# Making Databases Green: An Energy-Aware Software Approach

Yi-Cheng Tu[§]

Joint work with Xiaorui Wang,[°] Bo Zeng,[§] and Ken Christensen[§]

§ Department of Computer Science and Engineering, University of South Florida
° Department of Electrical Engineering, Ohio State University

# The Big Picture

- Electricity used to power up computers in US: **~150TWh**
- That is (roughly)
  - 15 billion US dollars
  - 2% of the global $CO_2$ footprint (comparable to the aviation industry)
- This number is still growing annually

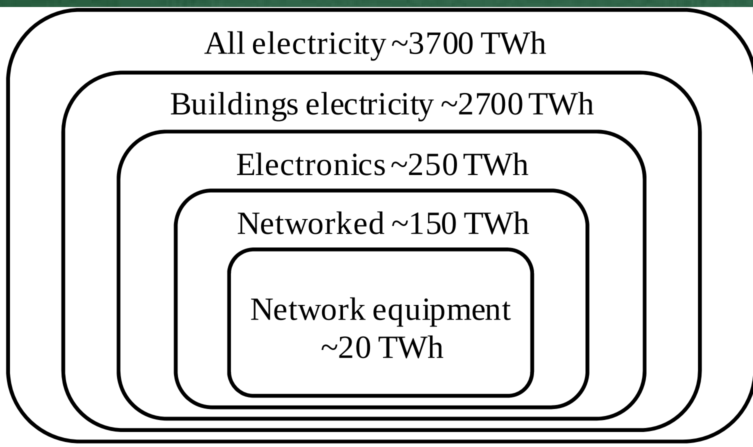



**Figure 1.** 2006 US electricity usage (not to scale)

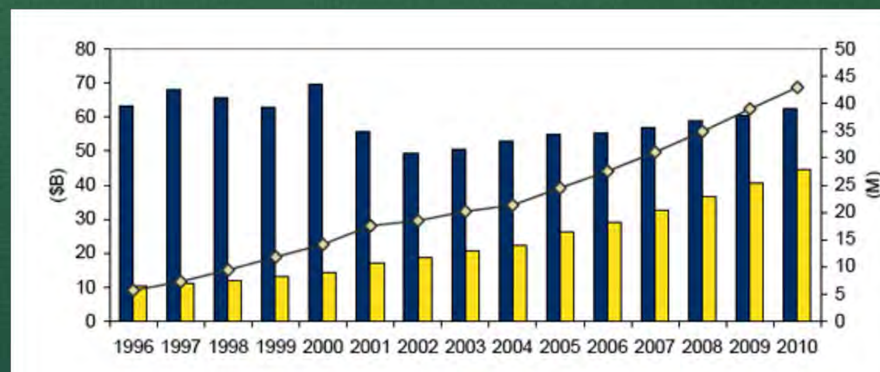



Figure 2. .Worldwide cost to power and cool servers. Blue bar: New server spend ($B), Yellow bar: power and cooling ($B), Dotted line: installed base (M). Source: Poess & Nambiar, 2008

# New Bottleneck for Computing: Energy consumption

- Second largest cost in IT manager's monthly bill
- Main factor to consider in building data centers
- Although energy is the key, power also plays an unique role
  - Directly impact total energy cost
  - Power capping required in high-density clusters
  - Hardware failure increases with power consumption

Google's data center in Hamina, Finland, which has cold climate and low electricity prices. (Source: Wall Street Journal)

Views inside a modern data center, showing racks of blade servers.

# Can We Save Energy in Computing?

- Yes, great opportunities exist.
- Computing resources are often overprovisioned
- Existing computing systems are designed with performance as the only goal

- Solutions require hardware and software coordination
- We work on databases they are the backbone  (and also the main energy consumers) in today's IT industry

