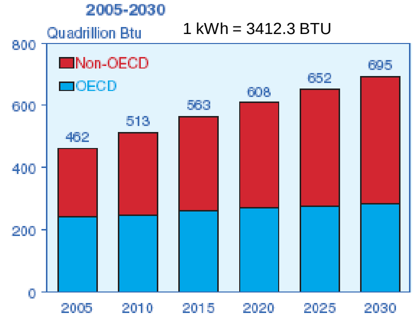
Green Computing:

Reducing the Carbon Footprint of Information Processing Systems in an Energy-Constrained World

Massoud Pedram
University of Southern California
Dept. of Electrical Engineering
August 3, 2009
ICCCN 2009

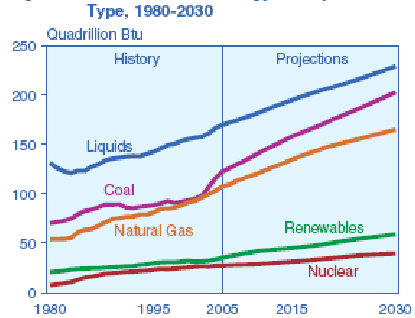
Energy Usage Worldwide

Figure 1. World Marketed Energy Consumption, 2005-2030



Sources: **2005:** Energy Information Administration (EIA), *International Energy Annual 2005* (June-October 2007), web site www.eia.doe.gov/iea. **Projections:** EIA, World Energy Projections Plus (2008).

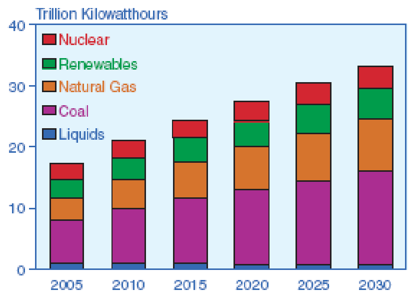
Figure 2. World Marketed Energy Use by Fuel Type, 1980-2030



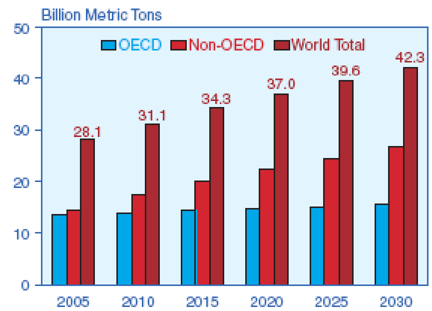
Sources: **2005:** Energy Information Administration (EIA), *International Energy Annual 2005* (June-October 2007), web site www.eia.doe.gov/iea. **Projections:** EIA, World Energy Projections Plus (2008).

World marketed energy consumption is projected to increase by 50 percent from 2005 to 2030. Total energy demand in the non-OECD countries increases by 85 percent, compared with an increase of 19 percent in the OECD countries.
- Energy Information Administration / International Energy Outlook 2008

World Electricity Generation and Carbon Dioxide Emission



Sources: **2005:** Energy Information Administration (EIA), *International Energy Annual 2005* (June-October 2007), web site www.eia.doe.gov/iea. **Projections:** EIA, System for the Analysis of Global Energy Markets/Global Electricity Module (2008).

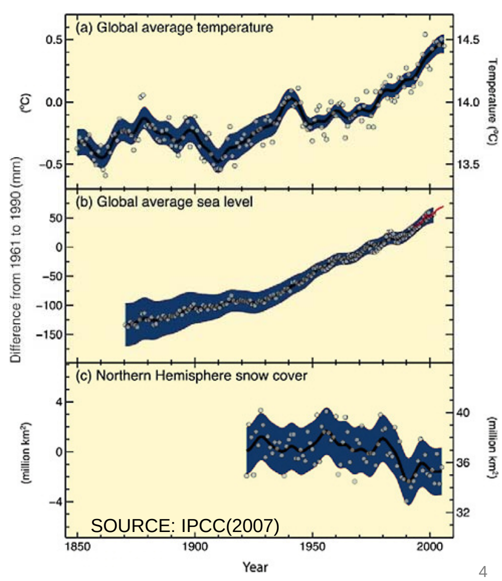


Sources: **2005:** Energy Information Administration (EIA), *International Energy Annual 2005* (June-October 2007), web site www.eia.doe.gov/iea. **Projections:** EIA, World Energy Projections Plus (2008).

Projections by the DoE's Energy Information Administration show that worldwide electric power demand will increase from the current level of about 2 Terawatts (TW) to 5 TW by 2050.

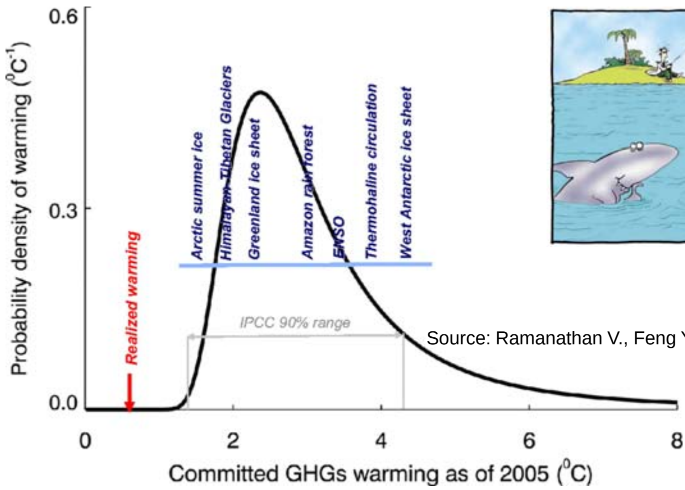
Evidence of Global Warming

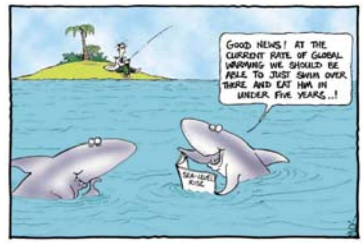
- 20th century saw about a 0.6-degree centigrade increase in global and ocean temperatures.
- Over that same period, sea level rose about 150 millimeters
- Deforestation and the burning of fossil fuels have increased the amount of carbon dioxide in the atmosphere from about 300 parts per million in 1900 to about 380 parts per million today.



