Emerald AI | NASEO-NARUC Webinar

August 2025

Agenda

1. About Emerald AI

2. Emerald AI's Technology

3. Deep Dive on Emerald AI's First Commercial Demo

4. Where do we go from here?

emerald ai

1. About Emerald AI

Emerald AI seeks to unlock multitrillion-dollar investments in AI data centers *today*, overcoming power grid constraints that threaten the AI revolution

- **Emerald AI transforms energy-intensive AI data centers into AI-powered grid allies through a software-only solution** by orchestrating compute power use to access electric grids and bolster grid stability, without compromising compute service.

- **The opportunity is massive:** Roughly 100 GW of data centers could be connected to existing grids *today* if data centers had just modest flexibility. Emerald AI aims to enable data centers & cloud providers to speed time-to-power and stay compliant.

- **Emerald AI successfully proved its technology suite in a first-of-a-kind commercial demonstration in Phoenix in May 2025 at a hyperscaler data center—power utilities and our cloud and AI partners have shared positive feedback**

- **Emerald AI is the only solution to get data centers access to power now** and will complement other solutions in future.

emerald ai

Our team combines expertise in AI, compute and energy

Leadership

Dr. Varun Sivaram, Founder & CEO

Education: Ph.D. Condensed Matter Physics, Oxford University (Rhodes Scholar); B.S., B.A., Stanford University

Boards: Atlantic Council, Stanford University Doerr School of Sustainability

Books: *Taming the Sun* (2018), *Energizing America* (2020), *Digital Decarbonization* (2018)

Awards: TIME 100 Next, World Economic Forum Young Global Leader, MIT Top 35 Innovators

Ørsted
Chief Strategy and Innovation Officer

White House
Sr. Advisor to Sec. John Kerry; Managing Director for Clean Energy, Innovation and Competitiveness

ReNew Power (NDAQ: RNW)
Chief Technology Officer

Council on Foreign Relations
Director, Program on Energy & Climate

McKinsey & Company
Consultant

Prof. Ayse Coskun, Chief Scientist

- *Professor, Boston University*
- *Director, Center for Information Systems and Systems Engineering*

Shayan Sengupta, Head of Engineering

- *Amazon Web Services EC2 Nitro Storage Leader*
- *Led 50+ engineers, managing 100,000 specialized GPU servers for AWS clients*

Aroon Vijaykar, SVP Commercial

- *Sunrun Head of Corp Dev; GM and leadership team of three businesses: AEE Solar; Snapnrack; & VPPs*
- *Investor at Partners Group; Consultant at Monitor Deloitte*

Core Team

Dr. Daniel Wilson

Dr. Philip Colangelo

Chris Williams

Dr. Ciaran Roberts

Jack Megrue

Ethan Tiao

emerald ai

In July, we announced our seed round & successful demonstration

Institutional Venture Investors

RADICAL VENTURES

NVIDIA

crv

NEOTRIBE

AMPLO

Select Individual Investors

Secretary John Kerry *68th U.S. Secretary of State*	**Tom Steyer** *Co-Chair, Galvanize*	**Mark Gallogly** *Co-Founder, Three Cairns Group*
Fei-Fei Li *Professor, Stanford University*	**Jeff Dean** *Chief Scientist, Google*	**John Doerr** *Chairman, Kleiner Perkins*
Malcolm Turnbull *29th Prime Minister of Australia*	**Kate Brandt** *CSO, Google*	**Rich Lesser** *Chairman, BCG*
Chase Lochmiller *CEO, Crusoe*	**Lukas Biewald** *CEO, Weights and Biases*	**Jonathan Frankle** *Chief AI Scientist, Databricks*
David Thorne *52nd U.S. Ambassador to Italy*	**Gerald Butts** *Vice-Chair, Eurasia Group*	**Markus Specks** *Managing Partner, Aventurine*

Select Advisors

Gina Raimondo *40th U.S. Secretary of Commerce*	**David Rousseau** *President, Salt River Project*
Tyler Norris *Duke University*	**Arushi Sharma Frank** *Fmr Energy Policy Lead, Tesla*

AXIOS

Nvidia stakes new startup that flips script on data center power

AI giant Nvidia and boldface names in tech and finance are backing a new startup that aims to transform data centers into flexible grid assets instead of liabilities…. There's growing interest in data centers' flexibility to lower power use for limited stretches.

The Economist

How Managing Energy Demand Got Glamorous

Emerald AI, an American startup, recently showed it can cut power use at AI data centres with software to manipulate computational loads without meaningful loss of performance. The economic logic is compelling.

POLITICO

Nvidia-backed startup wants data centers to be grid assets

A new software startup backed by chipmaker Nvidia aims to solve a crucial problem as data centers proliferate: how to stop the power-hungry operations from crashing the grid. Emerald AI…orchestrates and coordinates artificial intelligence workloads in real time to avoid straining the grid in times of peak demand.

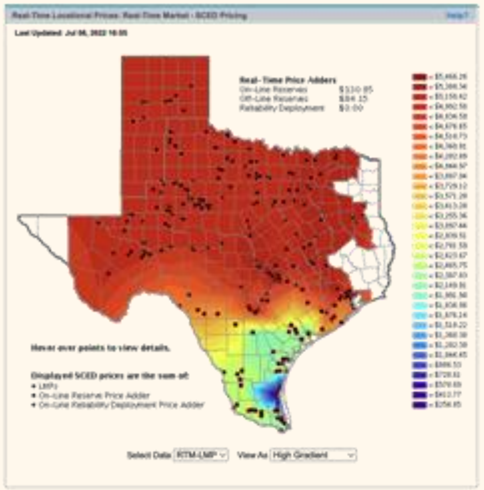
emerald ai

2. Emerald AI's Technology

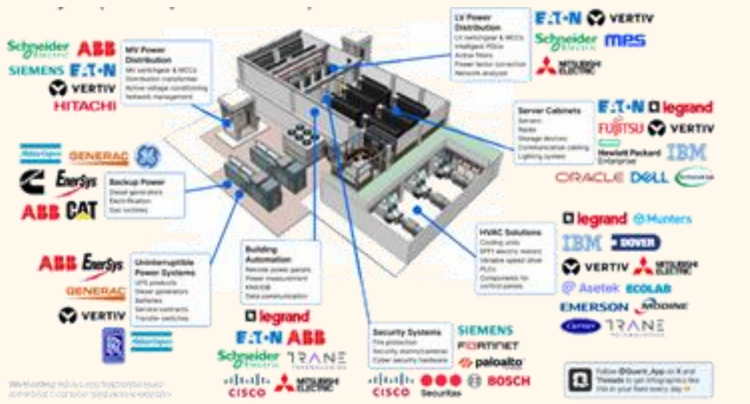
Emerald AI
Turning energy-intensive data centers into AI-powered grid allies