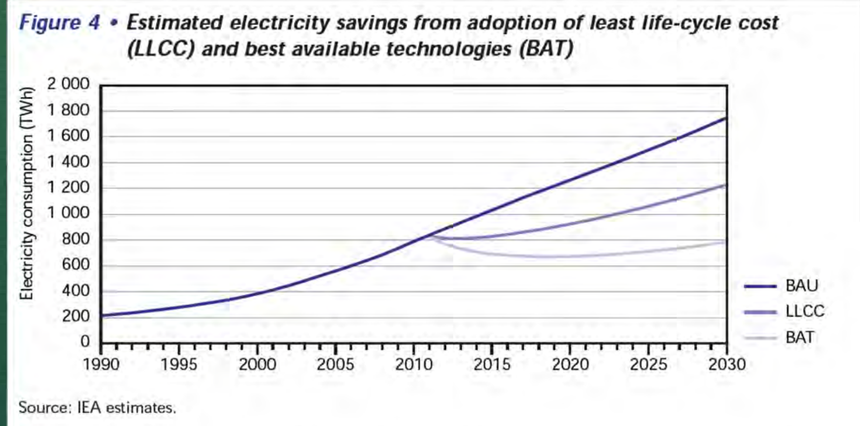



**Figure 3.** ICT and CE electricity savings potential [5]

# The Hardware Solution

- New hardware systems with low power/energy profiles
  - Solid State Drive (SSD)
  - Graphics Processing Units (GPU)

- Energy proportional hardware
  - Diffe
  - CPUs ... trol (DVF
  - Mult

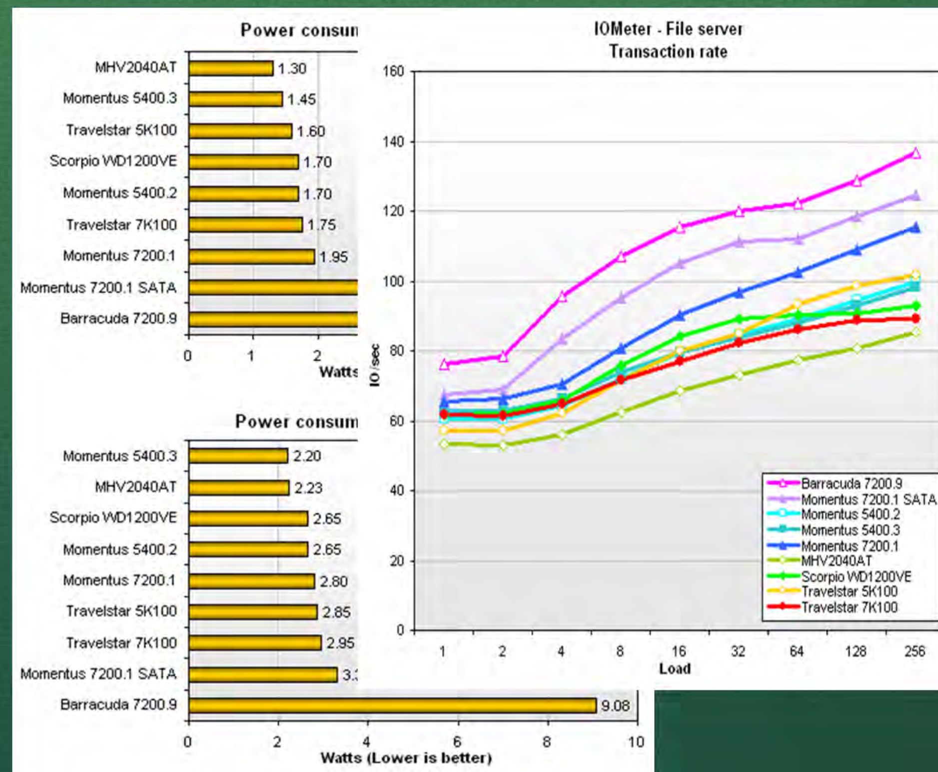
This needs software coordination and control – Dynamic Power Management (DPM)

# SSDs: Green Data Storage

No rotating parts - good for random reads and power efficient


| Intel SSD 700 Series | | |
|---|---|---|
| **Series** | 710 | 720 |
| **Codename** | Lyndonville | Ramsdale |
| **Capacities (GB)** | 100/200/300 | 200/400 |
| **NAND type** | 25nm MLC-HET | 34nm SLC |
| **Cache (DRAM)** | 64MB | 512MB |
| **Interface** | SATA 3Gb/s | PCIe 2.0 |
| **Read speed** | 270MB/s | 2200MB/s |
| **Write speed** | 210MB/s | 1800MB/s |
| **4KB read** | 35 000 IOPs | 180 000 IOPs |
| **4KB write** | 3 300 IOPs | 56 000 IOPs |
| **Power (active/standby)** | 4W/0.095W | 25W/8W |