The Planet Is Already Committed to a Dangerous Level of Warming

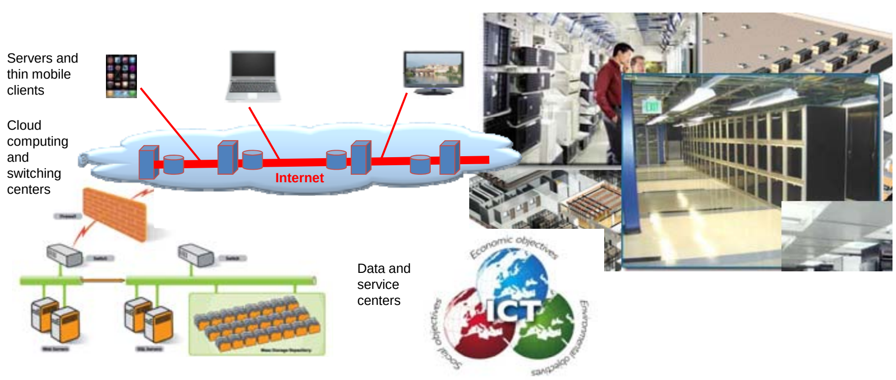
Probability distribution for the committed warming by GHGs between 1750 and 2005





Source: Ramanathan V., Feng Y. PNAS 2008

The ICT Ecosystem

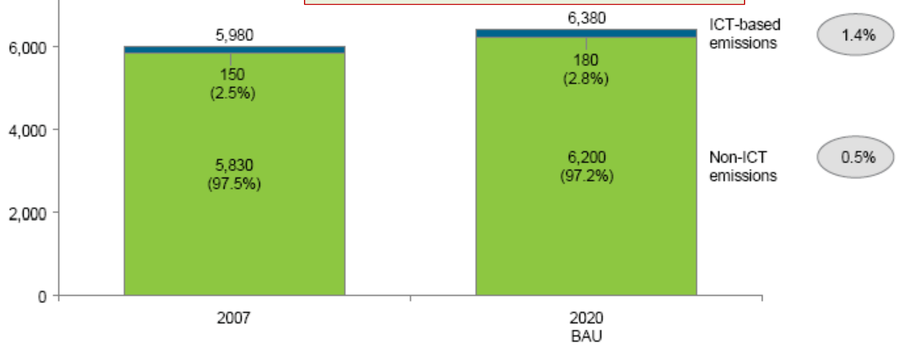


The ICT ecosystem encompasses the policies, strategies, processes, information, technologies, applications and stakeholders that together make up a technology environment for a country, government or an enterprise.

--Source: Harvard Law

ICT Industry's Carbon Emissions in the U.S.

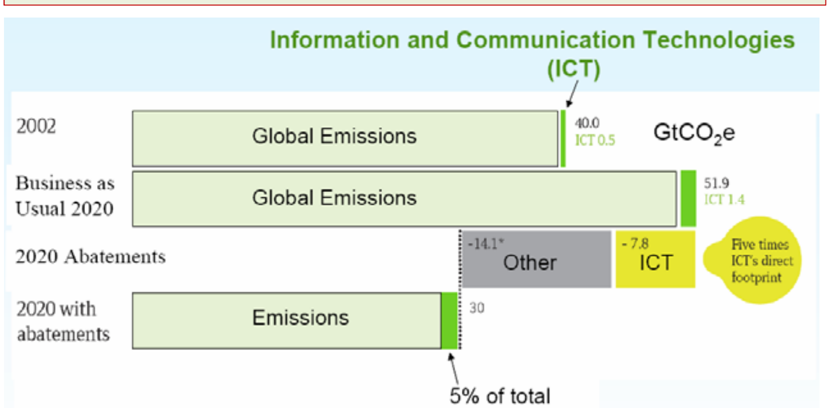
The ICT sector must ensure its own carbon footprint grows slower than the demand for its products.



Sources: GeSI and the Climate Group, "Global Smart 2020 report", 2008; EIA Annual Energy Outlook 2008.

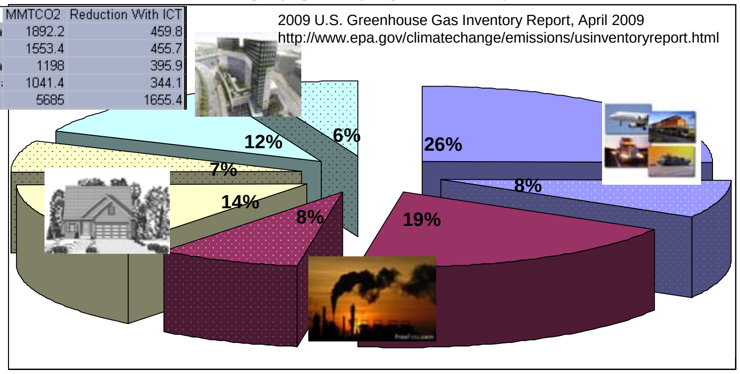
Target 2020

California's Global Warming Solutions Act of 2006
– Requires Reduction of GHG by 2020 to 1990 Levels
– 10% Reduction from 2008 Levels; 30% from BAU 2020 Levels
The European Union Requires Reduction of GHG by 2020 to 20% Below 1990 Levels (12/12/2008)



Green House Gases: Reduction Potential with IT

	MMTCO2	Reduction With ICT
Transporta	1892.2	459.8
Industrial	1553.4	455.7
Residentia	1198	395.9
Commerci	1041.4	344.1
Total	5685	1655.4



- Rising energy consumption and growing concern about global warming
  - World wide GHG emissions in 2006: 27,225 MMT of CO_2 Equivalent
- Major contributors: Transport, Industrial, Residential, Commercial
- Improved efficiency with Information Technology (IT) usage
  - 29% expected reduction in GHG
  - Equal to gross energy and fuel savings of $315 billion dollars
- Expect even wider utilization of IT in various sectors it is more energy-efficient