Emerald Conductor

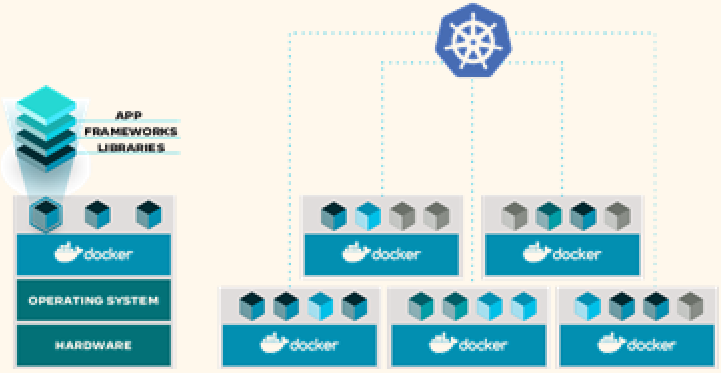
Power Markets & Grid Signals



Data Center Physical Infrastructure Controls



Computing Workload Orchestration



Source: ERCOT, Quartr, Palo Alto Networks

emerald ai

Building Intelligence at the Grid-Data Center Interface

Ultimately, Emerald AI's technology stack will address transient power fluctuations from AI workloads to hours-long grid demand response performance—and everything in between

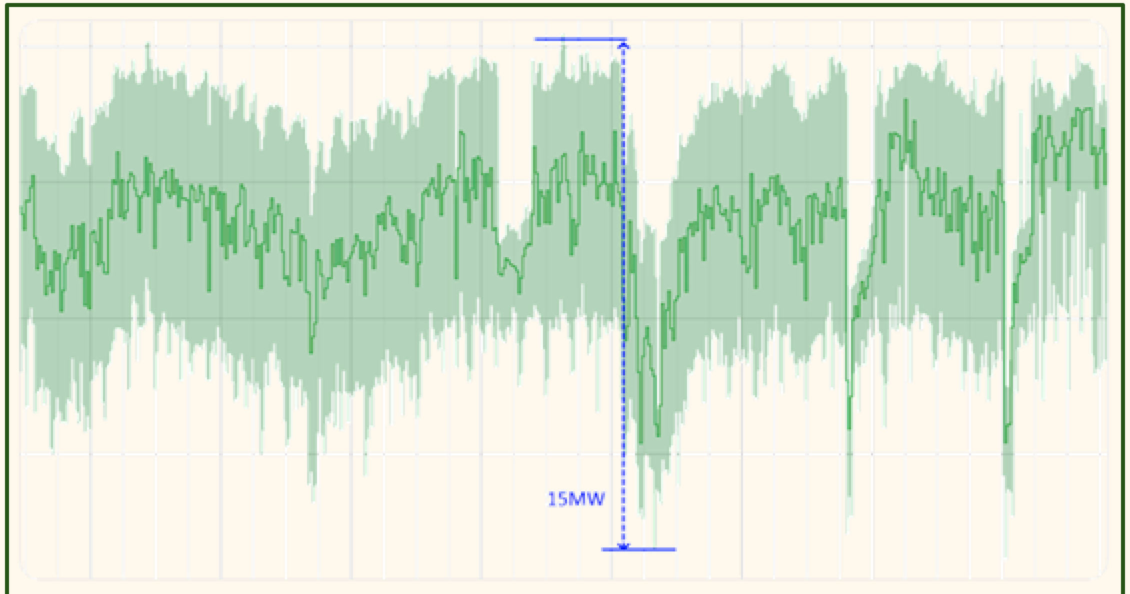
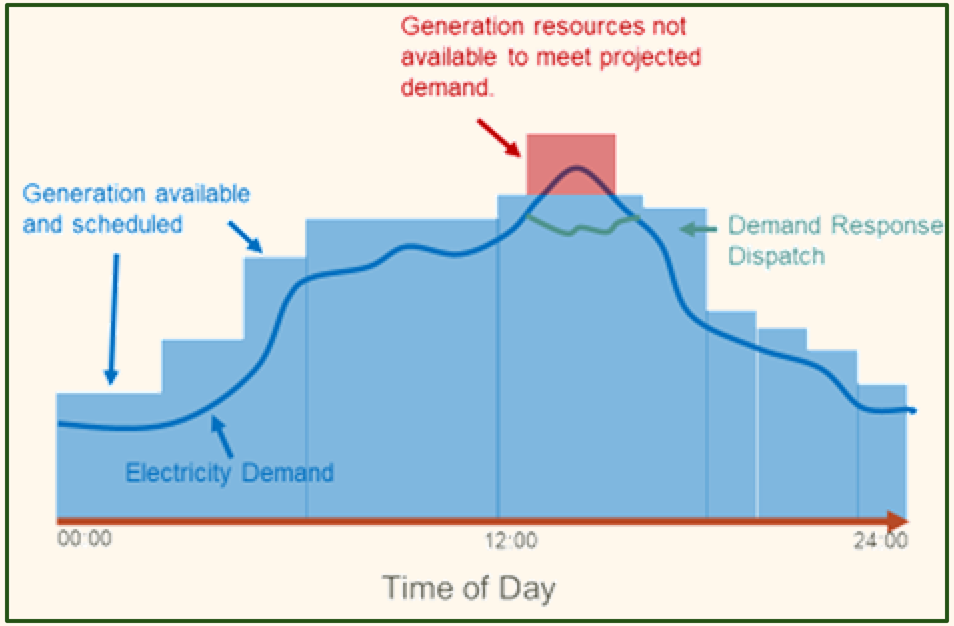


Fig. 1. Large power fluctuations observed on cluster level with large-scale synchronized ML workloads

