Energy and Thermal Management of Chips, Systems and Data Centers necessitates a return to fundamentals

Return to the way we were

Chandrakant D. Patel

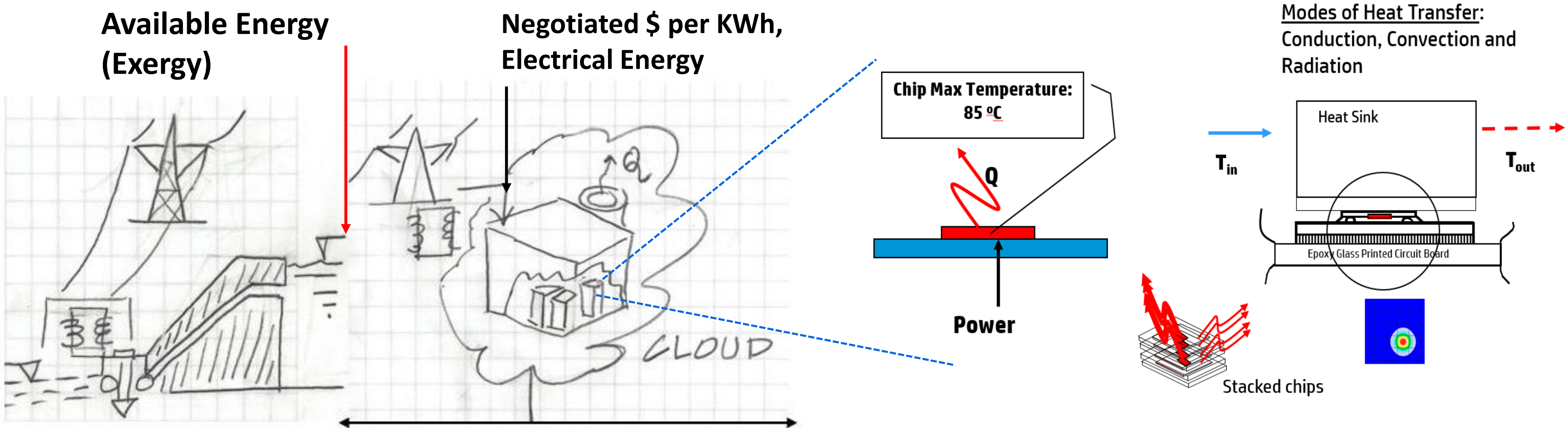
HP Inc. Senior Fellow

IEEE EDPS October 2024

Objective

A holistic* systemic fundamentals-based approach to energy and thermal management required in order to

- *Maximize the performance of chips and technologies*
- *Focus on available energy as opposed to pricing of electricity, the latter a fleeting proposition in the Age of AI*



Available Energy (Exergy)

Negotiated $ per KWh, Electrical Energy

Chip Max Temperature: 85 ºC

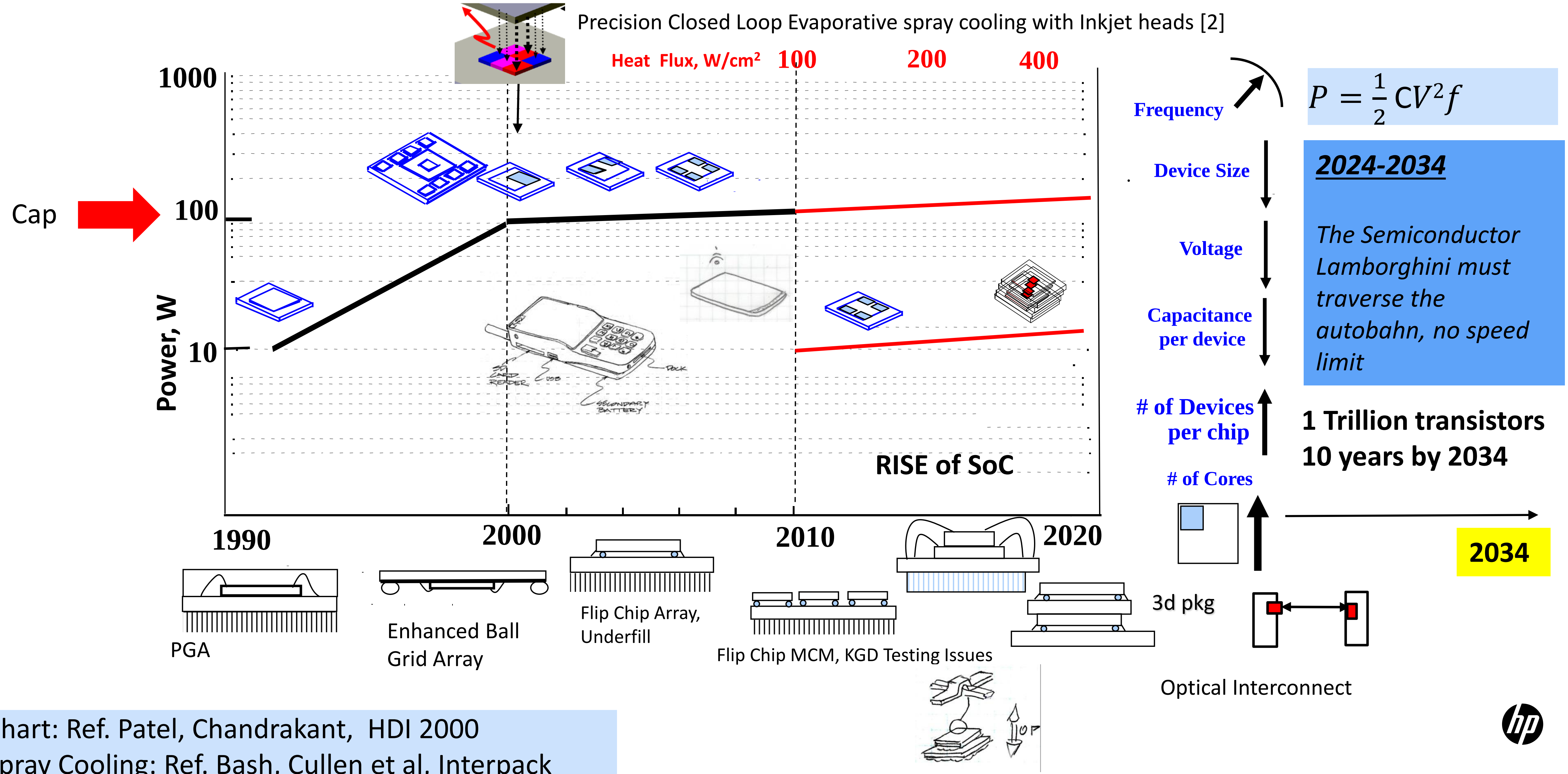
Q

Power

Modes of Heat Transfer: Conduction, Convection and Radiation

Cullen Bash, Jessica Bian, Dejan Milojicic, Chandrakant D. Patel, Luka Strezoski, Vladimir Terzija, and Dejan Milojicic. “Energy Supplies for Future Data Centers" Computer 57, 7 (July 2024), 126–134. https://doi.org/10.1109/MC.2024.3393248

History of Chip Heat Dissipation: Power Capping no longer an option, MCMs Return



$$P = \frac{1}{2} CV^2 f$$

2024-2034

The Semiconductor Lamborghini must traverse the autobahn, no speed limit

1 Trillion transistors 10 years by 2034

Chart: Ref. Patel, Chandrakant, HDI 2000
Spray Cooling: Ref. Bash, Cullen et al, Interpack

Units

SI Units

- Power, or rate of doing work, is measured in Joules per second or Watts
  - KW, MW, GW, TW

- Energy, in Joules, product of power and time
  - Energy use at home, office, etc is represented in KWh

- Heat dissipation is quantified in Watts

- Temperature in ^{o}C, K

- Mass Flow in kg/s

- Density in Kg/m^3

- Volume Flow in m^3/s (to facilitate conversation, we will convert to cubic feet per minute in some cases)

- Pressure in Pascals (N/m^2) (to facilitate conversation, we will convert to inches of water in some cases)

Organization

Energy & Sustainability

Overview of Energy and Thermal Management – Data Centers

Thermal Management of Chips & Systems

Framework for Devising Cooling Solutions

1. Joules of Available Energy is the Currency
2. Supply-Demand Framework Based on Available Energy
3. Tracing the Energy Flow from the Power Plant to Chip, and from Chip to the Cooling Tower
    - *Fundamentals Based Key Performance Indicator (KPI)*
4. Salient Elements of Thermal Management from Chip to Cooling Tower
5. Summary of Cooling Solutions

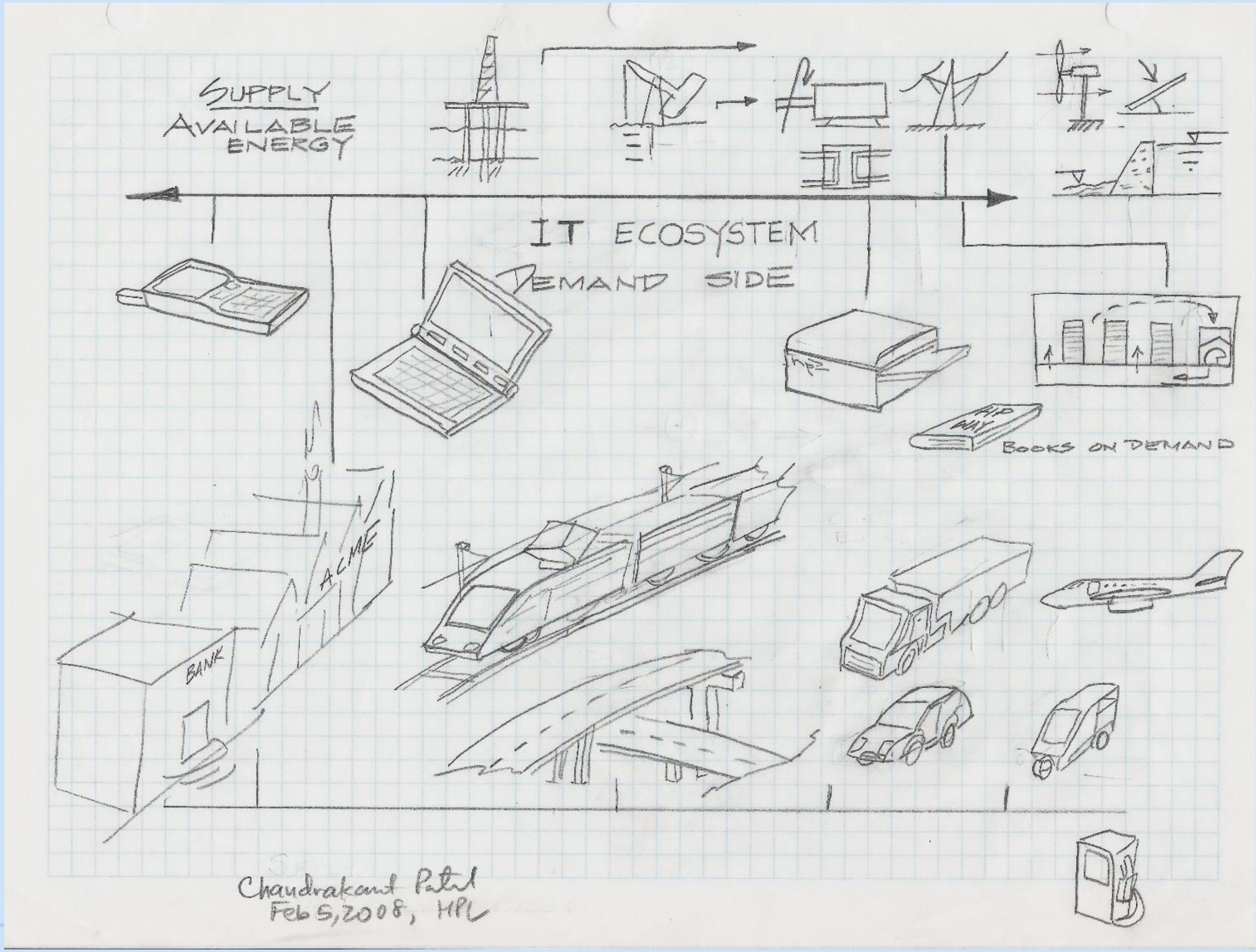
Energy: It is about Supply and Demand

Energy & Sustainability

Supply Side

Demand from IT

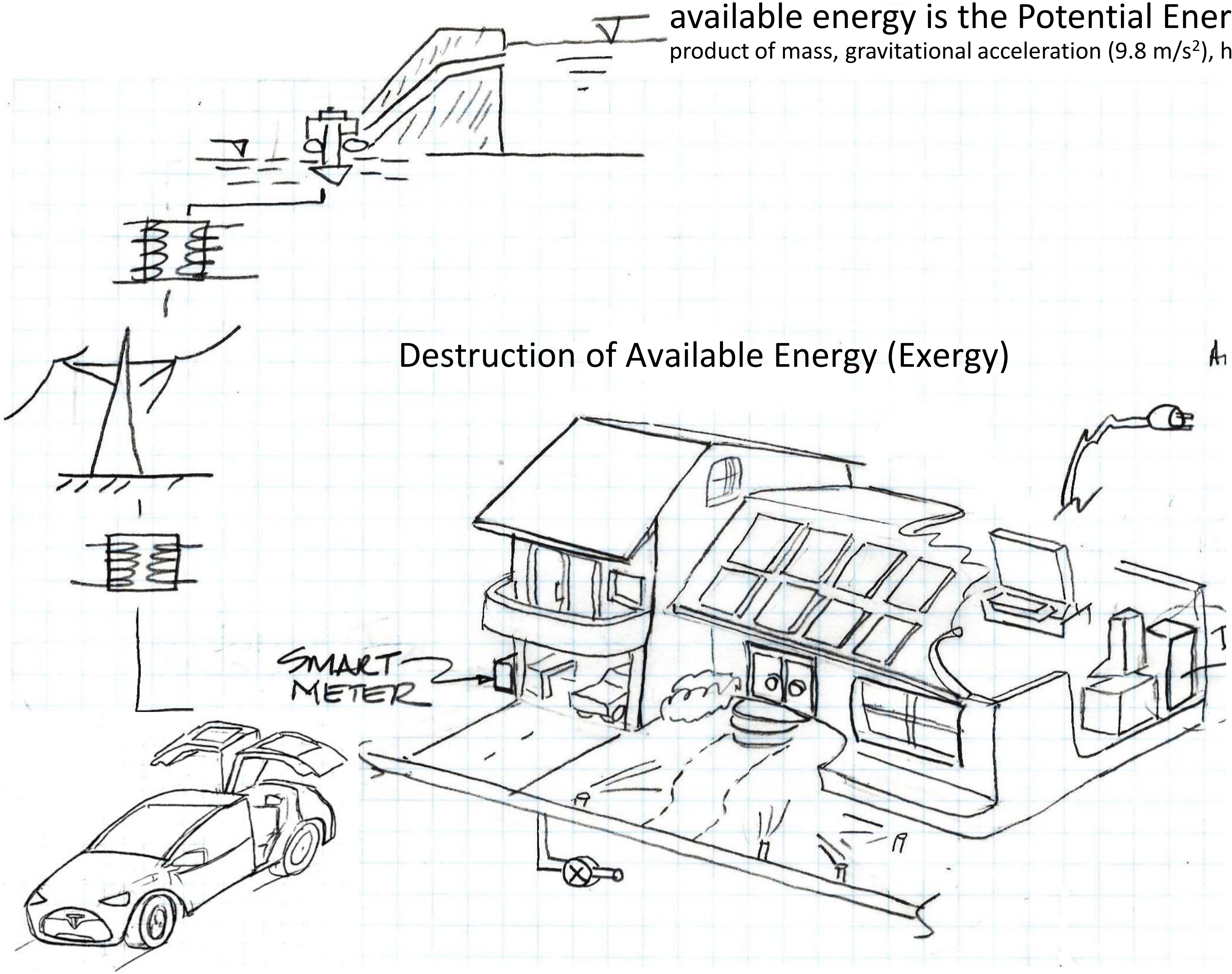
Demand from Basic Necessities



Tracing the Available Energy Flow to a Chip

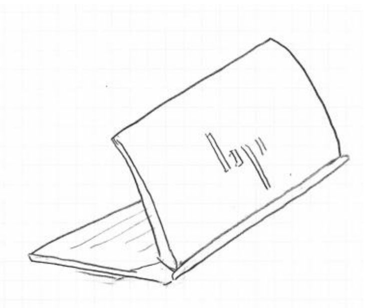
available energy is the Potential Energy PE = mgh

