



2024 U.S. Data Center Energy Use Report: Summary of Findings

Dr. Arman Shehabi
Staff Scientist

ELI Webinar on National Environmental Impacts of Data Center Proliferation

February 26, 2025



ENERGY TECHNOLOGIES AREA
BERKELEY LAB

Background

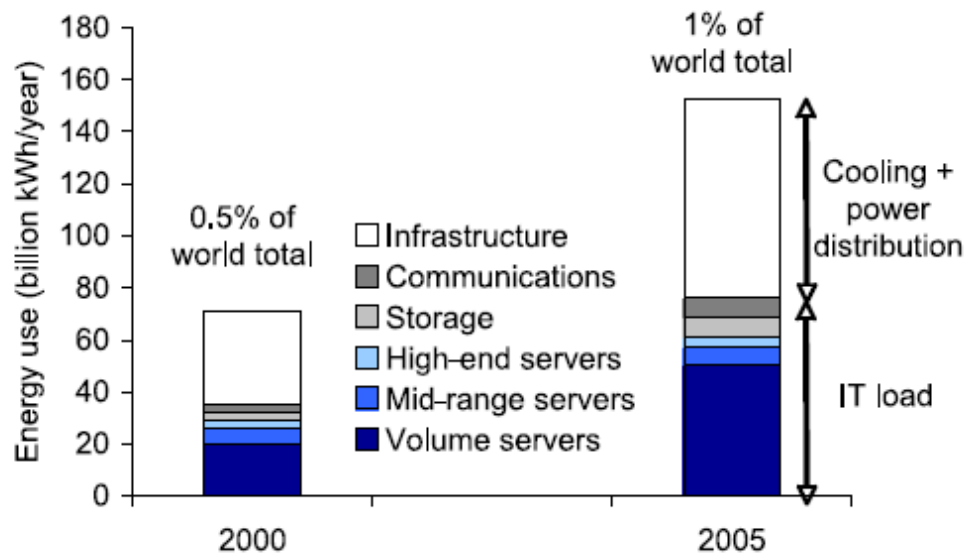


ENERGY TECHNOLOGIES AREA
BERKELEY LAB

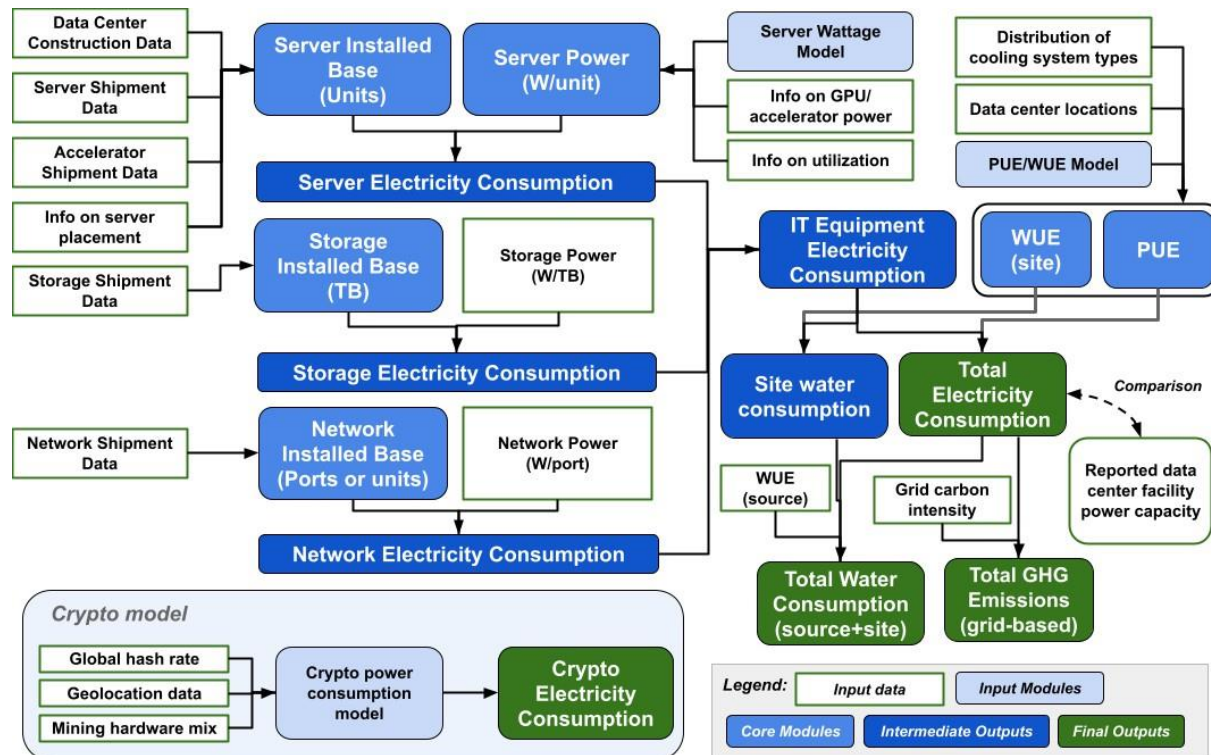
Berkeley Lab data center energy estimates in the 2000's