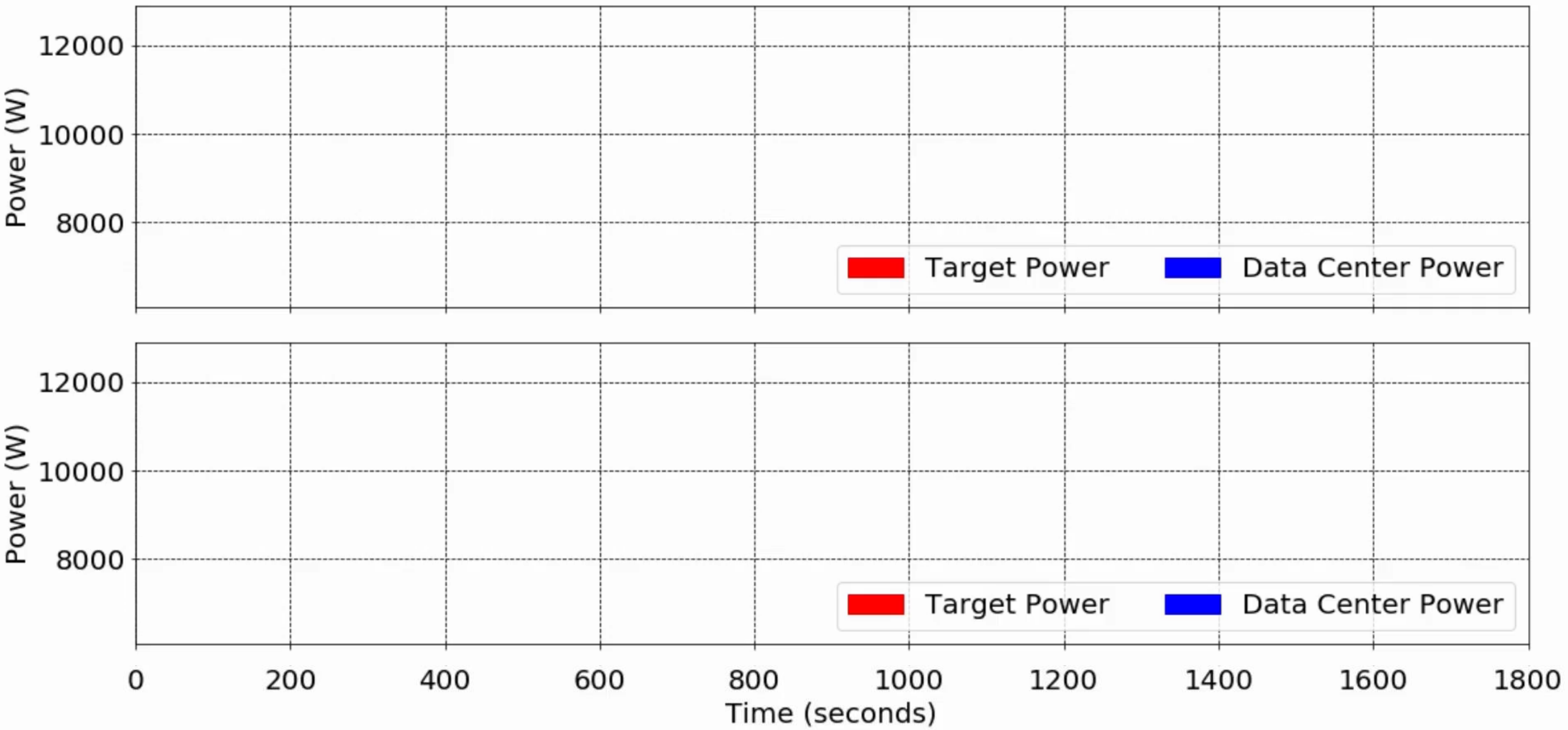


Milliseconds — *Hours*

Source: Google, CPower

emerald ai

The Emerald Conductor software platform orchestrated a real-world compute cluster to follow a 4-second PJM Regulation Reserve signal



3. Deep Dive: Emerald AI's First Commercial Demo

Overview of May 2025 Phoenix, AZ Technology Demonstration

The Phoenix demonstration convened Oracle (hyperscaler), Databricks (AI customer), NVIDIA, EPRI, and local power utilities, to orchestrate a GPU cluster to meet grid AND compute user needs

Hyperscaler Phoenix Data Center

AI Customer Cluster

databricks

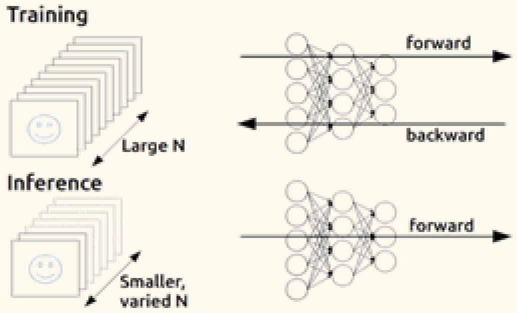
NVIDIA

ORACLE

256
A-100 GPUs

DCFlex

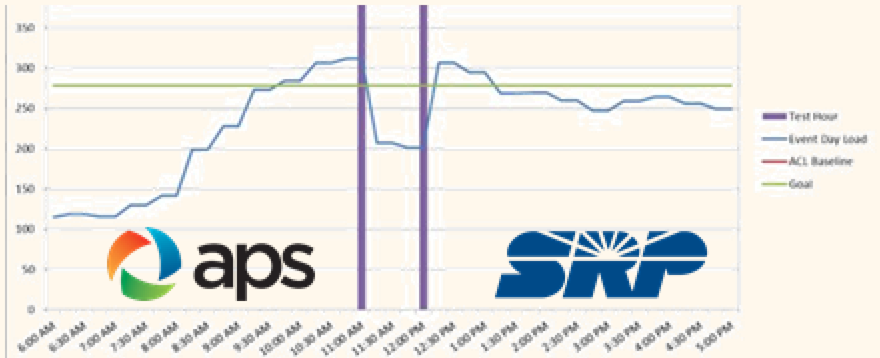
Orchestration of Representative AI Ensembles



Ensure acceptable Quality of Service (QoS)

- Runtime
- Throughput

Phoenix Utility Demand Response Demonstration



emerald ai

How it works: Flow from Event Trigger to Software Execution

Demand Response Event Trigger

Automated Response

- In response to a utility demand response signal, the Emerald AI software stack executes a response—**no manual intervention necessary.**
- In future demonstrations, we will build a full telemetry link to the utility to enable demand response on timescales from seconds to minutes.

