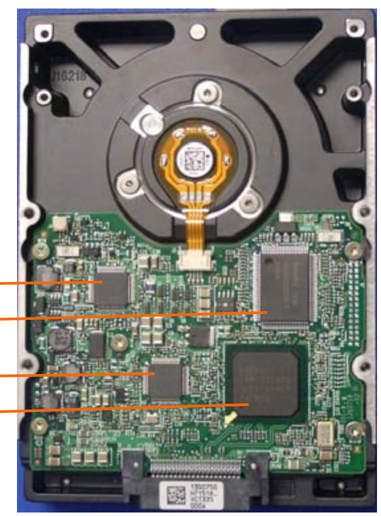


Mechanical

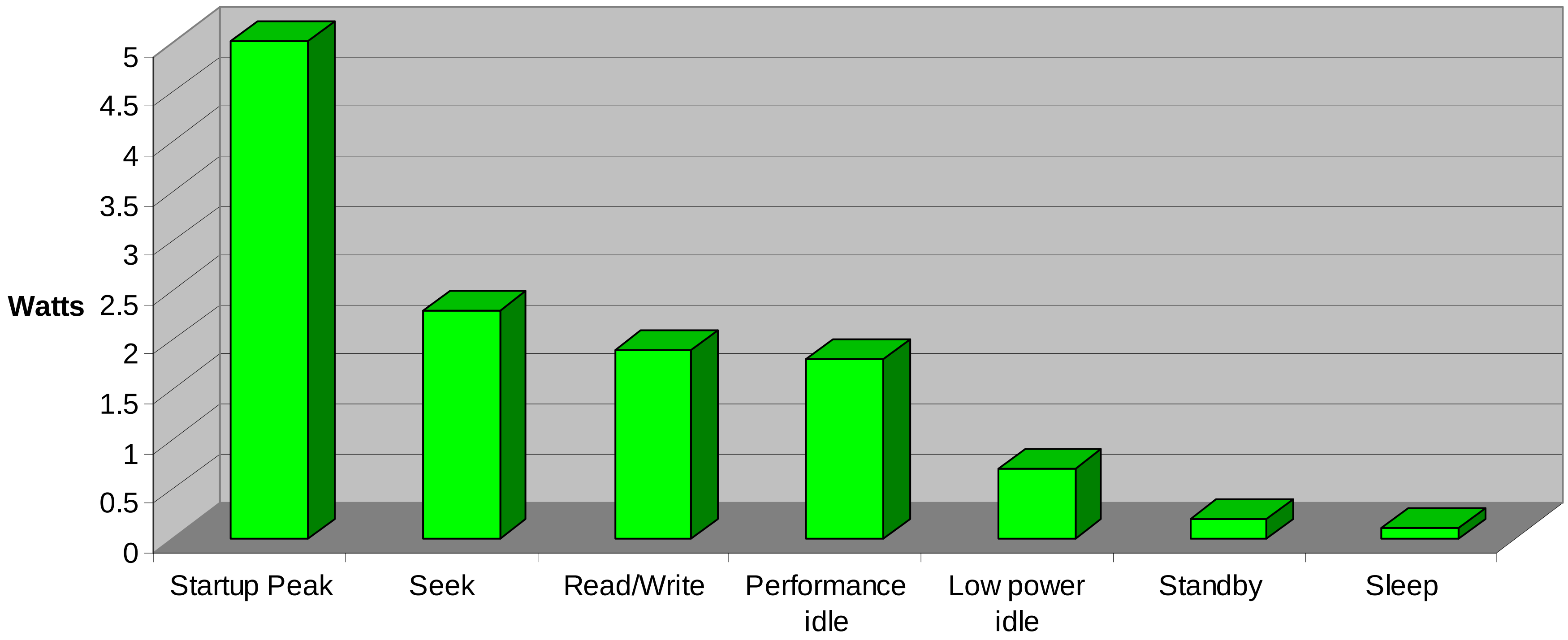
- Power to run spindle motor
- Power to move actuator
- Mostly derived from 12 V

