Architecting a Sustainable Planet

Srilatha (Bobbie) Manne
Principal Hardware Engineer
Microsoft

Outline

Background on carbon emissions

- It is real & companies are listening
- The little things add up

Data center history & analysis

- Bigger can be better: Efficiencies of scale
- Breaking it down: Where does the carbon go?

Sustainability ideas

- Demand shaping: The when, where and how of running jobs
- Sustainable Silicon

Innovative solutions & conclusion

- Thinking outside the box – May lead you back into a different box
- The bigger picture

Facets of Sustainability

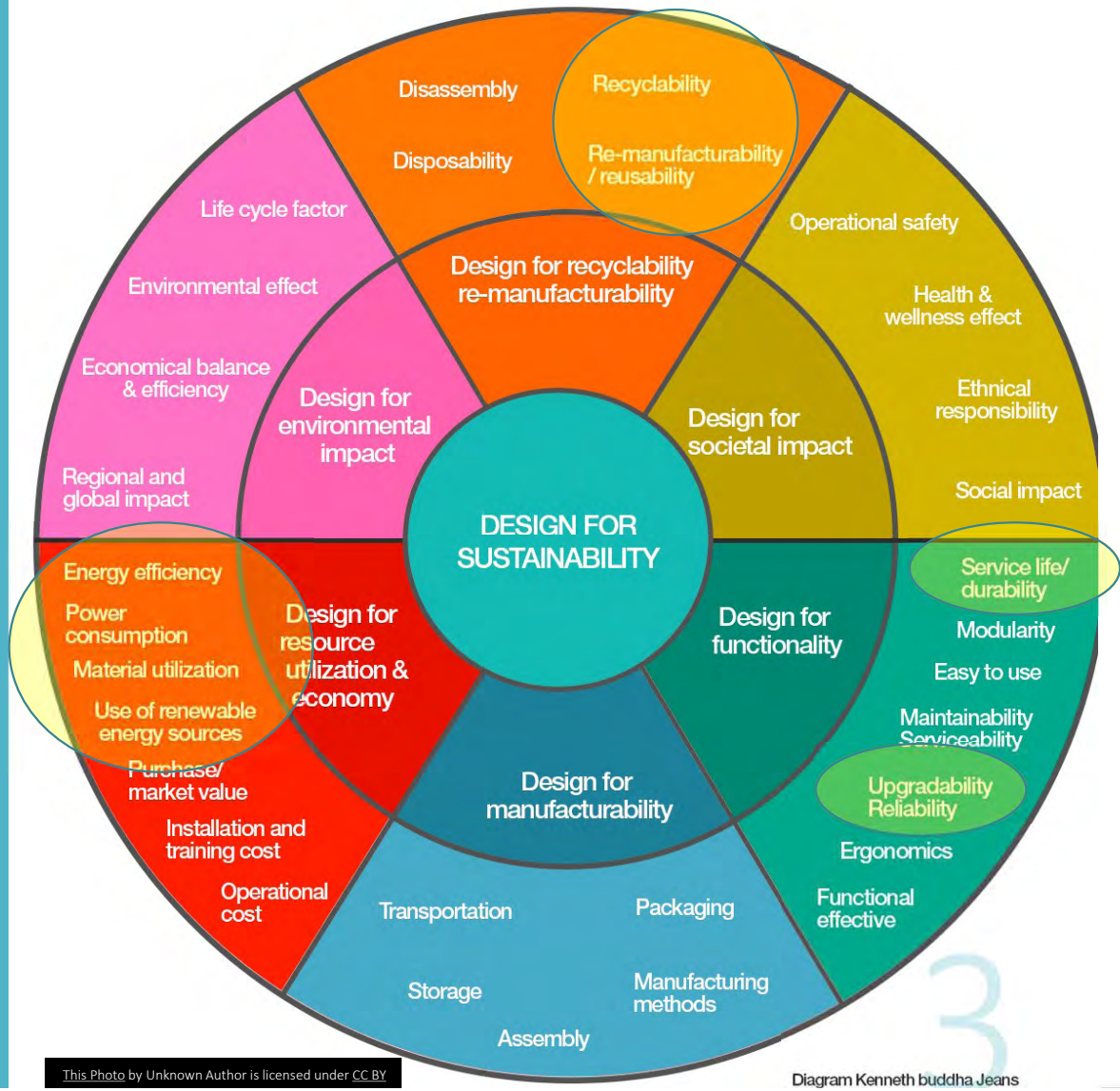


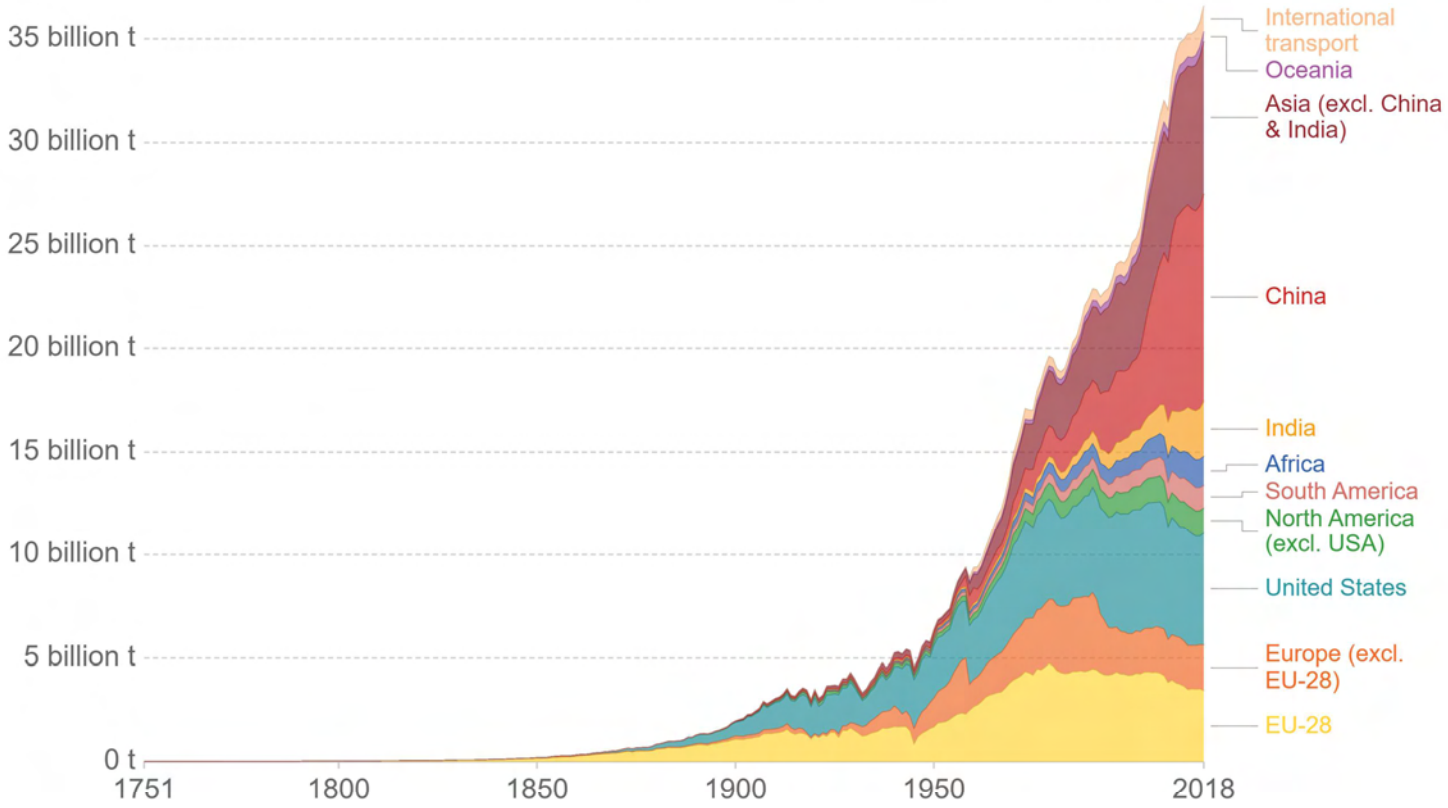
Diagram Kenneth buddha Jeans

It is real & companies are listening

A planet-sized challenge

Annual total CO_2 emissions, by world region

This measures CO_2 emissions from fossil fuels and cement production only – land use change is not included.



Source: Carbon Dioxide Information Analysis Center (CDIAC); Global Carbon Project (GCP)
Note: 'Statitistical differences' included in the GCP dataset is not included here.
OurWorldInData.org/co2-and-other-greenhouse-gas-emissions • CC BY

Data Centers

205TWh of electricity/year used †

Equivalent to 1 year of: *

31.3M cars

31,292 wind turbines

189.3M acres of forestland (Texas is ~172M acres)

145M metric tons CO2

† https://www.iea.org/reports/data-centres-and-data-transmission-networks

* Based on EPA Greenhouse Gas Equivalences Calculator

Walmart aims for zero greenhouse gas emissions by 2040

Axios, Sept 2020

Google, Facebook, Microsoft to Go Carbon-Free by 2030

Science Times, Sept 2020

Inside Apple's green revolution: Can it make a carbon neutral iPhone?

British GQ, Sept 2020

The little things add up

Examples

	Monthly Active User	# Monthly Active Users	Carbon/year (Metric Tons)
Facebook (2019)*	~0.1 kg CO2e	2.38 Billion	252,000 †

