THE ENERGY IMPACT OF NETWORKED SERVICES

Mar 15, 2023

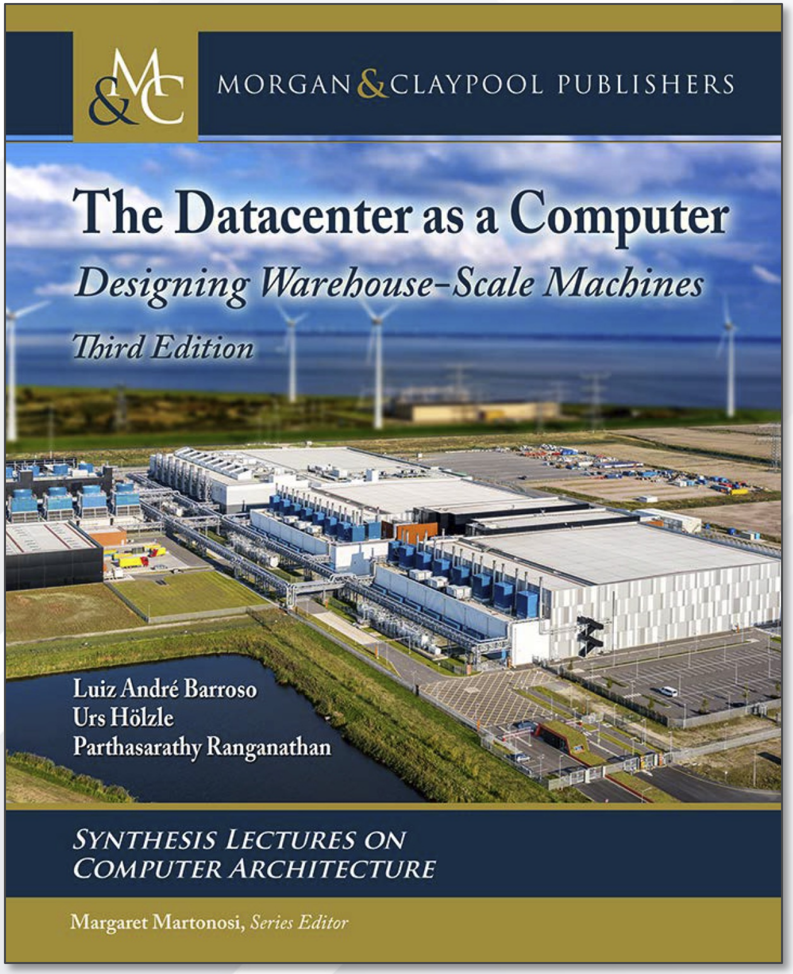
George Porter

UC San Diego

ATTRIBUTION

- These slides incorporate material from:
    - The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines, 2nd ed., by Barroso, Clidaras, and Hölzle.
    - Microsoft

Today's material available in canvas (optional reading)



POWER-PROPORTIONAL HUMANS

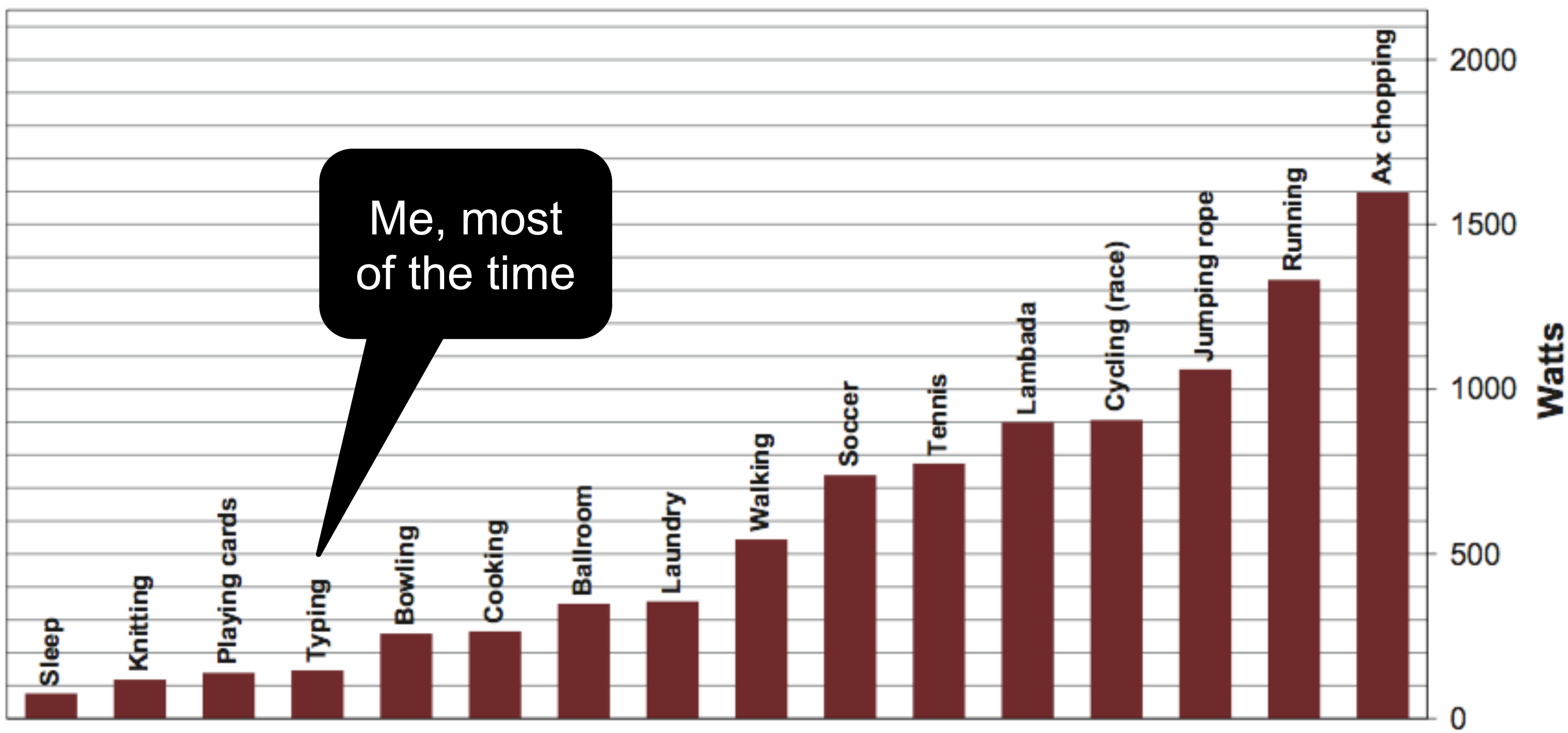


FIGURE 5.7: Human energy usage vs. activity levels (adult male) [52].

 Barroso and Holzle, *The Data Center as a Computer: An introduction to the Design of Warehouse-Scale Machines*

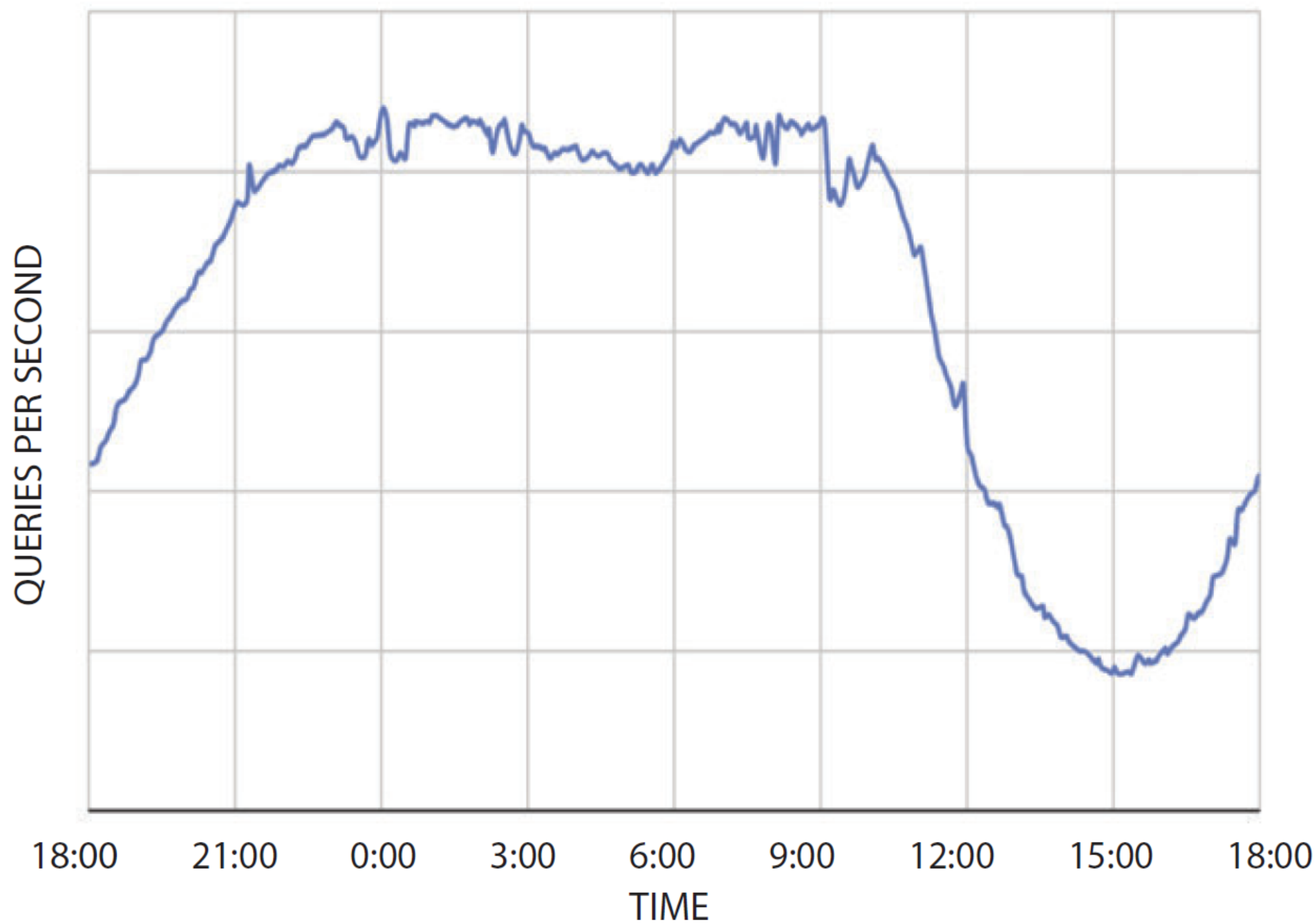


Figure 2.4: Example of daily traffic fluctuation for a search service in one data center over a 24-hr period.