Electronics

- Power for Electronic components
- Mostly derived from 5 V



Power consumption of a 2.5” SATA drive in various power modes



Power Reduction in Disk Drives

HITACHI
Inspire the Next

Electronics

- Silicon Technology and design improvements
  - Higher integration, lower voltages, low power libraries
  - Power reduction through clock gating and other low power design techniques
- Adaptive voltage schemes
- Frequency reduction in low power modes
- Improving power regulator efficiency

Firmware

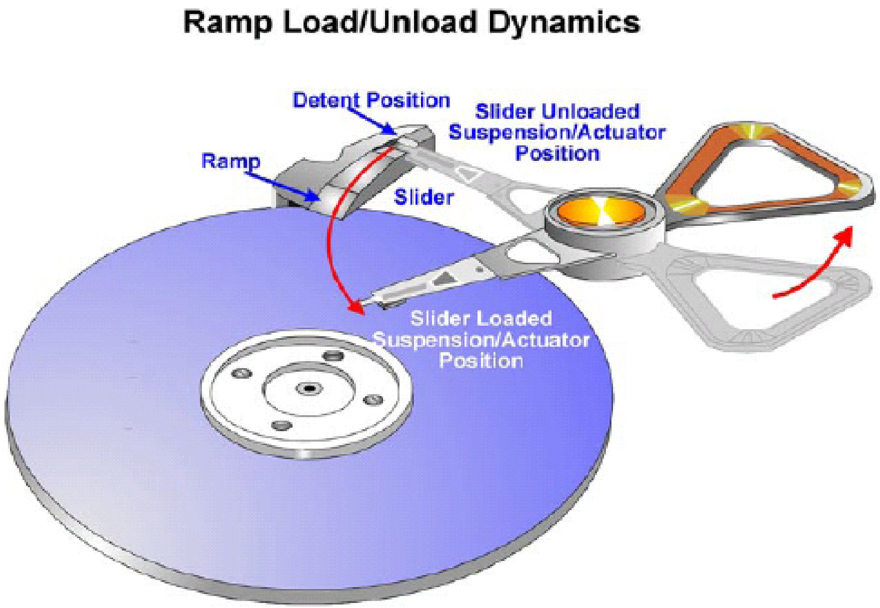
- Queuing algorithms to minimize rotational delays and maximize the amount of time in low power modes
- Algorithms to manage transition between active and low power modes

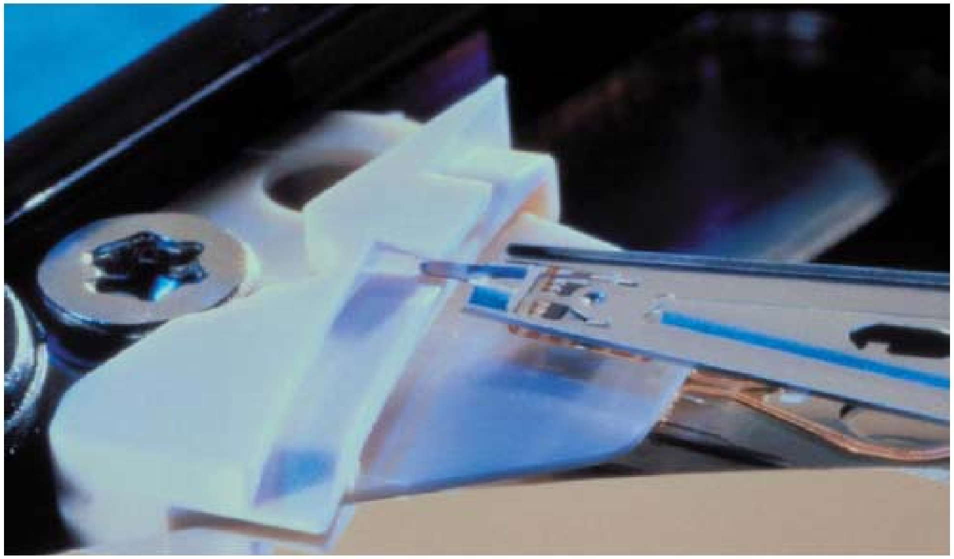
Servo mechanical

- Lighter materials for moving parts consume less power
- Less air drag on actuator and disk
  - Hermetically sealed drives filled with Helium
- Efficient Motor design – Spindle and VCM
- Algorithms to lower energy per seek (JIT Seek)
- Use of Load/ Unload during low power modes