M. Pedram, USC 9

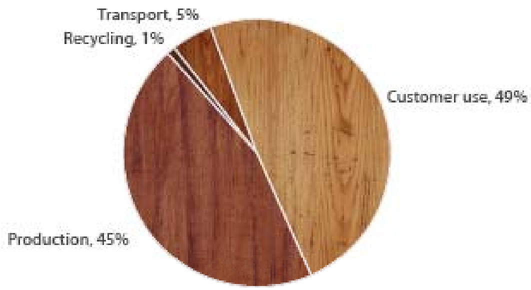
Life Cycle (Cradle-to-Grave) Analysis

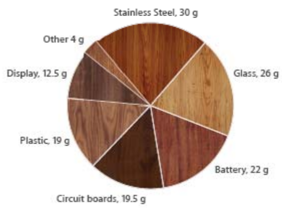
- The analysis of the full range of environmental impacts of a product requires the assessment of raw material production, manufacture, distribution, use, and disposal
- Life cycle energy analysis (LCEA) is an approach in which all energy inputs to a product are accounted for, not only direct energy inputs during manufacture, but also all energy inputs needed to produce components, materials and services needed for the manufacturing process
- Net energy content is the energy content of the product minus energy input used during extraction and conversion, directly or indirectly.
  - Different energy forms (heat, electricity, chemical energy etc.) have different quality and value
  - A joule of electricity is 2.6 times more valuable than a joule of heat or fuel

> We must find a way to meet the increasing demand for energy without adding catastrophically to greenhouse gases.
> —Ray Orbach, Undersecretary for Science, U.S. Department of Energy

M. Pedram, USC 10

Greenhouse Gas Emissions for iPhone 3G



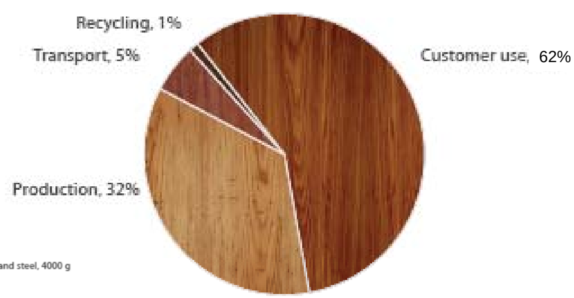


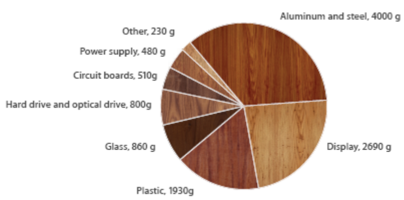
Total greenhouse gas emissions: 55 kg CO_2e

M. Pedram, USC

11

Greenhouse Gas Emissions for 24-inch iMac





Total greenhouse gas emissions: 1500 kg CO_2e

M. Pedram, USC

12

High Performance Computing (HPC)

Taking up 6,000 square feet and weighing in at 500,000 pounds, the latest version of IBM's roadrunner breaks the petaflop barrier (1.2Pflops). It consumes 2.5 MW, i.e., 488 Mflops/W.

Cores	Rmax (GFlops)	Rpeak (GFlops)
129600	1,105,000	1,456,704

1	Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz, Voltaire Infiniband
2	Jaguar - Cray XT5 QC 2.3 GHz
3	JUGENE - Blue Gene/P Solution
4	Pleiades - SGI Altix ICE 8200EX, Xeon QC 3.0/2.66 GHz

HPC applications:

Modeling of nuclear fusion

Weather modeling

Understanding climate change

Molecular modeling

Etc.

Cloud Computing

- "A style of computing where massively scalable IT-related capabilities are provided 'as a service' using Internet technologies to multiple external customers" - Gartner

Google App Engine,

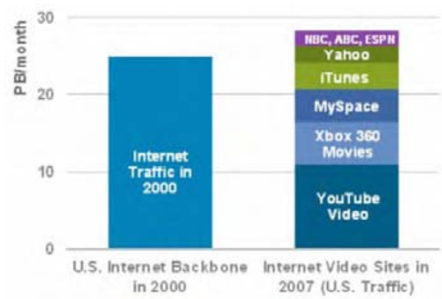
Amazon EC2 and S3,

Microsoft Windows Azure ,

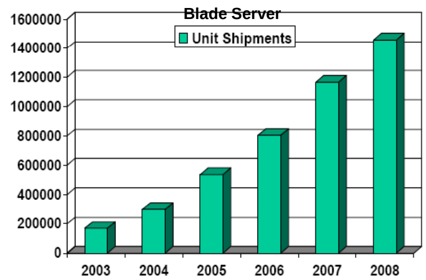
Yahoo! Cloud Services ,

…

Internet Growth is Driving Data Center Usage



Source: ADC

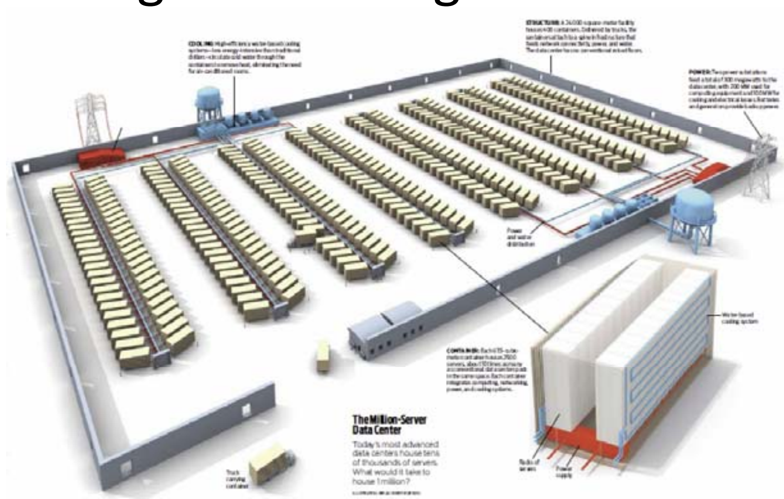


Increase from 500K to 1.45mn in 3 years

Source: Gartner

Corporate data growing 50 fold in three years.
—2007 Computerworld

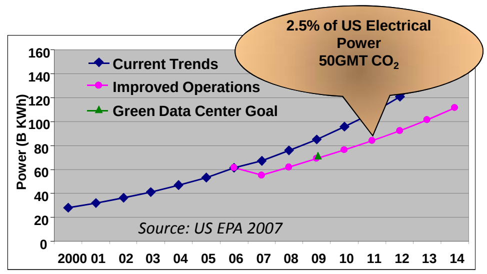
Emergence of Mega Data Centers



Microsoft's Chicago Datacenter: Entire first floor is full of containers; each container houses 1,000 to 2,000 systems; 150 - 220 containers on the first floor.