Emerald Simulator *"ES"*

Fully Simulated Future Courses of Action

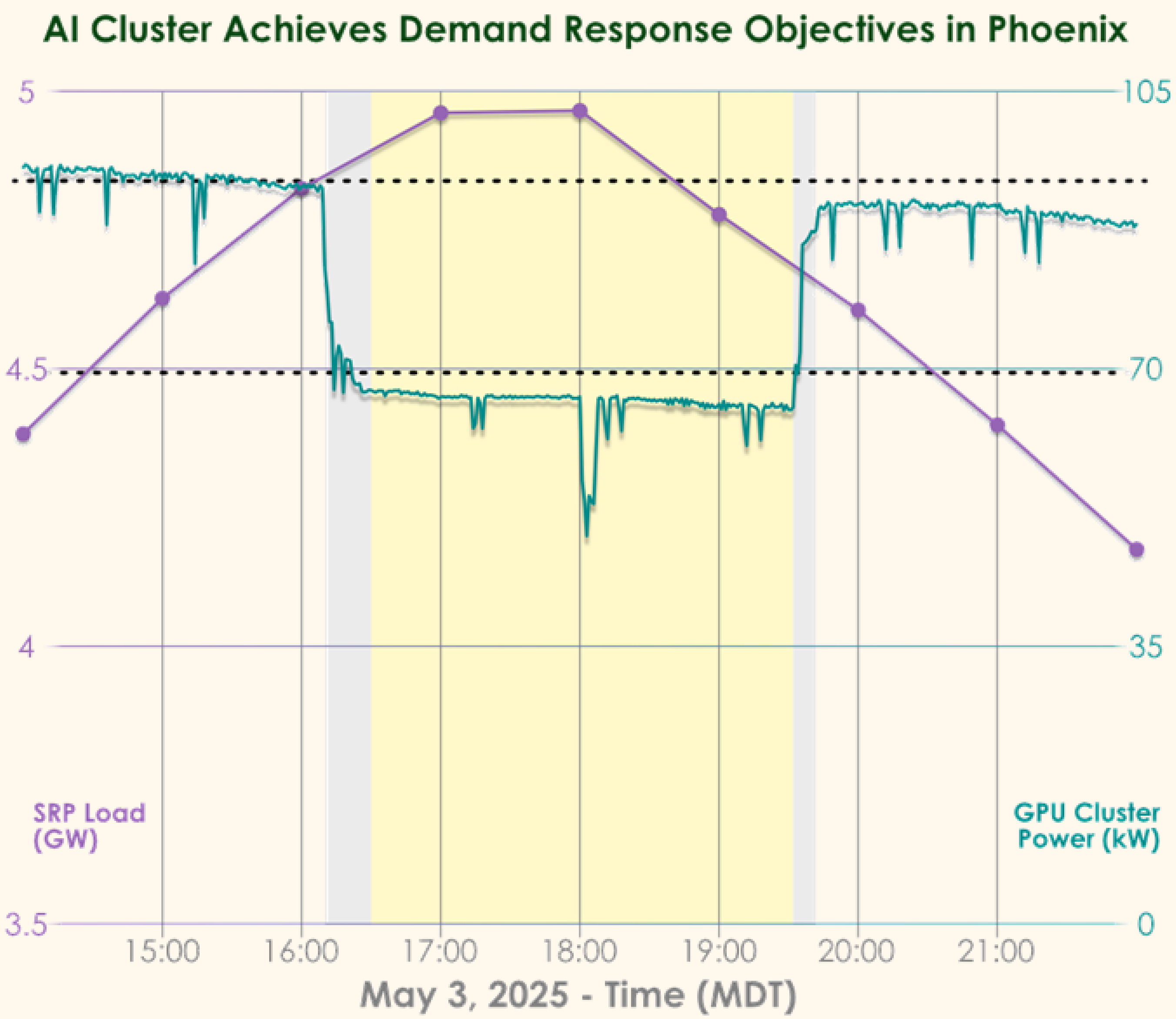
- Upon receipt of the demand response signal, **ES simulates numerous orchestration options for AI workloads**
- Understanding the performance constraints of the grid and of the AI jobs, the simulations optimize for the quality of service for the AI customer

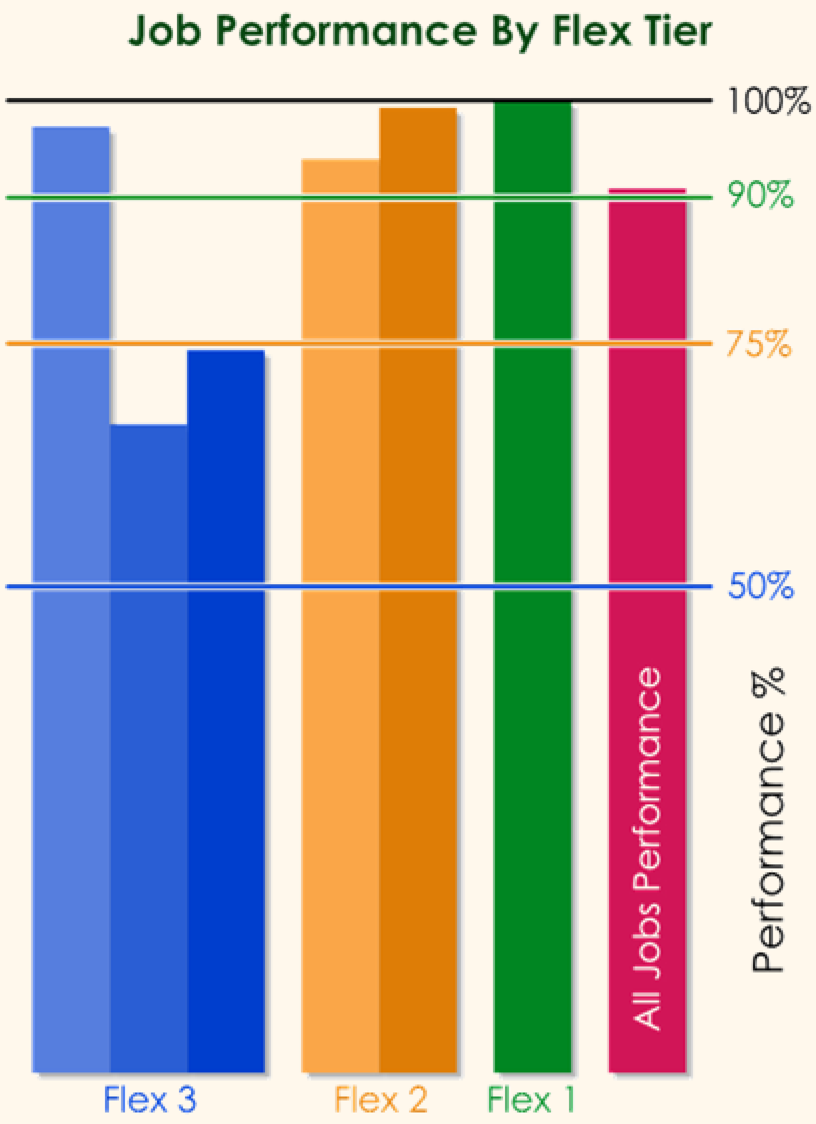
Emerald Conductor *"EC"*

Verifiable Achievement of Grid & AI Objectives

- EC leverages the ES results to orchestrate the data center workloads—and **the results match the simulation with high fidelity.**
- Conducts the **optimal outcome for both the power grid objectives and the compute customer's service level agreement**