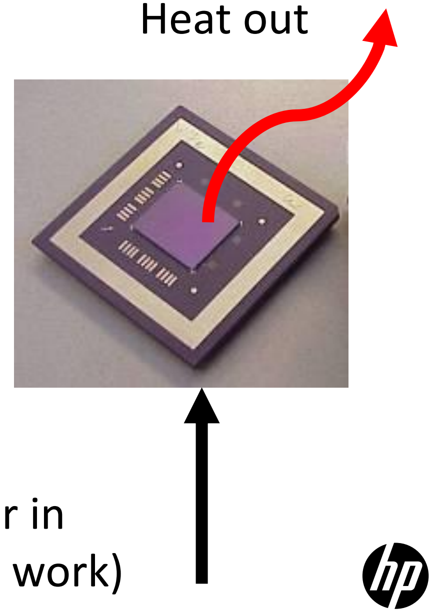
product of mass, gravitational acceleration (9.8 m/s^2), height in metres



Destruction of Available Energy (Exergy)

Conversion Losses in AC distribution, AC to DC, DC to DC conversion to various voltage levels





Heat out

Power in (pure work)

Available Energy in Waste Streams

1st Law of Thermodynamics

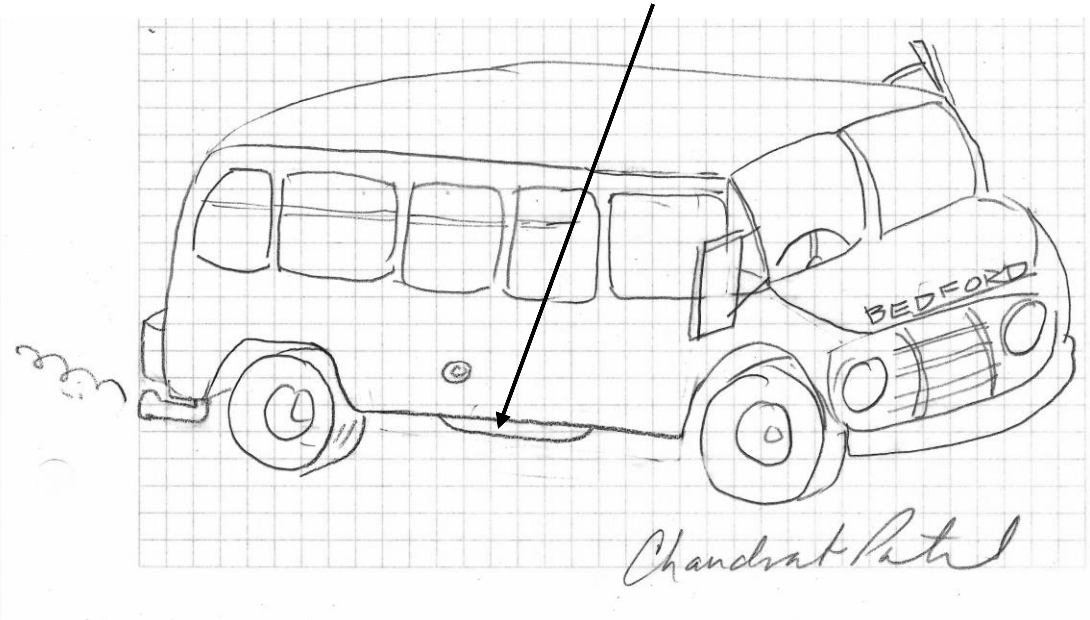
- Chemical Energy in Fuel = Kinetic Energy in Propulsion + Waste

2nd Law of Thermodynamics

- states that available energy or exergy (quality) is destroyed
  - conversion due to thermodynamic irreversibility (waste heat

- 1 Joule at 50 °C, with ref to 20 °C ambient, has 0.1 Joule available (exhaust hot air from a computer is low grade)
- 1 Joule at 500 °C, with ref to 20 °C ambient, has 0.6 Joules available [High Grade]

Diesel 45 MJ/kg



$$A_{waste\ heat} = \left(1 - \frac{T_a}{T_j}\right) Q$$

$$A_{waste} = \left(1 - \frac{T_a}{T_j}\right) Q = \left(1 - \frac{293}{773}\right) Q = 0.6Q$$

A, Available Energy, in Joules
Q, Heat energy in combustible waste products
T_a, Ambient Temperature or Cold Reservoir temperature, in Kelvins
T_j, Temperature of the hot gas from the exhaust, in Kelvins

Sustainability Framework based on Available Energy

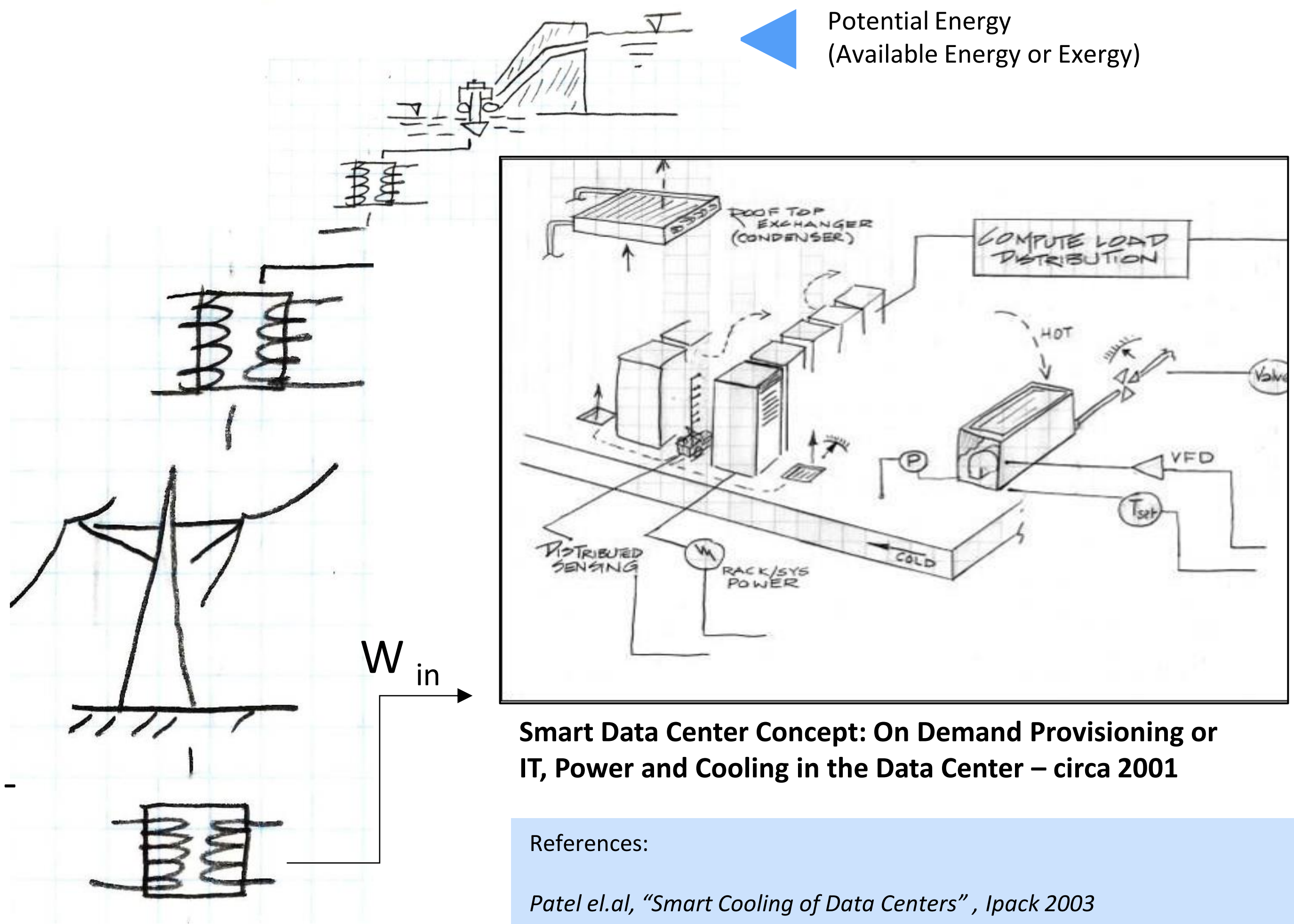
Integrated Supply-Demand Management

Supply Side

- Cradle-to-cradle available energy required (Joules) for extraction, manufacturing, waste mitigation (negative externalities), operation and reclamation
- utilize local sources of available energy to minimize destruction of available energy in transmission and distribution
- examine and utilize available energy in waste streams

Demand Side:

- Provision resources based on the needs of the user (on-demand)*
- sensing, communications, knowledge discovery, and policy-based control



Potential Energy
(Available Energy or Exergy)

Smart Data Center Concept: On Demand Provisioning or IT, Power and Cooling in the Data Center – circa 2001

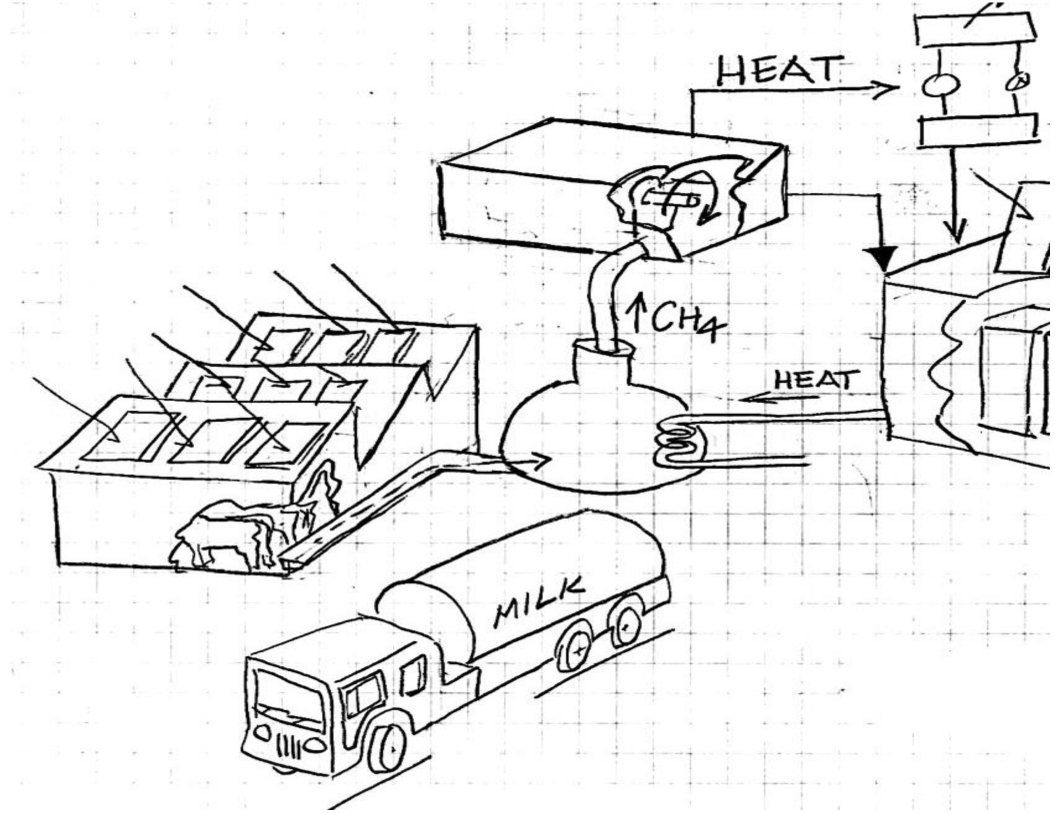
References:

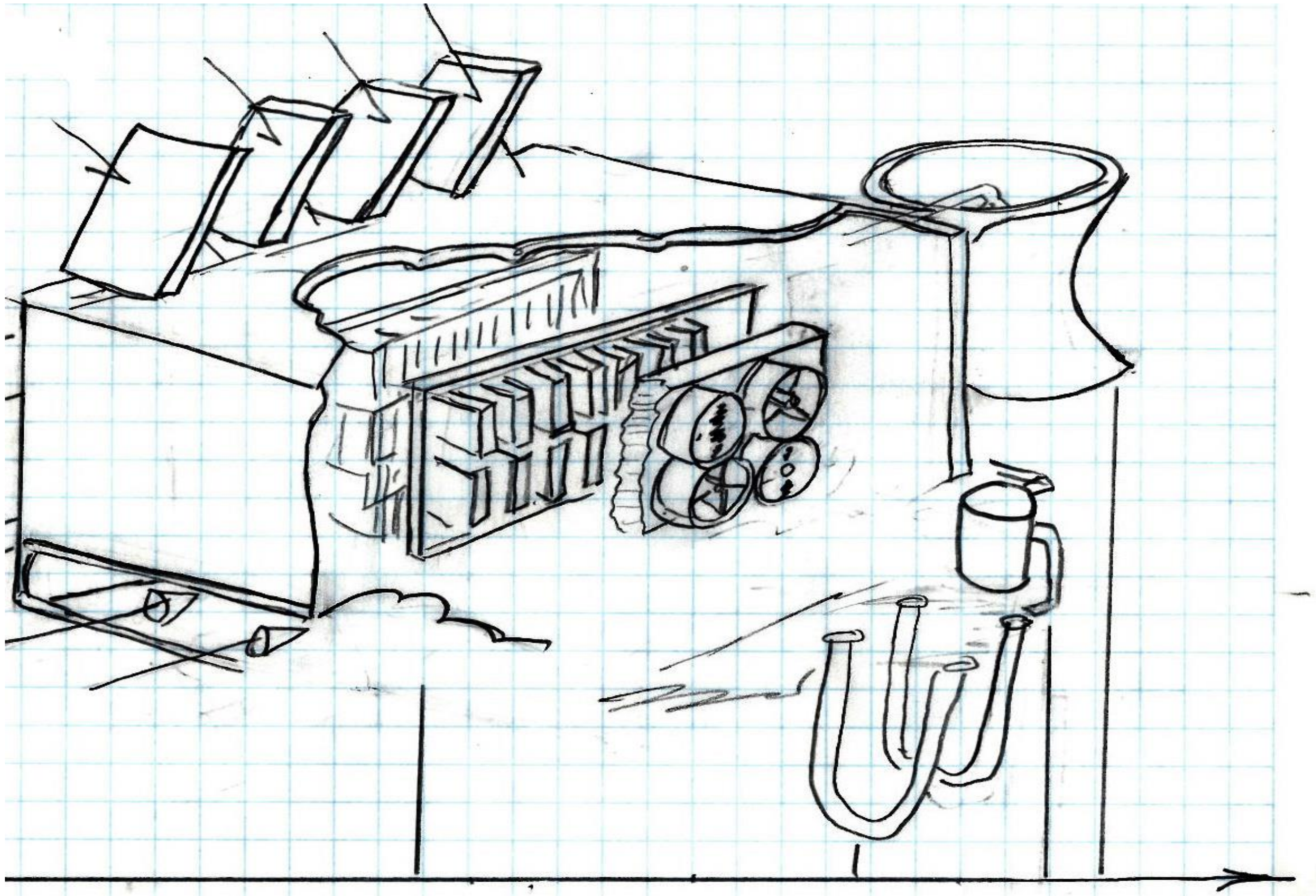
Patel el.al, "Smart Cooling of Data Centers" , Ipack 2003
Beitelmal and Patel, "Thermo-Fluids Provisioning of Data Centers, Journal of Dist and Parallel Databases,
Bash et. al, "Dynamic Thermal Management", Itherm 2006
Shah and Patel, "Data Center Total Cost of Ownership Model
Sharma et. al.,

