UPTIME INSTITUTE SYMPOSIUM

# Revolutionizing Data Center Efficiency

McKinsey & Company

## ACKNOWLEDGEMENT

McKinsey & Company would like to thank and recognize the important collaborative contributions of Kenneth Brill and The Uptime Institute to the development of this report and its recommendations. The Institute provided critical insight based on their many years of experience as well as proprietary data and analysis not previously made public

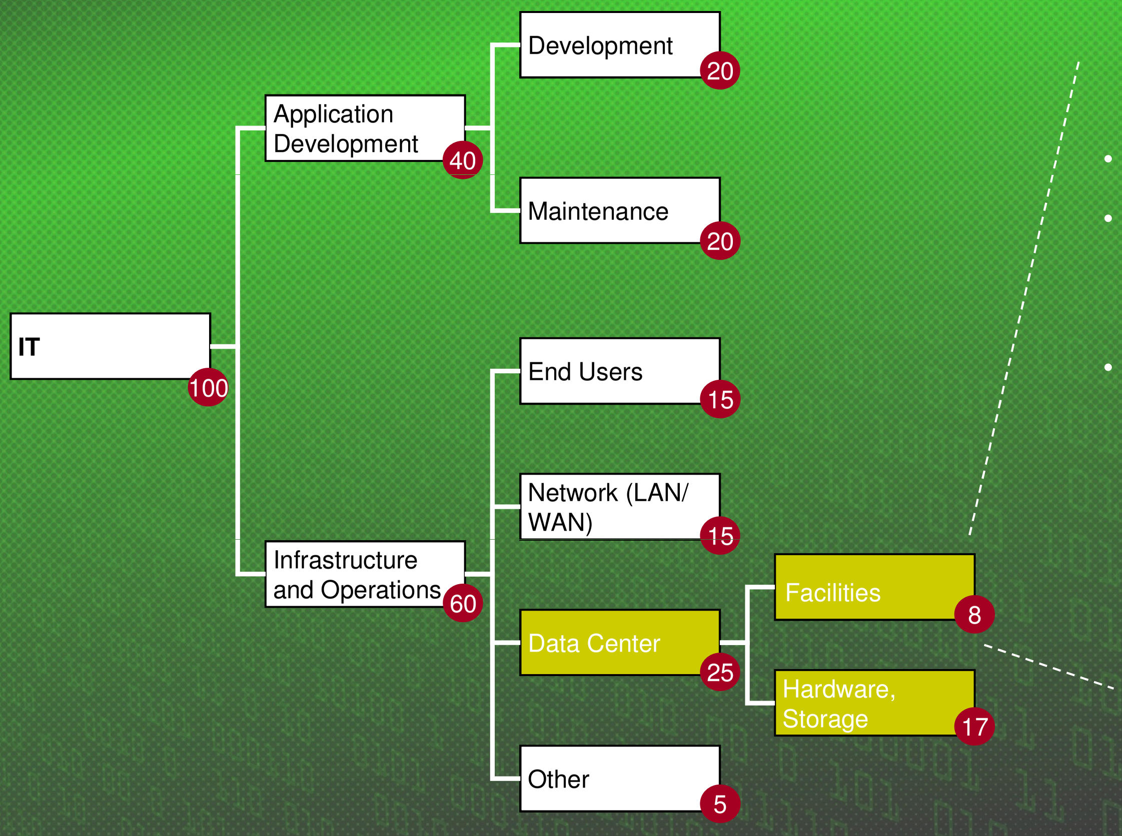
# EXECUTIVE SUMMARY

- **The rapid recent (and projected) growth in the number and size of Data centers creates two significant challenges for enterprises:**
  - **Data center facilities spend** (CapEx and OpEx) is a **large, quickly growing and very inefficient** portion of the total IT budget in many technology intensive industries such as financial services and telecommunications. **Some intensive data center users will face meaningfully reduced profitability if current trends continue**
  - **For many industries, data centers are one of the largest sources of Greenhouse Gas (GHG) emissions.** As a group, their overall emissions are significant, in-scale with industries such as airlines. Even with immediate efficiency improvements (and adoption of new technologies) **enterprises and their equipment providers will face increased scrutiny given the projected quadrupling of their data-center GHG emissions by 2020**
- The **primary drivers of poor efficiency** are:
  - Poor demand and capacity planning within and across functions (business, IT, facilities)
  - Significant failings in asset management (6% average server utilization, 56% facility utilization)
  - Boards, CEOs, and CFOs are not holding CIOs accountable for critical data center facilities CapEx and data center operational efficiency
- **Improving efficiency** is the best near term means to solving the twin challenges of rising spend and GHG emissions. We propose a three part solution **to double IT energy efficiency by 2012** and to arrest the growth of GHG emissions from data centers:
  - Mandate inclusion of **true total cost of ownership** (including data center facilities) in business case justification of new products and applications to throttle excess demand
  - Rapidly mature and integrate **asset management** capabilities to reach the same par as the Security function
  - Formally move accountability for data center critical facilities expense and operations to the CIO and appoint internal "Energy Czars" with an operations and technology mandate to **double IT energy efficiency by 2012**
- To achieve this doubling of energy efficiency CIOs, equipment manufacturers, as well as industry groups in dialog with regulators should quickly establish automotive style "CAFE" metrics that will **measure the individual and combined energy efficiency of corporate, public sector and 3rd party hosted data centers.** We propose one metric here for discussion and adoption. This metric would deliver immediate financial and transparency benefits to executive management of enterprises large and small and could become a government recognized measure of efficiency



# DATA CENTER COST IS APPROXIMATELY A QUARTER OF TODAY'S IT COSTS . . .

**Breakdown of average IT cash costs at a typical company, percent**



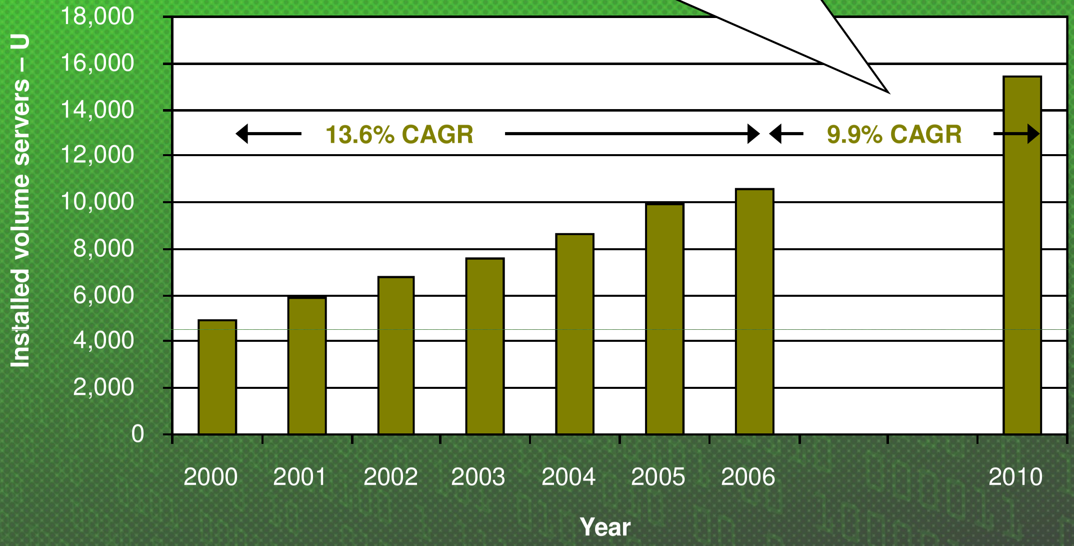

- Not all facilities within IT budget
- Unrealistically long depreciation timeframes artificially hide data center costs
- Current construction boom is not a one time catch-up investment. Server growth will require add'l new data center construction every 3-5 years

Note: Total IT budget is illustrative of a typical company
Source: McKinsey analysis

# AND DATA CENTER IT COSTS WILL CONTINUE TO GROW AS THE NUMBER OF SERVERS HOUSED WITHIN DATA CENTERS GROWS RAPIDLY . . .

**Servers hosted within data centers within USA**

- CAGR reducing due to increased use of virtualization
- Power consumption per server increasing even faster as newer machines consume much more power




- Data center spend is growing rapidly due to increased demand
- China and other developing countries are projected to grow even more rapidly
- Growing data center spend is putting pressure on other IT initiatives or functions (e.g., applications development, end user computing)

Note: Total IT budget is illustrative of a typical company
Source: EPA 2007 Report to congress



# . . . AND PLAYERS ACROSS INDUSTRIES CONTINUE TO MAKE MAJOR INVESTMENTS IN NEW FACILITIES

**citibank $735M**
- Invest $470 million in a new data center in Ohio
- Plans to invest EUR 170M to develop data center in Frankfurt to house over 4,000 servers; data center to support Citibank's operations in Europe, Middle east and Africa

**hp $1B**
- Consolidating 85 worldwide data centers into 6 major sites in US
- Each site with have 50,000 sqr feet raised floor space

**facebook $200M**
- Facebook plans to invest $200M to lease 86,000 sqr feet raised floor in Santa Clara, CA

**Microsoft $500M**
- Plans to spend $500M to build new 550,000 sqr feet data center in Chicago
- Plans to invest additional $500M to build server farm in Ireland

**at&t**
- Plans to add 186,000 sqr feet raised floor through its 28 global data center
- Opened a new data center in Chicago and one in Shanghai as joint venture with Shanghai Telecom to capture the enterprise managed service market