(source: anandtech.com)

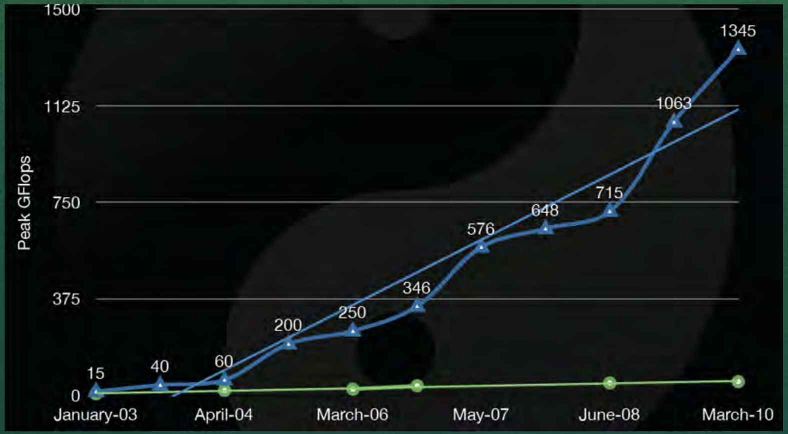


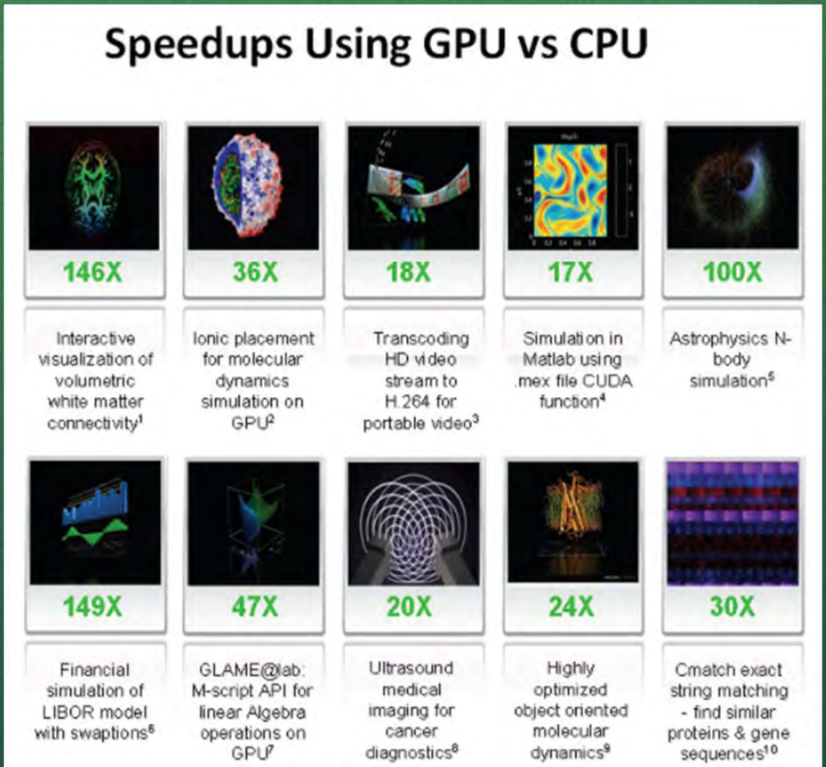

Up: Performance and power consumption of typical 2.5” hard drives (Source: http://techreport.com/articles.x/9859/10)

# GPUs: Powerful

# … and Power Efficient

- With 2-3 times of power consumption of a CPU, a GPU provides
  - Greatly enhanced computing capability
  - Application performance improvement of ~100 times




Left:. Computing capacity growth of GPUs vs. CPUs.
Up: Speedups of GPU computing in different application domains.
(Source: http://synergy.cs.vt.edu)