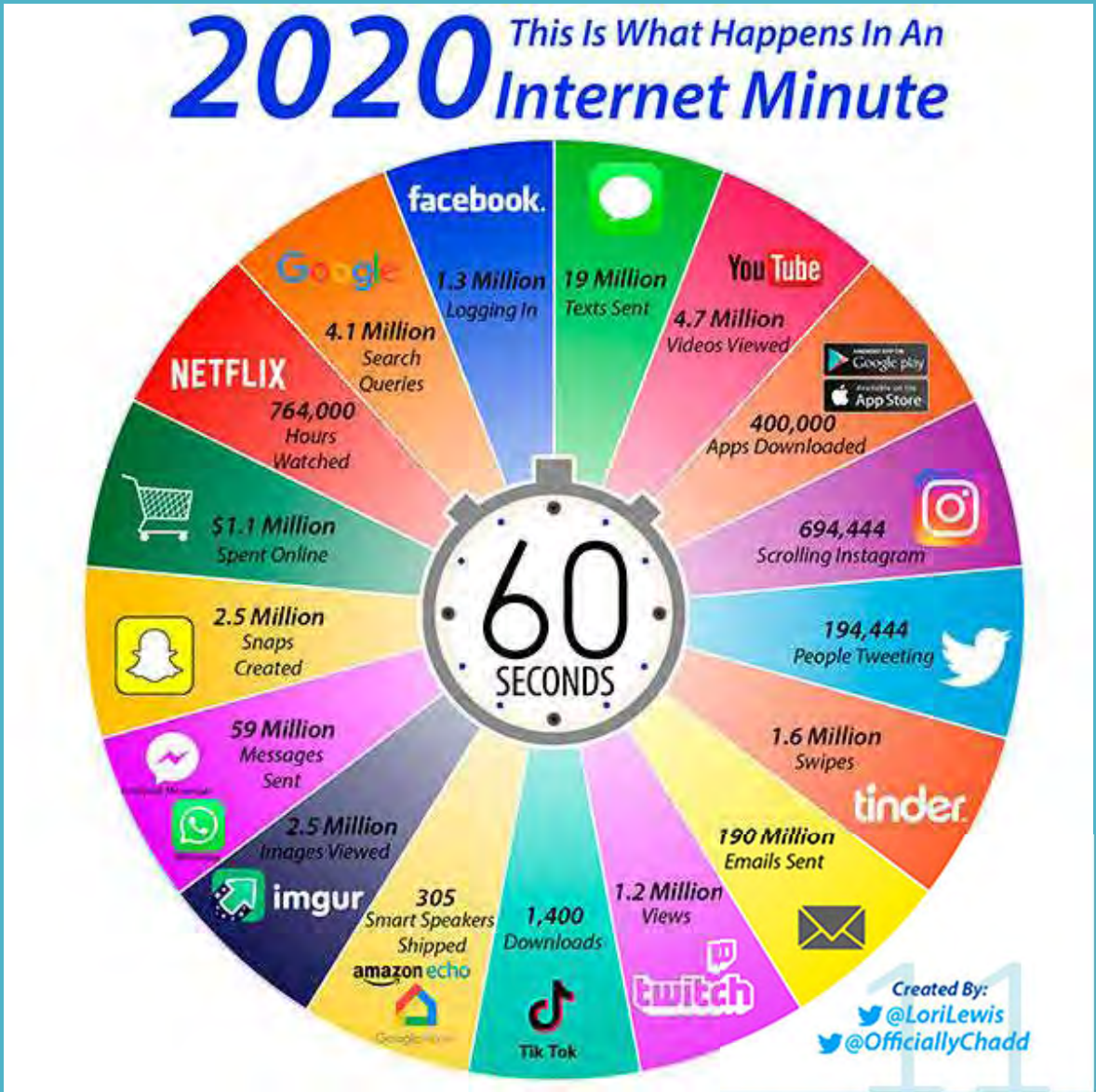
0.25 miles 54,443 cars

	Cost per hour (KWh)	Average Daily Hours	Tons CO2e/day	Carbon/year (Metric Tons)
Netflix-2019 + (Including Devices)	0.12 – 0.24	165,000,000	14,141-28,282	5,200,000 to 10,400,000

0.2 – 0.4 miles 1.1M – 2.2M cars

* https://sustainability.fb.com/report-pages/climate/
† 86% renewable energy mix
+ George Kamiya, "Factcheck: What is the carbon footprint of streaming video on Netflix?", https://www.carbonbrief.org/factcheck-what-is-the-carbon-footprint-of-streaming-video-on-Netflix

Death by a billion swipes



https://www.allaccess.com/merge/archive/31294/infographic-what-happens-in-an-internet-minute

“A cloud for everyone, on every device.”
Satya Nadella

Current World Population

7,810,981,656

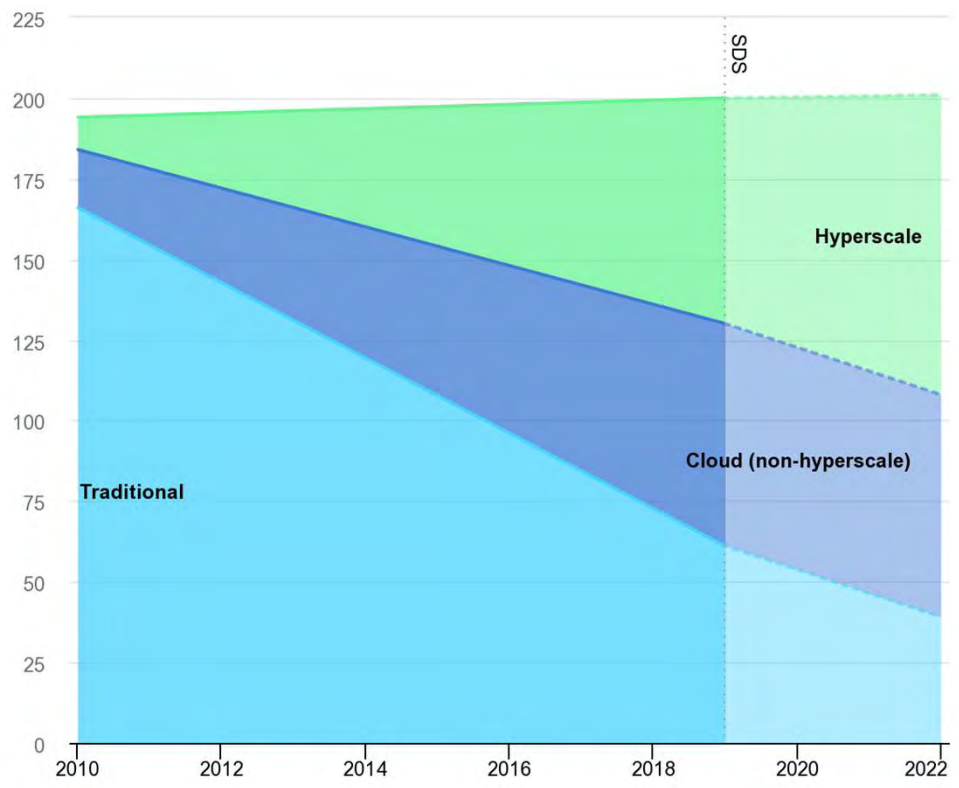
As of September 10, 2020

World Internet Users

4.57 BILLION

As of July, 2020

Bigger can be better: Efficiencies of scale

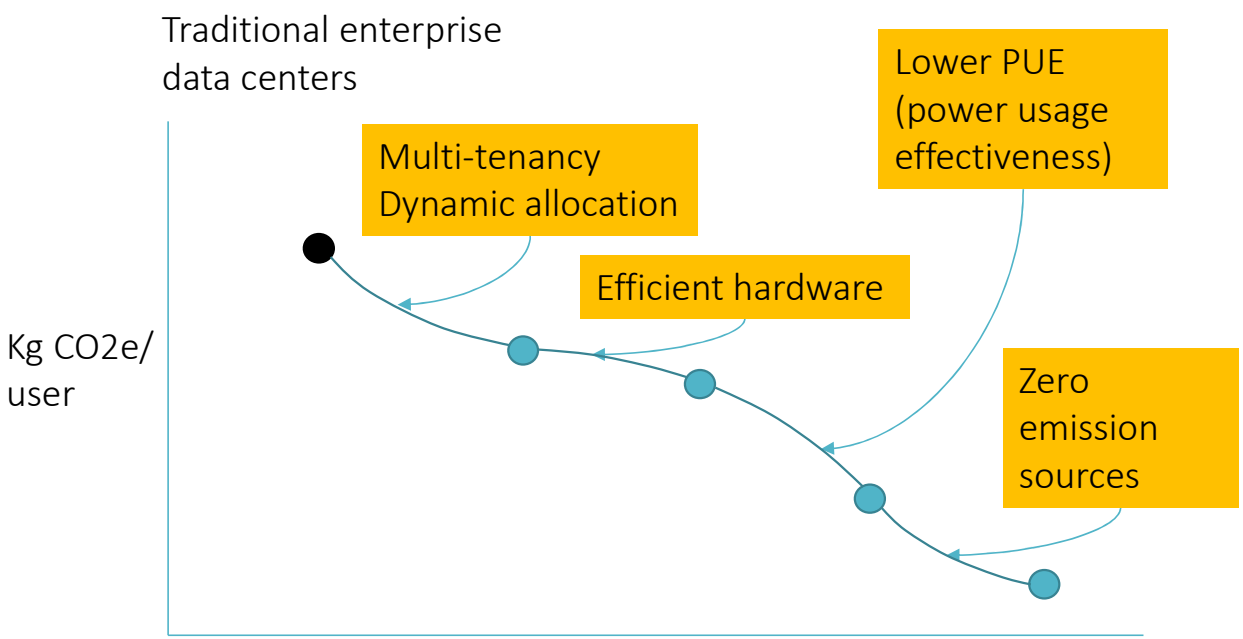


IEA, Global data centre energy demand by data centre type, 2010-2022, IEA, Paris https://www.iea.org/data-and-statistics/charts/global-data-centre-energy-demand-by-data-centre-type-2010-2022

Growth of hyperscale[1]

- 2X increase in users
- 12X increase in traffic
- 6% increase in energy use

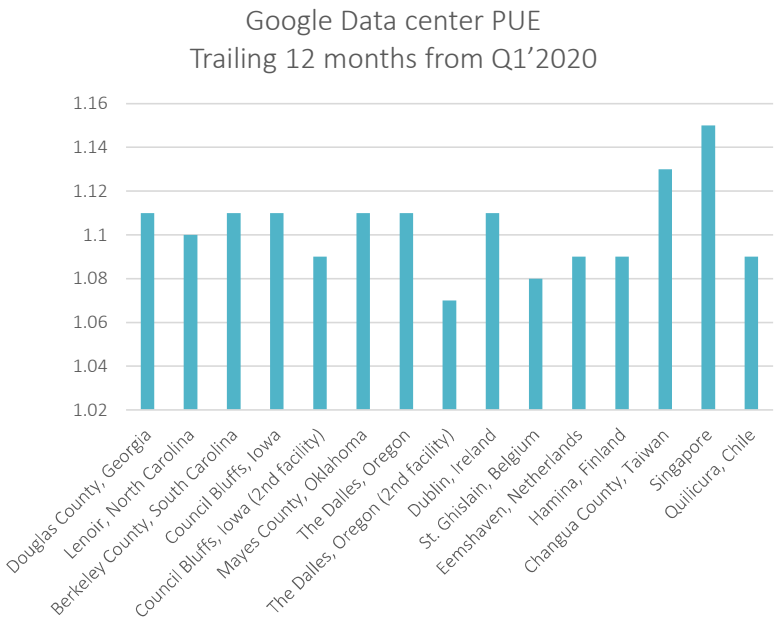
[1] https://www.iea.org/reports/data-centres-and-data-transmission-networks



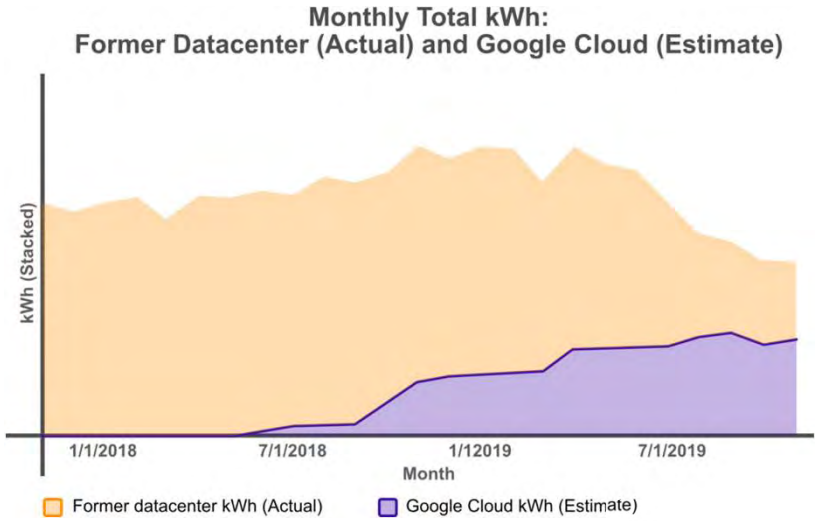
Benefits of hyperscale

"The carbon benefits of cloud computing: A study on the Microsoft Cloud in partnership with WSP", Microsoft 2020

Hyperscale efficiencies



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



E. Sommer et. Al, "Cloud Jewels: Estimating kWh in the Cloud", https://codeascraft.com/2020/04/23/cloud-jewels-estimating-kwh-in-the-cloud/

Breaking it down: Where does the carbon go?