Data Center with Supply Side Power Grid and Cooling Grid

Integrated Supply and Demand Management given the service level objectives

- **Local Power Grid** – Biogas (manure from dairy cows, sun)
- **Cooling Grid** – Ground Loop, Outside Air
- **Dynamic Managemen**t of Demand Given Supply & Service Level Agreement





Banerjee, P, Patel, C, Bash E., Shah, A. Arlitt M., “Towards a net-zero data centerJETC, (2012): 27:1-27:39.

Sharma, R., Christian, T., Arlitt, M., Bash, C., and Patel, C. Design of Farm Waste-Driven Supply Side Infrastructure. *4th ASME International Conference on Energy Sustainability*, 2010 Phoenix, AZ (2010)

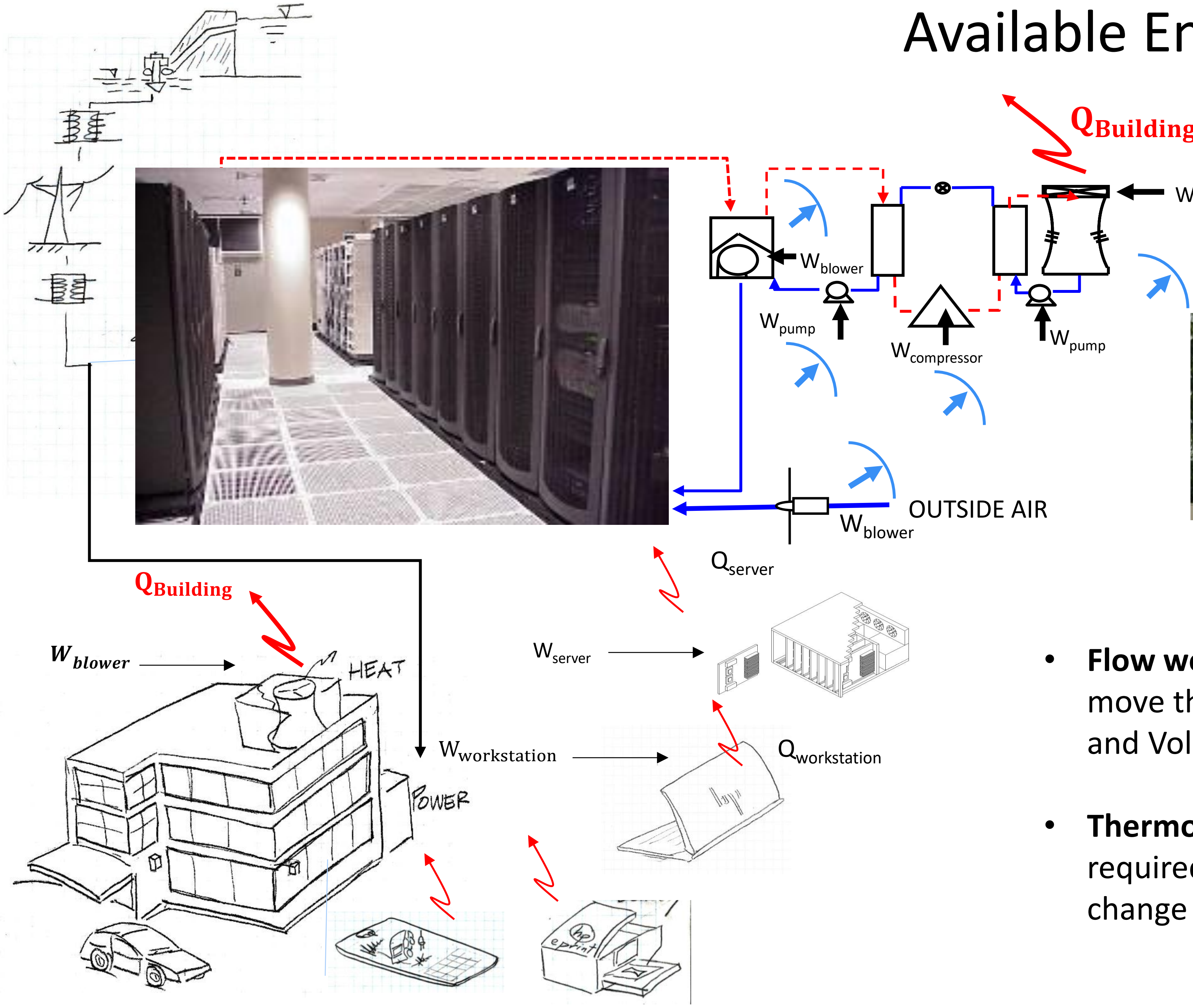
Media Coverage: Search “cow powered data centers”– NY Times, Wired, Los Angeles Times, 2010

Energy and Thermal Management necessitates a Holistic Perspective at the Data Center Level

Overview of Energy and Thermal Management

- Tracing the heat flow, and work required, from chip to cooling tower
- Introduce a Holistic Key Performance Indicator

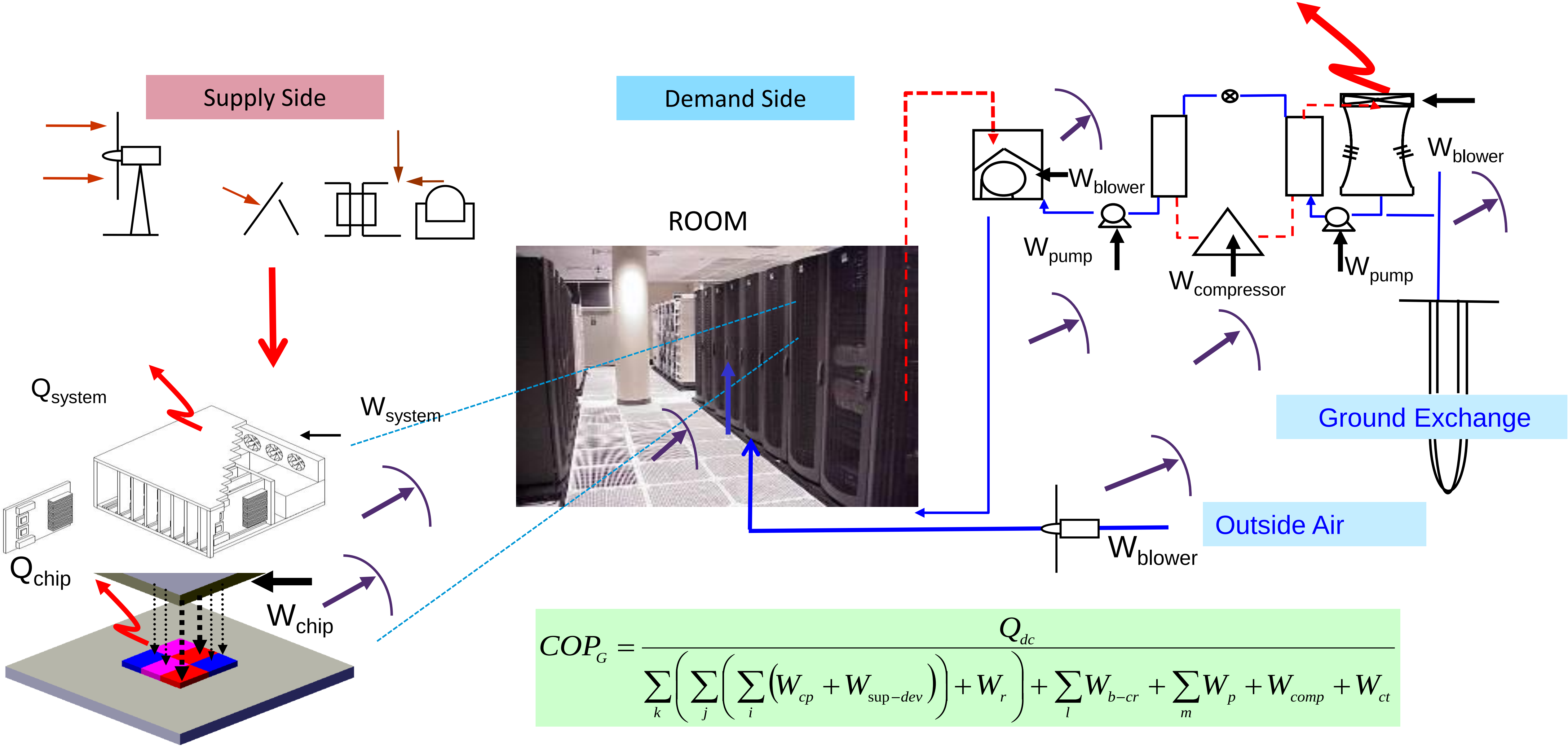
Available Energy Flow in the IT Stack



In removing Heat:
- Flow Work
- Thermodynamic Work

- **Flow work** (J/s or Watts) is the work required to move the fluid. Product of Pressure Drop (Pa) and Volume flow (m^3/s)
- **Thermodynamic work** (J/s or Watts) is the work required to reduce the temperature e.g. phase change

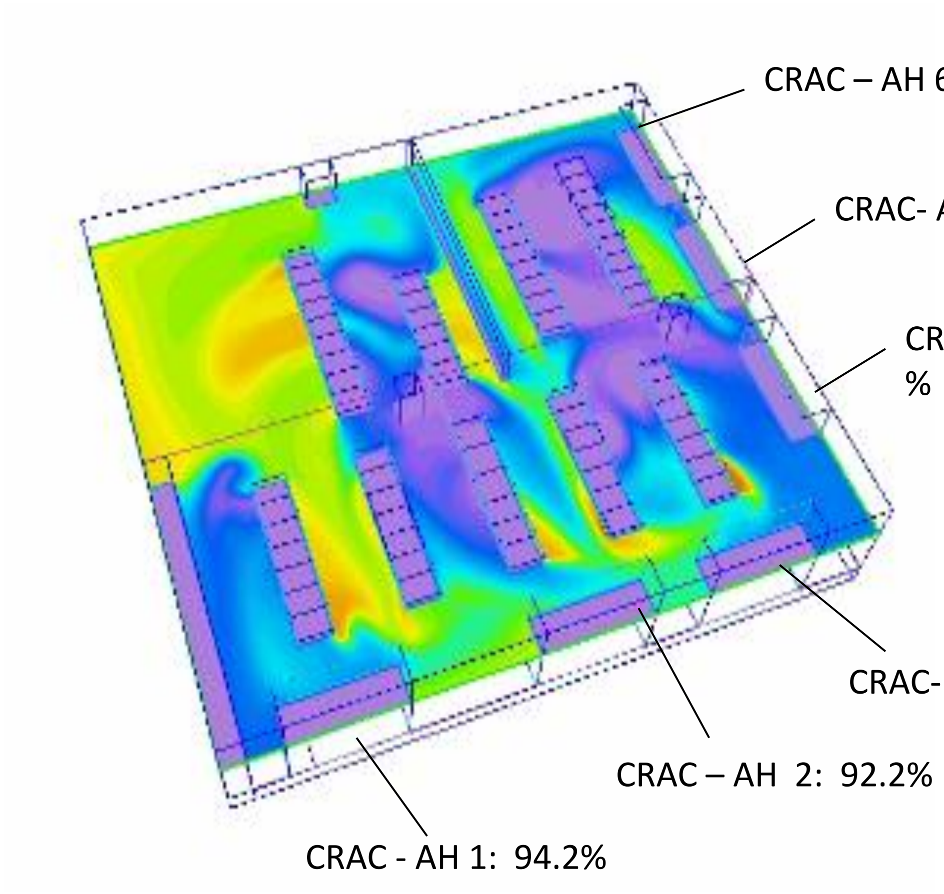
A Dimensionless KPI: Coefficient of Performance of the Ensemble



$$COP_G = \frac{Q_{dc}}{\sum_k \left(\sum_j \left(\sum_i \left(W_{cp} + W_{\sup-dev} \right) \right) + W_r \right) + \sum_l W_{b-cr} + \sum_m W_p + W_{comp} + W_{ct}}$$

Reference: Patel, C.D., Sharma, R.K., Bash, C.E., Beitelmal, M, “Energy Flow in the Information Technology Stack: Introducing the Coefficient of Performance of the Ensemble”, ASME International Mechanical Engineering Congress & Exposition, November 5-10, 2006, Chicago, Illinois