Green Data Centers: A Strategic Necessity

- Three drivers have led the datacenter crisis
  - Insatiable IT demand
  - Power-limited core technology
  - Increasing energy costs
- Datacenters consume ~ 2 % of all US electricity
- *Annual growth* (15%) is non-sustainable
- Datacenter power projected to be > 8 % of US power by 2020
- Need a paradigm shift in data center computing to put us on a more sustainable/scalable ICT energy efficiency curve



M. Pedram, USC 17

Peak Power Demands of Data Centers

- Apart from the total energy consumption, another critical component is the peak power; the peak load on the power grid from data centers is currently estimated to be approximately 7 gigawatts (GW), equivalent to the output of about 15 baseload power plants.
  - This load is increasing as shipments of high-end servers used in data centers (e.g., blade servers) are increasing at a 20-30 percent CAGR.
  - If current trends continue, power demands are expected to rise to 12 GW by 2011.
  - Indeed, according to a 2008 Gartner report, 50 percent of data centers will soon have insufficient power and cooling capacity to meet the demands of high-density equipment.

M. Pedram, USC 18

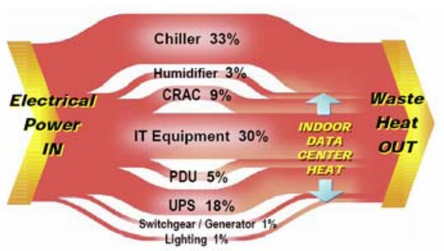
Reducing Oil Dependence

- US imports about 3.5million barrels of crude oil per day from the Middle East and Venezuela
  - **Energy content of a barrel of oil is ~1,700 kWh**

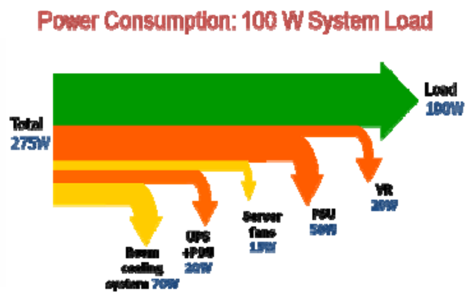
- As of 2006, the electricity use attributable to nation's servers and data centers is estimated at 61B kWh
  - This is projected to double by 2011 and triple by 2015
- Assume a moderate reduction of 25% in energy consumption of the U.S. data centers
  - Reduces U.S. foreign oil imports by 98,000 barrels a day!!
- With wider adoption of green ICT, even higher reduction of oil imports will be achieved
  - 1kWh energy consumed in a data center can help eliminate x >1 (e.g., 5 to 10) kWh of energy in other sectors
  - Although the carbon emission due to every kWh of electrical energy consumed in the U.S. varies depending on the type of power generator used to supply power into the power grid, **an average conversion rate of 0.433 kg CO_2 emission per kWh of electrical energy** may be assumed

M. Pedram, USC 19

Data Center Power Distribution



Source: IBM

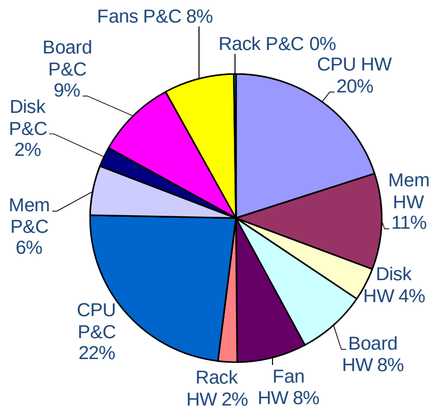


Source: Intel

More than 30 millions physical servers currently installed: growing 4X over the next 10 years.
For every $1 spent on server hardware, 50 cents is spent on power distribution and cooling.
— IDC, May 2008

M. Pedram, USC 20

Where Does the Power Go in a Blade System



Source: HP

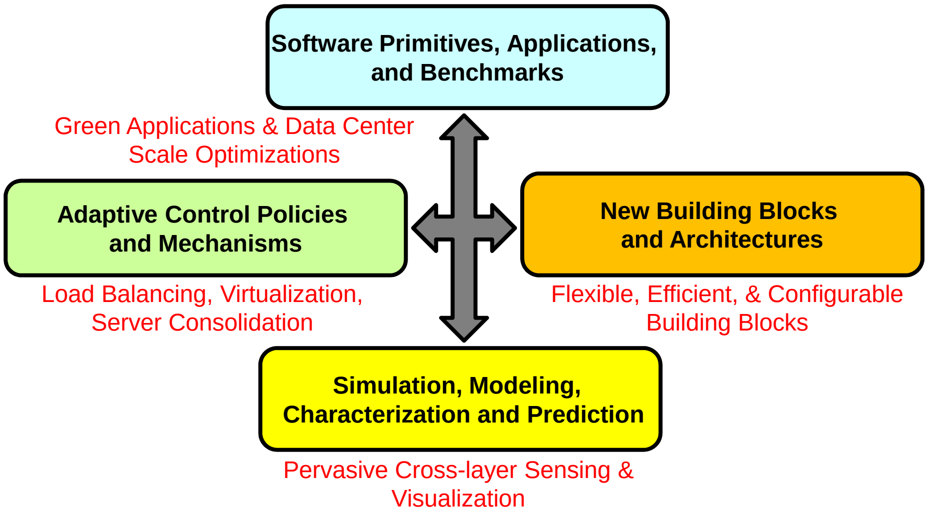
M. Pedram, USC 21

Approach to Energy Efficiency of ICT Solutions

- Need a multi-pronged approach to energy efficiency which is cross-tier, cross-area, and cross-discipline
  - Stake holders: users, industry, academics, government
  - Environmental constraints and socio-economic drivers
  - Policies and treaties
  - Technologies and applications
    - Hardware and server architectures
    - Storage
    - Networks and networking
    - Software and middleware
  - Physical infrastructure
    - Power generation (including renewable sources), The Grid – power distribution and delivery, demand shaping
    - Electromechanical, Electrochemical, lighting and Cooling/AC

