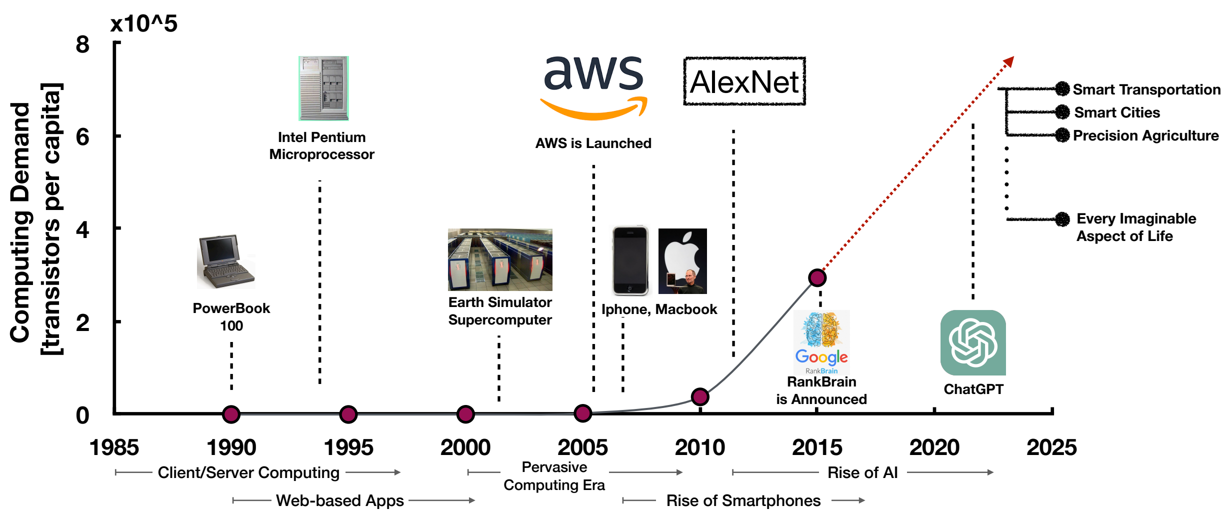
Sustainable Computing

& Computing for Sustainability

Noman Bashir - University of Massachusetts Amherst
nbashir@umass.edu
02/22/2023

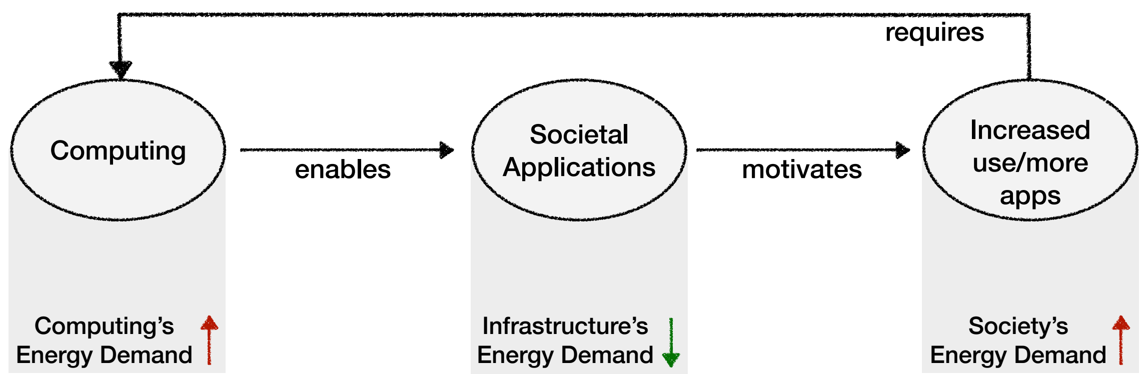
Computing's Demand is Growing Exponentially

