Energy Adaptive Computing

Krishna Kant

George Mason University

National Science Foundation

SMART 2012, Stuttgart, Germany

How do you make data centers environmentally smart?

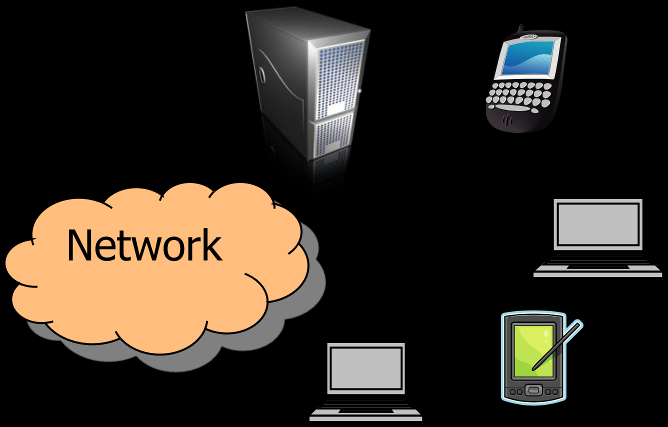
Smart energy mgmt is necessary but not sufficient

Computing Power is Growing

- 2020 projections
  - Clients: 8x in number, 3X in power
  - Data Centers: > 2X increase
  - Network: 3X increase

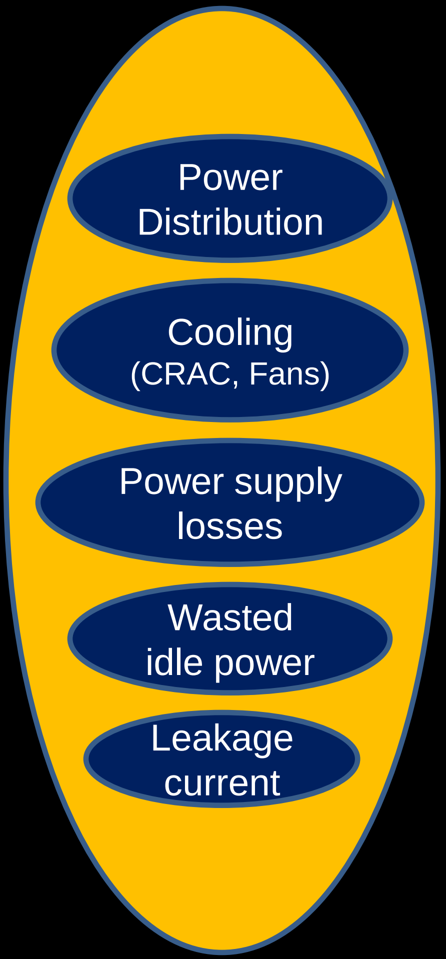
Transmission, conversion & distribution

Data Center



Smart Energy Mgmt is Essential

- Hardware Level
  - Aggressive power mgmt at each level
  - Coordination within & across levels
- Server Level
  - Fans, power supplies, OS, & app level power mgmt
- Data Center Level
  - Cooling & airflow management, placement, scheduling, …



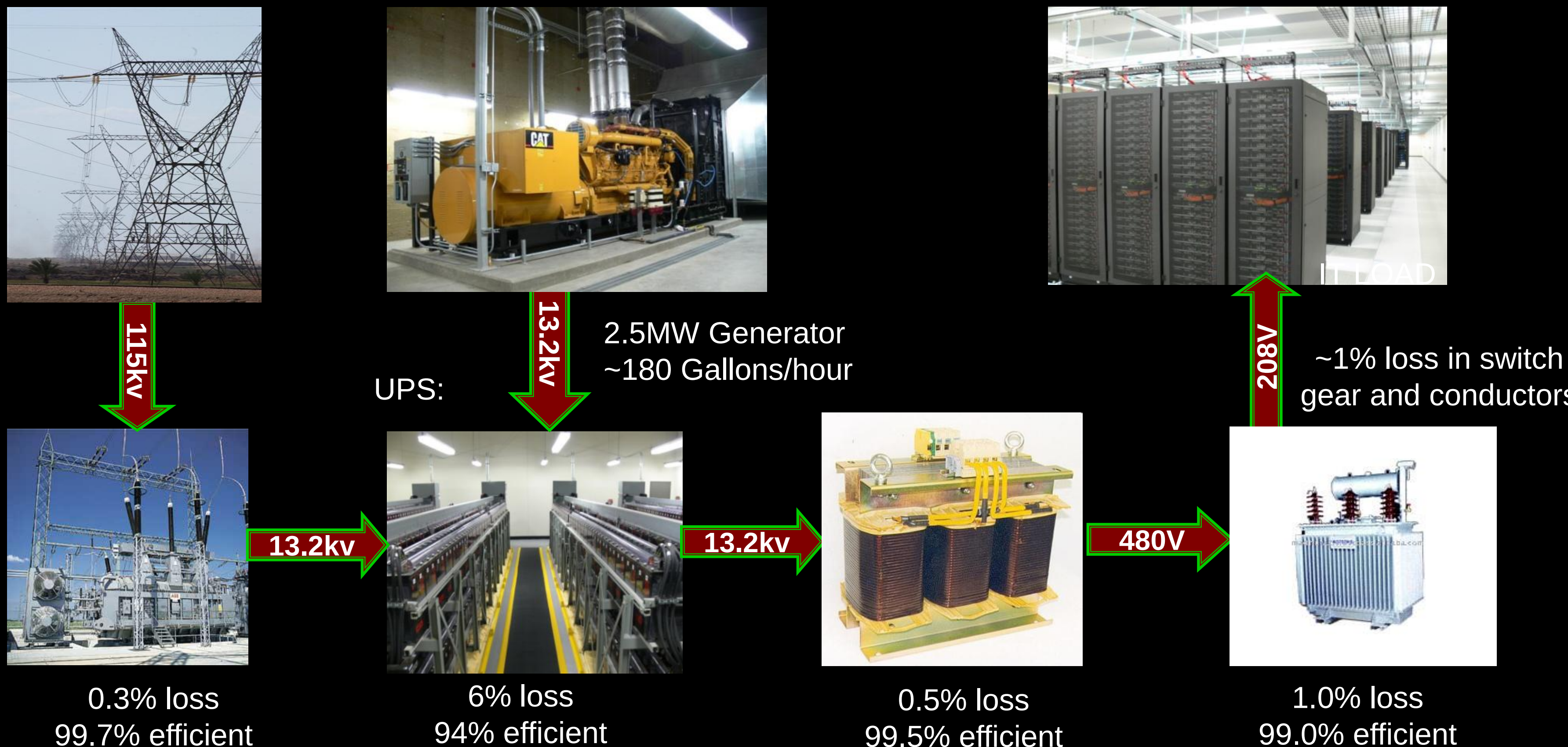
Is Energy Efficiency Enough?