M. Pedram, USC 22

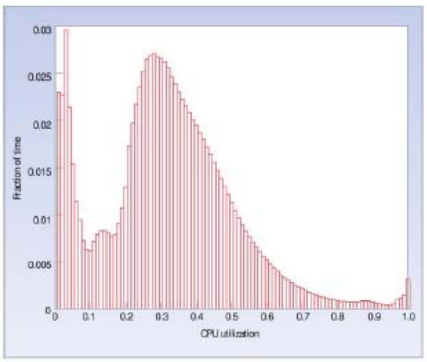
Green Data Center: Key Enablers



M. Pedram, USC 23

Server Utilization

"The Case for Energy-Proportional Computing,"
Luiz André Barroso,
Urs Hölzle,
IEEE Computer
December 2007



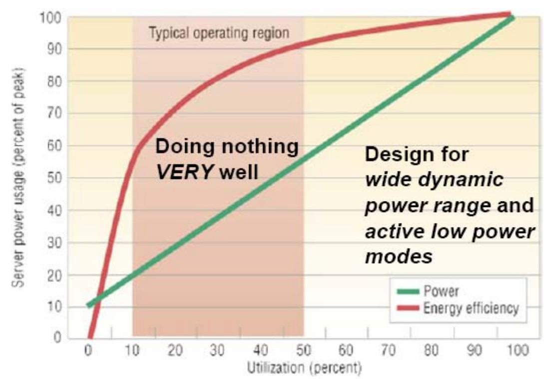
six-month period,
imum utilization,
their maximum

Servers are rarely completely idle and seldom operate near their maximum utilization, instead operating most of the time at between 10 and 50 percent of their maximum.

M. Pedram, USC 24

Energy-Proportional Servers

"The Case for Energy-Proportional Computing," Luiz André Barroso, Urs Hölzle, *IEEE Computer* December 2007

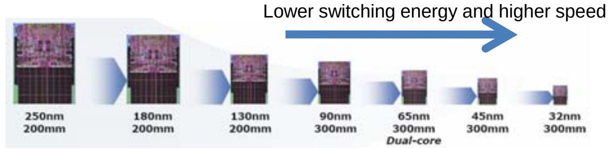


An energy-proportional server has a power efficiency of more than 80 percent of its peak value for utilizations of 30 percent and above, with efficiency remaining above 50 percent for utilization levels as low as 10 percent.

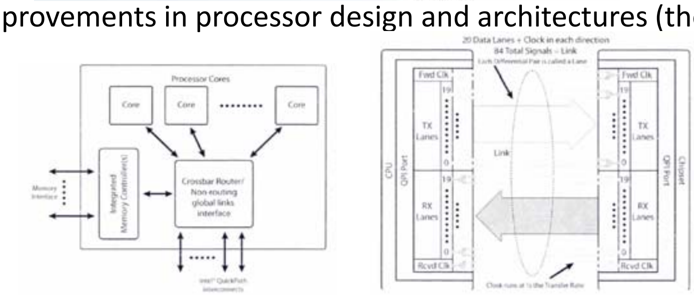
CMOS Process and Processor Design

- Moore's Law scaling (new CMOS process, the tick)

Lower switching energy and higher speed



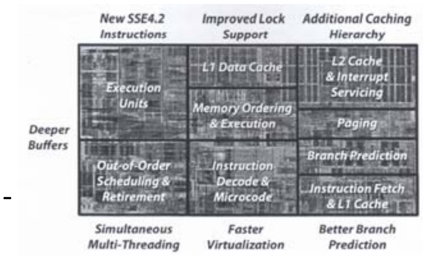
- Improvements in processor design and architectures (the tock)



Intel Quickpath Interconnect: differential current-mode signaling; Each link is composed of 20 lanes per direction capable of up to 26.5 GB/s.

Micro-architecture and Server Architecture

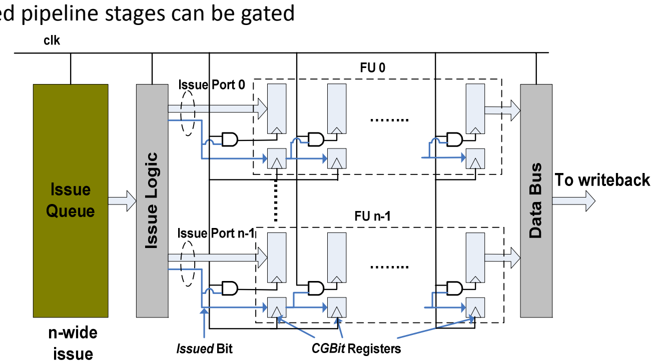
- Hyper threading (SMT) with appropriate Cache Hierarchy
- Integrated memory controller (support multiple channels of DDR3)
- PCI Express 3.0 (8Gbit/s per lane - will support 40 Gigabit Ethernet)
- On-chip(microcontroller-based) power management
- Selective speed boosting



Nahalem New Micro-architectural features

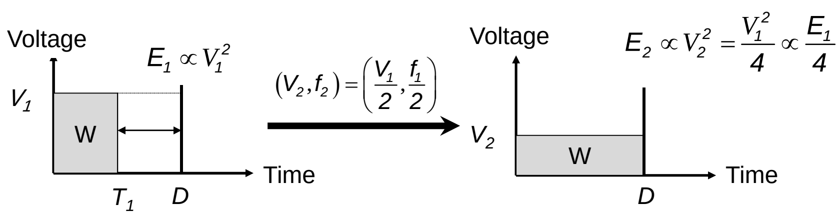
Pipelined Functional unit Clock Gating

- Not Enough Instruction Level Parallelism (ILP)
  - Typically, # of instructions issued per cycles < issue width
- Diversity of Applications
  - FU's usage varies across different programs
- Structurally pipelined FUs with multi cycle latencies
  - Not all the pipeline stages are always used
  - Unused pipeline stages can be gated



Dynamic Voltage and Frequency Scaling (DVFS)