emerald ai

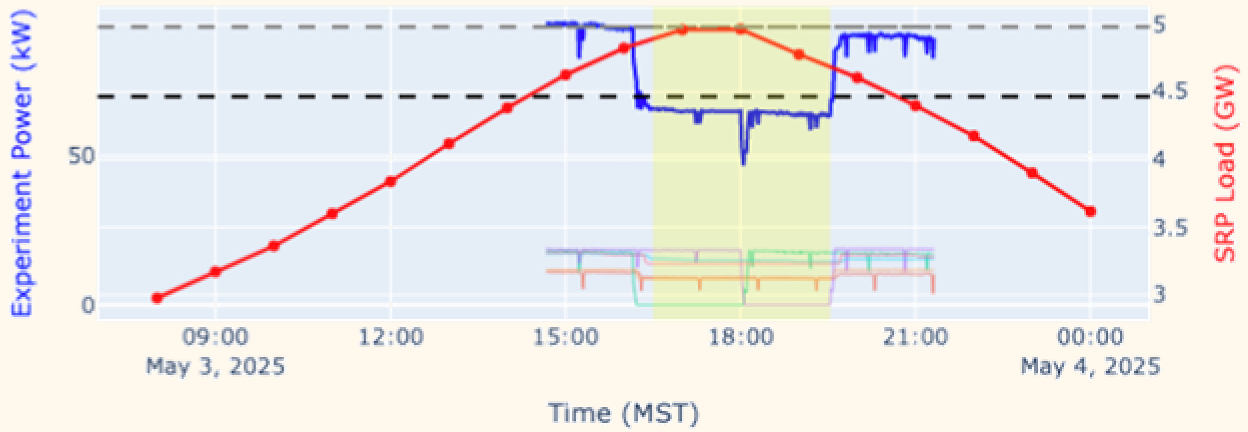

AI Cluster Achieves Demand Response Objectives in Phoenix
5
4.5
4
3.5
105
70
35
0
SRP Load (GW)
GPU Cluster Power (kW)
15:00
16:00
17:00
18:00
19:00
20:00
21:00
May 3, 2025 - Time (MDT)

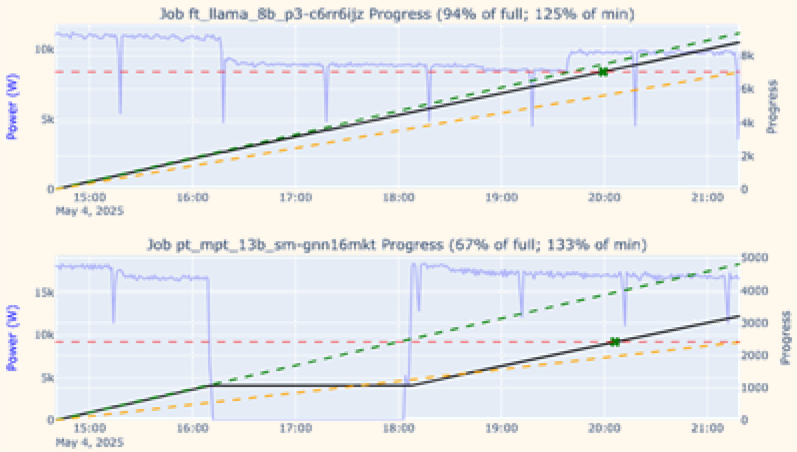

Job Performance By Flex Tier
100%
90%
75%
50%
All Jobs Performance
Performance %
Flex 3
Flex 2
Flex 1

emerald ai

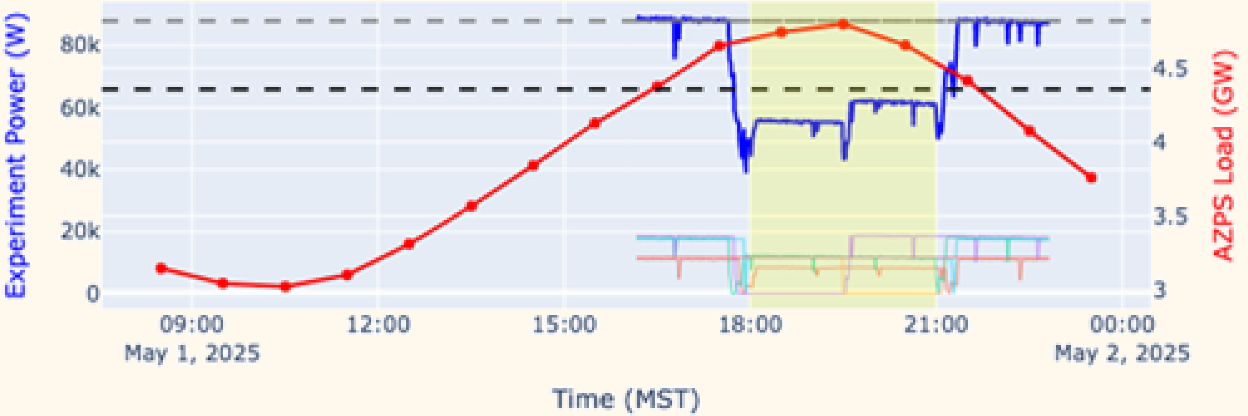
The aggregate load profile delivered is a combination of unique strategies deployed at the workload & GPU level

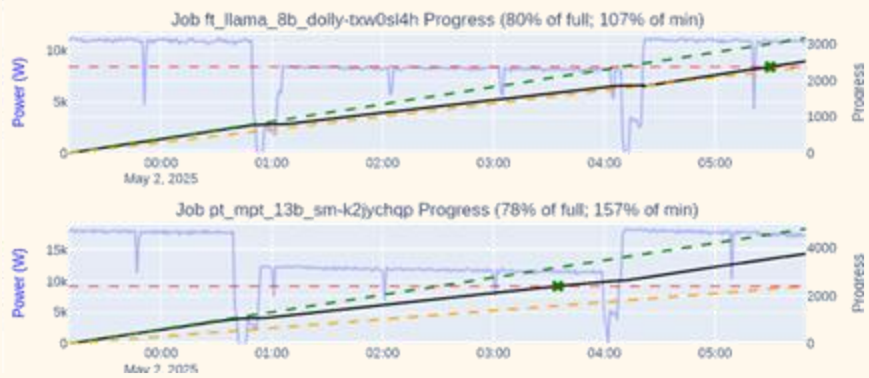
Salt River Project ("SRP"), May 3th, 2025