SERVER HARDWARE

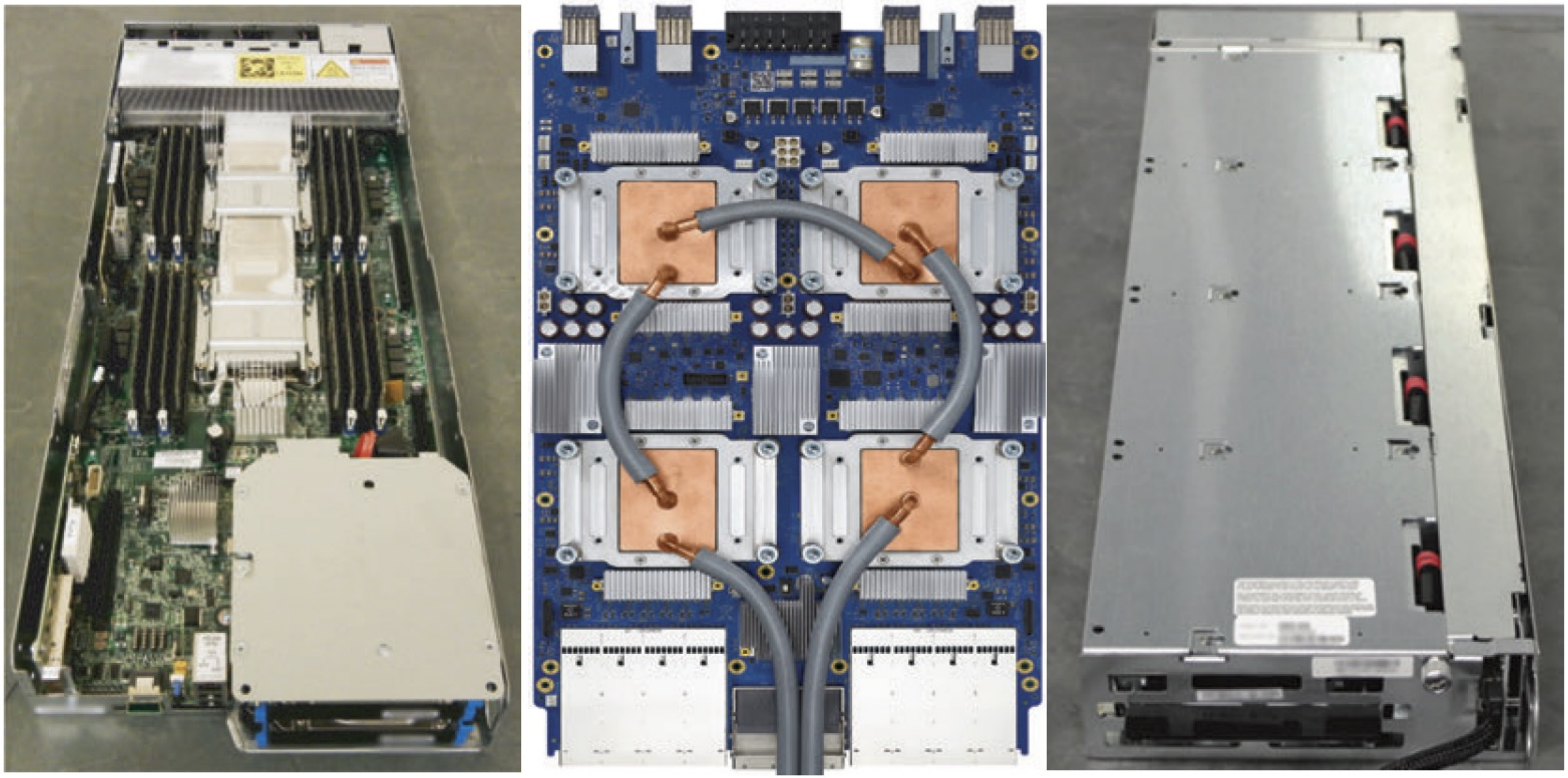


Figure 1.1: Example hardware building blocks for WSCs. Left to right: (a) a server board, (b) an accelerator board (Google's Tensor Processing Unit [TPU]), and (c) a disk tray.

DO DIFFERENT COMPONENTS SCALE SIMILARLY?

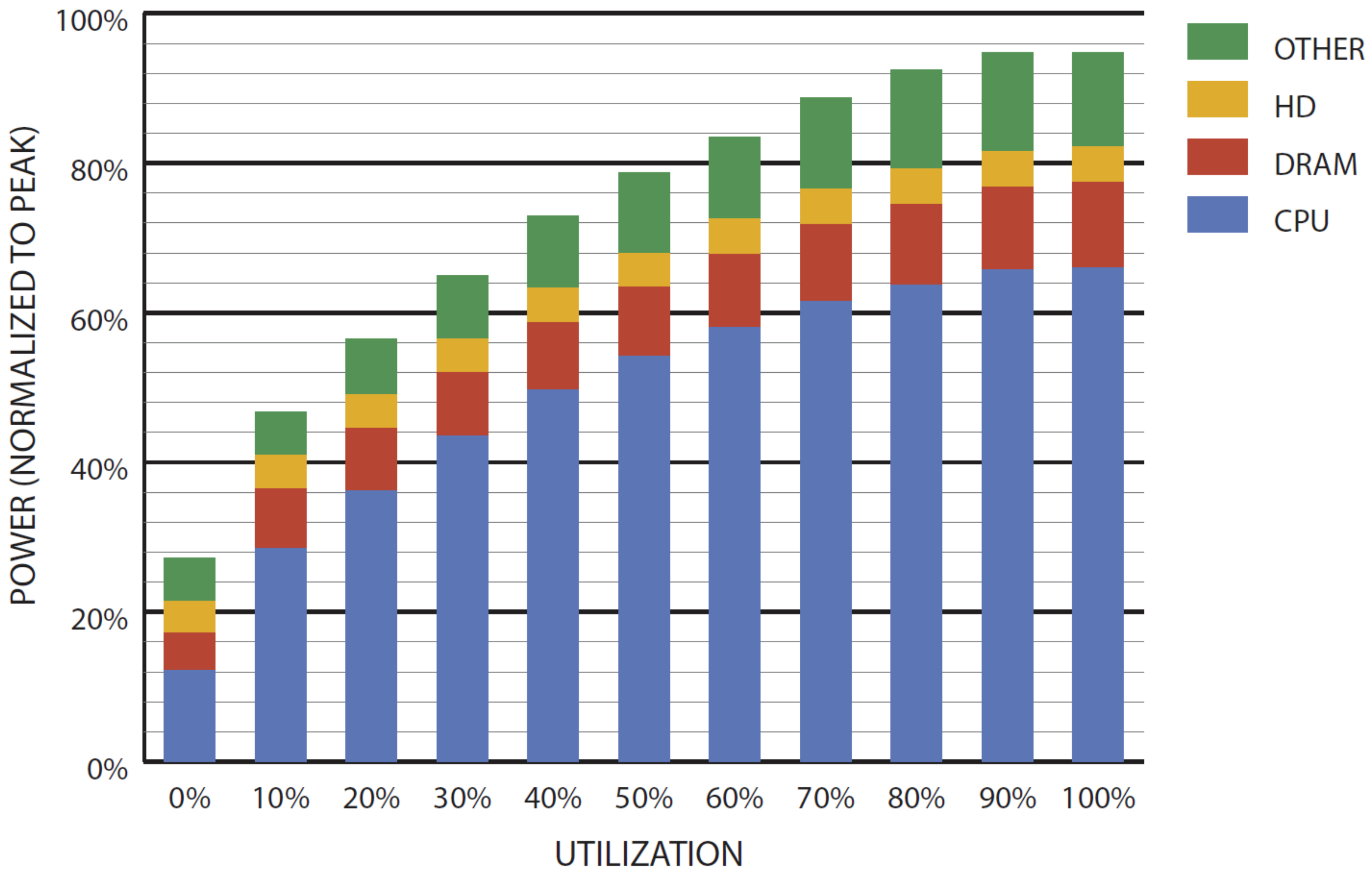


Figure 5.7: Subsystem power usage in an x86 server as the compute load varies from idle to full usage.

