SIFMA 2009

Securities Industry and Financial Markets Association Technology Conference June 2009

Data Center Design

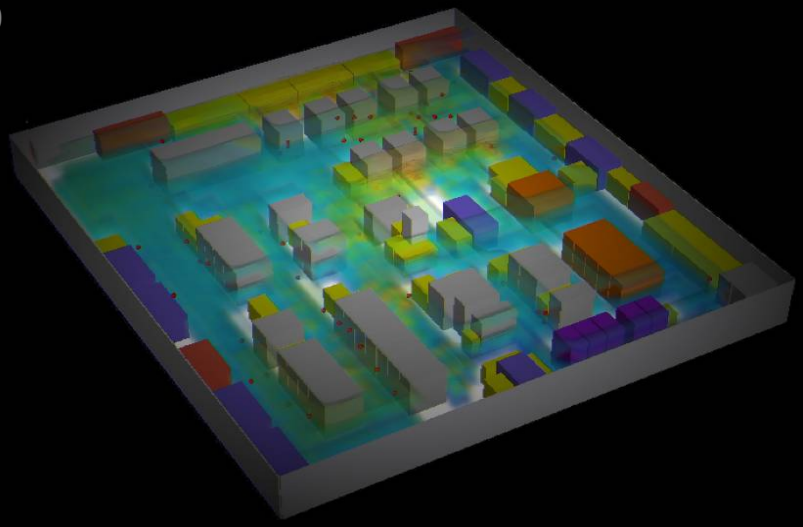
Tom Keller, Austin Lab

tkeller@us.ibm.com

Hendrik Hamann, T.J. Watson Lab

hendrikh@us.ibm.com

IBM Research



Presenters

Dr. Tom Keller
Distinguished Engineer
Austin Research Lab

- data center energy efficiency & tools
- led first system-level power management for IBM products
- technical lead IBM AIX performance
- led Los Alamos' performance evaluation of the Cray-1 serial #1 supercomputer
- at M.C.C. prototyped parallel database machine &
- led creation of the still surviving TPC-C benchmark
- Associate Director of the U. Texas Computation Center
- Chair, ACM Sigmetrics
- 50+ papers

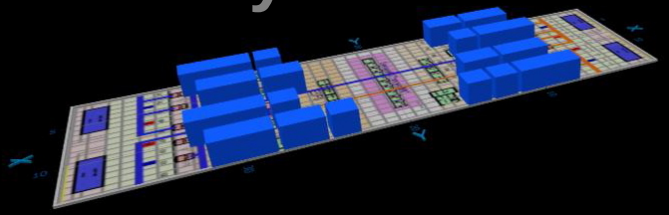
Dr. Hendrik F. Hamann
Manager Photonics and Thermal Physics
T.J.Watson Research Ctr

- physical aspects of thermal & energy management from the transistors to the data center
- nanoscale heat transfer research
- novel near-field optical microscopes to study single molecules at high spatial resolution, at joint institute between U. Colorado and NIST
- 30+ scientific papers
- 25+ patents with 25+ pending
- IBM Master Inventor
- National Academy of Sciences committee
- Industrial advisor to Universities

Agenda

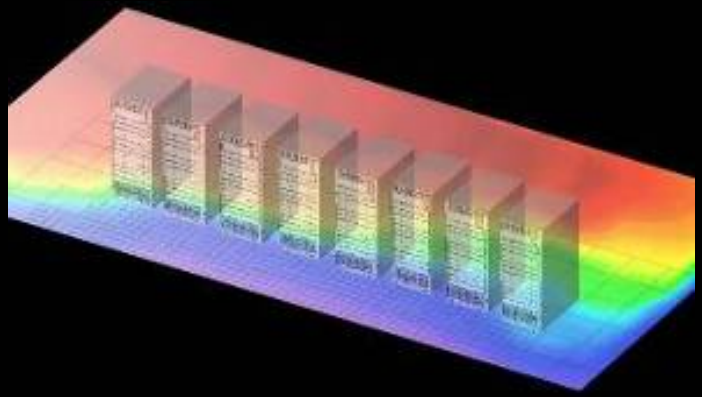
Case study from the financial Industry (Tom Keller)

- ACU Efficiencies
- Power Gap Analysis



DC Measurement and Management Technologies (Hendrik Hamann)

- How to measure, model and manage DCs
- Results & Savings

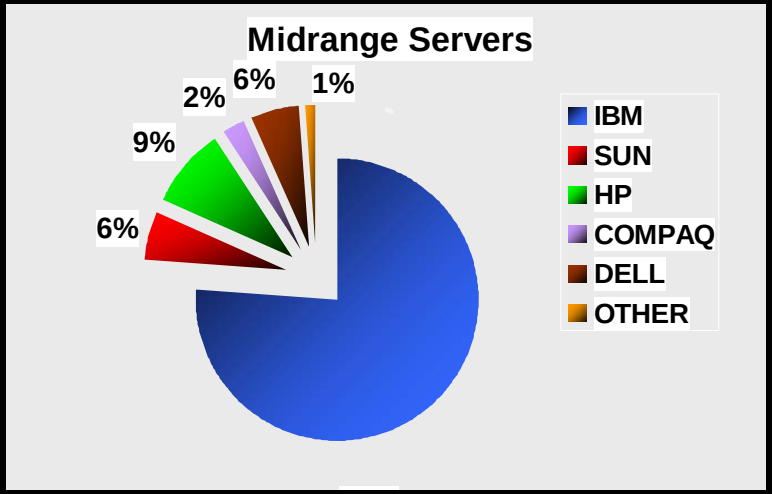


Data Centers

- Three drivers have lead to DC crisis
  - Insatiable IT demand
  - Power-limited core technology
  - Increasing energy costs
- DCs consume ~ **2 % of all US electricity**
- *Annual growth* (15 %) is non-sustainable
- DC power projected to be > 8 % of US power by 2020

- Every DC is *different*, DCs are *heterogeneous* and *change over time*
- DCs are **not as efficient** as they should
- Inefficiencies are caused by *lack of best practices*

IBM Data Centers



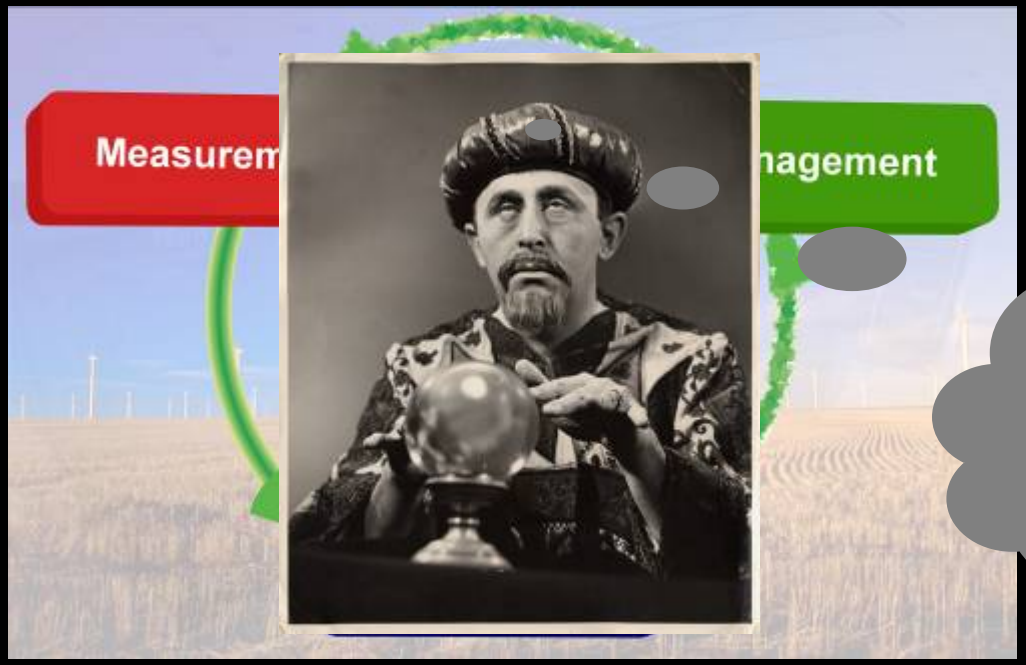
Data Centers	Americas	Pan Euro	Asia Pacific	Total
Count	135	110	170	**415**
Square Feet (Millions)	4.3	1.6	1.8	**7.7**

As of August, 2007

The Energy Efficiency Management Challenge

you know your DC could be more energy efficient, *but*

- Need “ongoing” energy efficiency measurements
  - “what you can’t measure you can’t improve “
  - no viable charge-back mechanism between IT and RESO
  - current efficiency metrics are weather snapshots
  - just too many good excuses/reasons not to implement
- Energy efficiency recommendations are like *New Year’s resolutions*



1.) How much energy ($) do I **really** save ?