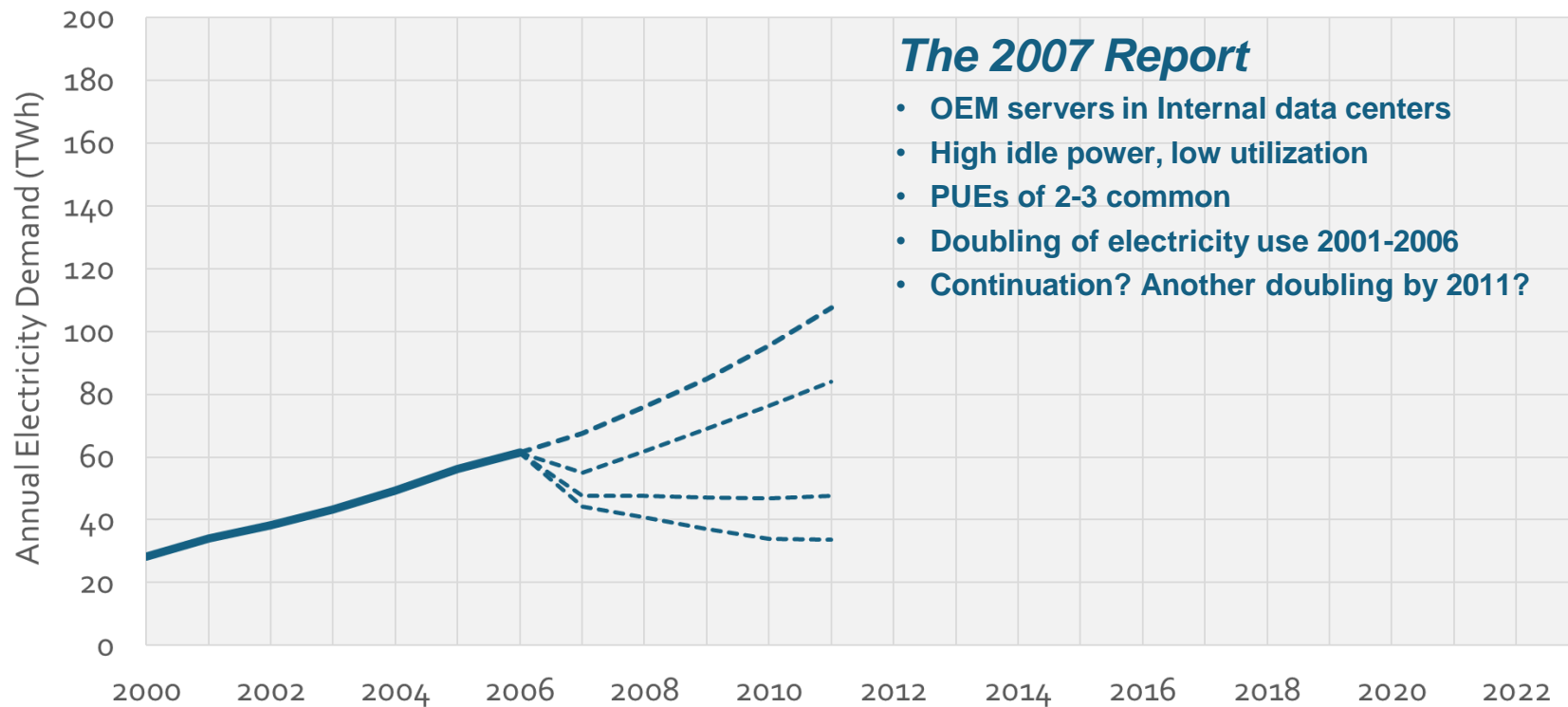
Data center electricity use doubled from 2000 to 2005