# Software Solutions

- Energy-efficient computational paths
- Active control of hardware mode
  - CPU (*e.g.*, via DVFS, Fig. )
  - Disks (sleep more, or spinning down)

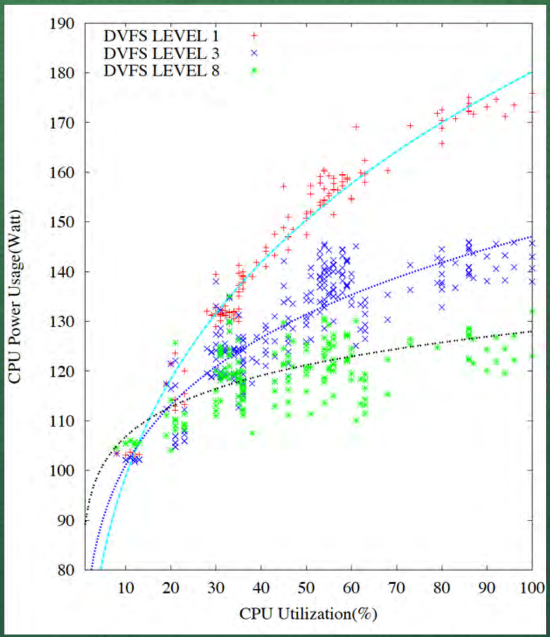



Figure 4 ->. Power consumption of an AMD Opteron 2222 SE CPU under different workload intensities and DVFS levels

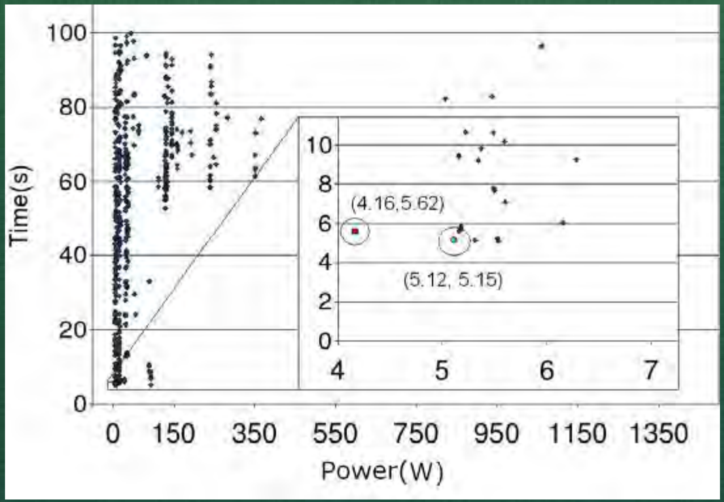



Figure 5. <- (Estimated) processing time and power consumption of plans evaluated by the PostgreSQL query optimizer [2]

# Overview of Our Approach

- A Power-Aware Database Management System (P-DBMS) with
  - High energy efficiency
  - Graceful degradation of performance
- Redesigning key components of existing DBMS with energy-saving functionalities
- Multiple components involved
  - A power-aware query optimizer
  - A dynamic CPU power mode controller
  - A storage management software that dynamically reorganizes data placement

