The Power Management Challenge:
Getting the next 100X improvement

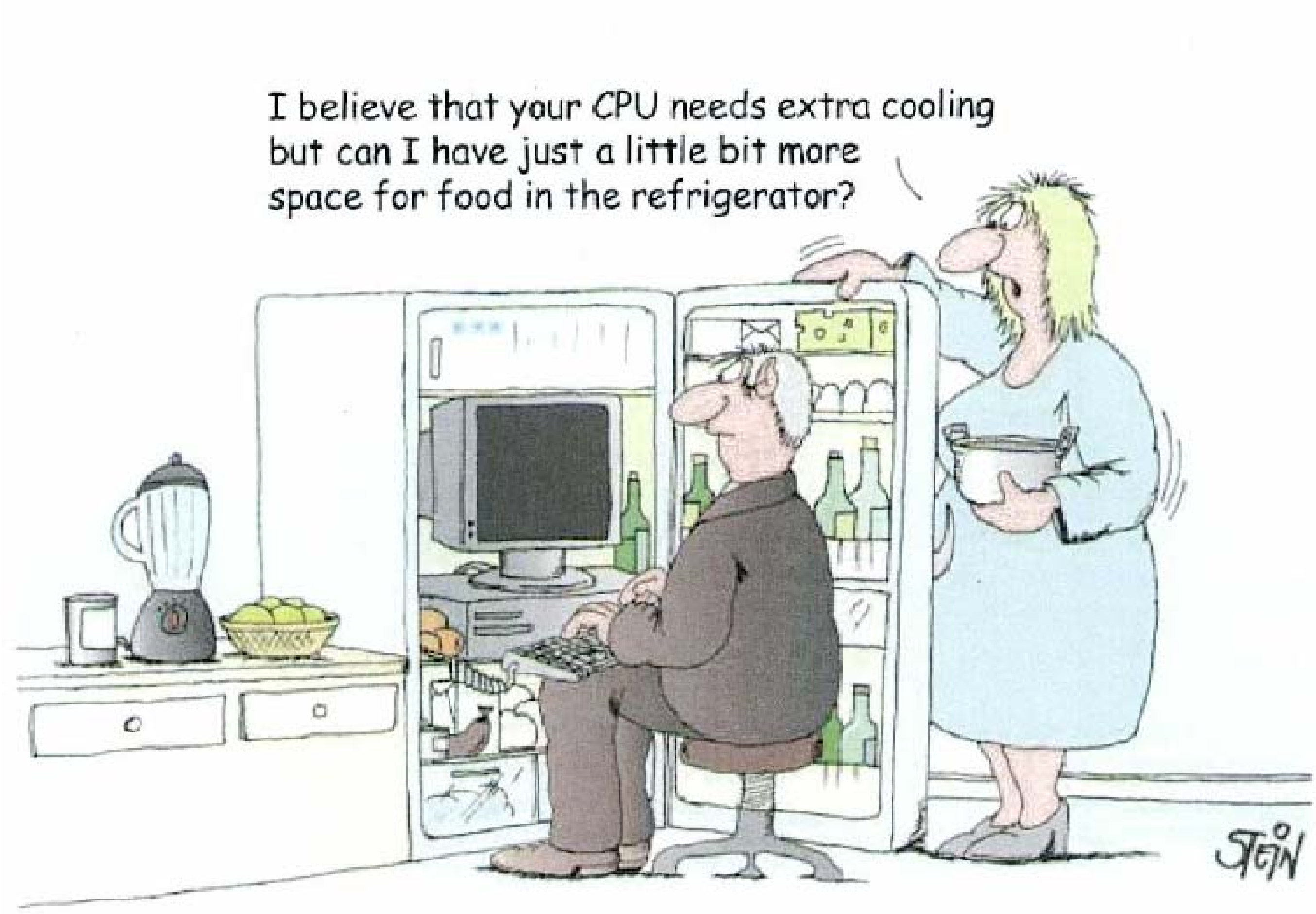
Partha Ranganathan
Senior Research Scientist, HP
Keynote, ODES-2, March 21, 2004

© 2004 Hewlett-Packard Development Company, L.P.
The information contained herein is subject to change without notice

Agenda

- The Power Management Challenge
  - Why is power important?
  - Mapping the power optimization landscape
  - Future challenges
- Opportunities for Innovation

Why is Power Important?



Why is Power Important? (1 of 2)

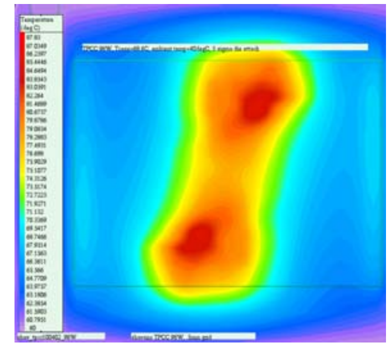
- Electricity consumption costs
  - For mobile devices, impacts battery life …
    - Gets worse with convergence/smaller form factors
  - For tethered devices, impacts electricity costs
    - Gets worse with larger data centers ($10M for 1000 racks)

- Environmental friendliness
  - Compute equipment energy use: 22M GJ + 3.9M tons CO2
  - EnergyStar (US), TopRunner (Japan), FOE (Switzerland),…

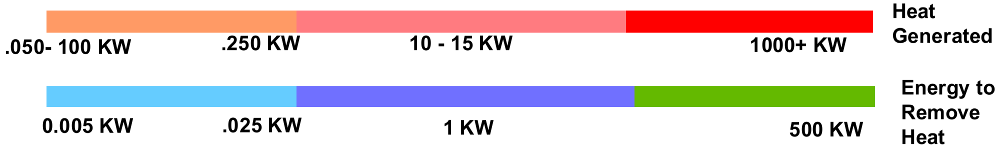
    "…goal to increase computer energy efficiency by 85% by 2005." Japan's "TopRunner" energy program, 2002

Money Isn't All You're Saving

Why is Power Important? (2 of 2)

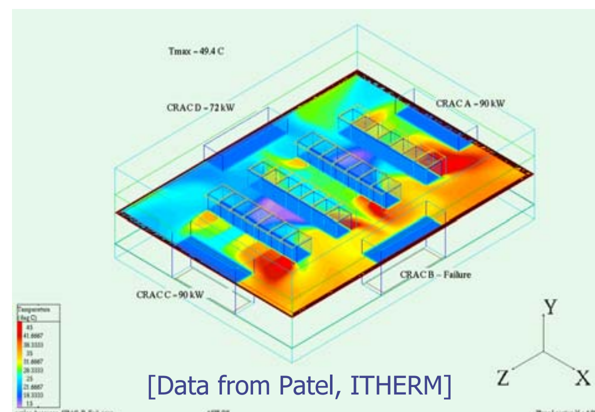


- Power delivery, packaging, and cooling costs
  - Increased cost of thermal packing
    - $1/W for CPUs > 35W [Tiwari, DAC98]
  - At high-end 1W of cooling for every 1W of power!