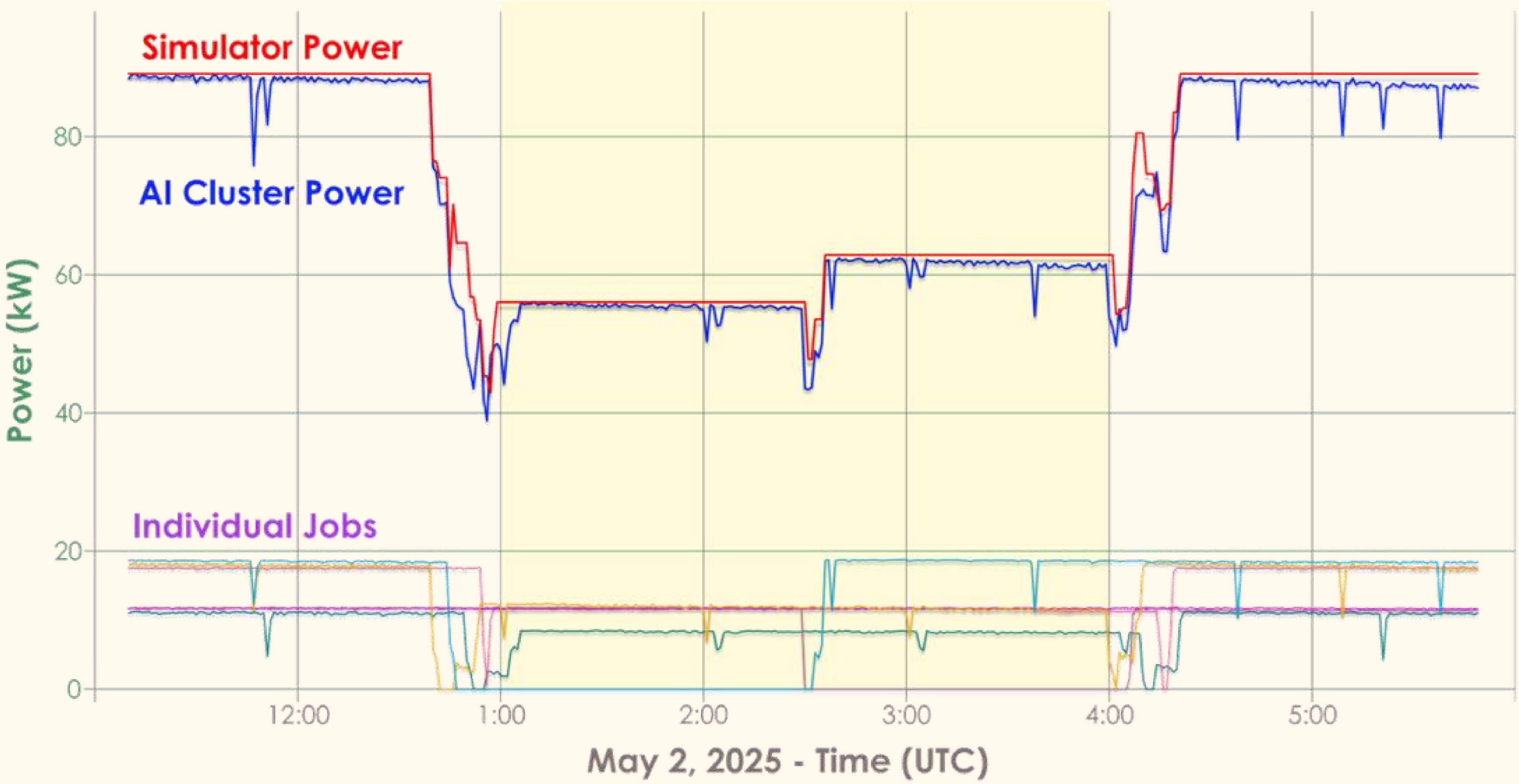
Arizona Public Service ("APS"), May 1st, 2025