- Society continues to find useful applications

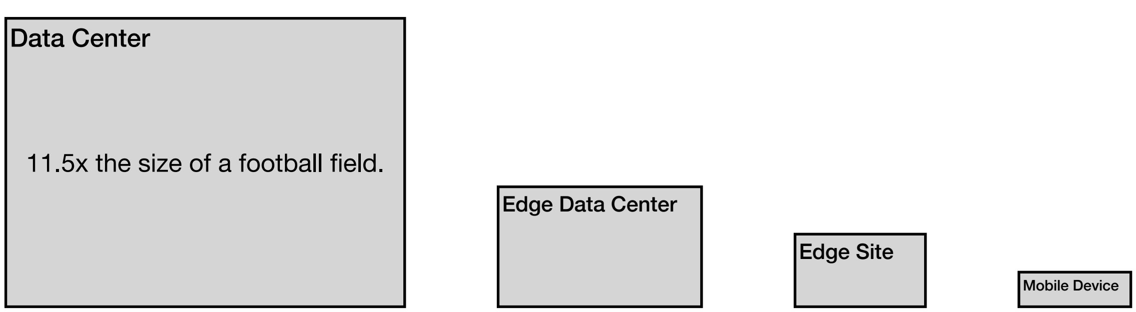


Source: "Unimaginable Output: Global Production of Transistors" - Darrin Qualman

Implications of Increasing Computing Demand



How is Computing Demand Served?



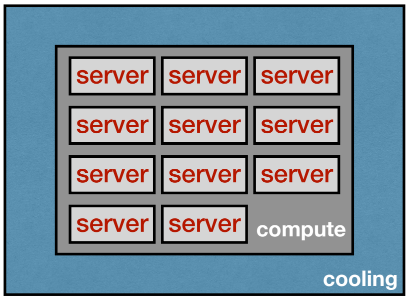
Thousands of servers and data storage, e.g., Google Dalles data center houses **~100k servers** and consumes **100MW** of power (enough for a small city)

10s-100s of servers and data storage, **1,000 sqft to 50,000 sqft** a **few kW** to a **few MW**

Mobile devices and small storage hand-held etc. **a few watts**

note: figures are not drawn to scale.

What Contributes to Data Center's Cost, Energy, Carbon Footprint?



Cost

- **Servers:** Cost a lot and are replaced every 3-5 years.
- **Building:** Capital investment, depends on location.
- **Energy:** Major cost of datacenter, depends on location.

Energy

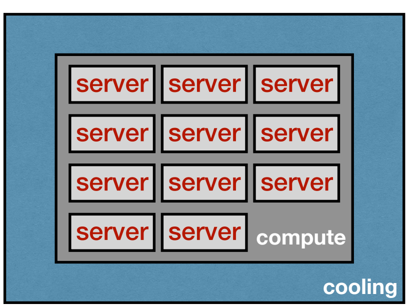
- **Computing:** Become more energy efficient over time.
- **Cooling:** Wasted energy, significantly reduced over years.

Carbon

- **Embodied:** Carbon emissions from manufacturing/building.
- **Operational:** Emissions from energy use for compute and cooling.

How to Serve Computing's Demand in a Sustainable Manner?

Sustainable —> least carbon intensive way.



Carbon

- **Embodied:** Carbon emissions from manufacturing/building.
- **Operational:** Emissions from energy use for compute and cooling.
  - From the energy used to **run** the servers.
  - From the energy used to **cool** the servers.

Reduce Embodied and Reduce Operational Emissions

$$\text{Carbon Footprint} = \frac{\text{Cycles per Unit Work} \times \text{Total Units of Work}}{\text{Computing's Energy Efficiency} \times \text{Energy's Carbon Efficiency}}$$

$$\text{Carbon Footprint} = \frac{\text{10 cycles per inference request} \times \text{100 inference requests}}{\text{5 cycles per kWh} \times \text{1 kWh per gCO2eq}}$$

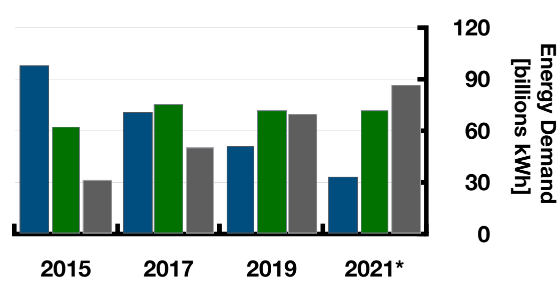
$$\text{Carbon Footprint} = \boxed{\text{200 gCO2eq}}$$