Microsoft Azure Regions

54 regions worldwide **140** available in 140 countries

Canada East
Canada Central
West US 2
US Gov Iowa
Central US
West Central US
West US
US Gov Arizona
North Central US
US DoD East
East US
East US 2
South Central US
US Gov Texas
US Gov Virginia
US DoD Central

Norway West
Norway East
West Europe
Germany West Central
UK South
Germany North
North Europe
Germany Northeast
UK West
Germany Central
France Central
Switzerland North
France South
Switzerland West

China North
China North 2
Korea Central
Korea South
Japan East
Japan West
China East 2
China East
East Asia

UAE North
UAE Central
West India
Central India
South India
Southeast Asia

Brazil South
South Africa North
South Africa West

Australia East
Australia Central
Australia Southeast
Australia Central 2

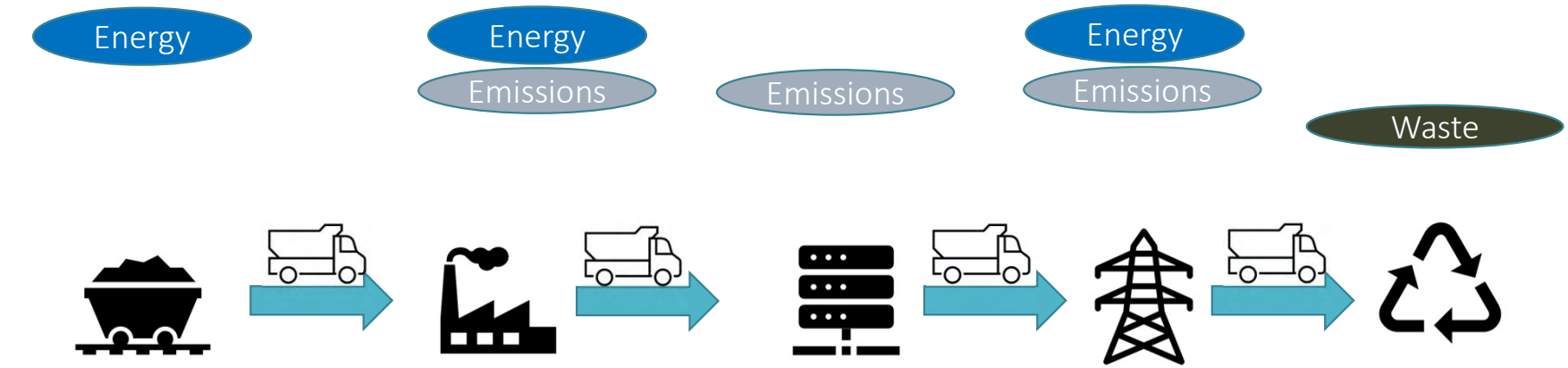
Available region
Announced region
Availability Zones

news.microsoft.com

"If the cement industry were a country, it would be the third largest emitter in the world… "

Jocelyn Timperley, "Why cement emissions matter for climate change"

Data center equipment lifecycle



Material extraction & manufacturing

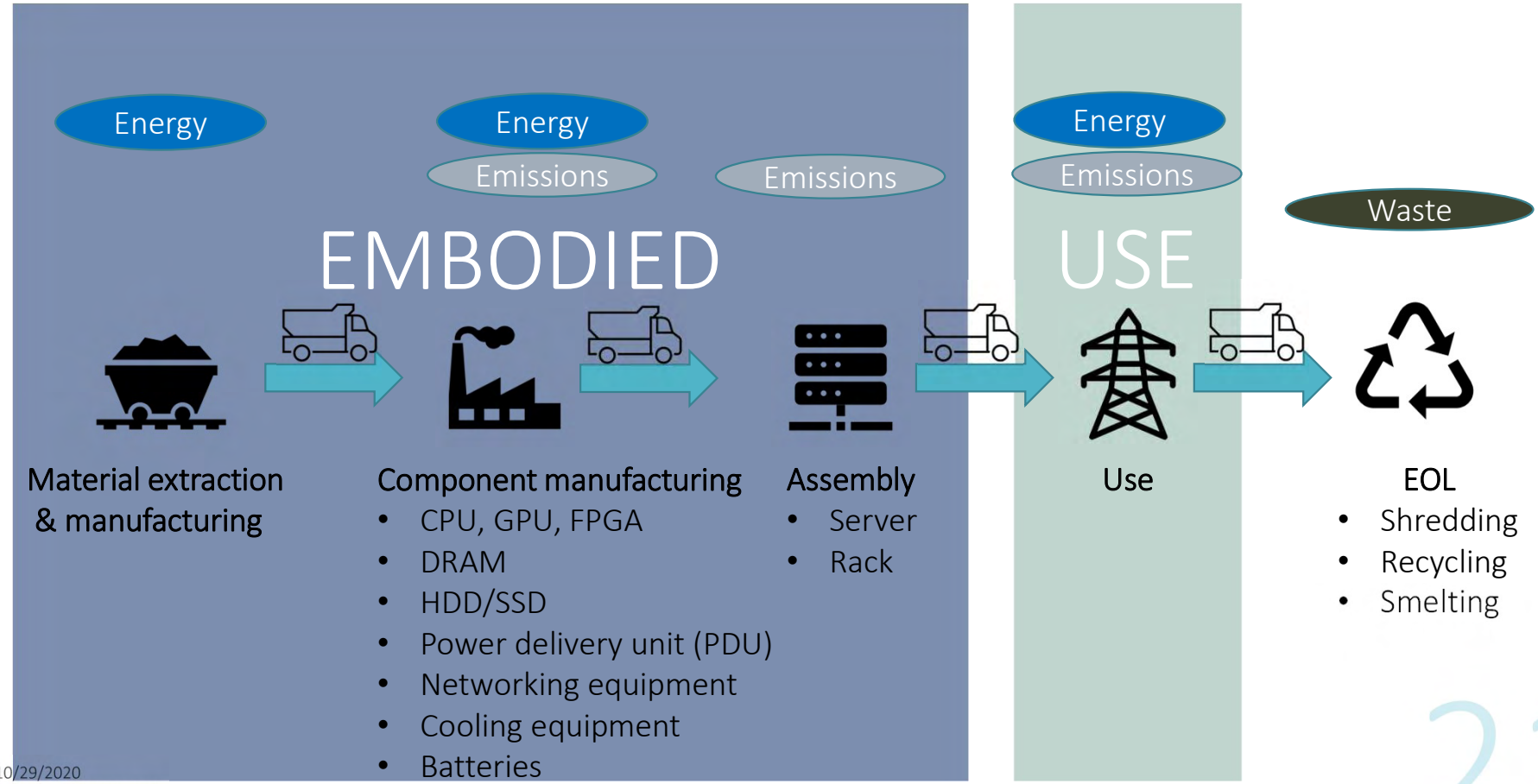
Component manufacturing
- CPU, GPU, FPGA
- DRAM
- HDD/SSD
- Power delivery unit (PDU)
- Networking equipment
- Cooling equipment
- Batteries

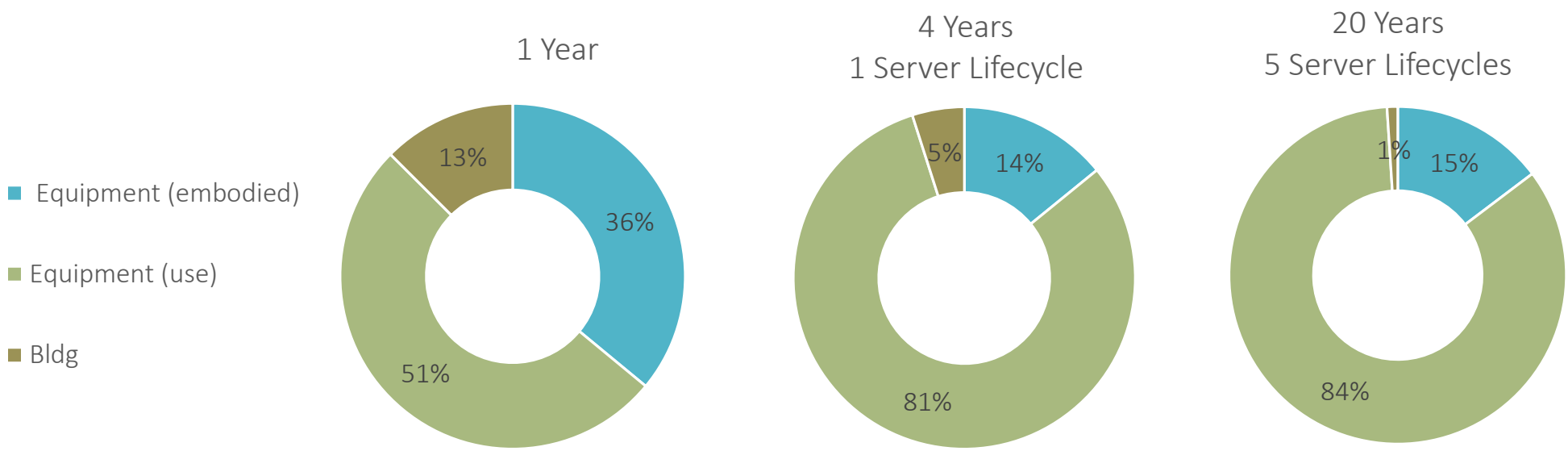
Assembly
- Server
- Rack

Use

EOL
- Shredding
- Recycling
- Smelting

Data center equipment lifecycle

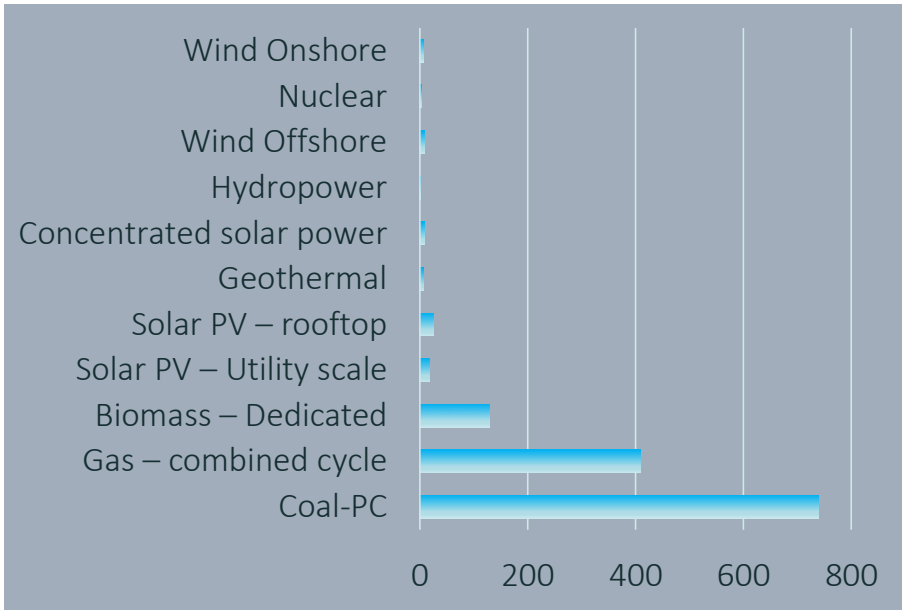




1MW Data Center

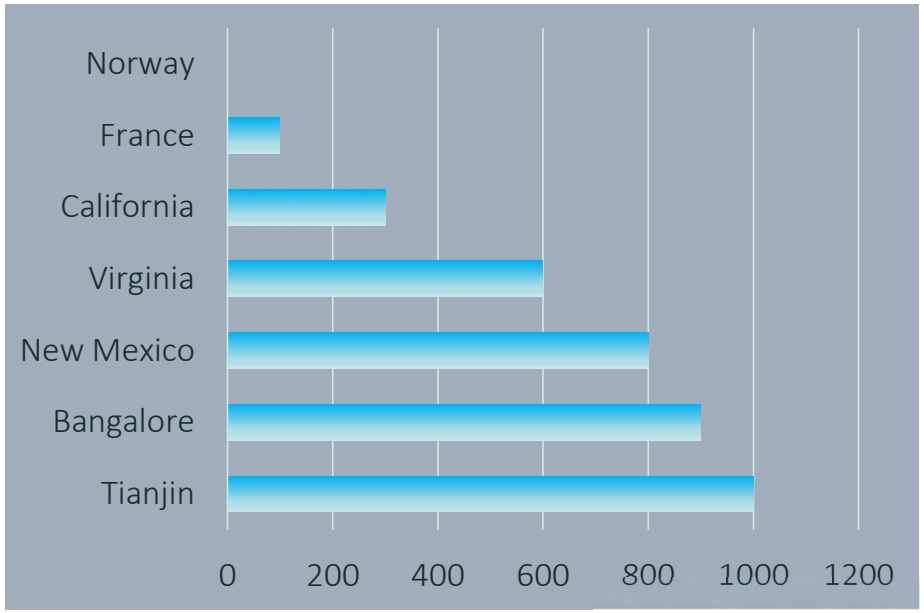
Assumptions: ~10,000 Sq. Ft building using data from "Quantifying the Carbon Footprint of Every Building", https://www.phnw.org/assets/2019Conference/Presentations/PHnw2019_Quantifying%20the%20Carbon%20Footprint%20of%20Every%20Project_Stacy%20Smedley.pdf, Dell Power Edge R640 embodied and use assumptions (https://i.dell.com/sites/csdocuments/CorpComm_Docs/en/carbon-footprint-poweredge-r640.pdf), 14KWRack, 1.1 PUE, 1 TOR per rack, 10% network power