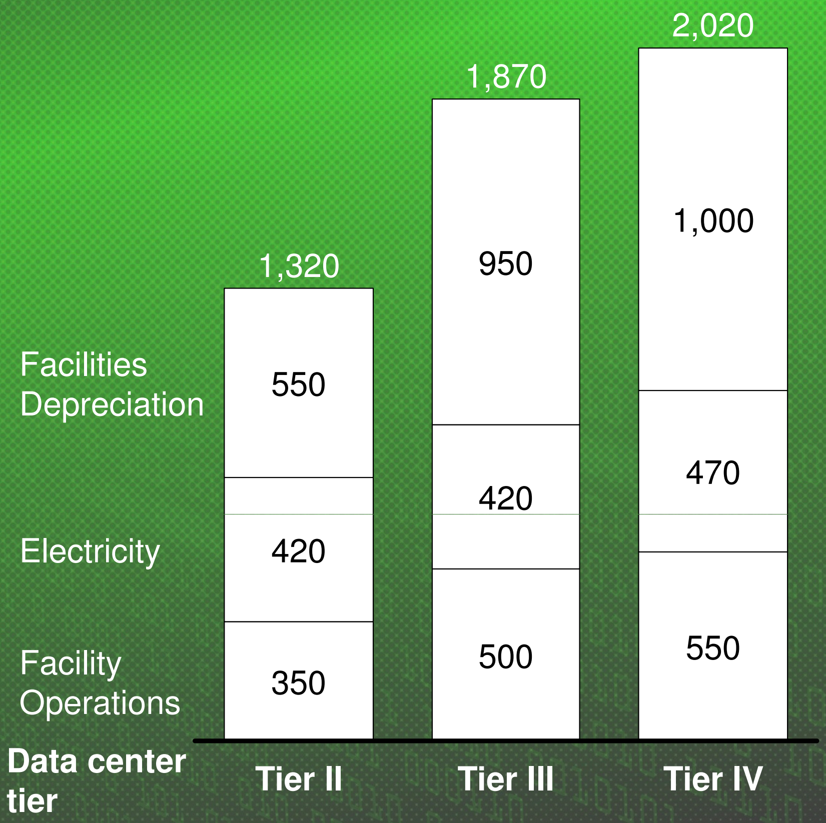
**WACHOVIA $400M**
- Investing $400M in new data center in Birmingham, AL as a part of 36 month program to reorganize data centers

**Google $600M**
- Plans to invest $600M to create a new data center in Berkeley county in SC and another $600M in Lenoir, NC
- Google also planning a data center near Columbia, SC

Source: Company reports; analyst reports; press release

## SERVERS AREN'T "CHEAP" BECAUSE THEY INCUR SUBSTANTIAL FACILITY (POWER AND COOLING) COSTS OVER THEIR LIFE

Annual OpEx to support a mid-tier ($2,500) server, dollars



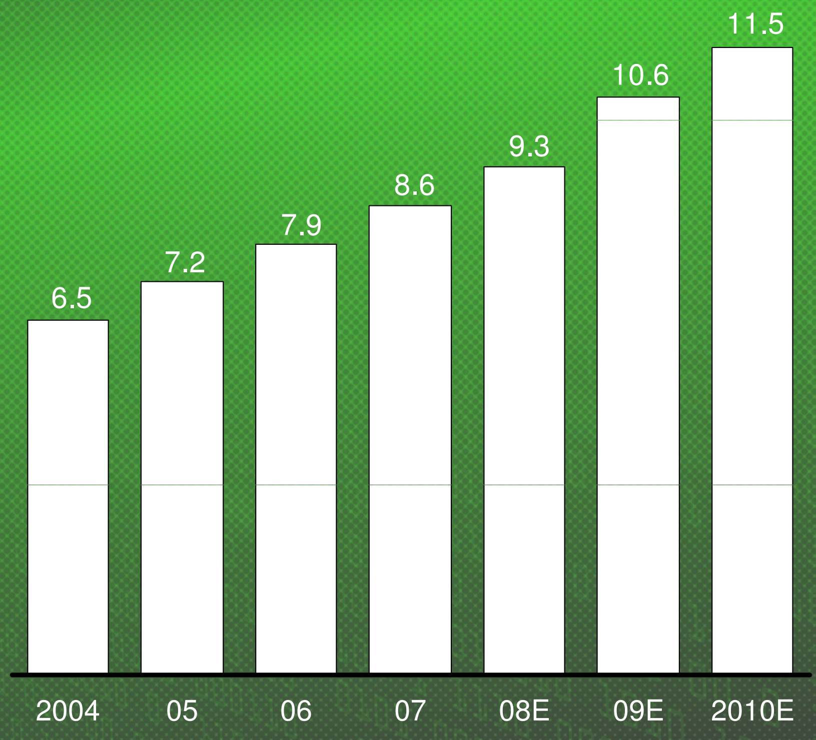

- True costs are often 4-5x the cost of the server alone over a 5-10 year lifetime of a server
- IT hardware energy consumption drives Facility costs
- Servers are often housed in a higher Tier Data Center than necessary, further driving Facility costs
- Facility costs are growing more rapidly (20%) than overall IT spend (6%)

Source: Uptime Institute



## HIGHER LOAD DENSITY ALSO CONTRIBUTES TO HIGHER ENERGY COSTS CURRENTLY INCREASING AT 16% PER YEAR

Total data centers energy bill, $ Billions




**3 Drivers of 16% CAGR Energy Cost Increase**

- Installed base on server is **growing by 16% and projected to** grow to 41-43 million servers worldwide by 2010
- Energy consumption per server is **growing by 9%** as growth in performance pushes demand for energy
- Energy unit price has **increased an average of 4%**

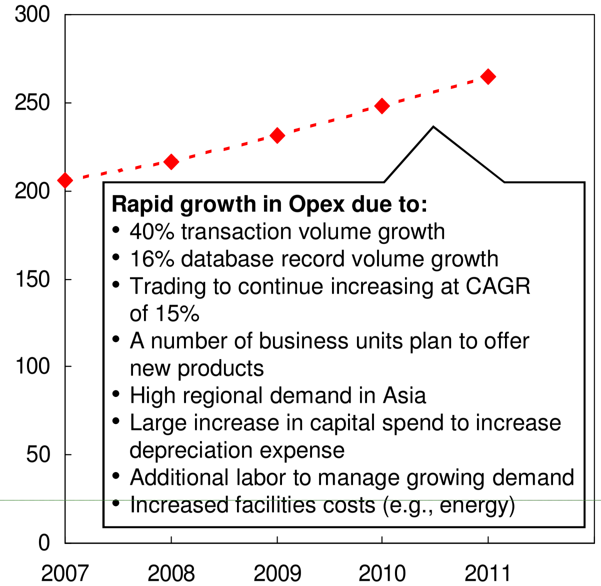
Note: Weighted average consumption for top selling volume servers

Source: IDC, "Estimating total power consumption by servers in the US and the world" from Jonathan G. Koomey, Ph.D.

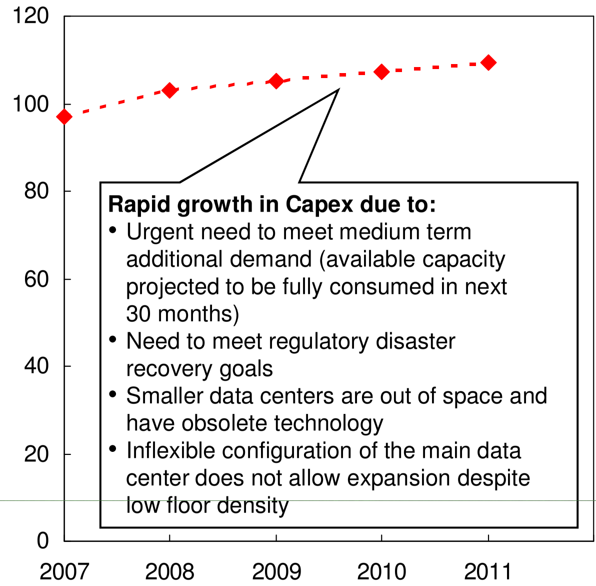
# WITHOUT RADICAL CHANGES IN OPERATIONS, MANY COMPANIES WITH LARGE DATA CENTERS FACE REDUCED PROFITABILITY

DISGUISED CLIENT EXAMPLE

**Opex projection**




**Capex projection**




- Data center cost as percent of total revenue all time high
- Data center cost growing twice as rapidly as revenue
- Data center construction investment significantly affects profitability for next two years

Source: McKinsey analysis

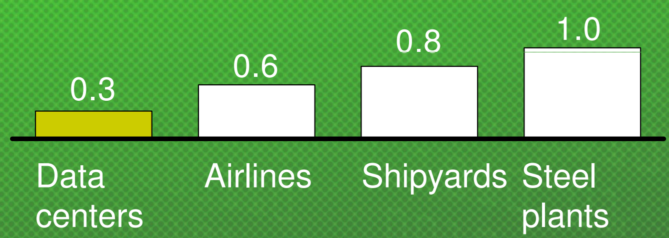


# DUE TO ENORMOUS ENERGY CONSUMPTION, DATA CENTERS' CARBON FOOTPRINT IS ALSO SURPRISINGLY HIGH AND GROWING

**Key points on data centers' greenhouse gas emissions**

- Data center **electricity consumption** is **almost .5% of world production***
- Average data center consumes energy equivalent to 25,000 households
- Worldwide energy consumption of DC doubled between 2000 and 2006
- Incremental US demand for data center energy between now and 2010 is equivalent of 10 new power plants
- 90% of companies running large data centers need to build more power and cooling in the next 30 months

**Carbon dioxide emissions as percentage of world total – industries**
Percent




**Carbon emissions – countries**
Mt $CO_2$ p.a.


* Including custom-designed servers (e.g., Google, Yahoo)

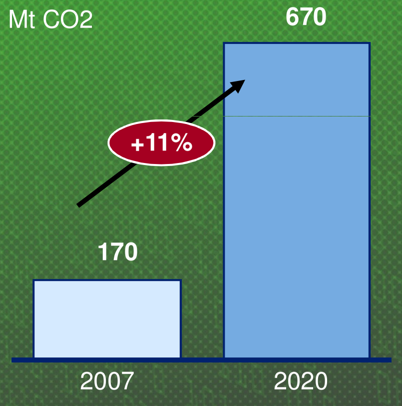
Source: Financial Times; Gartner report 2007; Stanford University; AMD; Uptime Institute; McKinsey analysis

## ONGOING INITIATIVES NOT WITHSTANDING, EMISSIONS WILL QUADRUPLE BY 2020 CAUSING INTENSE SCRUTINY FROM REGULATORS, ACTIVISTS AND CORPORATE BOARDS

**Emissions are set to quadruple by 2020**

- Due to higher performance per m2, the electricity consumption will grow faster than the number of servers
- Emission from data centers will surpass those from many industry such as Airlines