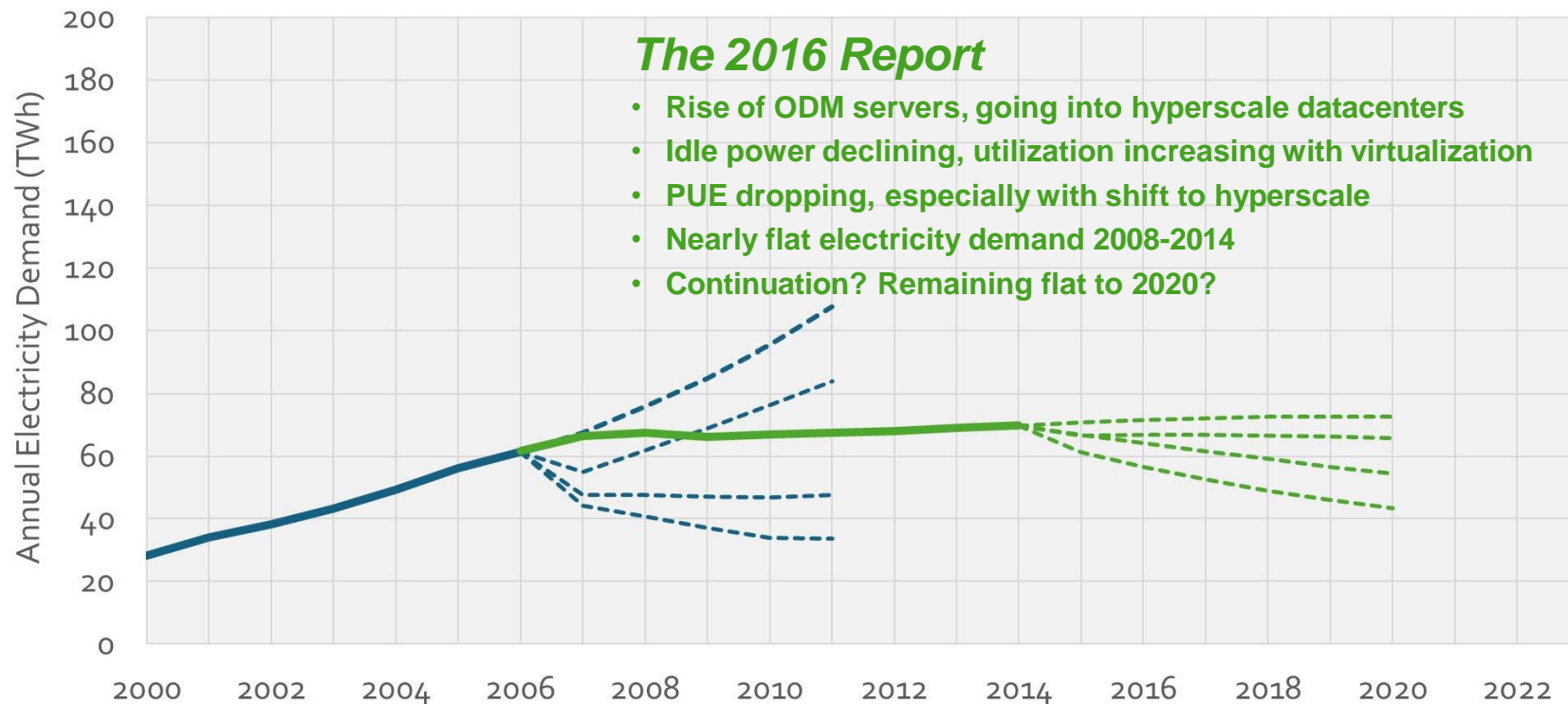
Berkeley Lab data center energy use modeling