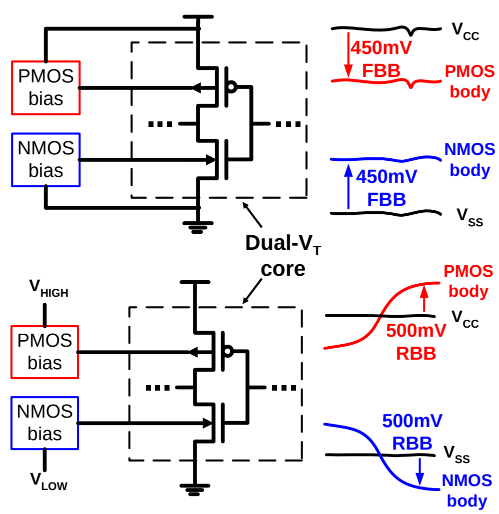
- Energy, *E*, required to run a task during T: $E = P \cdot T \propto V^2$
- Example: a task with workload *W* should be completed by a deadline, *D*



Assumes $E_{idle} = 0$

M. Pedram, USC 29

Dynamic Body Biasing for Active Mode Leakage Control



Active mode:

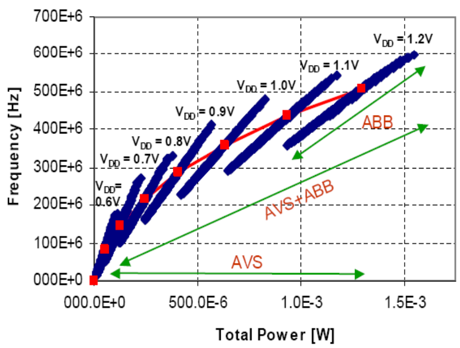
Forward body bias (FBB)

Idle mode:

Reverse body bias (RBB)

M. Pedram, USC 30

Example Performance and Power Dissipation Results



min V_{th} ≡ 0.6V forward body bias
max V_{th} ≡ 1.2V reverse body bias

AVS
- (V_{DD}= 1.2V,nom V_{th}) to (V_{DD}=0.6V,nom V_{th})
  - 24.4x power savings by 6.12x frequency reduction

ABB
- (V_{DD}= 1.2V,min V_{th}) to (V_{DD}=1.2V,max V_{th})
  - ±24% power tuning
  - ±22.6 % frequency tuning

AVS+ABB
- (V_{DD}= 1.2V,min V_{th}) to (V_{DD}=0.6V,max V_{th})
  - 127x power savings by 37.4x frequency reduction

Significant performance tuning with AVS and ABB

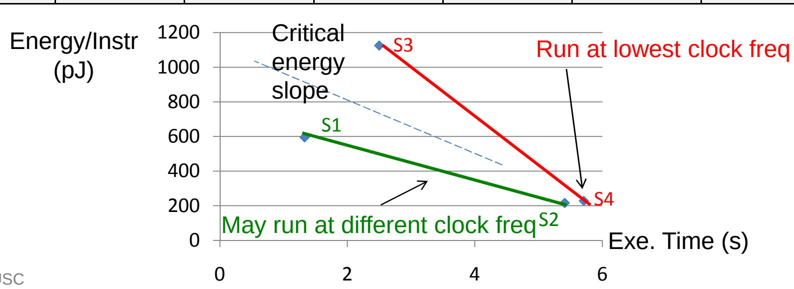
Ref CMOS090GP: M.Meijer, F. Pessolano and J. Pineda de Gyvez, "Technology Exploration of Adaptive Power and Frequency Scaling in 90nm CMOS", ISLPED2004, PR-MS-23.608

M. Pedram, USC 31

DVFS and IPC Value

- 1B instructions – Use a *critical energy slope* of 150pJ per instruction per sec

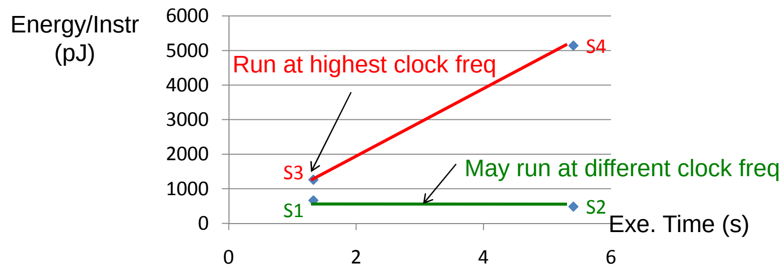
	IPC Value	Core clock frequency	MIPS	Core power	Execution Time	Energy per instruction
S1	High	600MHz	750	450mW	1.33s	595.5pJ
S2	High	180MHz	185	40mW	5.41s	216.4pJ
S3	Low	600MHz	400	450mW	2.5s	1,125pJ
S4	Low	180MHz	175	40mW	5.71s	228.4pJ



M. Pedram, USC 32

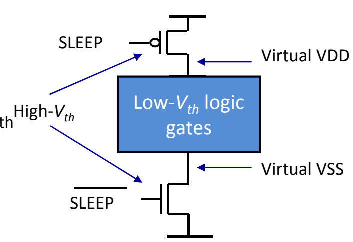
DVFS and Total System Power

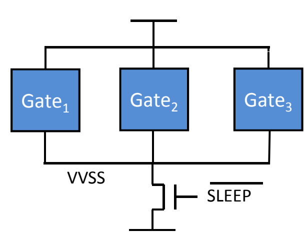
	Non-core power	Core clock frequency	MIPS (high IPC case)	Core power	Execution Time	Energy per instruction
S1	50mW	600MHz	750	450mW	1.33s	665pJ
S2	50mW	180MHz	185	40mW	5.41s	486.9pJ
S3	500mW	600MHz	750	450mW	1.33s	1,263.5pJ
S4	500mW	180MHz	185	40mW	5.41s	5139.5pJ



Power Gating (Multi-Threshold CMOS)