  - Temperature, noise issues …
  - Doctors are warning that laptop computers may inflict a burn even through clothed skin after... [The Age, November 22, 2002]

- Compaction, density, and reliability
  - Thermal failover limits



[Data from Patel, ITHERM]

Landscape of Power Optimizations

$P = C*Vdd^2*F_{0\to1} + Tsc*Vdd*Ipeak*F_{0\to1} + Vdd*I_{leakage}$

Average power, peak power, power density, energy-delay, …

CIRCUITS

- Voltage scaling/islands
- Clock gating/routing
  Clock-tree distribution, half-swing clocks
- Redesigned latches/flip-flops
  pin-ordering, gate restructuring, topology restructuring, balanced delay paths, optimized bit transactions
- Redesigned memory cells
  Low-power SRAM cells, reduced bit-line swing, multi-Vt, bit line/word line isolation/segmentation
- Other optimizations
  Transistor resizing, GALS, low-power logic

ARCHITECTURE

- Voltage/freq scaling
- Gating
  Pipeline, clock, functional units, branch prediction, data path
- Split instrucn windows
- SMT thread throttling

- Bank partitioning
- Cache redesign
  Sequential, MRU, hash-rehash, column-associative, filter cache, sub-banking, divided word line, block buffers, multi-divided module, scratch
- Low-power states
- DRAM refresh-control

- Switching control
  Gray, bus-invert, address-increment
- Code compression
- Data packing/buffering

COMPILER, OS, APP

- Switching control
  Register relabeling, operand swapping, instruction scheduling
- Memory access reduce
  Locality optimizations, register allocation
- Power-mode-control

- CPU/resource schedule
- Memory/disk control
  Disk spinning, page allocation, memory mapping, memory bank control
- Networking
  Power-aware routing, proximity-based routing, balancing hop count, …
- Distributed computing
  Mobile agents placement, network-driven computation

- Fidelity control
- Dynamic data types
- Power API

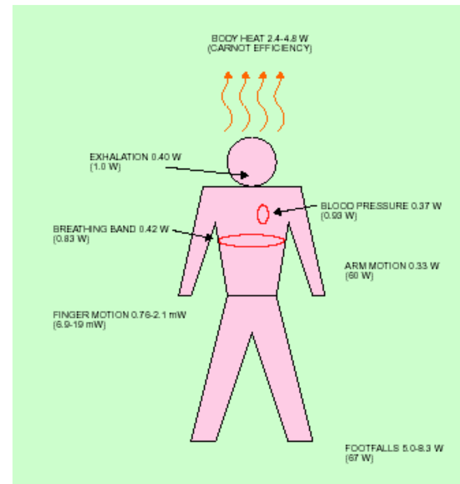
Huge body of previous work…

Is there anything left to be done?

Future Goals for Mobile Devices

Human-powered wearable computing [Starner, IBM Systems Journal,1996]

- Body heat: 2.4-4.8W
- Exhalation: 0.40W
- Breathing band: 0.42W
- Finger motion: 0.76-2.1 mW
- Blood pressure: 0.37W
- Arm motion: 0.33W
- Footfalls: 5.0-8.3W



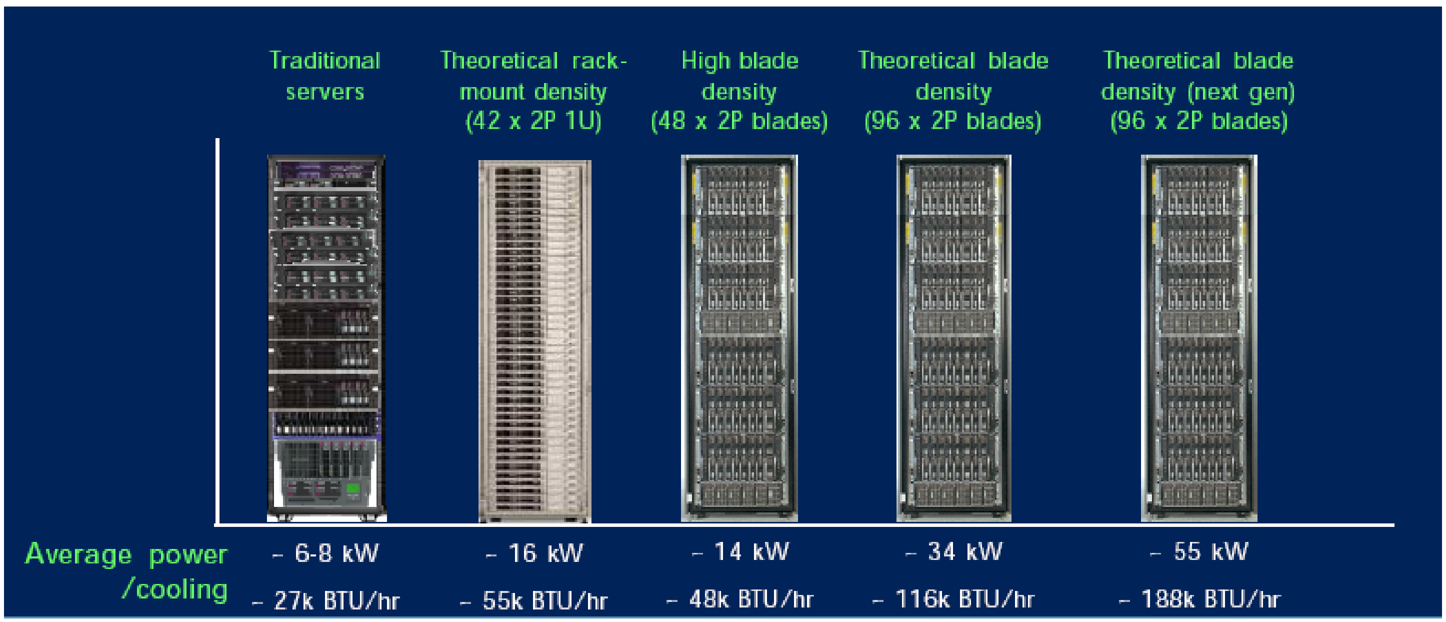
- Solar-powered [IBM2001]
  - Solar energy: 164 W/sqm
  - Conv. Efficiency 9%
  - 3 cm x 2 cm watch
  - 9.84 mW
    - Deration factor: 50 for non-sun periods