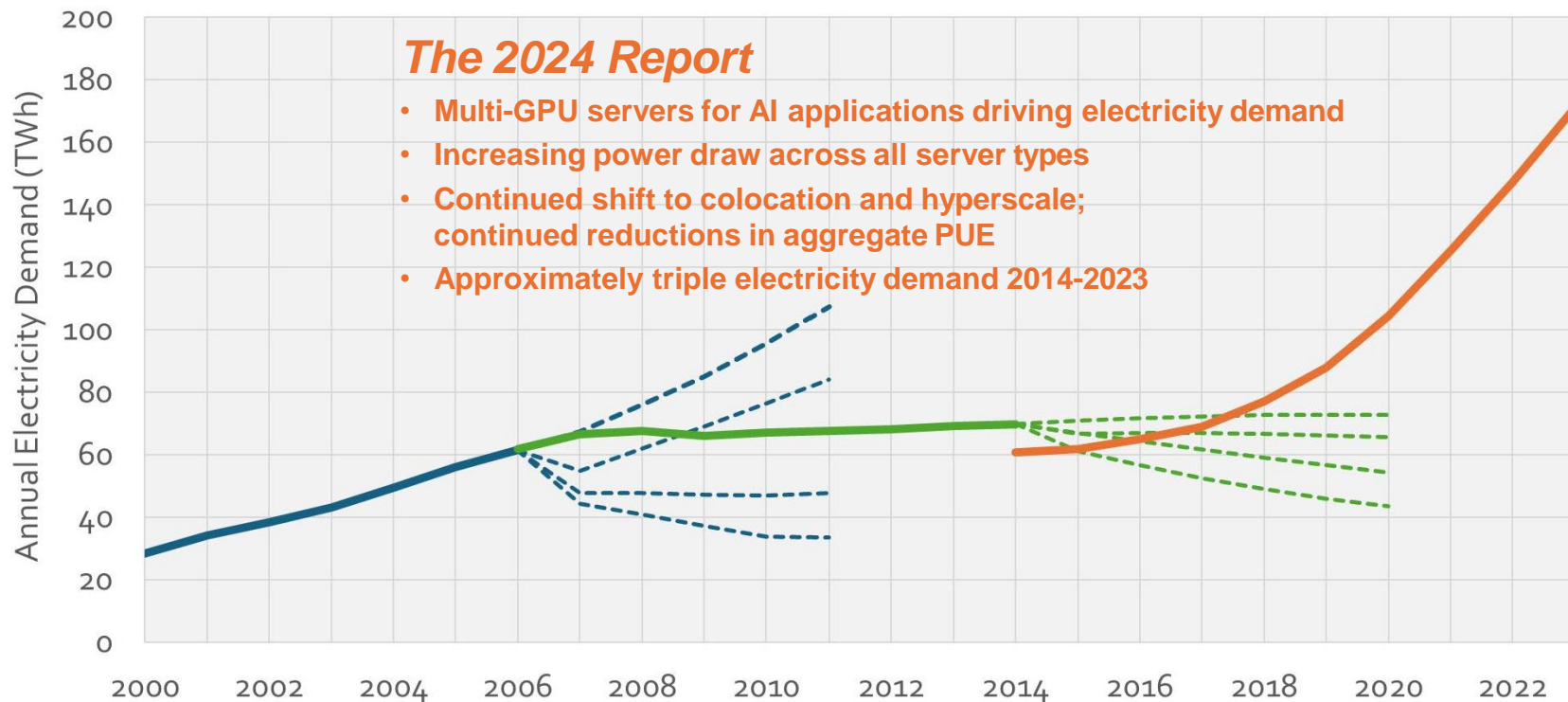
US Data Center Reports



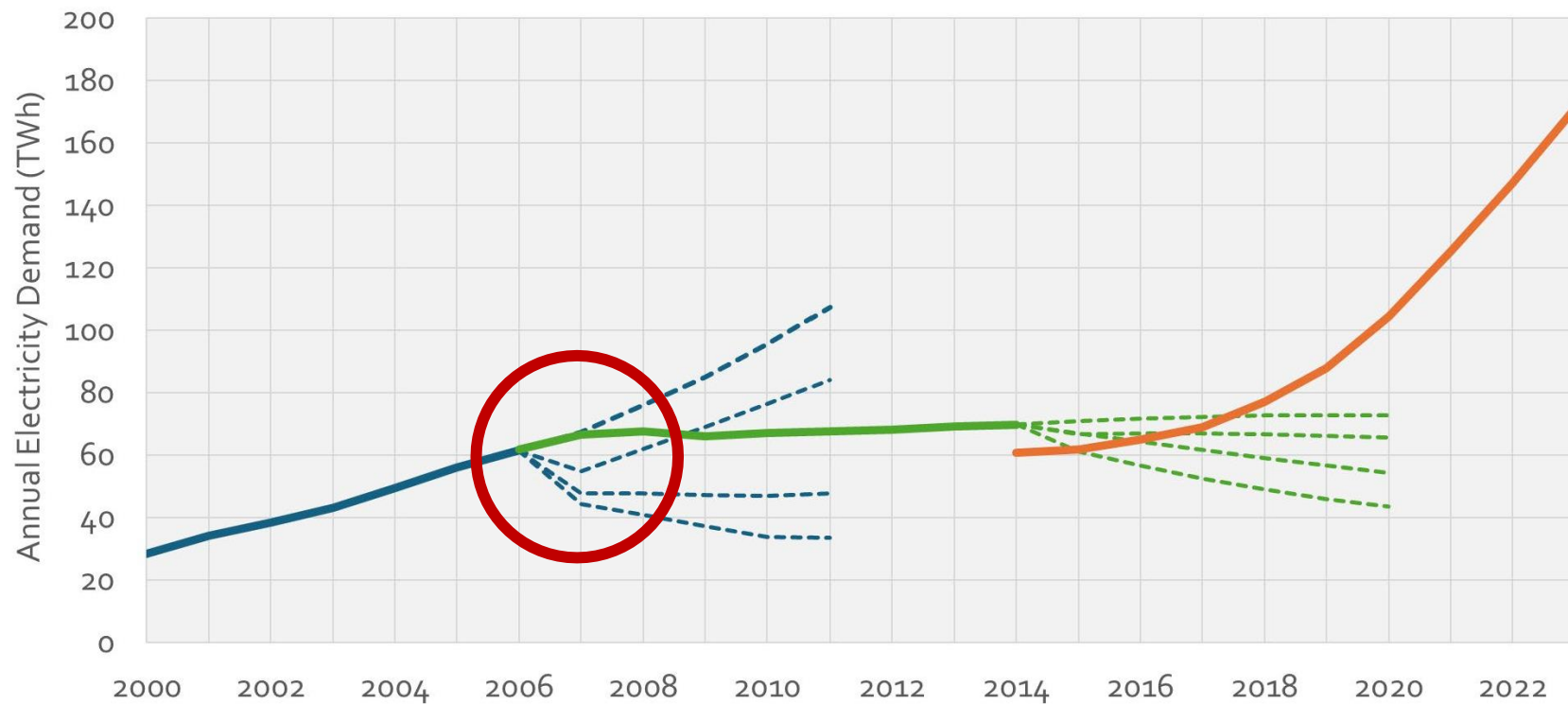
US Data Center Reports



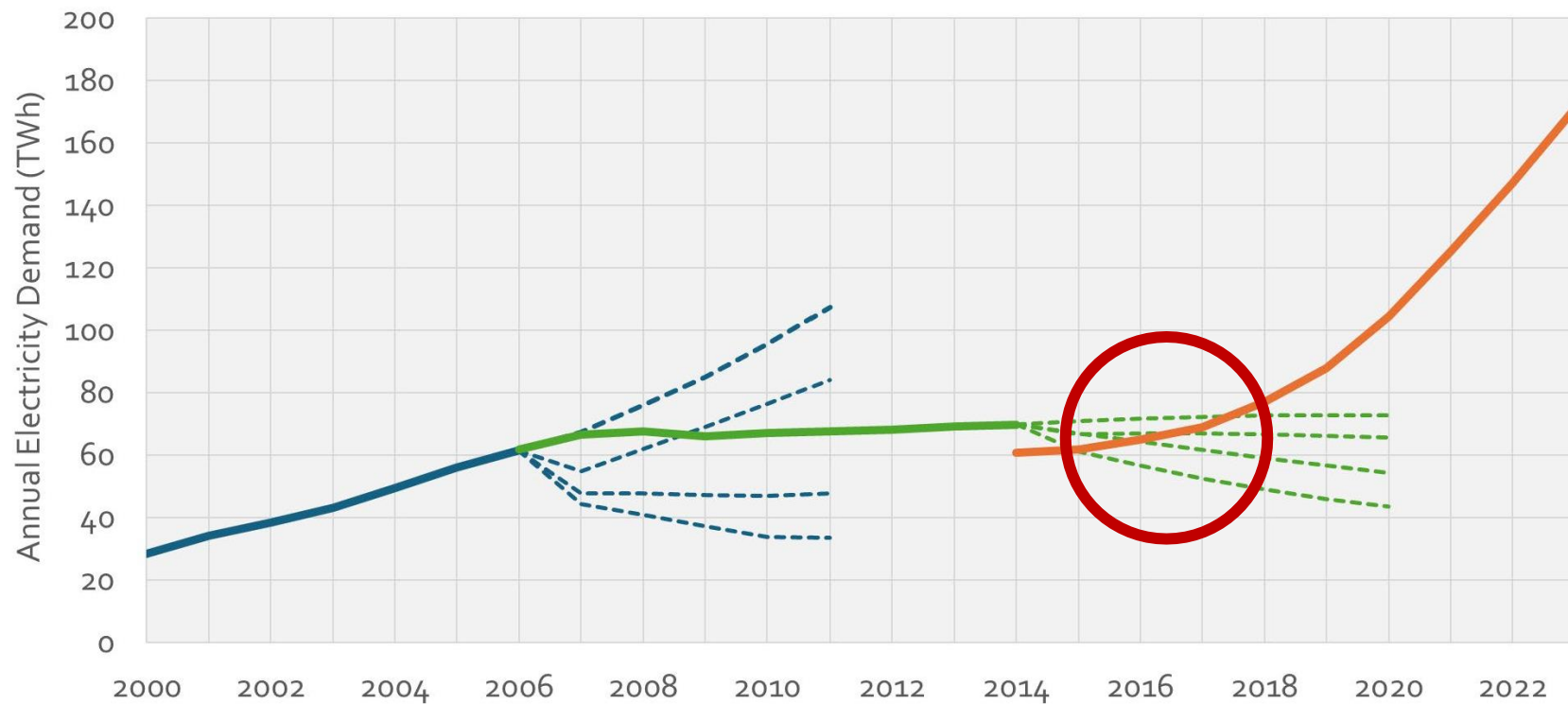
US Data Center Reports



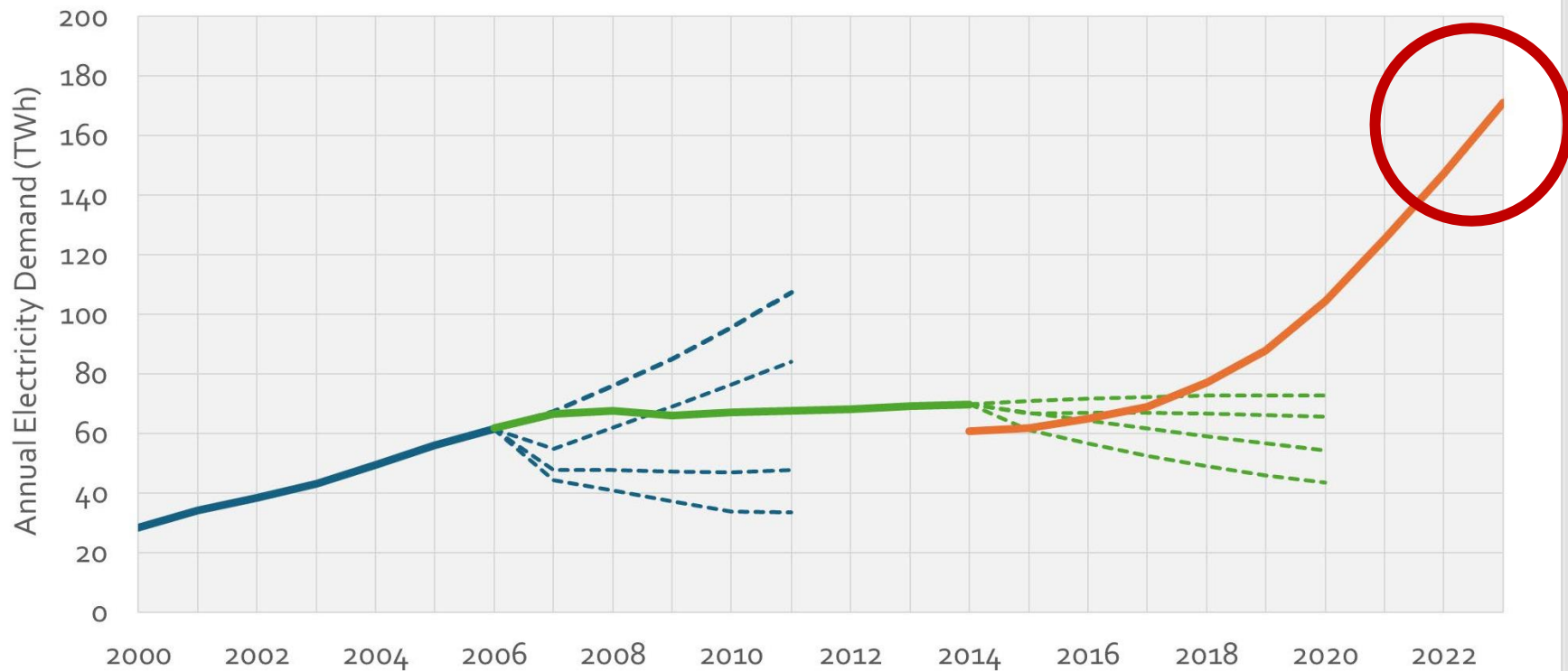
Turning points



Turning points



Turning points



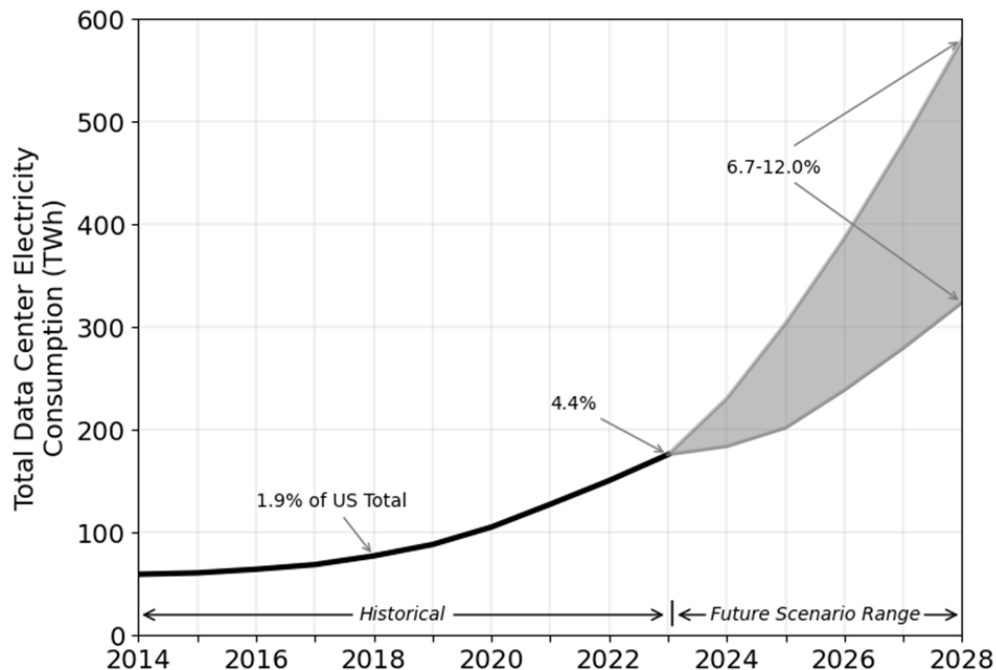
Executive Summary



ENERGY TECHNOLOGIES AREA
BERKELEY LAB

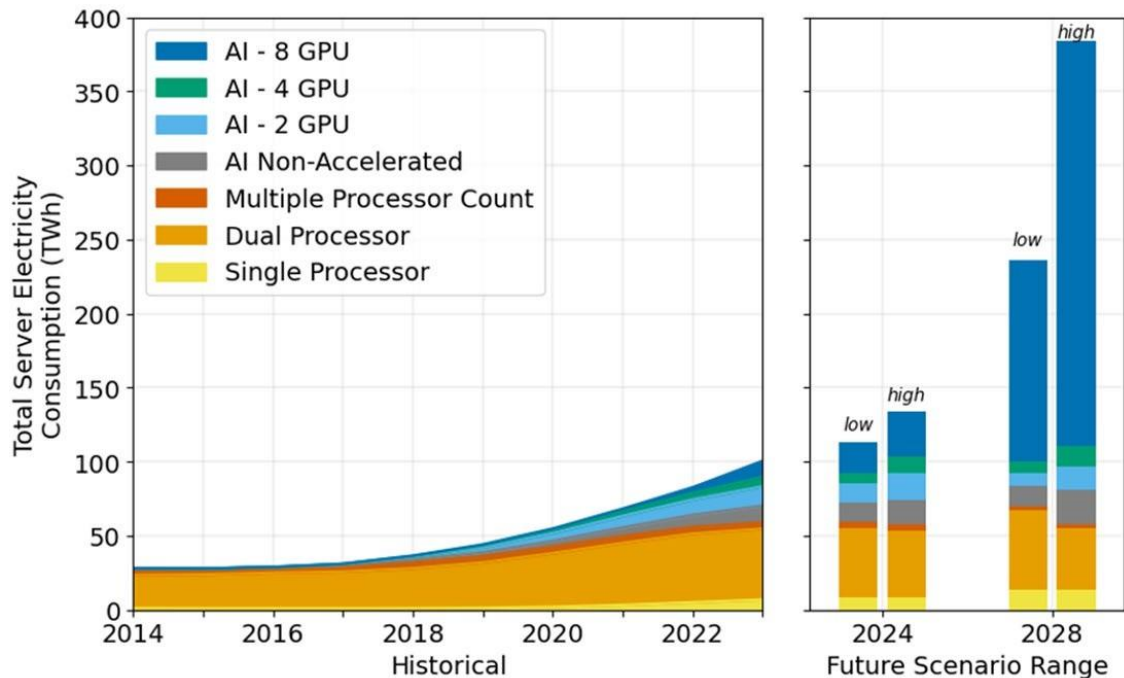
Total Data Center Estimates