Load/Unload

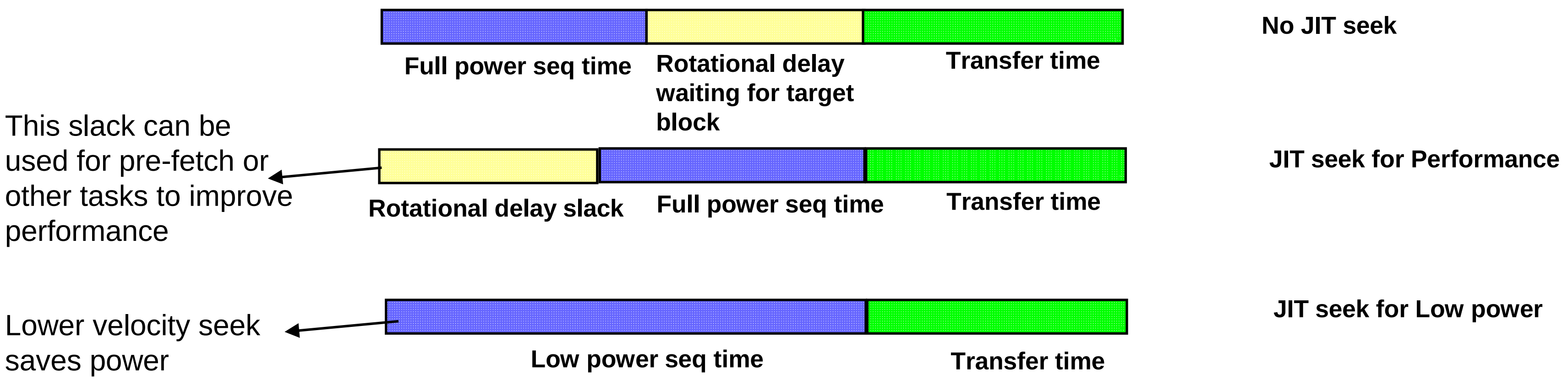
HITACHI
Inspire the Next





- **Prior to Load/ Unload technology Contact start stop was used where sliders land on disk during power down**
- **Load/ Unload allows enhanced power management modes – slider can be unloaded even while disk is spinning**
- **Reduces power by reducing aerodynamic drag during idle mode**
- **Increases number of start stop cycles for low power applications**
- **Elimination of stiction helps in improving reliability and areal density**