CPU UTILIZATION

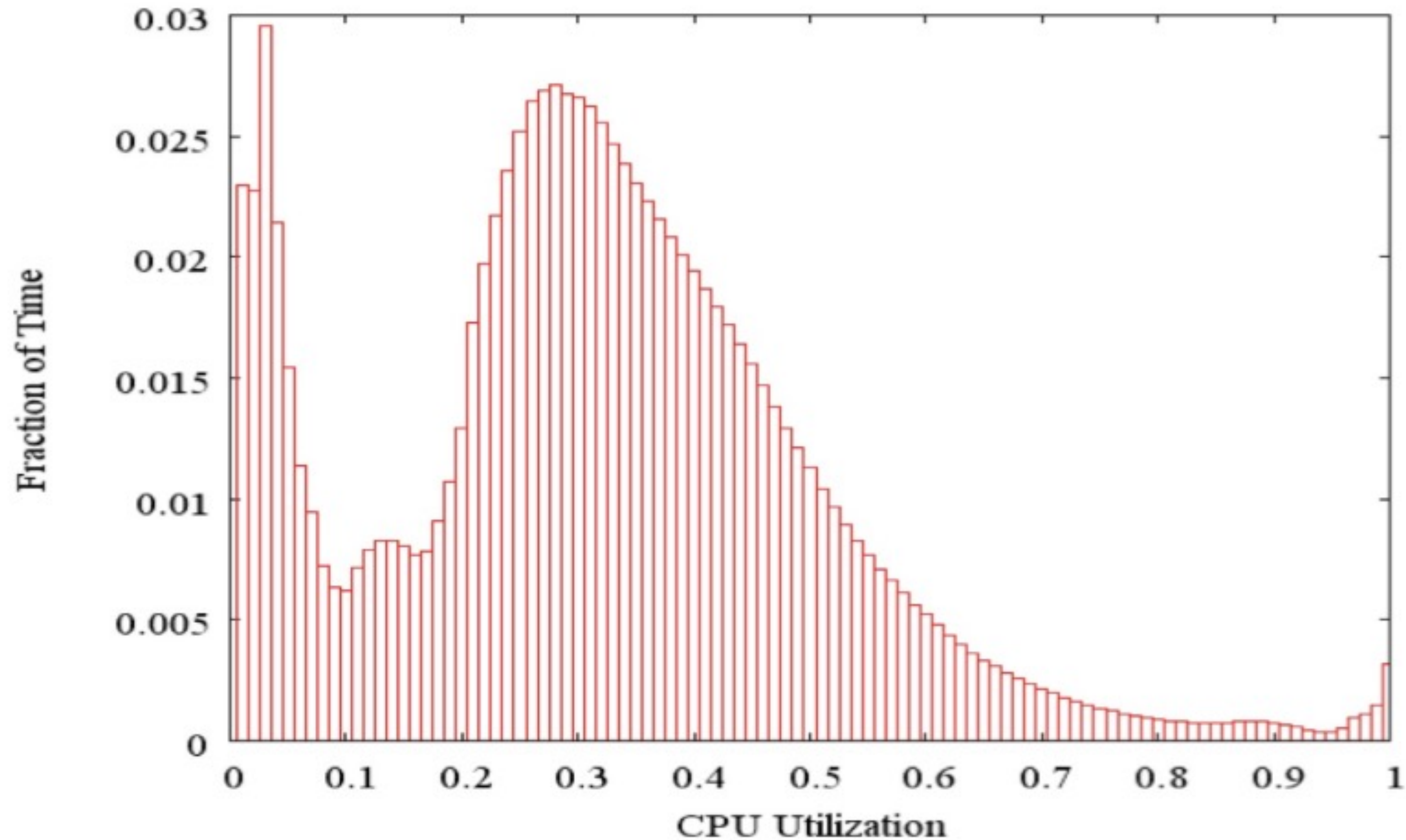


FIGURE 5.5: Activity profile of a sample of 5,000 Google servers over a period of 6 months.

 Barroso and Holzle, *The Data Center as a Computer: An introduction to the Design of Warehouse-Scale Machines*

WHAT ABOUT “POWER SAVING” FEATURES ON MODERN COMPUTERS?

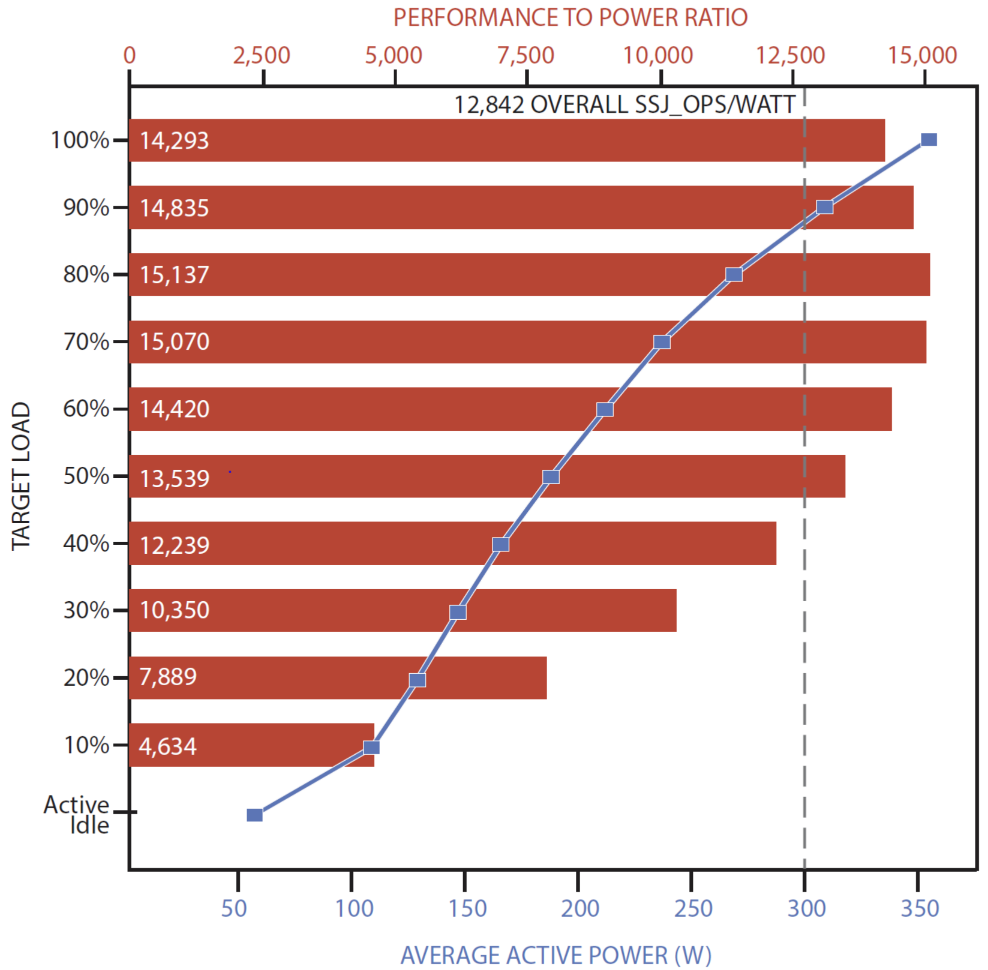


Figure 5.4: Example benchmark result for SPECpower_ssj2008; bars indicate energy efficiency and the line indicates power consumption. Both are plotted for a range of utilization levels, with the average energy efficiency metric corresponding to the vertical dark line. The system has two 2.1 GHz 28-core Intel Xeon processors, 192 GB of DRAM, and one M.2 SATA SSD.