- Energy efficiency less important, its carbon footprint really matters
- Energy efficiency may not reduce energy usage.
- Additional sustainability considerations
  - Use locally generated renewable energy
  - Reduce infrastructure & resource use (metals, water, …)

Cooling Infrastructure

- Cooling is very resource intensive
  - Lot of materials
  - Water, much of which evaporates

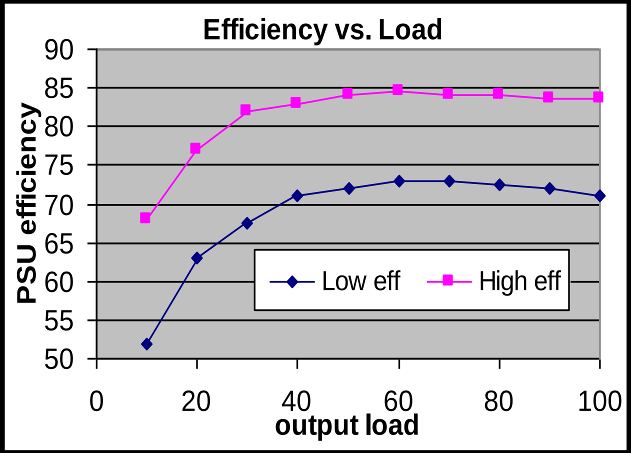
Power Distribution Infrastructure



- 9-10% distribution loss at power source
- Lots of earth’s resources used (metals, rare earths, …)

Overdesign

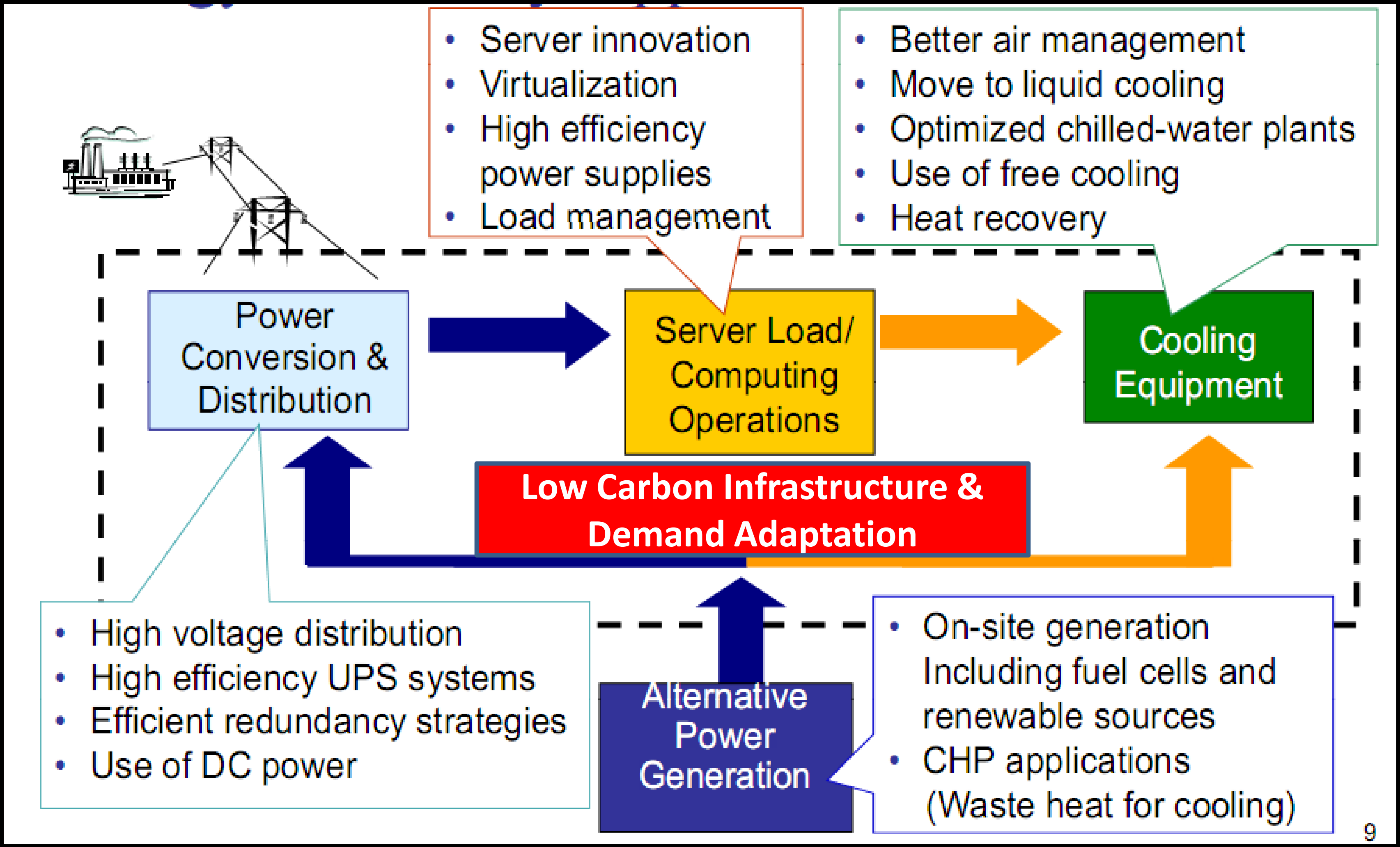
- Overdesign is the norm
  - Huge UPS, Generators, dist. frames, power supplies, fans, …
- Engineered for worst case
  - Huge waste of power, materials, …
- Example: Power Supply
  - Low utilizations, especially for duplex config ➔ Low efficiency
  - Voltage regulators: Similar issues



