CSCI 381/780
Cloud Computing

Data Center

Jun Li
Queens College

QUEENS
COLLEGE

Data center in cloud computing

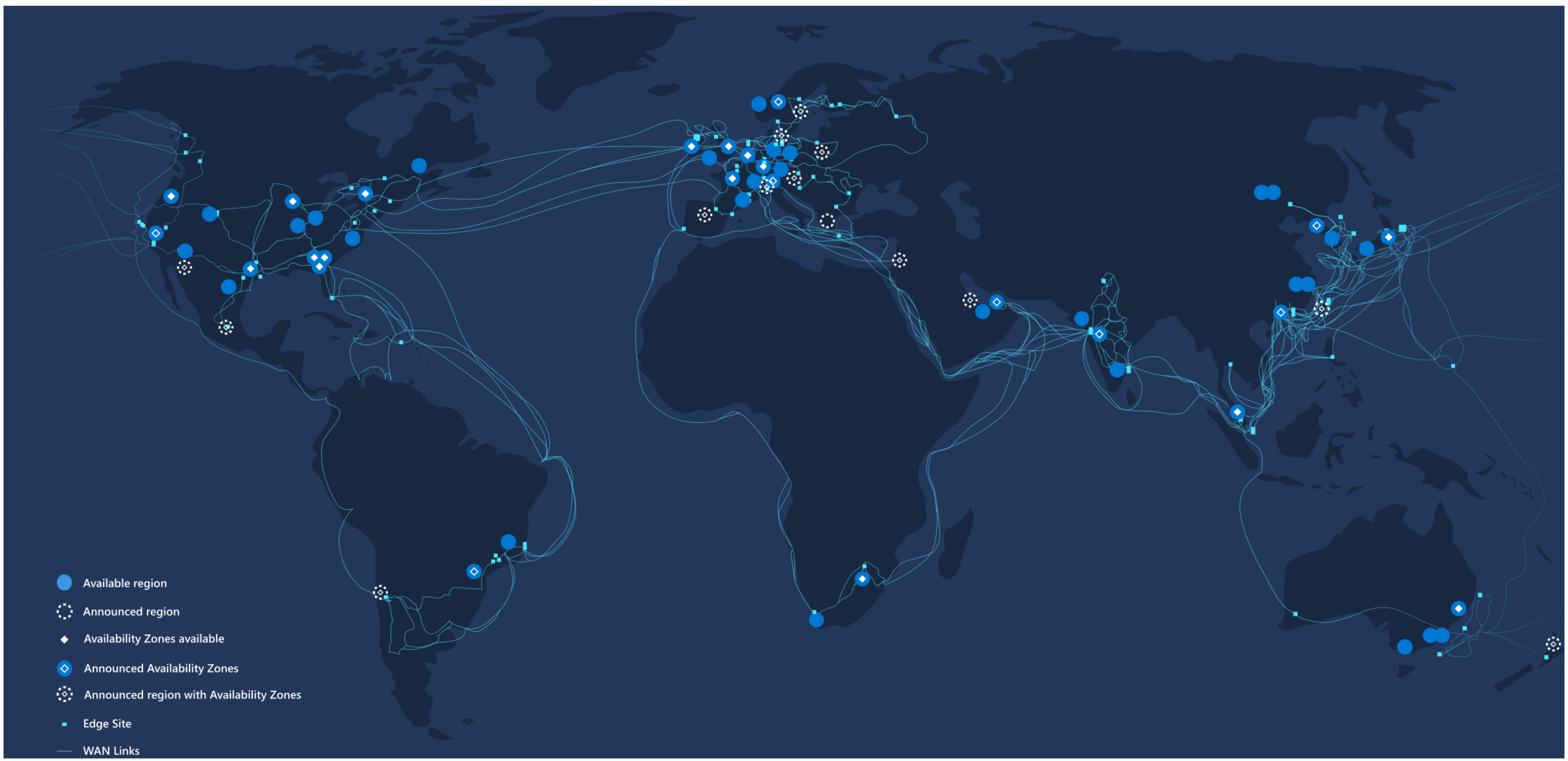
Azure's global data center

- As of 2021, Azure has 160+ data centers, arranged into regions, and linked by one of the largest interconnected networks on the planet.

- Data centers are usually located in places where electricity is cheap, or cooling is free, to keep costs down, and often using their own custom hardware and topologies to ensure high efficiency.

Azure’s global data center

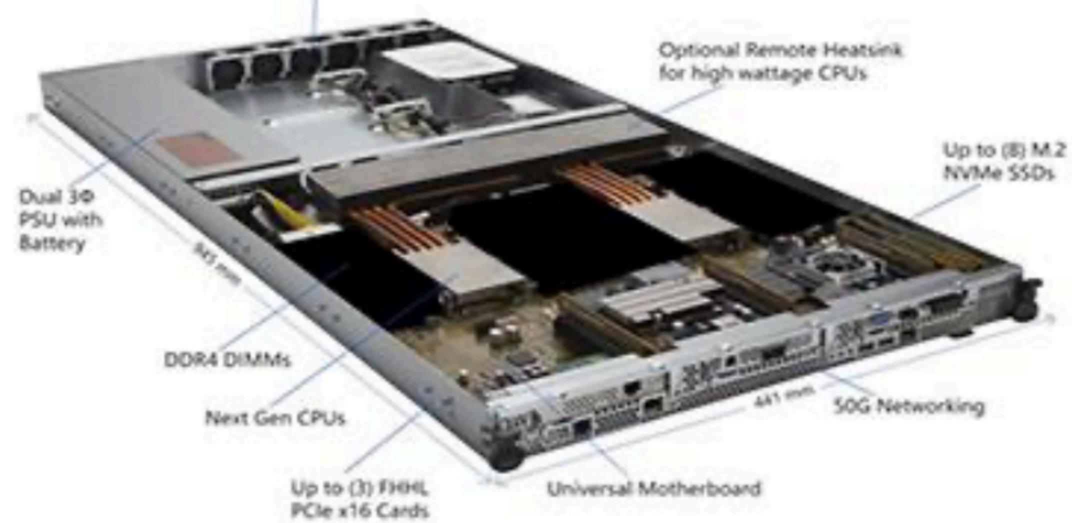
- As of 2021, Azure has 160+ data centers, arranged into regions, and linked by one of the largest interconnected networks on the planet.



Microsoft's data center

Microsoft's Chicago Data Center
(2,000 servers per container)

12-volt power supply



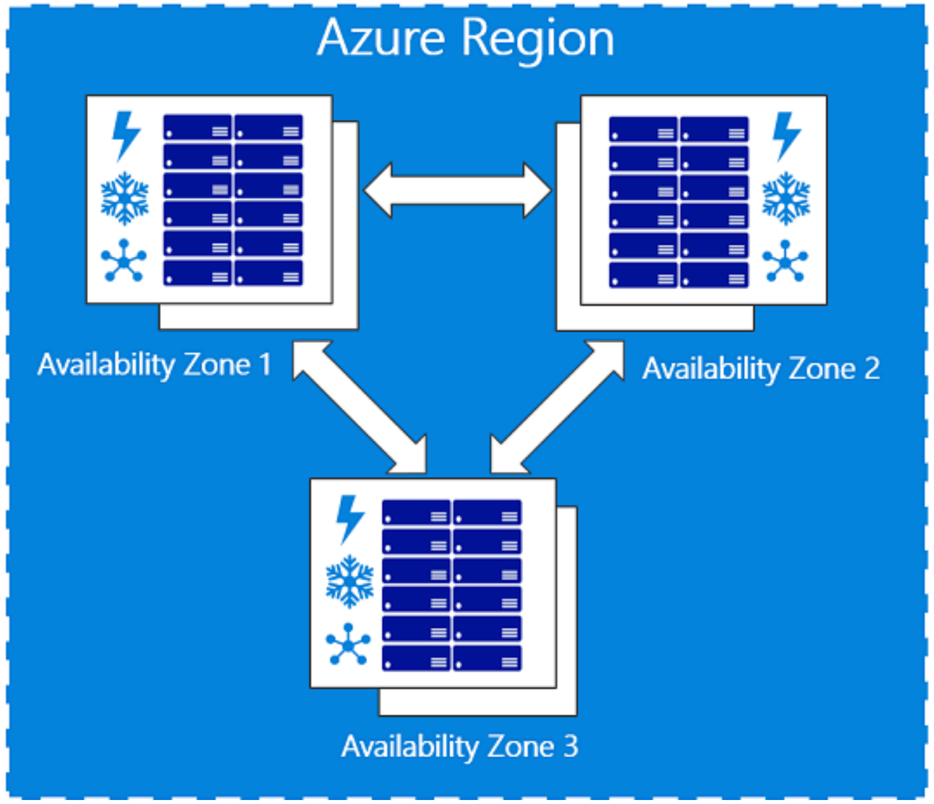
Nov. 2, 2016: **Microsoft Project Olympus: Releases Cloud Hardware Specs**

Availability Zones

- Availability Zones are **unique physical locations** within an Azure region. Each zone is made up of **one or more data centers** equipped with **independent power, cooling, and networking**. To ensure resiliency, there's a minimum of three separate zones in all enabled regions.

- The physical separation of Availability Zones within a region protects applications and data from **datacenter failures**. Zone-redundant services replicate your applications and data across Availability Zones to protect from single-points-of-failure.

- You can **synchronously** replicate your applications and data using Availability Zones for **high-availability** and **asynchronously** replicate across Azure regions for **disaster recovery protection**.



Project Natick

Target: long lived, resilient data centers that operate ‘lights out’ —nobody on site; go without maintenance for years at a time.

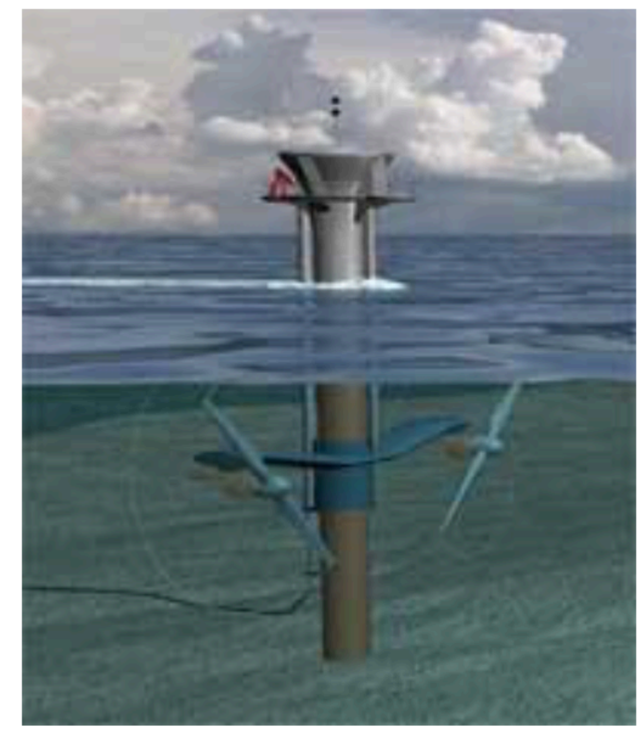
(June 2018) The vessel was deployed with 12 racks containing 864 servers and 27.6 petabytes of storage.