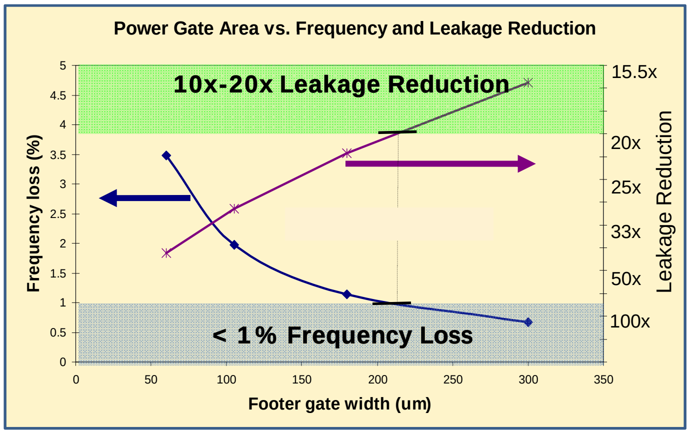
- High-V_{th} power switches are connected to low-V_{th} logic gates
  - Achieves high performance due to low-V_{th} logic gates
  - Reduces leakage power dramatically due to the series-connected high-V_{th} power switch
- Typically only a header or a footer sleep transistor is used, not both
- A single sleep transistor may be shared along several logic gates





Footer Gate Width Selection

A 2-stage pipelined 40-bit ALU (IBM)



M. Pedram, USC 35

Virtualization

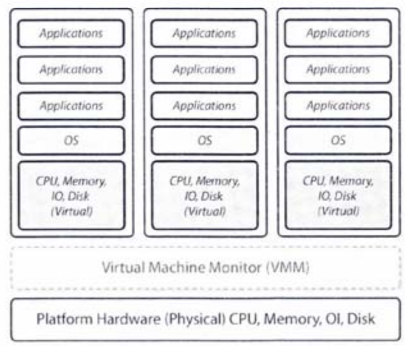
- Virtualization technologies offer a solution to two problems:
  - servers in large datacenters operate at utilization levels of 10-50%, and
  - mission-critical software systems require performance isolation

  by dividing a single physical machine into a number of virtual machines.
- Multiplexing a number of existing servers onto a single physical machine enables switching off unneeded hardware, while applications that have variable workloads can be scaled dynamically so that resource usage depend on the current service load.

M. Pedram, USC 36

Virtualization Support

- Hardware support
  - Various hardware assists (microarchitecture, chipset, networking and I/O)
  - Higher privilege ring for the hypervisor
  - Handoffs between the hypervisor and guest OS supported in hardware
  - Processor state info is retained for the hypervisor and for each guest OS in dedicated address space
  - Extended page tables
  - Virtual processor ID to avoid flushes on VM transitions



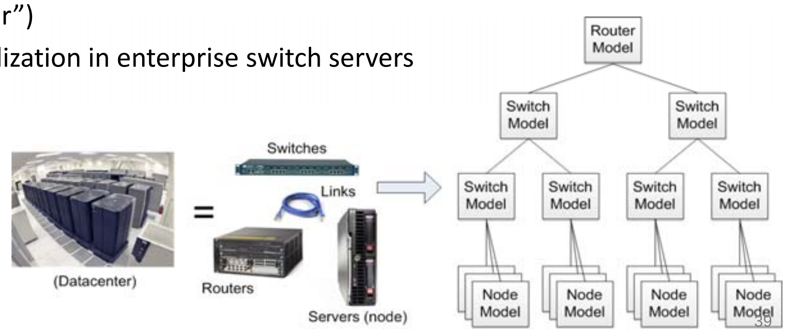
Storage

- Energy-efficient indexing structures and update algorithms, prediction/prefetching data schemes
- Storage virtualization, data de-duplication, storage tiering, and moving archival data to storage devices that can be powered down
- Shift from hard disk drive storage devices to smaller form factor disk drives and solid state memory and increasing use of serial advanced technology attachment (SATA) drives

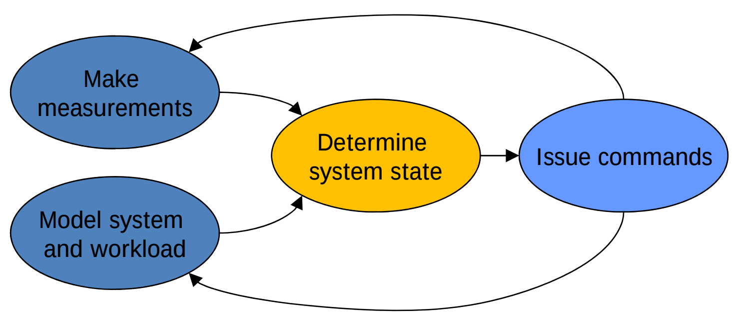
	Power dissipation	mW/GB
DRAM (1GB)	5W	5,000
15K RPM disk(300GB)	17.2W	57.33
Flash (SSD, 128GB)	2W	15.6

Networks

- Power-efficient network processors, switches, routers
- New end-to-end protocols for networks where nodes can go to sleep
- Novel design of the switching fabric using off-the-shelf components to improve the *bisection* bandwidth
- Energy-efficient routing protocols
- Computation vs. communication tradeoff (send me “bits” or send me “power”)
- Virtualization in enterprise switch servers



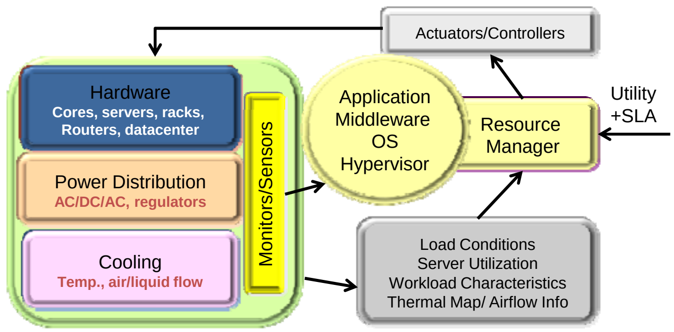
Generic Flowchart of a System-Level Power Manager