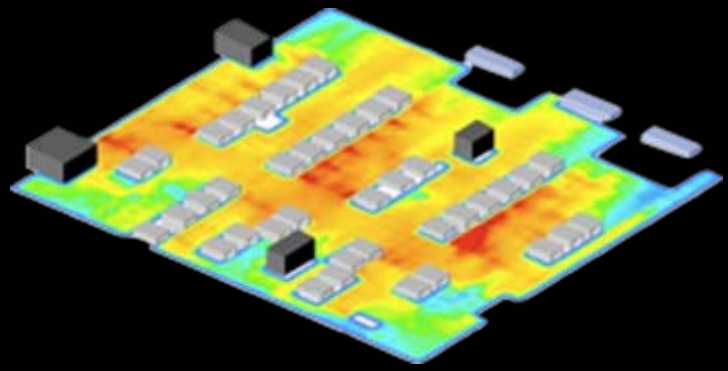
2.) Are we **really** executing on the recommendations ?

Energy Efficiency Proof of Concept Study

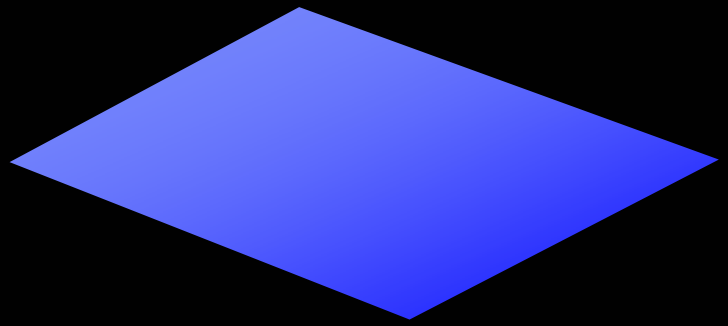
- **Tier 4 (cooling, power)**
- **40,000+ square feet**
- **Many hundreds of racks, plus**
- **Hundreds of freestanding racked machines**
- **Thousands of servers, SANs and routers**
- **40+ Computer Room Air Conditioners (ACUs)**
- **Cooling and actual power known to be underused**

** Exact measures are client confidential*

Using Research's Measurement & Management Technology, we expected above floor temperatures to look something like this

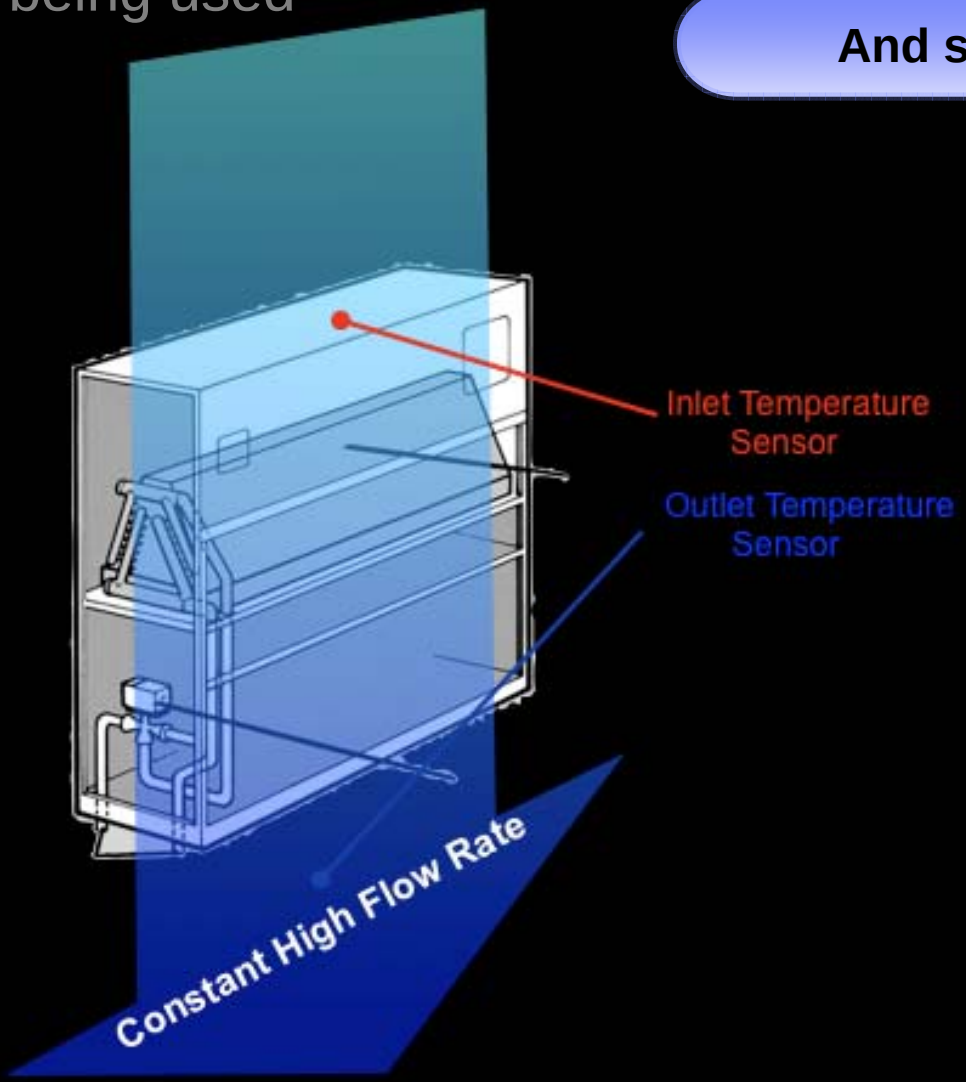


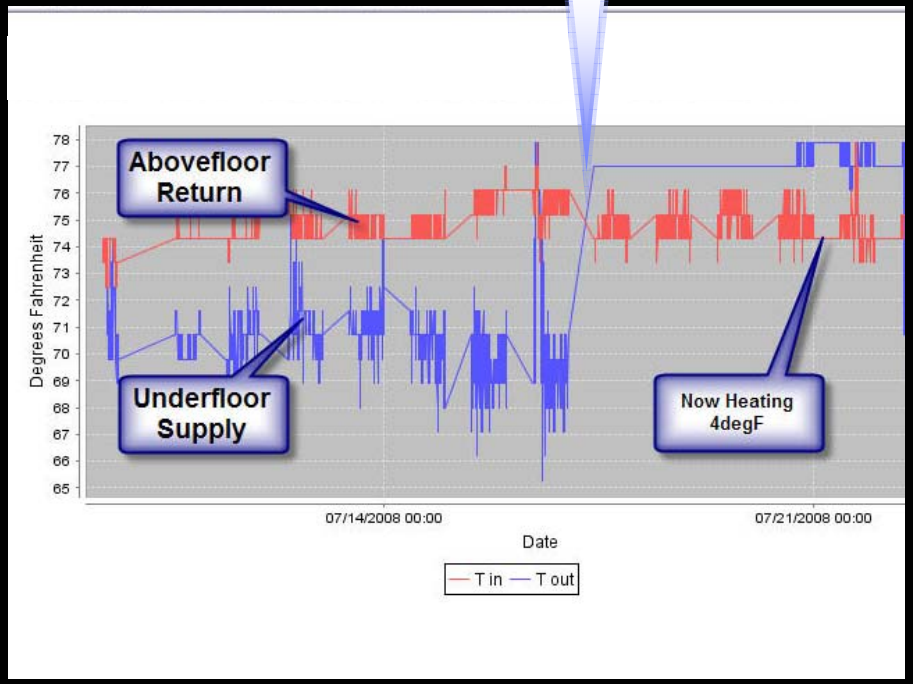
Instead, the machine room was uniformly cool at all heights



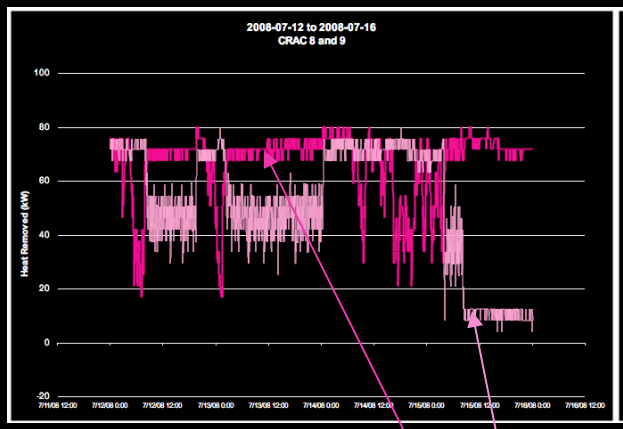
Instrumenting the ACUs revealed why -- excess cooling capacity being used