Thermo-Fluids Provisioning in the Room to minimize the work required



CRAC – AH 6: 27.7%

CRAC- AH 5: 36.7%

CRAC - AH 4: 38.8 %

CRAC- AH 3: 80.3%

CRAC – AH 2: 92.2%

CRAC - AH 1: 94.2%

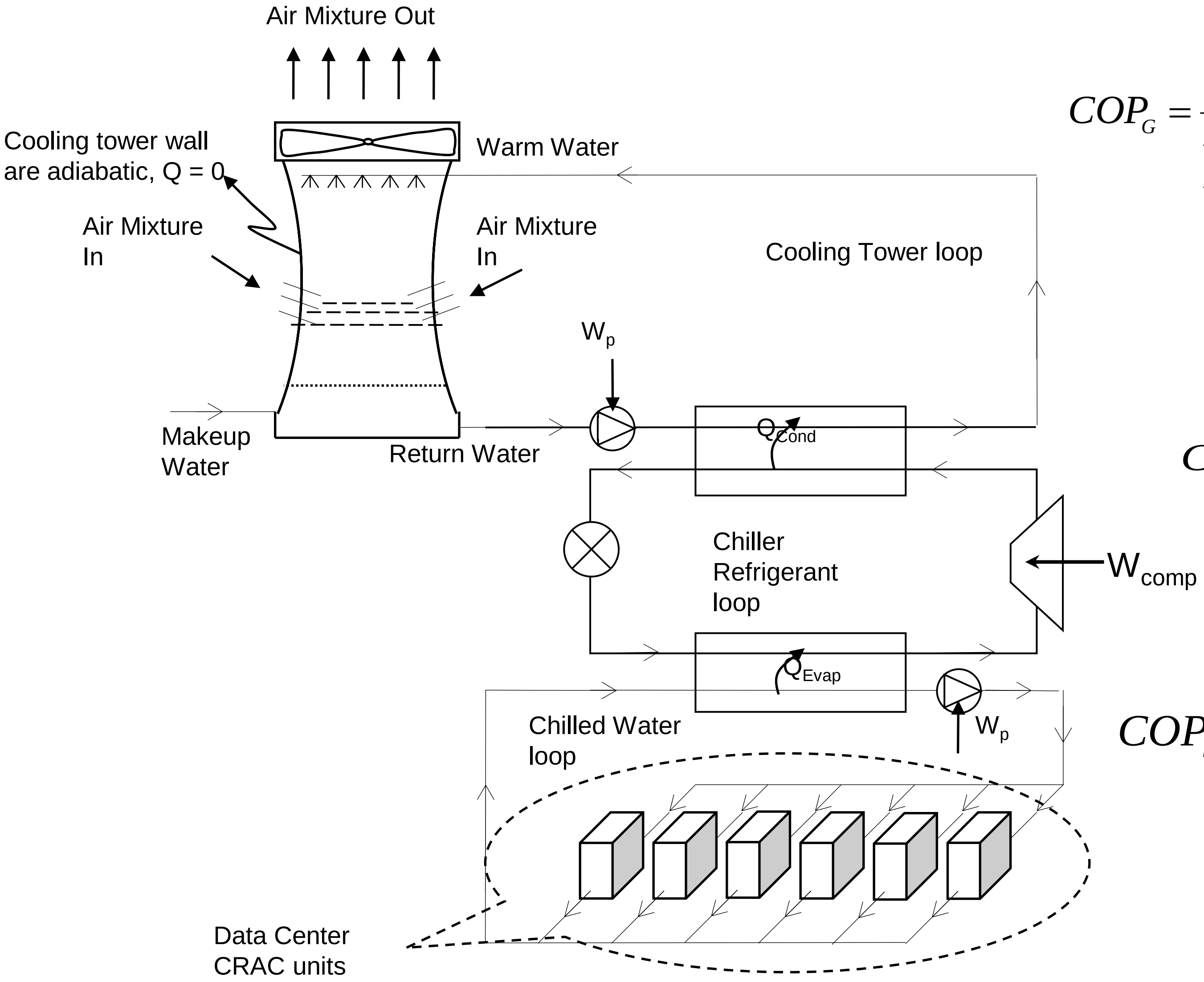
- Air Flow Distribution
- Temperature Distribution
- Computer Room Air Conditioning (CRAC) provisioning
- Energy Optimization in the Room
- Balance of Power through Workload Distribution to Drive Dynamic Thermal Management of Data Centers [2005]

Sharma, R. K., Bash. C, Patel, C.D, Friedrich R., Chase, J. “Balance of Power: Dynamic Thermal Management for Internet Data Centers” , IEEE Internet Computing, 2005

Beitelmal, A.H., Patel, C.D. Thermo-Fluids Provisioning of a High Performance High Density Data Center. *Distrib Parallel Databases* **21**, 227–238 (2007). https://doi.org/10.1007/s10619-005-0413-0

Estimating Coefficient of Performance of the Ensemble

Demand Side



$$COP_G = \frac{Q_{dc}}{\sum_k \left(\sum_j \left(\sum_i \left(W_{cp} + W_{\sup-dev} \right) \right) + W_r \right) + \sum_l W_{b-cr} + \sum_m W_p + W_{comp} + W_{ct}}$$

$$COP_{ch} = \frac{Q_{ch}}{W_{comp}}$$

$$W_{comp} = \frac{\dot{m}_{ref}\, n P_2 v_2}{\eta_p \eta_{motor} (n-1)} \left[\left(\frac{P_3}{P_2} \right)^{(n-1)/n} - 1 \right]$$

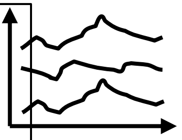
$$COP_{hydronics} = Q_{dc} \Big/ \sum \left(W_p \right)_{secondary}$$

Reference: Patel, C.D., Sharma, R.K., Bash, C.E., Beitelmal, M, “Energy Flow in the Information Technology Stack: Introducing the Coefficient of Performance of the Ensemble”, ASME International Mechanical Engineering Congress & Exposition, November 5-10, 2006,

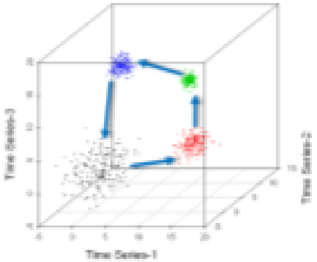
AI for Complex Systems: Temporal Data Mining of a Chiller System

pattern mining, inference and action

Multivariate time series data (r_i, t_i) (utilization, time) from 3 air cooled, 2 water cooled chillers
- *minimize electricity and water consumption in the ensemble*
- *minimize "short cycling"*



Cluster Analysis



aabaaaacaaabcccacggggaaa

single symbol sequence

Transition Encoding

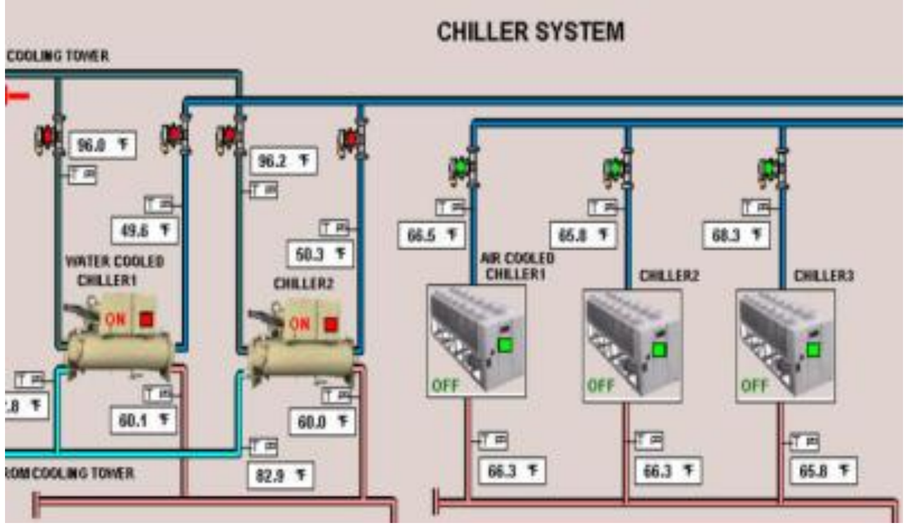
aabaaaacaaabcccacggggaaa

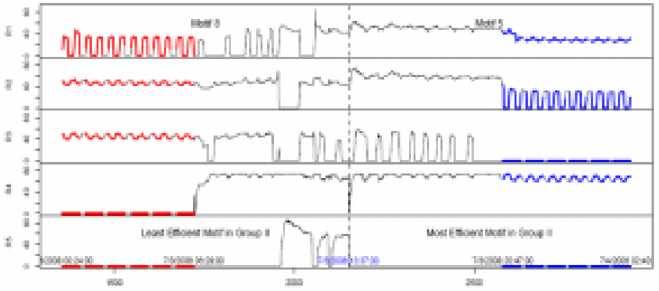
Motif Mining

example motif: ab->ac

Sustainability Characterization

22 motifs 3 months of data





Examine time sequence using physical domain knowledge e.g. COP

$$COP_G = \frac{Q_{dc}}{\sum_k\left(\sum_j\left(\sum_i\left(W_{cp}+W_{\text{sup}-dev}\right)\right)+W_r\right)+\sum_l W_{b-cr}+\sum_m W_p+W_{comp}+W_{ct}}$$

Reference: Patnaik et. al, KDD 2009, HP Smart Data Center Project, Bangalore India Production Data Center
Patel et. al, IMECE 2006, Energy Flow in the IT Stack: Introducing Coefficient of Performance of the Ensemble

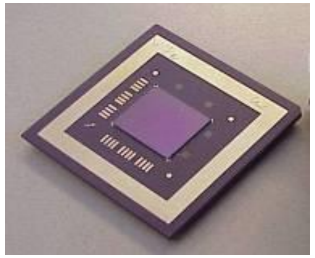
Salient Elements of Thermal Management of Chips and Systems

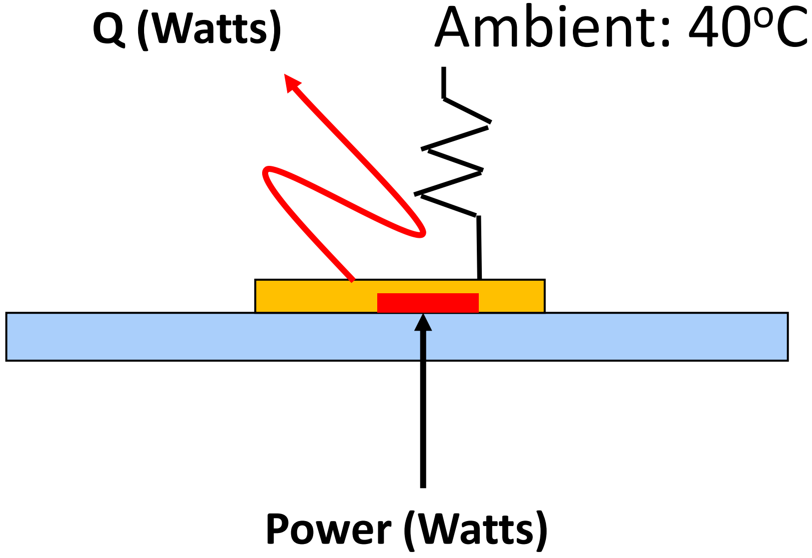
Thermal Considerations in Efficient Transfer of Heat from Chip to Ambient

1. Chip Packaging & Interconnect Considerations
2. System Considerations
   - Fluid Flow Required
   - Flow Resistance
   - Sizing the Fluid Mover
   - Fluid Mover Efficiency
3. Heat Flow from Chip to Ambient
   - Resistance to Heat Flow
   - Impact on Chip Core Temperature
   - Options for Heat Removal
4. Framework for Cooling Solutions

Problem Definition must include Chip Packaging and Interconnect

- The wall temperature of the chip, and the core of the chip, is higher than temperature of the ambient
  - Due to Thermal Resistance (° C/W) to heat flow (W)
  - Temperature of Chip Core must be kept at 85 °C to 100 °C

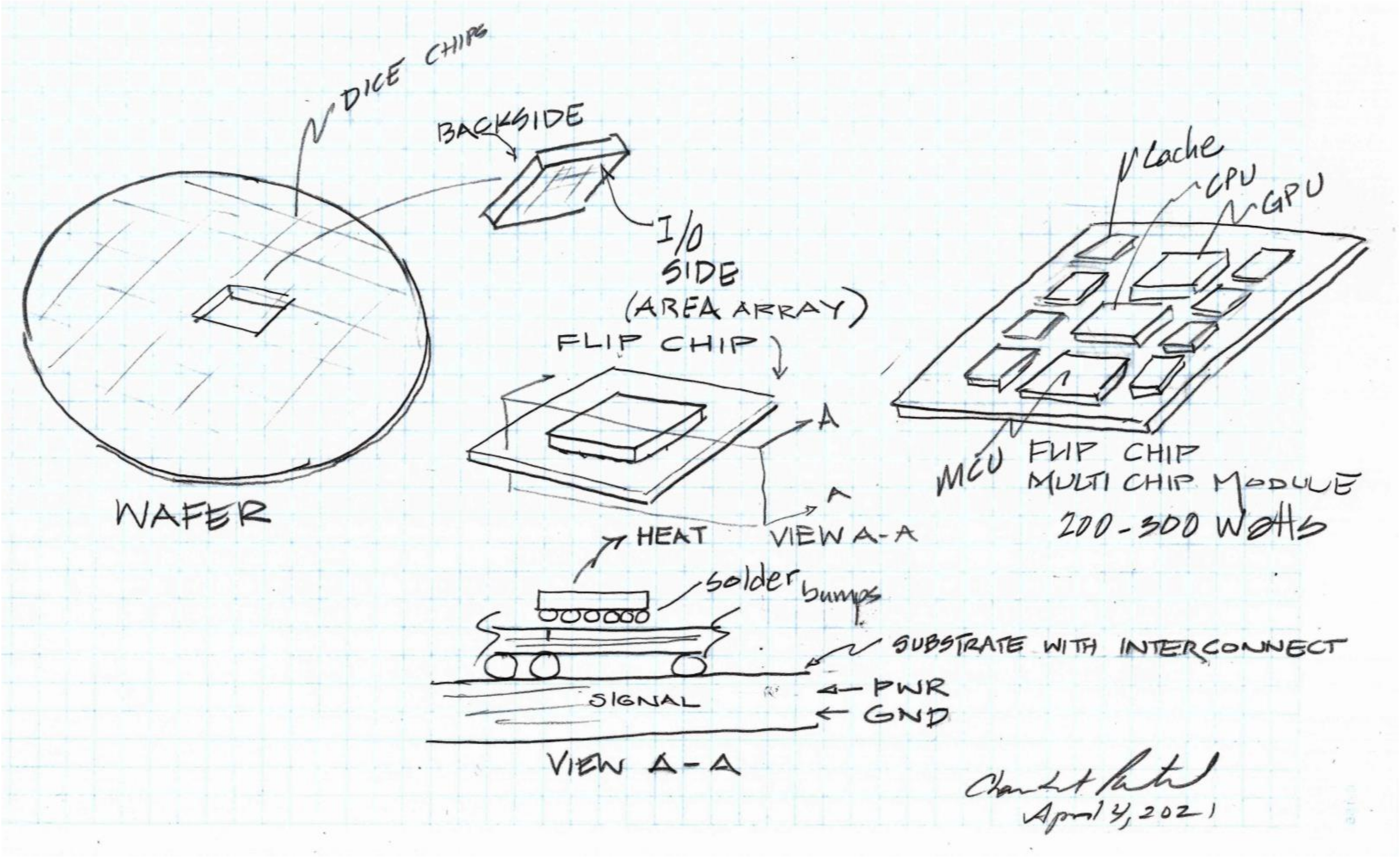




Q (Watts)

Ambient: 40°C

Power (Watts)



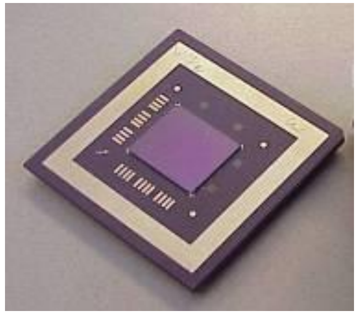
Understand the pathways to heat transfer