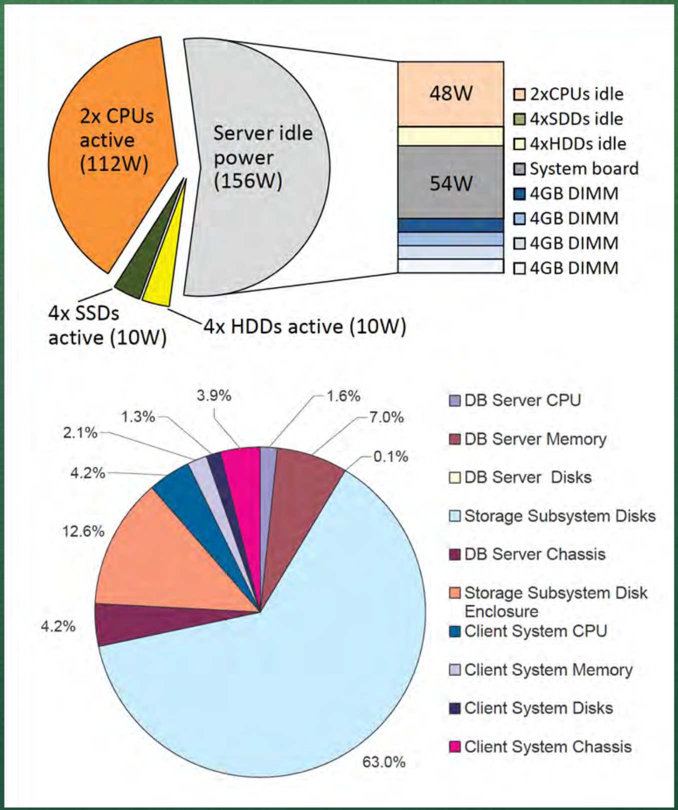



Figure 6. Power breakdown of two different server systems. Sources: Poess & Nambiar, 2008 (up), Tsirogiannis et al, 2010 (lower)

# Power-Aware Query Optimization

- Evaluates query plans by their performance and power cost
- Given (estimated) number of operations (operations vector **o**) and per-operation power cost **c** (cost vector), total power cost is?
- Modeling power cost

$$P = \mathbf{oc}^T$$

  - main challenge: estimations of **o** and **c** may be inaccurate
  - our solution: minimizing errors by *online model estimation*
- Query evaluation cost function – mapping the performance T and power cost P of a plan to a scalar for easy comparison
  - different cost functions – different tradeoffs between T and P (Fig. 3)
  - our proposal: $C = PT^n$
    - parameter $n$ - a handle for tuning P-DBMS towards desired tradeoff

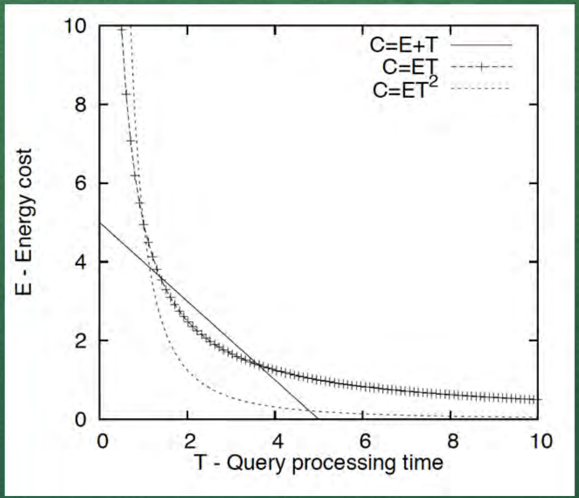



Figure 7. Pareto curves formed by different cost functions. Each plan is modeled as a point in this 2D space.

# Active CPU Mode Control

- We cannot (accurately) predict future CPU load
- A solution based on feedback control
  - System input: DVFS level, output: performance
  - Derive dynamic model by system identification
  - Basic PID controller
- Problem is more interesting if combined with query optimizer
  - Input becomes 2D: cost function parameter n and DVFS level
  - Our approach: MISO feedback loop using Model Predictive Control

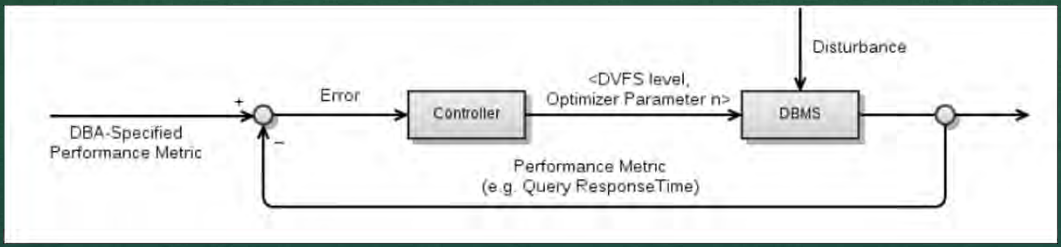



Figure 8. Feedback loop for energy-efficient CPU usage.

