THE ENERGY IMPACT OF NETWORKED SERVICES

Module 5
Fall 2020

George Porter

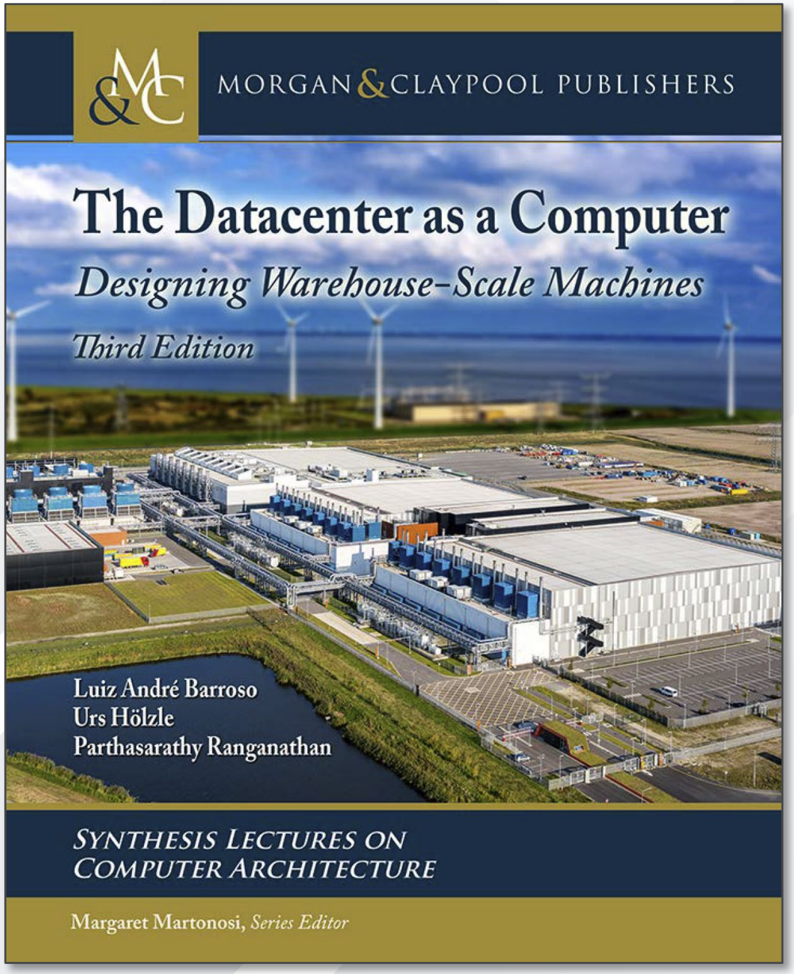
UC San Diego

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.

UC San Diego

Today's material available in canvas



POWER-PROPORTIONAL HUMANS

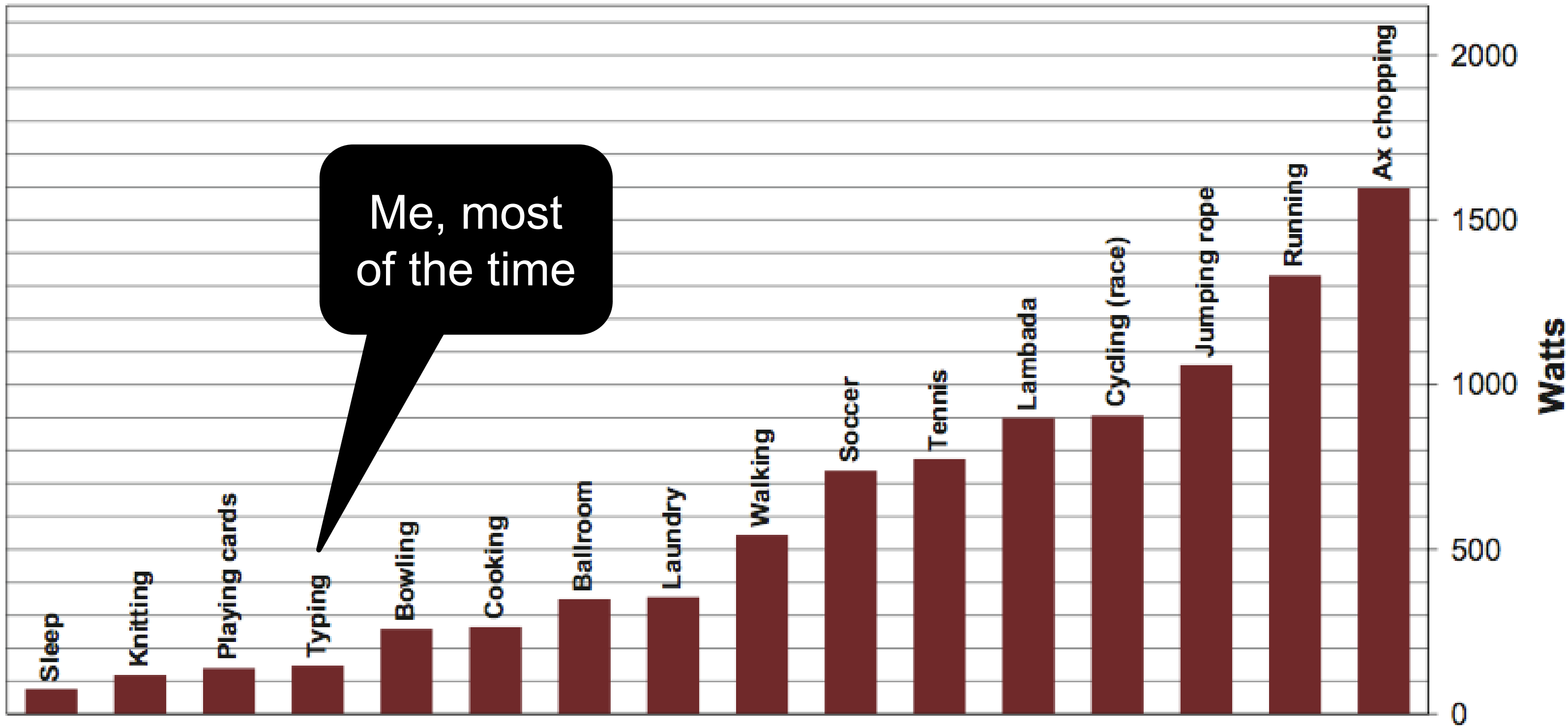


FIGURE 5.7: Human energy usage vs. activity levels (adult male) [52].