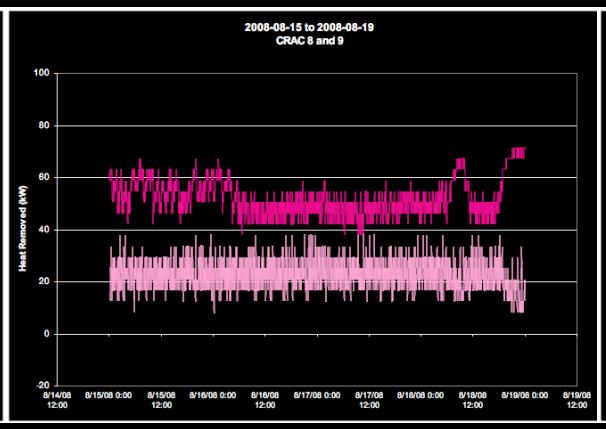
And sometimes ACUs went nuts

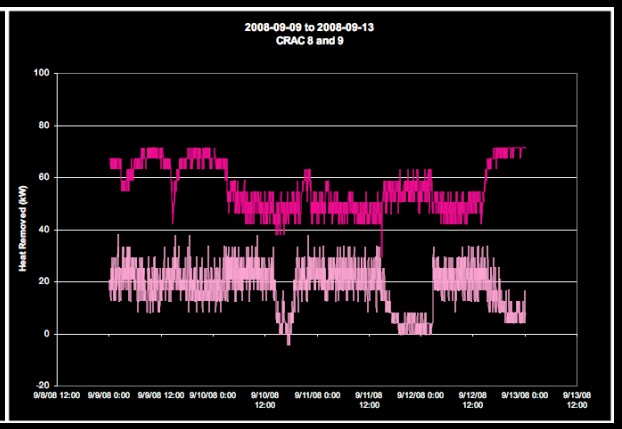


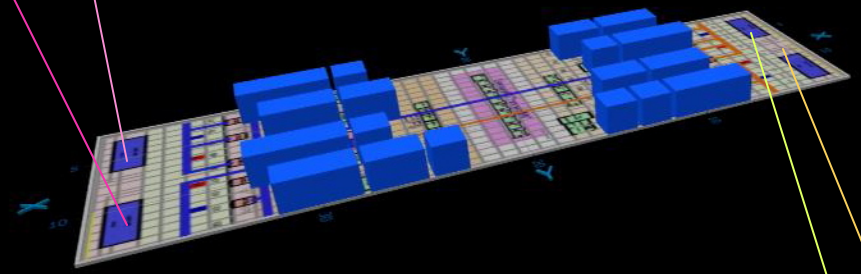


Cooling kWh histories of two pairs of ACUs

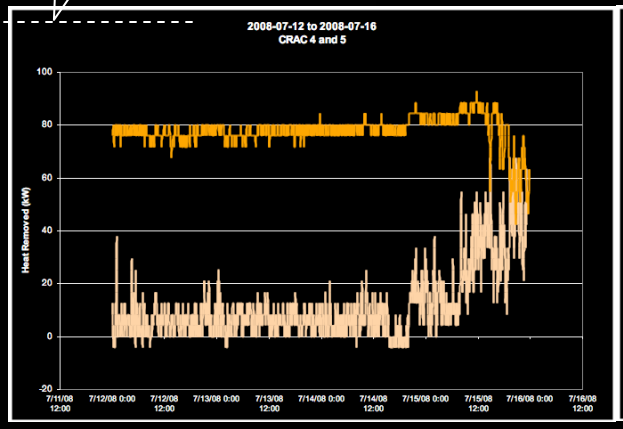


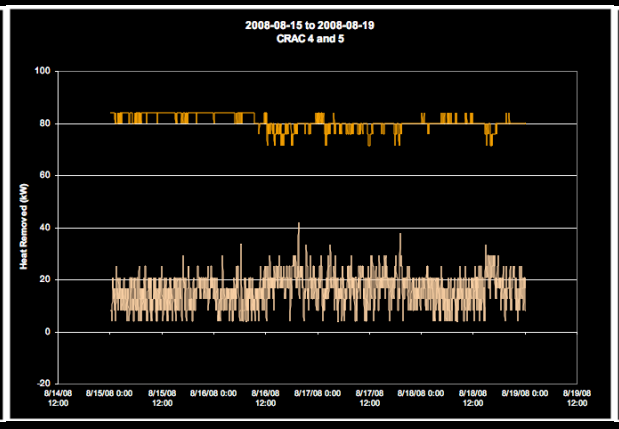


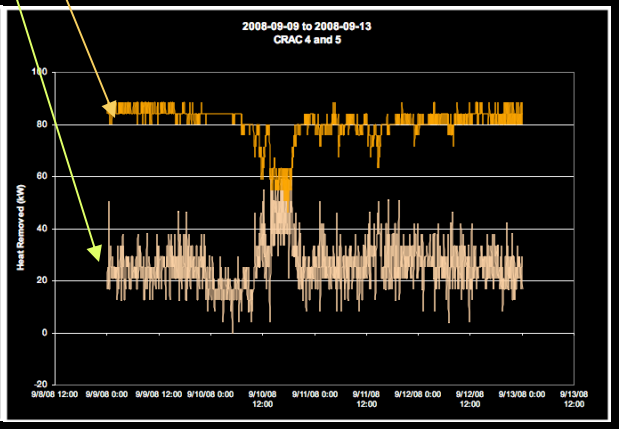




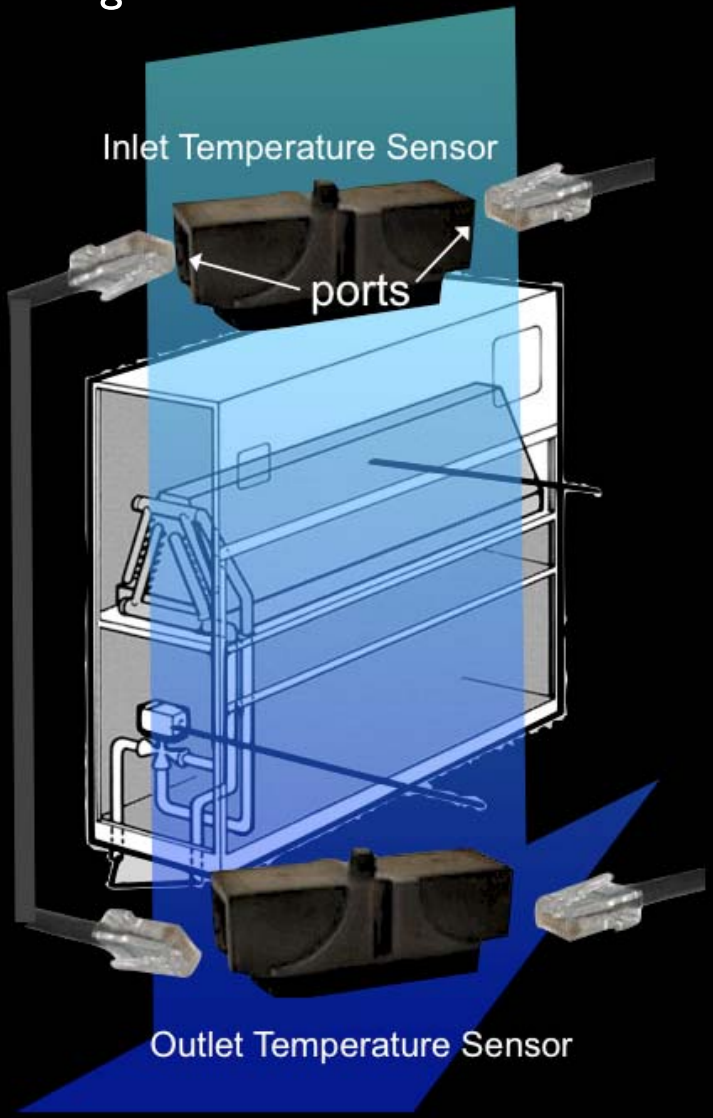
121 kWh 100% !