# Data Storage Management

- Goal: create energy-saving opportunities for low-level disk power management mechanisms
  - *Load consolidation* vs. *load balancing*
- A solution based on popularity-based data fragmentation
  - Organize data records into chunks based on their popularity
  - Periodically migrate data to consolidate load
  - Load consolidated disks run at high speed (*e.g.*, disk 1 in Fig. 9b)
  - Other disks can spin down (*e.g.*, disk 2 in Fig. 9b) or be turned off (*e.g.*, disk 3 in Fig. 9b)

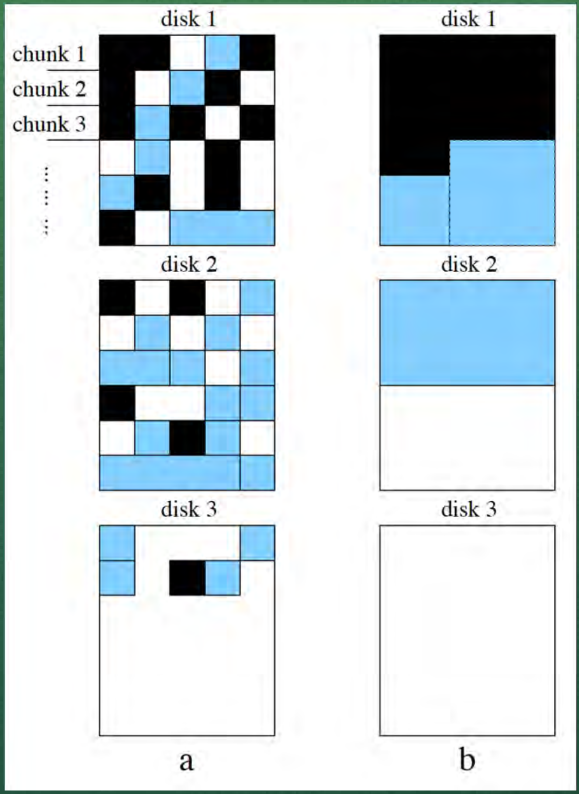



Figure 9. Data placement in (a) regular database systems and (b) PDBMS. Depth of color represents data popularity.

# Preliminary Results

- Single server, heavily loaded
- Workloads generated from well-established TPC-H and TPC-C benchmarks
- Power-Aware optimizer + DVFS controller
- Implemented in real-world DBMS – PostgreSQL
  - Up to 27% power reduction
  - Up to 13% total energy saving (ignoring cooling cost)
- Power-Aware storage manager
  - There are “sweet zones” where energy savings are large yet performance is acceptable