Total electricity use 176 TWh in 2023; modeled range 325-580 TWh in 2028



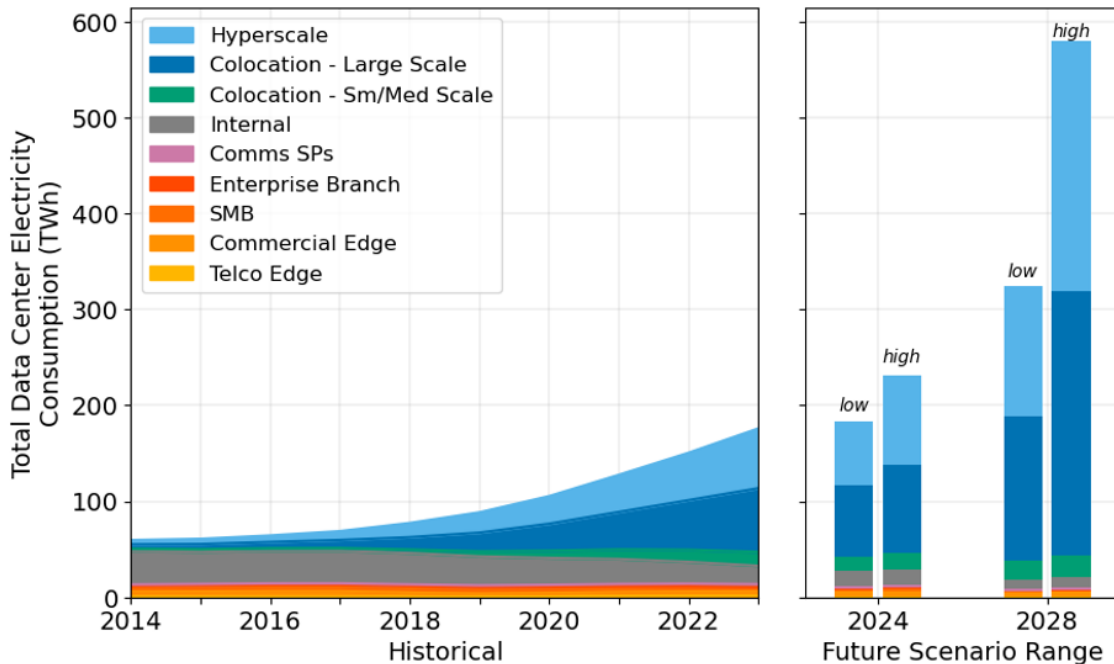
Server Data Center Estimates

Electricity use by server type based on report assumptions



Total Data Center Estimates

Electricity use by data center type based on report assumptions



Data Gaps

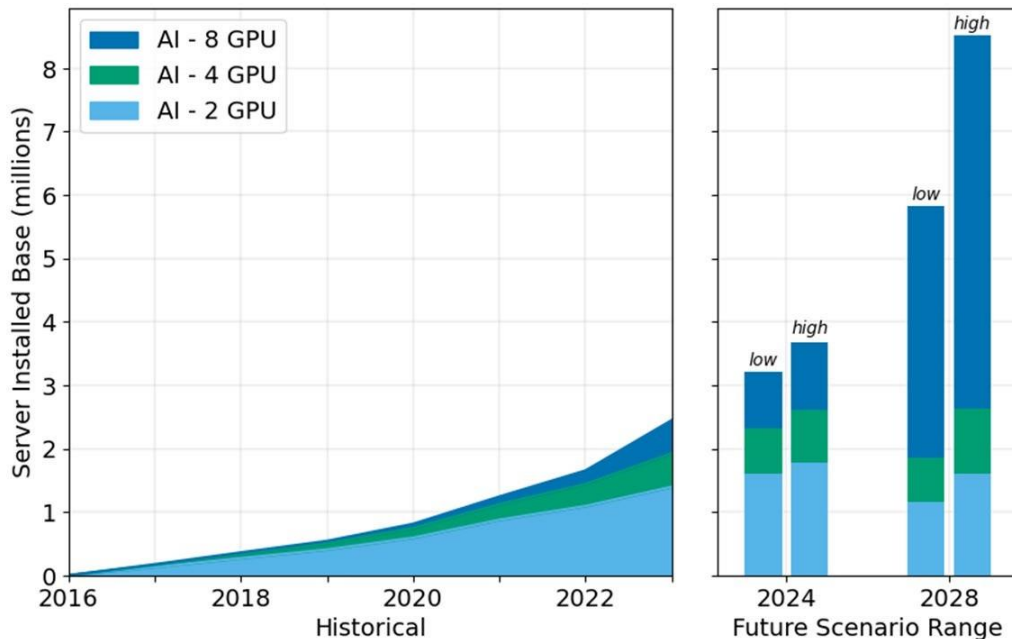


ENERGY TECHNOLOGIES AREA
BERKELEY LAB

AI Server Installed Base

Key Data Gaps

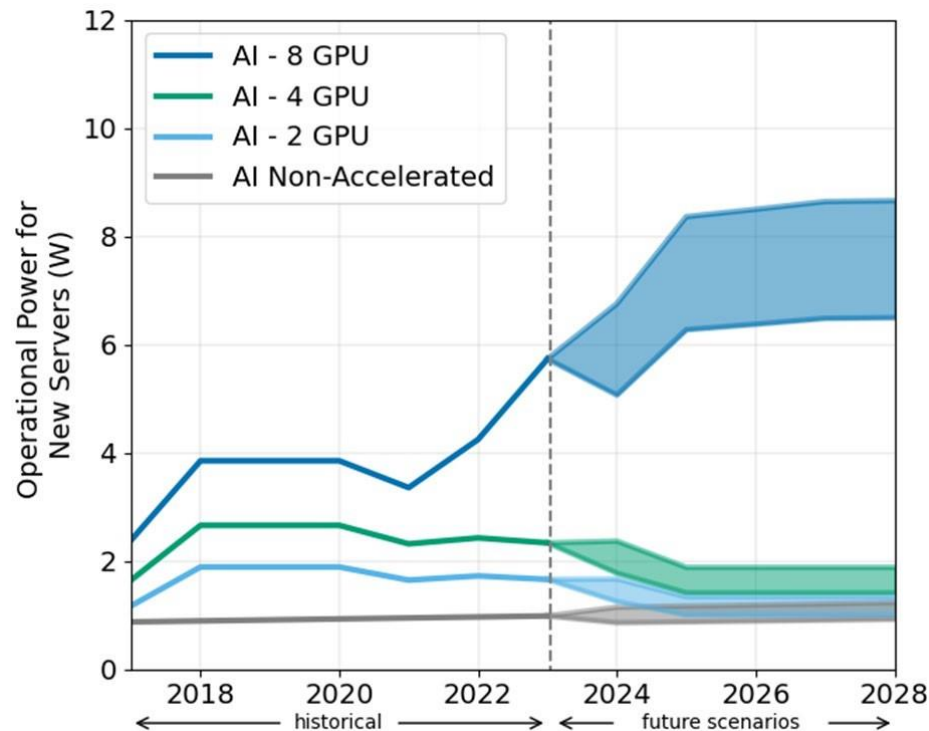
- **Future GPU shipments**