I/O

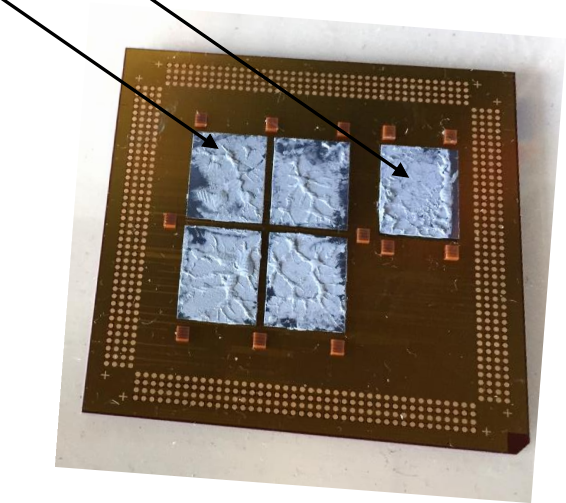
Cache
SRAMS

CPU

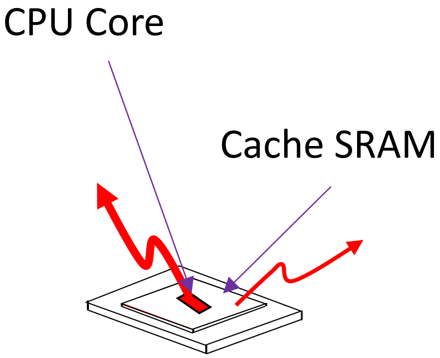
Update – IEEE Spectrum



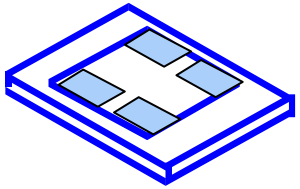
1991, 2 cm by 2 cm,
100 W
High performance
HP PA RISC
microprocessor



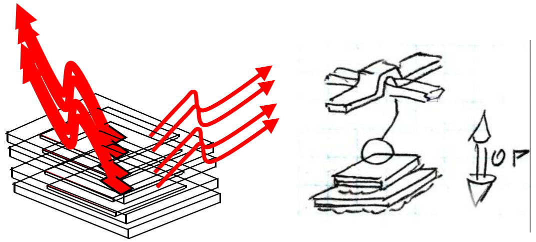
Multi-Chip Module
CPU + Cache SRAMS



Single Core
CPU with
Cache



Multi Core CPU
with Cache



- Multi Core CPU 3d stacked memory
- Cross bar memory
- Optical interconnects

System Considerations

Fluid Flow, Flow Resistance, Flow Work, Fluid Mover Efficiency

Air Flow Required to remove the heat is determined using the caloric equatio

start with a difference in temperature e.g. 15 °C

$$Q = \dot{m} C_p (T_{out} - T_{in})$$

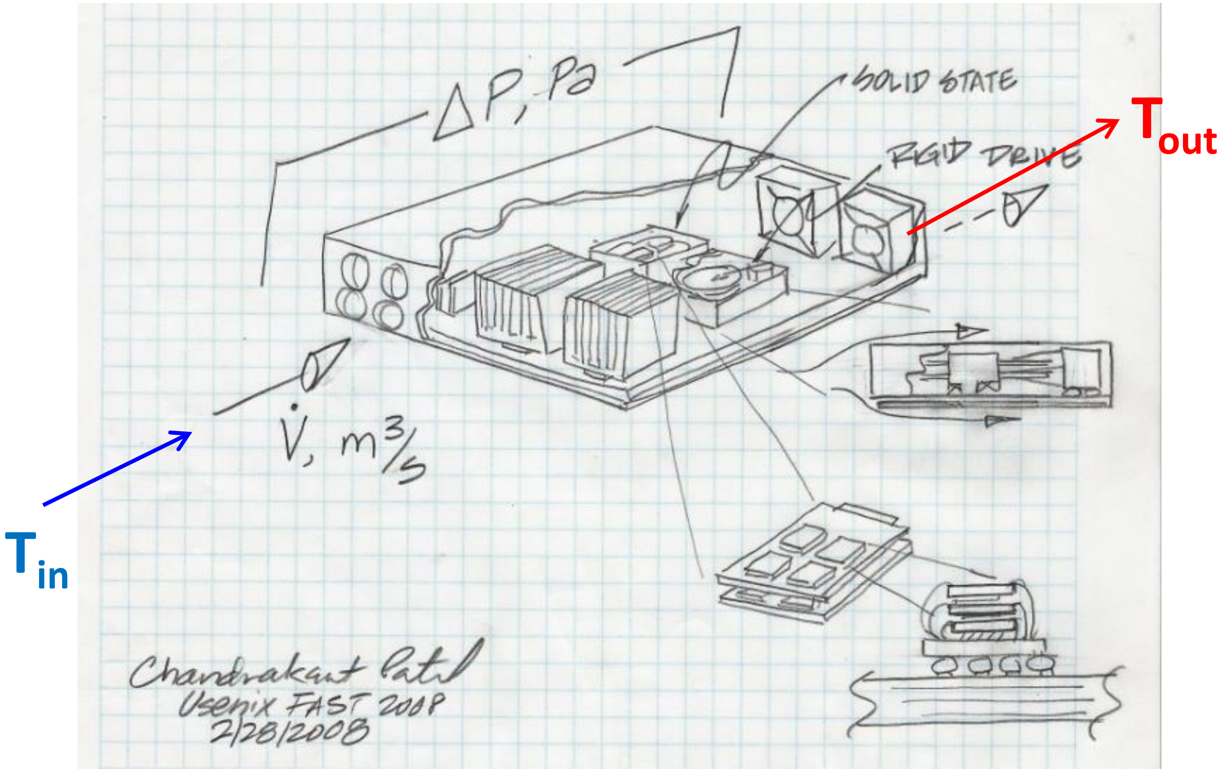
System Heat load in Watts

Kg/s
Mass flow

J/kg °C

Mass flow is product of density (kg/m³) and volume flow in m³/s

Volume flow needed in m³/s



Pressure Drop in Pascals (N/m²)

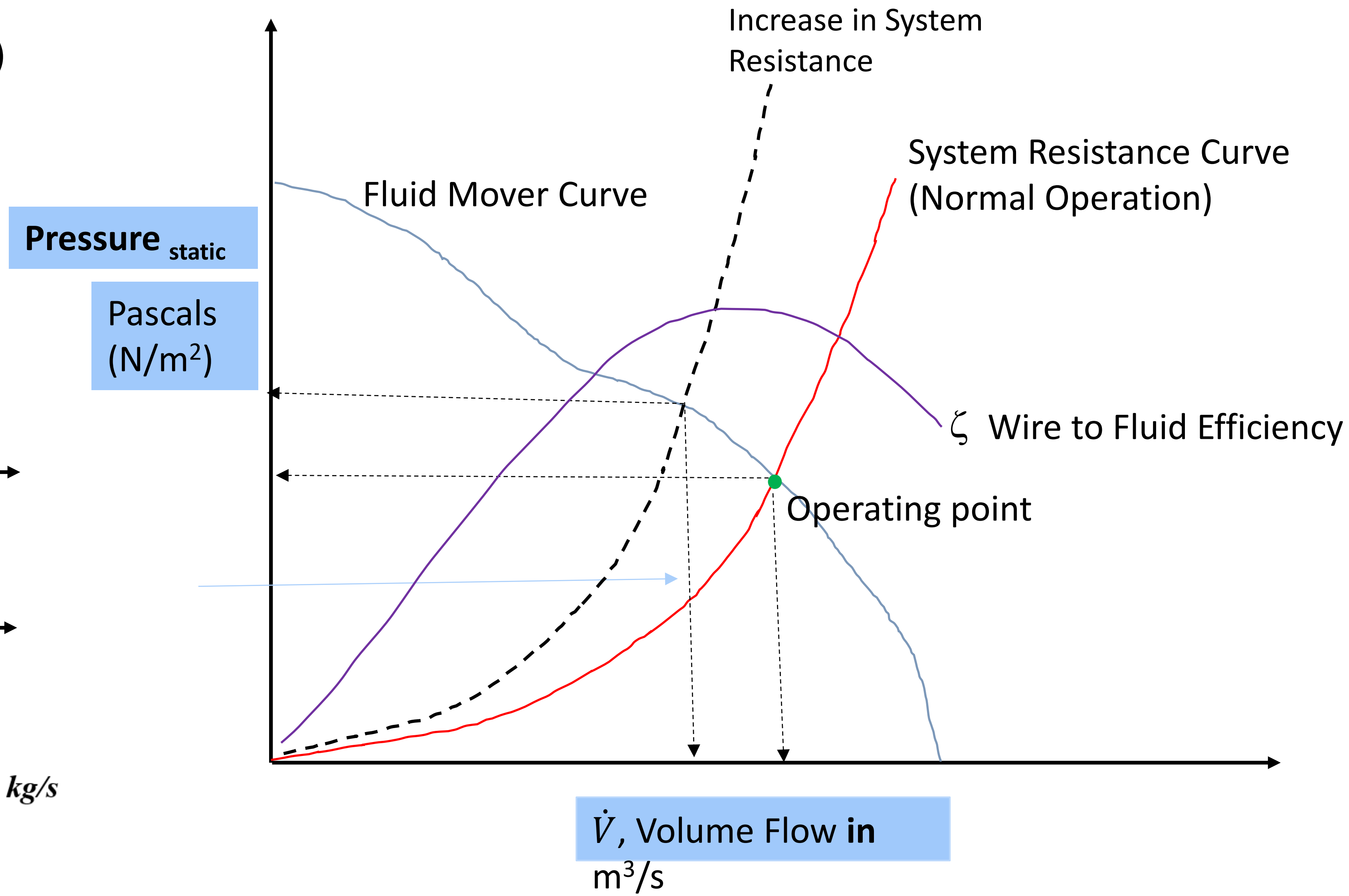
Fluid Mover Operating Point and Efficiency

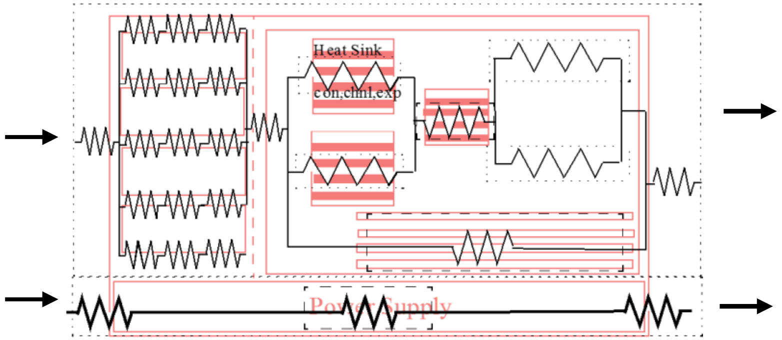
Sizing the Fluid Mover & Work Required to Operate the Fluid Mover

Flow Work (Pa. m³/s)