History: Driving Factors Behind Innovations in Data Centers

Cost of Energy Has Been Driving Innovation

- Assume 100,000 servers
- **Monthly cost of 1 server**
  - 500W server
  - Cost = (Watts X Hours / 1000) * cost per kWh
  - Always-on server monthly cost = **$50**
- **Monthly cost of 100k servers = $5M**
- What about the cost of cooling?
  - Use Power Usage Effectiveness (PUE)
  - PUE = 2 —> double the cost
  - **PUE =1.2 —> 10% extra on $5M ($6M)**

Shift from Traditional Data Centers to Cloud



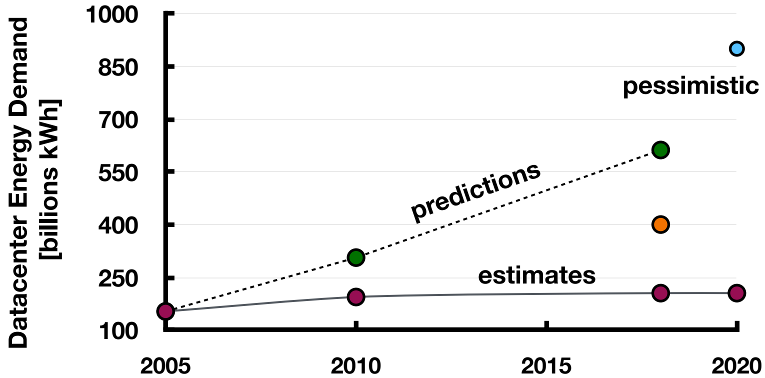
Traditional   Cloud (non-hyperscale)   Hyperscale

PUE: ~1.57   PUE: ~1.1

Source: Global data centre energy demand by data centre type, 2015-2021 - IEA

Energy Efficiency Gains Moving Forward

- Most optimistic estimates suggest 6% increase from 2010-2018



Demand Accelerating vs Energy-efficiency Gains Slowing Down

EPA Report to Congress on Server and Data Center Energy Efficiency (2007)
Recalibrating Global Data Center Energy-use Estimates - Eric Masanet (2020)
Efficiency Gains are Not Enough: Data Center Energy Consumption Continues to Rise Significantly - Ralph Hintemann (2018)

Algorithmic Efficiency can be further improved, but has limits

Industry has strong incentive to improve the **algorithmic efficiency**

Recent focus on ML training and Crypto-mining

[bounded]

[unbounded]

Datacenter capacity **increased by 6X** from 2010-2018

Crypto-mining and ML demand is **outpacing Moore's law**

Industry has strong incentive to **maintain and accelerate growth**

$$\text{Carbon Footprint} = \frac{\text{Cycles per Unit Work} \times \text{Total Units of Work}}{\text{Computing's Energy Efficiency} \times \text{Energy's Carbon Efficiency}}$$

[Koomey's Law: Energy efficiency doubles every 1.5-2.6 years**]** transition to cloud, dedicated hardware

[Laundar's Principle: Theoretical limit to be reached in 2050, practical sooner**]**

[Jevon's Paradox: Historically, gains in efficiency have not reduced demand**]**

[bounded]

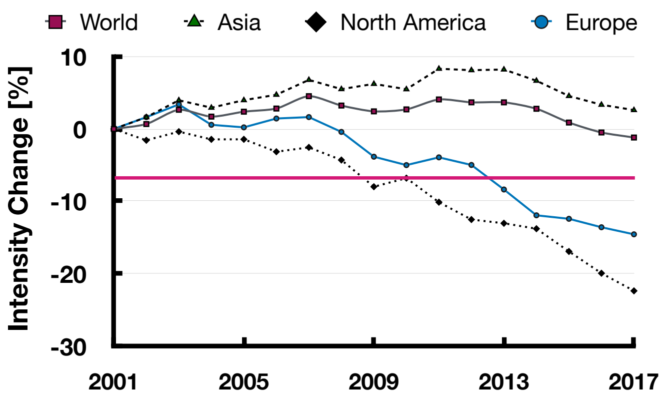
[unbounded]