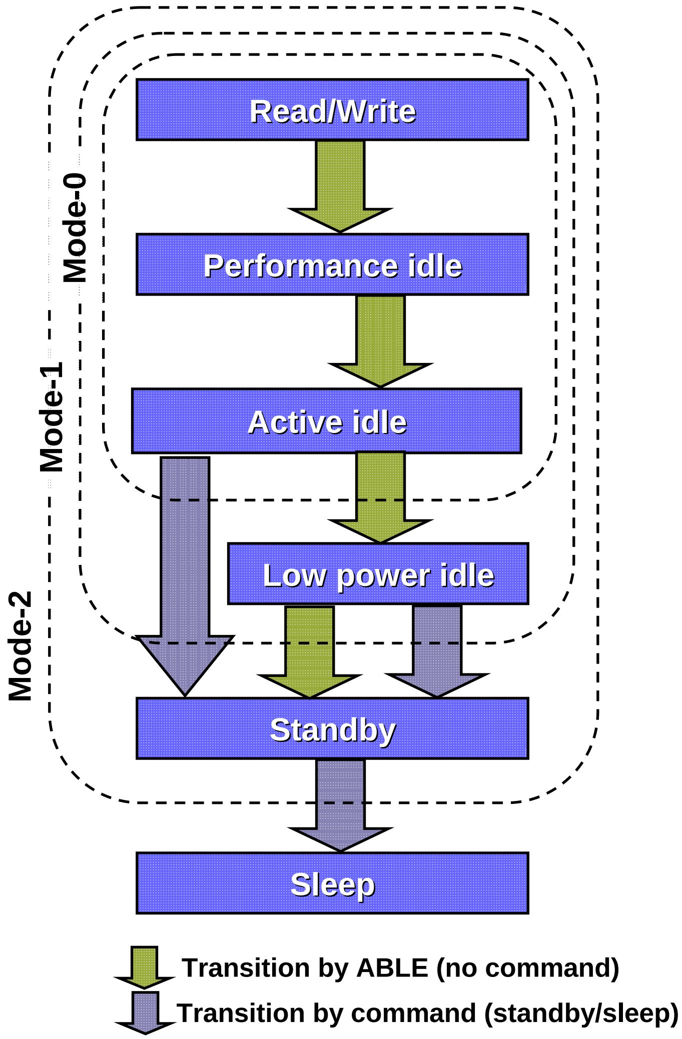
Just In Time (JIT) Seek



- **JIT Algorithm times the completion of seek just in time to start data transfer at target block**
- **Using velocity values that correspond to seek times saves power**
- **Algorithms calculate the estimated seek times and required velocity to optimize for power**

Adaptive Battery Life Extender (ABLE)

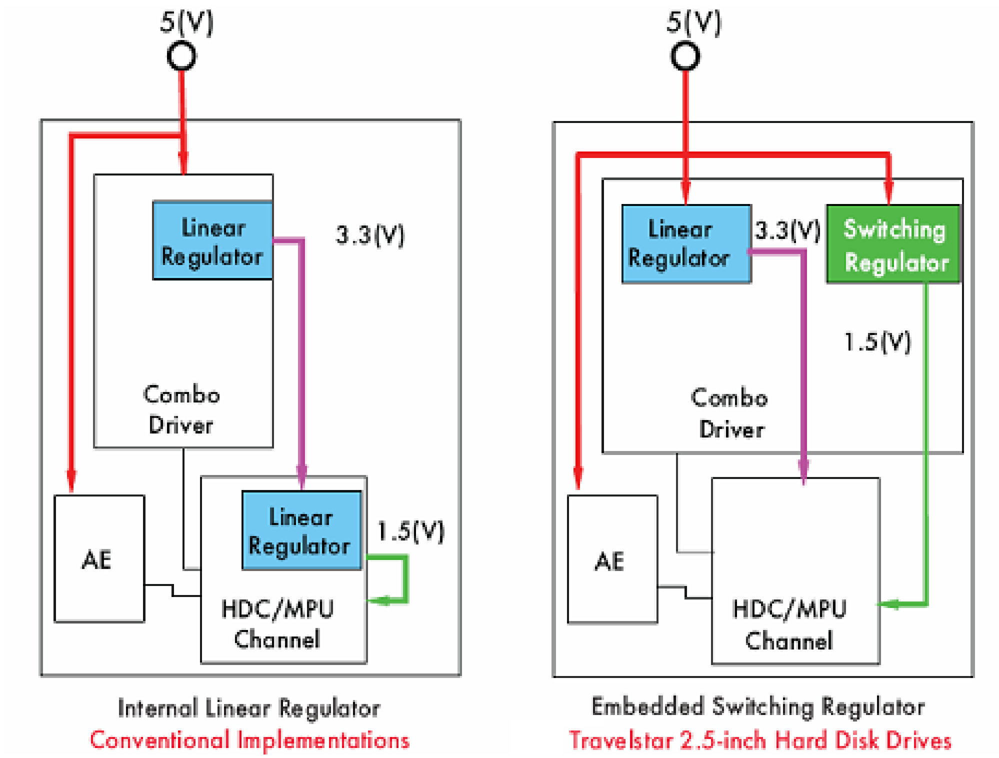
ABLE is an example of one of the earliest efforts at energy efficiency in HDDs
ABLE uses more than 4 low power modes and intelligently manages transition between modes



R/W	Servo	Heads	Spindle rotation	Elec-tronics	Power* (W)
On	On	Load	On	On	1.8
Off	On	Load	On	On	1.7
Off	Off	Load	On	On	0.8
Off	Off	Unload	On	On	0.55
Off	Off	Unload	Off	On	0.2
Off	Off	Unload	Off	Off	0.1

*Travelstar 5K320

- **Use of switching regulators to negotiate voltage reduction**
- **Minimizes power loss when converting input voltages to those used by the electronics components within the HDD**
- **Significantly reduces power consumption by the HDD without sacrificing performance of the base design**