Fluid Mover Power (W) = PV/ζ

Wire to fluid mechanical efficiency





System Resistance Model

$\Delta P = R\,\dot{m}^2$ *where* ΔP *is in Pascals, is* $\dot{m}$ *in kg/s*

Mass flow $_{\text{in}}$ = **Mass flow** $_{\text{out}}$ assuming no leaks

Slice through a given system – First order analysis

Heat Flow from Chip to Ambient

First Order Thermal Resistance Model – Single Chip Carrier

$R = (T_1 - T_2)/Q$, °C/W

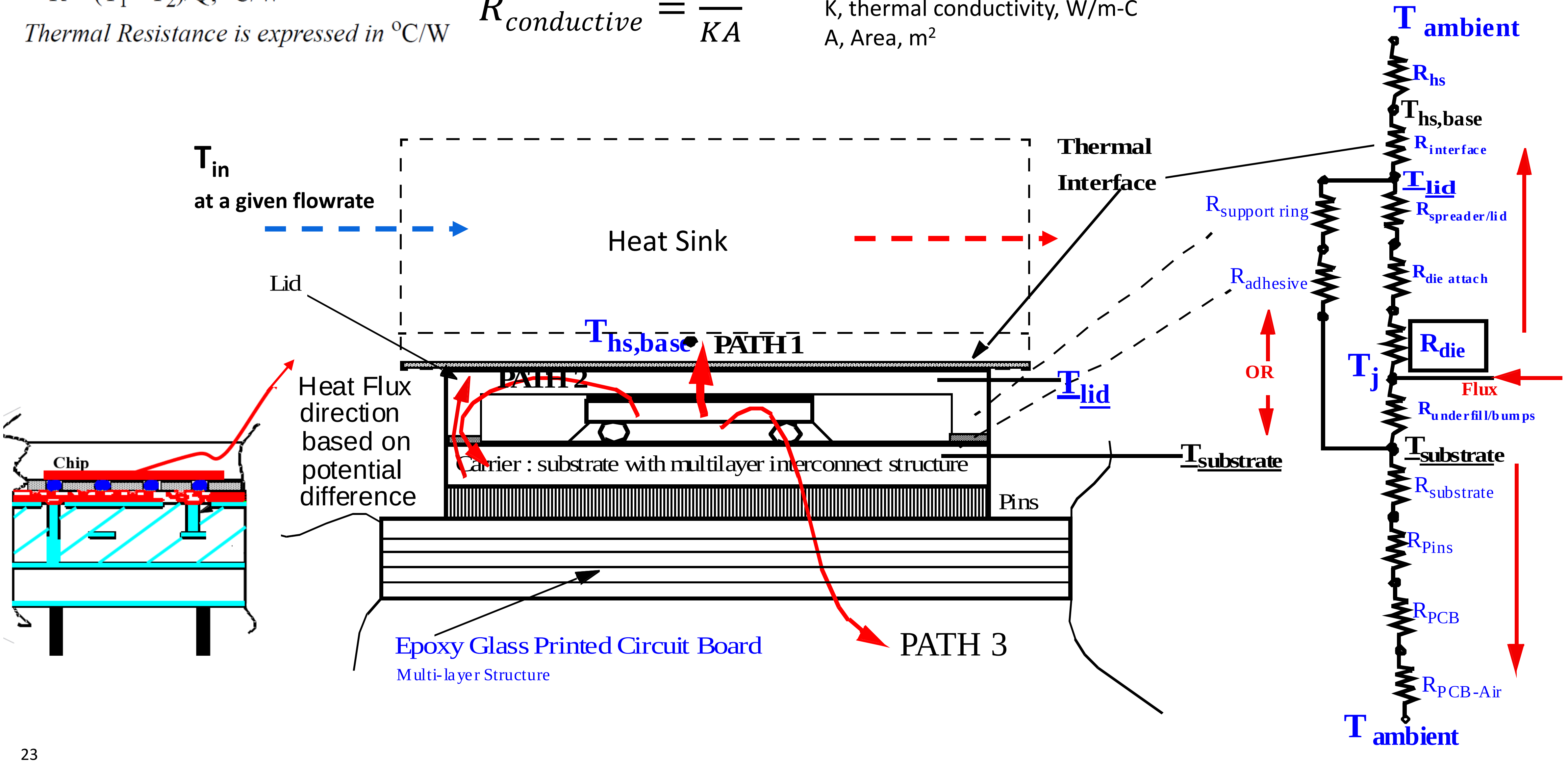
Thermal Resistance is expressed in °C/W

$$R_{conductive} = \frac{L}{KA}$$

L, thickness, in meters

K, thermal conductivity, W/m-C

A, Area, m^2

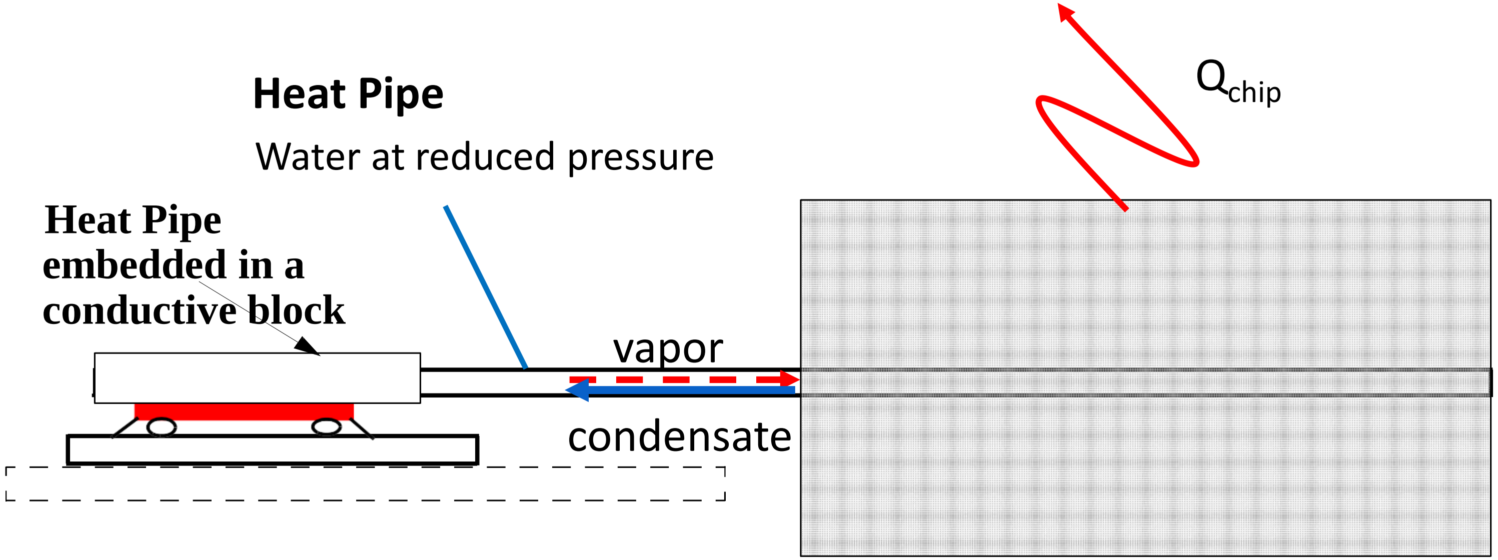


Heat Pipes for High Heat Dissipating Chips & Compact Systems Design

Capillary Pressure generated must be greater than liquid pressure drop, vapor pressure drop and gravitational pressure drop

$$\Delta P_c \geq \Delta P_l + \Delta P_v + \Delta P_g$$

Wick structure



Heat Pipe

Water at reduced pressure

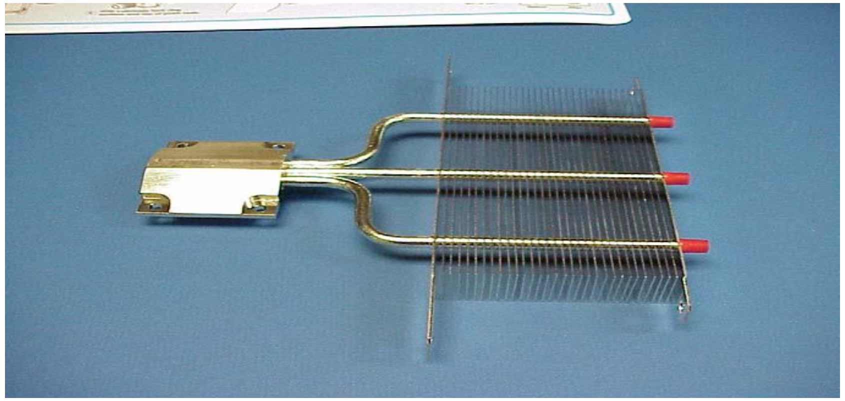
Heat Pipe embedded in a conductive block

vapor

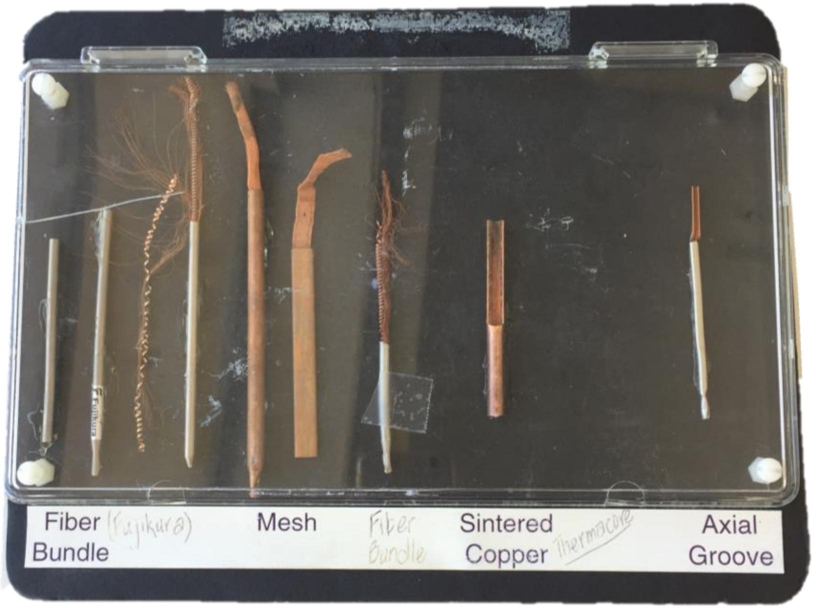
condensate

Q_{chip}

HP PA 8000 Heat Pipe System, 1994

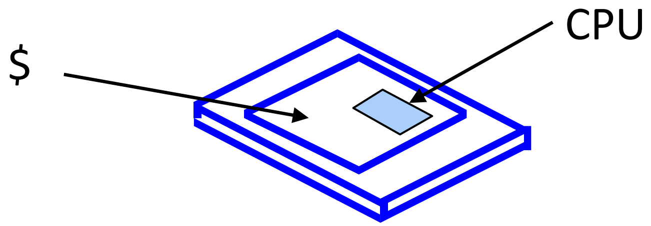


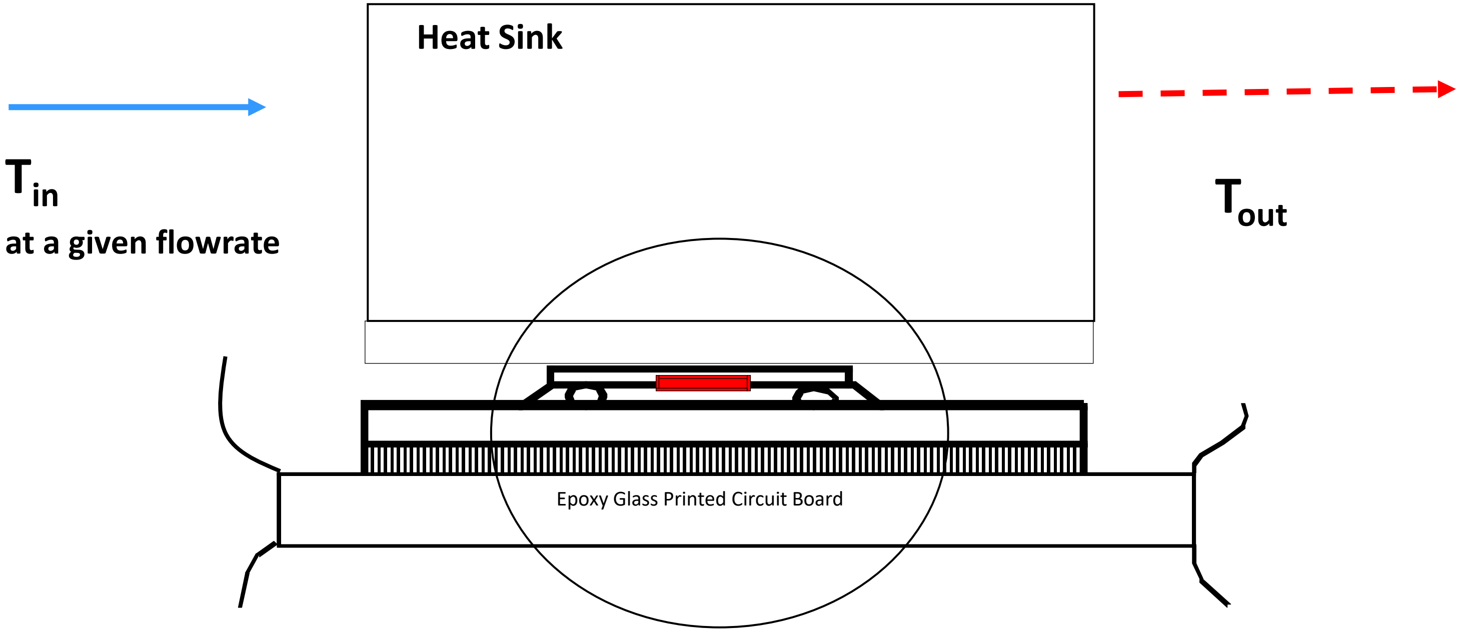
Heat Pipe Wick Structures



Challenge Faced by Small Single Core CPU Chip circa 2004 due to the concentrated heat source

Must consider spreading resistance

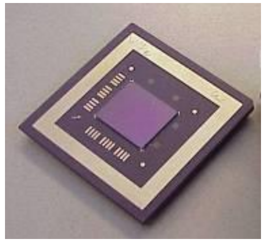




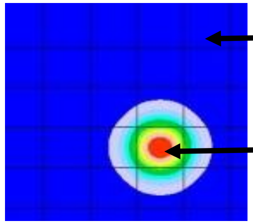
Core Temperature exceeds specifications

Choices:

- Limit Chip Power
- Design Better Cost Effective Thermal Solutions



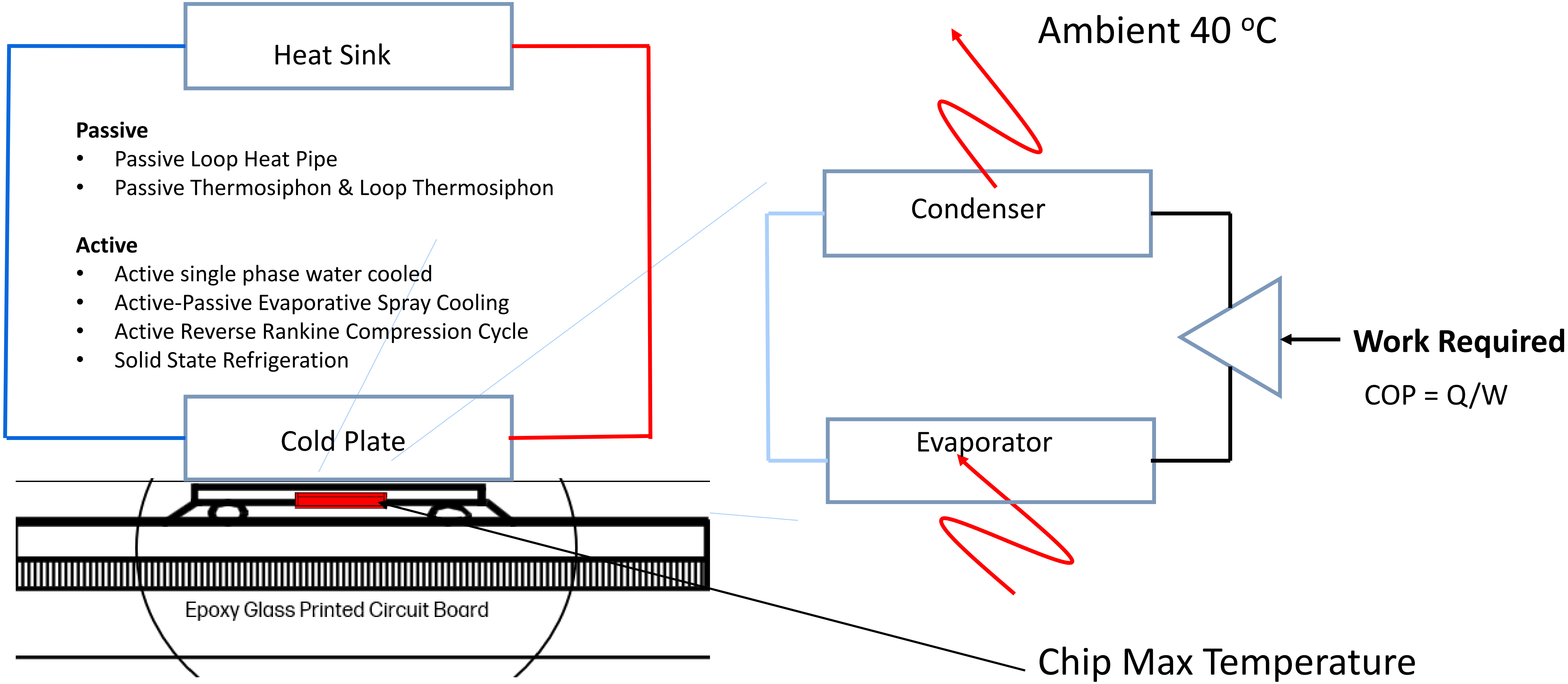
Non uniform
Heat source



T_{chip}

$T_{core} = T_{chip} + 50\ ^{o}C$

Driving the Lamborghini on the Autobahn requires active micro-mechanical means – single phase, two phase



Heat Sink

Passive

- Passive Loop Heat Pipe
- Passive Thermosiphon & Loop Thermosiphon

Active

- Active single phase water cooled
- Active-Passive Evaporative Spray Cooling
- Active Reverse Rankine Compression Cycle
- Solid State Refrigeration