**Emissions from Data Centers worldwide**
Mt $CO_2$




**Current technology focused initiatives will not be sufficient to reverse trend**

EPA driven initiative to reduce power consumption at homes, commercial buildings, and electronics

Global consortium to reduce energy consumptions of data centers

Third party hosting service provider based at Cheyenne, WY powered 100% by wind power

Renewable Fuels Association is a trade group of US ethanol industry that promotes policies, research, and regular to increase use of ethanol as fuel

- The carbon footprint has begun to attract scrutiny and legislation (e.g., US Public Law 109-431 requires EPA to submit a report on energy consumption of data centers to US congress)
- EPA has advocated use of separate energy meters for large data centers and development of procurement standards
- The European Union is developing a voluntary Code of Conduct for data centers proscribing energy efficiency best practices.
- Data center carbon footprint is expected to affect even the industries that are traditionally considered "clean" (e.g., telecom, media, technology)

Source: IDC U.S. and Worldwide Server Installed Base 2007-11 Forecast; McKinsey analysis

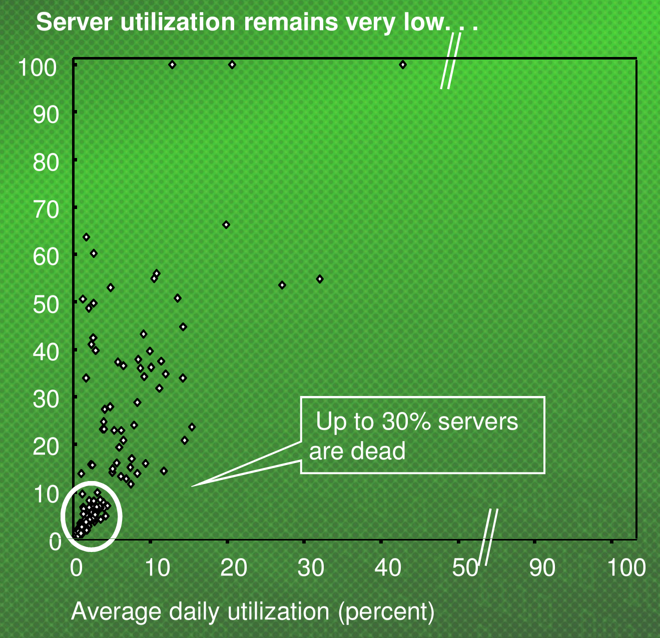


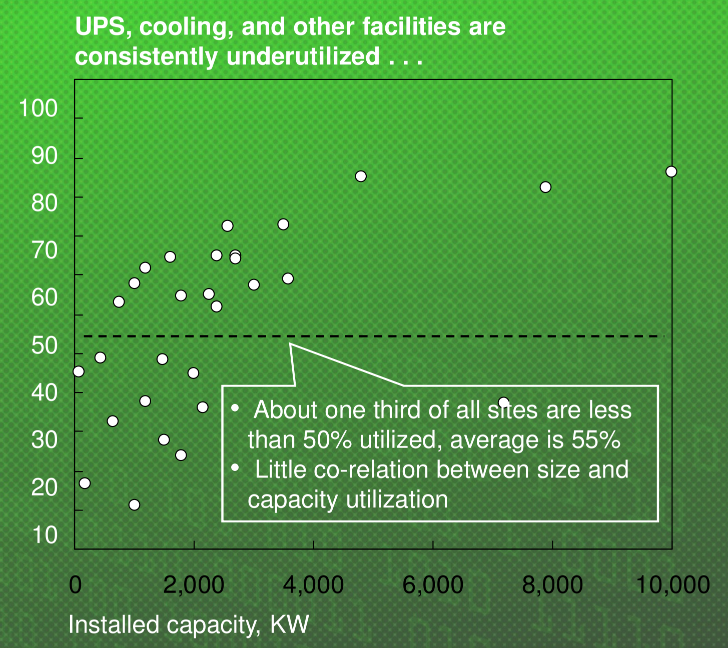
## EXECUTIVE SUMMARY

- **The rapid recent (and projected) growth in the number and size of Data centers creates two significant challenges for enterprises:**
  - **Data center facilities spend** (CapEx and OpEx) is a **large, quickly growing and very inefficient** portion of the total IT budget in many technology intensive industries such as financial services and telecommunications. **Some intensive data center users will face meaningfully reduced profitability if current trends continue**
  - **For many industries, data centers are one of the largest sources of Greenhouse Gas (GHG) emissions.** As a group, their overall emissions are significant, in-scale with industries such as airlines. Even with immediate efficiency improvements (and adoption of new technologies) **enterprises and their equipment providers will face increased scrutiny given the projected quadrupling of their data-center GHG emissions by 2020**
- The **primary drivers of poor efficiency** are:
  - Poor demand and capacity planning within and across functions (business, IT, facilities)
  - Significant failings in asset management (6% average server utilization, 56% facility utilization)
  - Boards, CEOs, and CFOs are not holding CIOs accountable for critical data center facilities CapEx and data center operational efficiency
- **Improving efficiency** is the best near term means to solving the twin challenges of rising spend and GHG emissions. We propose a three part solution **to double IT energy efficiency by 2012** and to arrest the growth of GHG emissions from data centers:
  - Mandate inclusion of **true total cost of ownership** (including data center facilities) in business case justification of new products and applications to throttle excess demand
  - Rapidly mature and integrate **asset management** capabilities to reach the same par as the Security function
  - Formally move accountability for data center critical facilities expense and operations to the CIO and appoint internal "Energy Czars" with an operations and technology mandate to **double IT energy efficiency by 2012**
- To achieve this doubling of energy efficiency CIOs, equipment manufacturers, as well as industry groups in dialog with regulators should quickly establish automotive style "CAFE" metrics that will **measure the individual and combined energy efficiency of corporate, public sector and 3rd party hosted data centers.** We propose one metric here for discussion and adoption. This metric would deliver immediate financial and transparency benefits to executive management of enterprises large and small and could become a government recognized measure of efficiency

# DESPITE RAPIDLY GROWING COSTS, DATA CENTERS ARE OPERATIONALLY VERY INEFFICIENT AND UNDERUTILIZED

DISGUISED CLIENT EXAMPLE




**A small number of organizations are starting to monitor server utilization, however very few organizations monitor facilities energy efficiency or utilization**