Retrieved for analysis after more than 2 years at the bottom of the ocean

Project Natick

- Power consumption 240 kW. All generated by on-shore wind and solar along with off-shore tide and wave power.

12 racks containing **864 servers with 27.6 petabytes of disk storage**.



Wave & Tidal Energy

Project Natick

Lessons learned?

- This more sustainable approach actually improves the performance and reliability of the datacenter when compared to land. **Project Natick had 1/8th the failure rate of land data centers.** Not only is a greener future possible, but it is economically practical.
  - The atmosphere of **nitrogen**, which is less corrosive than oxygen, and the **absence of people to bump and jostle** components, are the primary reasons for the difference.
- Operating with a highly efficient **PUE** (power usage effectiveness is total power divided by server power; lower values are better, 1.0 is perfect) of **1.07**

Google cloud

35
REGIONS

106
ZONES

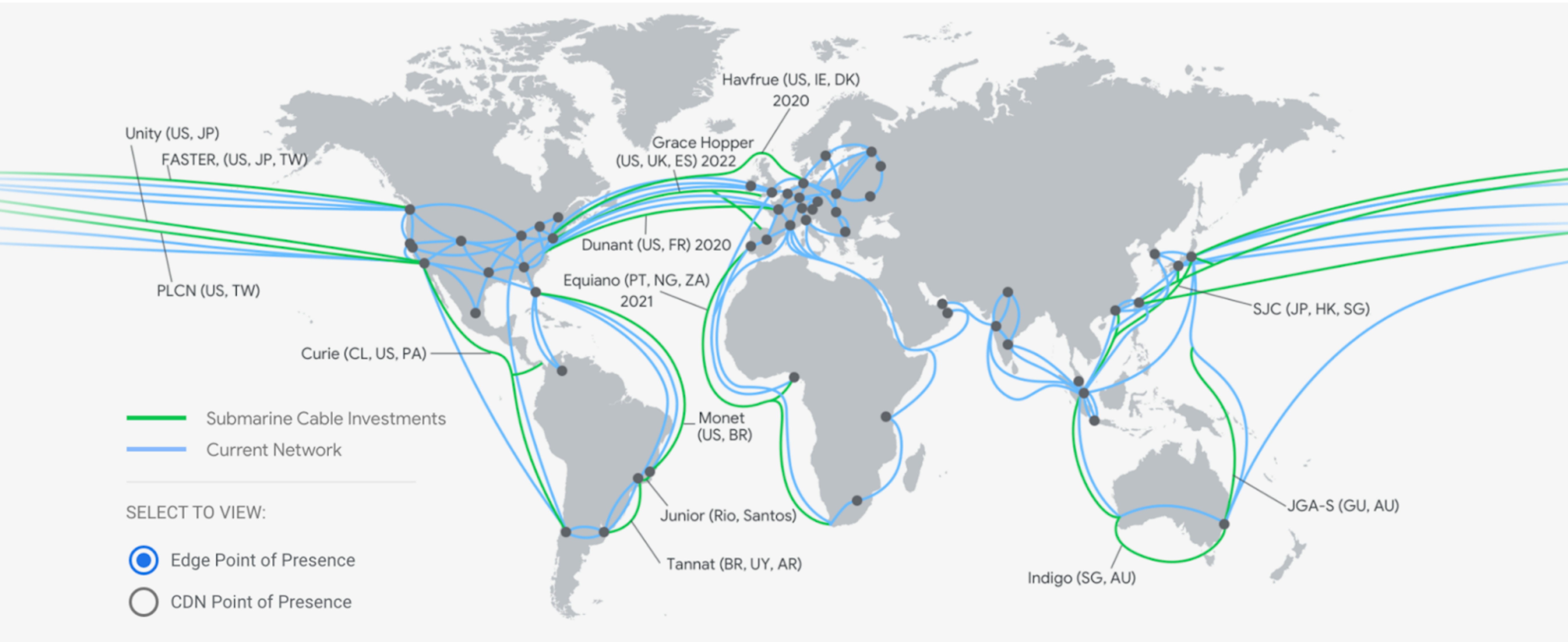
176
NETWORK EDGE LOCATIONS

AVAILABLE IN
200+
COUNTRIES AND TERRITORIES

Google cloud

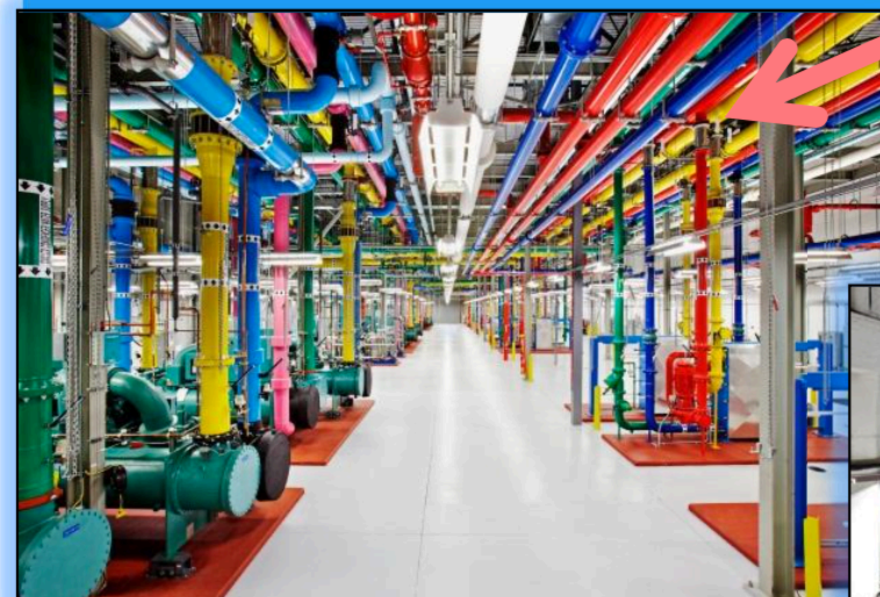
AVAILABLE IN

35	106	176	200+
REGIONS	ZONES	NETWORK EDGE LOCATIONS	COUNTRIES AND TERRITORIES

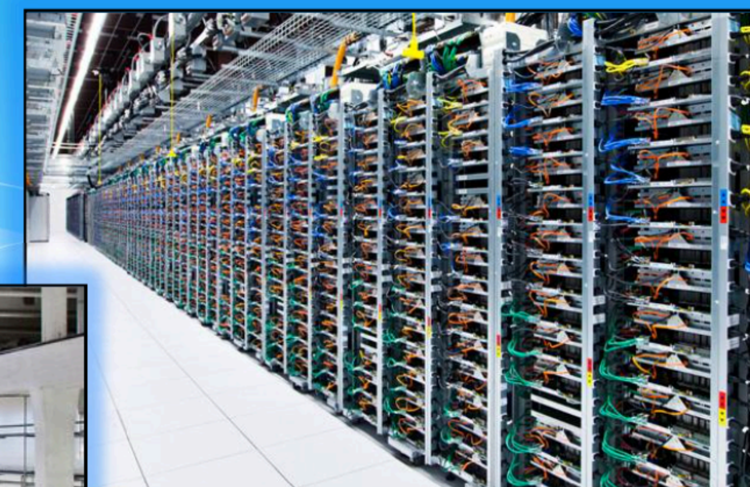


Google’s data center

Blue pipes supply cold water and the red pipes return the warm water back to be cooled !

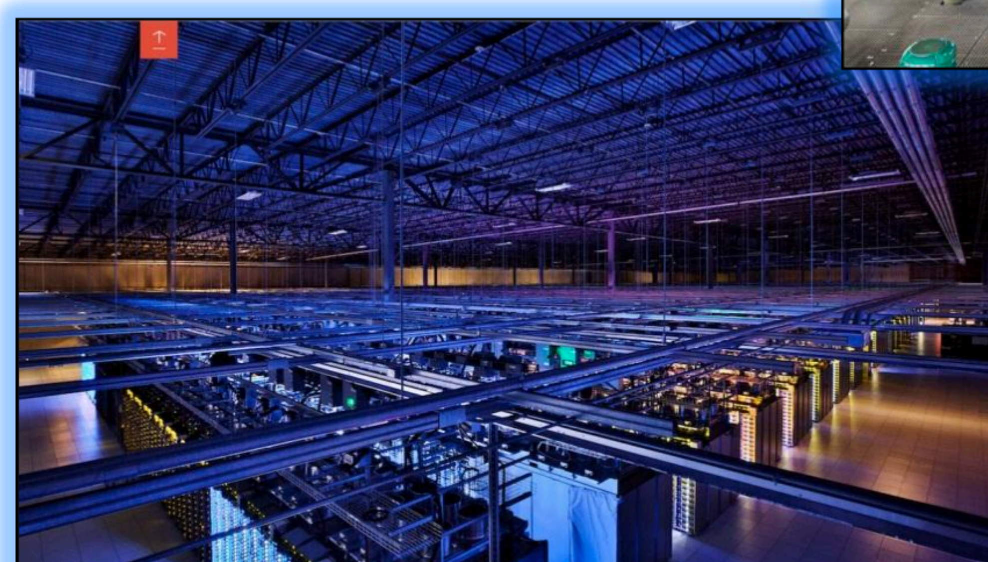


@Douglas County, Georgia (2003)



@Mayes County, Oklahoma (2011)