Sustainability Considerations

- Use of renewable energy
  - Must deal with variability & inadequacy of available energy
- Thrifty use of energy & materials
  - Free Cooling instead of CRAC
  - Reduce size of UPS, generators, power supplies, heat sinks, fans, …
- Smart adaptation to deal with under-capacity

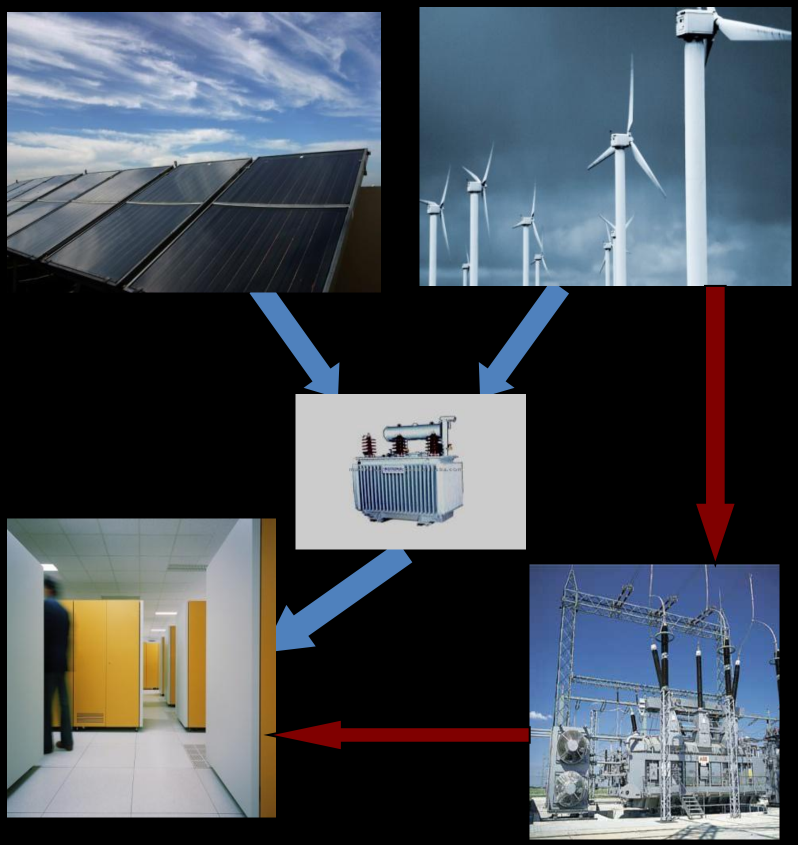
Data Center Energy Opportunities



Source: US DOE: Data Center Energy Efficiency Program

Renewable Energy Powered IT?

- Limit grid energy draw
  - Less infrastructure & losses, but variable supply
  - Impact on performance, QoS, SLA, …
- Variability Issues
- Reliability issues (small installations)



Need better power adaptability

High Temperature Operation

- Chiller-less data centers
  - Less energy/materials, but space inefficient

- High temperature operation of comm./computing equipment
  - Smaller $T_{outlet} – T_{inlet}$
  - Deal with occasionally hitting temp. limits.

Need smarter thermal adaptability

Energy Adaptive Computing

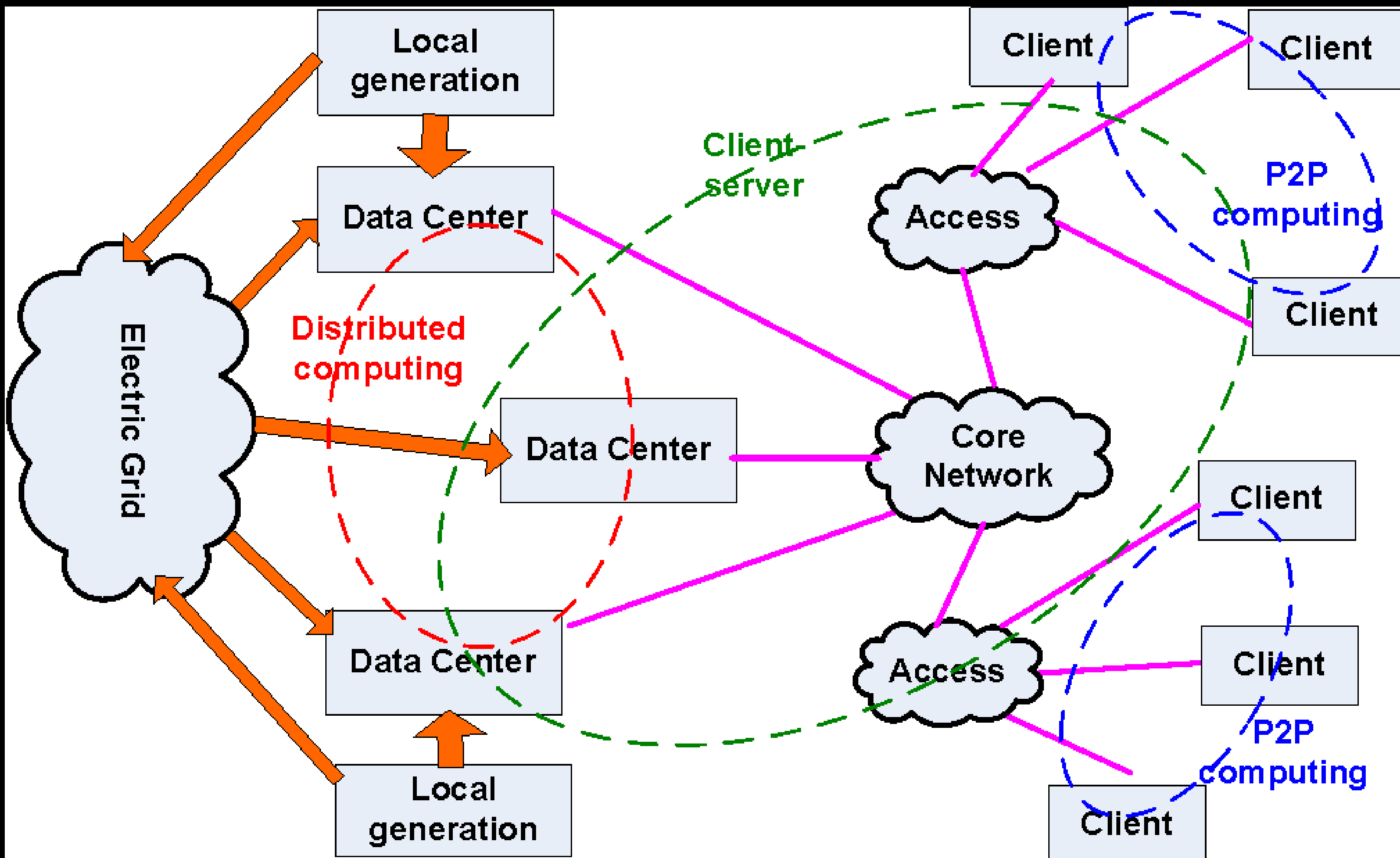
- Dynamic end to end adjustment to
  - Workload adaptation: What & how to run?
  - Infrastructure adaptation: Where & when to run?