After computational fluid dynamics modeling, recommended a schedule for turning ACUs on and off



Installed inexpensive ACU temperature sensing network

Procedure

- Preliminary data collection
- Deploy above and below floor sensor network
- Gather temperatures and pressure measurements
- Experimental evaluation of turning ACUs off
- Turn 2 ACUs off
  - Check for system safety
    - Sensor network to check temperatures and pressures
    - Spot check air flow in critical areas via flowhood
  - Compare results to model
  - Calculate ACU efficiencies
- Repeat until suggested number of ACUs are off
- Monitor system for 1 week to gather additional data
- Deployment
- Turn on additional ACUs when a ACU fails
- Turn on all ACUs during ACU maintenance

Savings

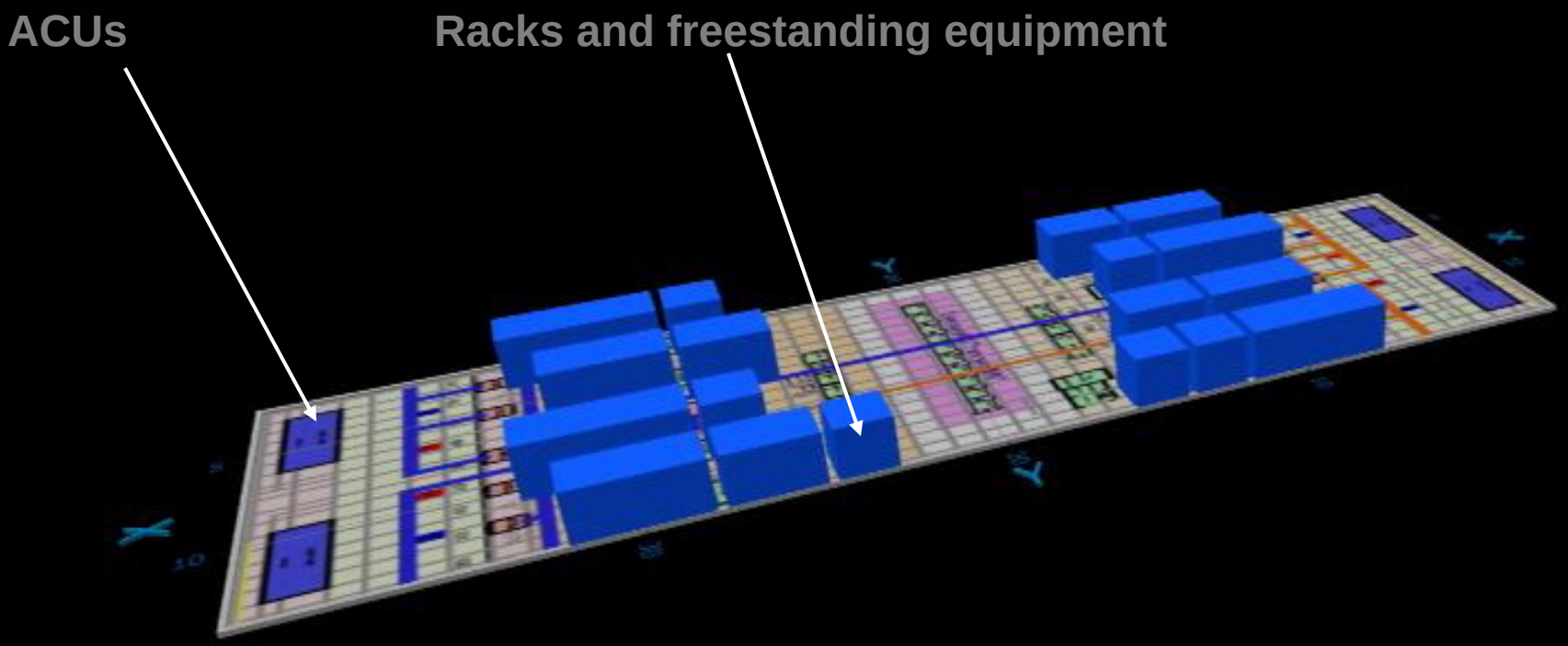
- Each off ACU will save $10,000 to $20,000 per year in electric bill

Power gap analysis

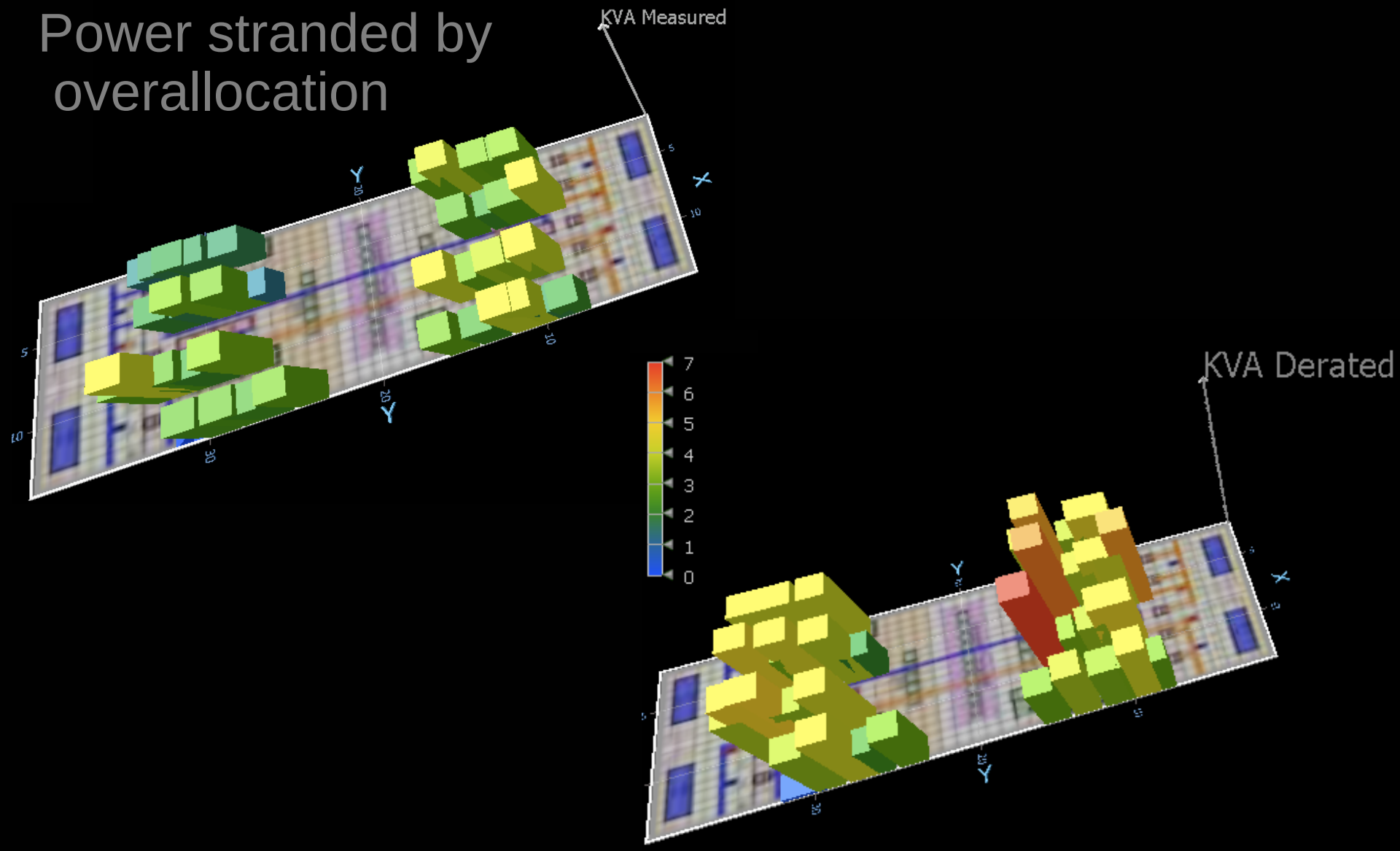
- **Displaying the difference between the actual power being used in a rack and the power allocated to the equipment in the rack showed opportunities for increasing the density of equipment in the data center**
- **Improvements in the 10’s of percent in equipment density can be made safely, deferring the construction of new data centers**