@Hamina, Finland: Cooling plants, with seawater from the Gulf of Finland

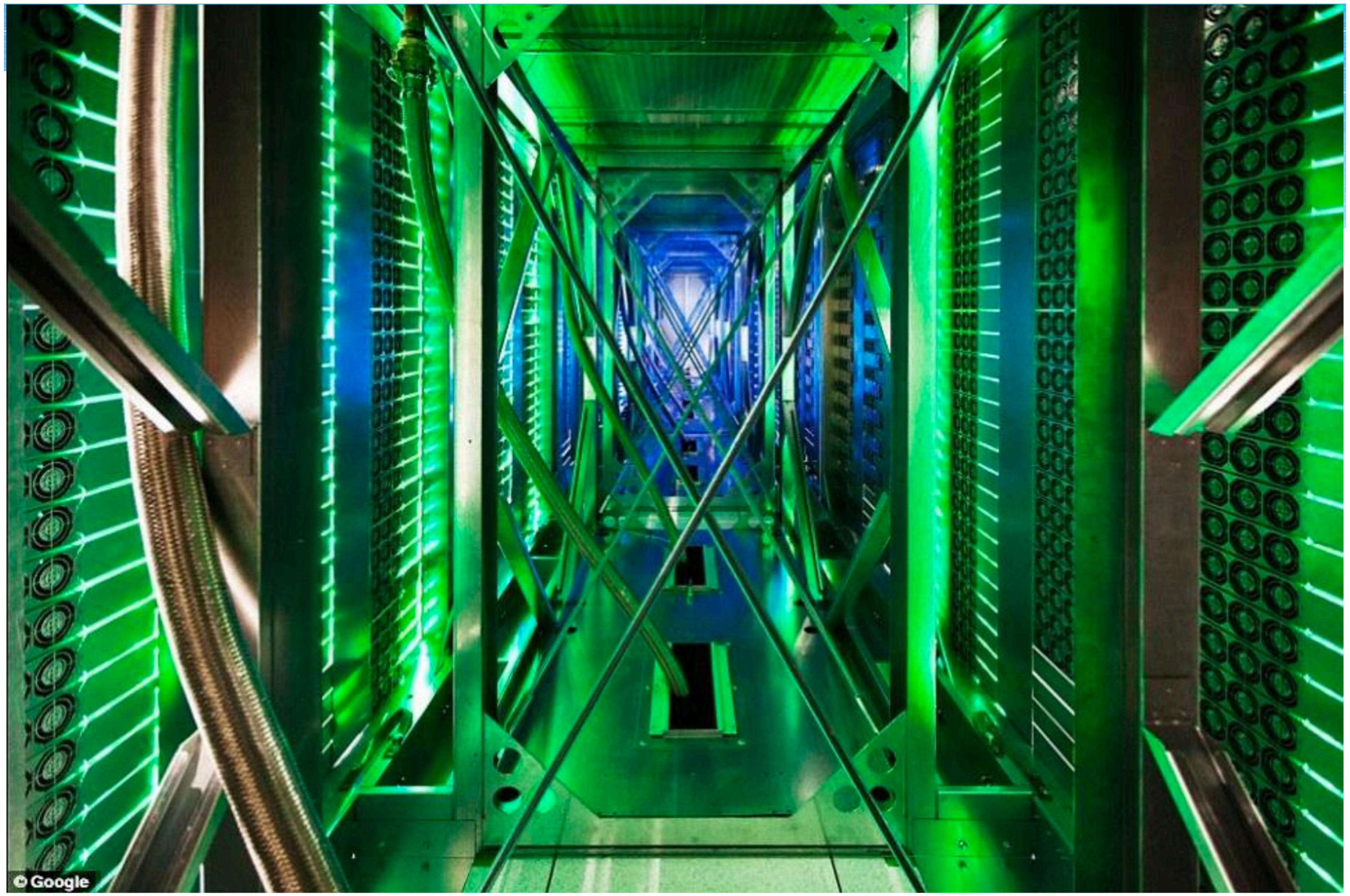


@Council Bluffs, Iowa (2009)

Google’s data centers use around 260 million watts of power which is enough to consistently power 200,000 homes.

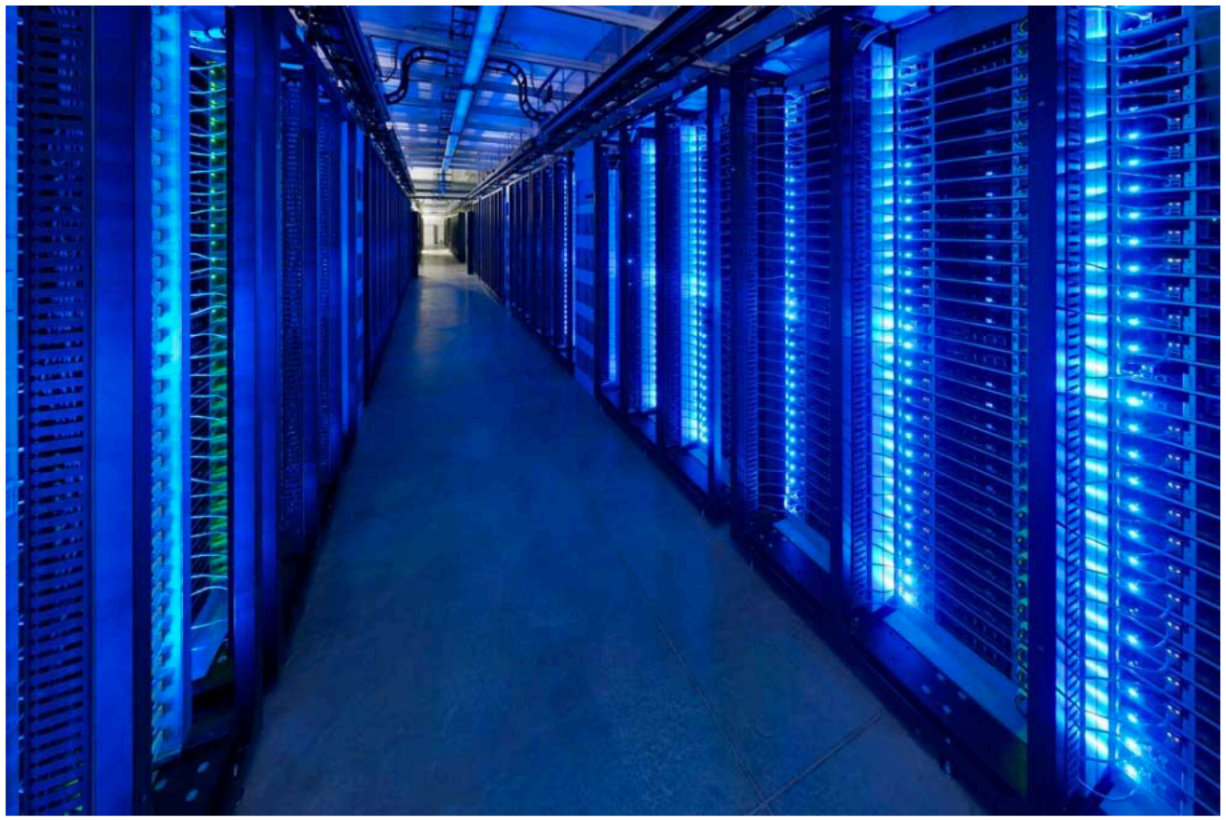
@Belgium: Chiller-less (just fresh air)

Google’s data center @Okalahoma



Hundreds of fans funnel hot air from the server racks into a cooling unit to be recirculated in. The green lights are the server status LEDs reflecting from the front of the servers.

Meta’s data center



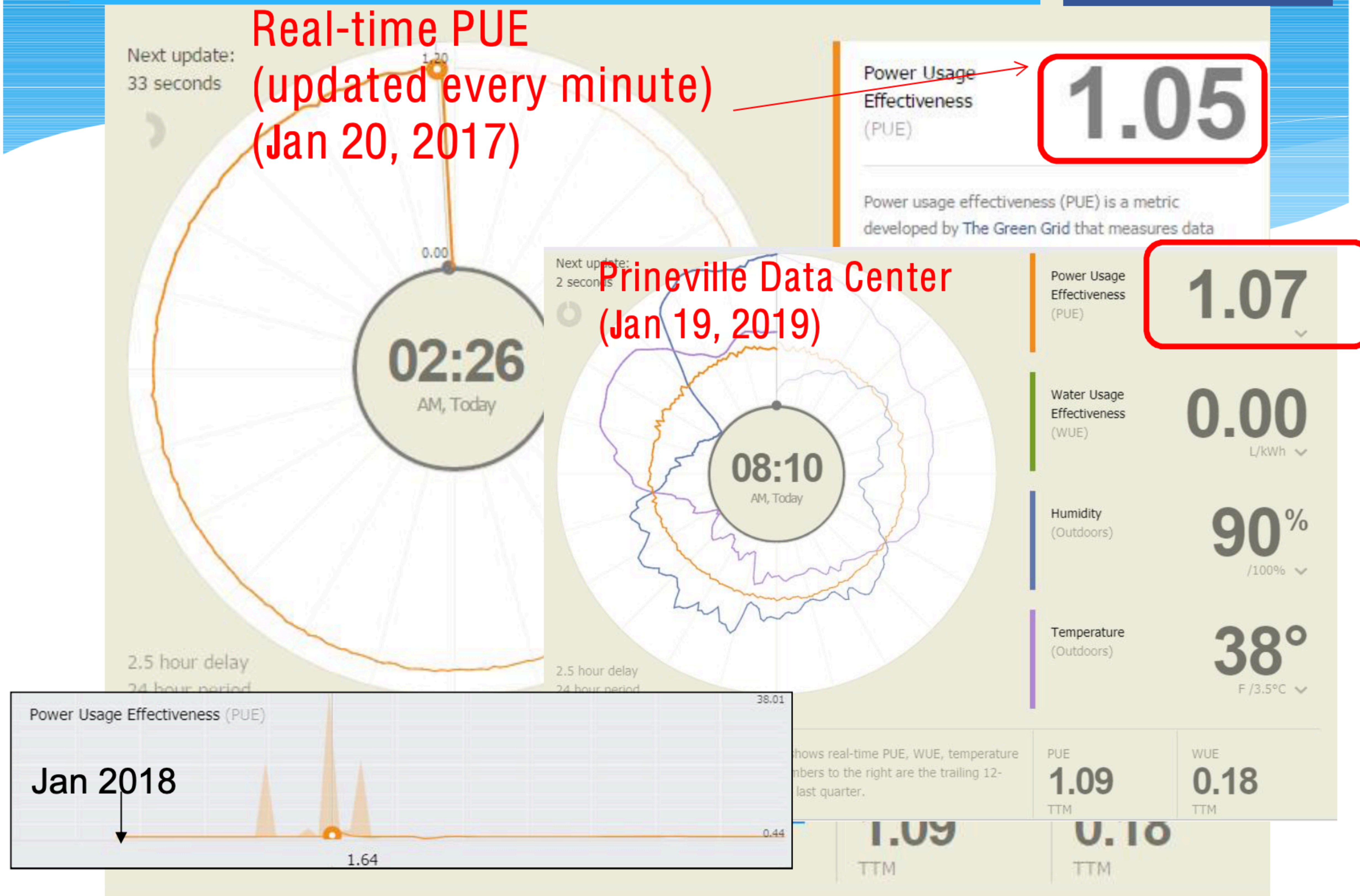
Facebook’s first Data Center @Prineville, Oregon (4/2011)
Managed by only 35 full-time employees

$$PUE = \frac{Total\ Facility\ Power}{IT\ Equipment\ Power}$$

Doesn't use traditional air conditioning, instead relying completely on "**outside air**".