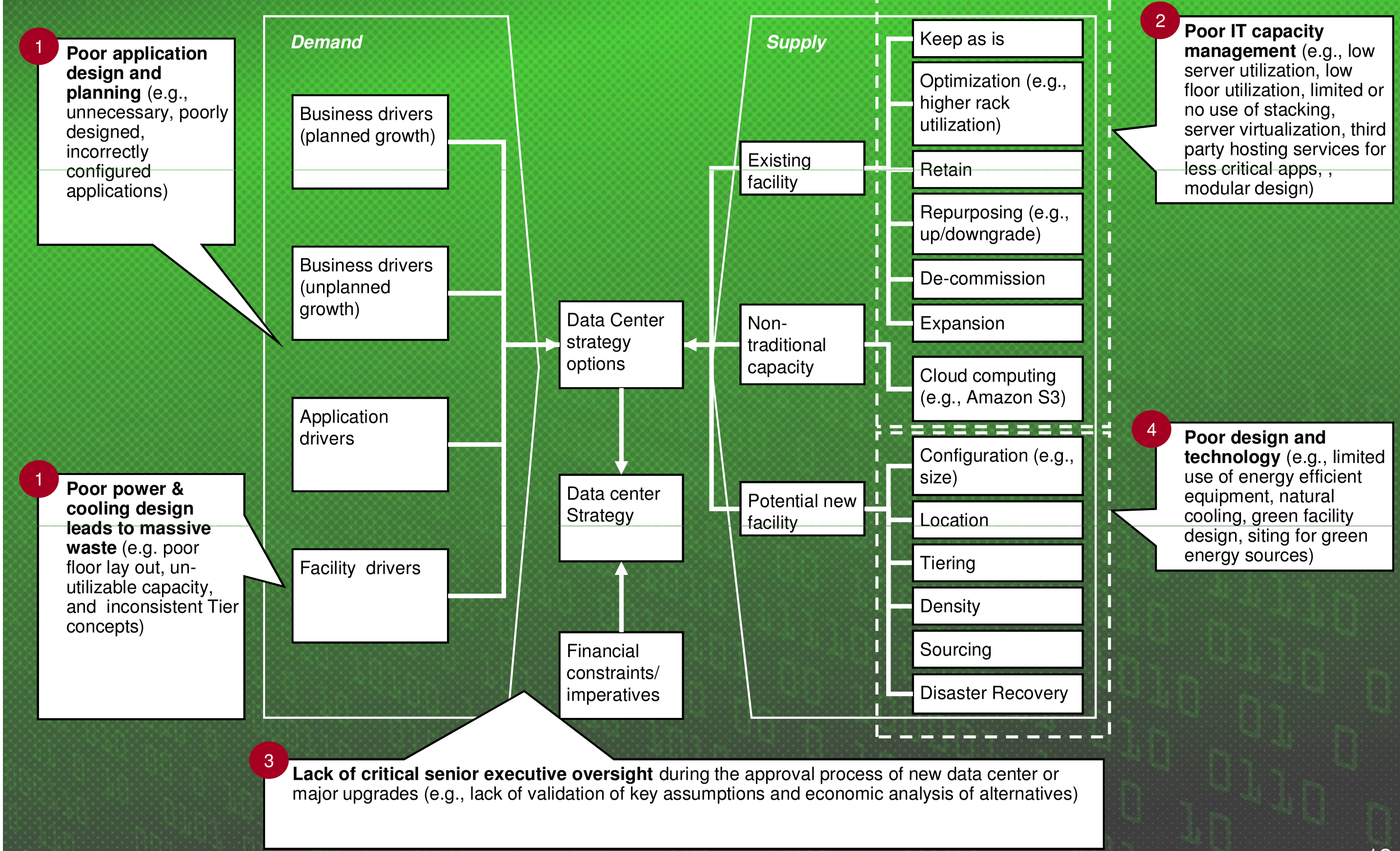
* Sample size – 45 data centers

Source: Uptime Institute



# THERE ARE FOUR PRINCIPAL CONTRIBUTORS TO DATA CENTER INEFFICIENCY ACROSS DEMAND SUPPLY FRAMEWORK

**Data center demand/supply framework**



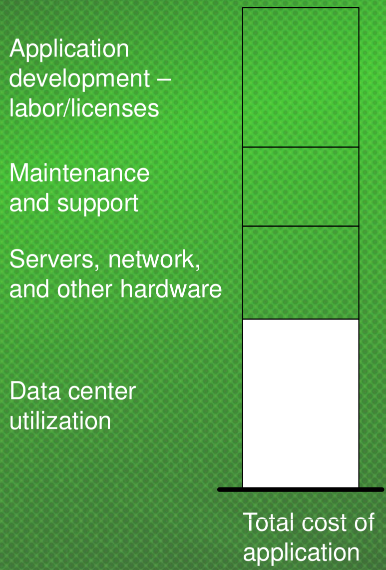

Source: McKinsey analysis

# 1. DECISIONS ABOUT APPLICATIONS AND INFRASTRUCTURE DO NOT ADEQUATELY CONSIDER THEIR IMPACT ON DC OPERATIONS AND COST

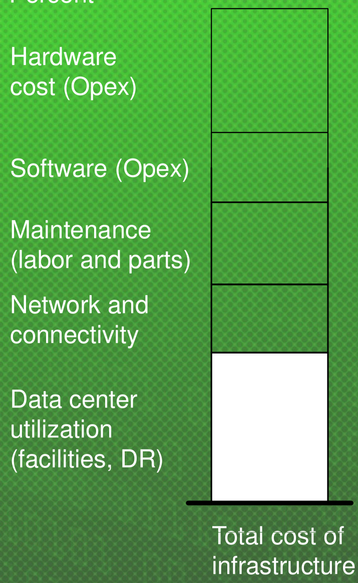
☐ Not considered in TCO business case for 'go/no-go' decision

ILLUSTRATIVE

**True Application TCO**
Percent




**True Infrastructure TCO**
Percent




- Limited understanding of data center TCO and limited access to relevant data
- Limited understanding of choices that can influence data center cost
- No representation of data center in design, planning, and approval process for new applications and hardware components

**Examples of poor application decisions…**
- Applications that don't reduce usage of monitors during off peak/closed hours
- Limited use of grid computing
- Computation load is not shifted among systems to maximize energy used

**Examples of poor infrastructure decisions…**
- Storage usage not maximized
- Limited use of MAID (massive array of idle disks)
- Poor layout design
- Equipment that is physically large

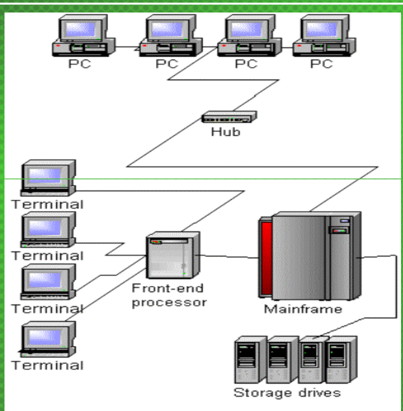
Source: Uptime Institute; EPA report; McKinsey analysis



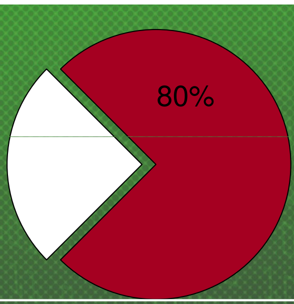
# 2. MANAGEMENT SOPHISTICATION HAS NOT KEPT UP WITH TRANSITION FROM MAINFRAMES TO DISTRIBUTED SYSTEMS

Utilized
Wasted

**From**

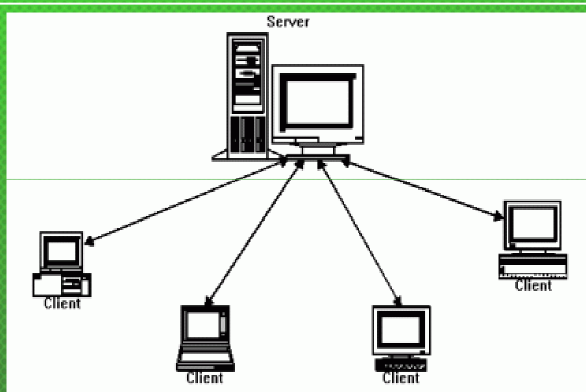



**Demand**

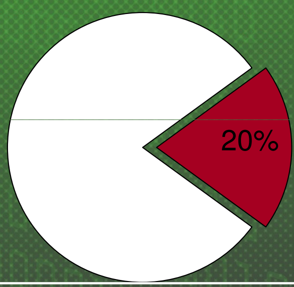



- In 1975-1985, mainframes with 70-80% utilization handled 80% of computing demand

**To**




**Demand**




- Today, 80% of computing demand is handled by distributed systems with 5-30% utilization

Source: IBM Energy Efficient Data Center Jun 2007; McKinsey analysis

# 3. LACK OF CIO/BOARD OVERSIGHT DURING TYPICAL CAPEX APPROVAL PROCESS FOR DATA CENTERS OFTEN RESULTS IN A SIGNIFICANT OVERSPEND

**Typical CapEx approval process for data centers**

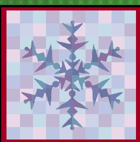
| Requirements | Design | Review and approval | Implementation |
|---|---|---|---|
| • No active decommissioning to free up existing facility capacity<br>• Assumes highest case demand projections<br>• Poor demand forecasting<br>• Alternate source of supply (e.g., third party hosting facility) not considered | • Gold plating to “future proof” data center capacity<br>• Limited use of future modular expansion capacity<br>• Lack of understanding or priority of IT and facility design choices that can significantly lower power requirements | • IT utilization data and demand projections are seldom challenged<br>• Unitary IT solutions as “fact accompli” assumptions and trade offs are difficult to validate<br>• CXOs and boards often are not suffic-iently knowledge-able to challenge assumptions or require alternative economic choices | • Items often missed in design phase (e.g., migration costs create project overruns)<br>• Specialized project management and cross functional oversight skills often are lacking resulting in delays and cost over runs |

Source: McKinsey analysis

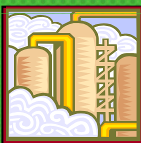


# 4. MOST DATA CENTER FACILITIES DO NOT FULLY USE ENERGY EFFICIENT DESIGN

SAMPLE CHALLENGES OBSERVED



**Temperatures in the cold aisle are much colder than required and can be increased to 74°. Similarly, the hot aisle should be hot (90° or even higher)**



**High density air cooling usually increases total facility CapEx for electrical and mechanical capacity as well as total energy consumption. Water cooling saves energy and is simpler and more reliable**

**All UPS modules, chillers, cooling units, etc. are installed initially instead of waiting until the center is more fully occupied**

**Efficiency focus is on 80% or higher loads instead of the 10-30% loads where most facilities operate for much of their lives**

**Winter free-cooling opportunities worth hundreds of thousands of dollars annually are not used because office building piping designs were used erroneously.**

Source: Uptime Institute

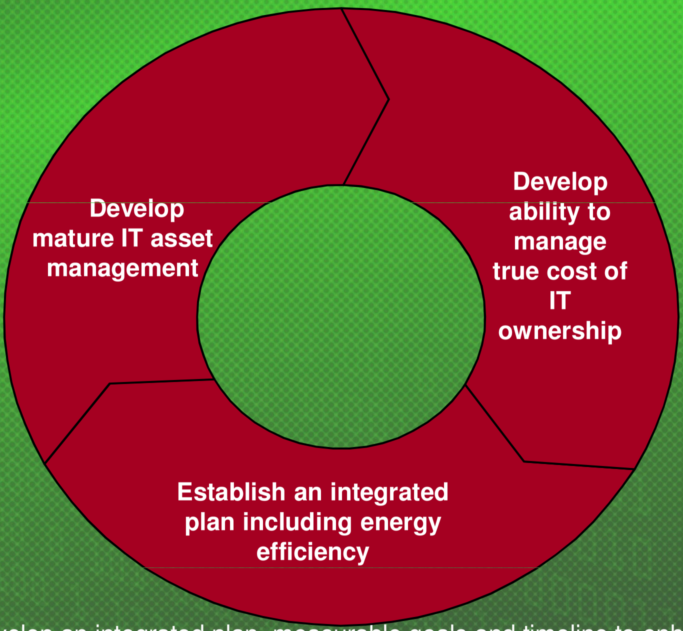
# EXECUTIVE SUMMARY

- **The rapid recent (and projected) growth in the number and size of Data centers creates two significant challenges for enterprises:**
  – **Data center facilities spend** (CapEx and OpEx) is a **large, quickly growing and very inefficient** portion of the total IT budget in many technology intensive industries such as financial services and telecommunications. **Some intensive data center users will face meaningfully reduced profitability if current trends continue**
  – **For many industries, data centers are one of the largest sources of Greenhouse Gas (GHG) emissions.** As a group, their overall emissions are significant, in-scale with industries such as airlines. Even with immediate efficiency improvements (and adoption of new technologies) **enterprises and their equipment providers will face increased scrutiny given the projected quadrupling of their data-center GHG emissions by 2020**
- The **primary drivers of poor efficiency** are:
  – Poor demand and capacity planning within and across functions (business, IT, facilities)
  – Significant failings in asset management (6% average server utilization, 56% facility utilization)
  – Boards, CEOs, and CFOs are not holding CIOs accountable for critical data center facilities CapEx and data center operational efficiency
- **Improving efficiency** is the best near term means to solving the twin challenges of rising spend and GHG emissions. We propose a three part solution **to double IT energy efficiency by 2012** and to arrest the growth of GHG emissions from data centers:
  – Mandate inclusion of **true total cost of ownership** (including data center facilities) in business case justification of new products and applications to throttle excess demand
  – Rapidly mature and integrate **asset management** capabilities to reach the same par as the Security function
  – Formally move accountability for data center critical facilities expense and operations to the CIO and appoint internal "Energy Czars" with an operations and technology mandate to **double IT energy efficiency by 2012**
- To achieve this doubling of energy efficiency CIOs, equipment manufacturers, as well as industry groups in dialog with regulators should quickly establish automotive style "CAFE" metrics that will **measure the individual and combined energy efficiency of corporate, public sector and 3rd party hosted data centers.** We propose one metric here for discussion and adoption. This metric would deliver immediate financial and transparency benefits to executive management of enterprises large and small and could become a government recognized measure of efficiency