- What's new?
  - Mandatory, rather than opportunistic power and thermal mgmt.
  - Coordination across compute, network & storage.
  - Integration of workload/infra adaptation

Adaptation Methods

- Workload Adaptation
  - Shut down low priority tasks
  - Lower resolution, precision, partial service, …
  - Pre-compute or pre-communicate
- Infrastructure Adaptation
  - Load consolidation & migration
  - QoS degradation
    - Higher delay (Batched service, mandatory sleep)
    - Lower tput (lower freq/voltage, “width” control, …)

EAC Instances

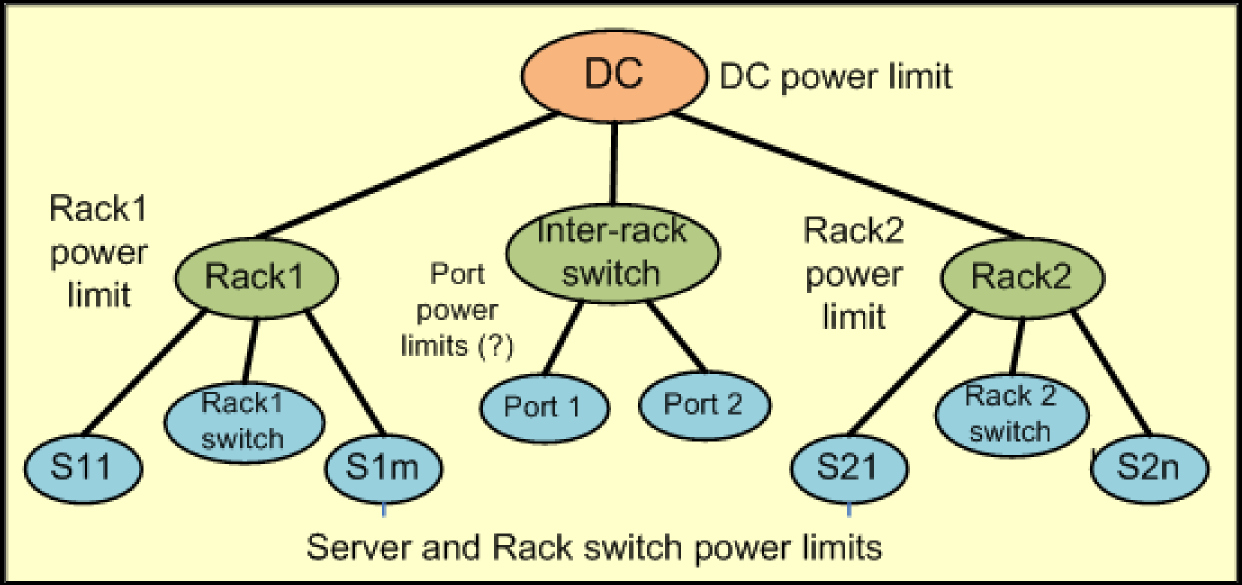


Adaptation Challenges

- Client-server adaptation
  - Transparently adapt to client energy states
  - Coordinated adaptation of client, network & servers
- Server side adaptation
  - Multi-level coordination: Server, rack & DC levels
- Adaptation among peers
  - Group adaptation to maximize overall utility