Power Measures

600 square foot snippet from machine room

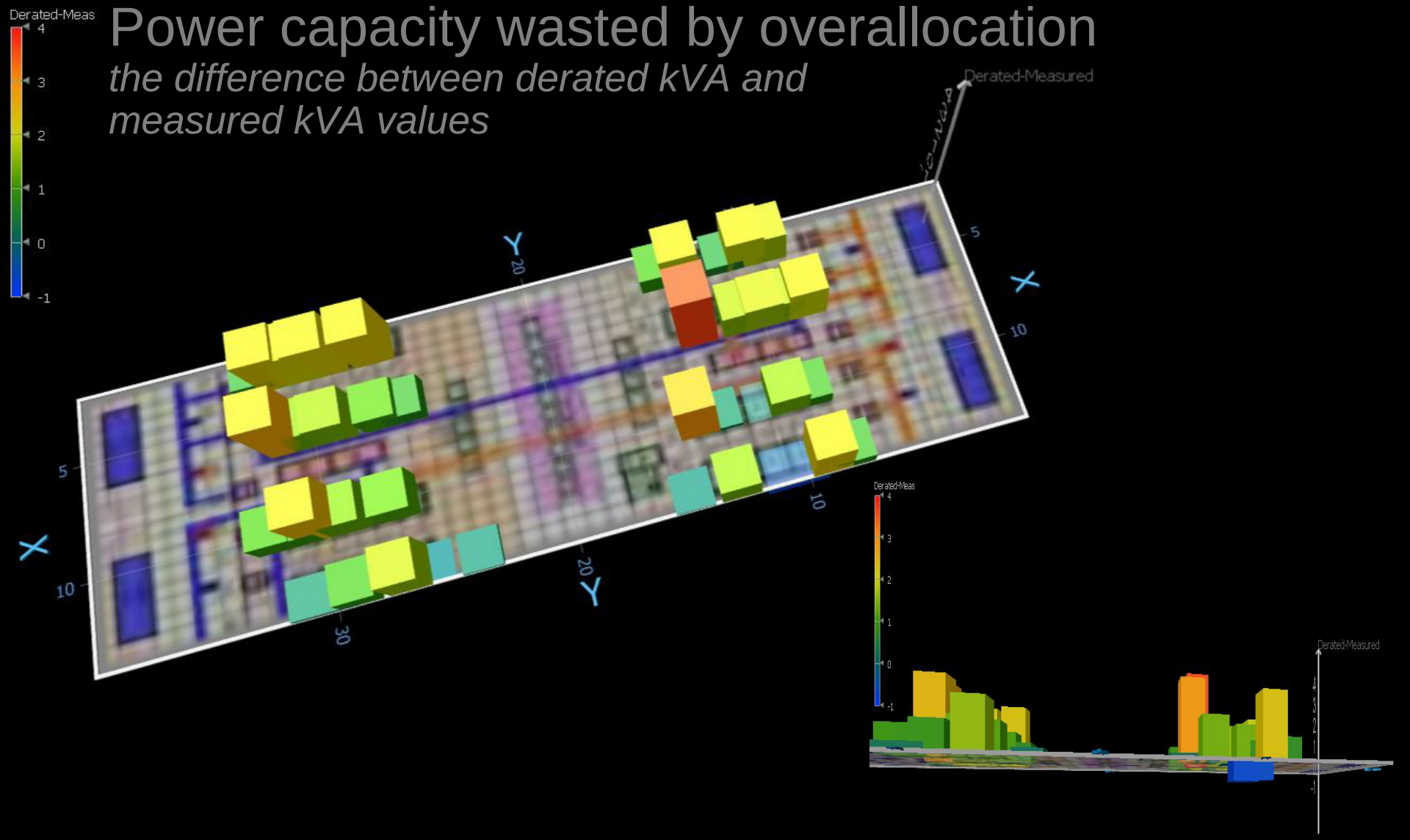


Power stranded by overallocation



Power capacity wasted by overallocation

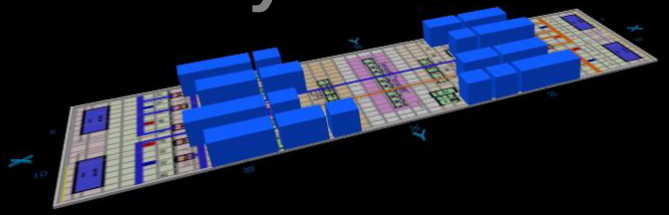
the difference between derated kVA and measured kVA values



Agenda

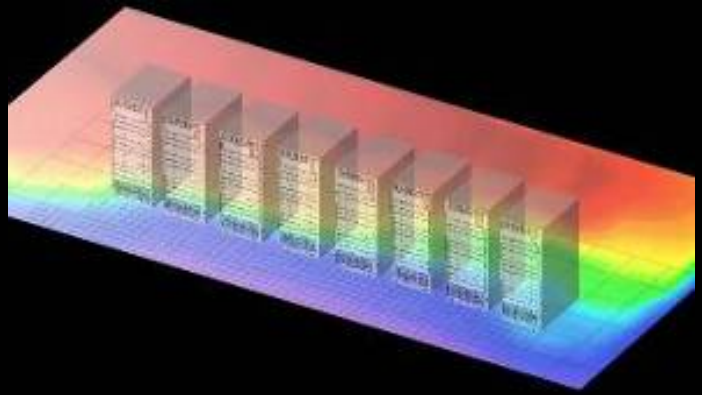
Case study from the financial Industry (Tom Keller)

- ACU Efficiencies
- Power Gap Analysis



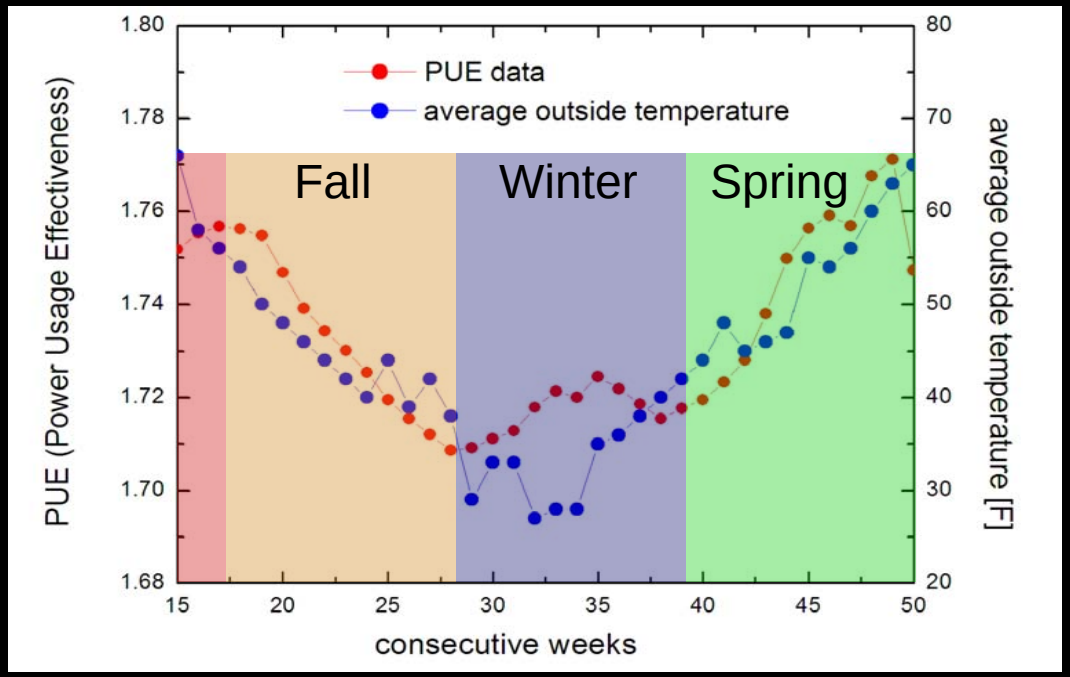
DC Measurement and Management Technologies (Hendrik Hamann)

- How to measure, model and manage DCs
- Results & Savings



Single Energy Efficiency Metric is Not Enough

PUE might be good for bragging rights, but it is only a start



PUE = Total DC Power / IT Power — widely used today

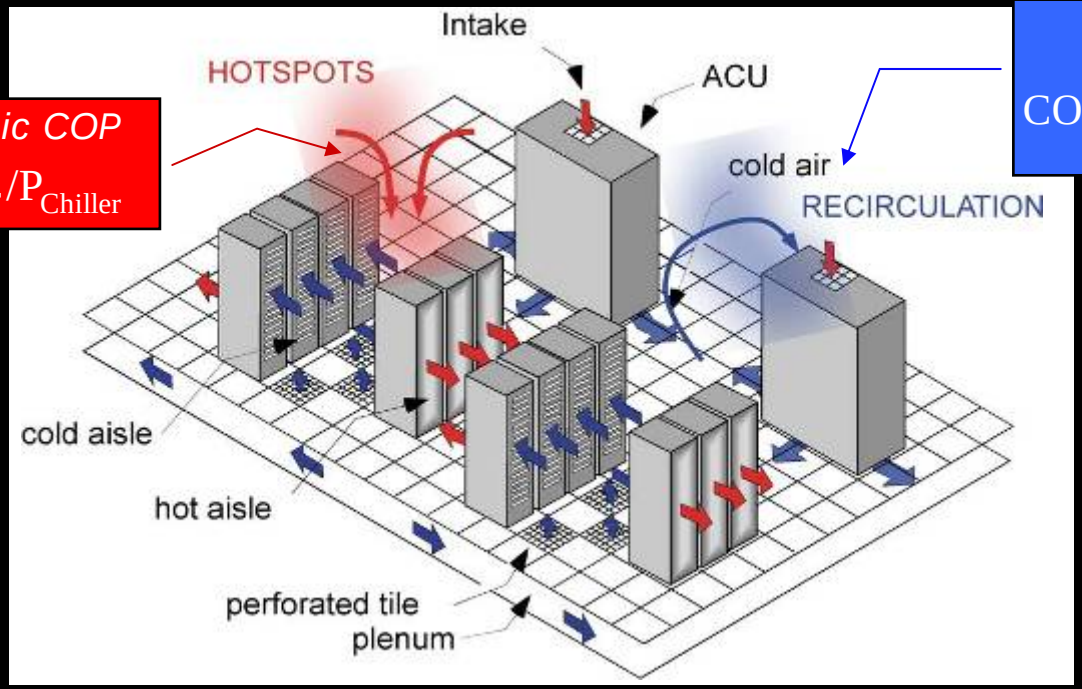
PUE metric can be problematic

- PUE is weather-, location-, application/tier-, and power density dependent
- does not include true IT performance
- metering is often not in place