Cold Plate

Epoxy Glass Printed Circuit Board

Ambient 40 °C

Condenser

Evaporator

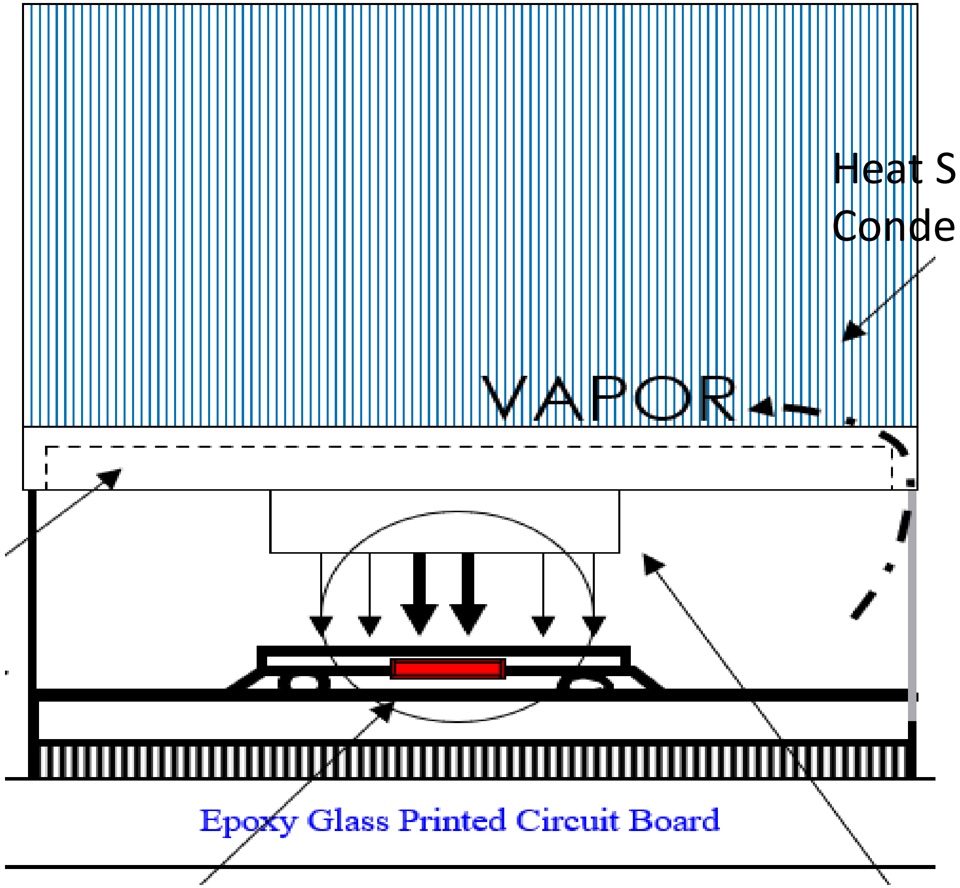
Work Required

COP = Q/W

Chip Max Temperature – 85 to 95 °C

Inkjet Assisted Spray Cooling – Chip Level

phase change evaporative cooling – precision inkjet



Heat Sink
Condenser

Inkjet Head with
reservoir for
condensate

Phase Change

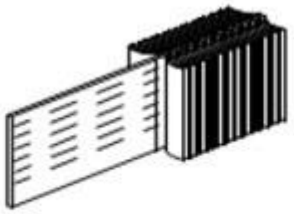
- Avoid pooling, bubble formation
- Avoid Dryout

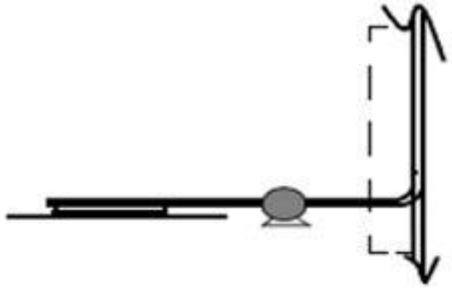
Precision “on-demand” provisioning of coolant

HP Labs 1997-2003, Best paper, Interpack 2003

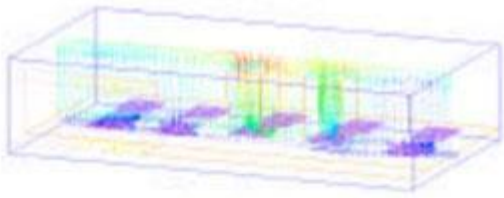
Summary of Cooling Solutions

High Fin Density
Embedded Heat Pipes

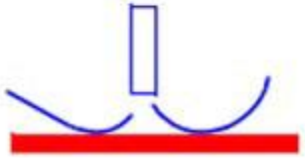


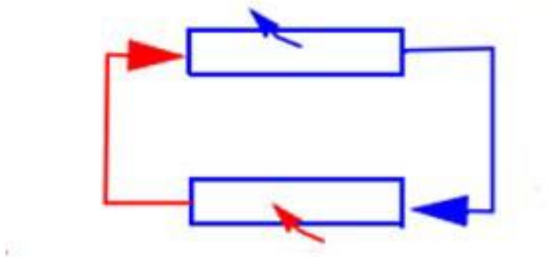


Cooling of Data Centers:
Static Optimization
Dynamic Closed Loop Control

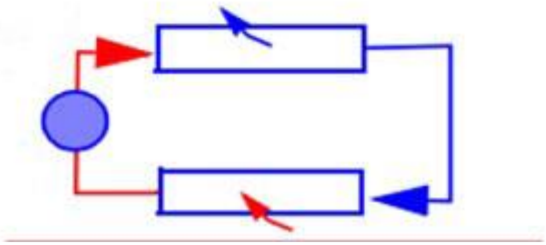


High Velocity
Airjet

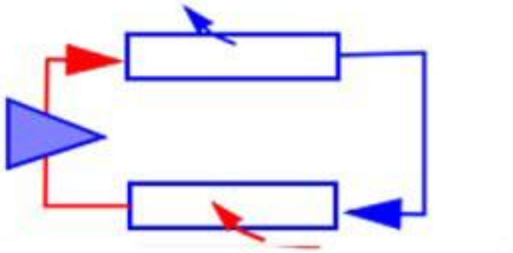




Passive Phase Change – Loop
Heatpipe and Thermosiphon

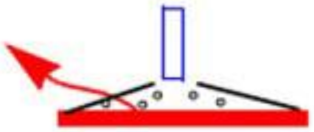


Single Phase Pumped Liquid
Loop



Reverse Rankine Vapor
Compression Cycle

Solid State Thermoelectric



Phase Change Spray
Cooling

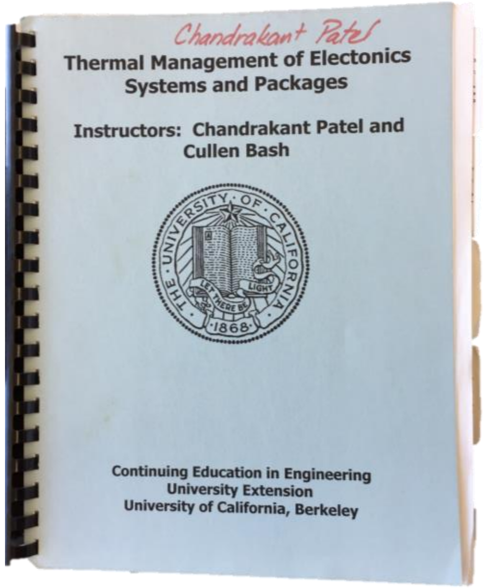
Immersion Pool
Boiling

Patel, 1999

To Learn More

1. Thermal Management of Electronic Systems and Packages, Patel and Bash
    - Graduate level course in thermal management

2. Google Scholar
    - search Patel Bash Beiltelmal Sharma Shah

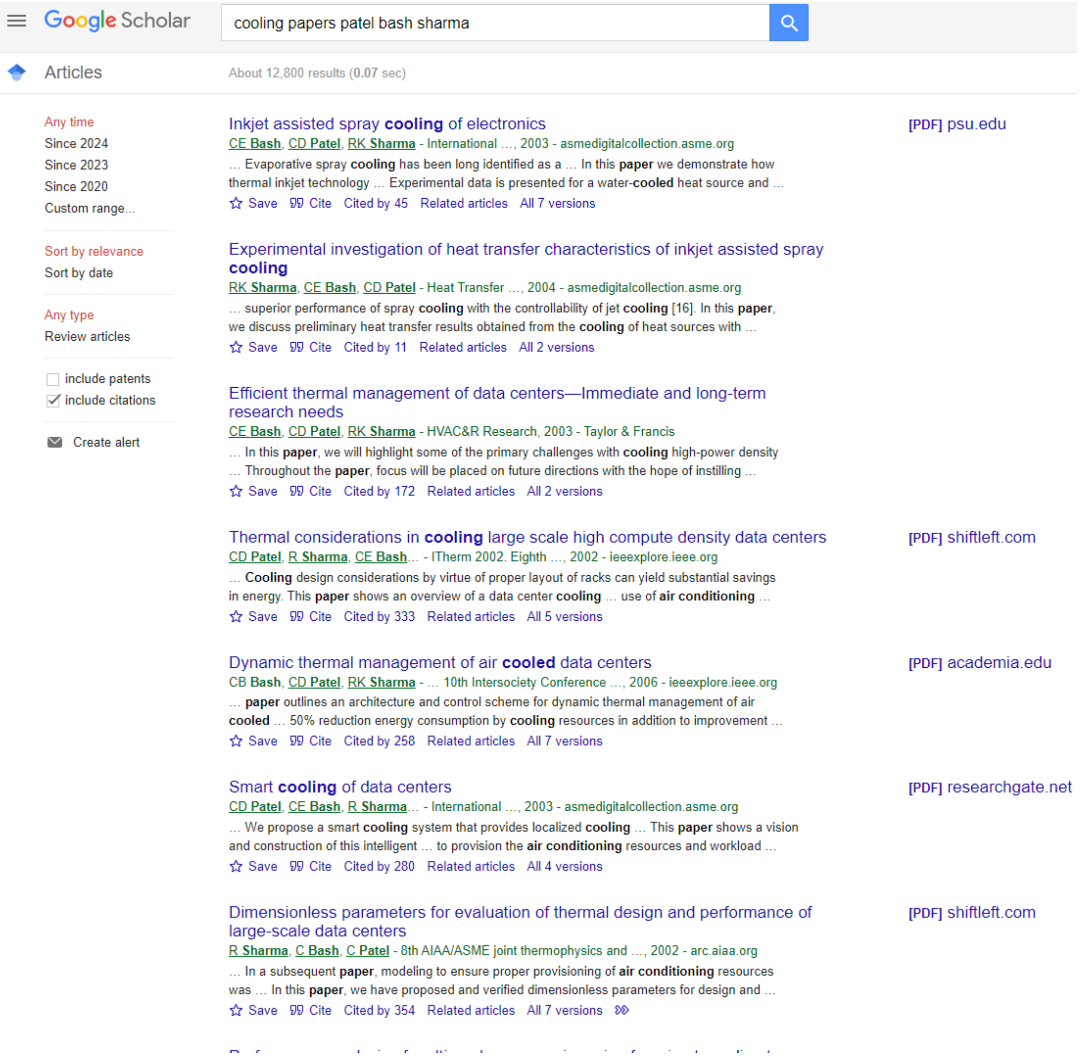


Chandrakant Patel

Thermal Management of Electonics Systems and Packages

Instructors: Chandrakant Patel and Cullen Bash

Continuing Education in Engineering
University Extension
University of California, Berkeley



cooling papers patel bash sharma

About 12,800 results (**0.07** sec)

Inkjet assisted spray **cooling** of electronics — [PDF] psu.edu
CE Bash, CD Patel, RK Sharma - International ..., 2003 - asmedigitalcollection.asme.org
... Evaporative spray **cooling** has been long identified as a ... In this **paper** we demonstrate how thermal inkjet technology ... Experimental data is presented for a water-**cooled** heat source and ...
Save Cite Cited by 45 Related articles All 7 versions

Experimental investigation of heat transfer characteristics of inkjet assisted spray **cooling**
RK Sharma, CE Bash, CD Patel - Heat Transfer ..., 2004 - asmedigitalcollection.asme.org
... superior performance of spray **cooling** with the controllability of jet **cooling** [16]. In this **paper**, we discuss preliminary heat transfer results obtained from the **cooling** of heat sources with ...
Save Cite Cited by 11 Related articles All 2 versions

Efficient thermal management of data centers—Immediate and long-term research needs
CE Bash, CD Patel, RK Sharma - HVAC&R Research, 2003 - Taylor & Francis
... In this **paper**, we will highlight some of the primary challenges with **cooling** high-power density ... Throughout the **paper**, focus will be placed on future directions with the hope of instilling ...
Save Cite Cited by 172 Related articles All 2 versions

Thermal considerations in **cooling** large scale high compute density data centers — [PDF] shiftleft.com
CD Patel, R Sharma, CE Bash... - ITherm 2002. Eighth ..., 2002 - ieeexplore.ieee.org
... **Cooling** design considerations by virtue of proper layout of racks can yield substantial savings in energy. This **paper** shows an overview of a data center **cooling** ... use of **air conditioning** ...
Save Cite Cited by 333 Related articles All 5 versions

Dynamic thermal management of air **cooled** data centers — [PDF] academia.edu
CB Bash, CD Patel, RK Sharma - ... 10th Intersociety Conference ..., 2006 - ieeexplore.ieee.org
... **paper** outlines an architecture and control scheme for dynamic thermal management of air **cooled** ... 50% reduction energy consumption by **cooling** resources in addition to improvement ...
Save Cite Cited by 258 Related articles All 7 versions

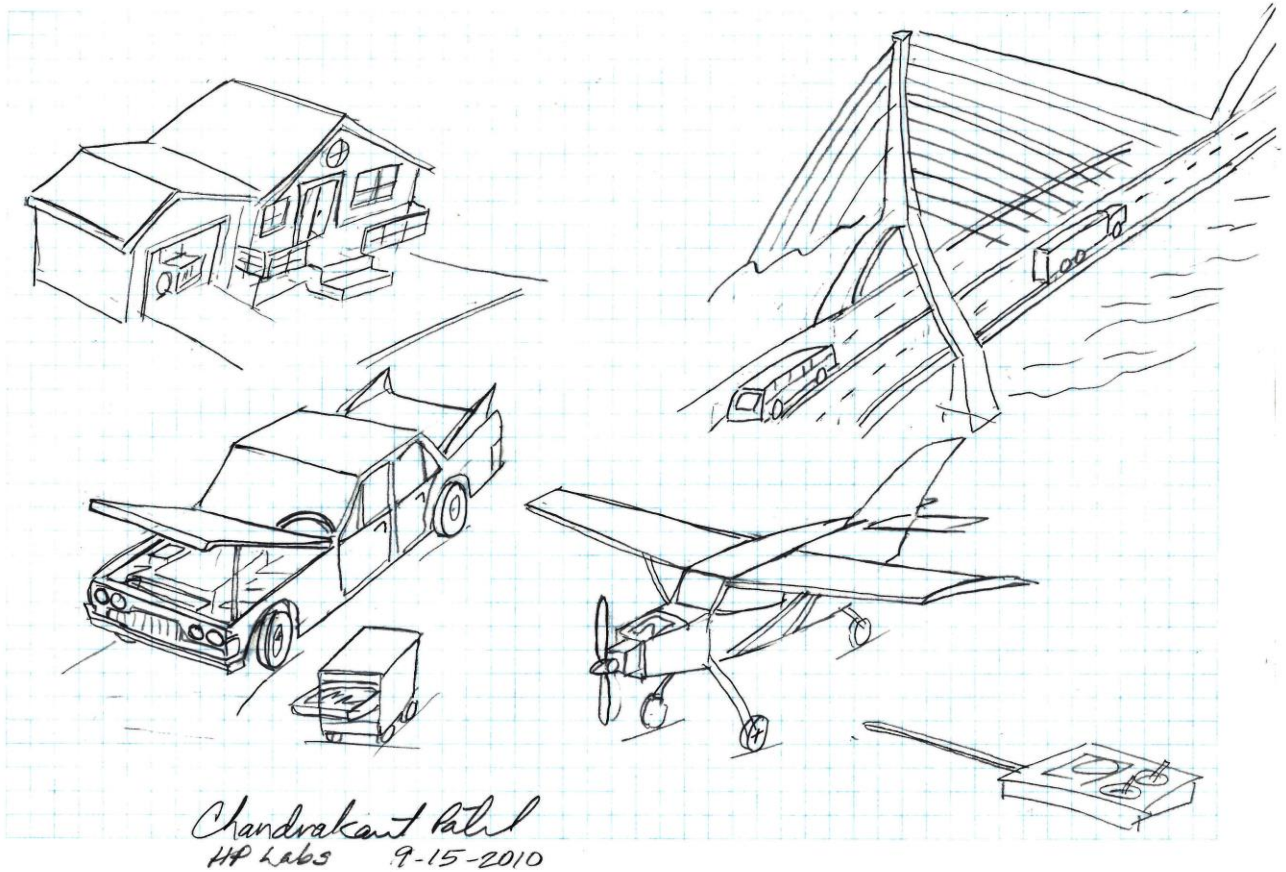
Smart **cooling** of data centers — [PDF] researchgate.net
CD Patel, CE Bash, R Sharma... - International ..., 2003 - asmedigitalcollection.asme.org
... We propose a smart **cooling** system that provides localized **cooling** ... This **paper** shows a vision and construction of this intelligent ... to provision the **air conditioning** resources and workload ...
Save Cite Cited by 280 Related articles All 4 versions