HDD Power Management Standards - SATA

- **T13 committee and SATA-IO working on power management standards for SATA**
- **SATA power management**
  - The ATA8-ACS standard describes four modes of power consumption for SATA products
    - 1. Active – The device is fully powered up and ready to send/receive data.
    - 2. Idle – The device is capable of responding to commands but may take longer. Drive is spinning.
    - 3. Standby – The device is capable of responding to commands but take longer (up to 30 seconds). Drive is spun down.
    - 4. Sleep – This is the lowest power mode. The device will exit the Sleep mode only after receiving a reset. Wake up time can be as long as 30 seconds.
  - The SATA specification defines three SATA interface power modes:
    - 1. PHY Ready (PHYRDY) - the SATA PHY is ready to send/receive data
    - 2. Partial - the PHY is in a reduced power mode; exit time can be up to 10 microseconds
    - 3. Slumber - the PHY is in a further reduced power mode; exit time can be as much as 10 milliseconds
  - An ATA device that implements power management has to support these functions
    - a) A Standby timer
    - b) CHECK POWER MODE command
    - c) IDLE command
    - d) IDLE IMMEDIATE command
    - e) SLEEP command
    - f) STANDBY command
    - g) STANDBY IMMEDIATE command

HDD Power Management Standards - SAS

- **T10 Committee is working on power management standards for SCSI (SAS) devices – final draft is pending approval**
- **SCSI Power Management - SPC-4 SBC-3 SAS-2.1**
  - Device power management
    - Behavior in the following states are not called out in the standard, Drive implementations may differ.
    - The following is an example
    - Active - Device is fully powered and ready to send/ receive data
    - Idle A – Idle mode. Disk spinning and Heads loaded. Some electronics shut off
    - Idle B – Disk spinning, Heads Unloaded.
    - Idle C – Disk spinning at lower rpm, Heads Unloaded
    - Standby Y - Disk spun down, Heads Unloaded
    - Standby Z – Most of the drive is shut off. Sleep mode
  - Interface Power management
    - Active – SAS PHY active and ready
    - Partial – PHY is in a reduced power mode; exit time can be up to 10 microseconds
    - Slumber - PHY is in a further reduced power mode; exit time can be as much as 10 milliseconds
    - Primitives PS_REQ, PS_ACK, PS_NAK defined
    - COMINIT, COMWAKE defined as in SATA
  - Timers and functions
    - Idle, standby and recovery timers
    - Statistics for transition from power modes