Cooling Efficiency requires a more detailed Look

Thermodynamic COP

$$COP^{*}_{thermo} \approx P_{RF}/P_{Chiller}$$



Transport COP

$$COP_{trans} \approx P_{RF} / \sum_{i=1}^{\#\text{ of active ACUs}} P^{i}_{ACU}$$

Data Center COP

$$1/COP \approx 1/COP_{thermo} + 1/COP_{trans}$$

*Average Chiller COP (throughout the year)

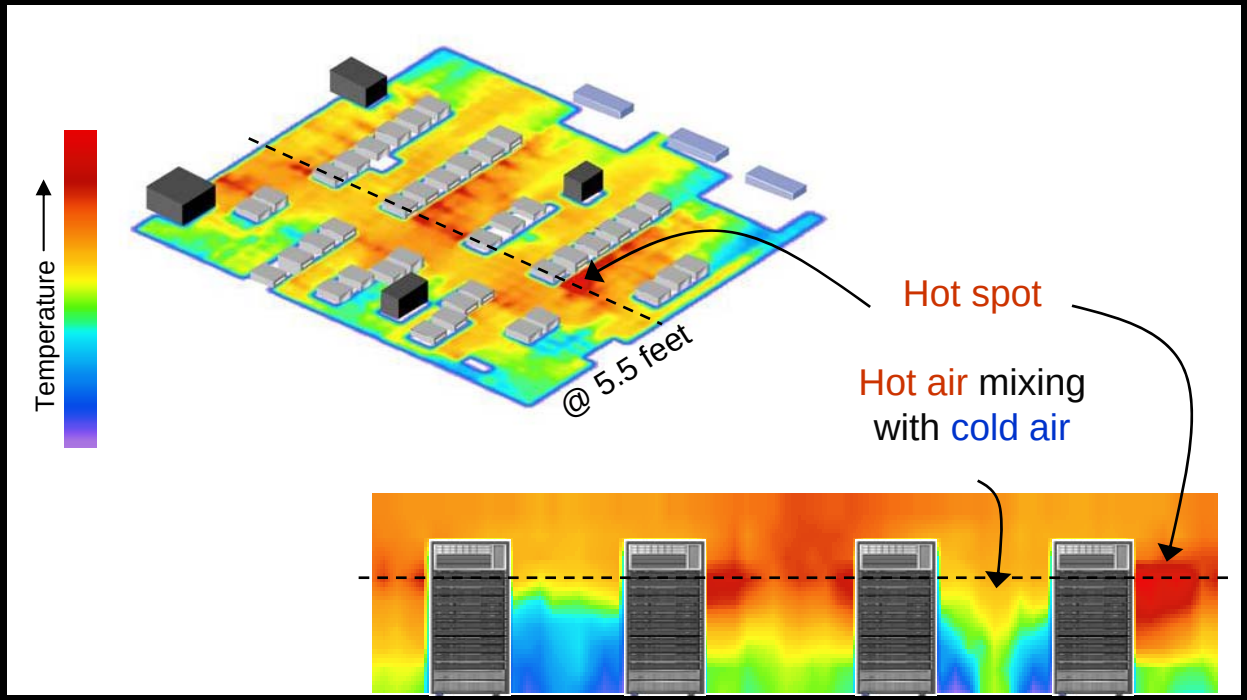
THERMODYNAMIC PART OF COOLING: HOTSPOTS / HIGH INLET TEMPERATURES IMPACT CHILLER EFFICIENCY (~ 1.7 % per F)

TRANSPORT PART OF COOLING: LOW ACU UTILIZATION IMPACTS ACU BLOWER CONSUMPTION (~ 5-8 kW/ACU)

Measuring and Managing DC Best Practices

IBM Measurement and Management Technologies (MMT 1.0)

- Optimize DC resources to reduce up to 15% of DC energy consumption
- Scans, digitize rapidly physical environment (temperature, flow, pressure etc..) of DC
- Cart tool comprises sensor network, where each sensor defines a virtual unit cell
- Integrates measurements, models and DC management



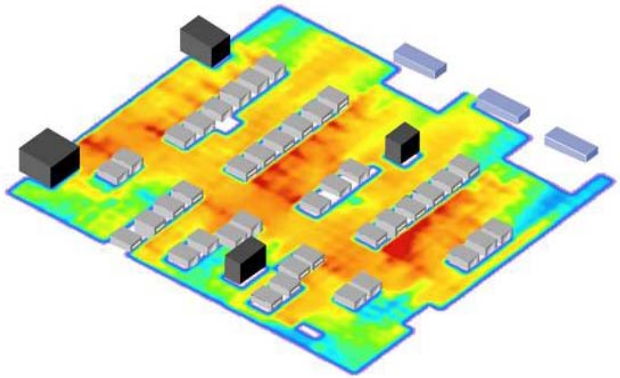
MMT 1.0 - Process

Solution Approach – *Three Steps*

1 Measure

- Capture *high resolution temperature data*, *air flow data* and *infrastructure & layout data*

2 Model



- To identify improvement opportunities *model the data center* and use optimization algorithms (“*best practices rules*”)

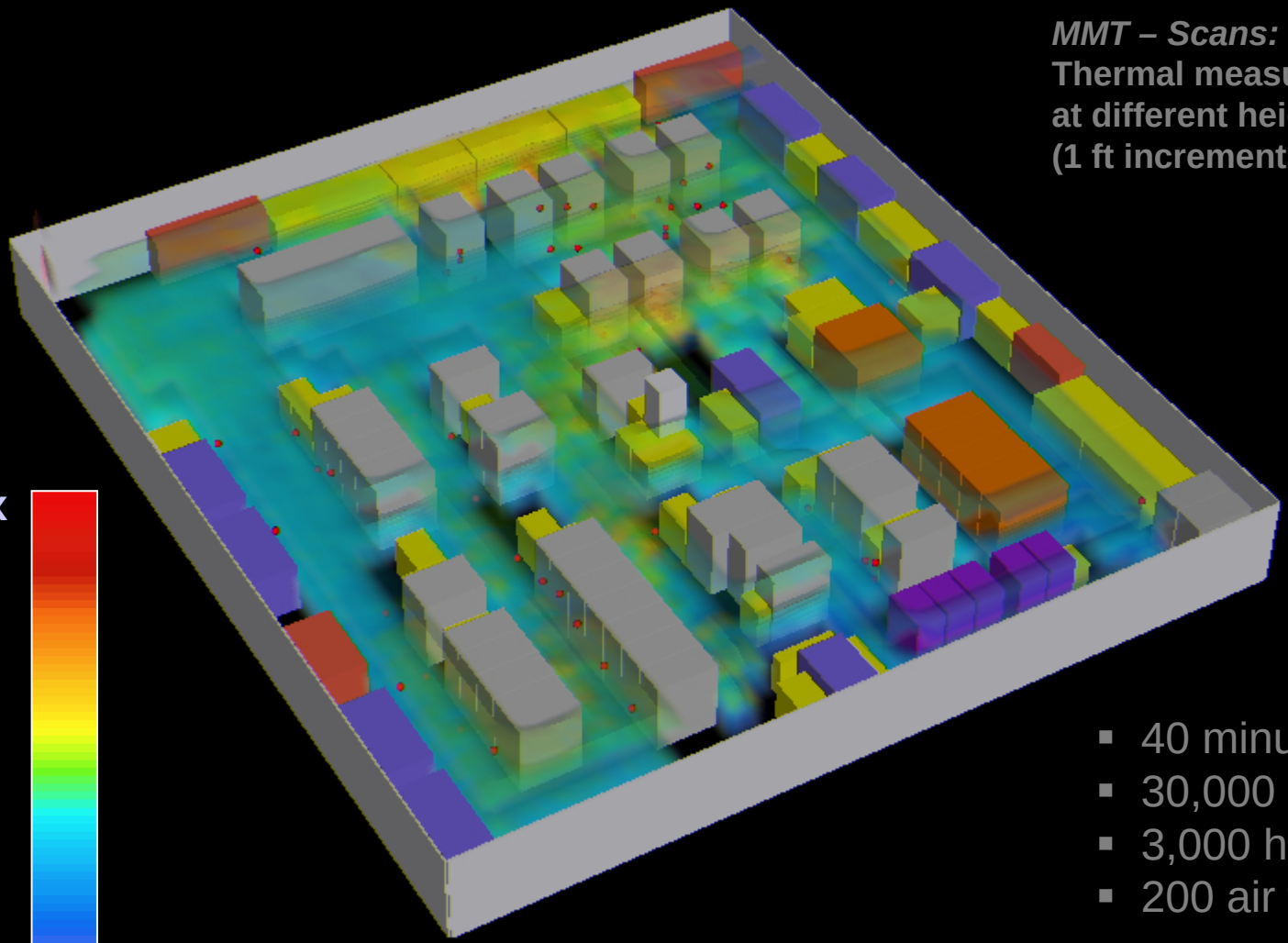
3 Manage “*Best Practices*”