Not all electricity is created equal



Grams CO2e per KWh

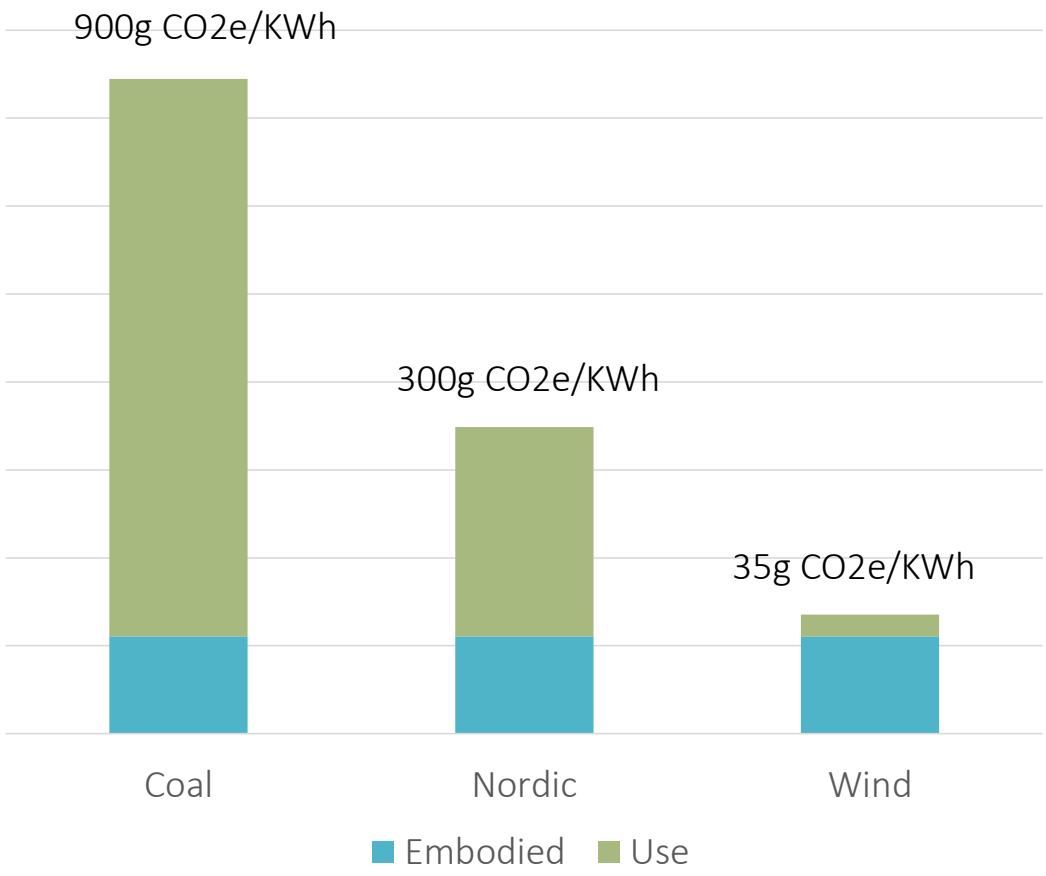
https://en.wikipedia.org/wiki/Life-cycle_greenhouse-gas_emissions_of_energy_sources



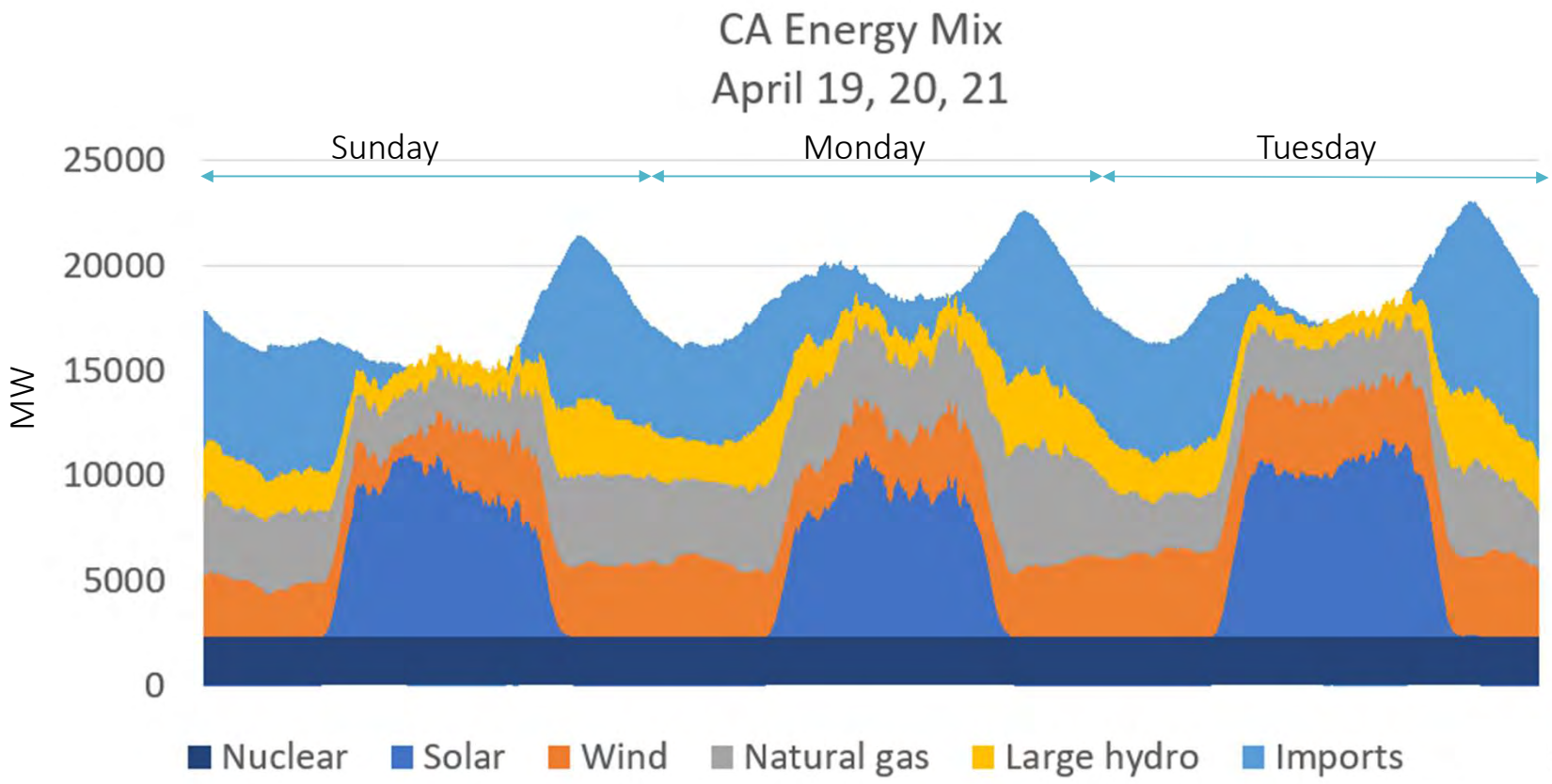
Grams CO2e per KWh

Fred Pearce, "Energy Hogs: Can World's Huge Data Centers Be Made More Efficient?, YaleEnvironment360, April 2018. https://e360.yale.edu/features/energy-hogs-can-huge-data-centers-be-made-more-efficient