# WE PROPOSE A THREE PART SOLUTION TO IMPROVING DATA CENTER EFFICIENCY

1
- Improve **IT asset management capabilities**
- Improve IT **demand forecasting** capabilities
- Promote regular dialog between business, IT, and Facilities
- Use new technology to increase **server utilization**
- Optimize current facilities utilization with a view on power cost
- Ensure that solutions are not over-designed
- Include **energy efficiency as an important criteria** in hardware procurement
- Implement facilities best practices




2
- Use **true total cost of ownership** (TCO) of a data center by incorporating facilities cost
- Compute TCO over entire life span of data center
- **Increase transparency** of data center costs
- Include data center TCO in application and infrastructure decisions

3
- Develop an integrated plan, measurable goals and timeline to enhance data center efficiency
- Move accountability for facilities expense (CapEx and OpEx) and facility operations to the CIO
- Appoint internal "**Energy Czars**" with a mandate to **improve data center efficiency while maintaining business availability and reliability needs**
- Implement chargeback for existing apps
- Improve large CapEx approval process for data centers
- Publicly commit to green house gas reduction targets

Source: McKinsey analysis

# DEVELOP MATURE ASSET MANAGEMENT AND IT PRODUCTIVITY CAPABILITIES

| | |
|---|---|
| Demand management | • **Ensure technical input from solution architect during RFP/RFI process**<br>• **Aggregate pipeline forecasts with solution architect and data center operations**<br>• **Use stage gate approach to qualify likelihood of demand** |
| Configuration/ location | • Build larger shells (or campuses of shells) by dividing floor space into smaller logical units ("fields") that are engineered to specific workloads and built without major M&E interruptions<br>• **Optimize current location portfolio with a view of operational and energy spend** |
| Layout/ Cabinet allocations | • **Rationalize cabinet allocation by eliminate/combine cabinets with few assets and discouraging allocations of space by whole cabinet to business units/ LOBs**<br>• Verify that allocated cabinets are used, don't report allocate cabinet as used automatically<br>• Utilize ITIL configuration mgt to track asset utilization/chargeback/de-commissioning |
| Density | • Reduce role of support infrastructure (routers/SANS) to contain density requirements<br>• Optimize rack utilization by eliminating unnecessary peripherals and fully loading each rack |
| Utilization | • **Virtualize/stack to reduce the number of physical servers; increase rack utilization**<br>• **Kill comatose servers and storage as up to 30% of server may be "dead"**<br>• Enable hardware power save features<br>• Eliminate network port redundancy |
| Sourcing | • Maintain internal control on most critical systems and co-locate less critical services<br>• Move non critical system to managed provider in a virtualized environment with expectation to move more as the services mature and establish better track record for reliability<br>• Include energy efficiency as an important criteria in hardware procurement |
| Facility operations | • Measure and report energy efficiency<br>• Optimize cooling unit set-points, balance number of cooling units running, number, and location of perforated tiles with actual load<br>• Optimize mechanical plant operation, raise chilled water supply temperature, eliminate "dueling" cooling units, utilize "free-cooling" opportunities, monitor humidification/dehumidification energy<br>• Seal cable openings and install blanking plates |

Source: McKinsey analysis





# ENHANCE DEMAND FORECASTING CAPABILITIES

| Best practices | Description |
|---|---|
| Improve forecast accuracy | • Track variation in forecast accuracy, incentivising business and IT to minimise deviations<br>• Use stage gate approach to qualify likelihood of demand<br>• Use tools and processes to capture and collate command |
| Build dynamic demand models | • Incorporate drivers to account for organic growth, unplanned business events and business cycles<br>• Use scenario models to understand how different potential scenarios drive data center capacity |
| Involve solutions architects | • Ensure technical input from architects during design process<br>• Ensure data center representation in projects approval process<br>• Design Applications and hardware to optimize computing |
| Aggressively pursue demand reduction | • Consider various ways to reduce data center space and power demands, from application and infrastructure sizing through to floor optimization.<br>• Instill culture of treating data center capacity as a scarce and expensive asset rather than as a bathtub to be filled |
| Establish business-technology dialog | • Ensure Technology teams present clear options trading off between key business drivers and underlying costs e.g., true cost of increments of availability, opportunity to acquire less floor space if businesses adopt wholesale virtualization, etc. |
| Draw economic connection between business demand and true TCO | • Develop analytic approach for connecting business demand to application requirements, application requirements to infrastructure requirements and infrastructure requirements to data center requirements |

**Value at stake from effective demand forecasting**

- 15-25% reduction in overall operational costs by avoiding overbuilds
- Delayed construction of incremental power and cooling capacity reduces CapEx

Source: McKinsey analysis

# OPTIMIZE CURRENT LOCATION PORTFOLIO WITH A VIEW ON OPERATIONAL AND ENERGY SPEND

DISGUISED CLIENT EXAMPLE

Buy – Prioritize for capability building; Hold – Invest to sustain; Sell – No investment exit

| Location | Ownership status | Tier 3/4? | >10K sq. ft? | Space/power available? | 12-18 months | 18-36 months |
|---|---|---|---|---|---|---|
| | Criteria for designation | | | | | |
| 1. US location 1, West | Leased | Y (3-) | Y (26K) | Y | Buy | Buy |
| 2. UK location 1, Europe | Leased | Y (3) | Y (16K) | N | Buy | Buy |
| 3. Location 1, Asia-Pac | Leased | Y (3+) | Y (11K) | Y | Buy | Buy |
| 4. Location 2 , Europe | Leased | N (2) | Y (12K) | Y | Hold | Hold |
| 5. Location 1, S America | Owned | N (2) | N (4K) | Y | Hold | Hold |
| 6. Location 2, Asia-Pac | Leased | Y (3) | N (6K) | N | Hold | Hold |
| 7. US location 2, Mid west | Owned | N (2) | Y (99K) | N | Hold | Sell |
| 8. US location 3, Mid west | Leased | N (1) | Y (23K) | Y | Hold | Sell |
| 9. Location 3, Europe | Owned | N (1) | Y (10K) | N | Hold | Sell |
| 10. Location 2, S America | Owned | N (1) | N (1K) | Y | Hold | Sell |
| 11. Location 2, Asia-Pac | Leased | N (2) | N (.8K) | N | Hold | Sell |
| 12. Location 2, Asia-Pac | Leased | N (1) | Y (12K) | Y | Hold | Sell |
| 13. Location 3, Asia-Pac | Leased | N (2) | N (5K) | N | Hold | Sell |
| 14. Location 4, Asia-Pac | Leased | N (2) | N (4K) | Y | Hold | Sell |

- Consider locations with natural cooling potential and access to clean power
- Optimize DC size to reduce power consumption and enhace floor density
- Consider migrating less critical applications to third party host to free-up capacity

Source: McKinsey analysis

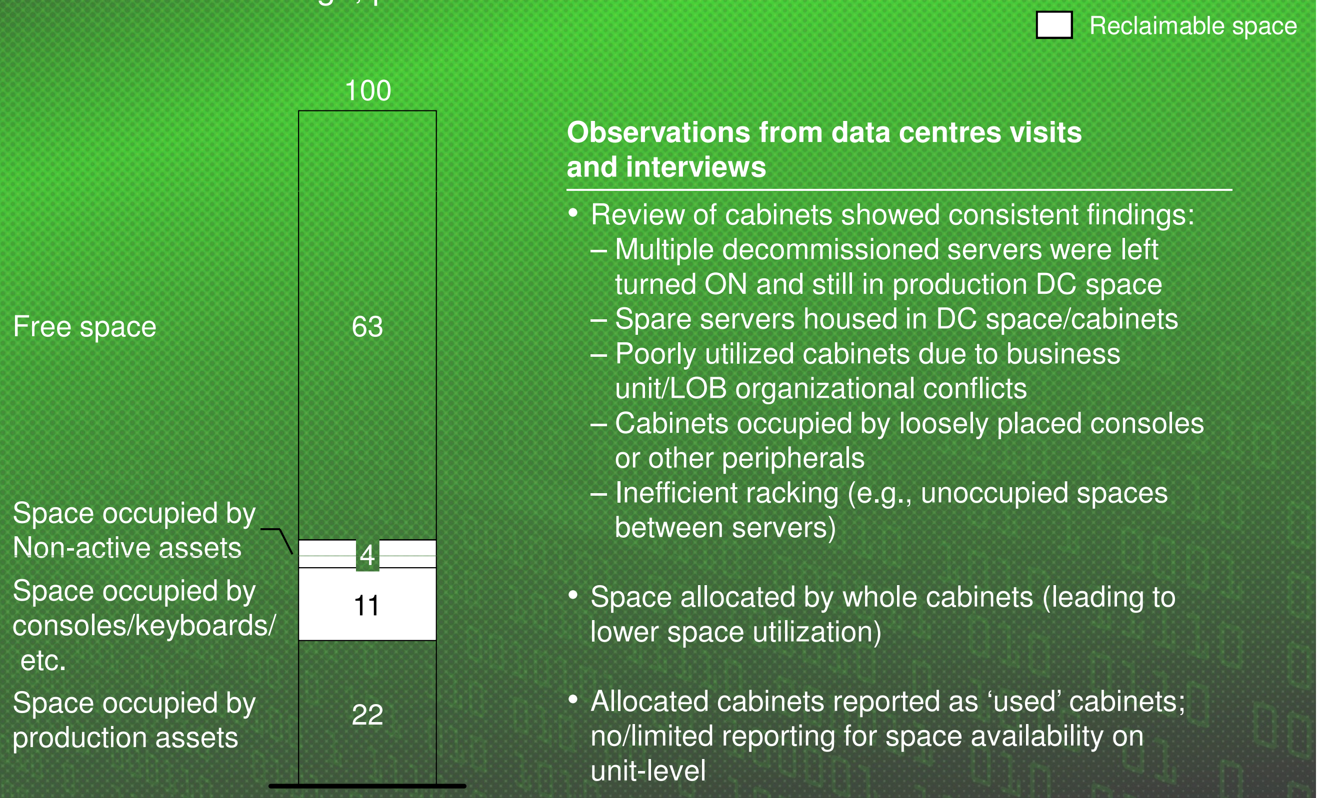




# RATIONALIZING CABINET ALLOCATION CAN SAVE UP TO 15% OF COMPUTER ROOM SPACE REQUIRED

DISGUISED CLIENT EXAMPLE

Physical audit of cabinets usage, percent




**Observations from data centres visits and interviews**

- Review of cabinets showed consistent findings:
  - Multiple decommissioned servers were left turned ON and still in production DC space
  - Spare servers housed in DC space/cabinets
  - Poorly utilized cabinets due to business unit/LOB organizational conflicts
  - Cabinets occupied by loosely placed consoles or other peripherals
  - Inefficient racking (e.g., unoccupied spaces between servers)
- Space allocated by whole cabinets (leading to lower space utilization)
- Allocated cabinets reported as ‘used’ cabinets; no/limited reporting for space availability on unit-level