- Current power consumption – Laptop (10W), handheld (1W)

- **Need orders of magnitude improvement!**

Future Goals for Tethered Systems

- Next generation servers are going to be blades



[Data from Mouton, Server Blade Summit, 2004]

- Some of these numbers => liquid cooling in data center!
- **Need orders of magnitude improvement!**

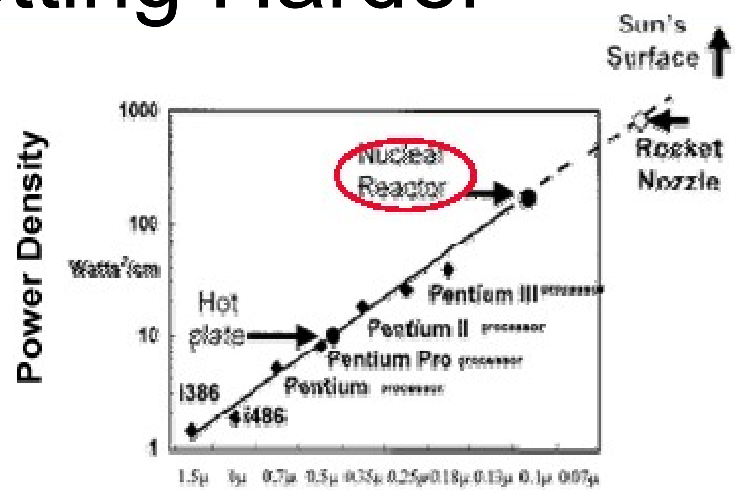
The Challenges are Getting Harder

- Technology benefits diminishing

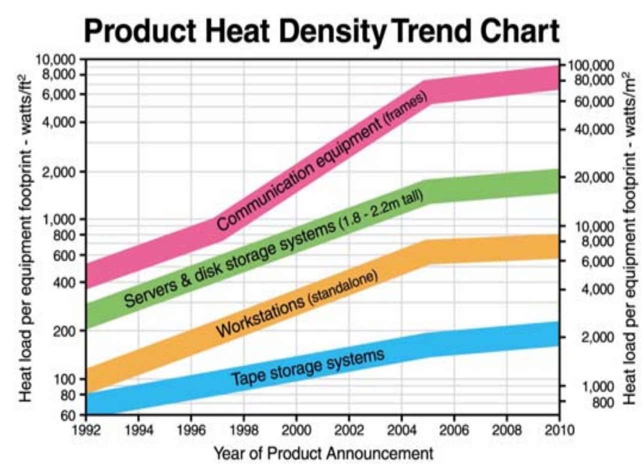
  "No Exponential is forever," Gordon Moore, ISSCC'2002

  - Technology scaling: 1/S; Voltage: 1/S; Power-dissipation/Gate: 1/S*S;
    - But overdrive, leakage, mobility, cooling limits

  - Technology scaling: 1/S; frequency increase: S; area/device: 1/S*S
    - Pipeline depth increases power; larger area means more power



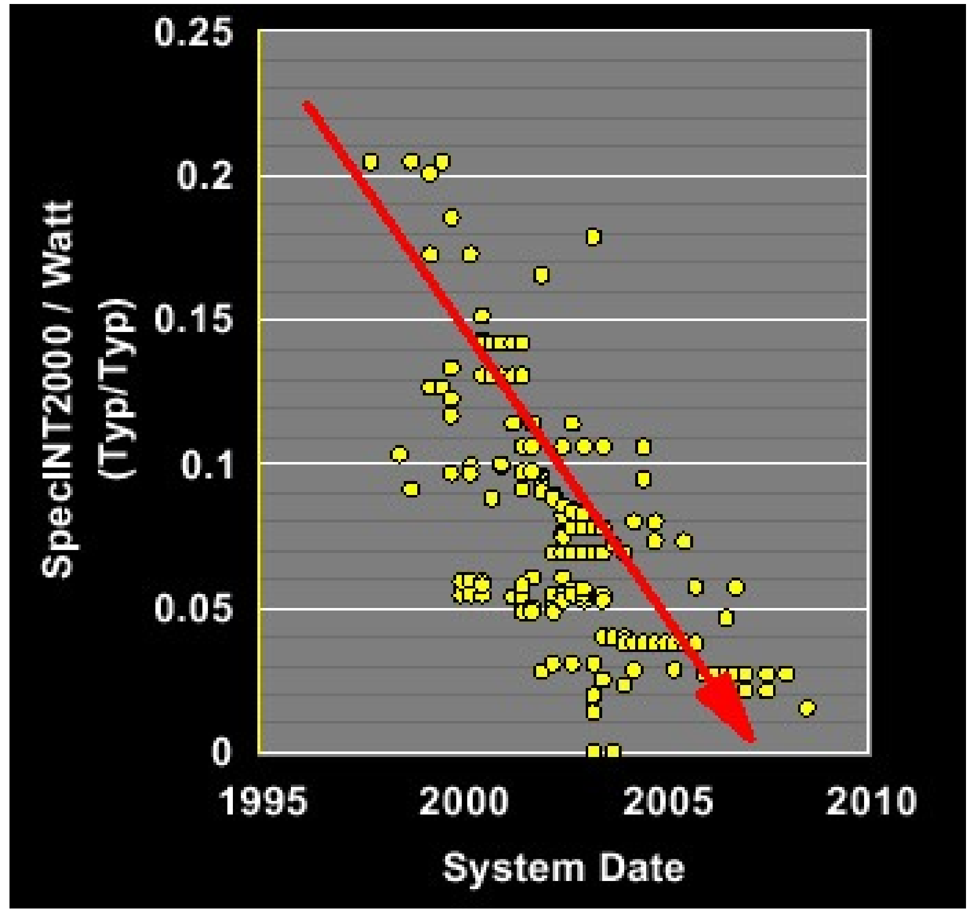
[Borkar, Intel]



[Uptime Institute]

The Challenges are Getting Harder

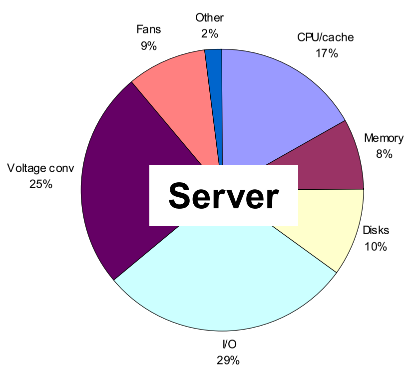
- Architecture benefits diminishing
  - Higher power overhead for lower performance deltas