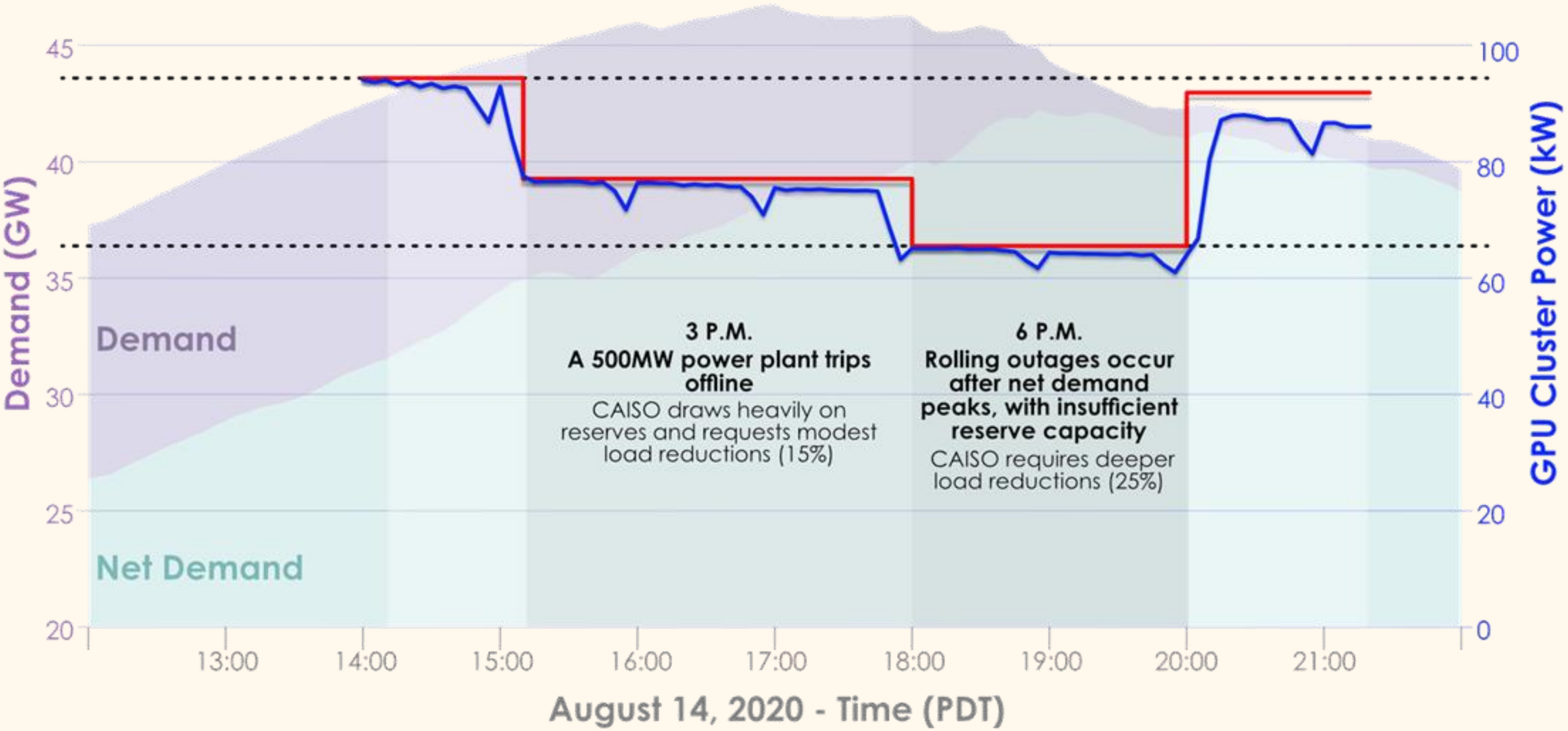
emerald ai

Emerald Simulator Accurately Predicts Real-World Power Consumption



emerald ai

Emerald AI Responds to Real-World CAISO Reliability Event Through Dynamic Control of AI Compute Load



emerald ai

4. Where do we go from here?

emerald ai

Here's how energy and AI leaders can collaborate to scale flexibility

Energy Sector

Develop accelerated interconnection processes for flexible AI factories

Design new approaches to study large flexible loads

Participate in demonstrations at larger scales alongside AI sector participants

AI Sector

Develop 'Power-Flex' SLAs

Develop reference designs and certifications for flexible AI factories

Participate in demonstrations at larger scales alongside power sector participants

Thank you

Contact Us:

Varun Sivaram, *Founder & CEO:* varun@emeraldai.co

Ayse Coskun, *Chief Scientist:* ayse.coskun@emeraldai.co

Aroon Vijaykar, *SVP Strategy & Commercial:* aroon.vijaykar@emeraldai.co

Shayan Sengupta, *Head of Engineering:* shayan.sengupta@emeraldai.co

Integrating Large Flexible Loads in US Power Systems

Tyler H. Norris

NASEO–NARUC Webinar:
Data Center & Large Load Flexibility to Meet Demand Growth

August 7, 2025

Duke NICHOLAS INSTITUTE *for* ENERGY, ENVIRONMENT & SUSTAINABILITY

GRACE Lab A Grid that is Risk-Aware for Clean Electricity | NICHOLAS SCHOOL *of the* ENVIRONMENT

Who we are

Duke NICHOLAS INSTITUTE *for* ENERGY, ENVIRONMENT & SUSTAINABILITY

The Nicholas Institute at Duke University accelerates solutions to critical energy and environmental challenges, advancing a more just, resilient, and sustainable world.

The Nicholas Institute conducts and supports actionable research and undertakes sustained engagement with policymakers, businesses, and communities—in addition to delivering transformative educational experiences to equip future leaders. The Nicholas Institute's work is aligned with the Duke Climate Commitment (climate.duke.edu).

GRACE Lab
A Grid that is Risk-Aware for Clean Electricity

Duke | NICHOLAS SCHOOL *of the* ENVIRONMENT

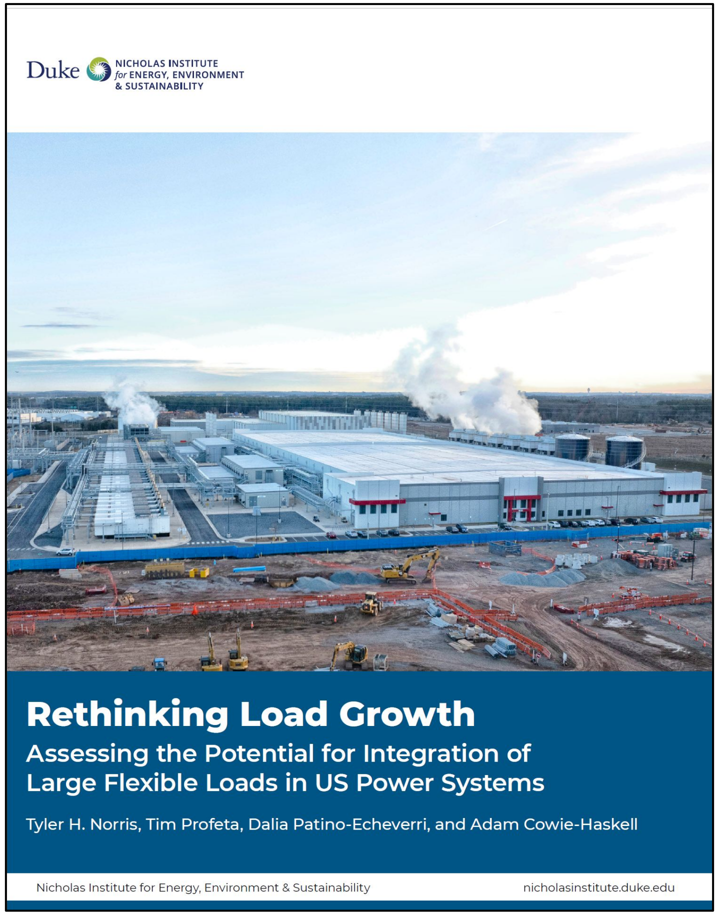
GRACE Lab's research explores, assesses, and proposes technological, policy, and market approaches to contribute to the pursuit of sustainability, affordability, and reliability in the energy sector.