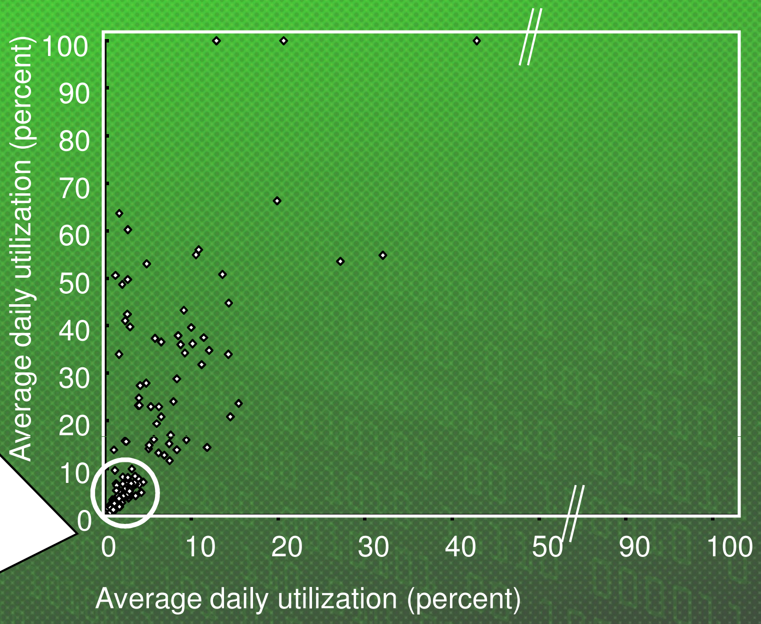
* Representative sample of 65 cabs

Source: McKinsey analysis

# POWER USAGE CAN BE REDUCED SIGNIFICANTLY - UP TO 30% OF INSTALLED SERVERS ARE POTENTIALLY 'DEAD'

DISGUISED CLIENT EXAMPLE

Peak vs. average utilization over 3-12 month period (n=458)*, percent




- 146 out of 458 servers (32%) could be dead, as they have peak average and average utilizations below 3%
- 99 servers (63%) have peak and average utilization below 10% suggesting significant overcapacity

- Easy to implement initiative
- Requires only consistent, up-to-date asset database and CPU utilization tracking tool
- Additional benefits of lower power/cooling cost, less monitoring, and recovery of cabinet space

* Sample of 4 DC production Wintel, Unix servers

Source: McKinsey analysis; Utilization measurement tool for 458 servers





# USE VIRTUALIZATION TECHNOLOGY TO IMPROVE UTILIZATION AND REDUCE NUMBER OF PHYSICAL SERVERS

ILLUSTRATIVE

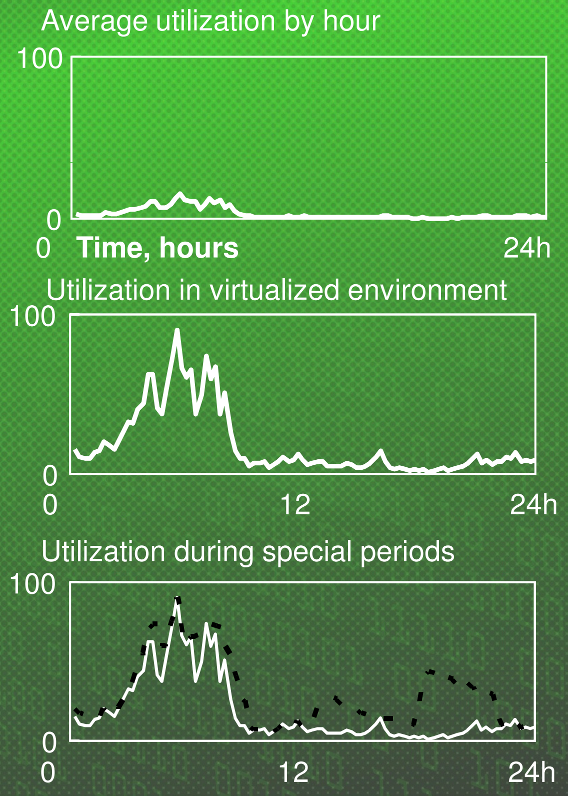
Average utilization by hour, percent

Average utilization is very low even during peak hours in non virtualized environment

Consolidate servers by stacking and virtualizing to increase average utilization

Consider weekly, seasonal and other (e.g., year-end) variation in utilization

Consider power saving for non-production hardware



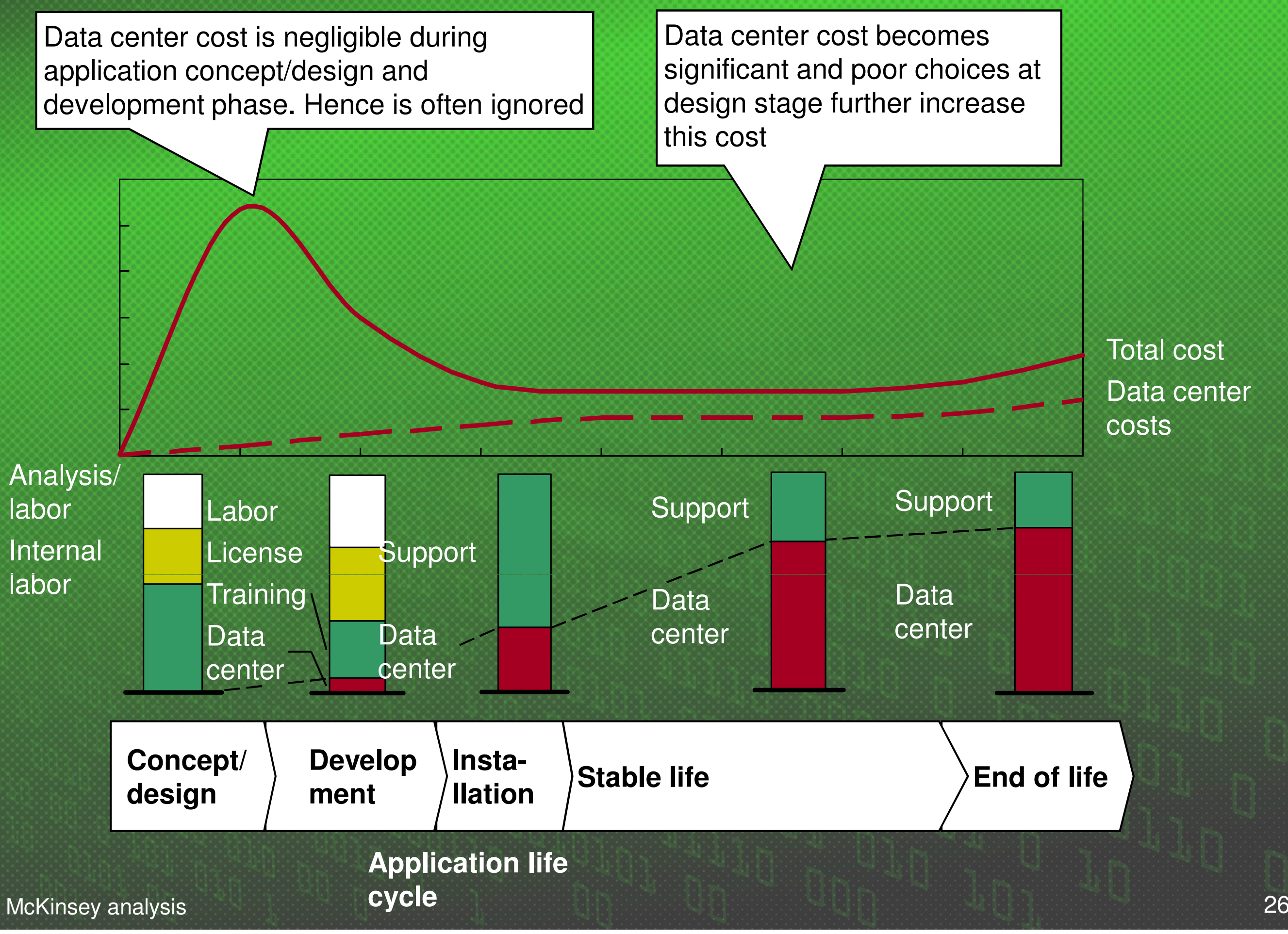

**Physical servers count**

| | Physical servers count |
|---|---|
| Today | 100 |
| Virtualized | 35 |

-65%

Source: McKinsey analysis

# OVER ENTIRE LIFE SPAN, DATA CENTER IS A SIGNIFICANT COMPONENT OF IT COST AND SHOULD BE INCLUDED IN APP/INFRA DECISIONS




Source: McKinsey analysis





# INDEPENDENT ARCHITECTURE AND STANDARDS COUNCIL REVIEWS ENSURE THAT IT SOLUTIONS ARE NOT OVERDESIGNED

| | |
|---|---|
| **Pre-production** | • Unnecessary software is not included<br>• Software are architecture is sized for effect and scale<br>• Hardware is correctly sized<br>• Software complies to all existing standards<br>• Hardware life cycle is clearly marked<br>• Life cycle costs of hardware and facilities are included in business cases |
| **Production** | • Production clearance is signed before moving to production environment<br>• Standards are maintained throughout the life of given hardware and software<br>• Applications and associated software are upgraded to ensure consistency<br>• As many applications moved to shared environment as possible |
| **Post-production** | • Software and hardware are decommissioned by due date<br>• All associated systems are decommissioned<br>• All data moved to tapes and shared storage! Purge any data not needed anymore<br>• Reuse rack space and IT kW capacity |