- Realize air transport energy savings
- Realize thermodynamic energy savings

→ *Achieve reduced energy consumption*

→ *Potential for deferring new investments*

MMT 3D Temperature Distributions

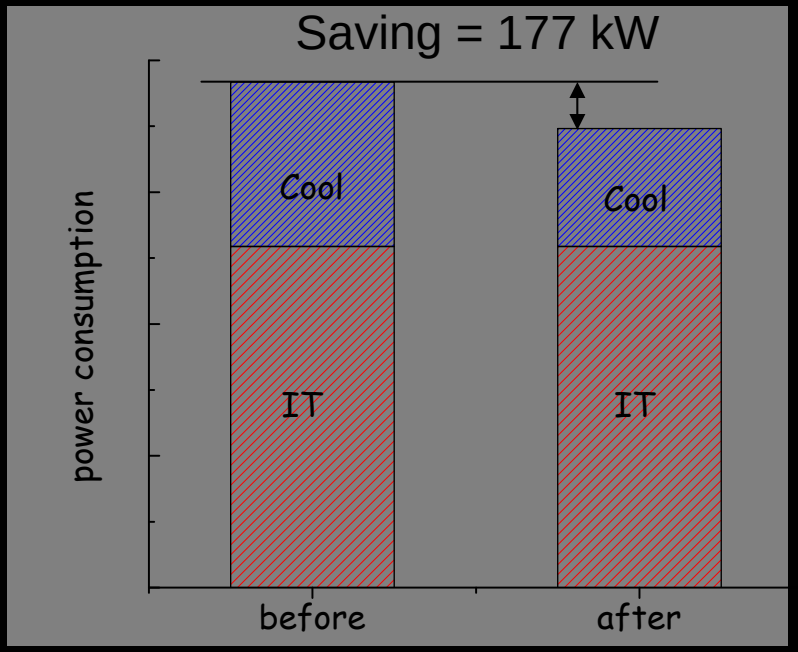


MMT – Scans:
Thermal measurements at different heights (1 ft increments in z)

- 40 minutes scan time
- 30,000 thermal readings
- 3,000 humidity readings
- 200 air flow readings

Example Energy Savings

- saved 177 kW with measurement / metrics driven best practices implementation
- typical 1-2 Month turnaround to realize savings
- improved DC COP 2.39 to 3.44
  - COP_{thermo} from 4.5 to 5.1
  - COP_{trans} from 5.3 to 9.8
- **total energy savings ~ $200k (ROI: 7 months)**



Finding / Metrics	Key Action / Solution
Horizontal hotspots (HH)	**change tile layout & deploy high throughput tiles**
Vertical hotspots (VH)	**snorkels / fillers**
Non-targeted air flow	**close leaks / cable cutouts**
Plenum temperatures	**service ACUs supply side / increase ACU utilization**
ACU utilization	**turn under-utilized ACUs off**
ACU flow	**remove blockage**

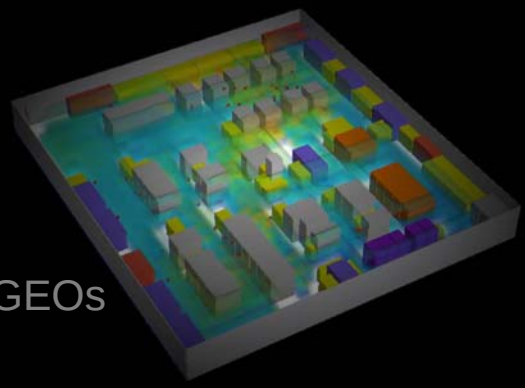
thermo

transport

Case Study: *DC Area = 20k sqf; Temp. Meas. = 200,000; Airflow Meas. = 1,200; Power density ~ 75 W / sqf*

MMT Historical Record / Scorecard

- MMT service provided to more than 50 DCs
  different sizes, power densities, locations etc.
- usually energy savings of > 10 % of IT power (< 1 y ROI)
- has delayed major DC upgrades / capital investments
- MMT is an WW IBM service offering with Research support in three GEOs



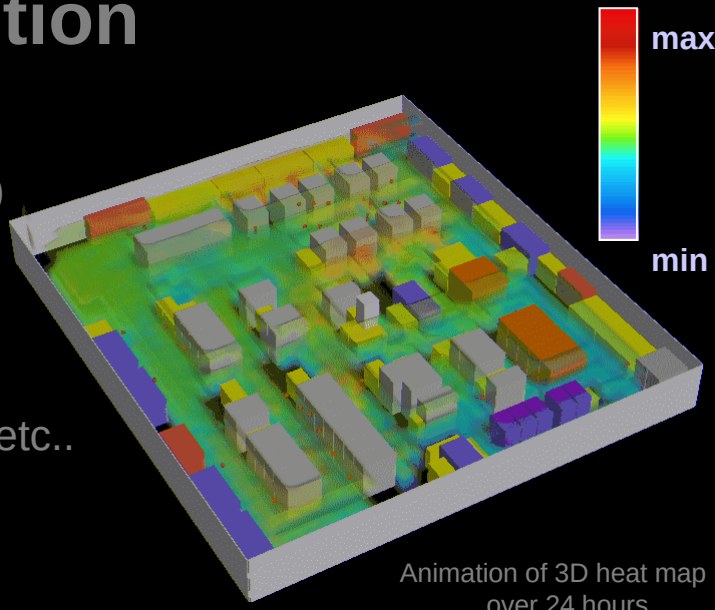
Example – MMT Savings

	Survey Date	RF power P_{RF} [kW]	DC area [feet^2]	% Savings
DC1	05/07	1400	42k	**13**
DC2	06/07	2316	84k	**11**
DC3	07/07	1917	55k	**9**
DC4	09/07	1822	57k	**9**
DC5	10/07	204	2k	**11**
DC6	10/07	418	10k	**12**
DC7	12/07	461	8k	**8**
DC8	12/07	1091	14k	**4**
DC9	01/08	442	11k	**10**
TOTAL				**10**

WW MMT Activities

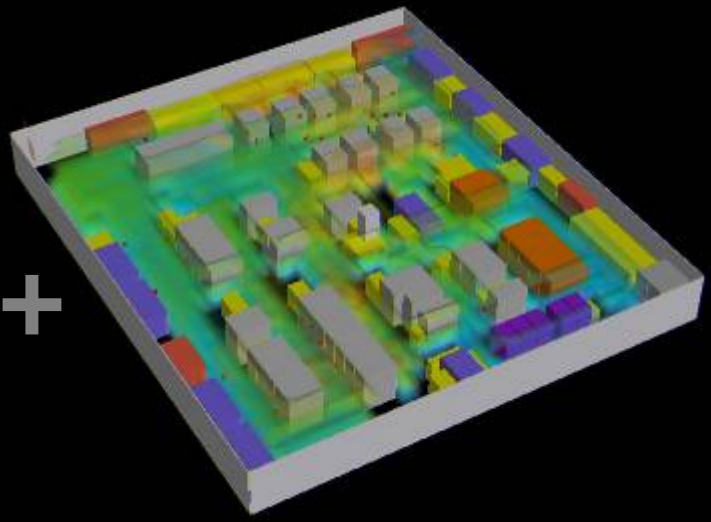
MMT 1.5: Move to a dynamic Solution

- DC can change over time
  - IT power levels can change (e.g., 10-15 % during a day)
  - cooling conditions change etc..
  - new racks / new servers / re-arrangement of tiles etc..
- MMT 1.5 provides high resolution combining
  - MMT 1.0 for **base model** generation, sensor placement etc..
  - real-time sensors for creating **dynamic models**