- **System level efficiency requires awareness of HDD power characteristics**

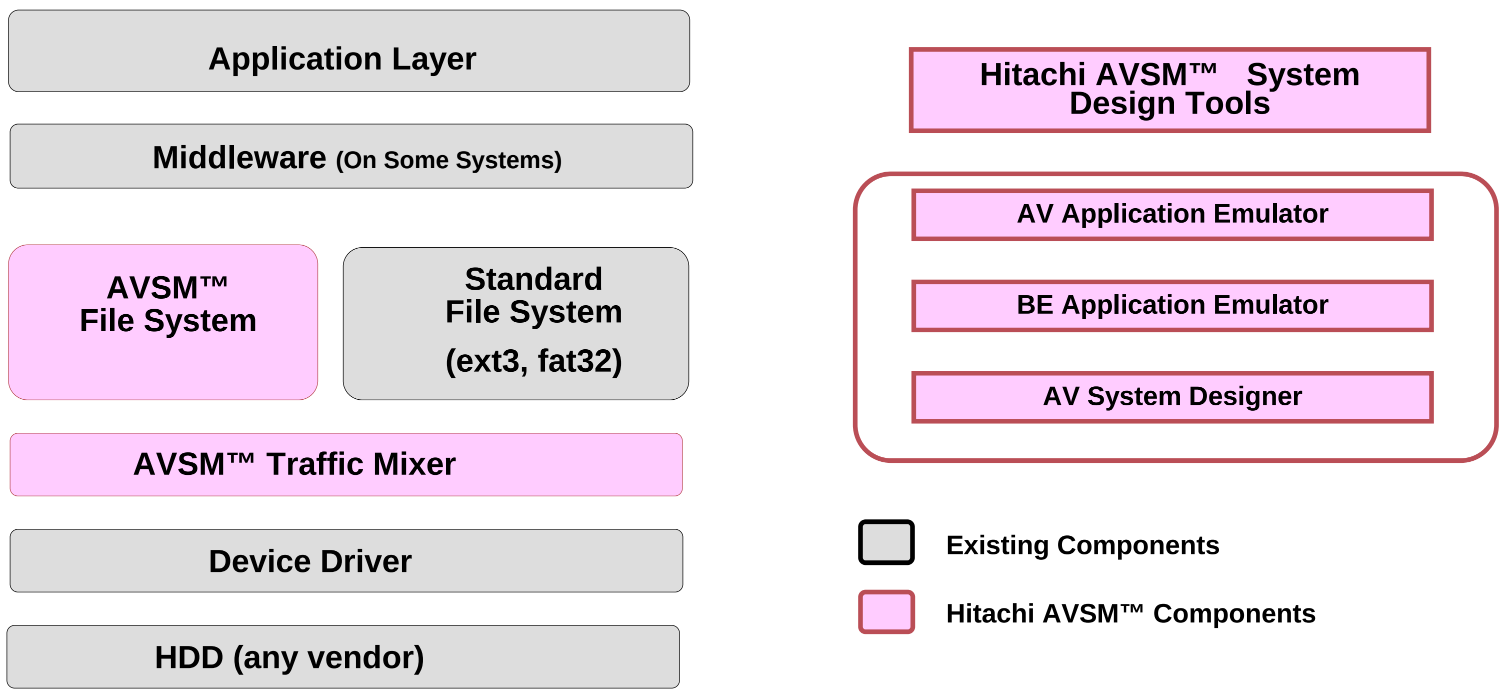
- **Efficient management of low power modes**
  - Energy saved in standby or sleep mode should be greater than energy required to spin up drives
  - Multiple drives in same rack spinning up from standby mode simultaneously raise peak current requirements
  - Manage workloads efficiently among multiple RPMs in a storage system
  - MAID – Massive Array of Idle Drives – designed for write once read occasionally applications uses low power modes very effectively by spinning down most of the drives in system for long periods

Efficient usage of storage capacity and data management

- Having large unused capacity is less power efficient
- Compression – data compression can reduce storage requirement
- Tiered storage – Tier 0 with solid state drives (5%), Tier 1 with high performance HDDs (15%) and Tier 2 with low power HDDs (80%)
- Data De-duplication – reducing number of duplicate copies of the same data reduces storage requirement

Audio-Visual Storage Manager™ (AVSM™) Overview

- **AVSM is an embedded software for multi-stream AV applications.**
- **AVSM consists of a set of software components and system design and analysis tools.**
- **AVSM Development by SJRC & Hitachi SDL has been completed in 2006.**
- **AVSM optimizes the system performance by minimizing duty cycle.**
  - With the reduced duty cycle, power, temperature and acoustic can be reduced and reliability can be improved, which are big benefits to customers like STB.
- **AVSM has proved to be very stable S/W through testing by partners.**