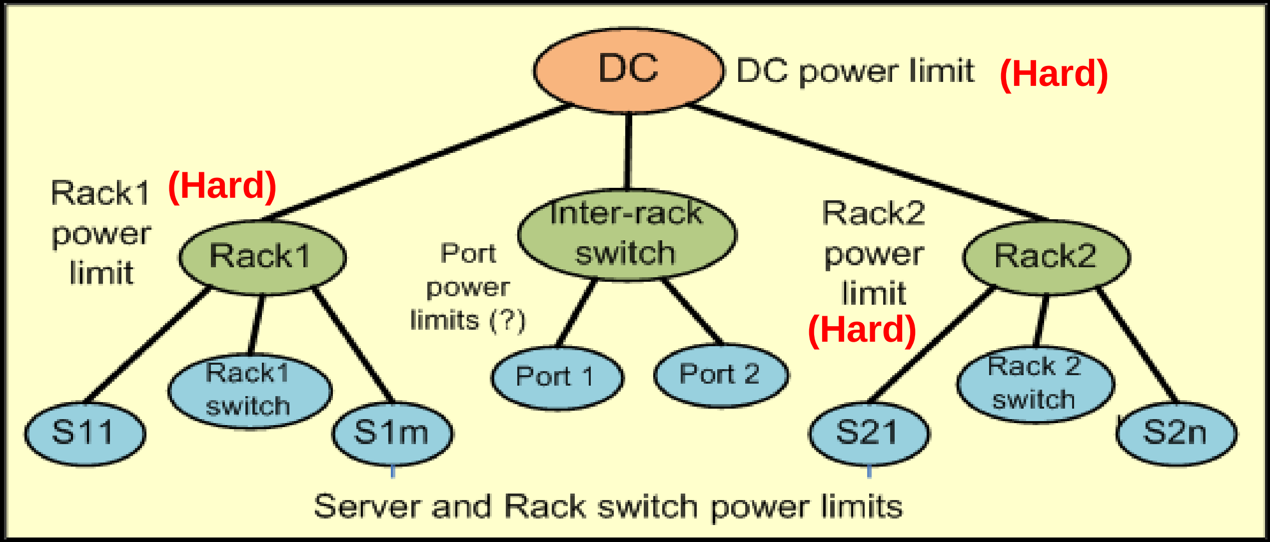
Data Center Adaptation

- Need a multilevel scheme –
  – Individual "assets" up to entire data center
- Need both supply & demand side adaptations



Hard vs. Soft Power Limits

- Hard limits
  - Energy availability, circuit limits, thermal limits, …
- Soft limits
  - Rationing at each level



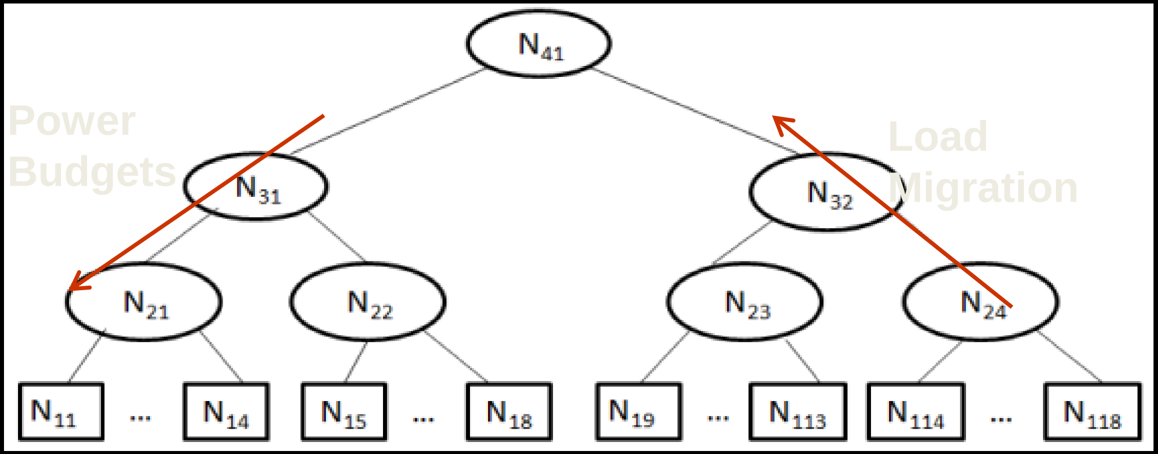
Adaptation

- Supply side: set soft limits as needed
- Demand side
  - Dynamic migration
  - Load consolidation

- Combined supply & demand side adaptation
  - Hierarchically organized scheme that
    - Minimizes imbalance and ping-pong
    - Minimizes error accumulation down the hierarchy.

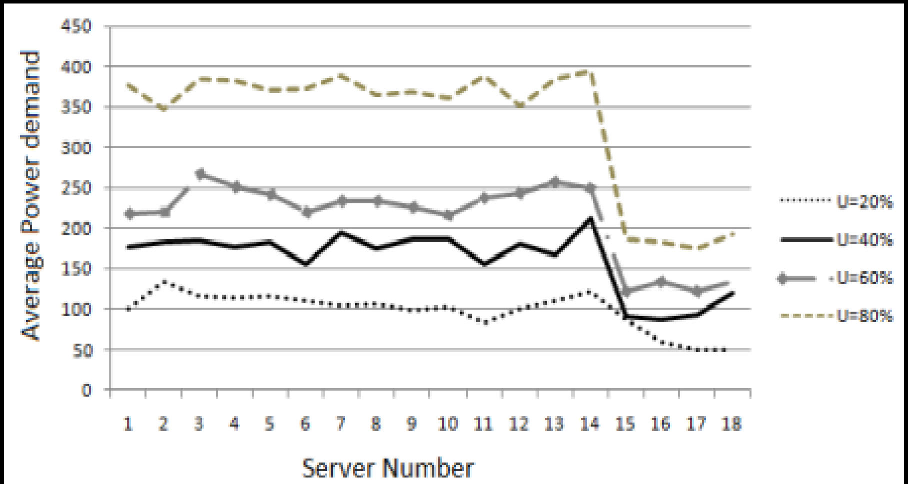
A Proposed Algorithm

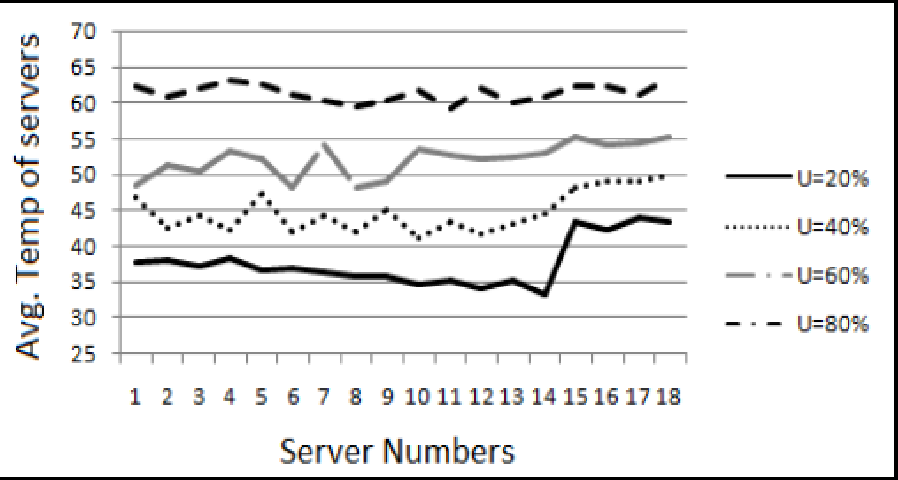
- Systematic control
  - Power budgets changes move downwards
  - Load migration moves up the hierarchy, from local to global.
  - Details available (IPDPS 2011 paper)