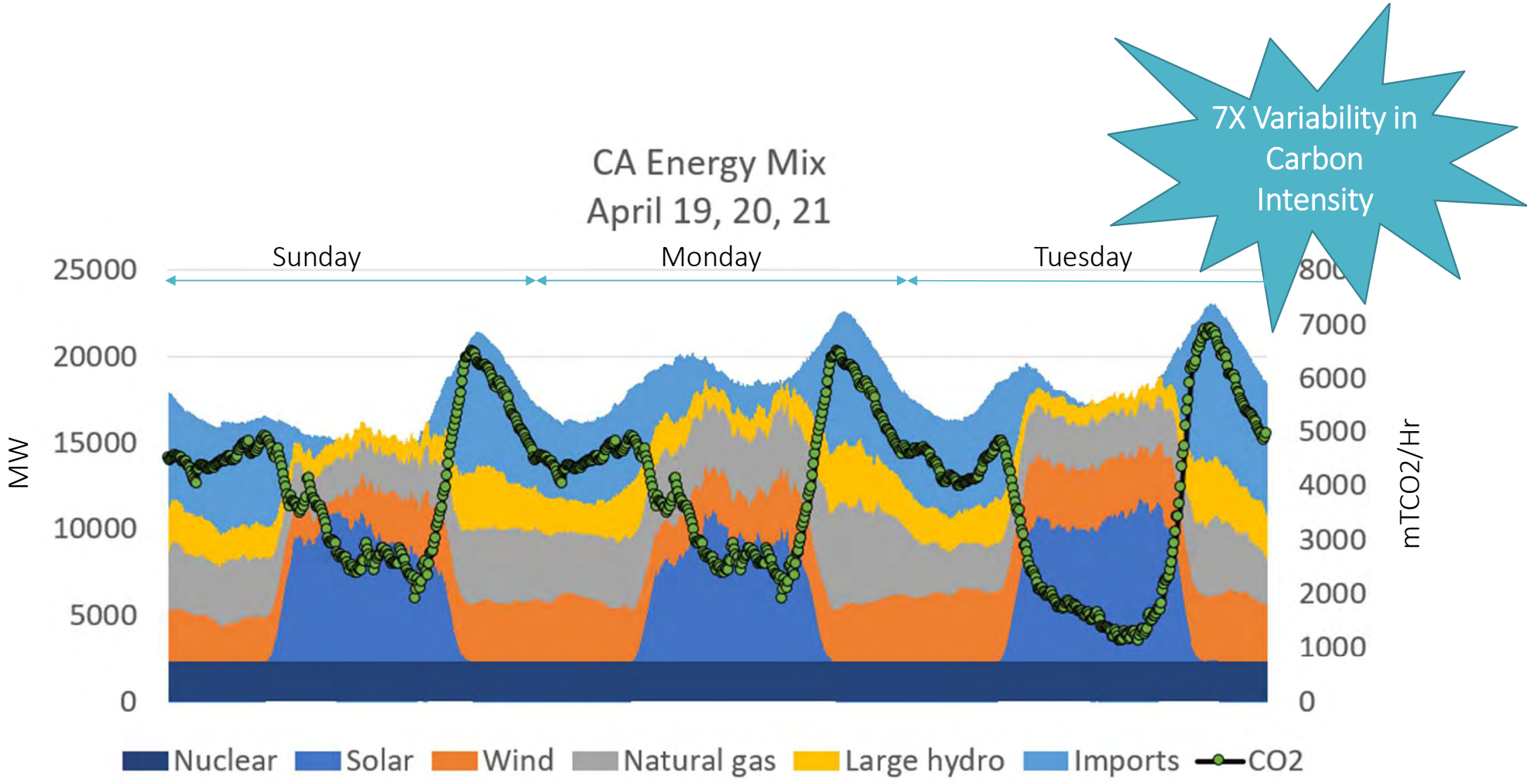
1MW data center & carbon intensity



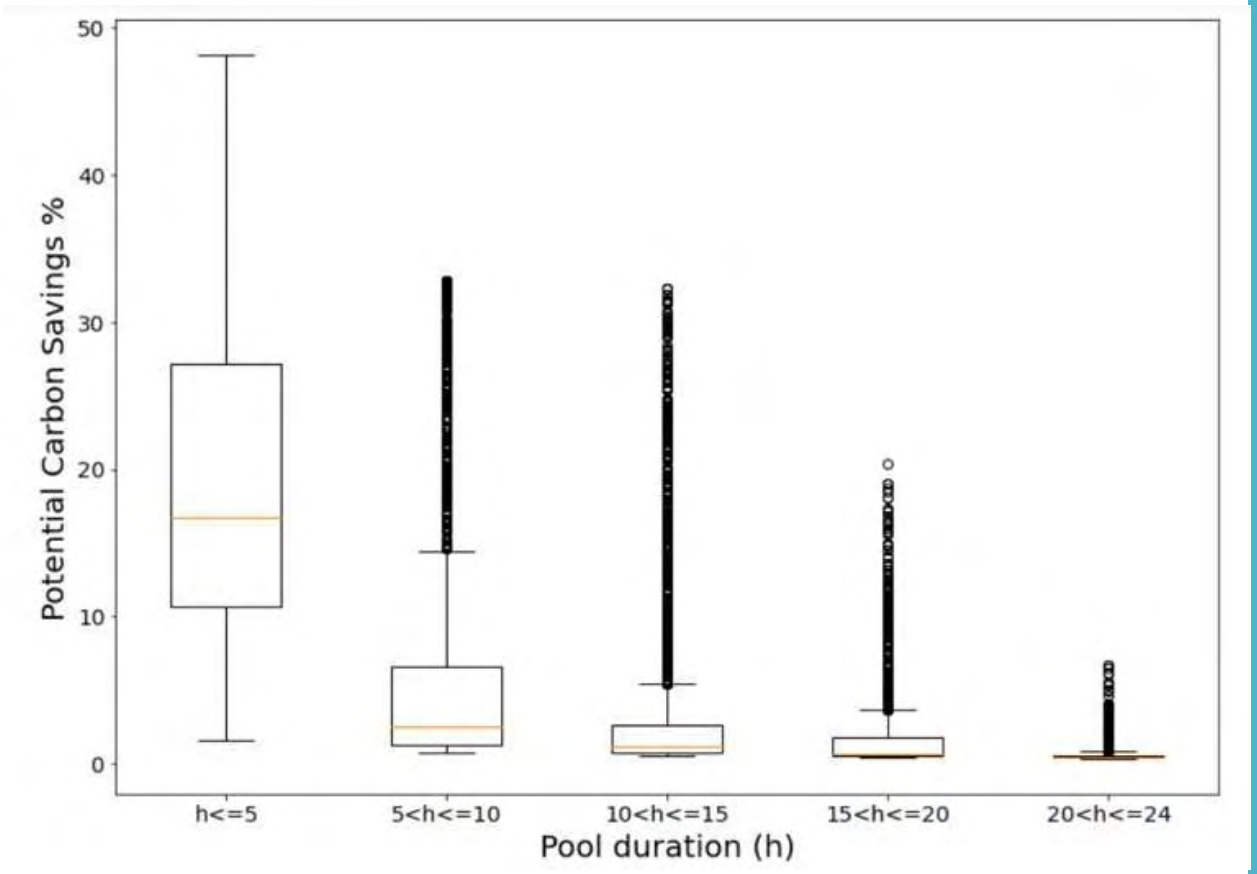
Demand Shaping: The *when*, *where*, and *how* of running jobs



Data from California Independent System Operator



Data from California Independent System Operator



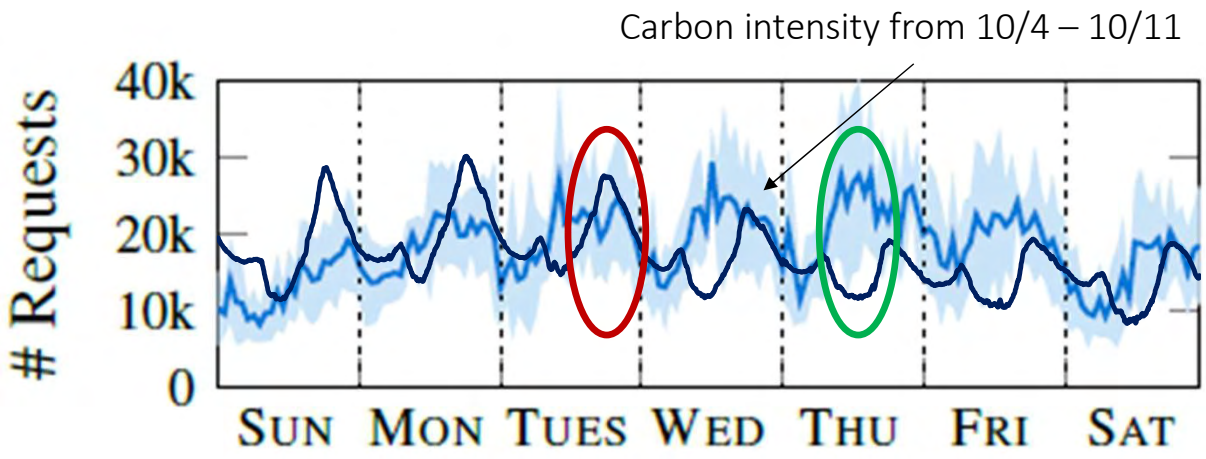
Example: Time shifting demand

A Microsoft hackathon project showed that the carbon footprint of AZ batch service in the Netherlands can be reduced by:

11% for jobs < 24 hrs

10% - 27% savings for jobs < 5 hours

VM Requests/hour
Averaged by day of week for 5 weeks



O. Hadary et. Al, "Protean: VM Allocation Service at Scale", OSDI, 2020

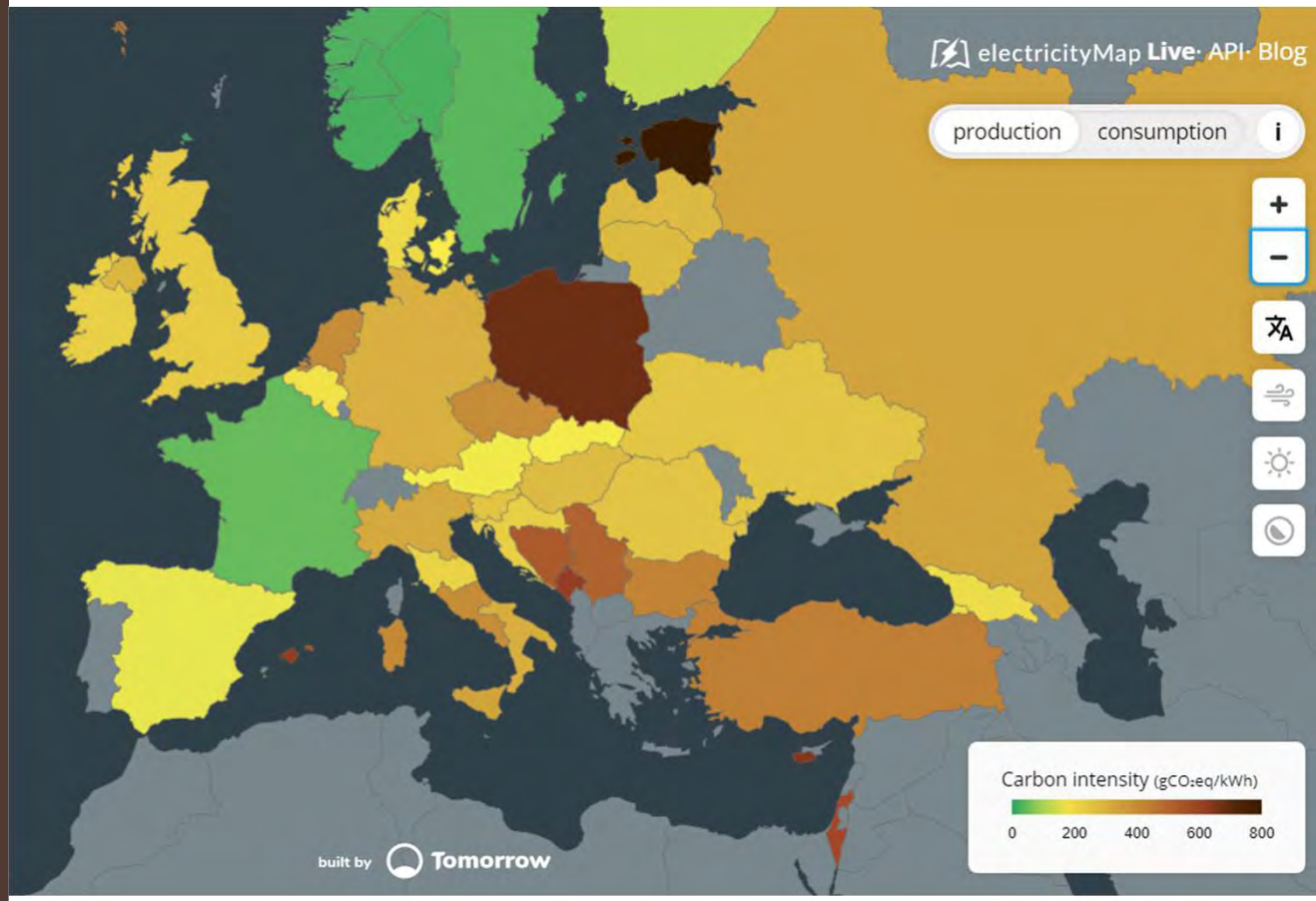
Load shifting complexities

Limited renewable availability

Overutilization of resources during low carbon intensity and/or in renewable friendly regions.

Underutilization in other data centers.