Zero-carbon energy means **carbon efficiency can be infinite**

Industry has helped subsidize zero-carbon energy

Grid's Carbon Intensity Has Been Decreasing

- Energy's carbon efficiency in the US has improved by 45.6% over 2001-2017



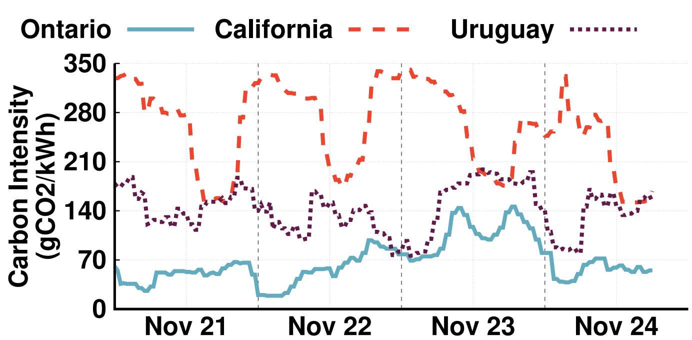
Carbon intensity may never truly reach 0gCO2eq per kWh.

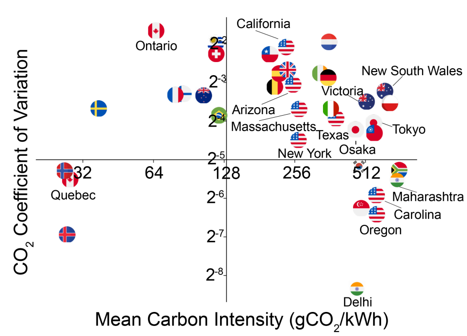
It may actually increase in parts of the world.

Source: Ember Global Electricity Review (2022)
Source: BP Statistical Review of World Energy
Source: Ember European Electricity Review (2022)

Clean Energy is Variable and Unreliable

- Carbon intensity variation: **less than 50g** to **more than 800g** across time and geographical regions.



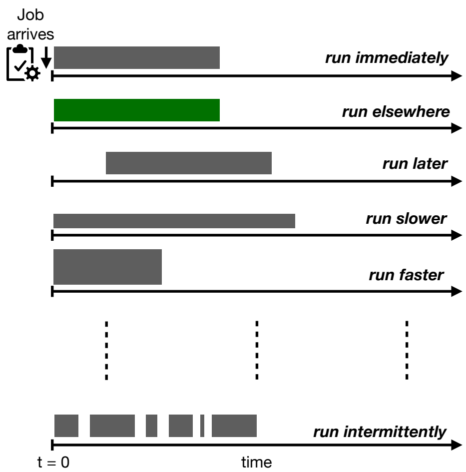


More regions in the world would look like Ontario in near future.

Source: electricityMap

The Good News: Computing's Unique Advantages

Driven by efforts to improve user experience & scale



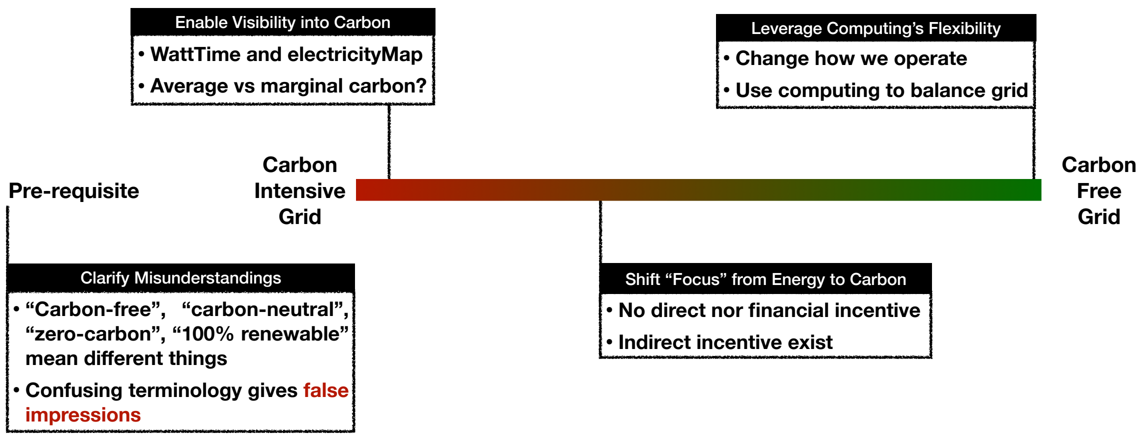
Accounting for and Reducing Embodied Carbon