Dimensionless parameters for evaluation of thermal design and performance of large-scale data centers — [PDF] shiftleft.com
R Sharma, C Bash, C Patel - 8th AIAA/ASME joint thermophysics and ..., 2002 - arc.aiaa.org
... In a subsequent **paper**, modeling to ensure proper provisioning of **air conditioning** resources was ... In this **paper**, we have proposed and verified dimensionless parameters for design and ...
Save Cite Cited by 354 Related articles All 7 versions

A Parting Thought on Career

Point-of-View, Passion, Pivot, Practice grounded in fundamentals of engineering

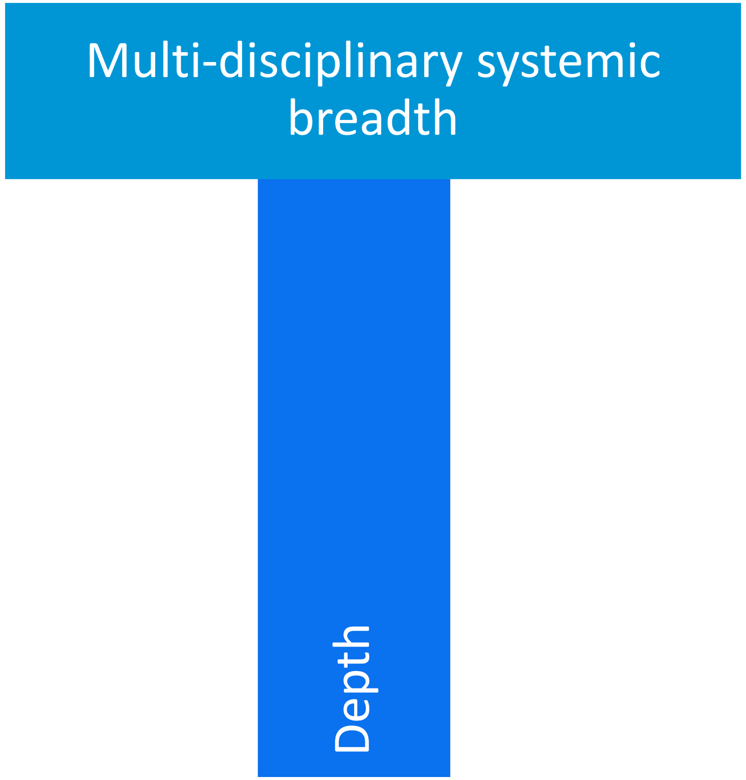
America, the land of Tinkerers, a childhood perspective



In the Valley of Tinkerers, as an early career Silicon Valley Contributor

With Bill Hewlett, Founder of HP, Inc in the Tinker Tank Interconnect and Thermal Sciences Laboratory

T-Shaped Cyber Physical Contributors



Multi-disciplinary systemic breadth

Depth

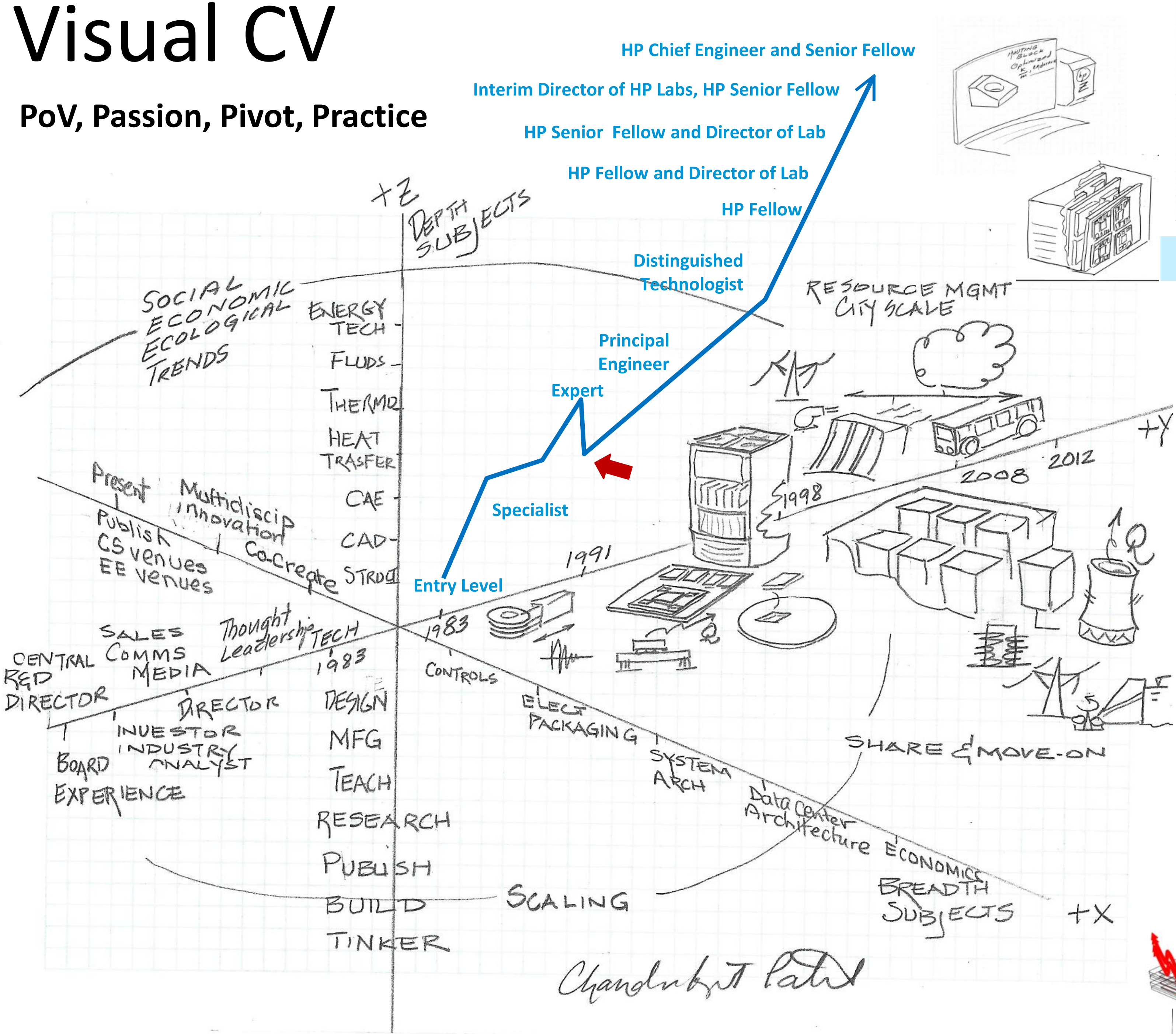
Success necessitates:

1. Depth in physical sciences
2. Breadth in multiple disciplines – cyber sciences, economics, social sciences, art, history, culture
3. Learn-by-doing

References:
Patel and Baveja, "Rise of Cyber Physical Systems", National Academies of Engineering 2023;
Patel, "Opening Doors to Opportunity", ASME Foudation, 2021

Visual CV

PoV, Passion, Pivot, Practice



HP Chief Engineer and Senior Fellow

Interim Director of HP Labs, HP Senior Fellow

HP Senior Fellow and Director of Lab

HP Fellow and Director of Lab

HP Fellow

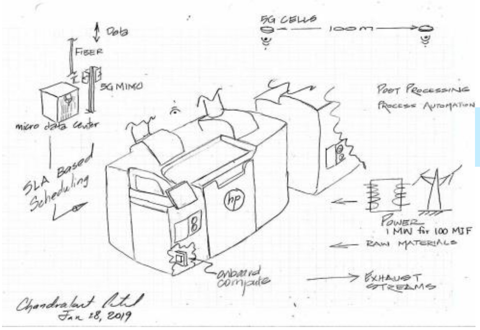
Distinguished Technologist

Principal Engineer

Expert

Specialist

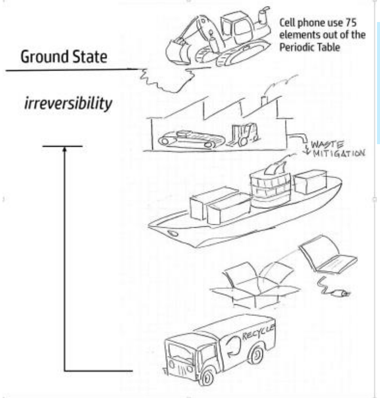
Entry Level



Cyber Physical Integration

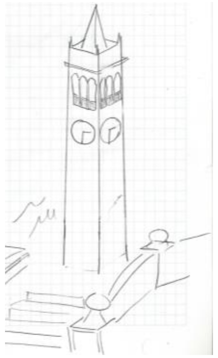
Operational Technologies + Information Technologies

Edge Computing at the source of data

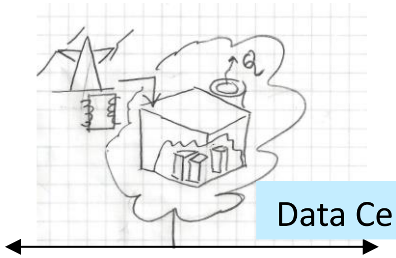


Lifetime Joules as the Currency

Net Positive Impact

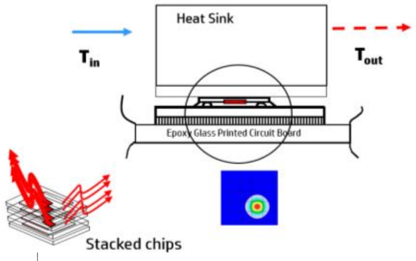


Local Workforce development



Data Center is the Computer

TCO = f(power, ping, pipe)



Thermal Management of Chips, Systems and Data Centers

Thank You

Chip Backside (non I/O side) Thermal Interface Challenge

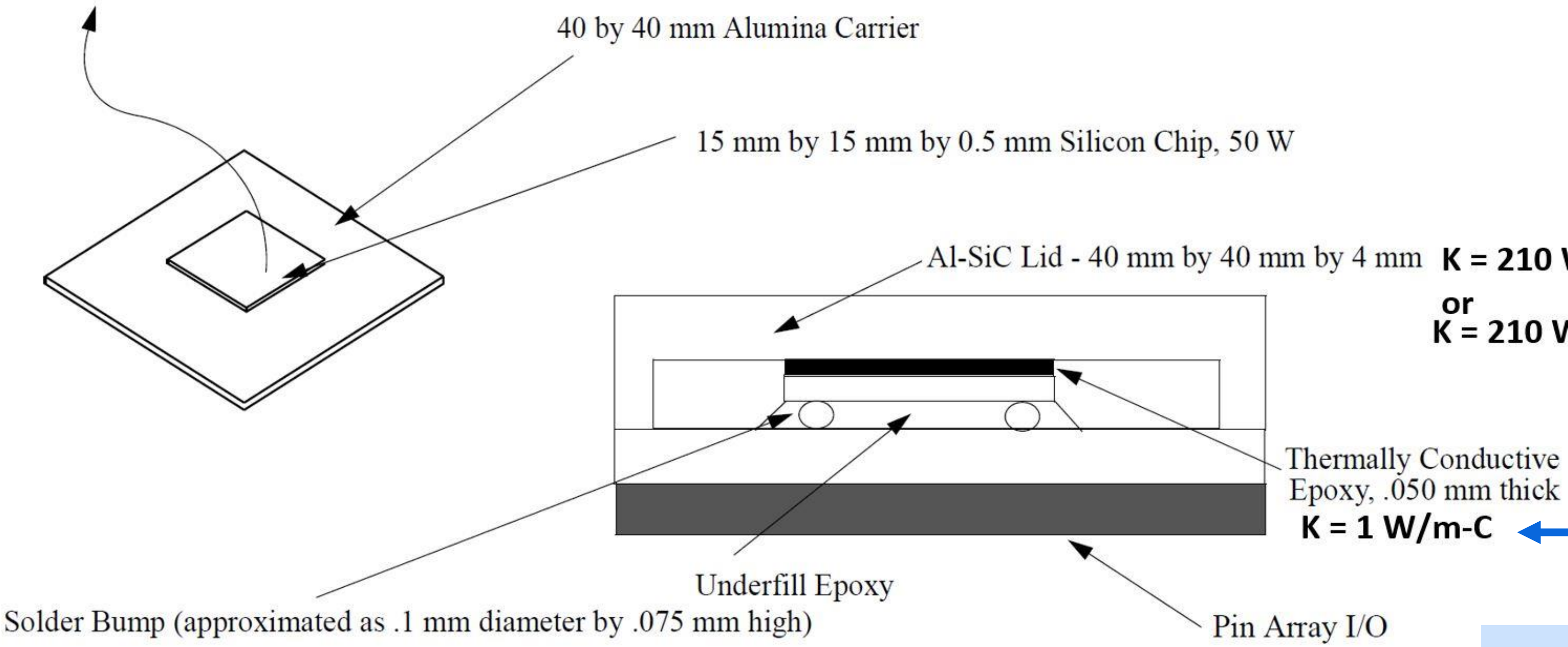
R, thermal resistance, in °C/W

$$R = \frac{L}{KA}$$

L, thickness, in meters
K, thermal conductivity, W/m-C
A, Area, m²

$$R_{chip\ interface\ layer} = \frac{L}{KA} = \frac{0.050E-3\ m}{(1\frac{W}{m}/^{o}C)(20E-3\ m * 20E-3\ m)}$$

Thermal Conductivity increased by greater volume loading and balancing rheology



K = 210 W/m-K
or
K = 210 W/m-C (in centigrade)

K = 1 W/m-C

$R_{chip\ interface\ layer} = 0.125\ ^{o}C/W$