Location, Location, Location



Take aways

Renewable mix

Customer requirements

Hardware availability

Application needs

Data center utilization

VM power estimation
Demand shaping algorithm

Carbon intensity prediction
Job duration estimation

Checkpointing mechanisms
Delay tolerant applications

Sustainable silicon

Silicon manufacturing

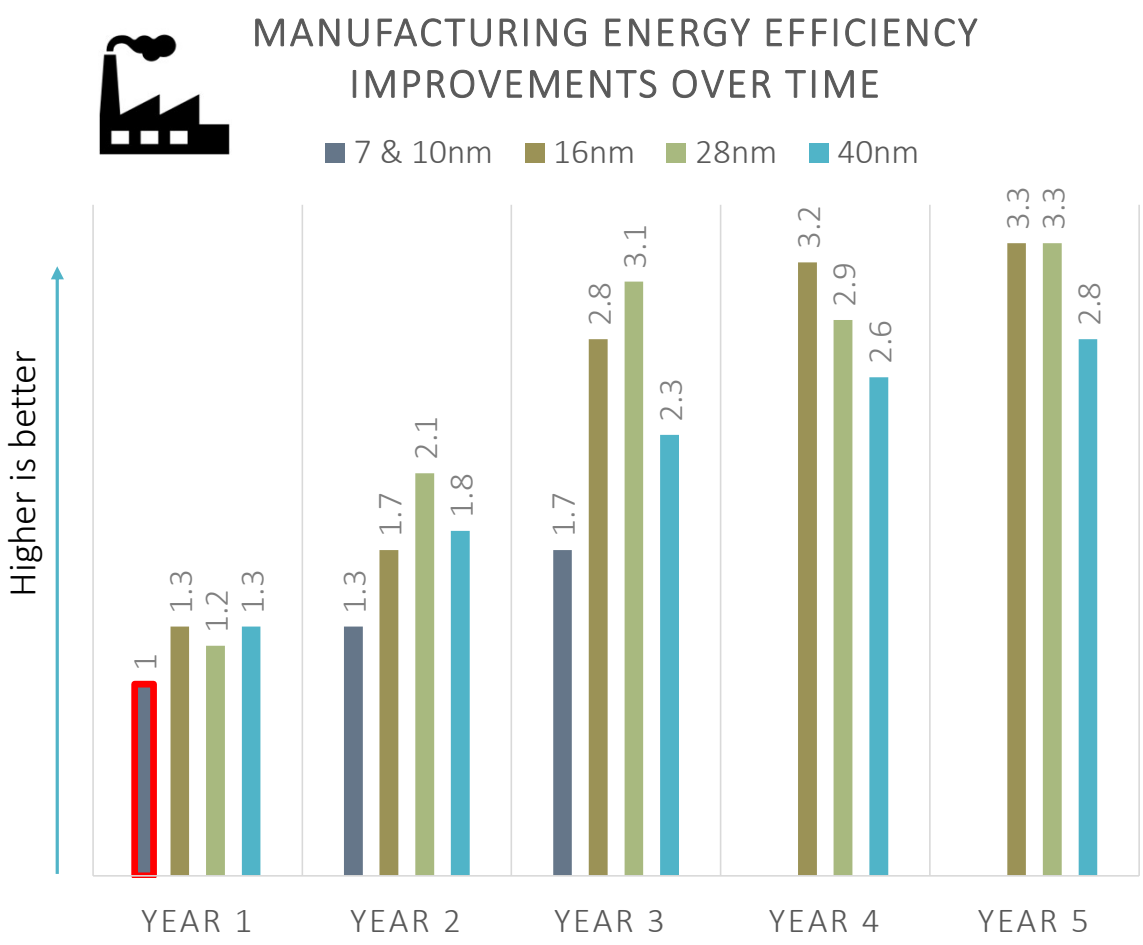
TSMC (2019)
- 14.1M tons CO2e
- 75% electricity (7% renewable)
- 25% from chemicals with high global warming potential (GWP)

Increasing cost and complexity

Diminishing power and performance improvements

Metric	16nm -> 7nm +	7nm – > 5nm*
Power	65% ↓	30% ↓
Performance	35% ↑	15% ↑
Density	3.3X ↑	1.84X ↑
Wafer Cost †	2.34X ↑	1.82X ↑

+ 7nm numbers from https://en.wikichip.org/wiki/7_nm_lithography_process
* Andrei Frumusanu, "TSMC Details 3nm Process Technology: Full Node Scaling for 2H22 Volume Production", Aug 2020, AnandTech.com (https://www.anandtech.com/show/16024/tsmc-details-3nm-process-technology-details-full-node-scaling-for-2h22)
† Saif M. Khan and Alexander Mann, "AI Chips: What They Are and Why They Matter" (Center for Security and Emerging Technology, April 2020), cset.georgetown.edu/research/ai-chips-what-they-are-and-why-they-matter/."AI Chips: What They Are and Why They Matter", Saif Khan, Alexander Mann



Annual improvement in manufacturing energy efficiency

TSMC Corporate Social Responsibility Report, 2019

USE IT LONGER

MAKE IT CHEAPER
(IN TERMS OF CARBON)

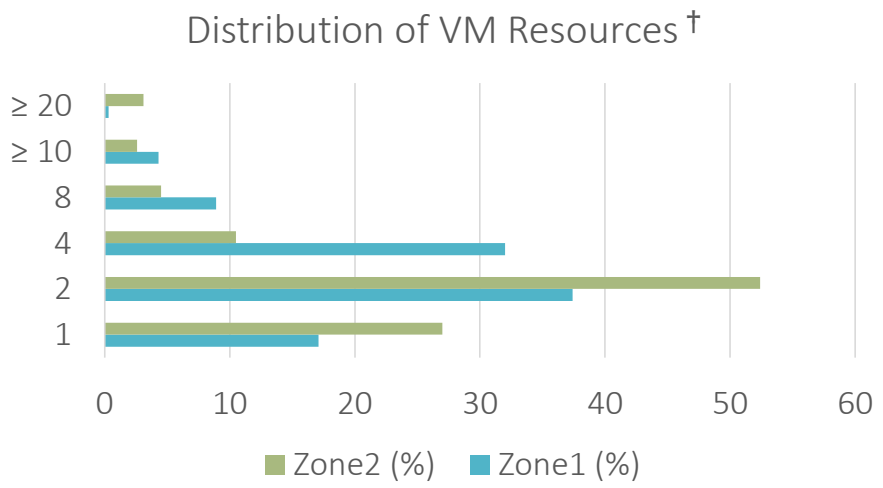
DO MORE WITH LESS
(INCREASE UTILIZATION)

Key points

Hardware vs VM viewpoint

Hardware Perspective:

- 64+ cores
- Coherent LLC
- 8+ memory channels
- 256+ GB memory capacity
- Dual socket systems



85% of VMs are <= 4 virtual cores
70% of VMs require < 4GB memory*

† O. Hadary et. Al, "Protean: VM Allocation Service at Scale", OSDI, 2020 (https://www.usenix.org/conference/osdi20/presentation/hadary)
* E. Cortez et. Al, "Resource Central: Understanding and Predicting Workloads for Improved Resource Management in Large Cloud Platforms", SOSP'17

Don’t throw the baby out with the bathwater

Resilient hardware and software

Increase usability by adding spares or tolerating defects

- Spare cores
- Spare cache capacity

Detect, isolate, and manage faulty components

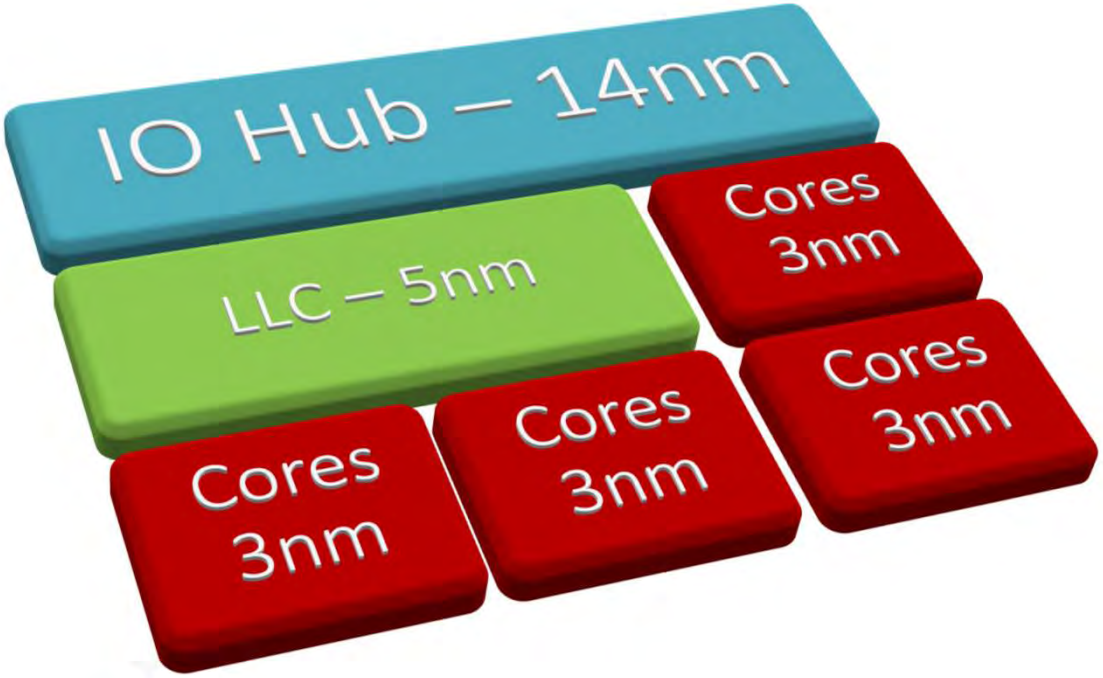
- Graceful repair and recovery
- Remove faulty components

Resource pooling & disaggregation

- Pool SSDs, DRAM, accelerators

Manage feature updates

- Modular HW design
- Software shim layer and/or microcode



Disaggregate the processor

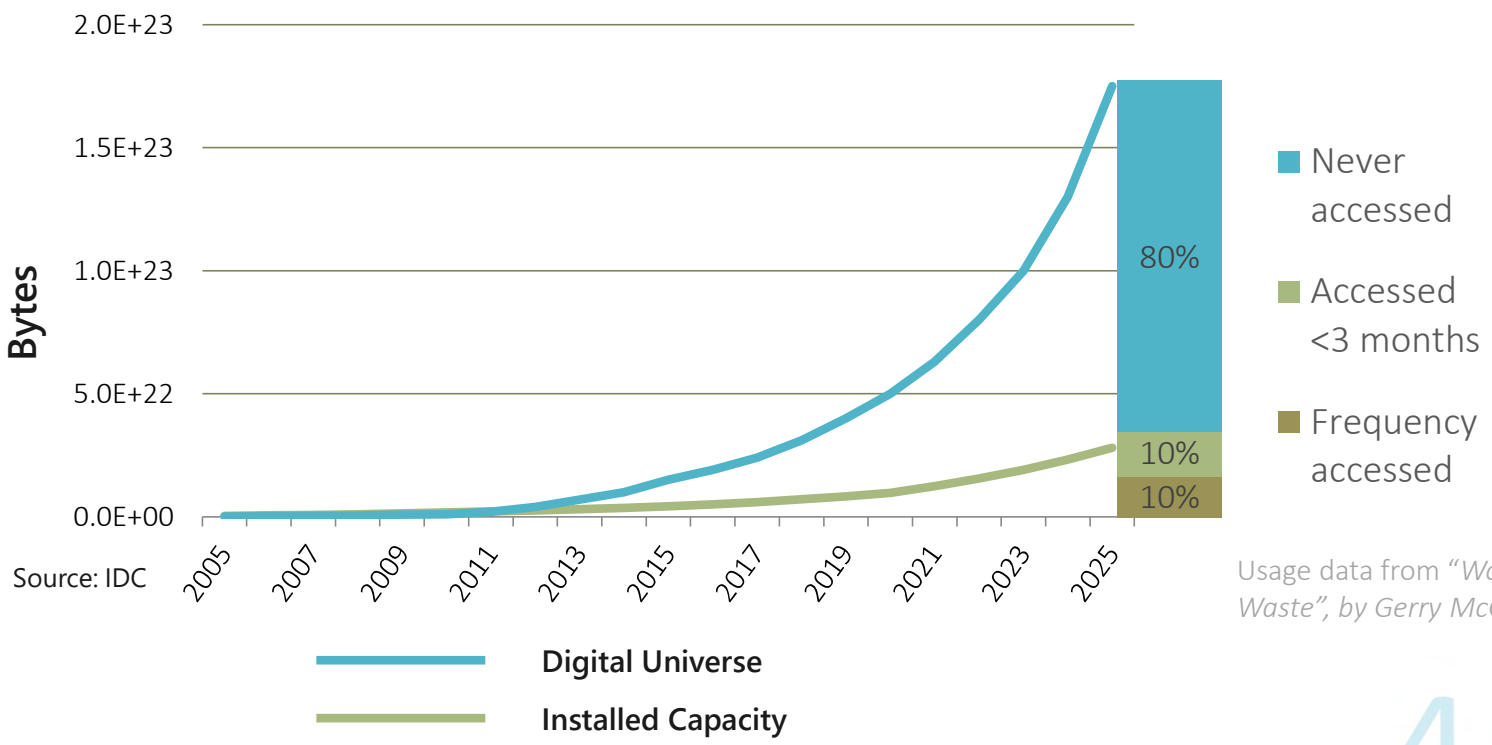
Sustainability benefits of chiplets:

- Less waste due to defects
- Optimal process
- Higher re-use

Examine architectural decisions through the lens of sustainability

Thinking outside the box… might lead you to different box

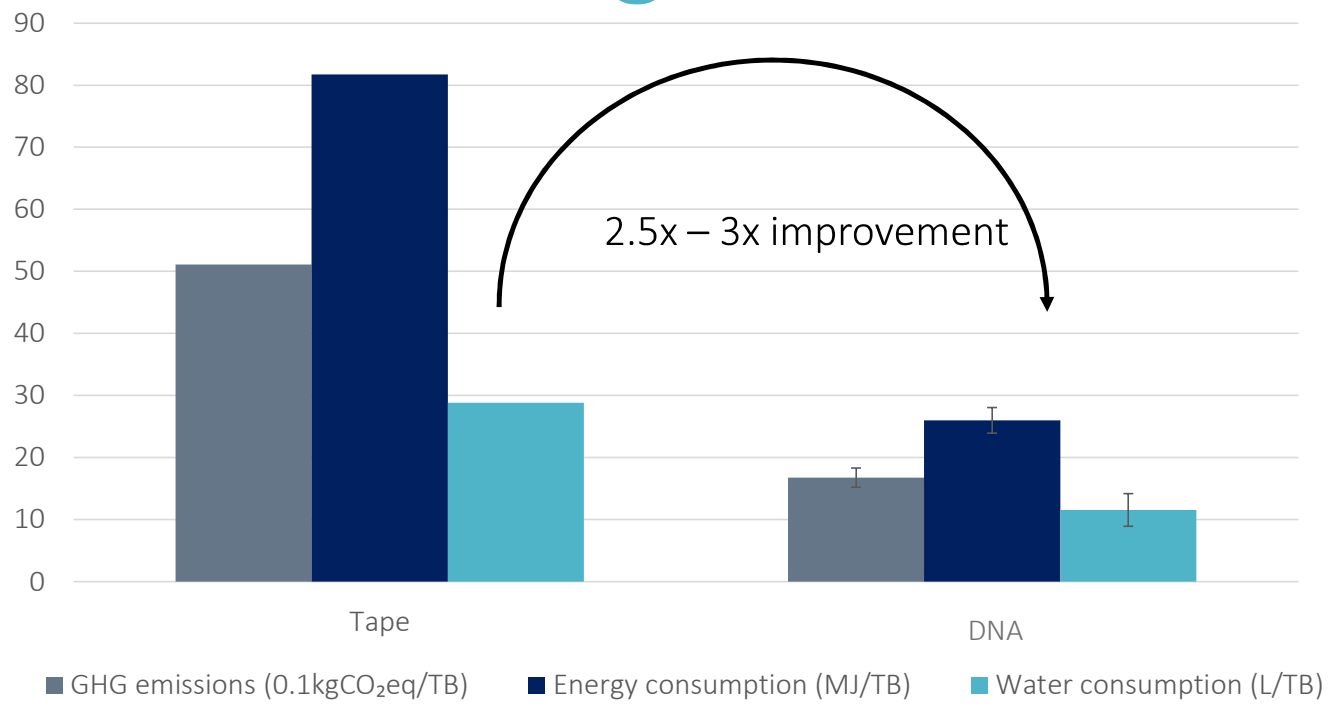
Storage requirements



Source: IDC

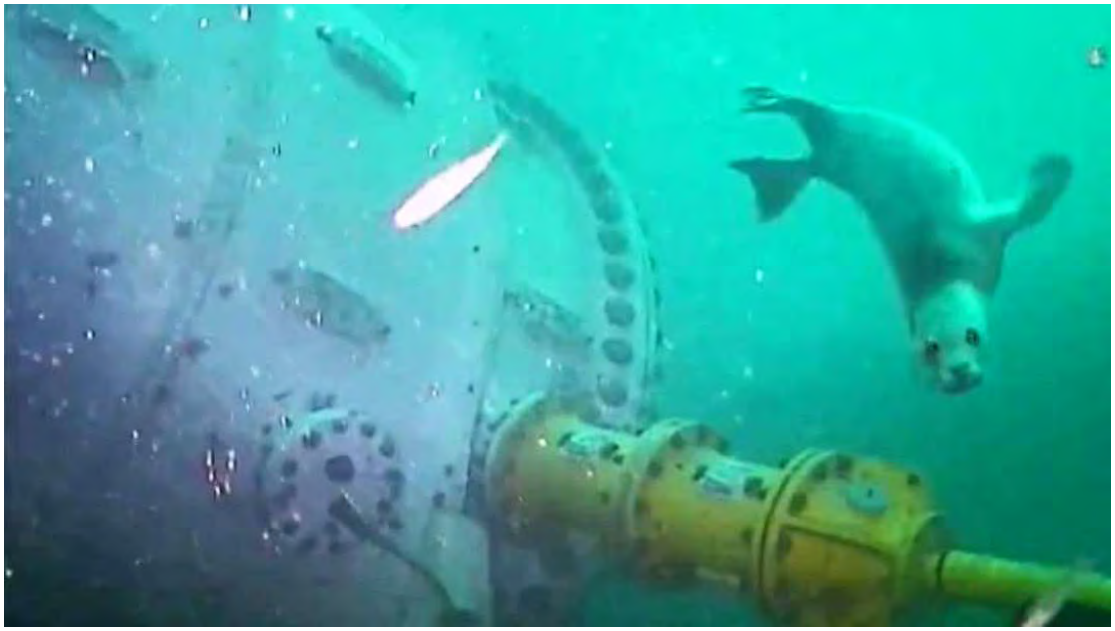
Usage data from "*World Wide Waste*", by Gerry McGovern

Sustainable storage



- DNA promises to be significantly more sustainable than tape

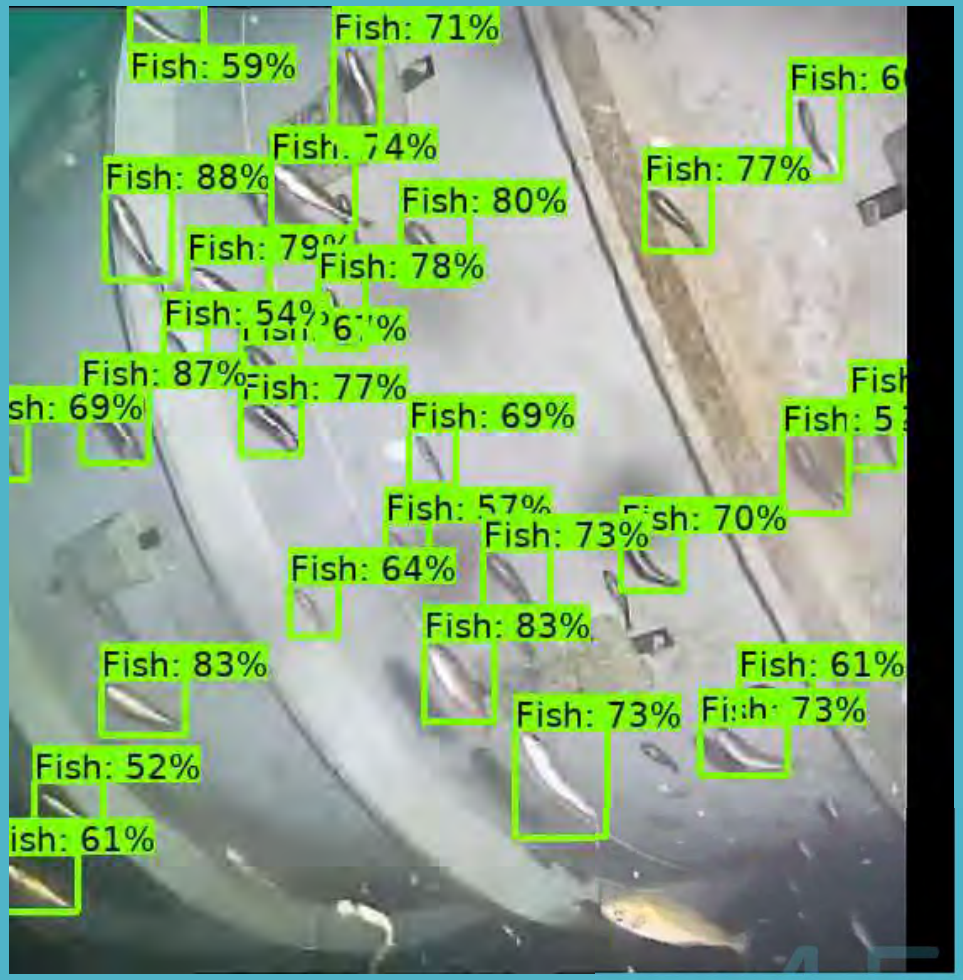
Nguyen et al., Electronics Goes Green, 2020



Project Natick

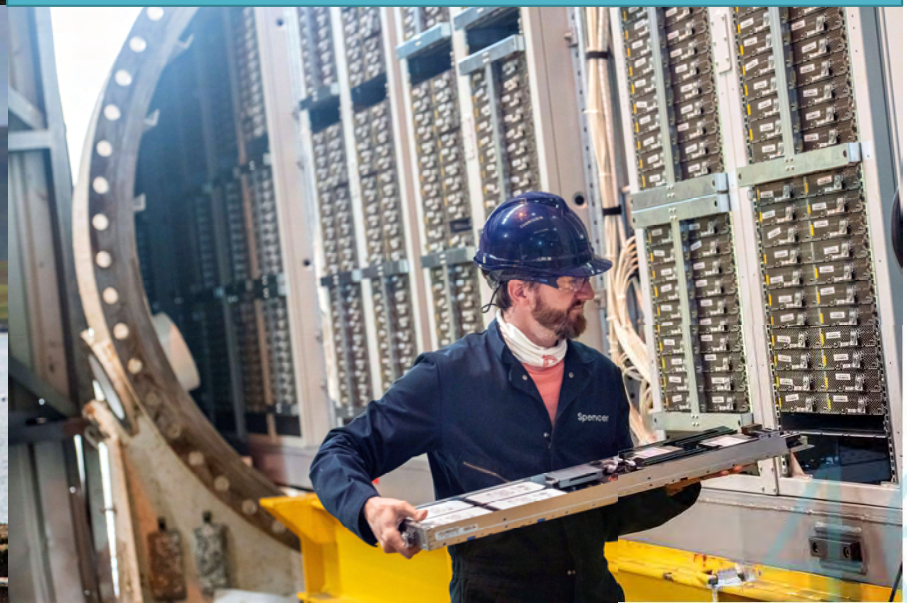
- 864 servers
- Sealed tank filled with nitrogen
- Deployment took 1 day
- 117 feet underwater