- Carbon emissions from producing products or services, e.g., buildings facilities, manufacturing servers

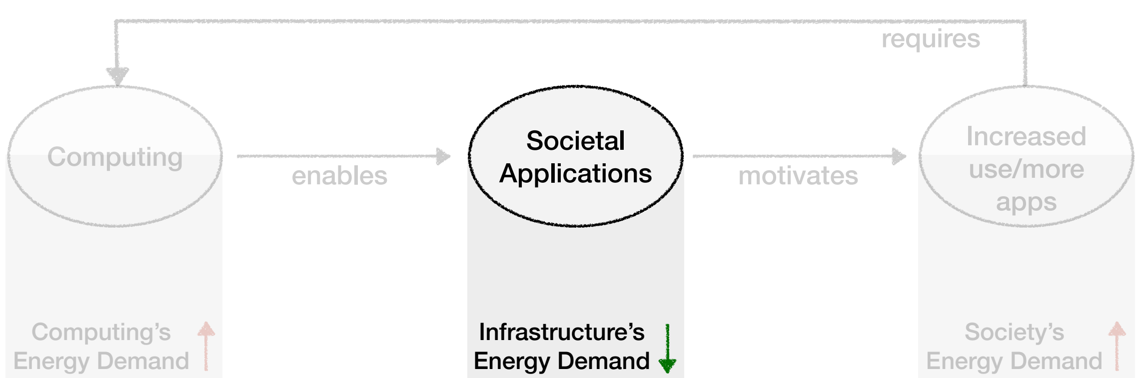
Embodied	Operational
• Your embodied is someone else's operation • Incentivizes buying less or buying different	• Operational is completely under your control • Operational emissions are not a solved problem

- Embodied and operational emissions are **NOT additive**
- One is **NOT** more **important** than the **other**
- Focus on embodied can **distract** from operational

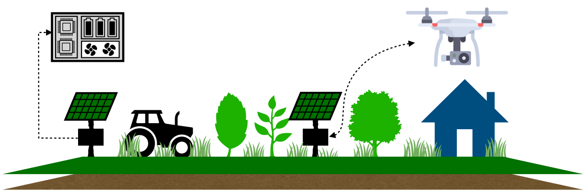
Implications for Sustainable Computing



Computing for Sustainability



Computing Use Cases



Improving Buildings and Transportation Sectors

- Building as an example of a distributed system
  - **Sense** monitor energy, temperature, occupancy etc.
  - **Analyze** data using computational tools.
  - **Control** lights, HVAC, doors to reduce energy usage.

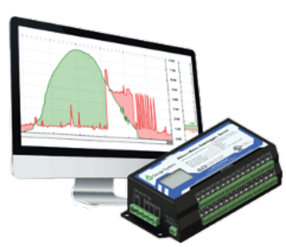
- Transportation as an example of a distributed system
  - **Sense?**
  - **Analyze?**
  - **Control?**

- Agriculture as an example of computing use case
  - **Sense?**
  - **Analyze?**
  - **Control?**

Building Monitoring

- Power metering at different levels
  - Outlet-level monitoring
  - Meter-level monitoring