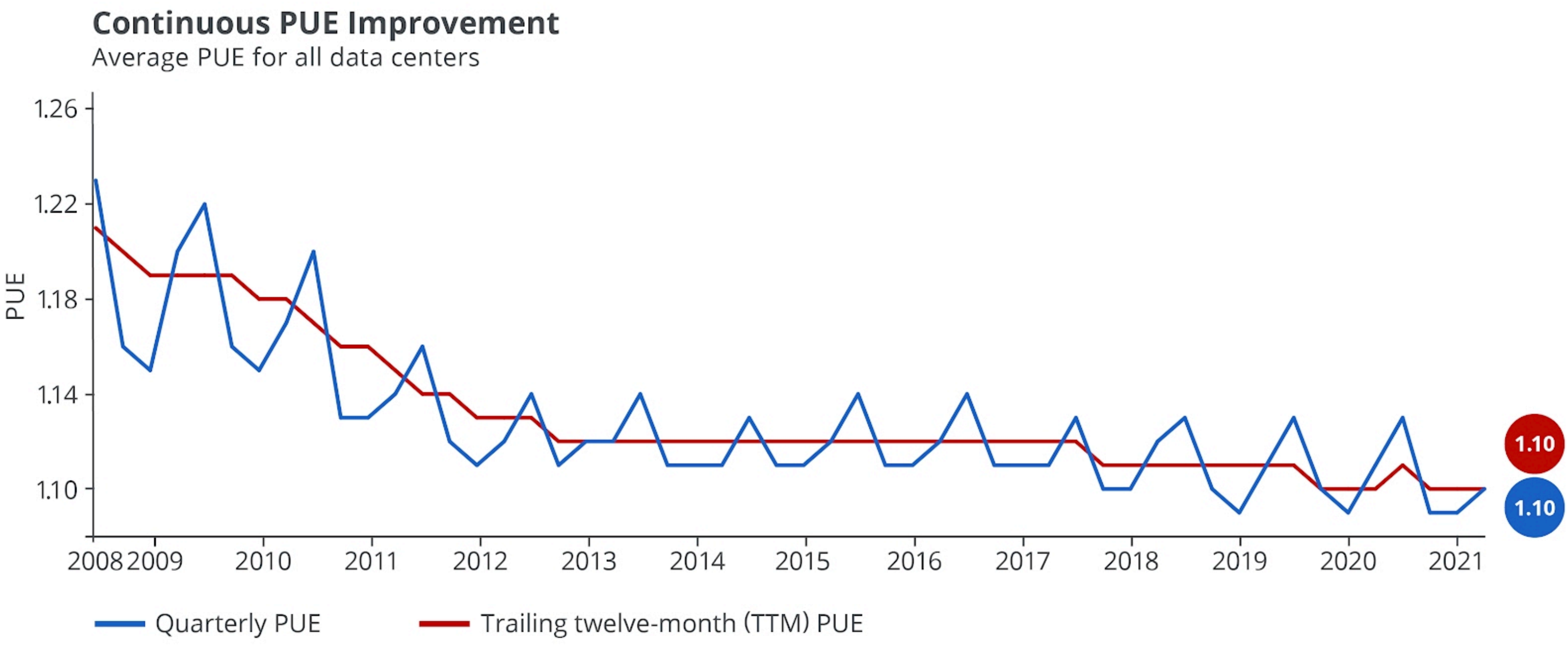
Power Usage Effectiveness (PUE) @Prineville is only 1.07 much more efficient than the industry average (1.5).

Facebook Prineville Data Center



Google Data Center PUE performance

Our fleet-wide PUE has dropped significantly since we first started reporting our numbers in 2008. The TTM energy-weighted average PUE for all Google data centers is 1.11, making our data centers among the most efficient in the world.



Yearly performance report

2021 2020 2019 2018 2017 2016 2015 2014 2013 2012 2011 2010 2009 2008

QUARTER 01 QUARTER 02

Fleet wide PUE	**Quarterly PUE**	**Trailing twelve-month (TTM) PUE**
Fleet	1.09	1.10

Campuses	**Quarterly PUE**	**Trailing twelve-month (TTM) PUE**
Douglas County, Georgia	1.08	1.10
Lenoir, North Carolina	1.08	1.09
Berkeley County, South Carolina	1.08	1.11
Montgomery County, Tennessee	1.09	1.14
Jackson County, Alabama	1.11	1.16
Loudoun County, Virginia	1.10	*
Loudoun County, Virginia (2nd facility)	1.14	*
Council Bluffs, Iowa	1.11	1.11
Council Bluffs, Iowa (2nd facility)	1.08	1.09
Mayes County, Oklahoma	1.08	1.11

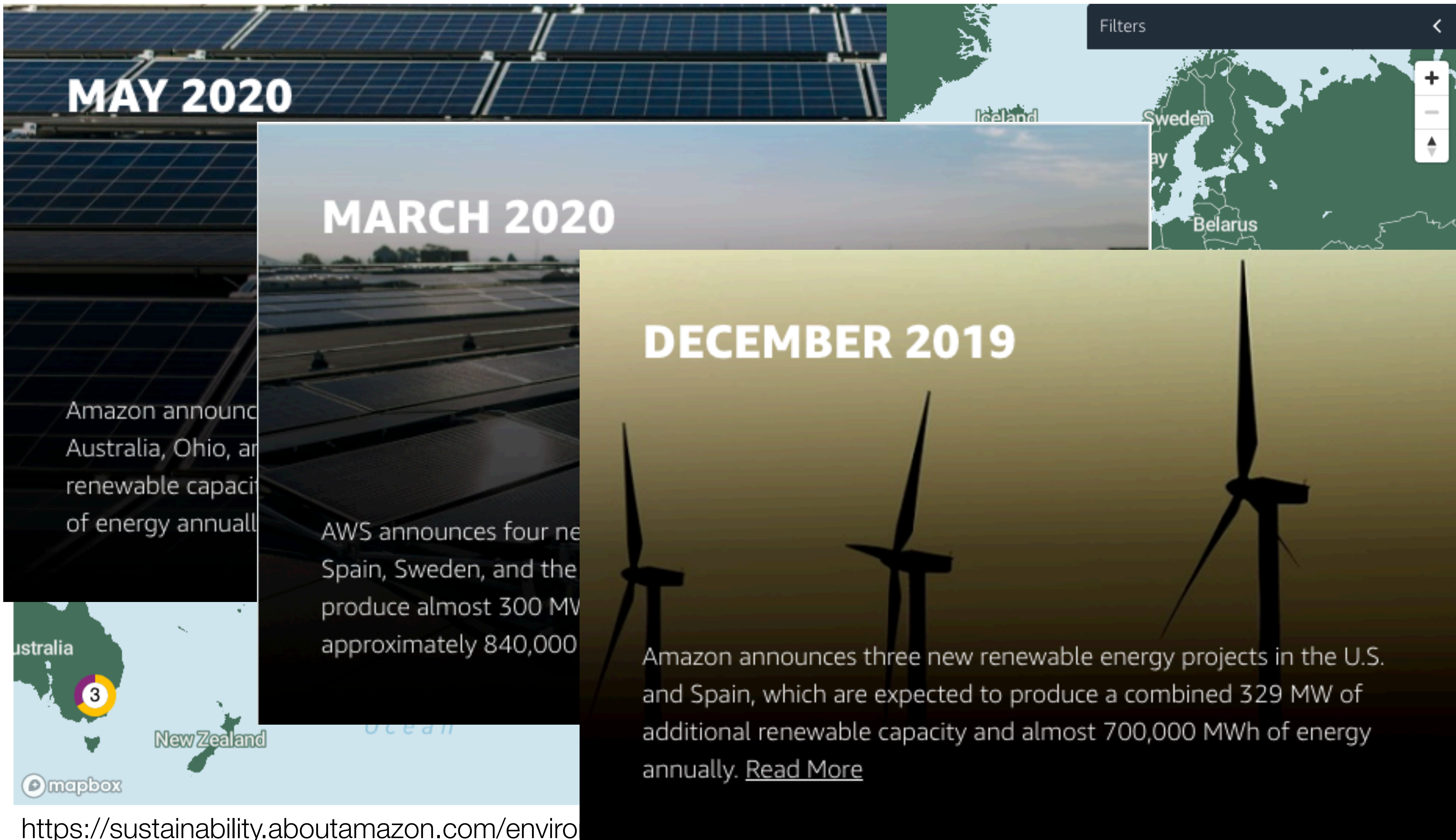
https://www.google.com/about/datacenters/efficiency/

AWS's data centers

The AWS Cloud spans 84 Availability Zones within 26 geographic regions around the world.