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

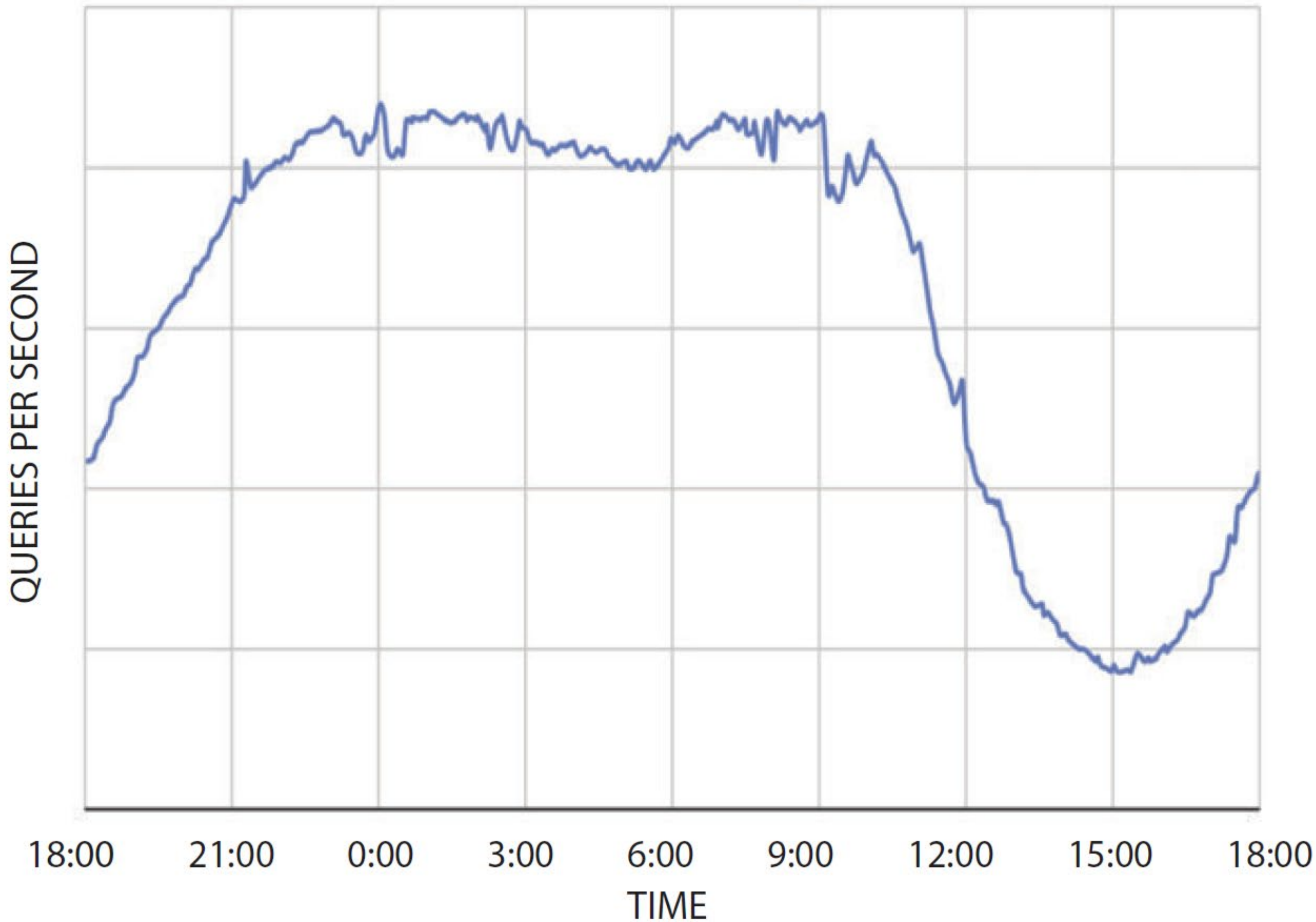


Figure 2.4: Example of daily traffic fluctuation for a search service in one data center over a 24-hr period.

SERVER HARDWARE

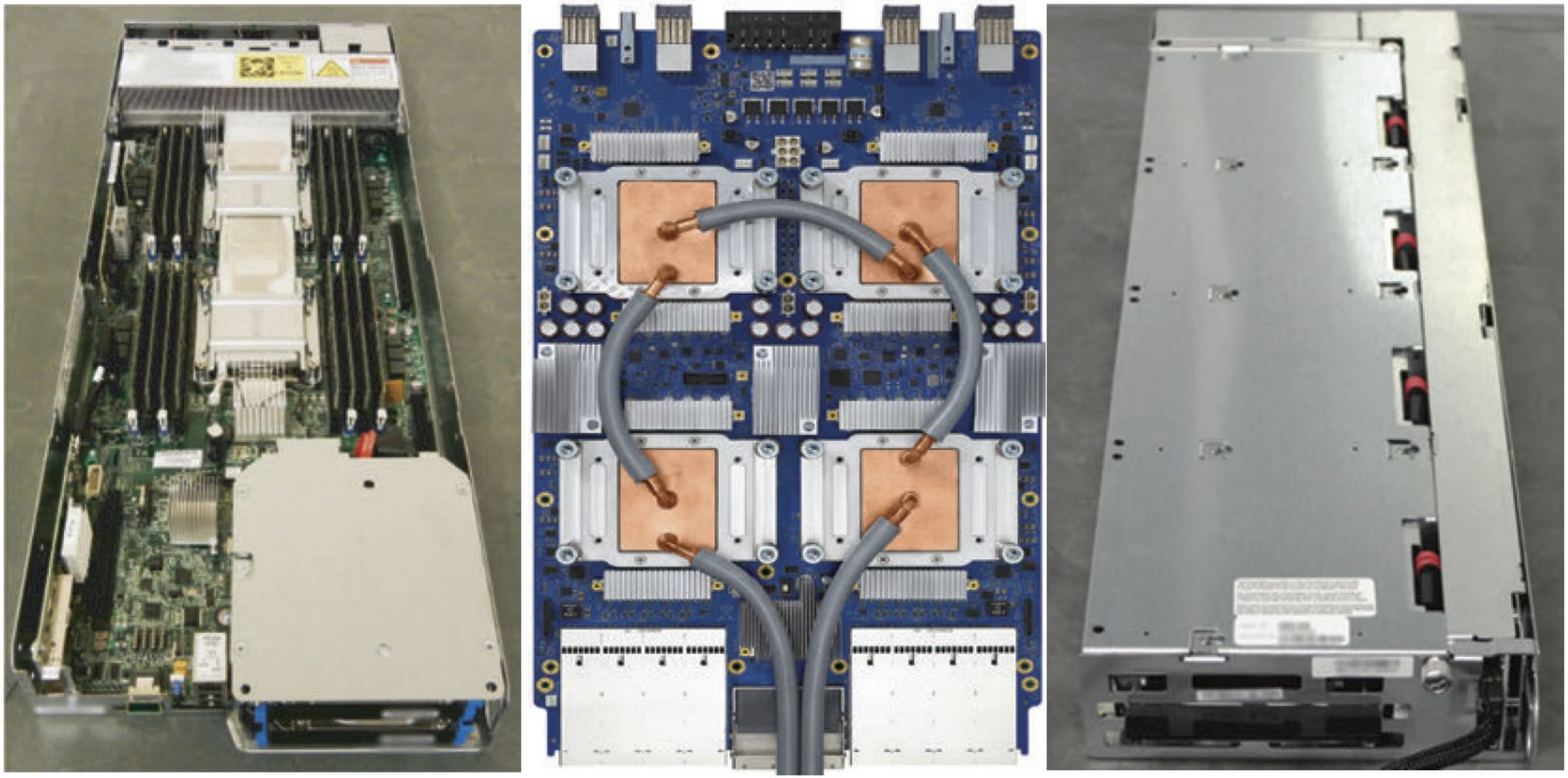


Figure 1.1: Example hardware building blocks for WSCs. Left to right: (a) a server board, (b) an accelerator board (Google's Tensor Processing Unit [TPU]), and (c) a disk tray.

DO DIFFERENT COMPONENTS SCALE SIMILARLY?

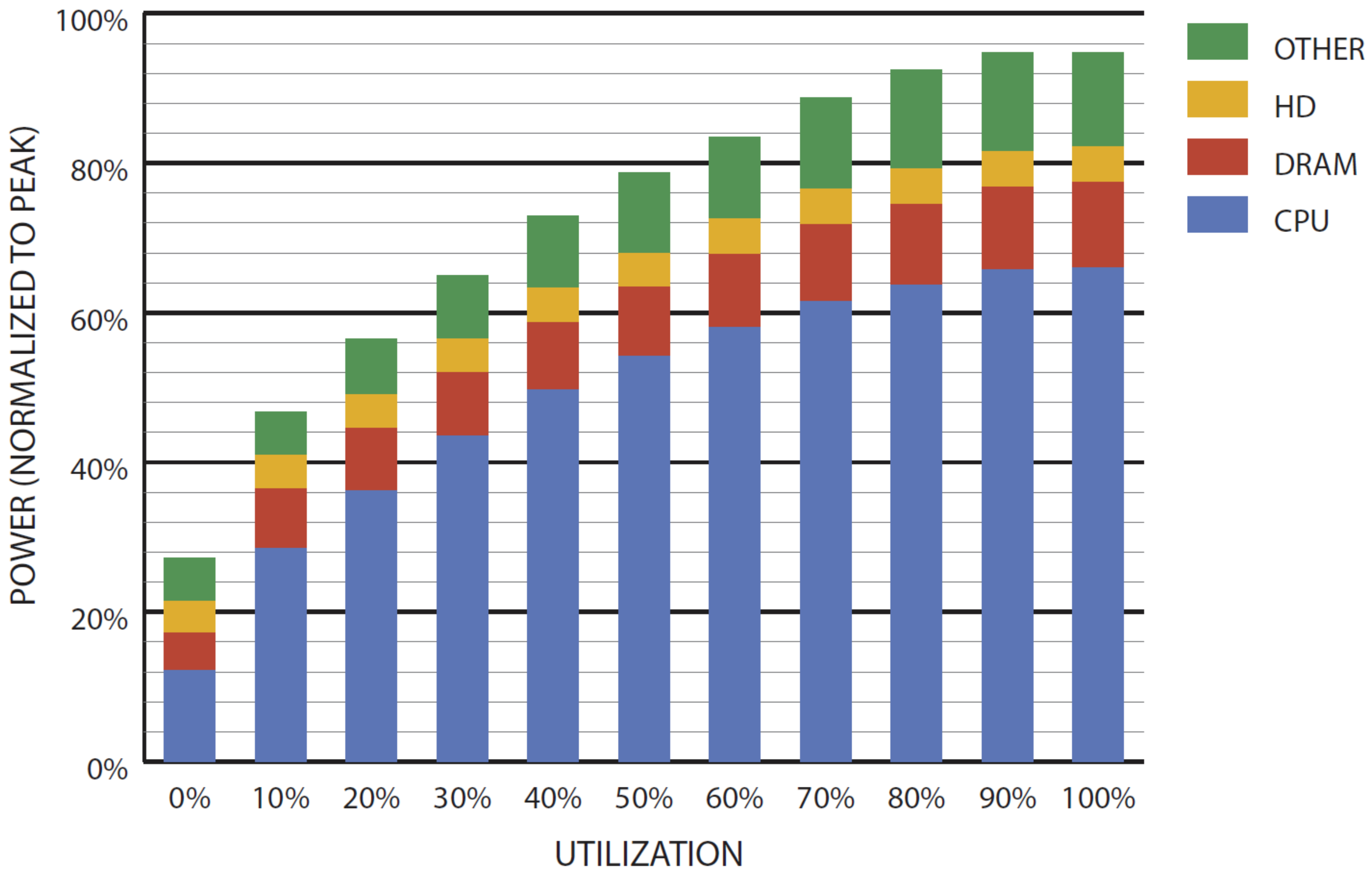


Figure 5.7: Subsystem power usage in an x86 server as the compute load varies from idle to full usage.