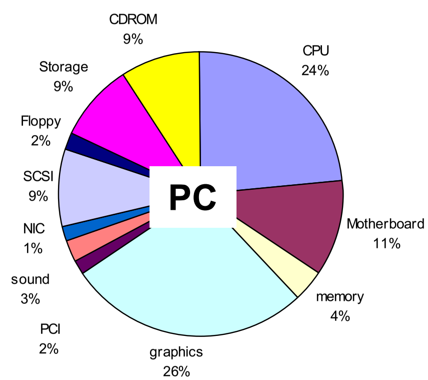


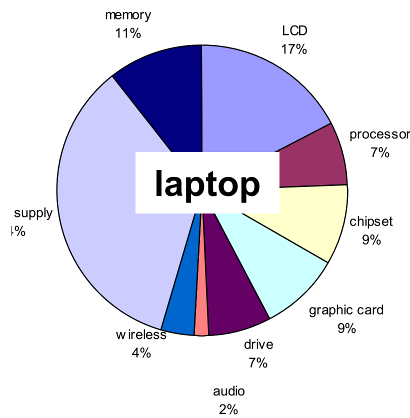
Data from [Bernstein, MICRO2003]

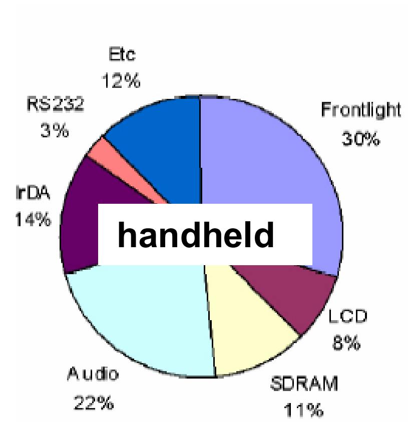
The Challenges are Getting Harder

- Non-CPU system components more power-critical
  - Displays, fans, power supply, routing, wireless radio, …
  - Many don't follow semiconductor scaling trends!









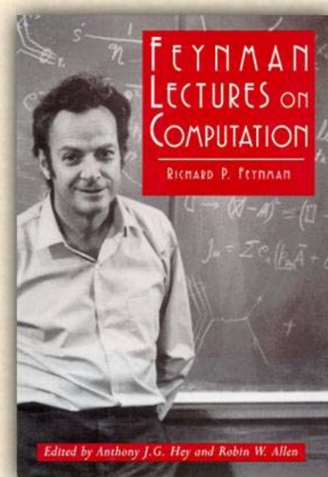
Note on data: pie-charts represent one arbitrarily chosen point in design space; percentages can vary based on vendor, generation, and configuration.
Data from http://pads.east.isi.edu/presentations/misc/sjcho-pm-report.pdf, http://www11.brinkster.com/bayup/doarticles.asp?d=a&a=61, [Bose2003, PACS]

Hard goals and hard challenges…

Is there anything to be done?

You can't beat physics, but you can work with it!

- Limit to power cost of information transfer quite low
- Huge difference between theoretical & current numbers
  - Why?