https://aws.amazon.com/about-aws/global-infrastructure/

Renewable energy in AWS

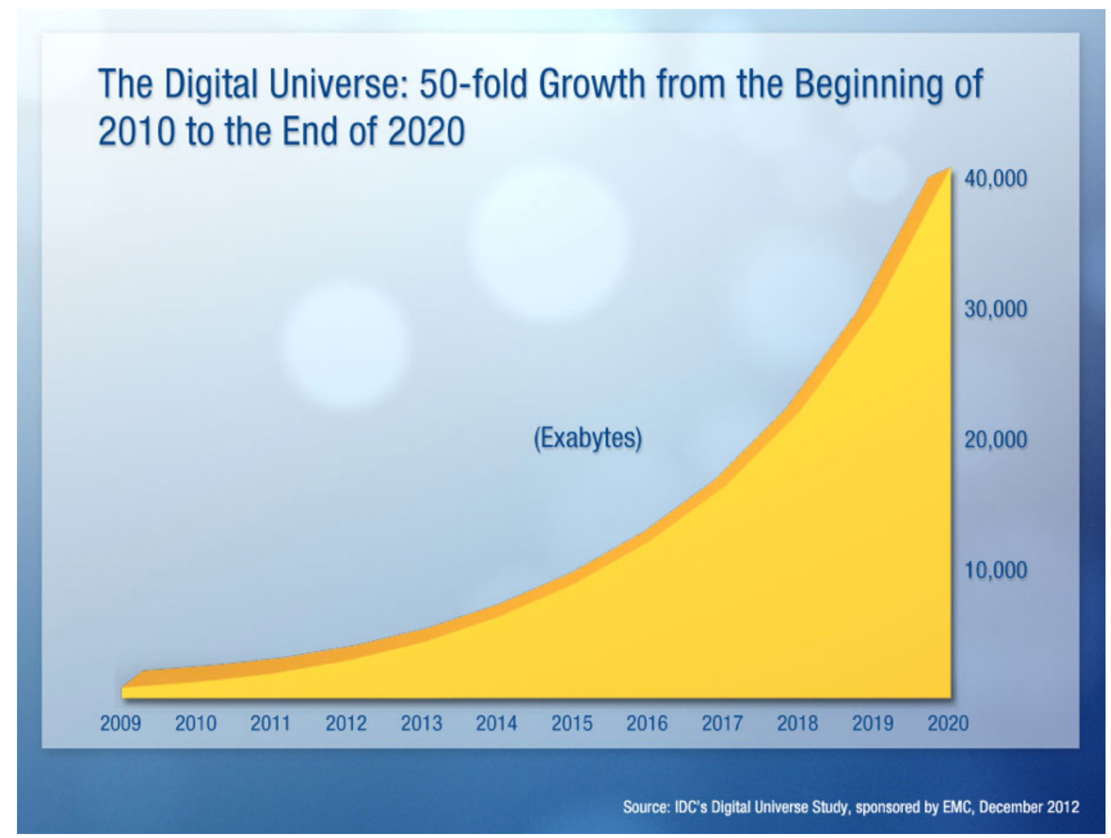


https://sustainability.aboutamazon.com/enviro

Data center as a computer

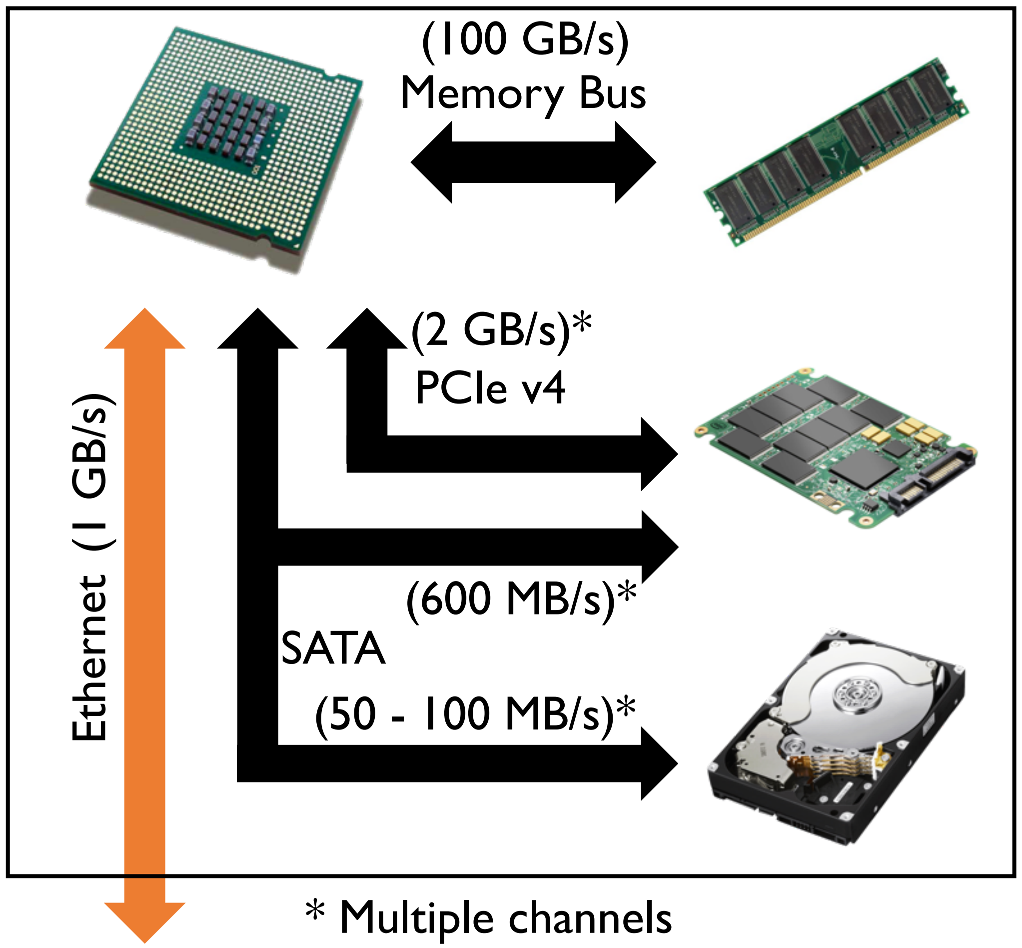
Why is One Machine Not Enough?

- Too much data
- Too little storage capacity
- Not enough I/O bandwidth
- Not enough computing capability



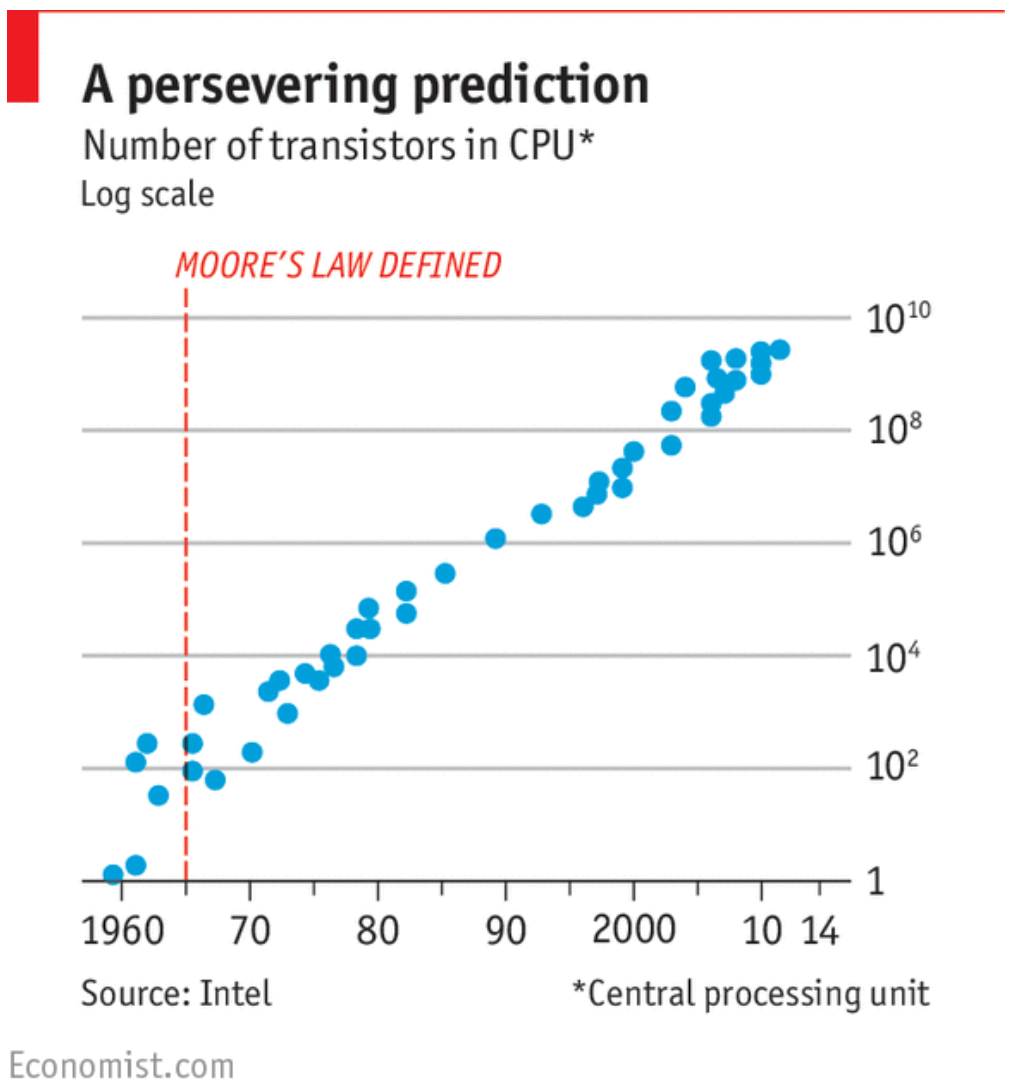
What's in a machine?

- Interconnected compute and storage resources
  - Different bandwidth and latency constraints
- Simplified diagram
  - Doesn't include GPUs, accelerators such as FPGAs, faster networks such as RDMA, dedicated GPU interconnects such as NVlink, etc…



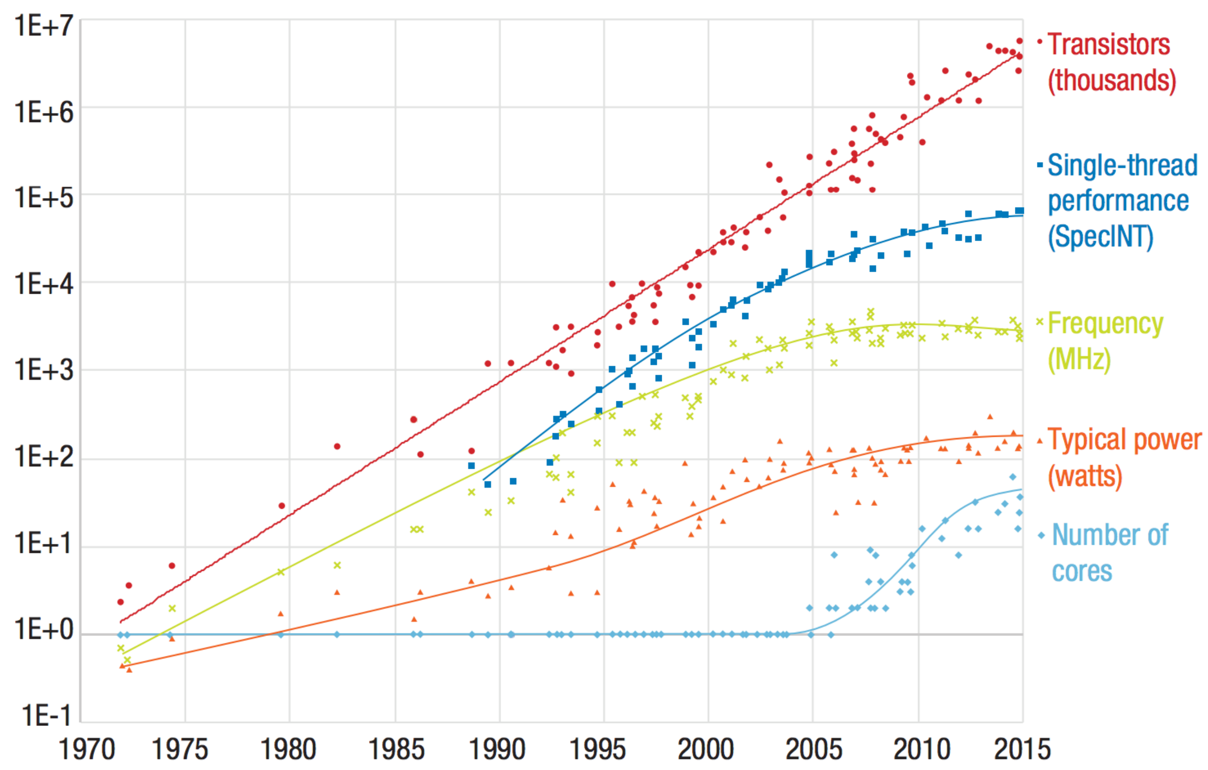
Scale Up: Make More Powerful Machines

- Moore's law
  - Stated 50+ years ago by Intel founder Gordon Moore
  - Number of transistors on microchip double every 2 years
  - Today "closer to 2.5 years" – Intel former CEO Brian Krzanich