Wemo smart plug



eGauge meter with interface

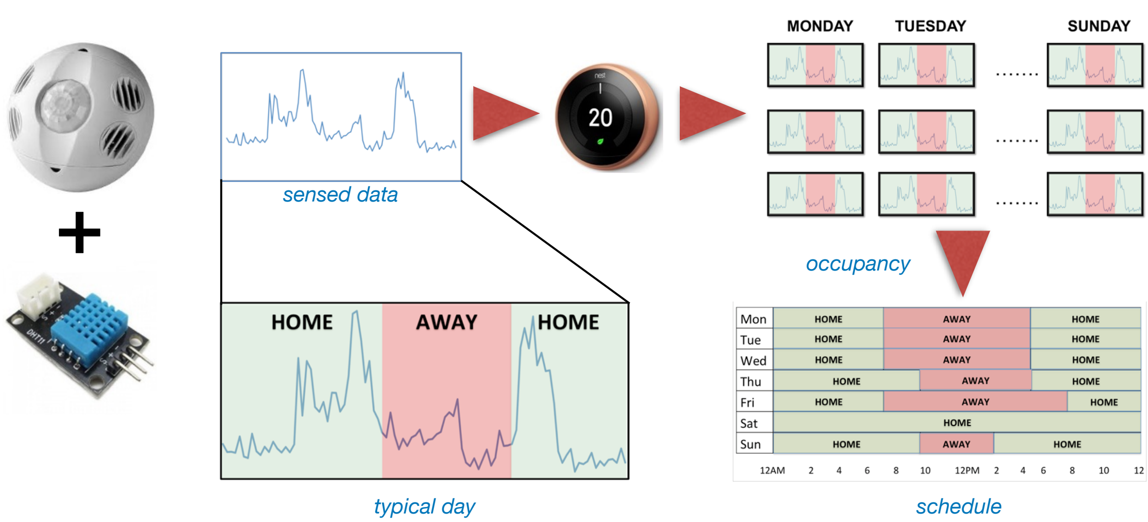
smart meter

Analyzing the data

- Energy monitors / sensors provide real-time usage data
  - Building monitoring systems (BMS) data from office / commercial buildings

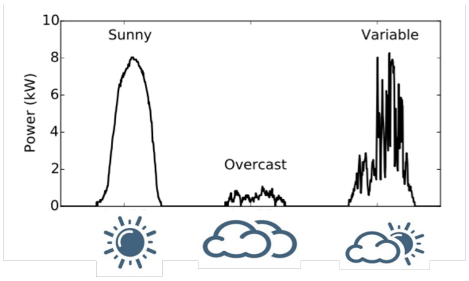
- Modeling, Analytics and Predictions
  - Use statistical techniques, machine learning and modeling to gain deep insights
  - Which homes have inefficient furnaces, heaters, dryers?
  - Are you wasting energy in your home?
  - Is an office building’s AC schedule aligned with occupancy patterns?
  - When will the aggregate load or transmission load peak?

Reduce Energy Use —> Learning Thermostat



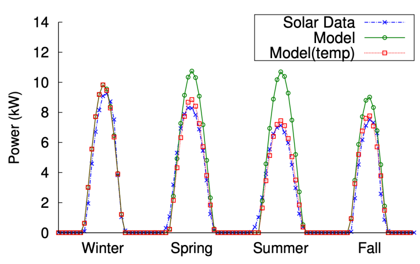
Use Low Carbon Energy —> Use Solar Power

- Significant growth in renewable energy adoption
  - Roof top wind turbines, solar PV, solar thermal (water heating)
- Highly intermittent
  - Impacted by cloud cover, temperature, environmental variables



Forecasting Solar Energy

- Predictive analytics to model and forecast solar energy generation
  - Use machine learning and NWS weather forecasts to predict solar generation



- Better forecasts of near-term generation; “Sunny load” scheduling

Use Case - EV Charging

- Solar panels installed in parking lots, rest areas, paid garages
  - Possible use case in offices and car rental services
- Assumptions
  - Arrival/departure times for EVs
  - Accurate solar predictions
- Intelligent charging
  - When to charge?
  - Which EV to charge?
  - How much to charge?

Summary

- Sustainable Computing
  - Demand for computing is growing
  - Need to serve the demand sustainably
  - Energy efficiency gains reducing
  - Computing has unique advantages
  - Try to optimize computing's carbon efficiency
  - Reduce operational as well as emobodied carbon
- Computing for Sustainability
  - Leverage computing to reduce energy consumption
  - Leverage computing to enhance use of low carbon energy