Sample Results
Adaptation to Thermal Profile

- Scenario
  - 3 levels, 18 servers
  - 3 apps (25 app instances)
- Adaptation to handle hot-spots
  - Servers 1-14: T_a=25^o C
  - Servers 15-18: T_a=40^{o}C
  - Temperature limit: 65^{o}C



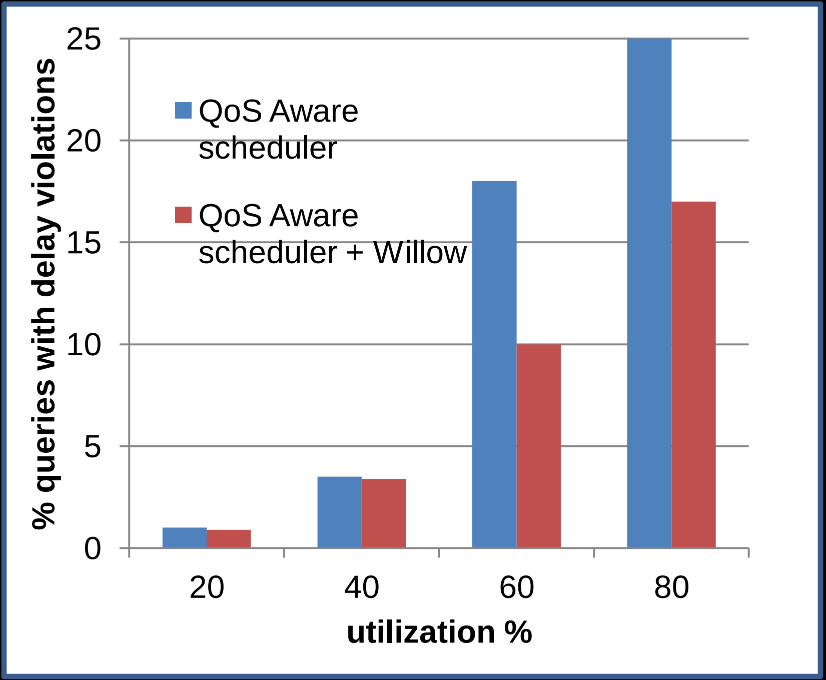


Recent Results (with QoS)

Application Type	SLA Requirement	Mean Runtime
Type I	Average Delay $\leq 120ms$, cannot be migrated	10 ms
Type II	Average Delay $\leq 180ms$, can be migrated	15 ms
Type III	Average Delay $\leq 200ms$, can be migrated	20 ms

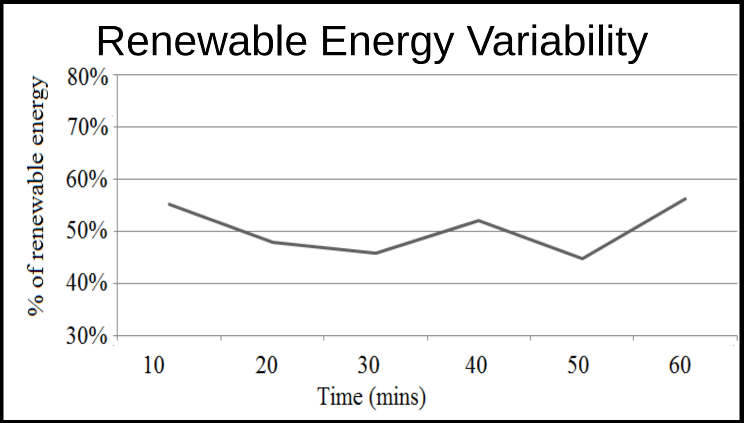
- 3 types of queries w/ different QoS needs
- Willow: Our adaptation mechanism
- Performs better than just QoS aware scheduling