# Savings from CPU Control

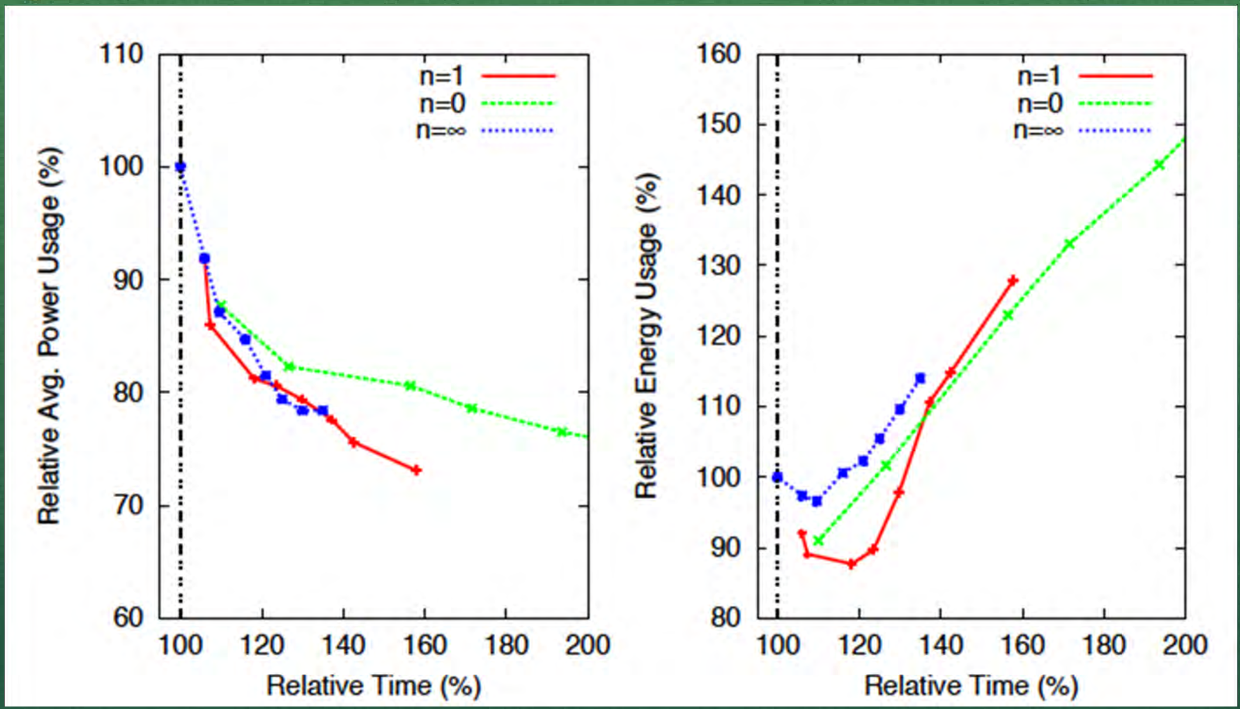



Figure 10. Power (left) and energy (right) consumption of processing the same workload under different optimizer parameters. Each experiment is plotted as one point in the 2D space consisting of the power/energy cost and workload processing time. All numbers are relative to that of the baseline experiment where the original PostgreSQL is tested.