- GPU-server operational practices
- Limited data center location/size and electricity provider



AI Server Operational Power

Key Data Gaps

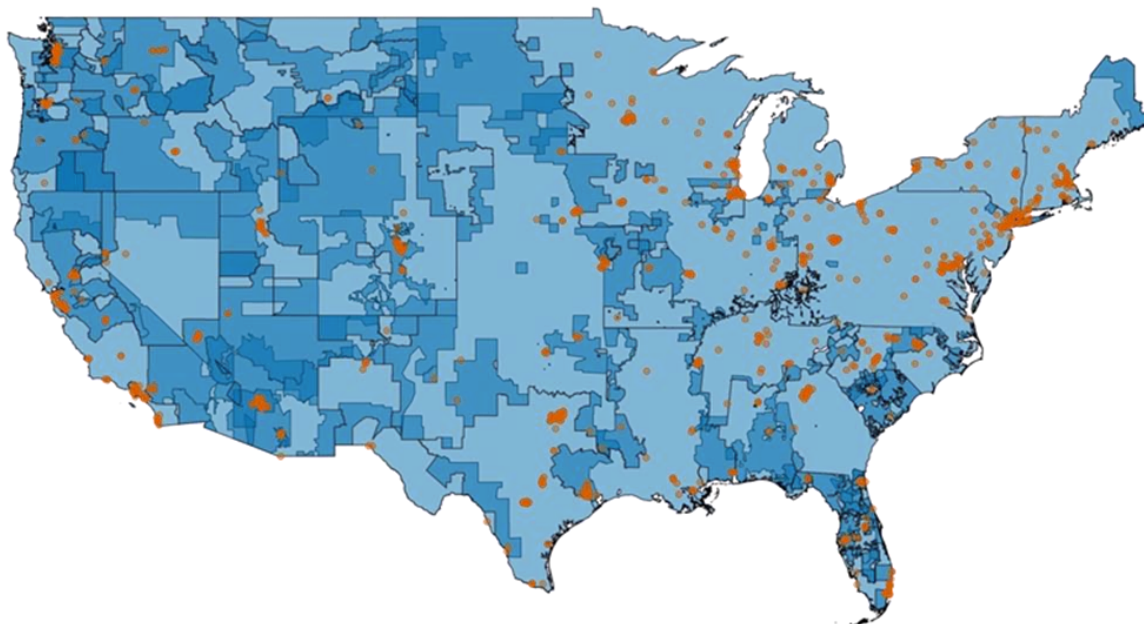
- Future GPU shipments
- **GPU-server operational practices**
- Limited data center location/size and electricity provider



Data Center Location and Electricity Provider

Key Data Gaps

- Future GPU shipments
- GPU-server operational practices
- **Limited data center location/size and electricity provider**



Limited locations of identified data centers (orange dots) in each U.S. balancing authority

Looking forward



ENERGY TECHNOLOGIES AREA
BERKELEY LAB

Looking forward

- Forecasts are very uncertain
 - Short timeline (3-4 years) can be based on orders placed for chips, facilities under construction, and so on.
 - Longer timelines are anyone's guess, especially when considering local demands
 - Data gaps reduce ability to best support power needs, especially as AI hardware development expands
- Computational efficiency has a high rebound effort
 - Improvement in GPU efficiency (currently) translates directly to larger models
 - Previous turning points identified waste and a rethinking of technology and operations
- Context is important
 - Data center demand growth is real, and it is happening fast. But other demand growth in **building electrification, EVs, and industry** is expected. These should all be considered holistically.