CPU UTILIZATION

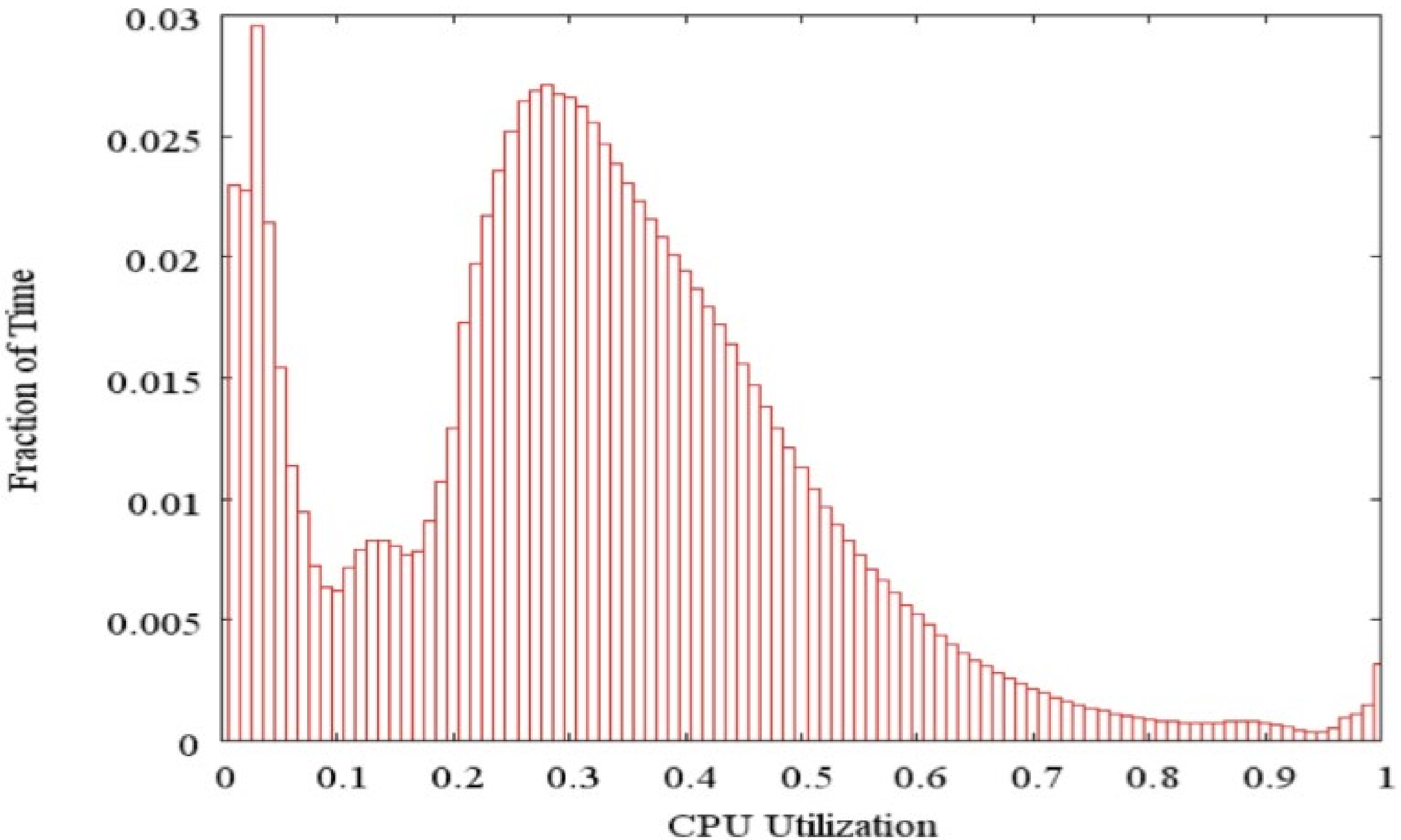


FIGURE 5.5: Activity profile of a sample of 5,000 Google servers over a period of 6 months.

 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?

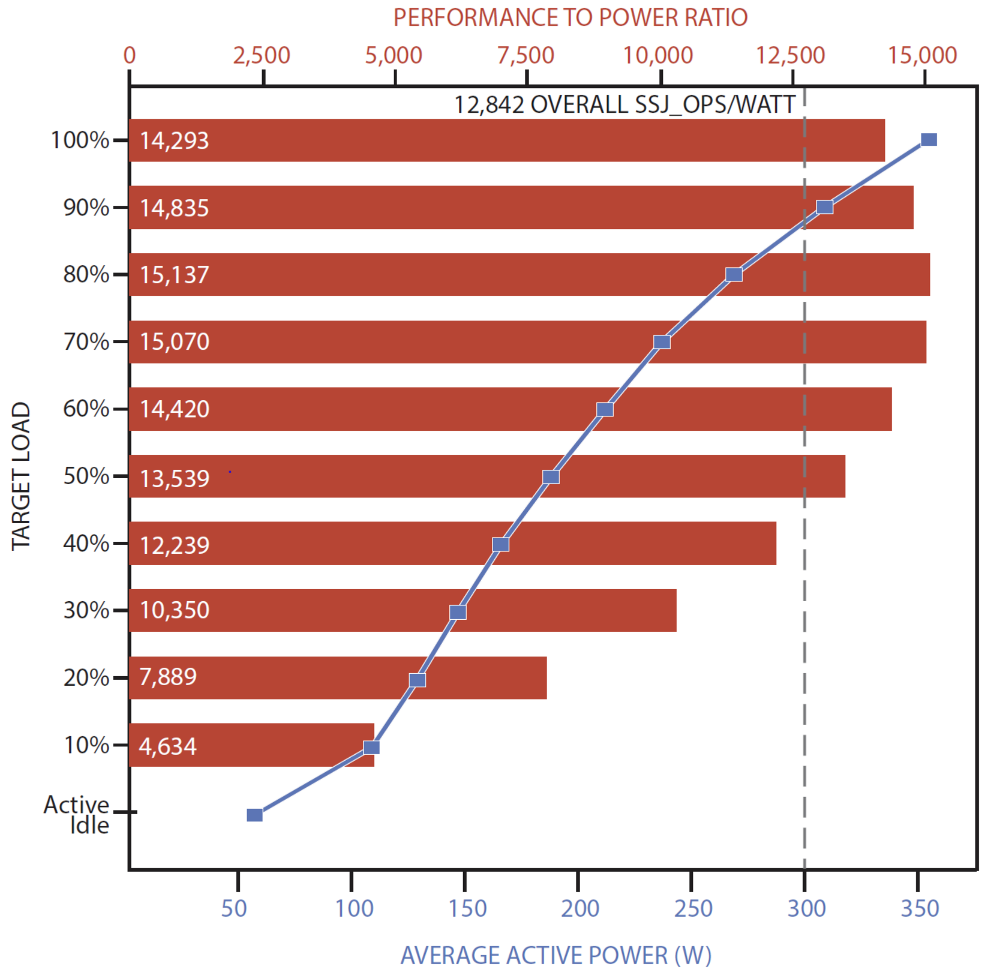


Figure 5.4: Example benchmark result for SPECpower_ssj2008; bars indicate energy efficiency and the line indicates power consumption. Both are plotted for a range of utilization levels, with the average energy efficiency metric corresponding to the vertical dark line. The system has two 2.1 GHz 28-core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