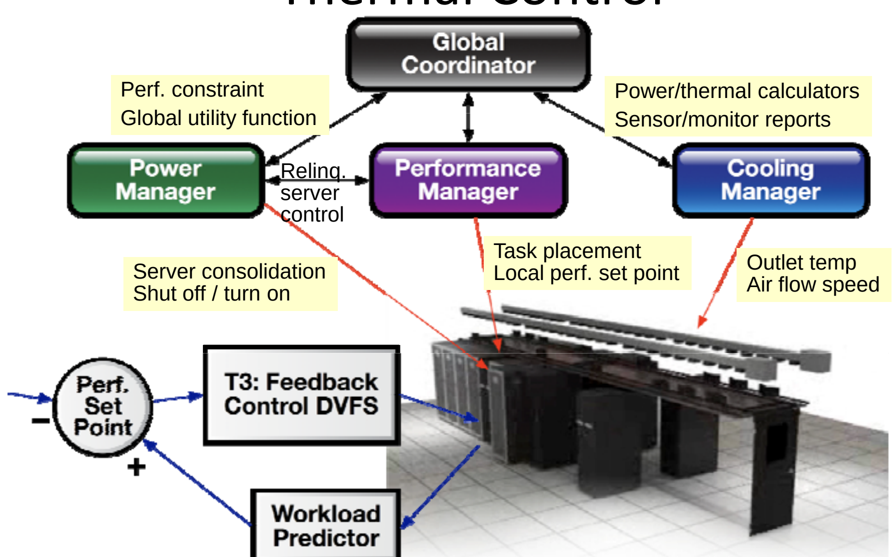
- Use hardware performance counters to gauge processor activity
  - Analyze program phases and adapt processor state accordingly
  - Recognize power/thermal hotspots and take (possibly preventive) action

Power Management Stack in Datacenters

- Power management
  - Hierarchical, adaptive (learning-based), rigorous (control theory, game theory), and robust (models and manages uncertainty)
  - Receives runtime information from various instrumentation and holistically manages all datacenter resources
  - Utilizes various power/energy optimization levers provided in hardware, hypervisor, middleware, software

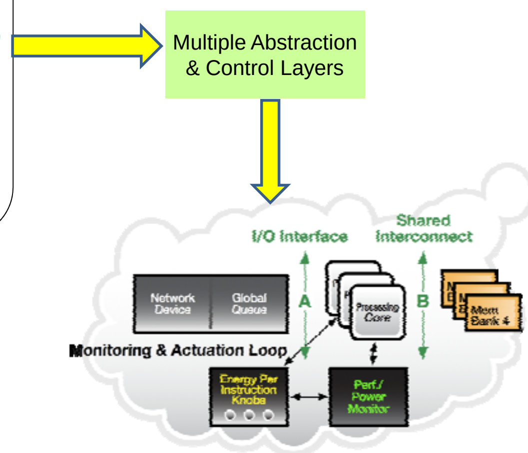


Multi-Agent, Multi-Tiered Power & Thermal Control



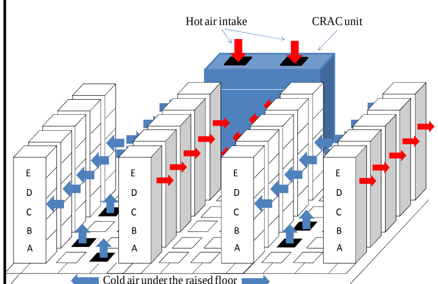
Energy-Aware Application Design

Start
Initialization
Read Data
PragmaStart(Mem-bound & Perf-Critical)
Until ? do Kernel1
PragmaEnd
Data transformation
Until ? do Kernel2
....
Prepare Output
End



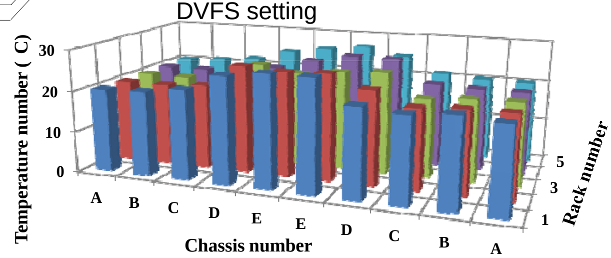
M. Pedram, USC 43

Dynamic Workload Placement based on Cooling Efficiency



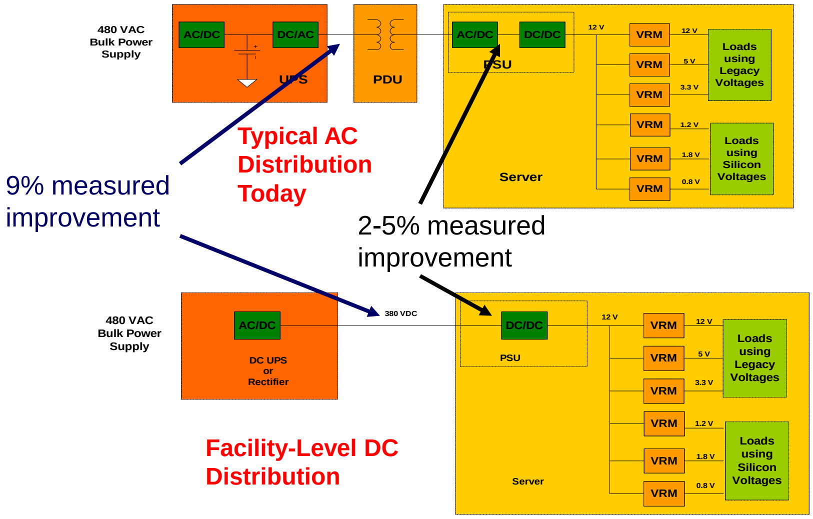
<u>Premise:</u>
- Hotspots exist that impact HVAC efficiency and energy consumption

<u>Approach:</u>
- Place workload in more power-efficient locations
- Perform server consolidation and DVFS setting



M. Pedram, USC 44

DC Power Distribution in a Data Center



It is About More than Watts …

- Environmental impact of IT is a growing worldwide concern
- Management of available energy is required in order to run cost effective IT operation
- A holistic, cross-layer approach to data center design and management is necessary to improve its energy efficiency
- The real metric is services per joule per dollar
- Many interrelated ideas:
  - Application efficiency and energy management software
  - Workload variations and service level agreements
  - Micro-architecture and system design
  - Storage and network bandwidth and cost
  - Power availability and cost, power distribution and conversion efficiency