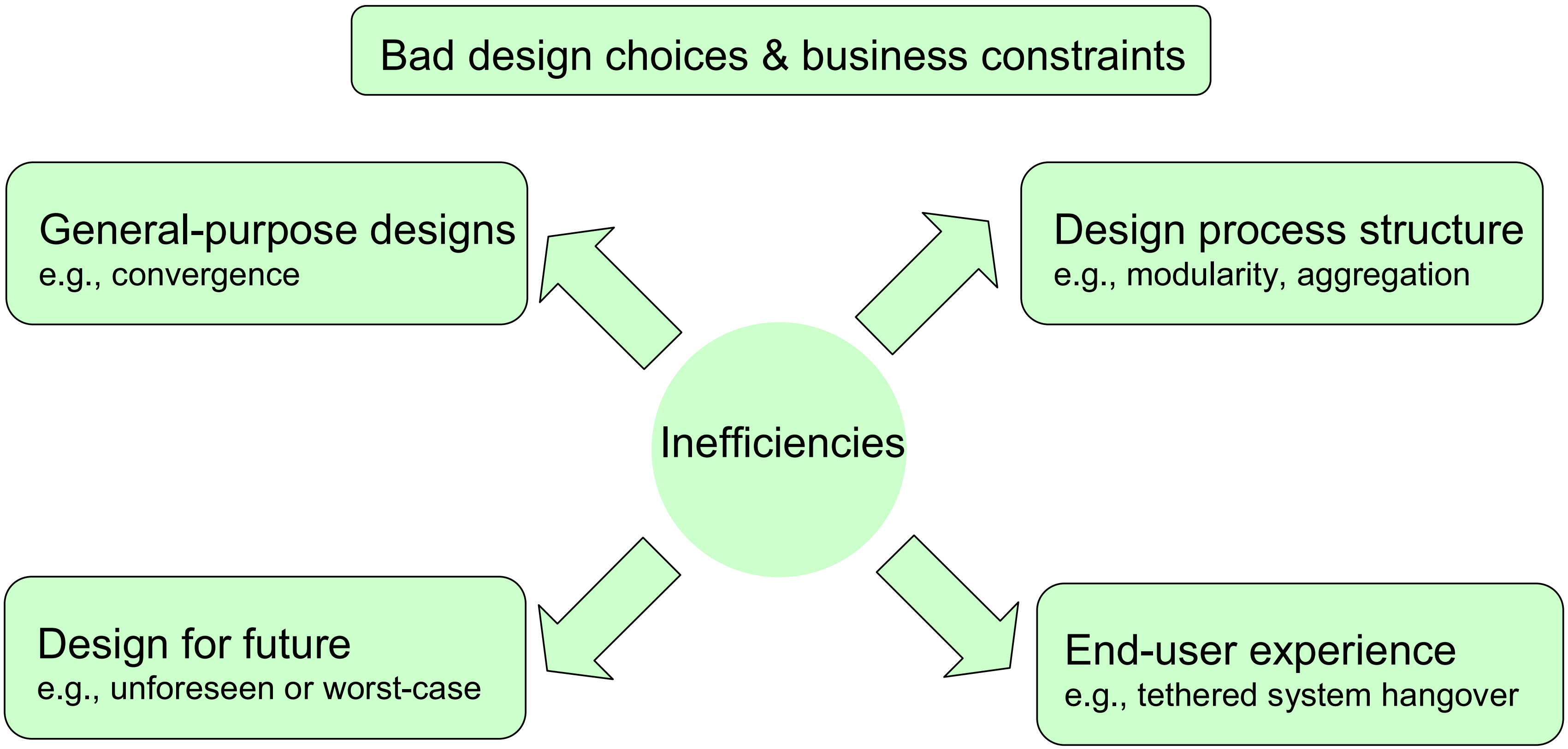
$$P = k_B T \frac{d}{c} n^2$$

$$n = \sqrt{\frac{cP}{k_B T d}}$$

10^{18} bit-ops/sec for 1 W

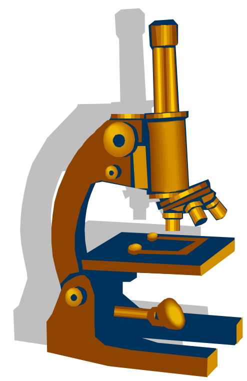
1 billion Pentiums in a handheld device!

It is all about inefficiencies!!



Getting the Next 100X!

- Discover inefficiencies
  - Understand and quantify trends



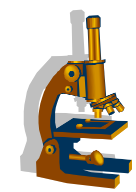
- Invent new system designs
  - Mechanisms and policies to leverage trends

Addressing inefficiencies

Special-purpose devices as surrogate lower bounds

A Recipe to Quantify Inefficiencies

- Blue-sky characterization of inefficiencies very hard
  - Different users, different needs, different times, different applications, different components, different engineering costs, different metrics, …
  - Feynman equation is too abstract



- Use special purpose device optimized for task as reference
  - Representative of acceptable tradeoffs in functionality, performance/power metrics and choice of components
  - Incorporate insights on user needs and application behavior
  - Help identify opportunities for policies/mechanisms to exploit trends

A Quick Experiment

- Energy comparison for a spectrum of mobile devices

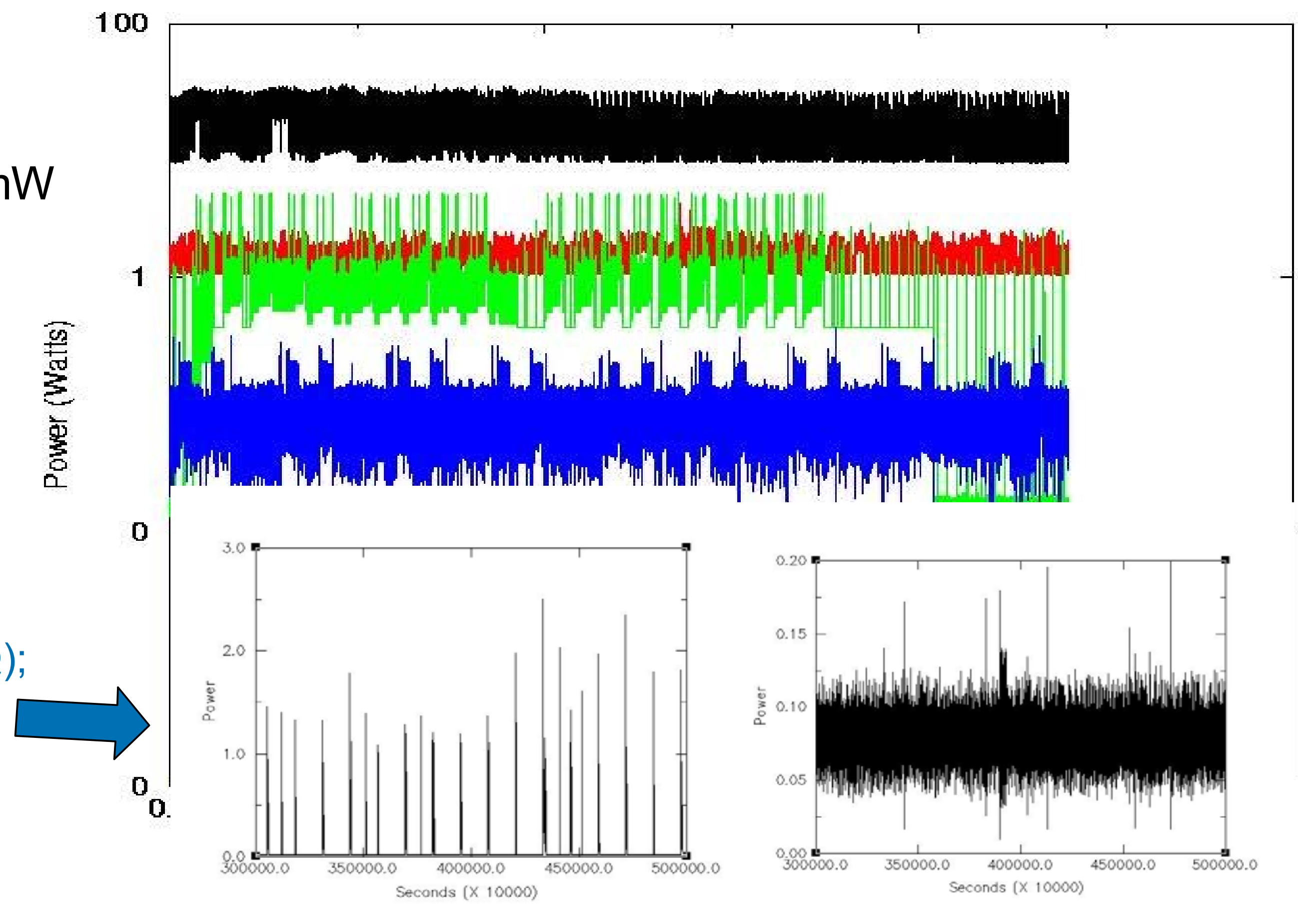
- Devices:
  - Laptop (Armada M300), PDA (iPAQ 3630)
  - Cell phone (Nokia 8260), Pager (Blackberry W1000), High-end MP3 (Nomad jukebox), low-end MP3 (ipaq PA1), voice-recorder (VoiceltVT90)

- Benchmarks representative of typical mobile workloads
  - Email, text messaging, phone calls, web browsing
  - MP3 play-back, text notes, audio notes, games, idle mode
  - Benchmarks structured to have core functionality consistent

- Measurement – data acquisition of current/voltage
  - Total energy for task
  - Temporal power signatures

Energy Comparison for Email

- Laptop: 165X
- Handheld: 15X
- Cell phone: 6X
- RIM pager: 92 mW



Data from [PACS2003]