INCREASING POWER PROPORTIONALITY OVER TIME

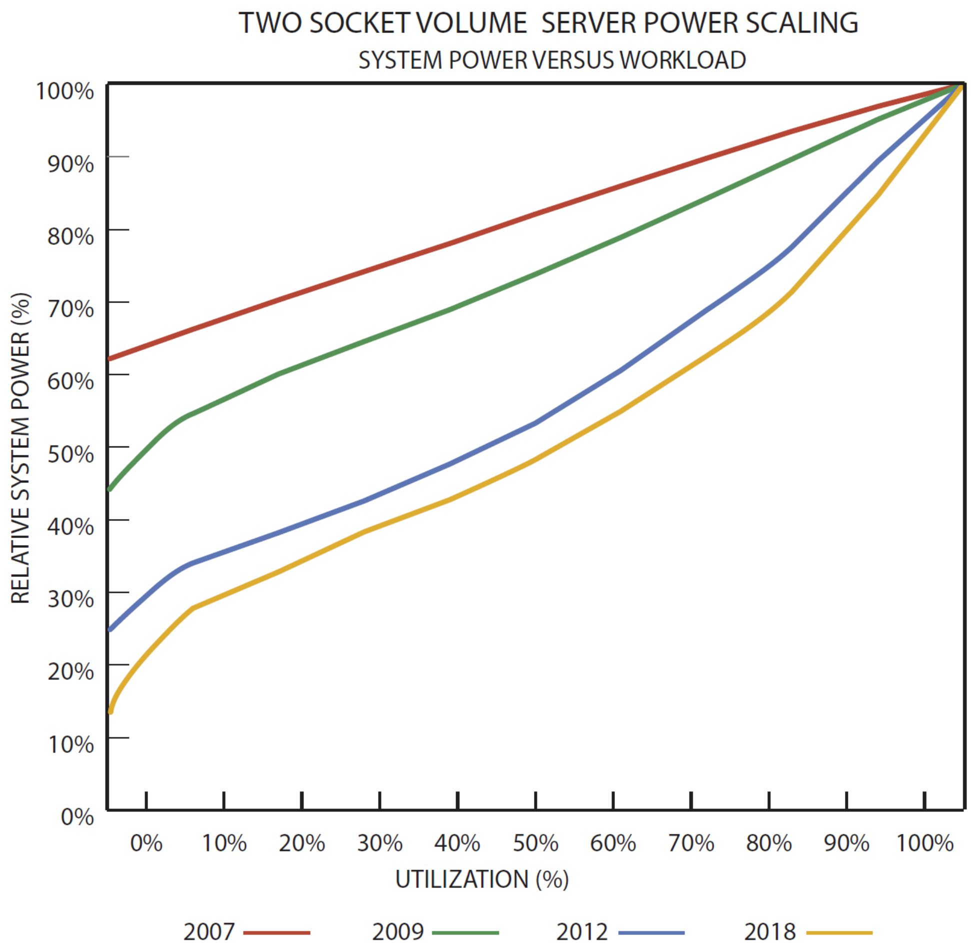


Figure 5.8: Normalized system power vs. utilization in Intel servers from 2007–2018 (courtesy of David Lo, Google). The chart indicates that Intel servers have become more energy proportional in the 12-year period.