INCREASING POWER PROPORTIONALITY OVER TIME

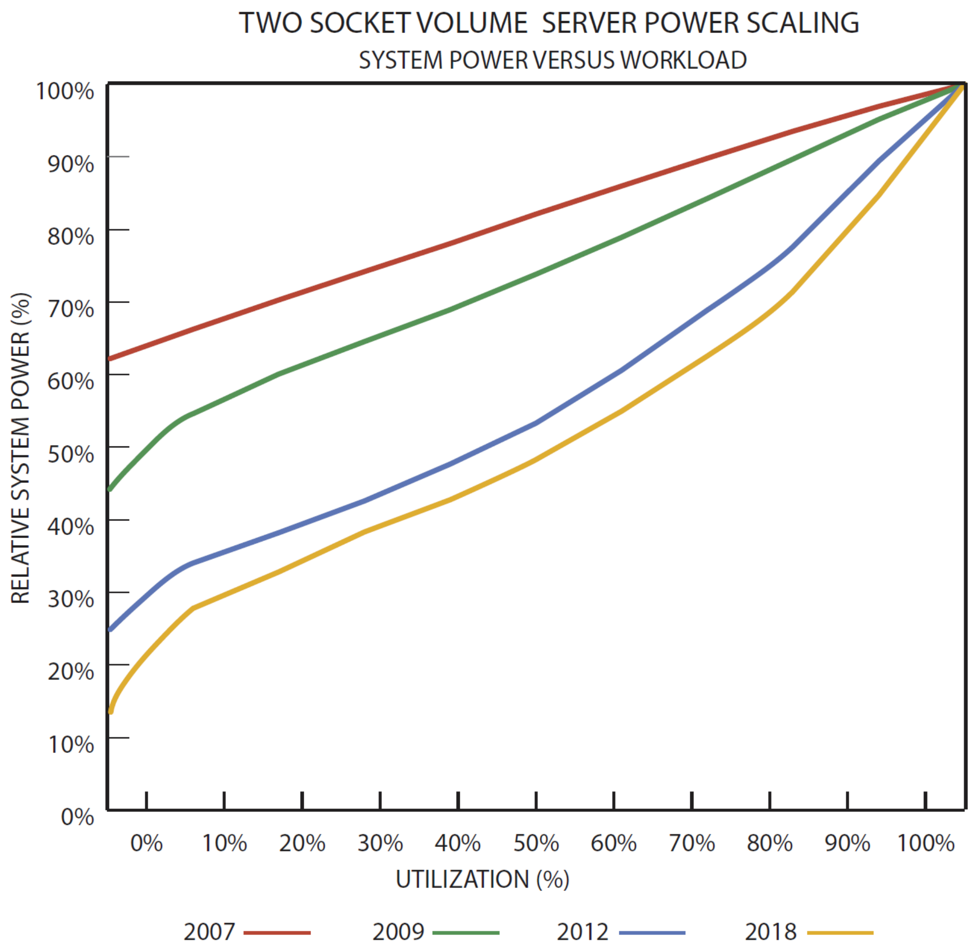


Figure 5.8: Normalized system power vs. utilization in Intel servers from 2007–2018 (courtesy of David Lo, Google). The chart indicates that Intel servers have become more energy proportional in the 12-year period.

EXAMPLE: “DATA PARALLEL” PROCESSING

```
for (iterations = 1 to 100){
  work = new Task[10000];
  for (i = 1 to 10000) {
    rpc.asyncCall(server[i],
  }
  waitForAllAsyncCallsToComp
}
```

Assume latency drawn from N(mu,sigma)

with 1 in 100 requests >= 1 second

What happens to the ~9,900 servers that finished “quickly”?

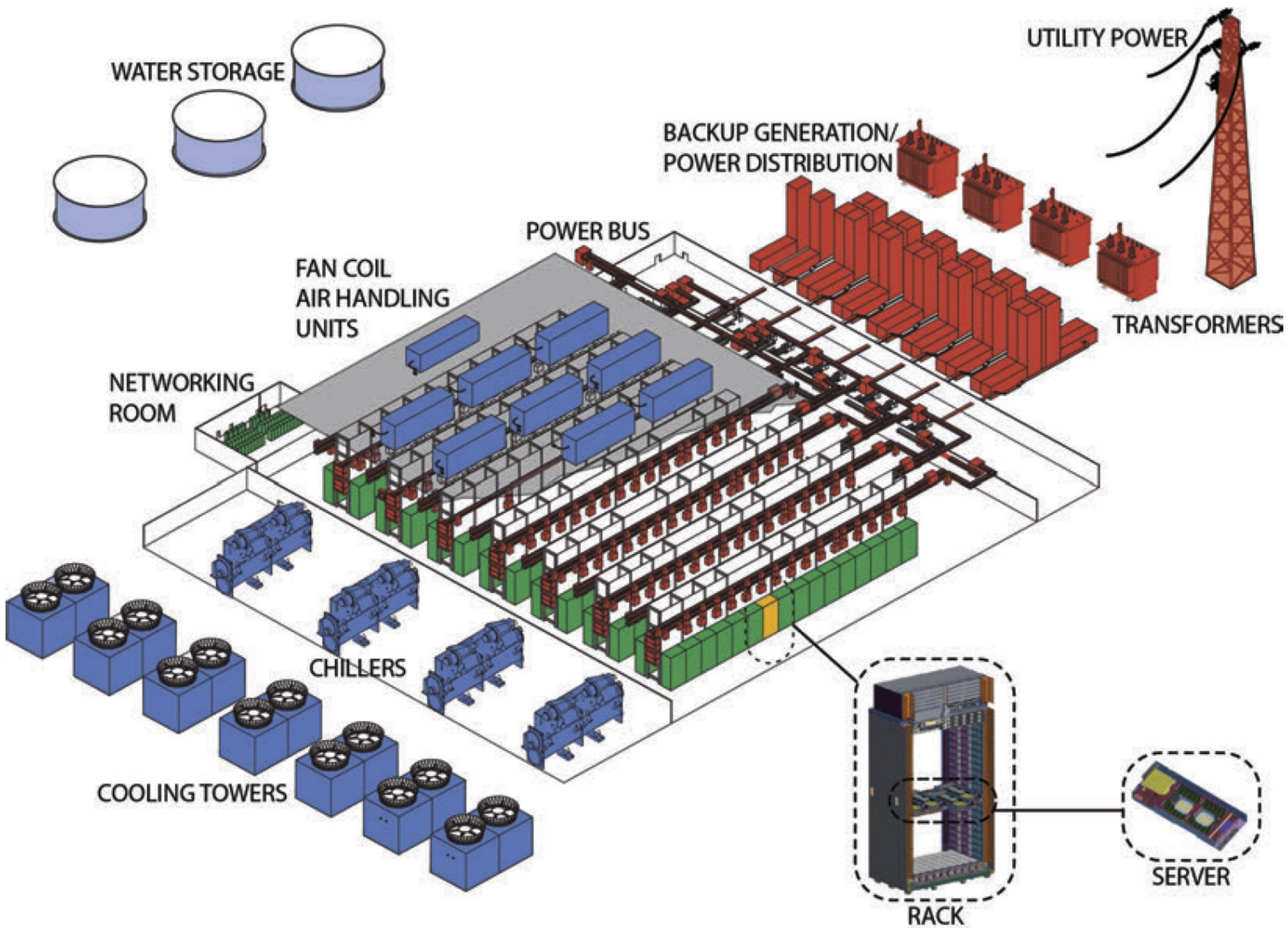


Figure 4.4: The main components of a typical data center.

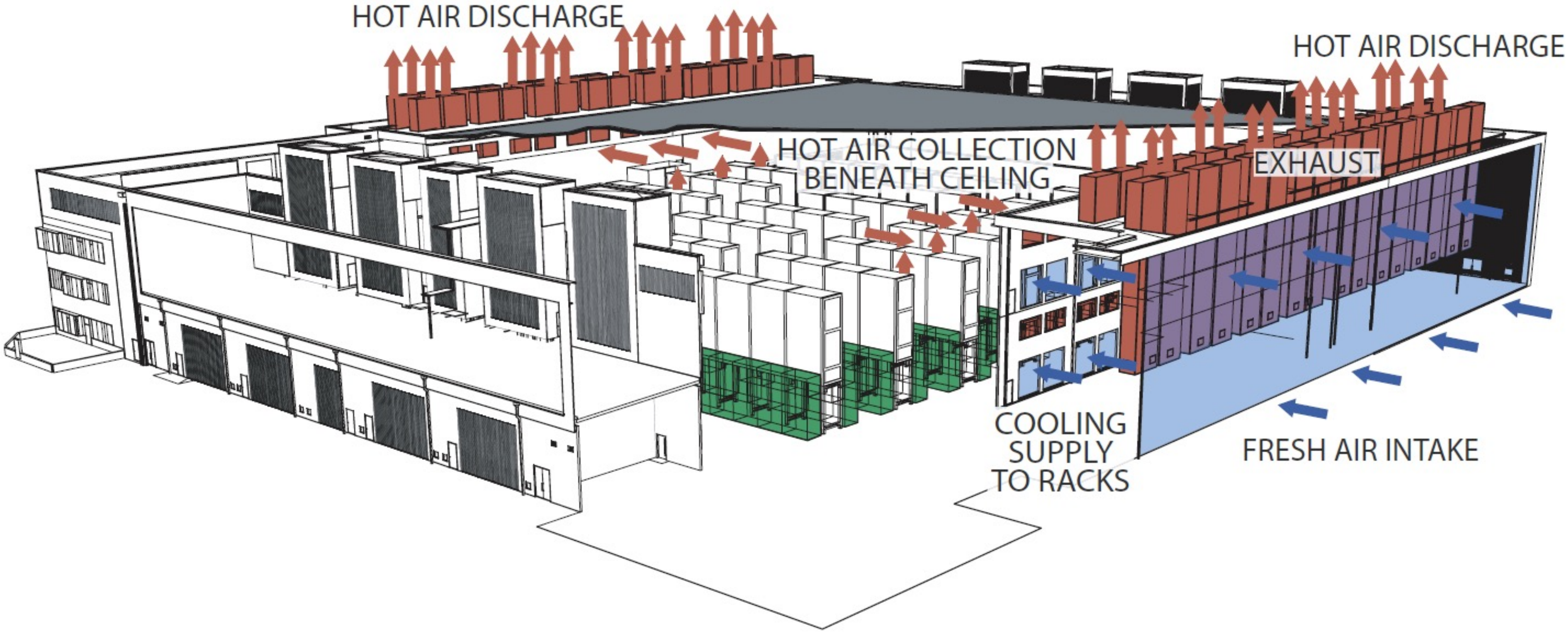


Figure 4.8: Airflow schematic of an air-economized data center.