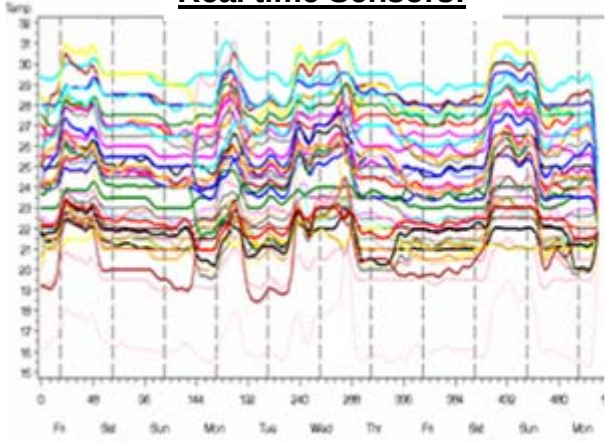
Animation of 3D heat map over 24 hours

+

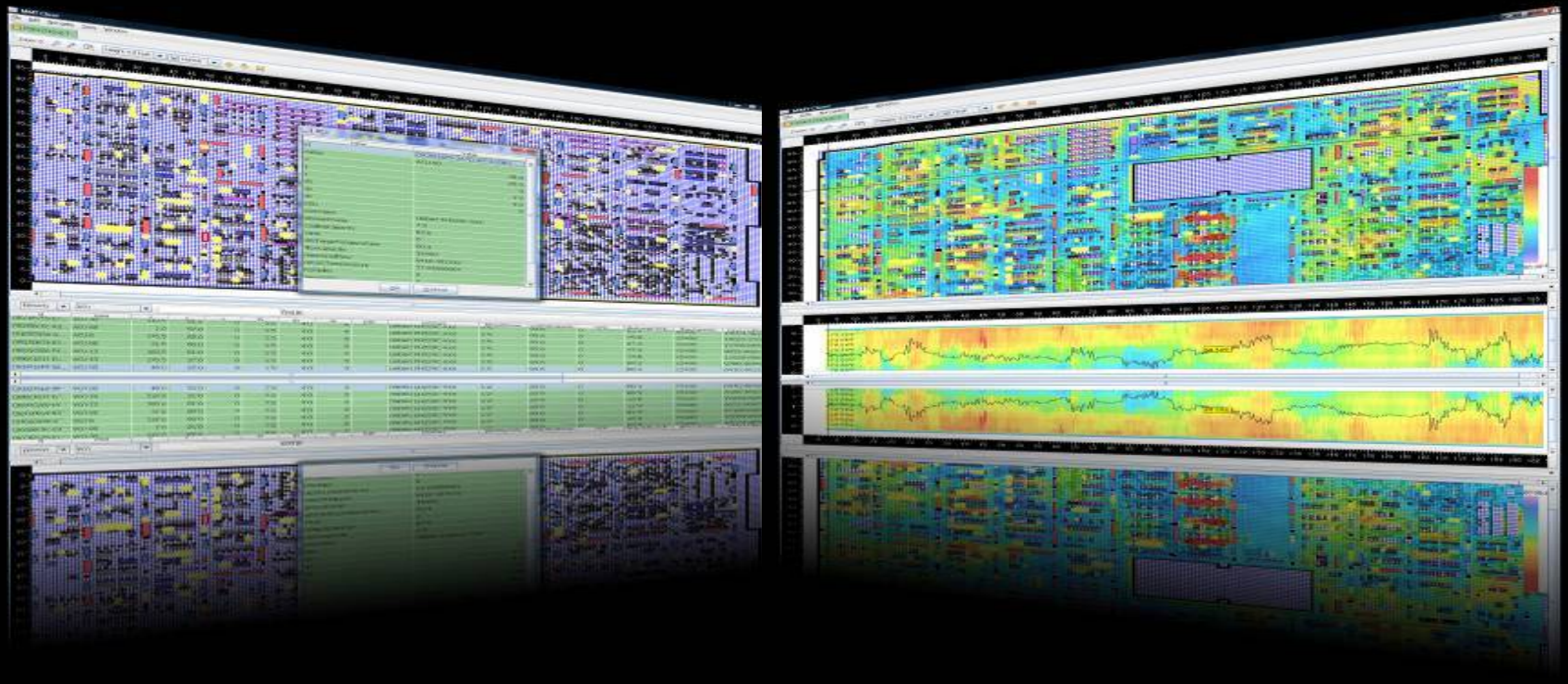


+

Real time Sensors:



MMT 1.5 Client

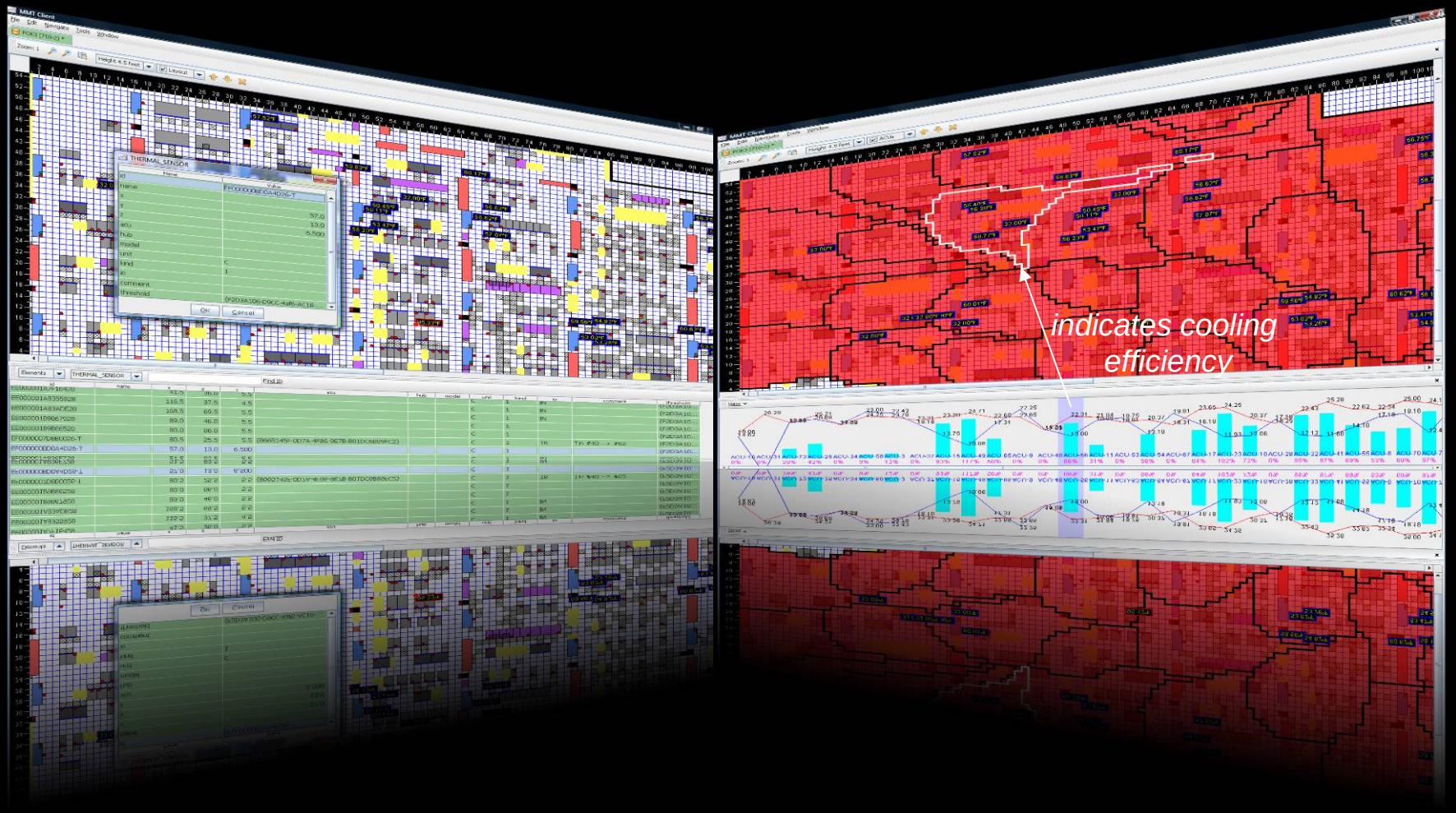


FEATURES:

- efficiency of each cooling zone in real-time
- data analysis capabilities
- energy efficiency reporting
- detailed layout editor
- alarm services / hotspot services
- real-time temperatures / 3D capabilities
- current being integrated into Tivoli and Maximo

https://researcher.ibm.com/mmt2/launch.htm

MMT 1.5 Client - continued



- provides alarm/threshold settings for sensors
- determines fan efficiencies in real-time and provides corresponding cooling zones
- provides weekly reports of all logged temperatures

MMT IBM Internal Deployment

MMT 1.0 successfully WW deployed in > 1 M square feet of DC space

- ~2MW of savings so far
  - $1.5M savings in utility bill savings
  - > $10M delayed capital cost
- over 150 ACUs turned off / decommissioned
- ~ 8-9 % PUE improvements

MMT 1.5 currently being rolled out WW

- active management large-scale dynamic DC
- large scale deployment in EMEA

MMT 1.5 part of leadership DC design and architecture

- full deployment in leadership DC in Raleigh
- integrated into ITD data models & asset management systems