EXAMPLE: “DATA PARALLEL” PROCESSING

```
for (iterations = 1 to 100){
   work = new Task[10000];
   for (i = 1 to 10000) {
      rpc.asyncCall(serve  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”?

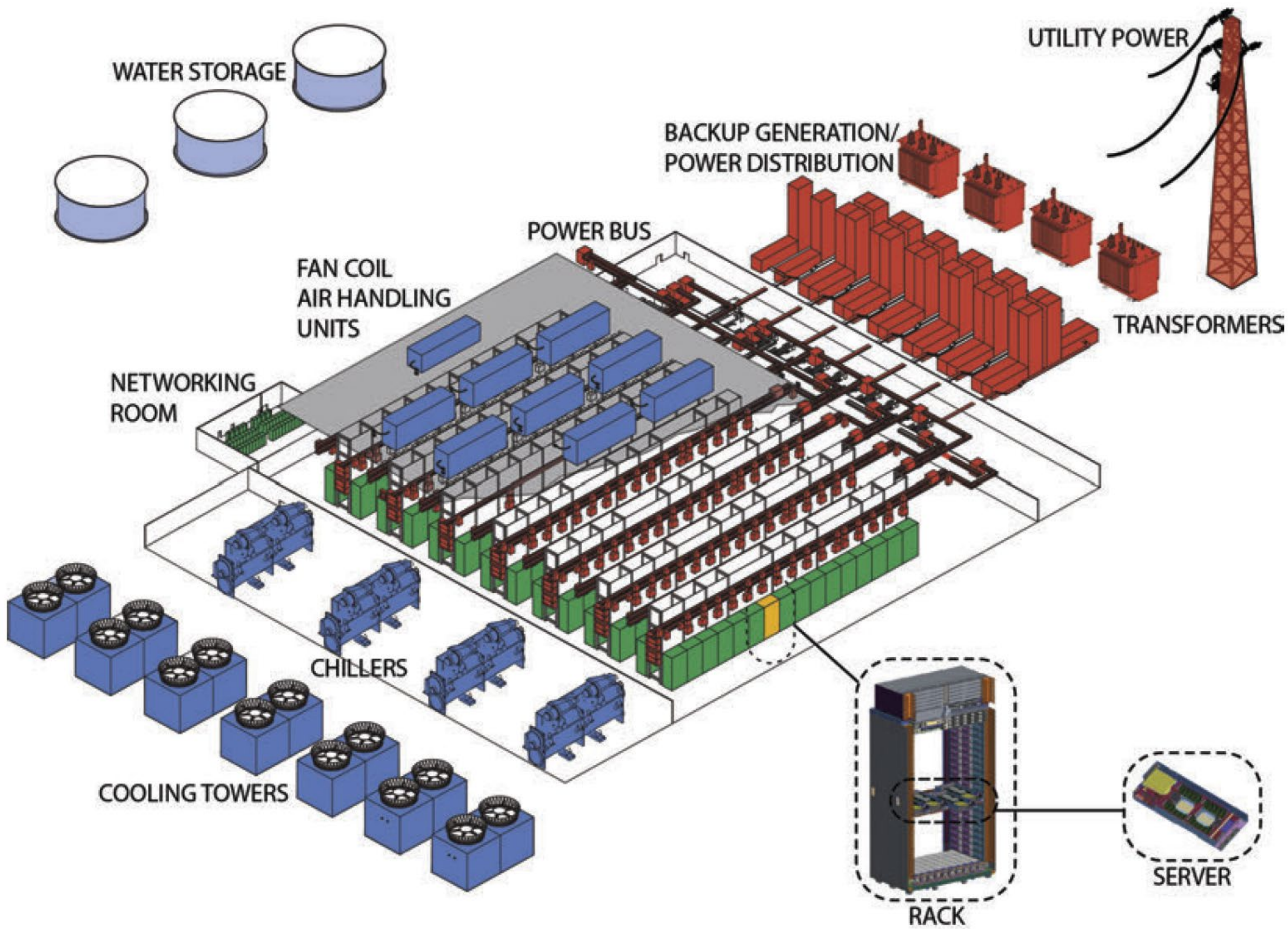


Figure 4.4: The main components of a typical data center.

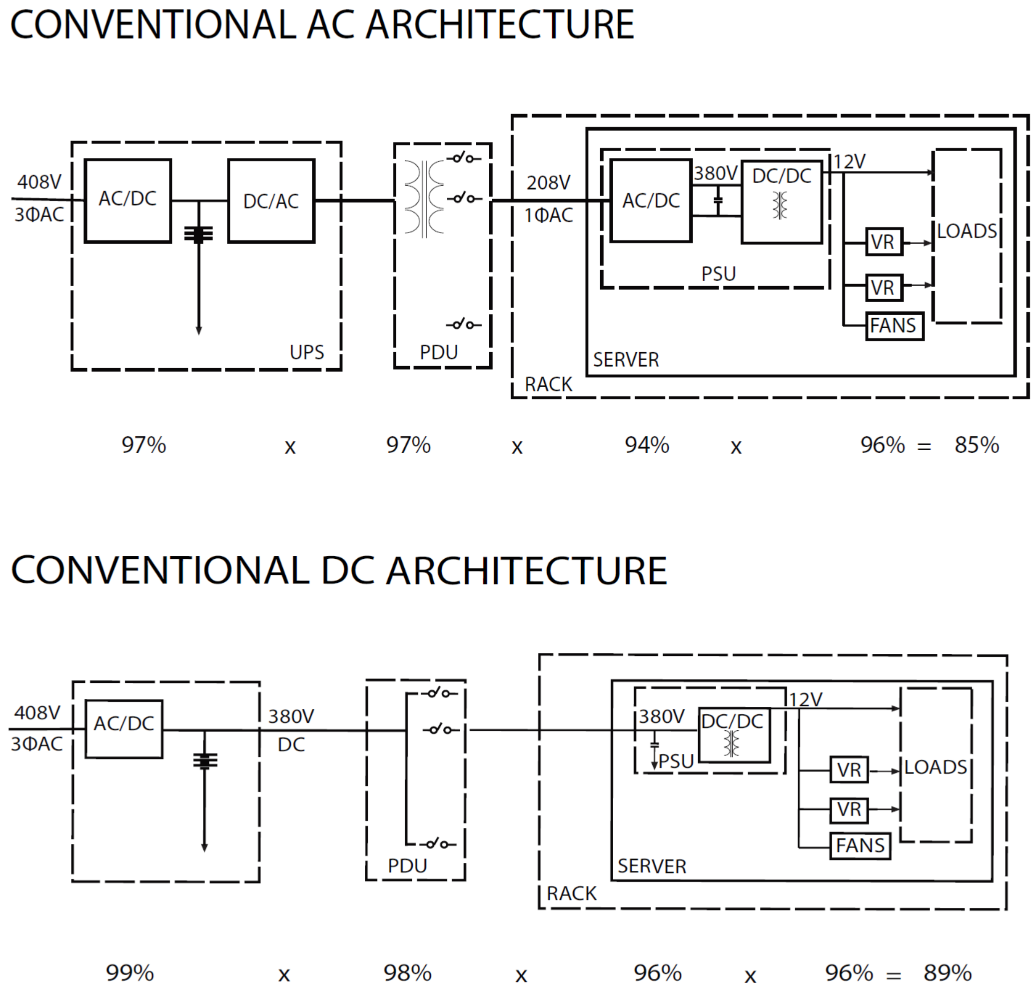


Figure 4.5: Comparison of AC and DC distribution architectures commonly employed in the data center industry.

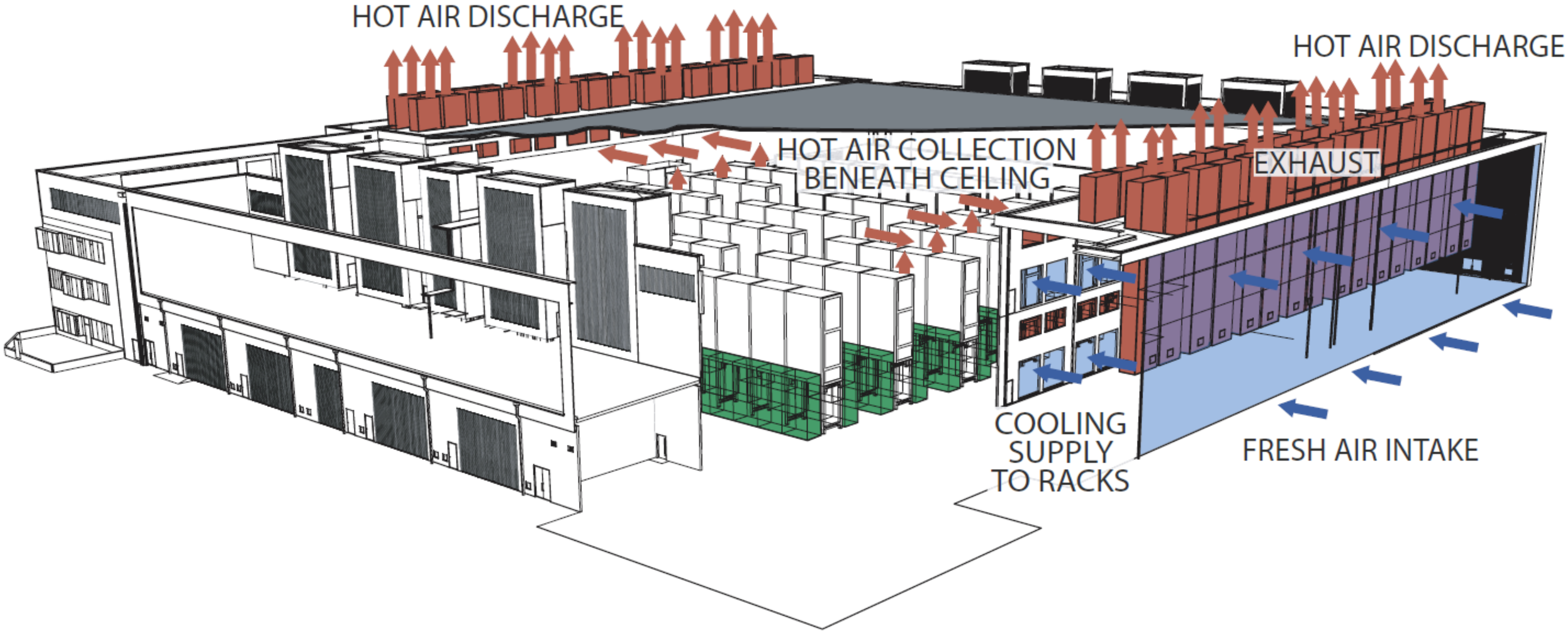


Figure 4.8: Airflow schematic of an air-economized data center.

ENERGY “OVERHEAD”: A/C AND COOLING

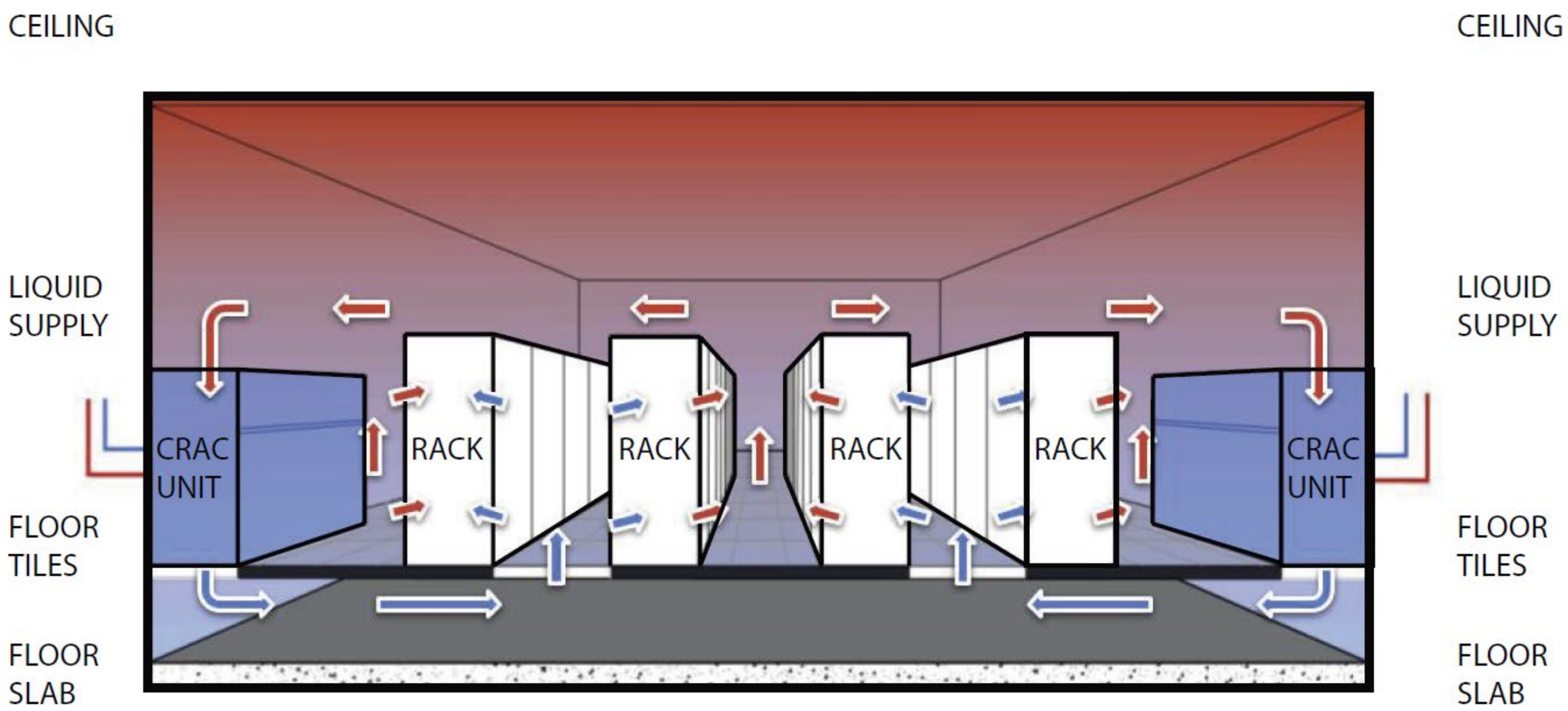


Figure 4.9: Raised floor data center with hot-cold aisle setup (image courtesy of DLB Associates [Dye06]).

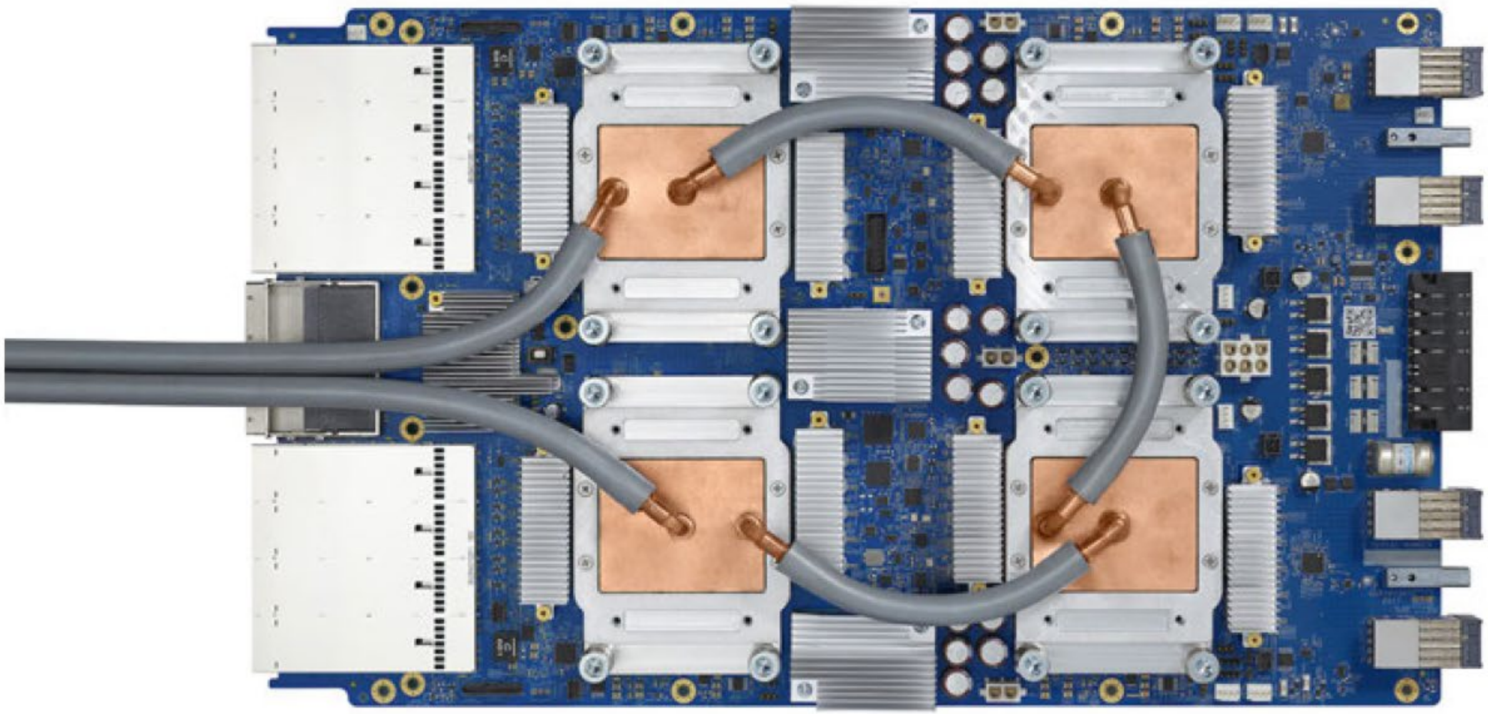
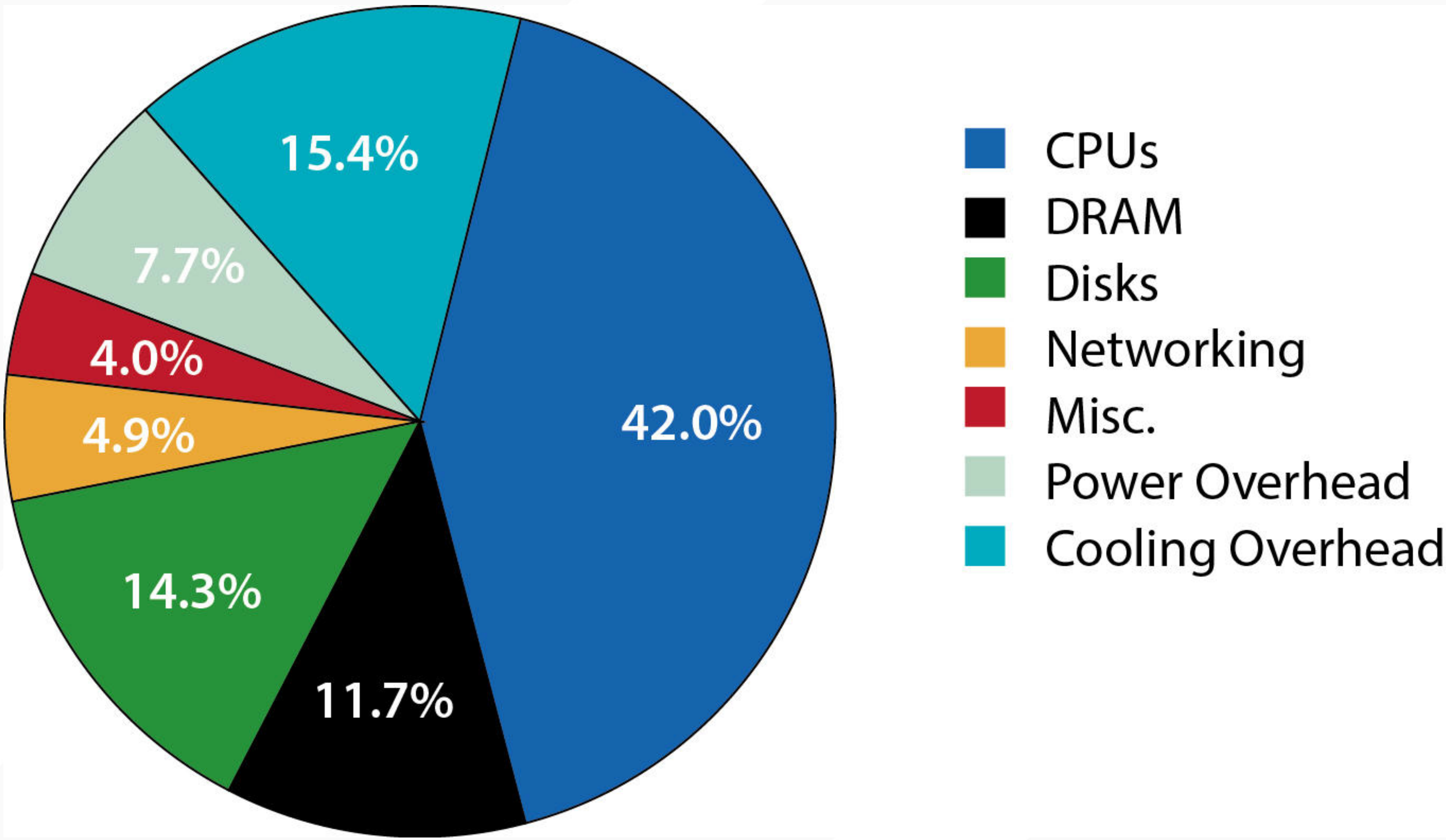


Figure 4.16: Copper cold plates and hose connections provide liquid cooling for Google's third-generation TPU.