ENERGY “OVERHEAD”: A/C AND COOLING

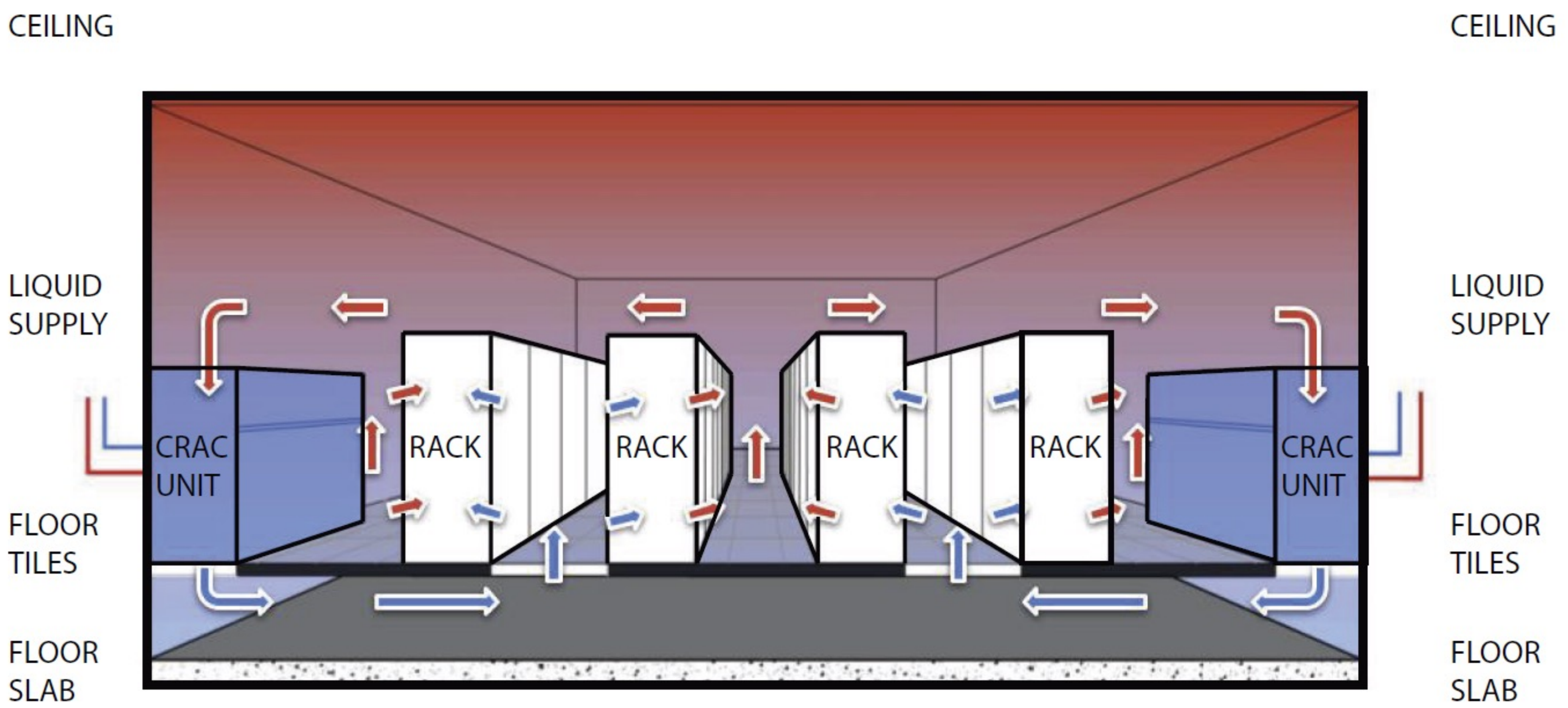


Figure 4.9: Raised floor data center with hot-cold aisle setup (image courtesy of DLB Associates [Dye06]).

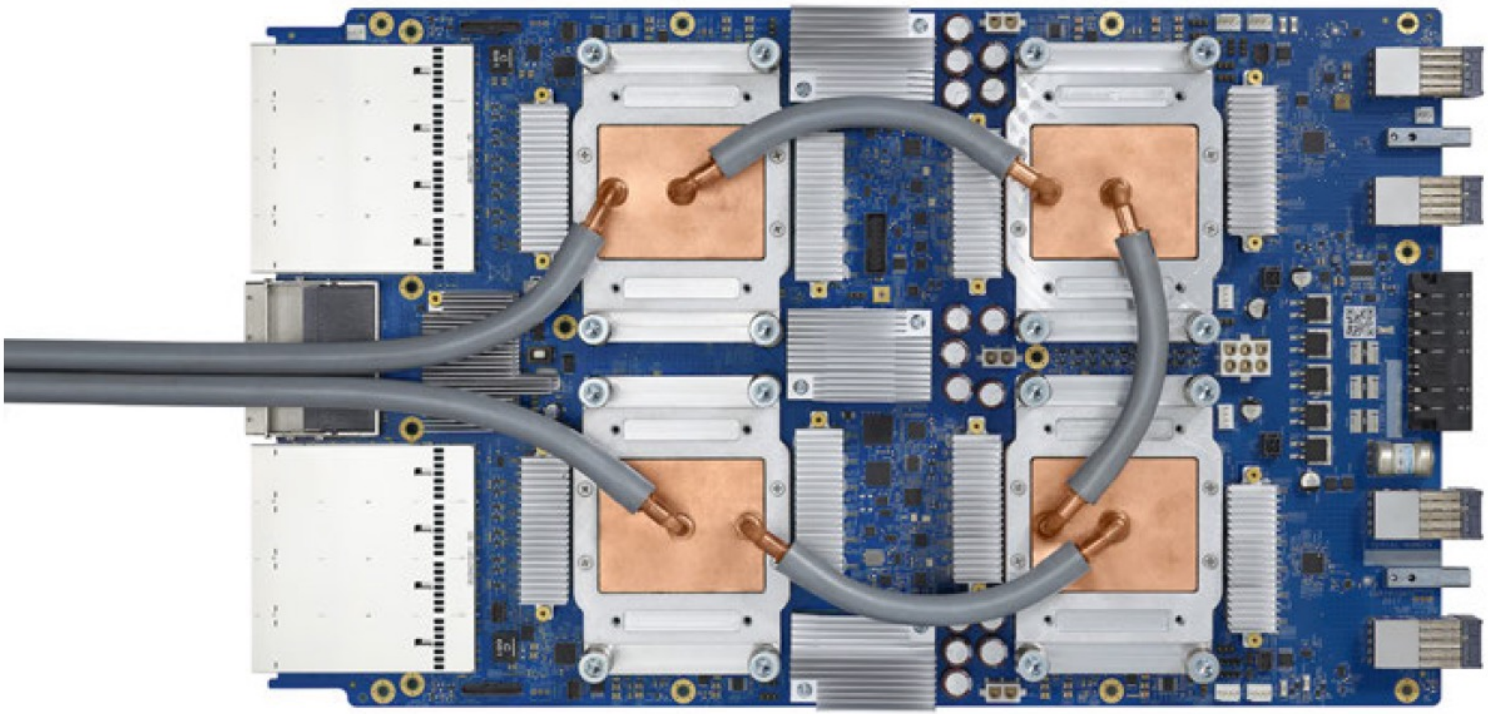
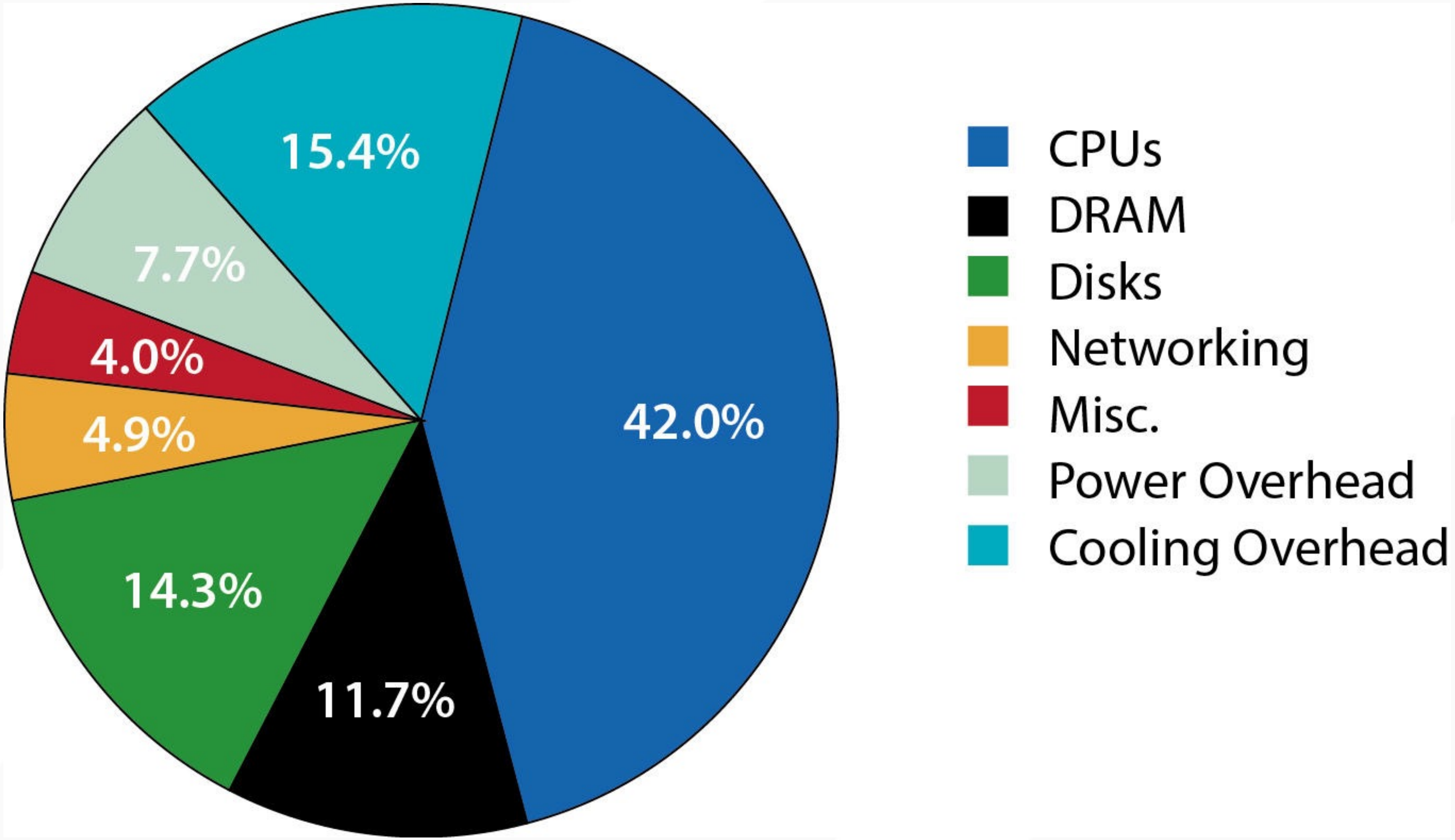


Figure 4.16: Copper cold plates and hose connections provide liquid cooling for Google's third-generation TPU.