Radio wakeup:
- 100ms (iPAQ);
- 1.2 sec (cell)
- 5 sec (RIM)

Other Benchmarks

	Email		MP3			Notes		Messaging		
Device	Rcv	Reply	Speaker	Headphone	Browse	Text	Audio	Text	Audio	Idle
Laptop	15.16 W	16.25 W	18.02 W	15.99 W	16.55 W	14.20 W	14.65 W	14.40 W	15.50 W	13.975 W
Handheld	1.386 W	1.439 W	2.091 W	1.700 W	1.742 W	1.276 W	1.557 W	1.319 W	-	1.2584 W
Cellphone	539 mW	472 mW	-	-	-	-	-	392 mW	147 mW	26 mW
Email Pager	92 mW	72 mW	-	-	-	78 mW	-	-	-	13 mW
High-end MF	-	-	-	2.977 W	-	-	-	-	-	1.884 W
Low-end MP	-	-	-	327 mW	-	-	-	-	-	143 mW
Voice Recor	-	-	-	-	-	-	166 mW	-	-	17 mW
variance	*16496%*	*22727%*	*861%*	*4890%*	*950%*	*18252%*	*8825%*	*3673%*	*1351%*	*107500%*

- Wide variation in power
  - iPAQ 5X-10X higher energy; Laptop 10X-100X higher energy
  - Huge potential from addressing energy inefficiencies in laptop and handheld
- Using special-purpose devices as surrogate lower bounds
  - Variations related to better task-specific system design
    - Avoid wasted energy (CPU scaling)
    - Redefine problem for lower-power (compute acceleration)
    - Lower power with same user experience (fidelity)

Putting it all together

Energy Scale-down

Energy Scale-Down

Energy reductions without compromising user/task requirements

Match energy consumption to user requirements through adaptivity in hardware/software

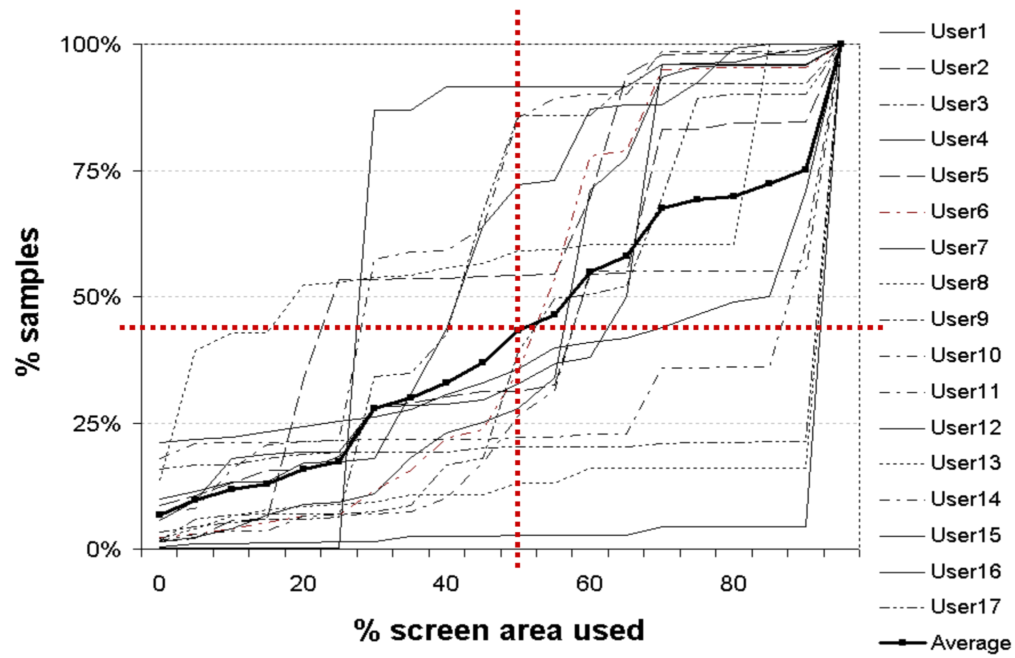
Example: Display Scale-Down

- Displays consume significant power in mobile systems
  - 50% on laptops[7], 61% on handhelds[1]

- Previous approaches:
  - Turning off the entire display
  - Using lower quality or smaller sized displays

- Scale-down approach: energy-adaptive displays
  - Match display power consumption to screen content for user/task

Discovering User Requirements



[User study: study screen usage behavior of 17 Windows users ~100 days, Mobisys2003]

- Display capacities are not fully utilized
  - On average, ~60% of screen area used (window-of-focus)
  - Other functions of display are not used always (color, res., …)
- BUT display power constant all the time!

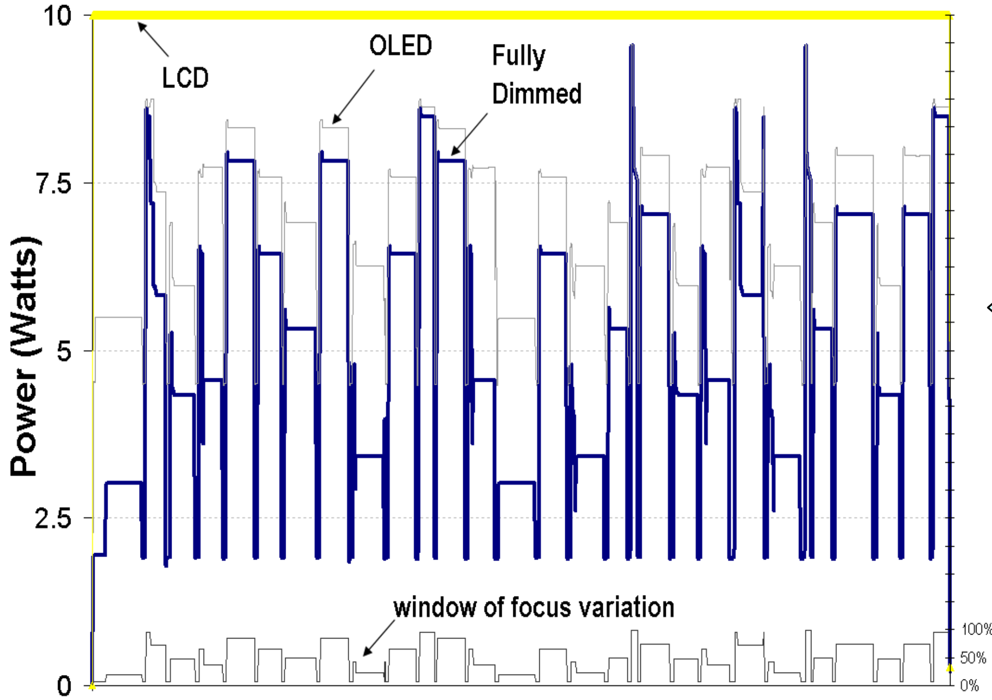
Designing Energy-Adaptive Displays

- Energy-adaptive Hardware Design
  - Organic-light-emitting-diode (OLED) displays
    - Power based on pixel value (brightness, color)
    - Currently in cell phones, expected in handhelds/laptops 2004-5