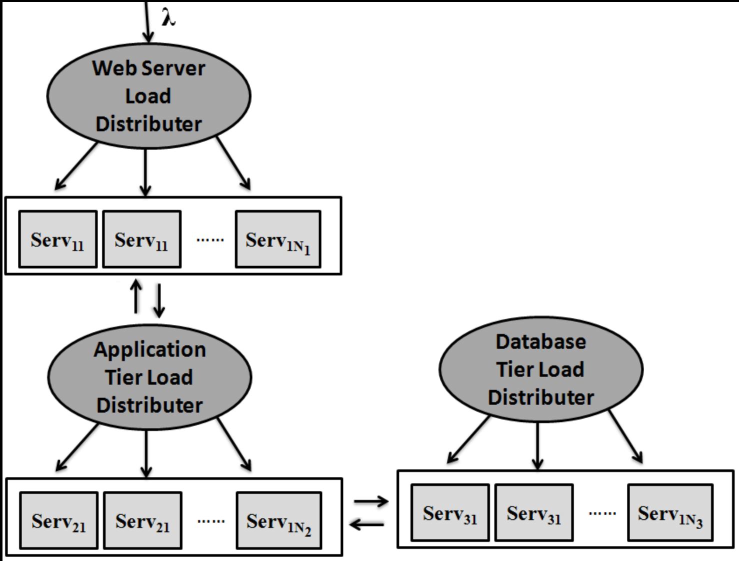
Results in ACM JETC



Adaptation in Multi-Tier Systems

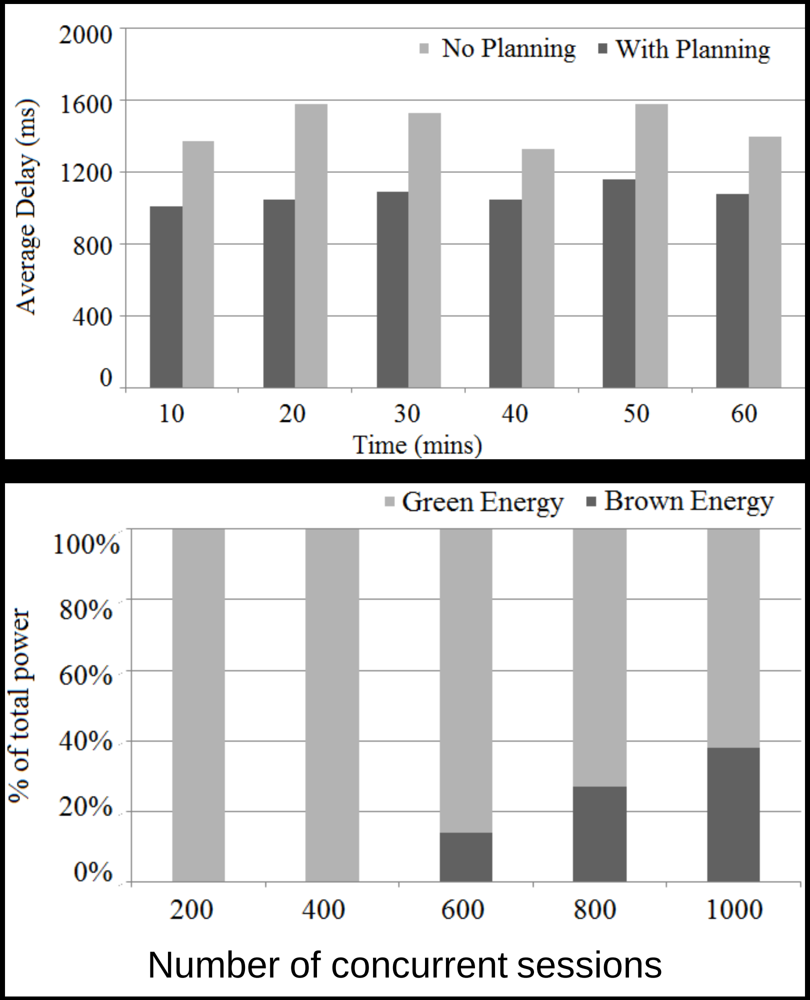
- Typical 3-tier system
  - Heterogeneous servers
  - Some fraction of power is renewable
- Reallocate power budget to
  - Balance delays across tiers
  - Consolidation in each tier
  - Minimize pwr state changes for servers & switches
- Results in ITJ paper





Sample Results

- Careful planning of power state changes
  - Minimizes state changes & control delays
- Maximization of green energy use
  - Requires specially designed power infra.
- At low utilization only green energy is used.



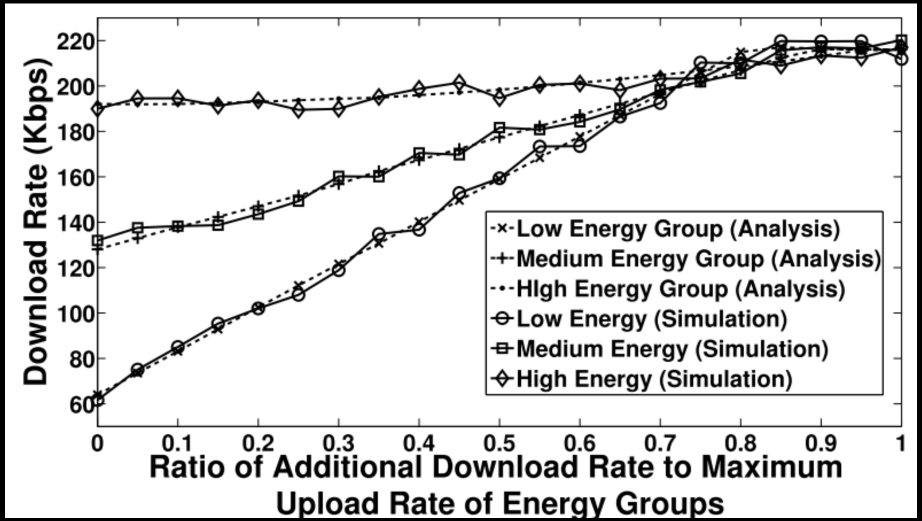
Energy Adaptation in P2P Systems

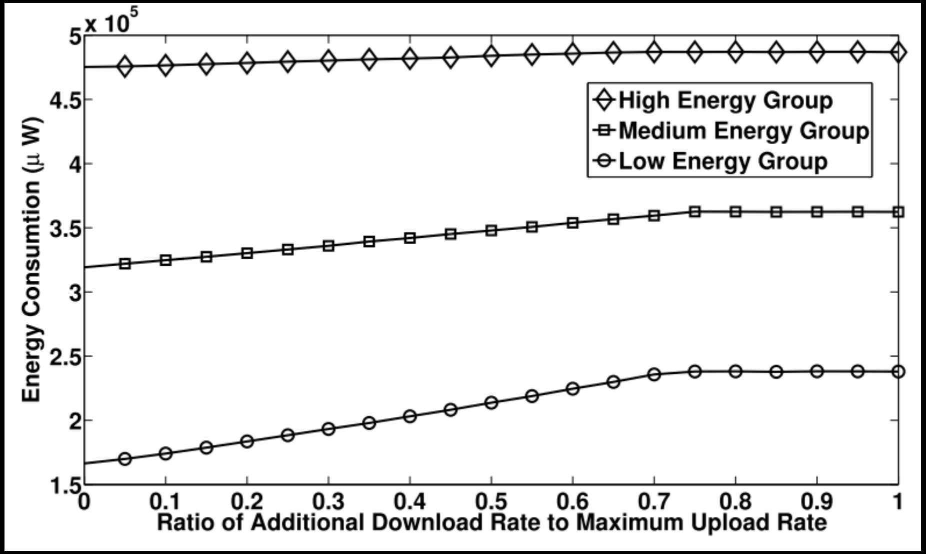
- Multiple energy groups
  - Joined based on remaining battery of mobiles
- Break the normal tit-for-tat
  - Download rate α upload rate only within a group