NUMBERS FROM JAMES HAMILTON (EARLY 2010S) (MSFT, AMAZON)



GOOGLE, ~2012 OR SO

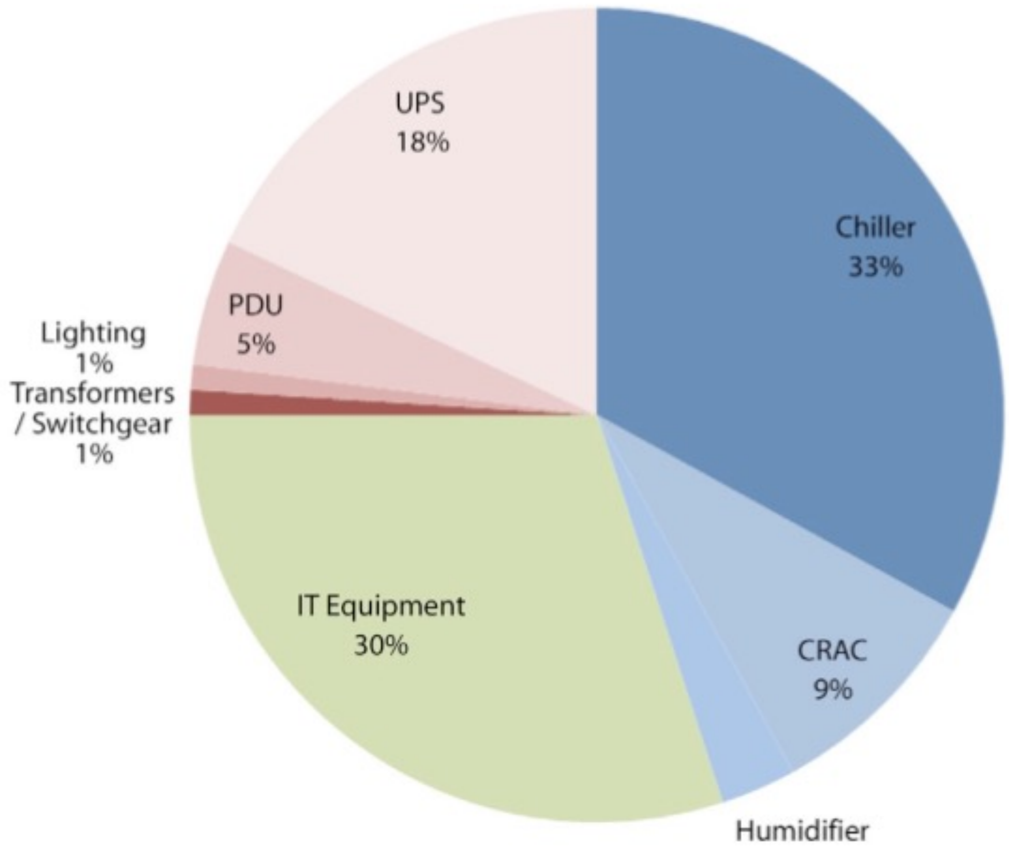


FIGURE 5.2: Breakdown of datacenter energy overheads (ASHRAE).

 Barroso and Holzle, *The Data Center as a Computer: An introduction to the Design of Warehouse-Scale Machines*

GOOGLE, ~2016

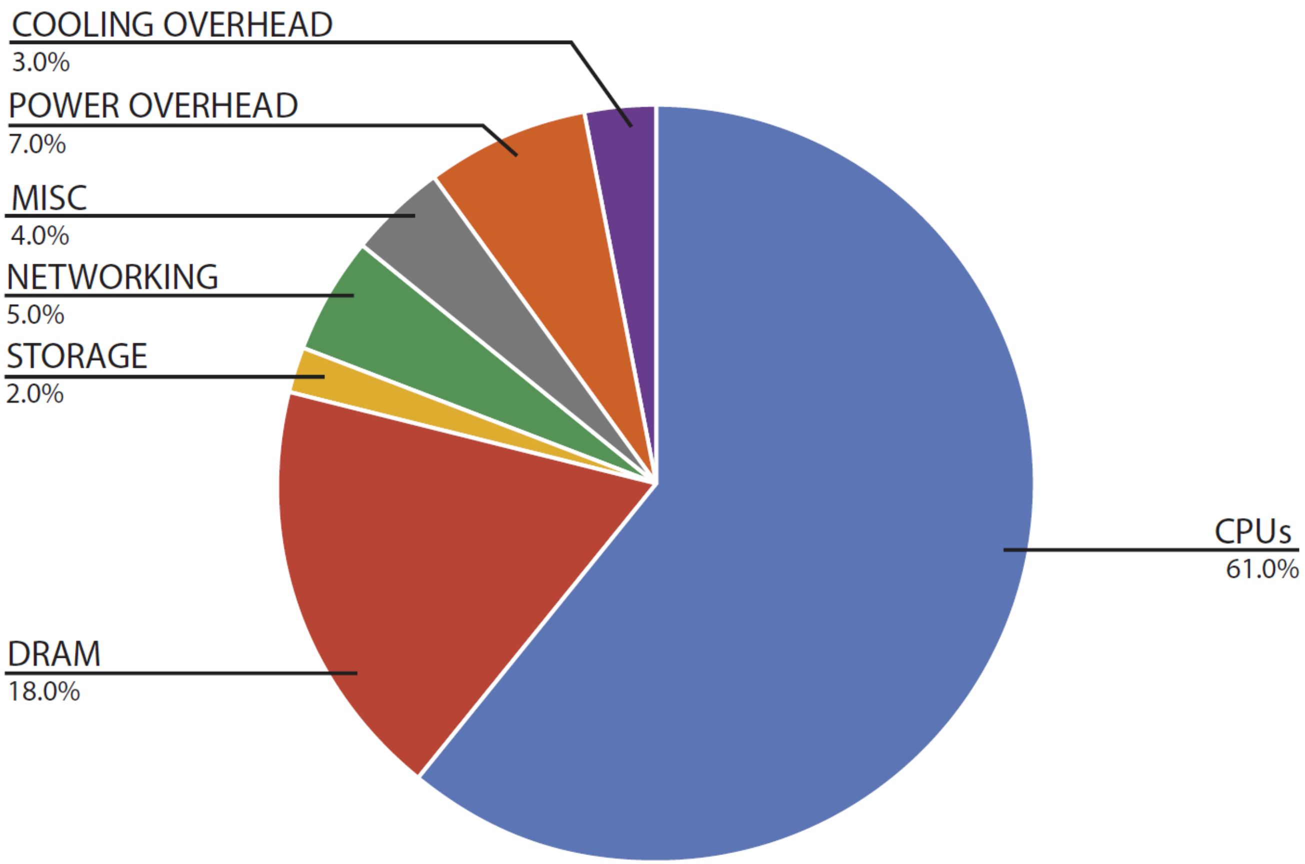


Figure 1.8: Approximate distribution of peak power usage by hardware subsystem in a modern data center using late 2017 generation servers. The figure assumes two-socket x86 servers and 12 DIMMs per server, and an average utilization of 80%.

QUANTIFYING ENERGY-EFFICIENCY: PUE

- PUE = Power Usage Effectiveness
- Simply compares
  - Power used for computing vs Total power used
- Historically cooling was a huge source of power
  - E.g., 1 watt of computing meant 1 Watt of cooling!