Dennard: Scaling is the Problem

- Suggested that power requirements are proportional to the area for transistors
  - Both voltage and current being proportional to length
  - Stated in 1974 by Robert H. Dennard (DRAM inventor)
- Broken since 2005

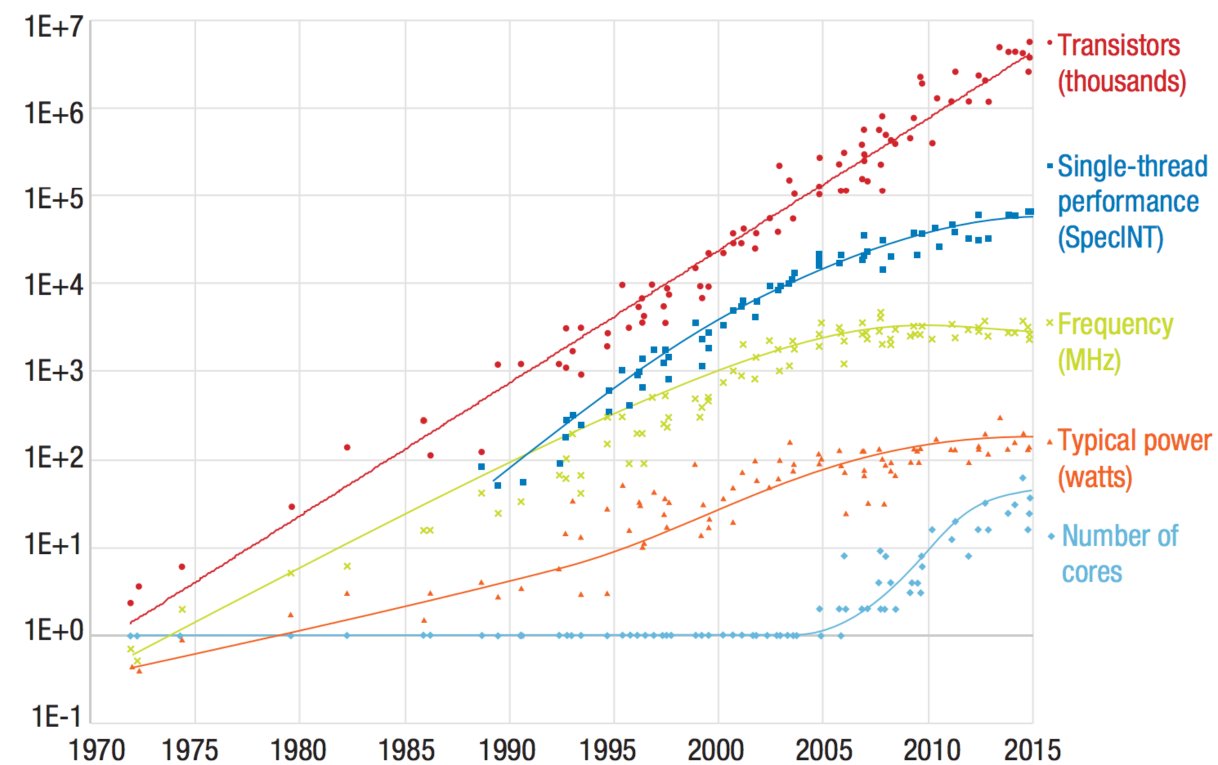


"Adapting to Thrive in a New Economy of Memory Abundance," Bresniker et al

At small sizes, current leakage poses greater challenges and also causes the chip to heat up, which creates a threat of thermal runaway and therefore further increases energy costs.

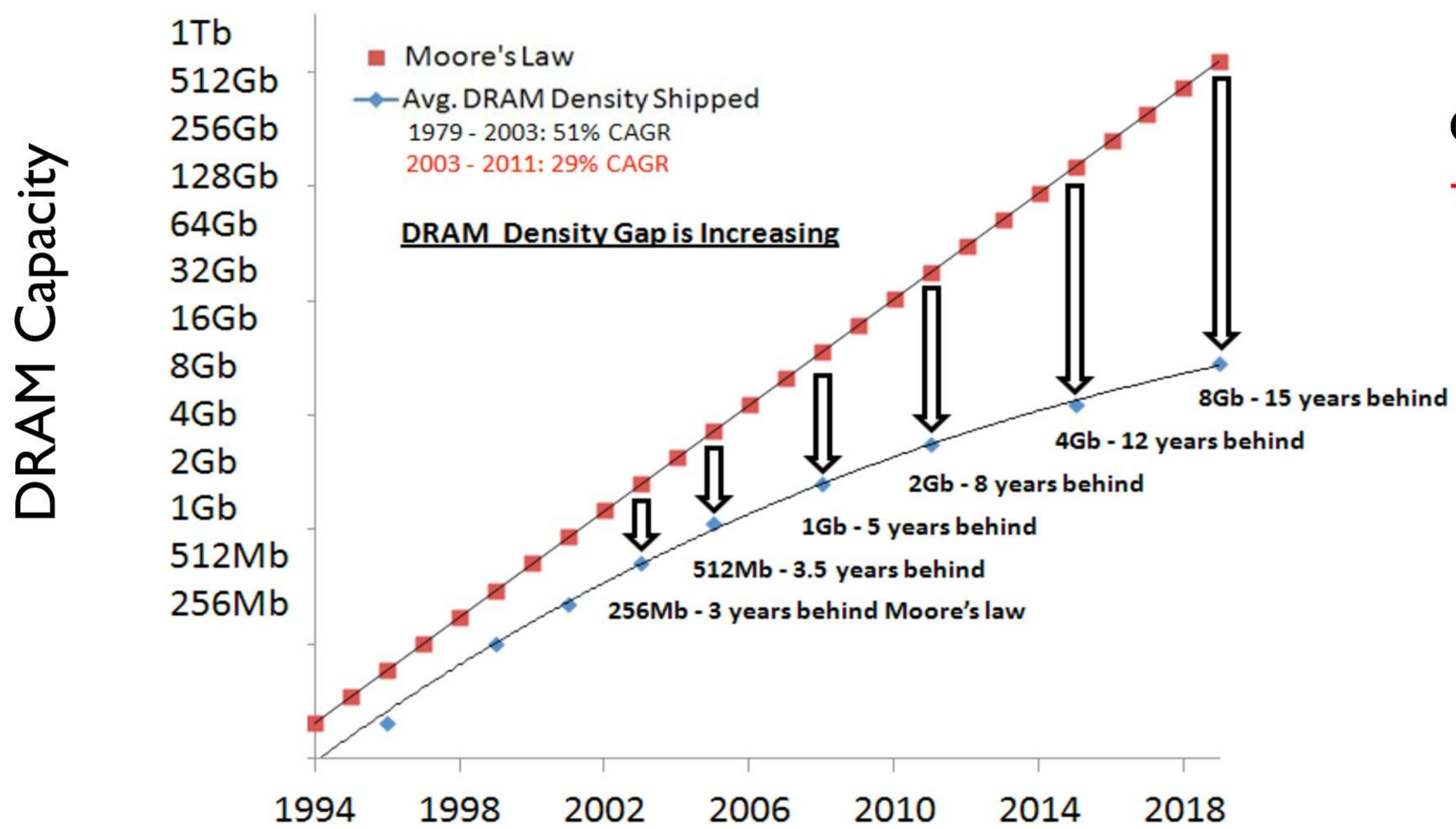
Dennard: Scaling is the Problem

- Performance per-core is stalled
- Number of cores is increasing



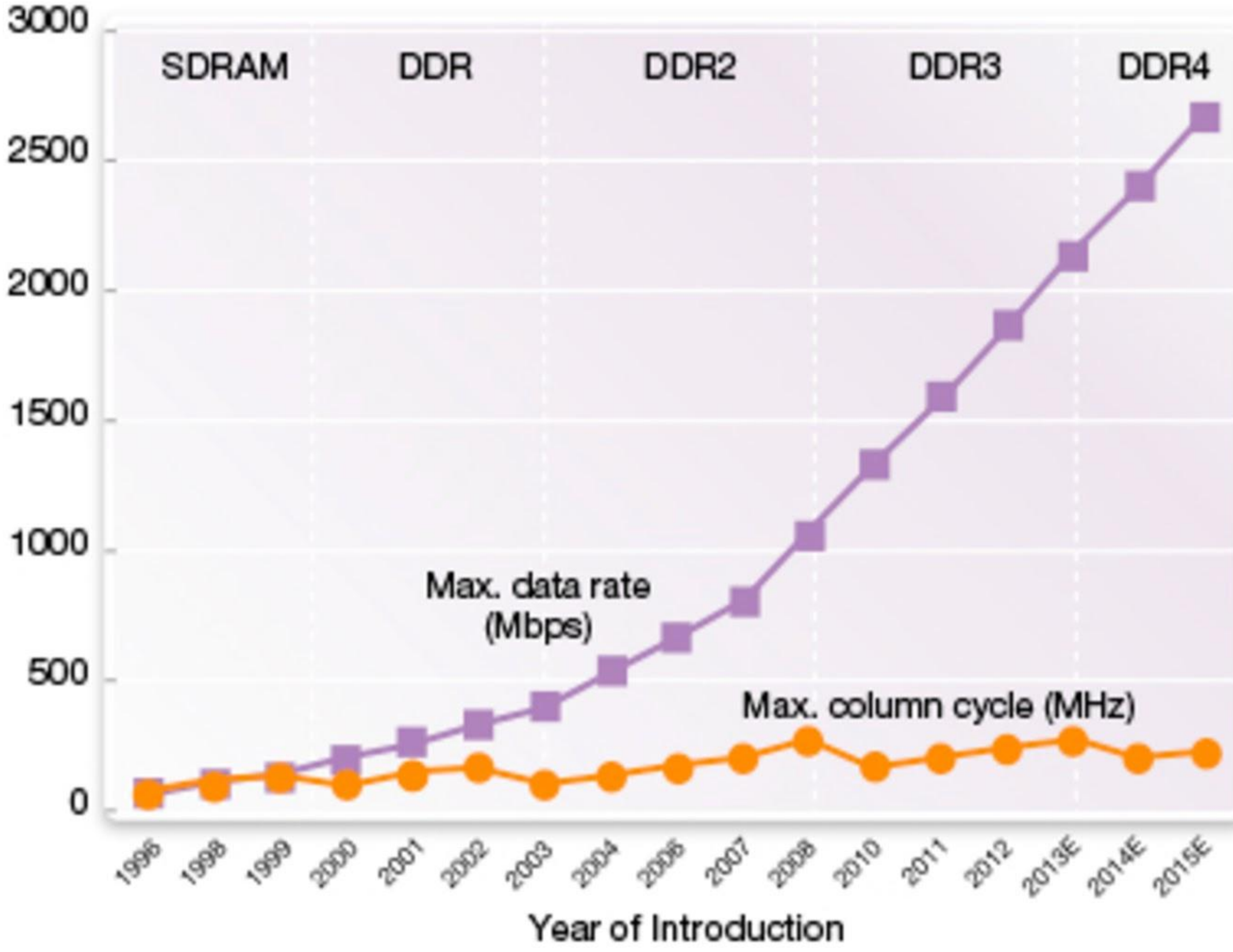
“Adapting to Thrive in a New Economy of Memory Abundance,” Bresniker et al