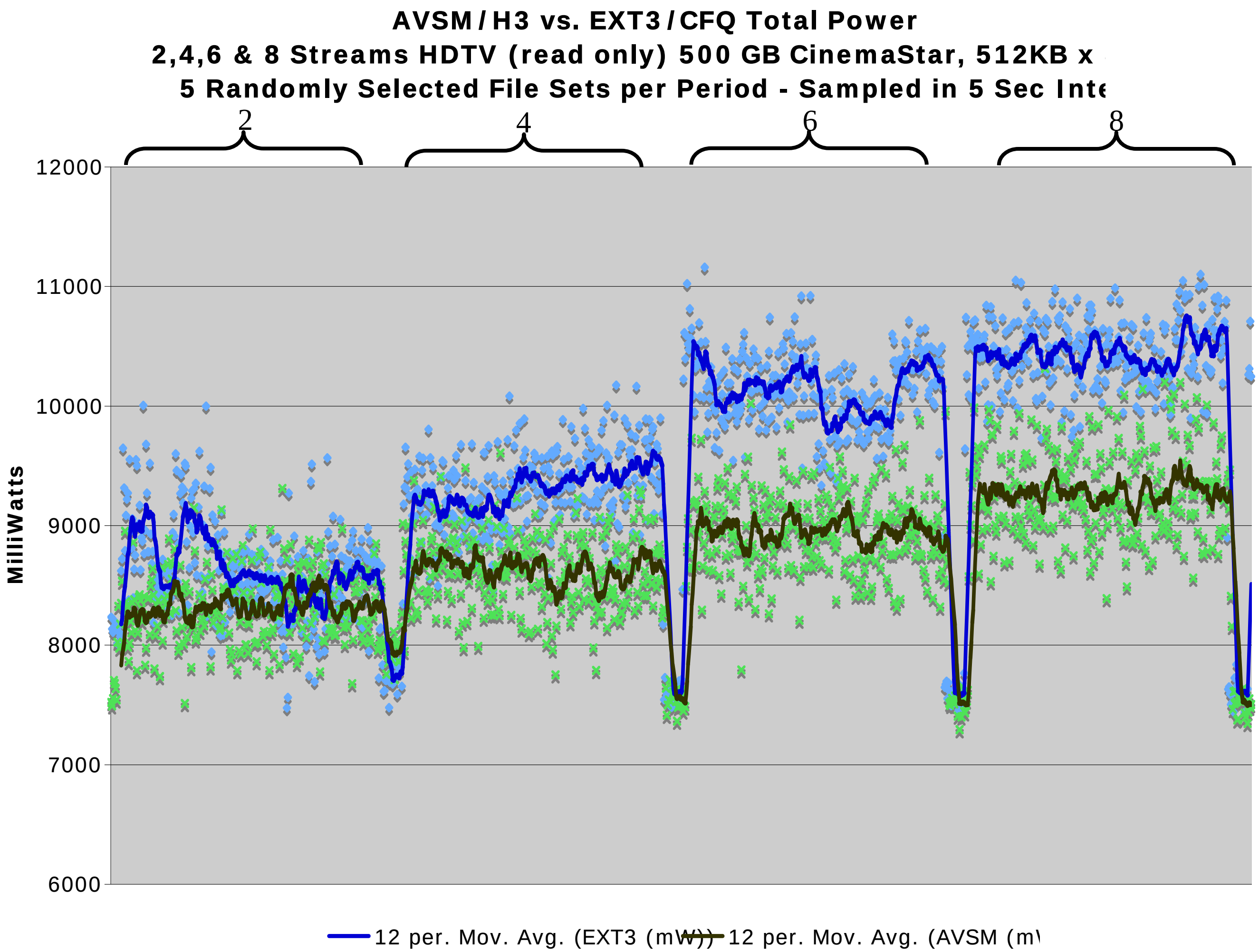
Measured Power Improvements with AVSM (3.5")

HITACHI
Inspire the Next

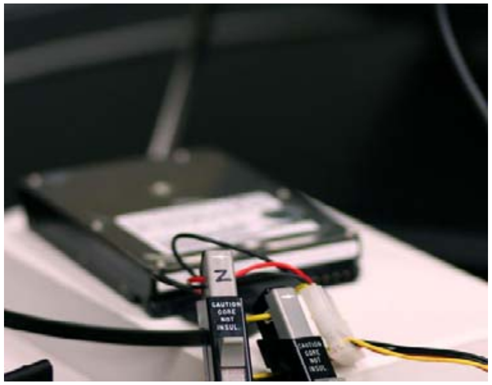
Improved Performance / QoS
(Improved differentiation and lower risk)

Better Reliability/Robustness
(Lower failure rates & support costs)

Simple Integration Process
(Lower design & manufacturing costs)



The effect of duty cycle minimization is shown in the amount of power consumed by the HDD at various workloads



Work Load HDTV Streams	AVSM Power Advantage
2	**~ 0.2 - 0.5 W**
4	**~ 0.5 - 0.7 W**
6	**~ 1.0 - 1.2 W**
8	**~ 1.2 W**

Summary

- **Disk storage is a significant portion of data center energy consumption**
- **Continued growth expected in digital data storage requirements**
- **Industry wide efforts to reduce power consumption of disk drives**
- **Standardization key to achieving system level efficiencies**
- **System level power management solutions necessary for overall energy efficiency**

Q & A