PUE = (Facility Power) / (Computing Equipment power)

LBNL PUE SURVEY (2007)

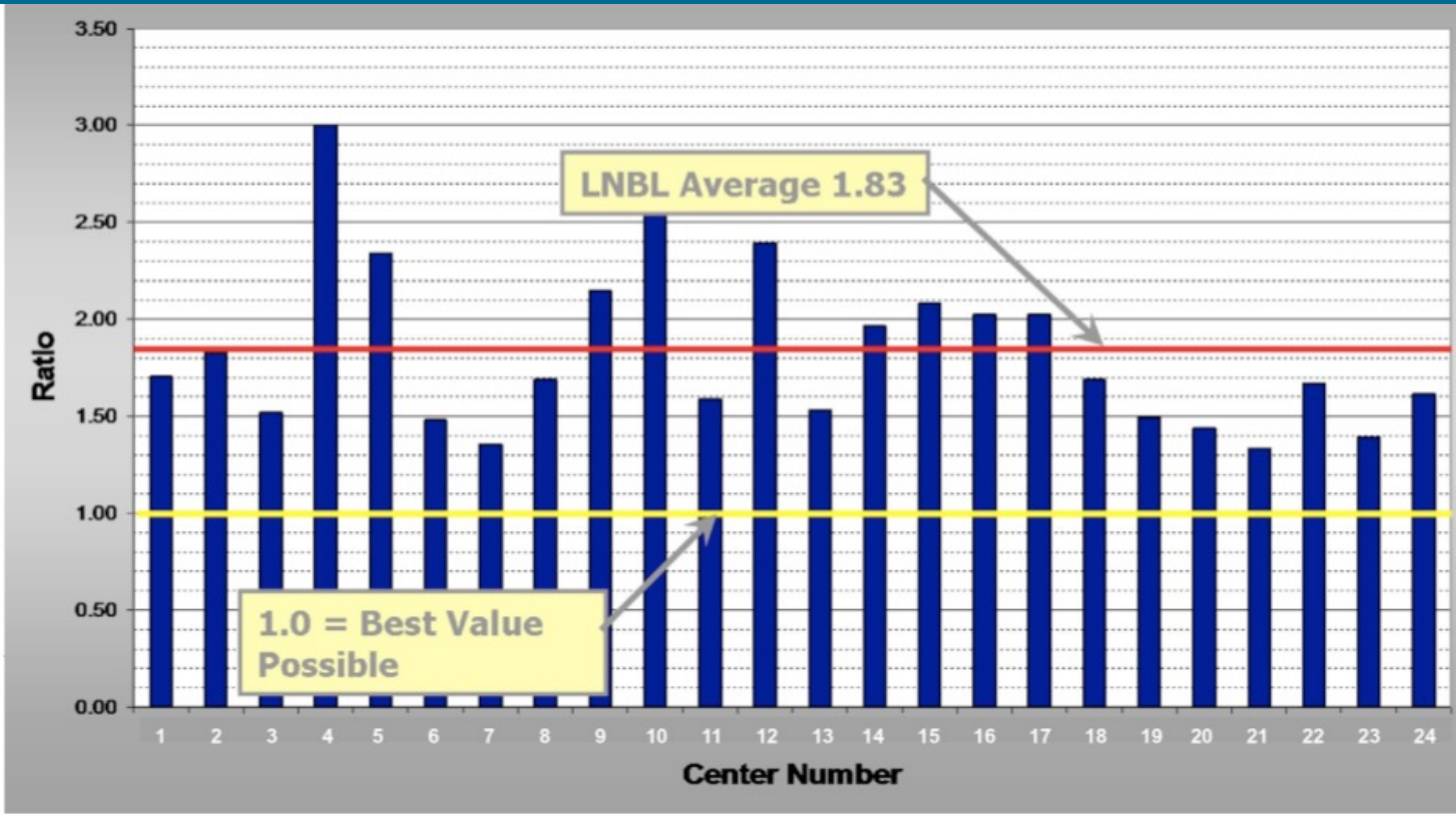


FIGURE 5.1: LBNL survey of the power usage efficiency of 24 datacenters, 2007 (Greenberg et al.) [41].

 Barroso and Holzle, *The Data Center as a Computer: An introduction to the Design of Warehouse-Scale Machines*

LBNL PUE Survey (2013)

Table 1: Summary of power/energy data

Data center IT devices (W/device)	
Volume server	235
Midrange server	450
External HDD spindle	26
Data center PUE	
Server closet	2.5
Server room	2.1
Localized	2
Mid-tier	2
Enterprise-class	1.5
Cloud	1.1
Network data transmission (μJ/bit)	
Wired	100
Wi-Fi	100
Cellular (3G/4G)	450

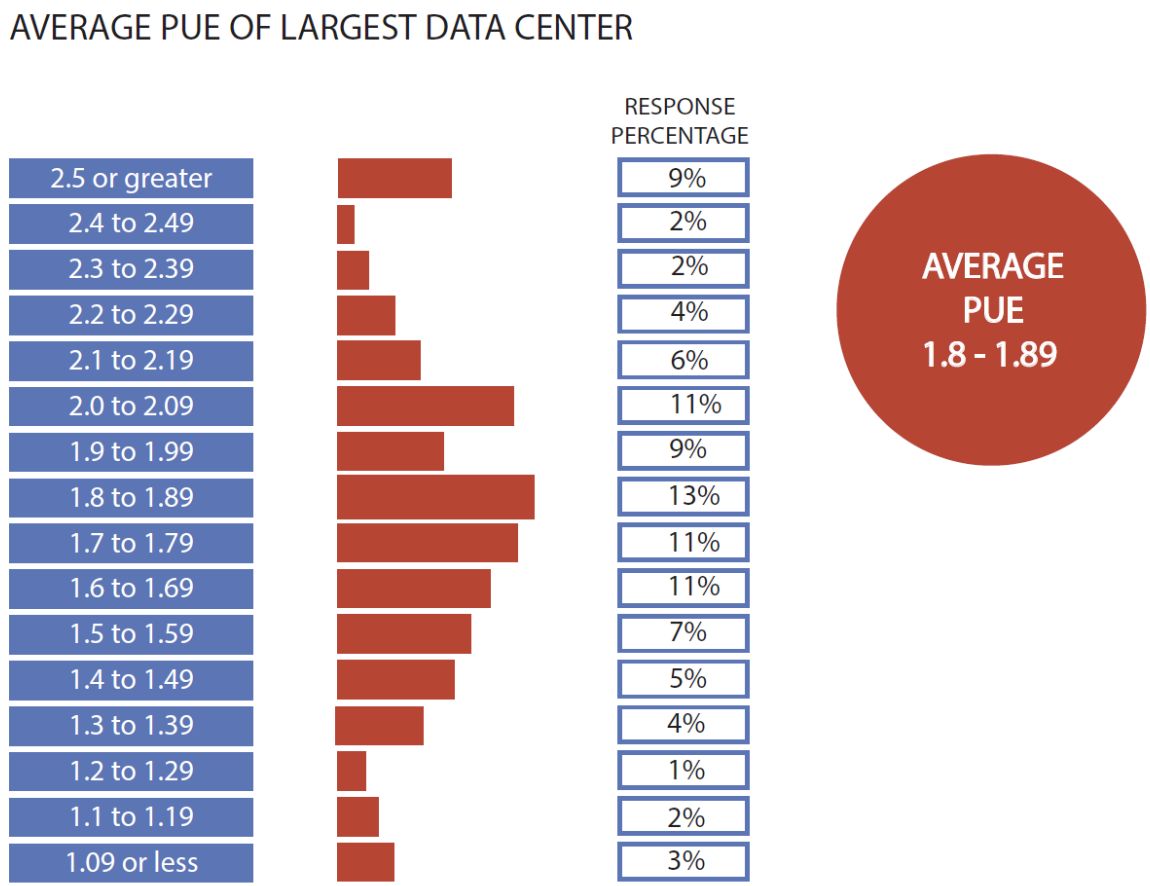


Figure 5.1: Uptime Institute survey of PUE for 1100+ data centers. This detailed data is based on a 2012 study [UpI12] but the trends are qualitatively similar to more recent studies (e.g., 2016 LBNL study [She+16]).