MEMORY CAPACITY



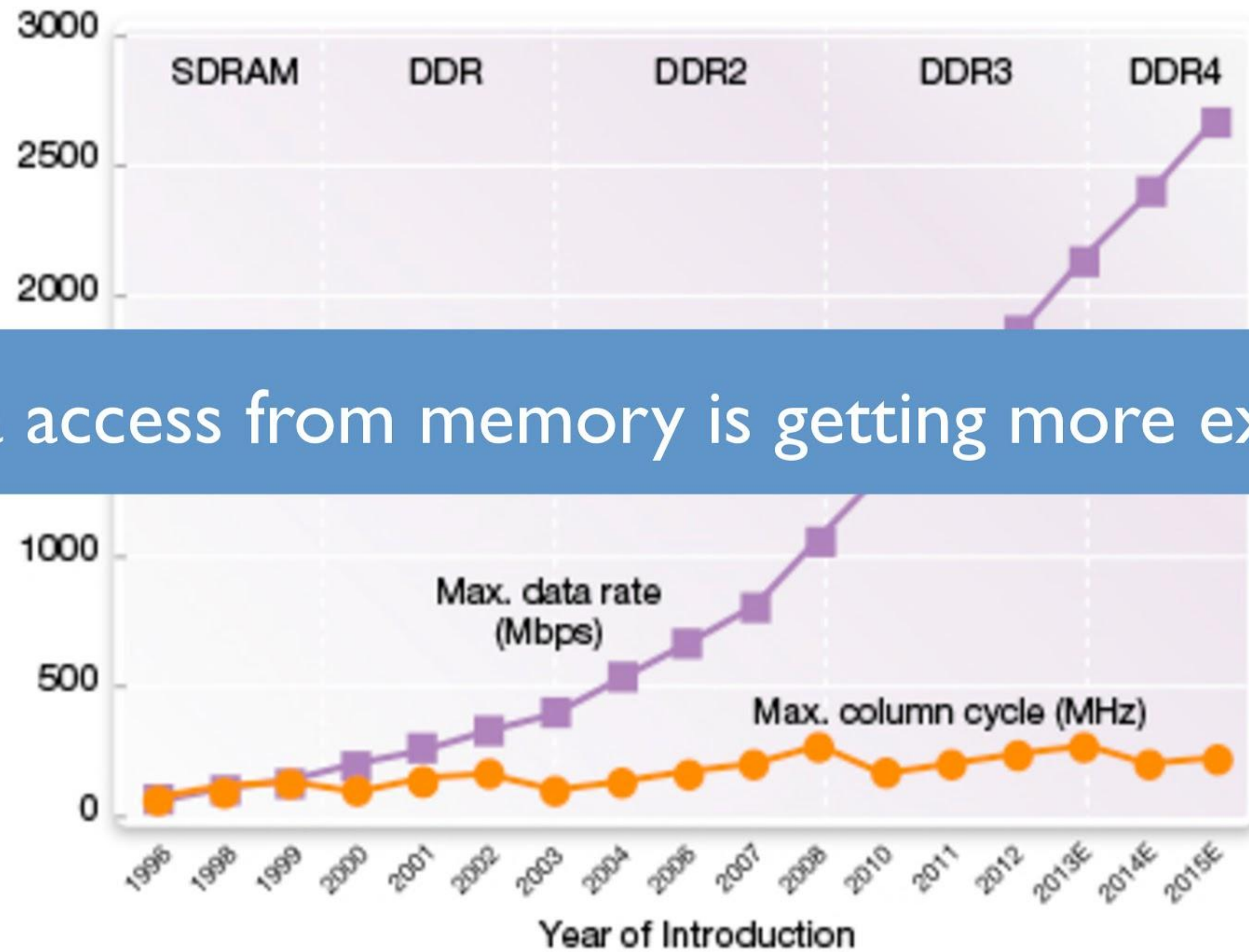
Growing by
+29% per year

MEMORY BANDWIDTH



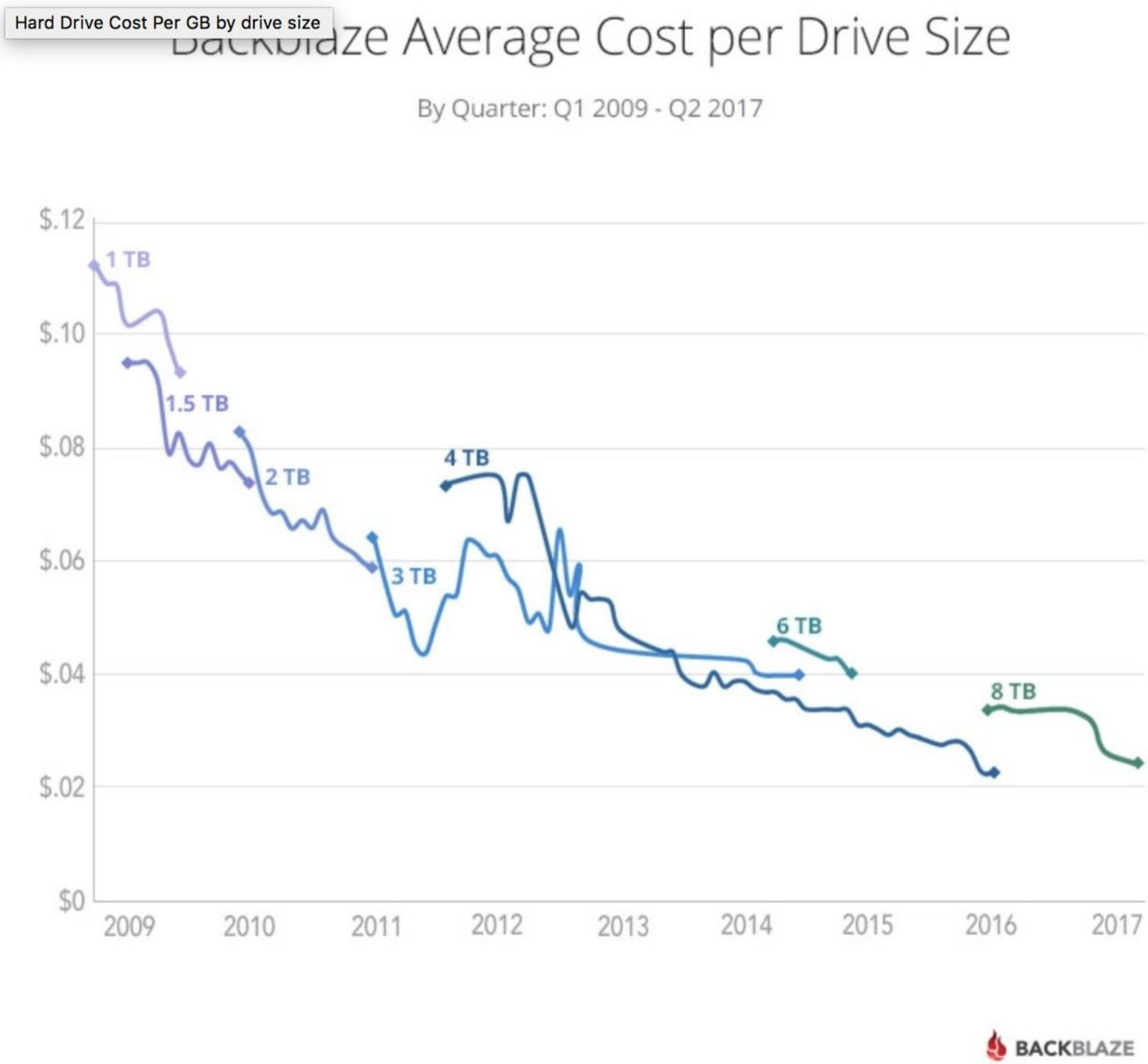
Growing
+15% per year

MEMORY BANDWIDTH



Data access from memory is getting more expensive !