- Energy-adaptive Software Design
  - Energy-aware screen usage with less emphasis on low-interest areas
    - Four simple interfaces assuming window of focus user interest

background dimmed, background dark, background gray-scale, background green-scale

Benefits: Average User Trace

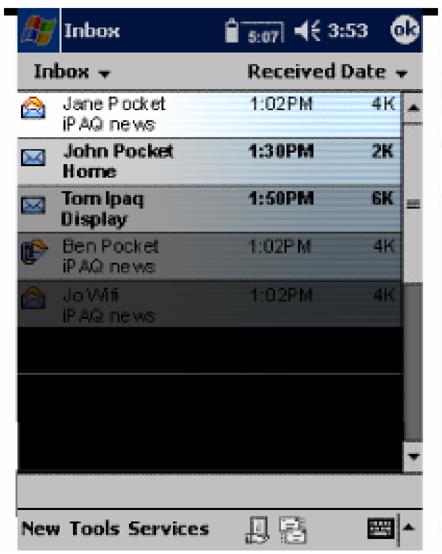


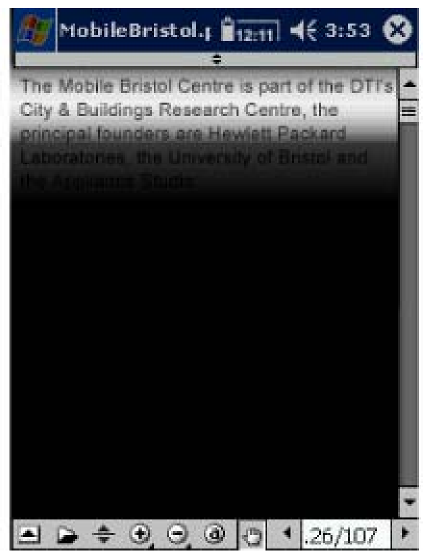
Display energy proportional to screen usage

- Significant energy savings
- Energy adapts to use scenarios

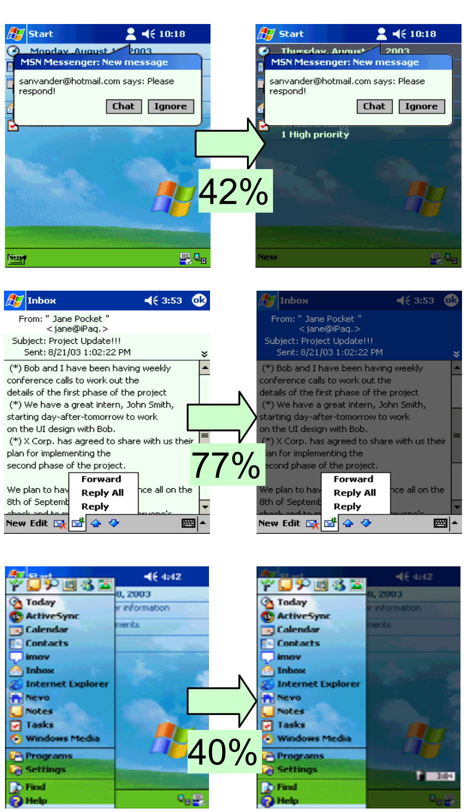
Other Energy-Aware Screen Designs

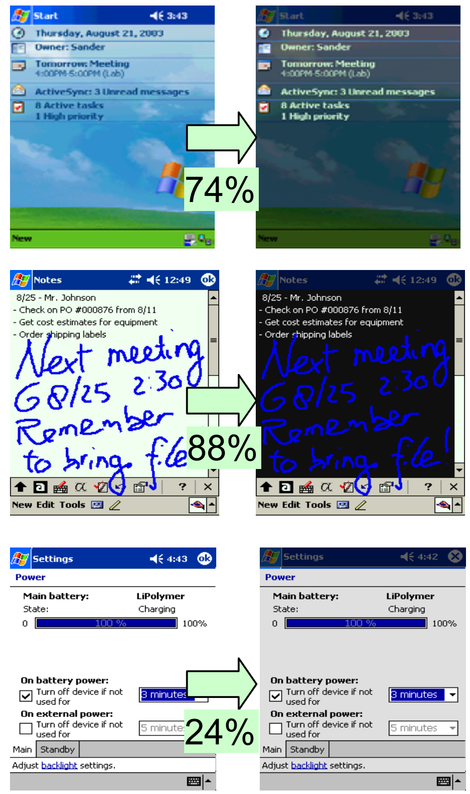
- Going beyond simple energy-aware user interfaces
  - [Spatial] Focus on informational content
  - [Temporal] Focus on content of interest at given time

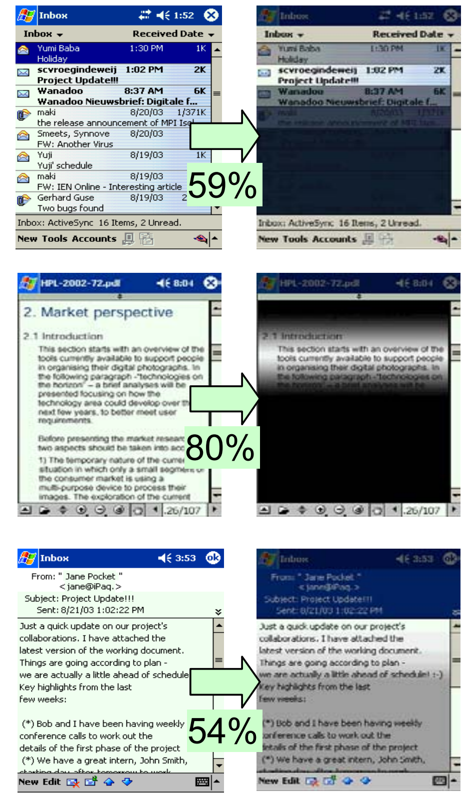




User Acceptance Study







High acceptance

High-Med acceptance

Med-Low acceptance

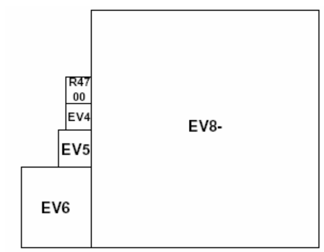
Data from [CHI2004]