Source: McKinsey analysis

# DEVELOP AN INTEGRATED PLAN TO SIGNIFICANTLY IMPROVE DATA CENTER EFFICIENCY BY 2012

| | Summary | Rationale | Benefits |
|---|---|---|---|
| **Move full DC operations respo-nsibility to CIO** | • Centralize accountability for spend and performance | • Currently accountability is divided between facilities/corporate real state and IT which distorts total cost view | • Allows CIO to make rational decisions on facilities<br>• Brings all DC cost under a single standard reporting<br>• Ensures single point responsibility |
| **Appoint "energy czars"** | • Integrate and prioritize energy-efficiency measures | • Energy is "nobody's" business today<br>• Lack of awareness about the design choices to optimize energy usage | • Include energy consumption and facility costs as a key criteria for IT project ROI analysis and decision making<br>• Bring accountability ? |
| **Double energy efficiency by 2012** | • Quickest and easiest way to improve return on assets and reduce GHG emissions | • Process improvement and current technology can drive energy efficiency significantly higher | • Sets clear directions for the company<br>• Significantly lowers cost |
| **Publicly commit to emission targets** | • Raise commitment to and profile of targets within organization | • Many companies can reduce GHG emissions without adversely affected their day-to-day business | • Proactive addresses a political issue which otherwise might be mandated by regulators, boards, or NGOs |

Source: McKinsey analysis





# MAKE CIO ACCOUNTABLE FOR EFFICIENCY OF DATA CENTERS

**Siloed organizations**

- **Facilities and IT teams have limited interactions** when **designing or** efficiently **operating data centers** leading to multiple layers of conservatism and waste. There is little cross-functional learning and coordination
- Executive decision makers are **not provided with sufficient facility economic outcomes and alternatives** resulting from IT application investment decisions

**Limited transparency**

- Facilities have **intelligence on IT power consumption**, but **no insight into how IT equipment being utilized,** how efficiently power within IT hardware is being utilized, nor what the future is. This leads to over provisioning
- The data center **electrical bill** is likely to be **included within a larger electrical bill** and the bill typically does not go to IT
- **Tools for modeling IT electrical consumption** are not widely available and are **not commonly used during data center design**

**Misaligned metrics**

- **Facility costs (both OpEx and CapEx) not clearly linked to any particular IT application decision nor IT operating practices**. They are therefore viewed as inevitable
- Few, if any, metrics link facilities and corporate real estate groups with IT/CIO efficiency metrics

Source: APC "Implementing energy efficient data centers"

# PUBLICLY COMMIT TO EFFICIENCY AND EMISSIONS IMPROVEMENTS AND TARGETS FOR IT (INCLUDING DATA CENTERS)

**Some leading firms have begin to establish GHG emission targets . . .**

BT has unveiled plans to reduce carbon emissions by 80% by adopting a number of initiatives including using clean energy, providing flexible to 17,000 employees.

EDS will reduce carbon emissions by 25% in Australia/New Zealand by 2010 using video conference technology and cutting air travel by a third

HSBC has become carbon neutral in 2005 by using a four step program that includes measuring carbon footprint, reducing energy consumption, buying clean energy, and offsetting CO2 emissions

**. . . Some firms have are also creating data center specific targets . . .**

Barclays will use dynamic smart technology to reduce energy consumption in its new data center in Gloucester, UK to save 13% of total energy used

IBM will use green technology at Boulder, CO based data center by involving high density computing, and energy efficiency cooling to double DC capacity without increasing energy used or emissions

Citibank will improve DC efficiency by building a LEED certified DC by using clean energy and reduce water usage. Overall, Citi will reduce GHG emissions by 10% by 2011

**Existing efforts focus on component-level efficiency but are not yet mature or integrated enough to drive system level results**

- Intel
- Climate Savers
- Green GRO
- Energy Star

Source: Press reports; analyst reports



# 5TH GENERATION DATA CENTER DESIGN IS ESPECIALLY IMPORTANT IN NEW DATA CENTERS AND SHOULD REFLECT KEY IDEAS OF GREEN

| | Conventional Design | 5th Generation Green DC Design |
|---|---|---|
| **Concept** | • High density computing minimizing size of raised floor (e.g., Google)<br>• Campus located for energy prices | • Medium/high density<br>• Campus located for natural cooling, e.g. "cold" location, cold water available, etc. |
| **Hardware** | • High density (e.g., blades)<br>• DC conversion done per device/rack | • Controlled moderate density to allow natural cooling (but less coincidental dead space)<br>• Broad operating temperature envelope, e.g., 5-40˚C<br>• DC power conversion done for entire DC |
| **Cooling** | • Chilled air cooling (UPS backed) | • Fresh air cooling with occasional refrigerated backup<br>• Usage of direct water cooling, where possible |
| **Electricity supply** | • AC power<br>• 200 W/sq ft<br>• 6,000 W/rack | • DC power<br>• 125 W/sq ft<br>• 3,750 W/rack<br>• Supplied mostly with renewable, C02 free energy |
| **What you must believe** | • Raising power density is efficient means of meeting demand<br>• ICT equipment will utilize available power | • New technologies such as large DC converters and fresh air cooling deliver significant savings<br>• Carbon footprint important design principle in addition to total energy consumption |

**Bits and pieces of these measures applied in individual DCs, however coherent, systematic "green" DC design yet to emerge**

Source: McKinsey analysis

# WHILE NEXT GEN DESIGN HAS RECEIVED GREAT PRESS, IMPROVING EFFICIENCY OF EXISTING SITES WILL LOWER ENERGY USAGE AND REDUCE GHG FASTER AND MORE SIGNIFICANTLY WITH LESS COST

**Typical scenario for “green data center**

| | |
|---|---|
| **Concept** | • Site located for natural cooling<br>• Site located for green energy |
| **IT Hardware** | • Broaden reliable temperature band, e.g., 5-40˚C<br>• Direct current power input |
| **Cooling** | • Direct chilled water cooling to chips<br>• Increased efficiency at partial load<br>• Fully utilize free-cooling |
| **Electrical** | • Increased efficiency at partial load |
| **Core belief** | • Carbon footprint important design principle in addition to total energy consumption |

- Direct current and/or water cooling requires industry wide technology shift
- Larger temp band requires industry consensus
- Most applicable for new data centers
- Medium to long-term timeframe

**Improving operational efficiency**

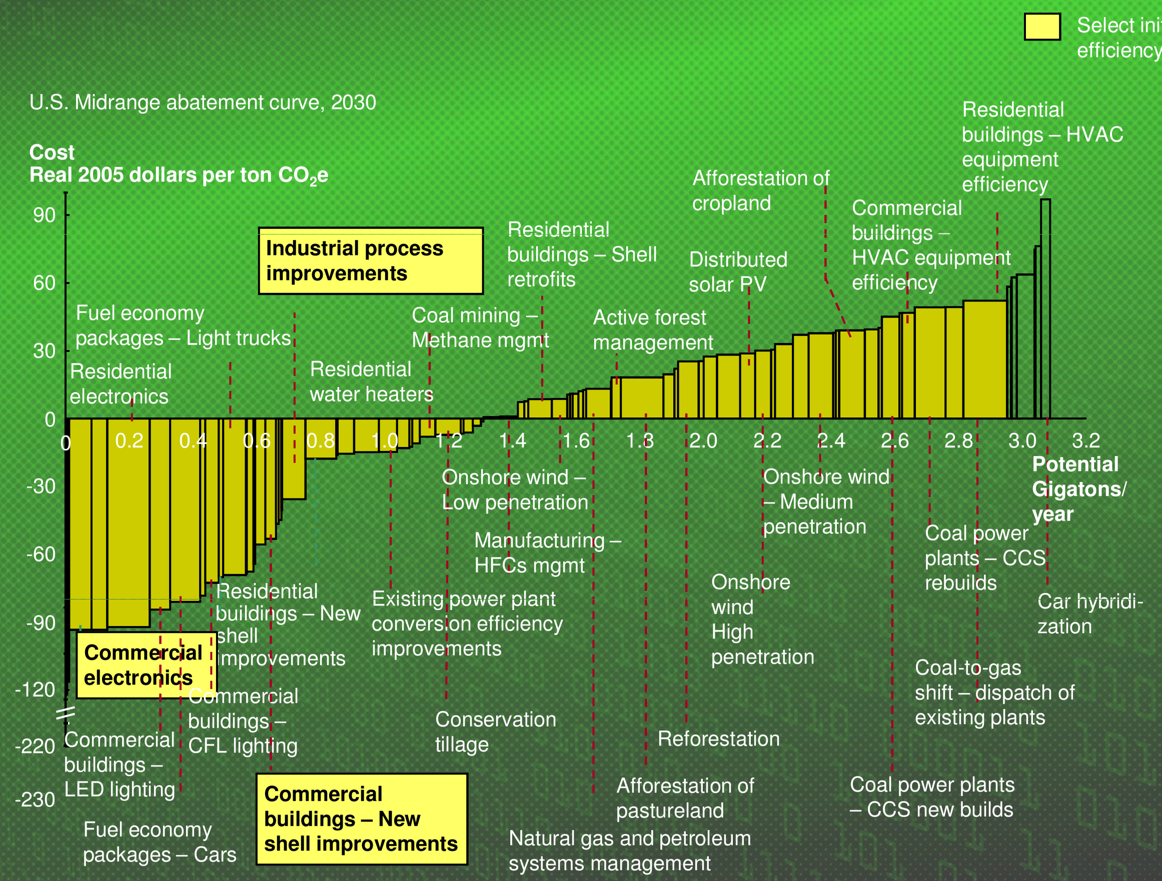
| | |
|---|---|
| **Demand management** | • Rationalize IT demand<br>• Reduce/eliminate unnecessary applications |
| **Smart “Tier” sourcing** | • Focus internal control on most critical systems; source others from co-lo (e.g., HR) |
| **IT asset efficiency** | • Increase server utilization<br>• Virtualize servers<br>• Decommission redundant server, and eliminate network port redundancy<br>• Buy energy efficient replacement hardware |
| **Computer room utilization** | • Divide floor space into smaller building bays engineered to specific density workloads<br>• Reduce IT infrastructure (routers/ SANS) to contain density |