HDD CAPACITY



HDD Bandwidth

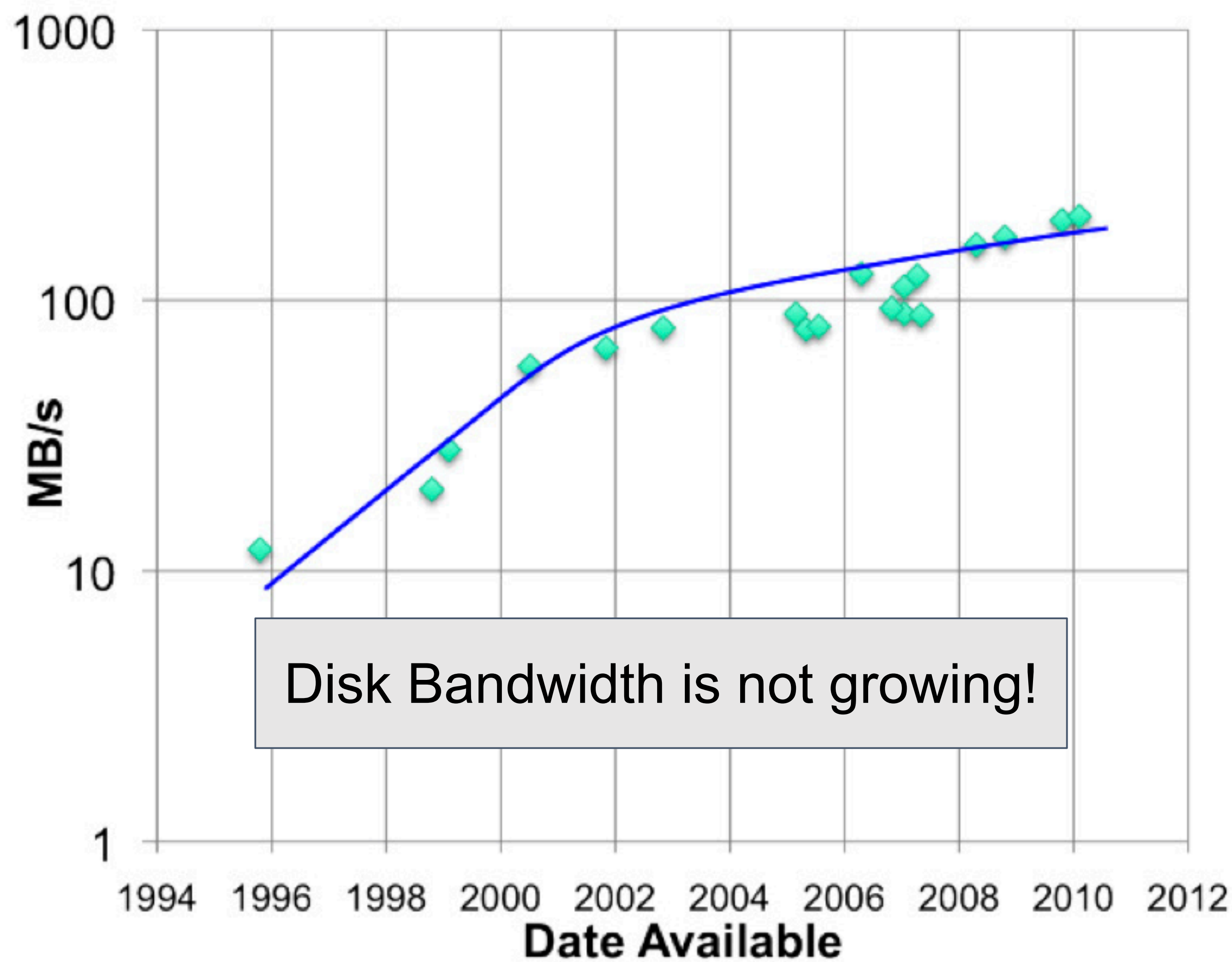
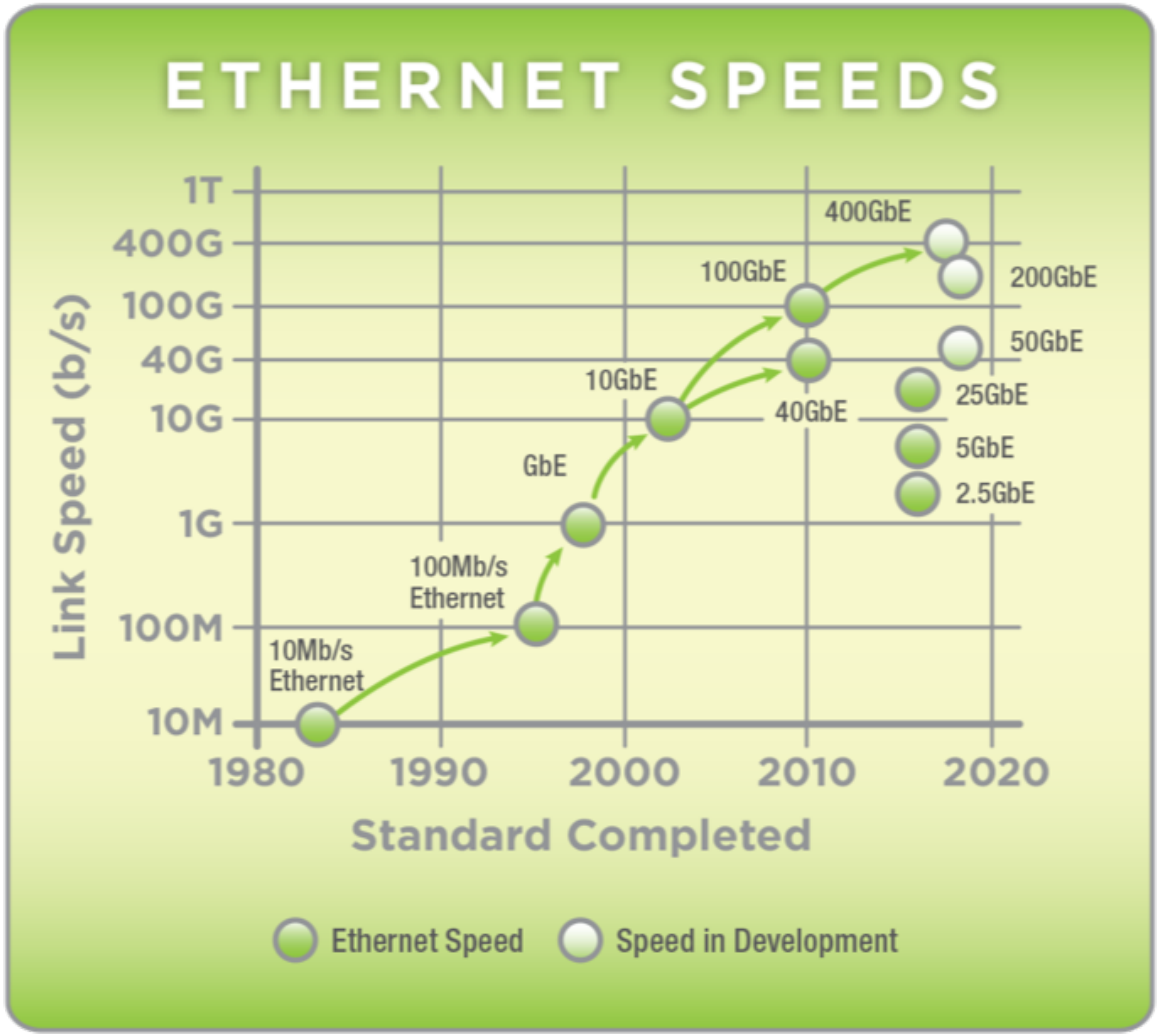


Figure 4: Maximum sustained bandwidth trend


ETHERNET SPEEDS
1T
400G
100G
40G
10G
1G
100M
10M
Link Speed (b/s)
10Mb/s
Ethernet
100Mb/s
Ethernet
GbE
10GbE
40GbE
100GbE
400GbE
200GbE
50GbE
25GbE
5GbE
2.5GbE
1980
1990
2000
2010
2020
Standard Completed
Ethernet Speed
Speed in Development

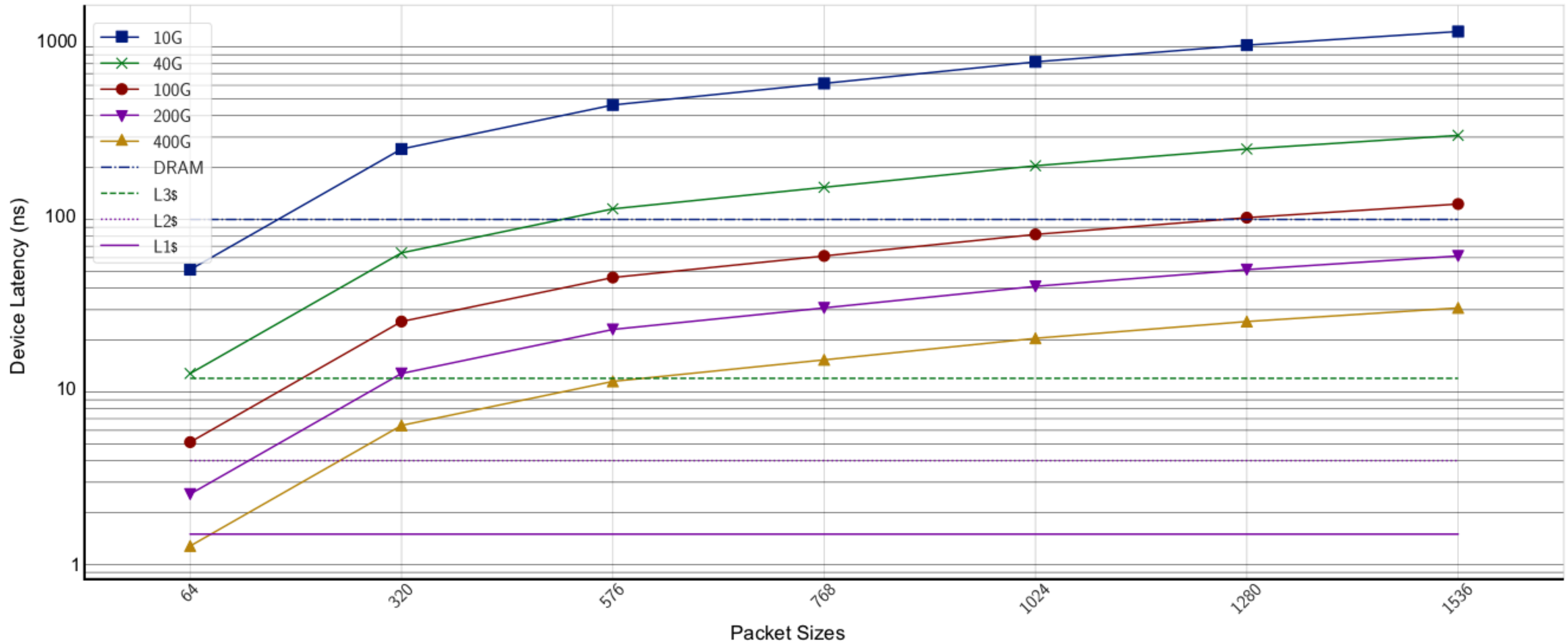


Figure 1: Evolution of Ethernet link speeds and interpacket gaps compared to CPU memory hierarchy latencies. As link speeds improve, the interpacket gap decreases. 40GbE represents the first time where several small packet sizes are just below DRAM latency. At 100GbE almost all packet sizes are under the latency line. This represents a new regime for the future where bottleneck is now the memory and cache hierarchy. If a system takes longer to service a DRAM miss than it does to obtain a new packet packets will drop. This forces network operators to run next generation NICs at current generation speeds and increases system cost as the only way around this problem is to increase core count.

Scaling Up is Unlikely to Help

- Physical limits
- Economical constraints
- What about fault-tolerance?
- How would you upgrade?
- Anything else?

Scale Out: Warehouse-Scale Computers

- Single organization
- Homogeneity (to some extent)
- Cost efficiency at scale
  - Multiplexing across applications and services
  - Rent it out!