AI for sustainability



1/8 the failure rate

- Very tight operating envelope
- No humidity and nitrogen filled tank
- Isolated conditions



Full Immersion

Liquid Immersion Optimized Servers

M. Shaw, H. Alissa, “Liquid Immersion Optimized Servers”, 2020 OCP Virtual Summit

The bigger picture

Water

Reduce water use

Restore wetlands

Remove contaminants

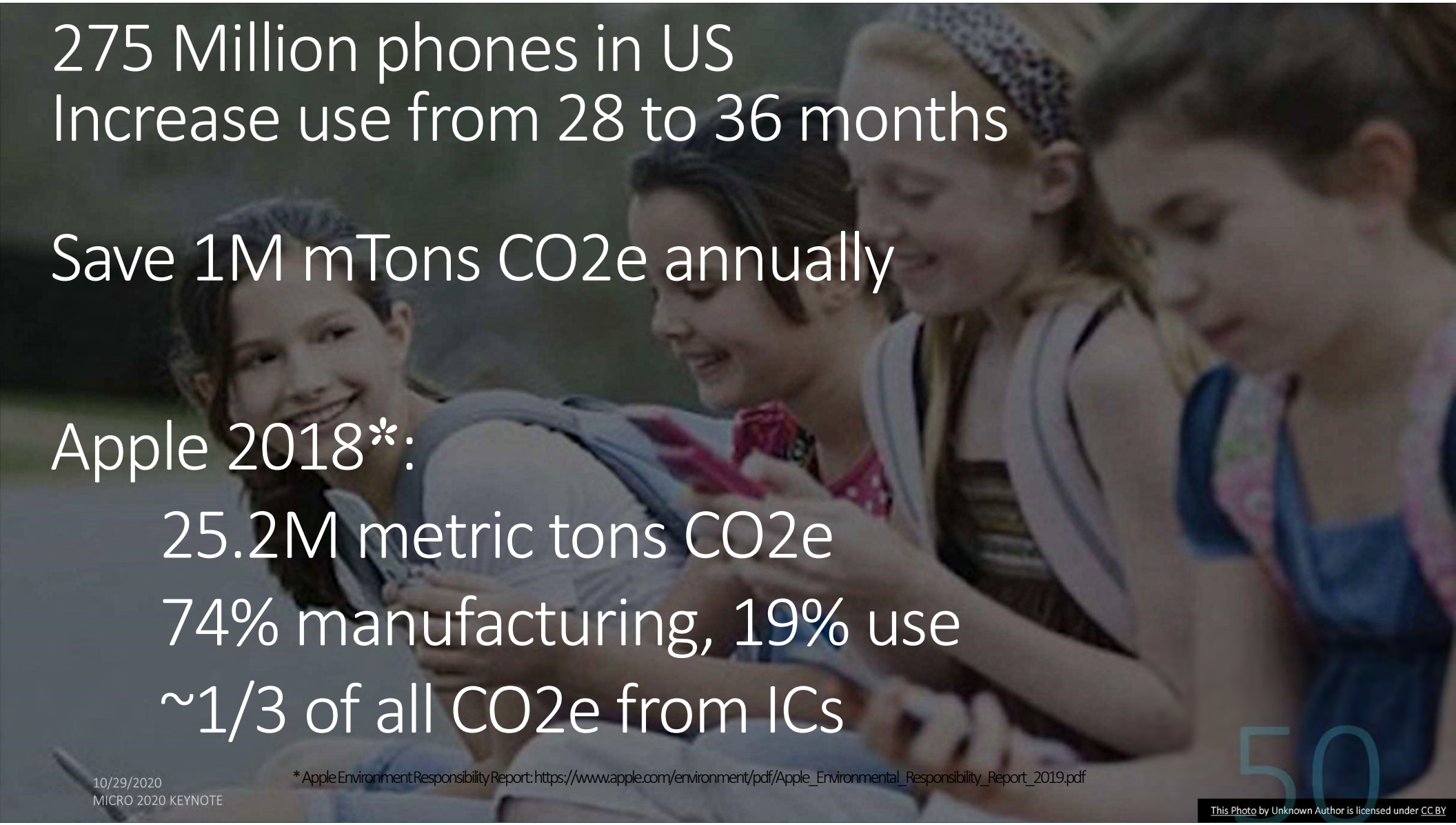
Waste

56M metric tons of e-waste generated every year

Electricity distribution

24x7 carbon free

In addition…



275 Million phones in US
Increase use from 28 to 36 months

Save 1M mTons CO_2e annually

Apple 2018*:

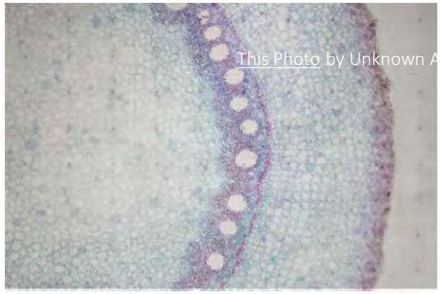
- 25.2M metric tons CO_2e
- 74% manufacturing, 19% use
- ~1/3 of all CO_2e from ICs

* Apple Environment Responsibility Report: https://www.apple.com/environment/pdf/Apple_Environmental_Responsibility_Report_2019.pdf

Datacenters reimagined

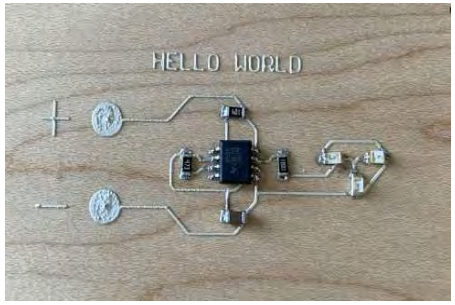
DNA data storage

Green concrete



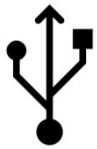
Biofueled electricity

Compostable electronics



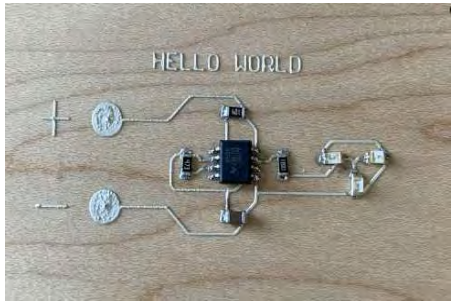
Nguyen et al., Electronics Goes Green, 2020

Datacenters reimagined

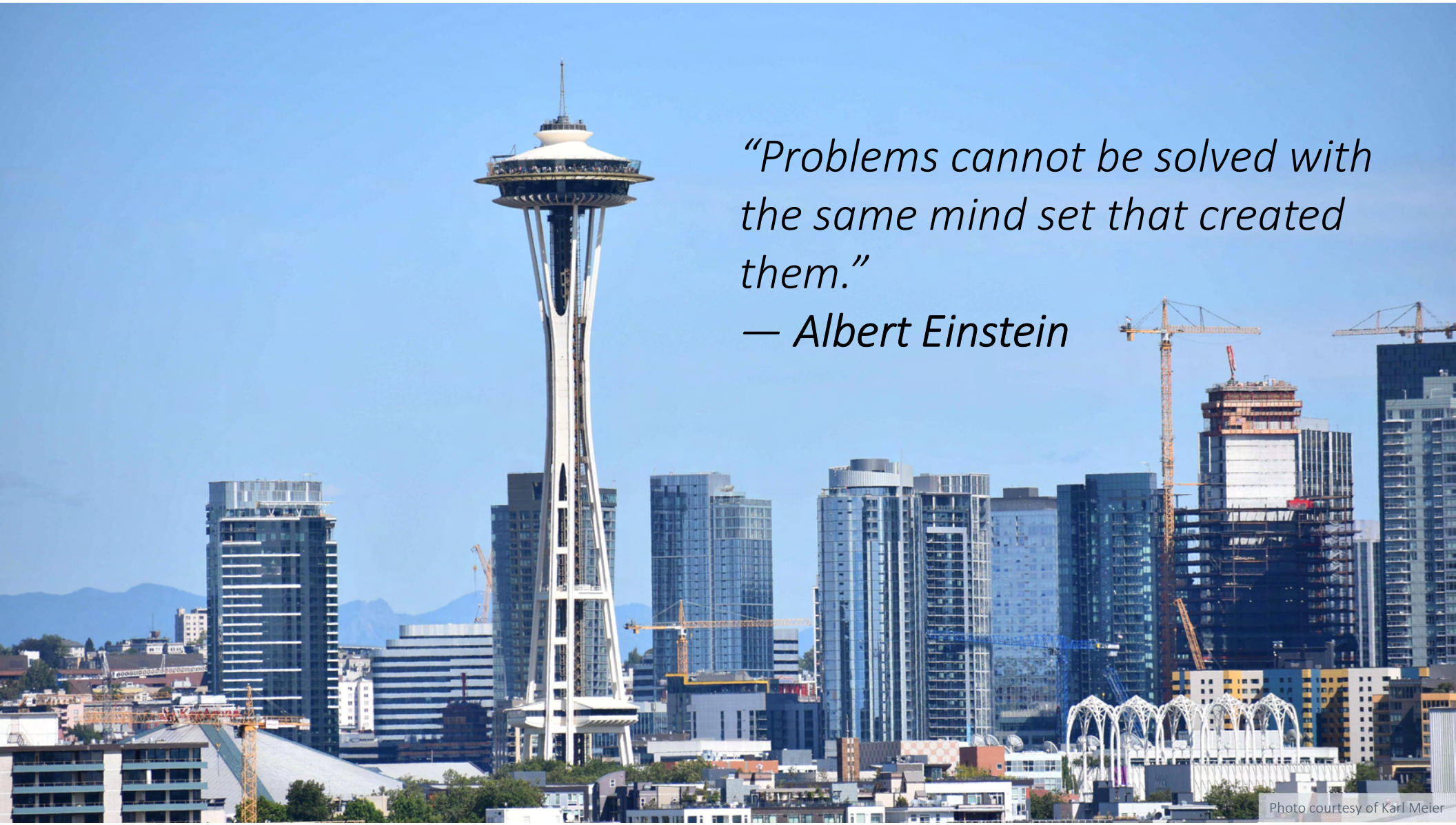


DNA data storage

Compostable electronics



Nguyen et al., Electronics Goes Green, 2020


“Problems cannot be solved with the same mind set that created them.”
— Albert Einstein
Photo courtesy of Karl Meier

Materials

https://devblogs.microsoft.com/sustainable-software/

http://vlsiarch.eecs.Harvard.edu/research/sustainable-computing/

Udit Gupta (Harvard University)
Gu-Yeon Wei (Harvard University)
David Brooks (Harvard University)
Carole-Jean Wu (Facebook)
Hsien-Hsin Sean Lee (Facebook)

Acknowledgements

Doug Carmean
Ehsan Nasr
Husam Alissa
Will Buchanan
Kari Lio
Winston Saunders
Sean James
Bill Johnson
Alex Bitiukov
Fred Constantino
Lorraine Tew
Spencer Fowers
Ben Cutler
Kesava Viswanathan
Robin Smith
Ram Huggahalli
Taniya Saddiqua
Ishwar Agarwal
Robert Henry
Gagan Gupta
Anant Deval
Esha Choukse
Bharadwaj Yadavalli
Asim Hussain
Scott Chamberlin
Conor Kelly
Abhisek Pan
Anjana Krishnan
Karin Strauss
Bichlien Nguyen