Power Management for Storage Systems
Miriam Allalouf and Ronen Kat
IBM Haifa Research Lab

1. Abstract
• Reducing the energy consumption of IT systems is a major challenge. To date, power management in storage systems has been limited.

Project goals
– Study the factors that affect power usage in storage systems and attempt to understand the trade-offs between performance and power consumption.
– Power modeling
  • Embed power calculations with capacity planning techniques.
  • Identify the power consumption components for storage controllers.
  • Provide trade-offs between performance and power.
– Develop fine-grained and accurate power metrics
  • Design and build a benchmark that measures the power consumed by a single disk drive under various disk utilizations.
  • Understand the trade-offs of I/O rate and response time.

Next steps
– Focus on developing power-aware techniques and algorithms.
– Test and validate techniques using power-aware models.

Capacity planning tools: Tools that predicts the expected performance level given an expected host workload and storage configuration. Such tools predict the utilization of the various storage components and compute the workload response time.

2. Motivation

New problem: Datacenter energy bills are skyrocketing.
The digital universe is expanding exponentially:
• New digital data is constantly being created.
• A large part of the data is stored “online.”
• Traditional business data becomes a minority in the storage space.
The server/storage performance gap is increasing:
• Servers are becoming larger and faster.
• Disks are becoming larger and faster.
• However, disk performance increases lag behind capacity increase.

IT heat density is continuously increasing:
• Everything is smaller, faster, and therefore hotter.

Storage data characteristics are changing:
• Low cost SATA drives become the majority in the enterprise.

3. Storage Power Road Map

• Develop models, metrics, and benchmarks to measure power efficiency.
• Define energy efficient storage.
• Design architectures that support:
  – Power vs. performance trade-offs.
  – Power vs. resiliency trade-offs.
• Expose power consumption details to the application layer.

Measure performance and power counters:
• Export to out-of-band analysis tools, e.g., data warehouse, data mining tools.
• Export to software optimizers (online/offline).

4. Power Modeling for Capacity Planning

Model the power consumption of different storage components for various workload characteristics:
– Focus on high-end controller (not simple RAID).

The majority of the storage controller power consumption is attributed to the disk arrays:
– About ⅓ of the power is for spinning the platters. The platters are spinning even when “idle.”
– About ⅓ of the power is for serving I/O requests.

How can we add power metrics to current performance metrics?
1. Identify “power hooks” in the performance model. Some components can be relevant to both performance and power, others may not be.
2. Find the power consumption of each controller component.
3. Compute per-disk power consumption in different I/O utilization levels using workload-dependent power measurements.
4. Integrate the per-disk performance information from the model with the power utilization measurements.

For power-integrated planning tool workload-dependent power measurements are needed!!

5. Measurement Benchmark: Disk Power Under Workload

• Understand disk power consumption in idle vs. active scenarios.
  – Reduce disk power by changing periods of disk idle.

Idle mode: A disk is spinning at its full speed, but no disk activity is taking place. Staying in the idle mode when there is no disk request provides the best possible access time since the disk can immediately service requests.

Benchmarks and Software Tools:
– Iometer sends performance data to the monitoring station.

6. Academic Review

Power-Aware Data Allocation / Migration
Place “hot” data on faster disks and “cold” data on larger, low-cost and even idle media.
– Caching-based MAID (Colarelli and Grunnwald):
  • Cache frequently accessed data together.
  • Migration-based MAID (Pinheiro and Bianchini, Zhu et al.):
  • Move data according to access frequency.

Increasing Disk Idle by Caching
Uses cache management to increase the idle periods of disks to allow disks to enter power-saving modes (standby, spin-down, etc.).
• Cache partition and eviction policies (Zhu and Zhou).
• Write caching using NV (flash) storage (Blisson et al., Chen et al.).
• Prefetch using application hints (Papathanasiou and Scott).

Power-Aware Redundancy
• Tradeoff number of RAID strips with desired power/performance needs (Wendt et al.).
• Shutoff redundant disks (Pinheiro et al.).

Multi Speed Disks (DRPM)
DRPM are shown by academic studies to be very power efficient; however, none are available yet.

Contact: Miriam Allalouf, Ronen Kat miriama, ronenkat@il.ibm.com