- Simple, incremental change, known technology
- Low incremental capital investment, fast payback
- Applicable for existing and new data centers
- Short to medium-term timeframe

Source: McKinsey analysis; Uptime Institute



# OPERATIONAL EFFICIENCY IMPROVEMENTS IN DATA CENTERS ARE SOME OF THE BEST OPTIONS FOR CARBON FOOTPRINT ABATEMENT




- Most data center efficiency initiatives fall to the left size of sustainability curve, suggesting high cost effectiveness and very high returns on investment
- Efficiency benefits from existing data centers are available immediately with little or no waiting period

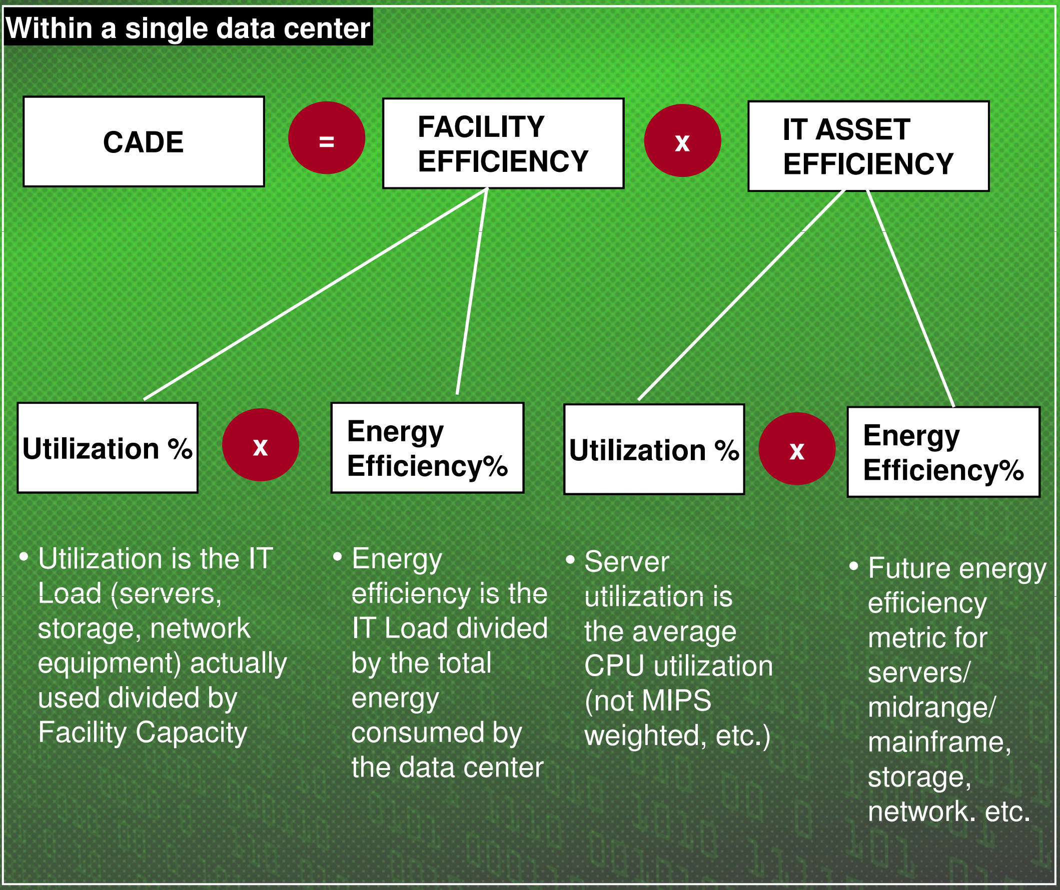
Source: McKinsey round table GHG report

# EXECUTIVE SUMMARY

- **The rapid recent (and projected) growth in the number and size of Data centers creates two significant challenges for enterprises:**
  – **Data center facilities spend** (CapEx and OpEx) is a **large, quickly growing and very inefficient** portion of the total IT budget in many technology intensive industries such as financial services and telecommunications. **Some intensive data center users will face meaningfully reduced profitability if current trends continue**
  – **For many industries, data centers are one of the largest sources of Greenhouse Gas (GHG) emissions.** As a group, their overall emissions are significant, in-scale with industries such as airlines. Even with immediate efficiency improvements (and adoption of new technologies) **enterprises and their equipment providers will face increased scrutiny given the projected quadrupling of their data-center GHG emissions by 2020**
- The **primary drivers of poor efficiency** are:
  – Poor demand and capacity planning within and across functions (business, IT, facilities)
  – Significant failings in asset management (6% average server utilization, 56% facility utilization)
  – Boards, CEOs, and CFOs are not holding CIOs accountable for critical data center facilities CapEx and data center operational efficiency
- **Improving efficiency** is the best near term means to solving the twin challenges of rising spend and GHG emissions. We propose a three part solution **to double IT energy efficiency by 2012** and to arrest the growth of GHG emissions from data centers:
  – Mandate inclusion of **true total cost of ownership** (including data center facilities) in business case justification of new products and applications to throttle excess demand
  – Rapidly mature and integrate **asset management** capabilities to reach the same par as the Security function
  – Formally move accountability for data center critical facilities expense and operations to the CIO and appoint internal "Energy Czars" with an operations and technology mandate to **double IT energy efficiency by 2012**
- To achieve this doubling of energy efficiency CIOs, equipment manufacturers, as well as industry groups in dialog with regulators should quickly establish automotive style "CAFE" metrics that will **measure the individual and combined energy efficiency of corporate, public sector and 3rd party hosted data centers.** We propose one metric here for discussion and adoption. This metric would deliver immediate financial and transparency benefits to executive management of enterprises large and small and could become a government recognized measure of efficiency



# CORPORATE AVERAGE DATA EFFICIENCY (CADE) V1.0 MEASURES DATA CENTER EFFICIENCY




**CADE can also measure across a the enterprise footprint**
- Each data center is measured independently
- A weighted-average value is determined by weighting data centers based upon installed facility capacity
- CADE can be used in conjunction with each DC's energy source(s) to determine efficiency of GHG emissions

Note: CADE, Version 2.0 is expected to include storage and networking measurements additionally
Source: McKinsey analysis

# WE PROPOSE CADE RATINGS TO FALL INTO FOUR BANDS TO NORMALIZE RATINGS AND SET TARGETS FOR IMPROVEMENT

Expected target range for most data centers: Level 3 and Level 4

| CADE level | CADE band |
|---|---|
| **Level 1** | 0-5% |
| **Level 2** | 5-10% |
| **Level 3** | 10-20% |
| **Level 4** | 20-40% |
| **Level 5** | >40% |

- CADE tiering will set efficiency targets for data center management (e.g., increase from CADE level 2 to 3 in 18 months)
- CADE bands will flex over time as companies begin standardizing on its measurement
- Additional updating when server, storage and networking energy efficiency are included in the measurement

Source: McKinsey analysis



# EXAMPLE OF ACTUAL EXISTING SITE AND IN-FLIGHT IMPROVEMENT TO EVENTUALLY DOUBLE CADE RATING

ACTUAL DATA

**Within an actual data center currently on CADE band 1**

**CADE: 1.3% Band 1** = **FACILITY EFFICIENCY: 27%** x **IT ASSET EFFICIENCY: 5%**

**FACILITY EFFICIENCY** = **Asset Utilization: 54%** x **Energy Efficiency: 44%**

**IT ASSET EFFICIENCY**: **Asset Utilization: 5%**

- Actual energy delivered to IT of 7,550kW divided by Facility capacity of 14,000 kW
- Energy reaching servers, storage and networking equipment of 7,550 kW divided by actual incoming data center energy of 17,120 kW
- IT asset utilization is average across 20,000 servers as 5%

Project underway to remove dead servers, expected to remove 6,000 servers and increase CADE by 30%

Server virtualization project expected to remove 4,000 servers and move load to other servers, increasing CADE by 40%

Sealing of cable cutout openings allows cooling units to be turned off increasing CADE by 3%

Removal of 10,000 servers worsens CADE by 42%

Source: Uptime institute

## 10 GAME CHANGING IMPROVEMENTS TO DOUBLE EXISTING EFFICIENCY

| Facility Efficiency | CADE Impact |
|---|---|
| 1 Create data center energy dashboard, harvest obvious | • 5-5% |
| 2 Seal cable cutouts, turn off excess cooling units | • 3-5% |
| 3 Increase cold aisle temp, eliminate cooling unit "dueling" | • 5-5% |
| 4 Implement free cooling | • 0-10% |
| 5 IT load reduction reduces utilization, worsens efficiency | • -3-5% |

| IT Asset Efficiency | CADE Impact |
|---|---|
| 6 Remove dead servers | • 10-50% |
| 7 Select standalone rightsizing (existing equipment) | • 30-70% |
| 8 Shared virtualization | • 30-70% |
| 9 Demand management:reduce and rightsize new demand) | • 20-50% |
| 10 Upgrade older equipment | • 20-100% |

- CADE impact of each improvement depends substantially on the "As-Found" conditions within the data center
- Goal of doubling data center efficiency is attainable by almost everyone

Source: McKinsey analysis



## MORE INFORMATION

For more information, please contact:
**William Forrest**
william_forrest@mckinsey.com
+1 (312) 551-3500

For press inquiries, please contact
**Charles Barthold**
charles_barthold@mckinsey.com
+1 (203) 977-6915

For information on the Uptime Institute, please contact:
**Bruce Taylor**
btaylor@uptimeinstitute.org
+1 (505) 946-3440