# Savings from Storage Systems

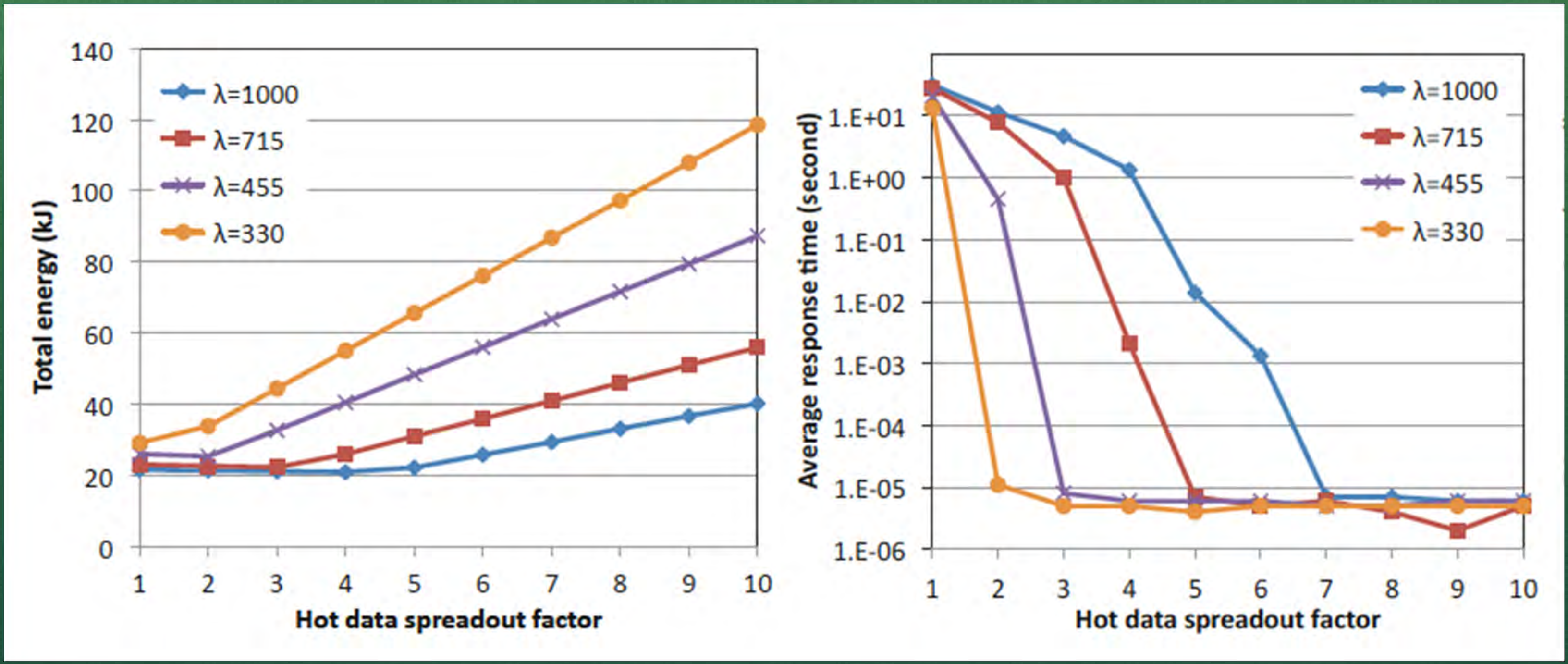



Figure 11. Energy consumption (left) and performance (right) of a random read workload under different consolidation levels and workload intensities. An array of 10 identical disks are simulated. Data access follows a 90/10 model, i.e., 10% of data tuples are hot (popular).

# Summary

- Green computing is of high economic and environmental significance
- Energy savings in databases can be made possible by
  - Identifying and executing energy-efficient plans
  - Low-power modes of hardware
- Our design of P-DBMS encompasses power-aware query optimization and dynamic control of hardware modes
- Significant energy savings are observed

# Acknowledgements


FESC USF Seed grant (Tu & Christensen)

NSF Grant IIS-001117699 (Tu, Wang, & Zeng)

USF Graduate Student Challenge Grant 2011

Student participants:

- Zichen Xu
- Peyman Behzadnia
- Wei Yuan


# Main References


1. "Report to Congress on Server and Data Center Energy Efficiency Public Law 109-431," U.S. EPA Energy Star Program, August 2007. URL: http://www.energystar.gov/ia/partners/prod_development/downloads/EPA_Datacenter_Report_Congress_Final1.pdf.
2. "EPA Announces New Computer Efficiency Requirements," Release date: 10/23/2006, Contact: Enesta Jones. URL: http://yosemite.epa.gov/opa/admpress.nsf/4a3d7e51caf96c7a85257359003f533e/113b0c0647fee41585257210006474f1!OpenDocument.
3. From Bruce Nordman, LBNL, 2009.
4. "Gartner Estimates ICT Industry Accounts for 2 Percent of Global CO2 Emissions," Gartner Group press release, 2007. URL: http://www.gartner.com/it/page.jsp?id=503867.
5. *Gadgets and Gigawatts: Policies for Energy Efficient Electronics*, International Energy Agency, Organization for Economic Cooperation and Development, 2009.
6. Meikel Poess and Raghunath Othayoth Nambiar. Energy cost, the key challenge of today's data centers: a power consumption analysis of TPC-C results. *Proceedings of the Very Large Data Bases* (*VLDB*) *Endowment*, Vol. 1, No, 2, pp. 1229-1240, 2008.
7. Z. Xu, Y. Tu, and X. Wang. "Exploring Power-Performance Tradeoffs in Database Systems," to appear in *Proceedings of 26th International Conference on Data Engineering* (*ICDE*), Long Beach, CA, March 2010.
8. W. Lang and J. Patel. "Towards Eco-friendly Database Management Systems," *Proceedings of Conference on Innovative Database Research* (*CIDR*). January 2009.
9. J. Blackburn and K. Christensen, "A Simulation Study of a New Green BitTorrent," *Proceedings of the First International Workshop on Green Communications* (in conjunction with the *IEEE International Conference on Communications*), June 2009.
10. B. Nordman and R. Brown, "Networks: Tomorrow's Big Challenge for IT Energy Use," Woodrow Wilson Center Workshop on Environment and the Information Economy: Next Steps for Research, March 15, 2004.
11. IEEE P802.3az Energy Efficient Ethernet Task Force, 2009. URL: http://www.ieee802.org/3/az/index.html.
12. TC32-TG21 – Proxying support for sleep modes, Ecma International, 2009. URL: http://www.ecma-international.org/memento/TC32-TG21.htm.