NUMBERS FROM JAMES HAMILTON (MSFT, AMAZON)



GOOGLE, ~2012 OR SO

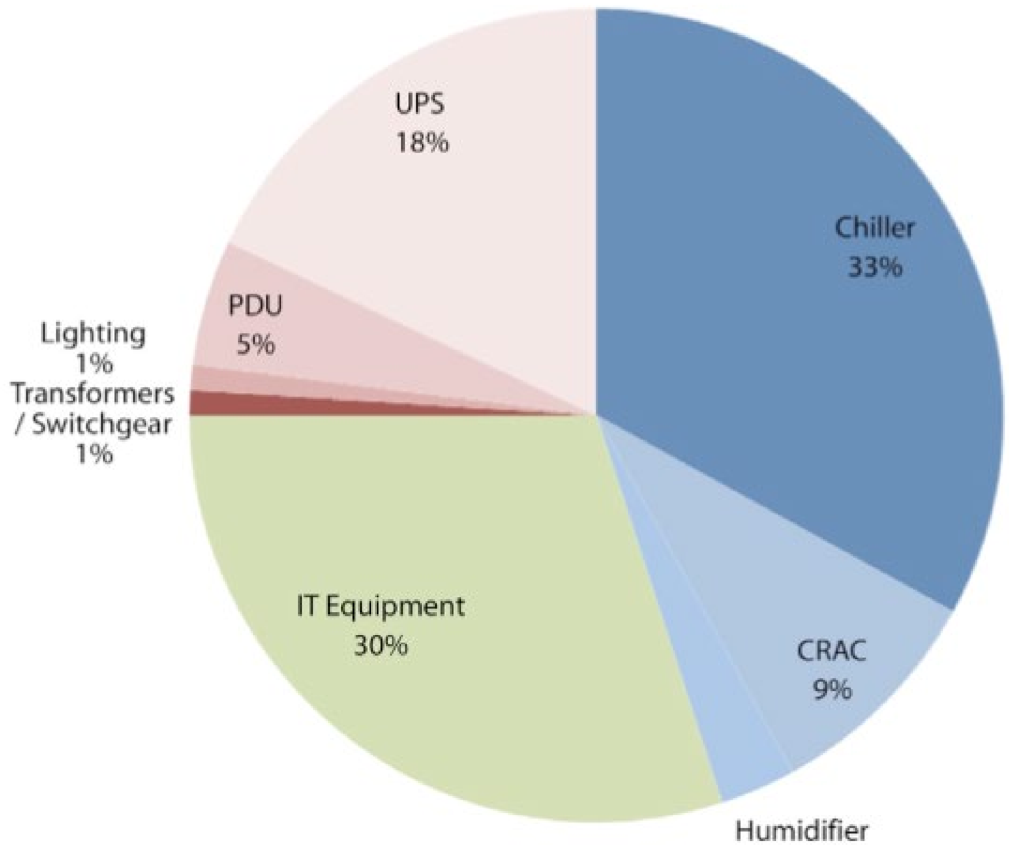


FIGURE 5.2: Breakdown of datacenter energy overheads (ASHRAE).

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

GOOGLE, ~2016

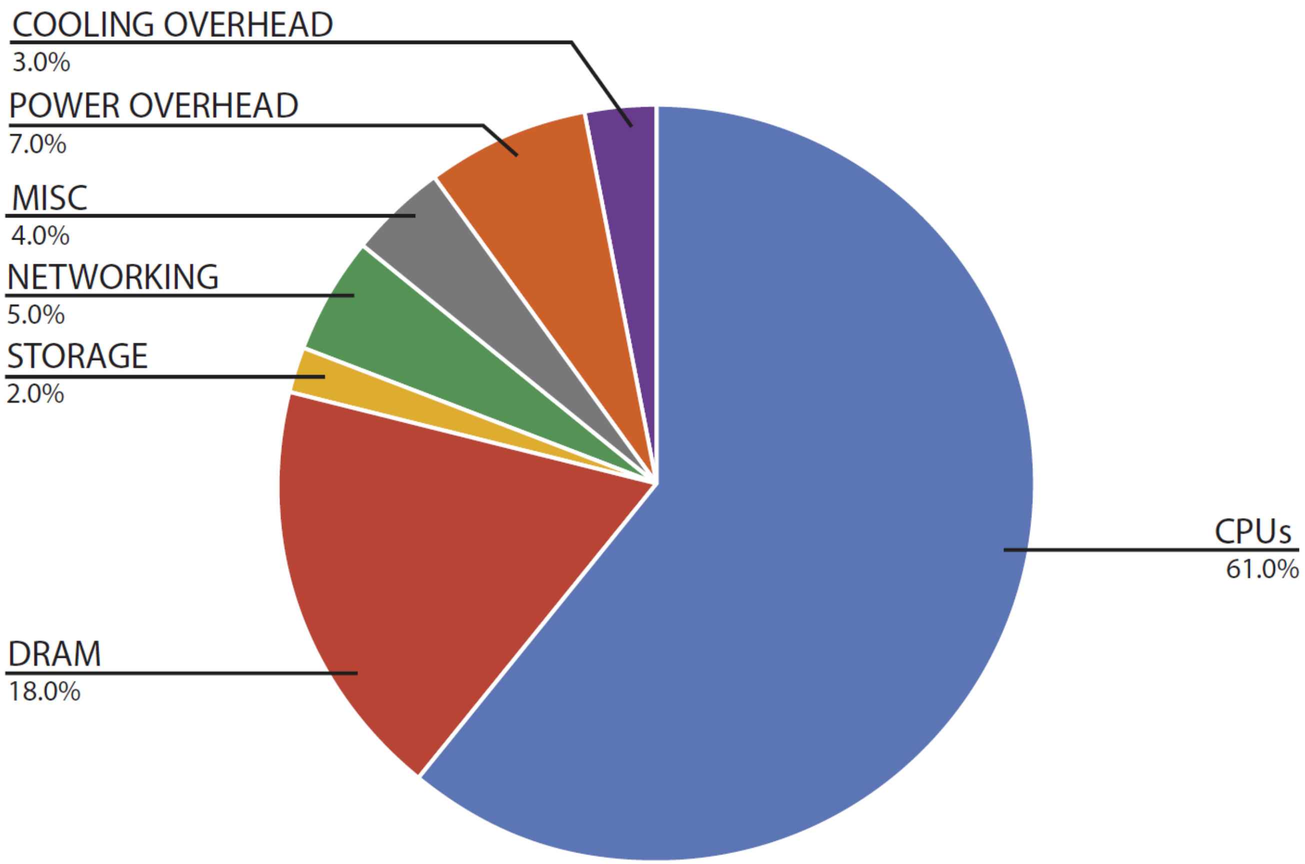


Figure 1.8: Approximate distribution of peak power usage by hardware subsystem in a modern data center using late 2017 generation servers. The figure assumes two-socket x86 servers and 12 DIMMs per server, and an average utilization of 80%.

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)

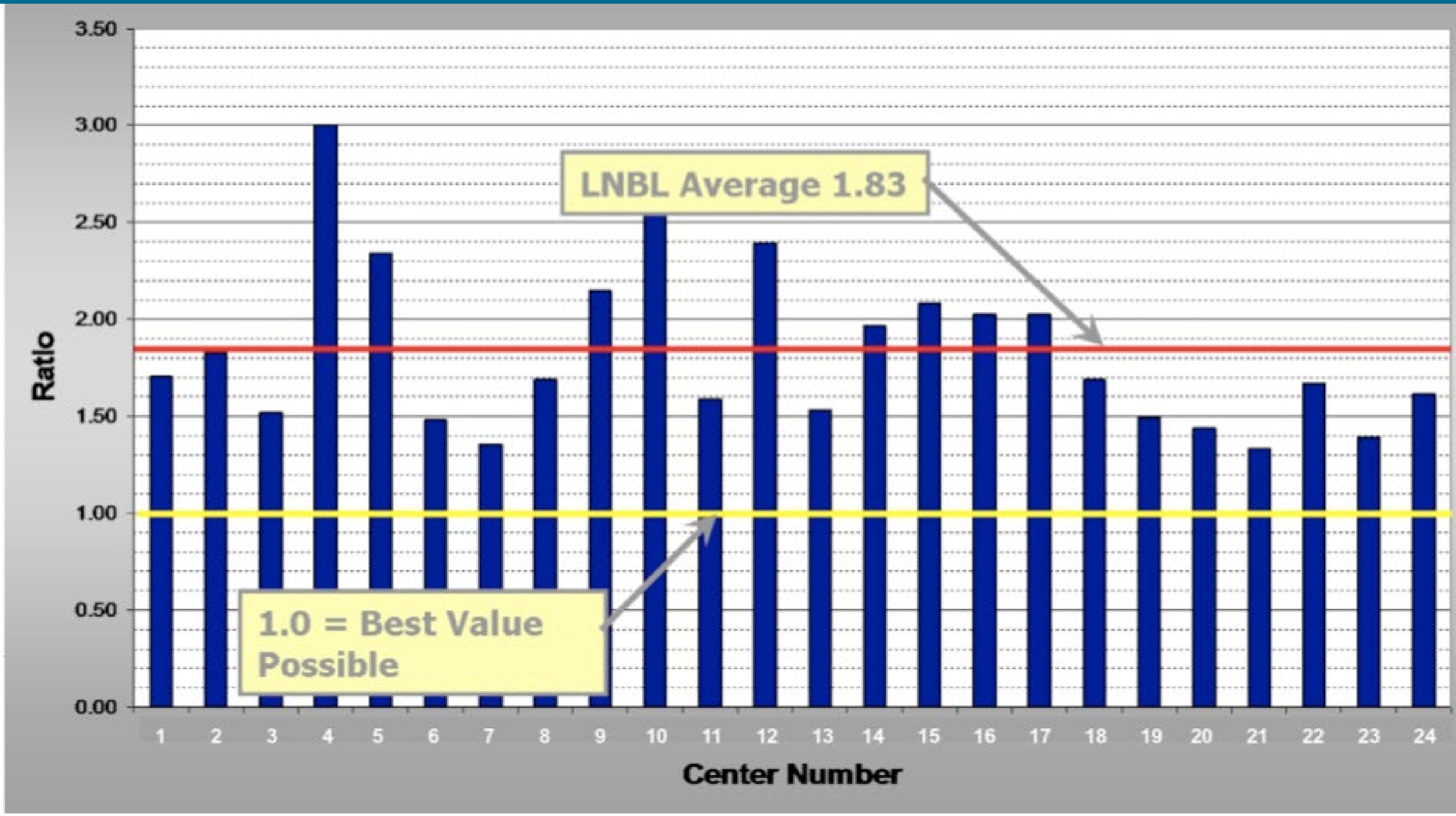


FIGURE 5.1: LBNL survey of the power usage efficiency of 24 datacenters, 2007 (Greenberg et al.) [41].

 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

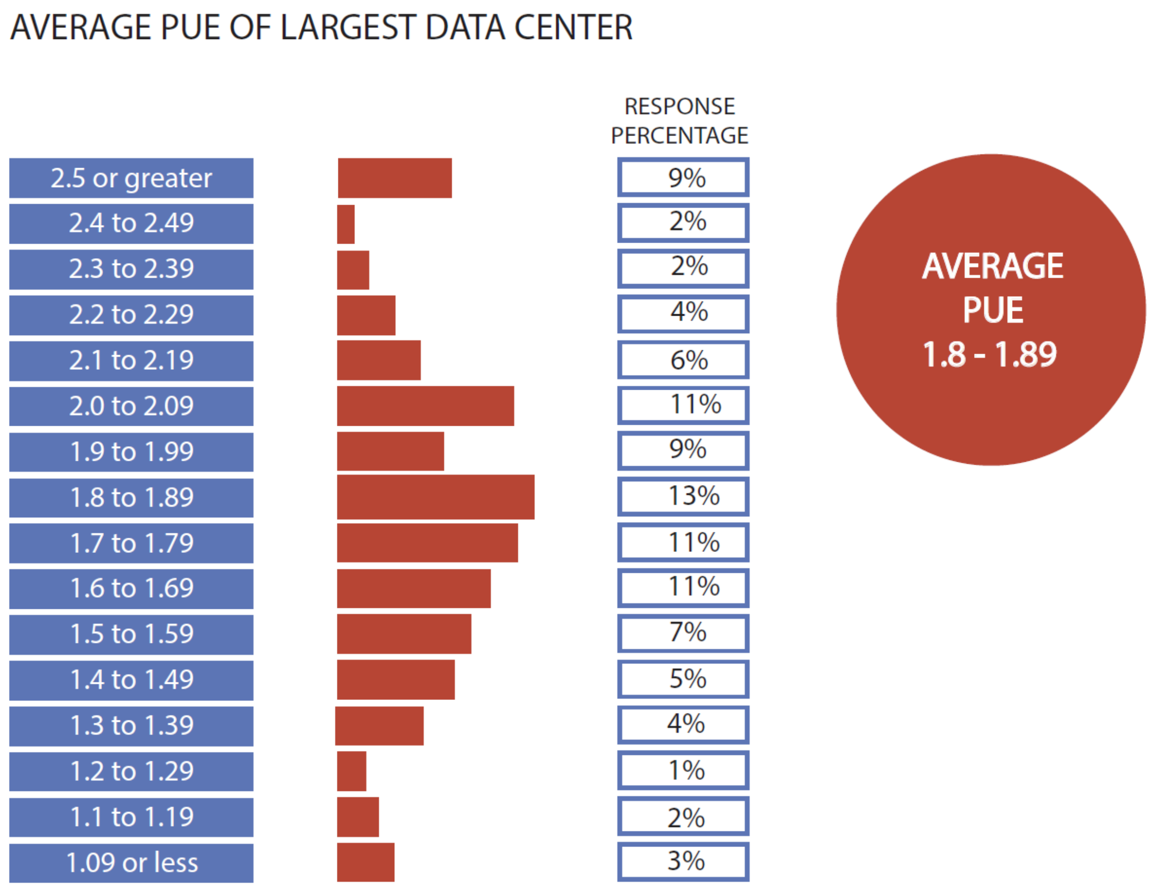


Figure 5.1: Uptime Institute survey of PUE for 1100+ data centers. This detailed data is based on a 2012 study [UpI12] but the trends are qualitatively similar to more recent studies (e.g., 2016 LBNL study [She+16]).