- Exploit transmit energy >> receive energy
  - Low battery: Low upload rate, but high download
  - Extra downloads from higher energy groups

P2P Adaptation Results

- P2P Adaptation
  - High download rate at low energy!
  - Need a credit mechanism used to avoid abuse





Mandatory Sleep

- Blink architecture [ASPLOS'11]
  - Define a duty cycle for each server
  - Adjust sleep durations based on current power availability.
  - Proactive workload mgmt to deal with sleep
    - Migrate tasks away before the sleep begins.
    - Migrate tasks in just in time for wakeup
- Characteristics
  - Another form of energy adaptive computing
  - Mandatory sleep for all servers, instead of keeping some servers down ➔ More overhead

Future Challenges

Power Estimation Challenges

- Notion of effective power?
  - Additive relationship: Workload → power
  - Why is this hard? Interference

- Available power
  - Determined by power, thermal & perhaps other issues (noise).
  - Required at multiple levels: facility, enclosure, machine, …

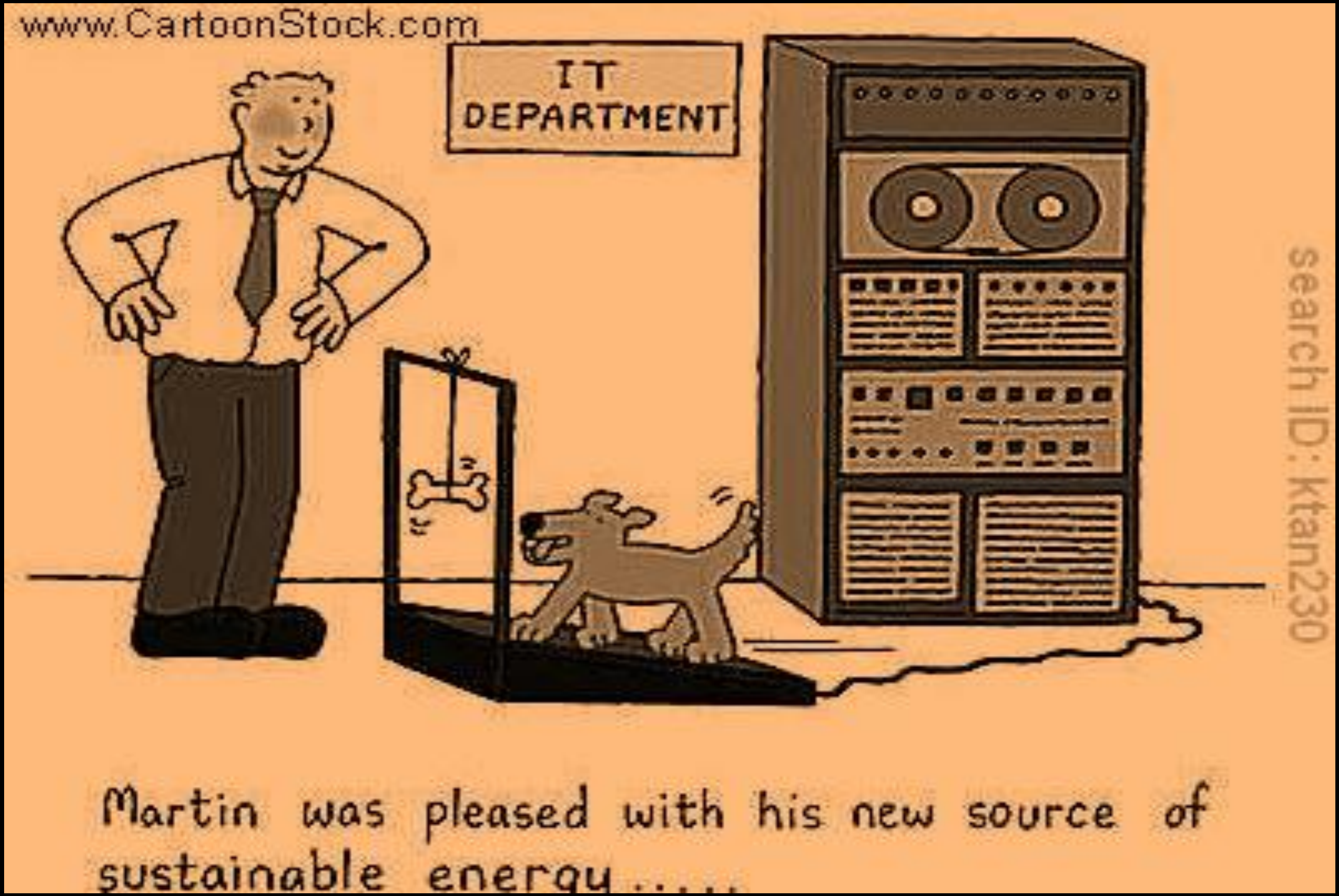
Network Role in EAC

- Energy Adaptation
  - – Aggressive control of switch/router ports
    - Speed, state & width controls
  - – Traffic consolidation across paths
- Adaptation induced congestion
  - – Propagation (e.g., ECN, EBCN) & response
    - Computation – communication tradeoff ?
    - Redirection ?
- Network protocol support for adaptation?

Other Issues

- Storage adaptation
  - Storage devices, controllers & network.
- Preprocessing
  - More work during energy plenty times in anticipation of deficit
- EAC Security
  - Attacks on power sources
  - Energy Attacks on IT, e.g.,
    - Demanding too much, cyclic demands, …
- Coordinated end to end control is hard!
  - Formal models to understand impact of energy adaptation.

Thank You!


www.CartoonStock.com
IT
DEPARTMENT
search ID: Ktan230
Martin was pleased with his new source of sustainable energy.....