New displays reduce energy and have better ease-of-use

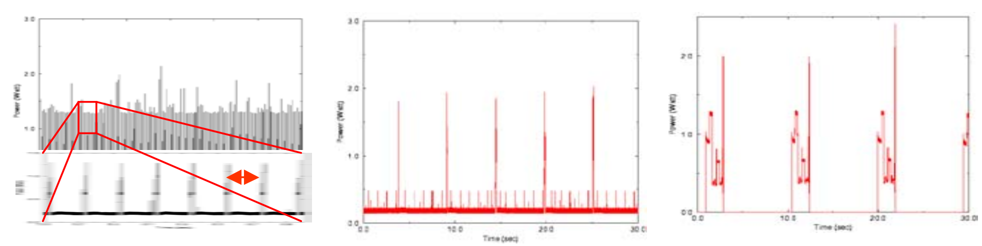
Other Energy Scale-Down Approaches

- Energy-aware user interfaces: energy benefits with acceptable user interfaces

- Heterogeneous CMPs: energy benefits with acceptable performance

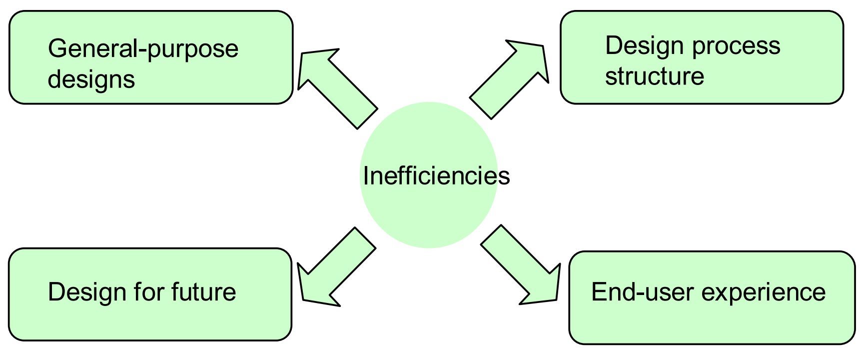


- Wireless scale-down: energy benefits with acceptable response delays



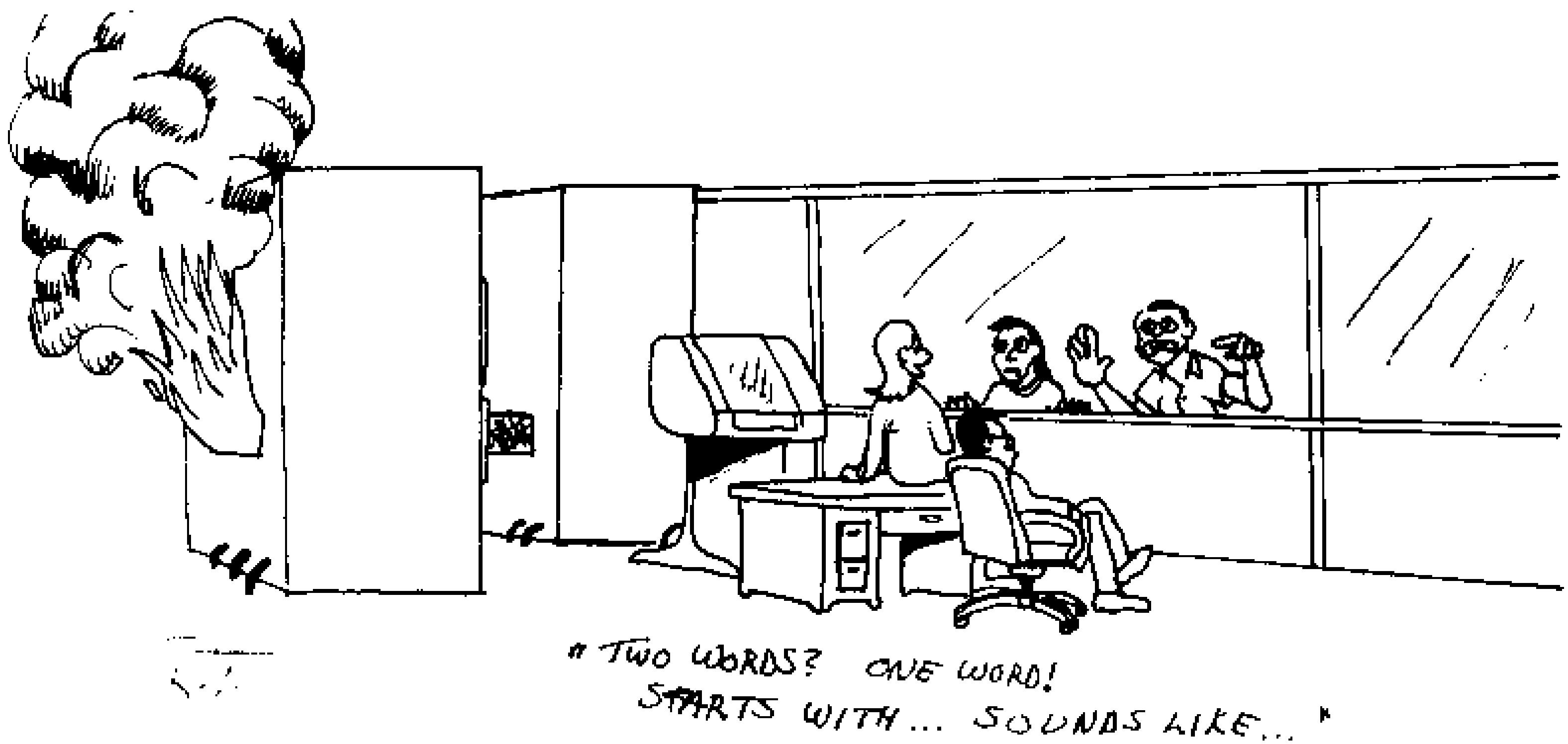
Summary

- Power critical from embedded systems to data centers
  - Huge body of previous work, BUT ways more to go
- Huge opportunity: *"it is all about inefficiencies!"*
  - Space for new discoveries and new inventions



- Energy Scale-Down: lower power & better user experience
  - Use special-purpose implementations as best practices' reference

Questions?



asuwlink.uwyo.edu/ ~jimkirk/