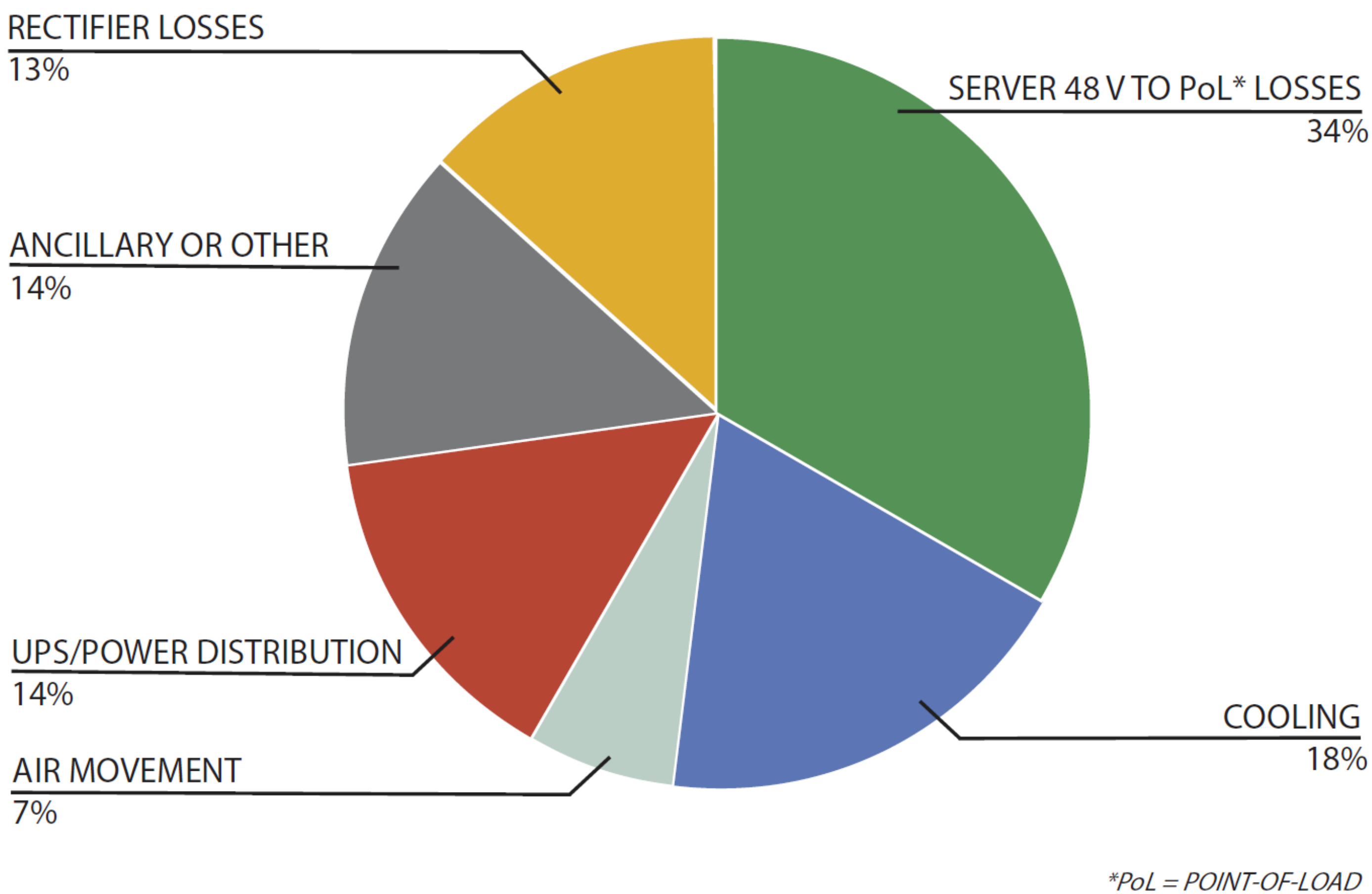


Figure 5.3: A representative end-to-end breakdown of energy losses in a typical datacenter. Note that this breakdown does not include losses of up to a few percent due to server fans or electrical resistance on server boards.

Course review

- What have we covered in the last 10 weeks?

COURSE CONCEPTS

- Networking fundamentals
  - IP addresses, hostnames
  - Layering and protocols
  - Framing and parsing
- Socket programming
  - Sockets API
  - Semantics of the calls
  - When does blocking occur? How does the client "know" to send at a rate that the other endpoint can handle?
  - Server socket vs client socket
- Domain name system
- Performance
  - Bandwidth, latency
- HTTP protocol
- Concurrency and parallelism
- Network file system (NFS)
  - Caching / cache consistency

COURSE CONCEPTS

- Remote procedure calls
  - Stub compiler
  - Lifecycle of an RPC call
  - XML-RPC
  - Failure semantics
    - Maybe vs at-most vs at-least
- Physical time
  - Time sources
  - Cristian's algorithm
  - Berkeley algorithm
  - Network time protocol
- Logical time
  - Lamport clock rules
  - Vector clock rules
  - Message board example
  - Totally ordered multicast (TOM)
- Transactions
  - Two-phase commit
  - Safety and liveness
- DynamoDB
  - Eventually consistent databases
- Virtualization and cloud platforms

COURSE CONCEPTS

- Video
  - MPEG-DASH
  - MPD file (high-level, not low-level details)
- CDNS, Akamai
- Availability and performance
- “Tail at scale” article
- Overlay networks
- Chord problem/homework
- Power-proportional computing
- PUE

FINAL MESSAGE TO YOU

- You now have a strong foundation for "backend/server" computing
  - Cloud, distributed systems, client-server, …
- Good luck on all your final exams
  - Please take care of yourself though!
  - Reach out if you need help/support
- Thanks for the great class!

UC San Diego