DATACENTER WAREHOUSE

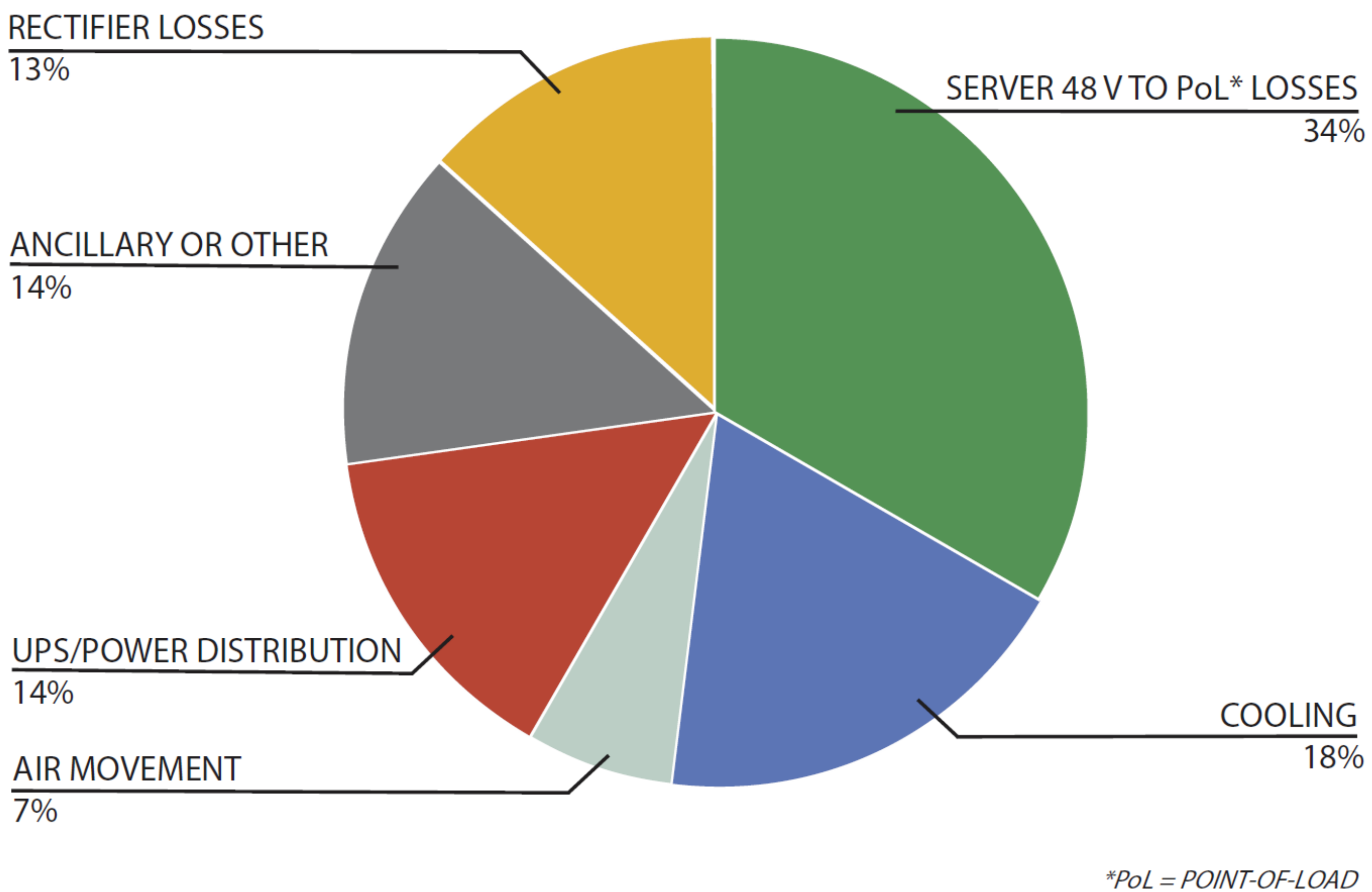


Figure 5.3: A representative end-to-end breakdown of energy losses in a typical datacenter. Note that this breakdown does not include losses of up to a few percent due to server fans or electrical resistance on server boards.

MICROSOFT’S PROJECT NATICK

The Economist

Menu | Weekly edition | Search

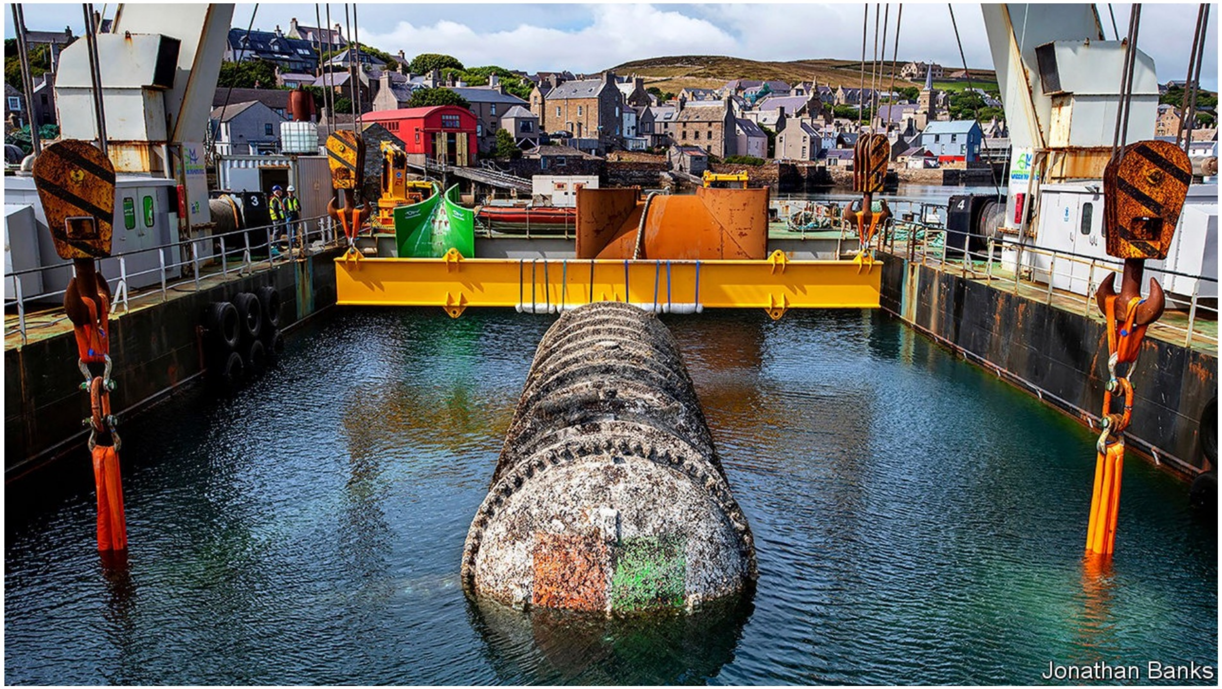
Science & technology

Sep 19th 2020 edition ›

Cloud computing

Davy Jones’s data-centre

Building server farms underwater is less crazy than it sounds



MICROSOFT NATICK TOUR

- https://youtu.be/LmfvUiJ6tB8

A note about CAPEs

- **Only** source of feedback to UCSD about Professor & TA teaching
- **Anonymous**
- **Optional**
- **Extremely important**
  - Determines whether faculty get promoted, get tenure, keep their jobs
  - Determines whether TAs become TAs in the future

- **Please make your voice heard!**
  - Usually only students who ***love*** or ***hate*** the class fill them out
  - We appreciate the few minutes it takes to make your opinions/voice heard, esp for the TAs

Thanks!

EXAM LOGISTICS

- “Cheat sheet”
  - Four 8.5x11” paper (both sides)
- No calculator needed
- Covers all material including projects, readings

- Performance metrics such as bandwidth and latency. Tools like ping and iperf for measuring latency/bandwidth. Which factors are more important for small transfers? For large transfers? What does "latency sensitive" and "bandwidth sensitive" mean?
- Network layering. What is a "layer 3" protocol? What is a "layer 4" protocol? What is layer 7?
- Network semantics. What semantics does the IP protocol offer to applications? What semantics does UDP? TCP?
- Addressing. Format of an IPv4 vs IPv6 address. Subnets and subnet masks. CIDR prefix notation. What does x.x.x.x/p mean? Address aggregation. How to work with IP address prefixes. How to carve them up into subnets? Classful vs Classless address assignments
- TCP ports (source and dest)
- Network address translation.
- Sockets API. Socket, Listen, Connect, Bind, Close, etc. Which of these calls blocks? Under what conditions? How does the send and receive buffer sizes related to blocking behavior? Can deadlock occur between two endpoints of a TCP connection?
- Framing and parsing. Delimeters vs explicit length fields. Algorithms for framing variable length messages with delimiters
- HTTP protocol, including requests, responses, structure of messages, key/value pairs, protocol structure.

- Web architecture more generally. Reverse vs forward proxies. Caching. Virtual hosting. How does a cache know when data is out of date?
- The domain name system. TLD vs root vs authoritative name servers. Resource record types. How does caching work within DNS?
- Content distribution networks. How does DNS support tranparent caching in a CDN? What aspects of the HTTP protocol are used to support CDNs? How does Akamai work? What is the Stable Marriage problem and how does it relate to content distribution networks
- Basics of the Go language. I will not ask you to write lots of code on the exam, but you may need to write a line or two of code in a function. A brief summary of the parameters and return types of any Go functions you will need will be provided to you as needed. You may be shown a snippet of code and asked questions about it.
- Knowledge/experience/content in projects, as well as any demos we covered in lecture.

PROTOCOLS AND NETWORKING BASICS

- Performance (prop delay, latency, transmit time, serialization latency, queuing, …)
- Framing vs. parsing
- Delimiter vs. length-value
- Server-side protocol handling
- Deep dive on HTTP
- Encoding (in Go, JSON, XML, protobufs)
- Protobufs and gRPC IDLs and tools

REPLICATION, CONSISTENCY AND FAULT TOLERANCE

- Two-phase commit
- Quorums
- RAFT paper
- Dynamo DB paper (main topics)
- Akamai paper (sections referred to in class)
- Akamai CDN + DNS

DATA CENTERS AND CDNS

- Round-robin DNS vs load balancers. advantages and disadvantages of each
- Terms: MTTR, MTBF, availability, yield, harvest, DQ principle
- Tail latency vs. average latency
- Consistent hashing
- Energy and environmental impact of datacenters (slides not the optional reading on canvas)

UC San Diego

GM

Pretend that you are a college professor and write a short but encouraging note for your students who have final exams next week

Dear students,

As we approach the final exams next week, I know many of you are feeling the pressure and stress that comes with this time of the semester. But I want you to know that I believe in you, and I have confidence in your abilities to succeed.

You have all worked incredibly hard throughout the semester, and your dedication and commitment to your studies have not gone unnoticed. Remember that this is your chance to show what you have learned, and to demonstrate your knowledge and understanding of the material.

I encourage you to take some time to review your notes, organize your thoughts, and practice any areas that you feel less confident in. And most importantly, take care of yourselves both physically and mentally in these final days leading up to the exams.

Remember that I am here to support you, and I want to see you all succeed. You have what it takes to do well on these exams, and I wish you all the best of luck.

Sincerely,

[Your Name]

UC San Diego