Primary research areas:

- Characterizing sources of uncertainty that increase the financial and reliability risk of power systems, and designing risk management strategies
- Examining the possibilities and advantages of designing flexible policy mechanisms
- Assessing the economic, environmental, and reliability potential of low-emissions energy technologies

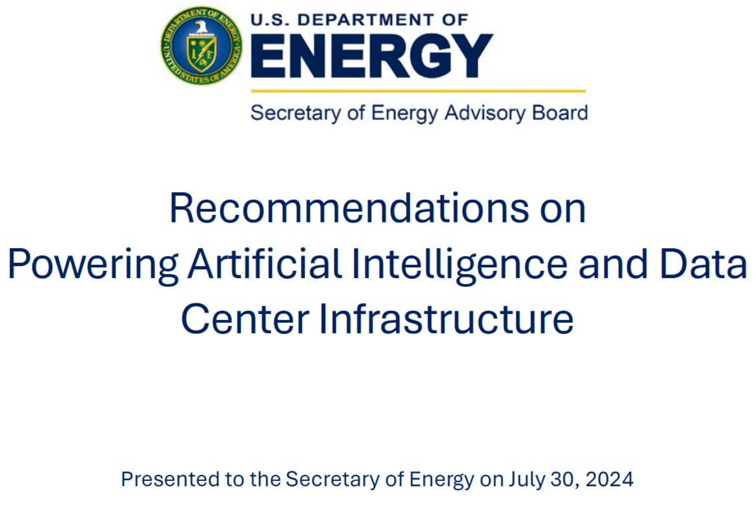
Goals for *Rethinking Load Growth*

- Support regulators and stakeholders in identifying strategies to accommodate load growth without compromising reliability, affordability, or progress on decarbonization
- Provide informational resources and a first-order estimate of the potential for accommodating new loads while mitigating or deferring capacity expansion
- Motivate additional analysis to more precisely quantify headroom in each balancing authority



Load flexibility offers a near-term solution

The US Secretary of Energy Advisory Board (SEAB) and the Electrical Power Research Institute (EPRI) have highlighted a solution: load flexibility.


U.S. DEPARTMENT OF ENERGY
Secretary of Energy Advisory Board

Recommendations on Powering Artificial Intelligence and Data Center Infrastructure

Presented to the Secretary of Energy on July 30, 2024

"For immediate impact, the Secretary should direct relevant offices across DOE to explore opportunities for temporal and spatial ***flexibility in AI training and inference****..."*

SEAB Recommendations (PDF)

"EPRI's DCFlex Initiative will demonstrate how data centers can ***support and stabilize*** *the electric grid while improving interconnection and efficiency."*

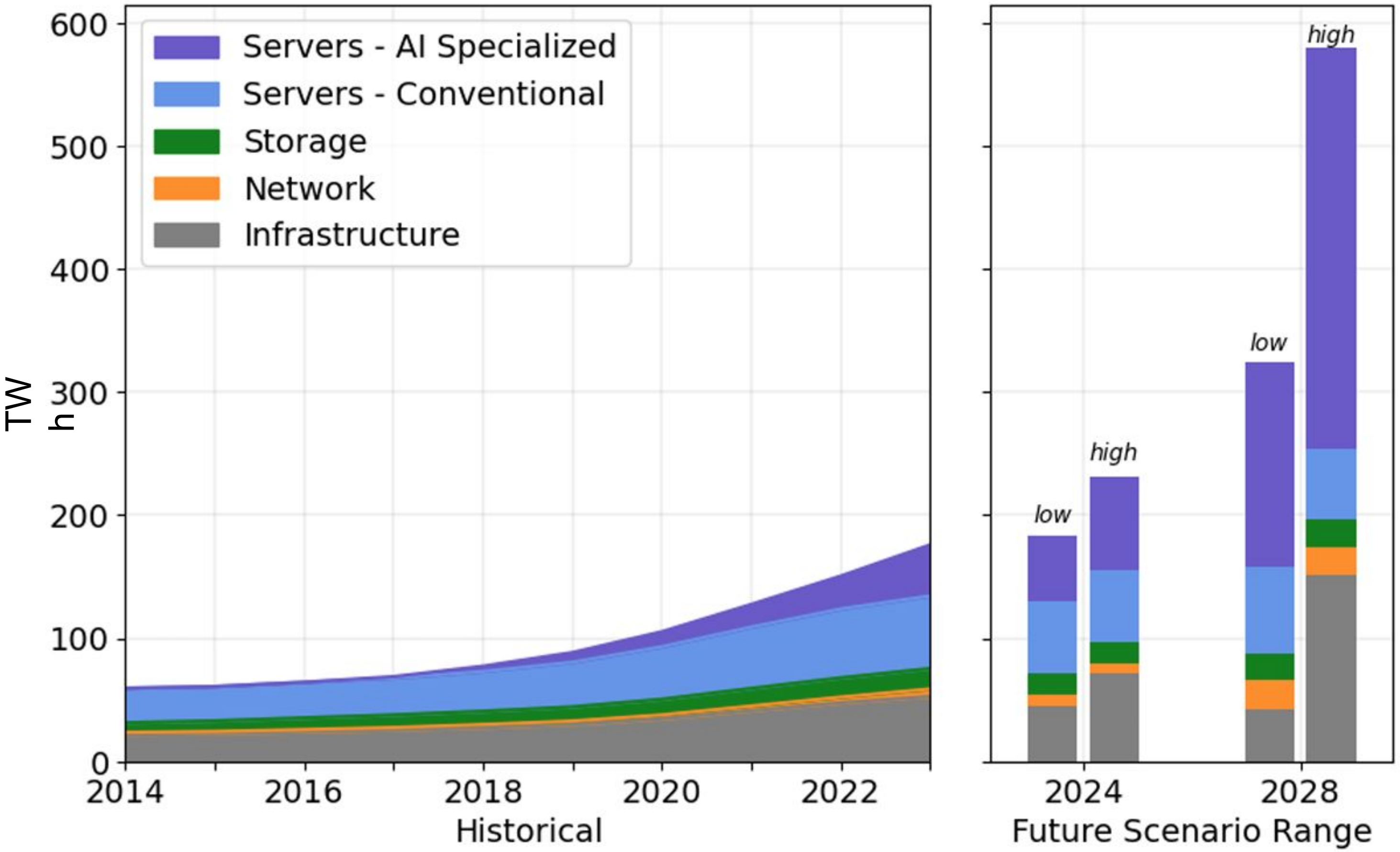
https://msites.epri.com/dcflex

Background

AI data centers lead US electricity load growth

- Data centers could account for up to 44% of US electricity load growth through 2028
- AI workloads are projected to represent 50% to 70% of data center demand by 2030
- Hyperscale and large-scale colocation data centers account for the vast majority of growth

US data center electricity use by equipment type, 2014-2028

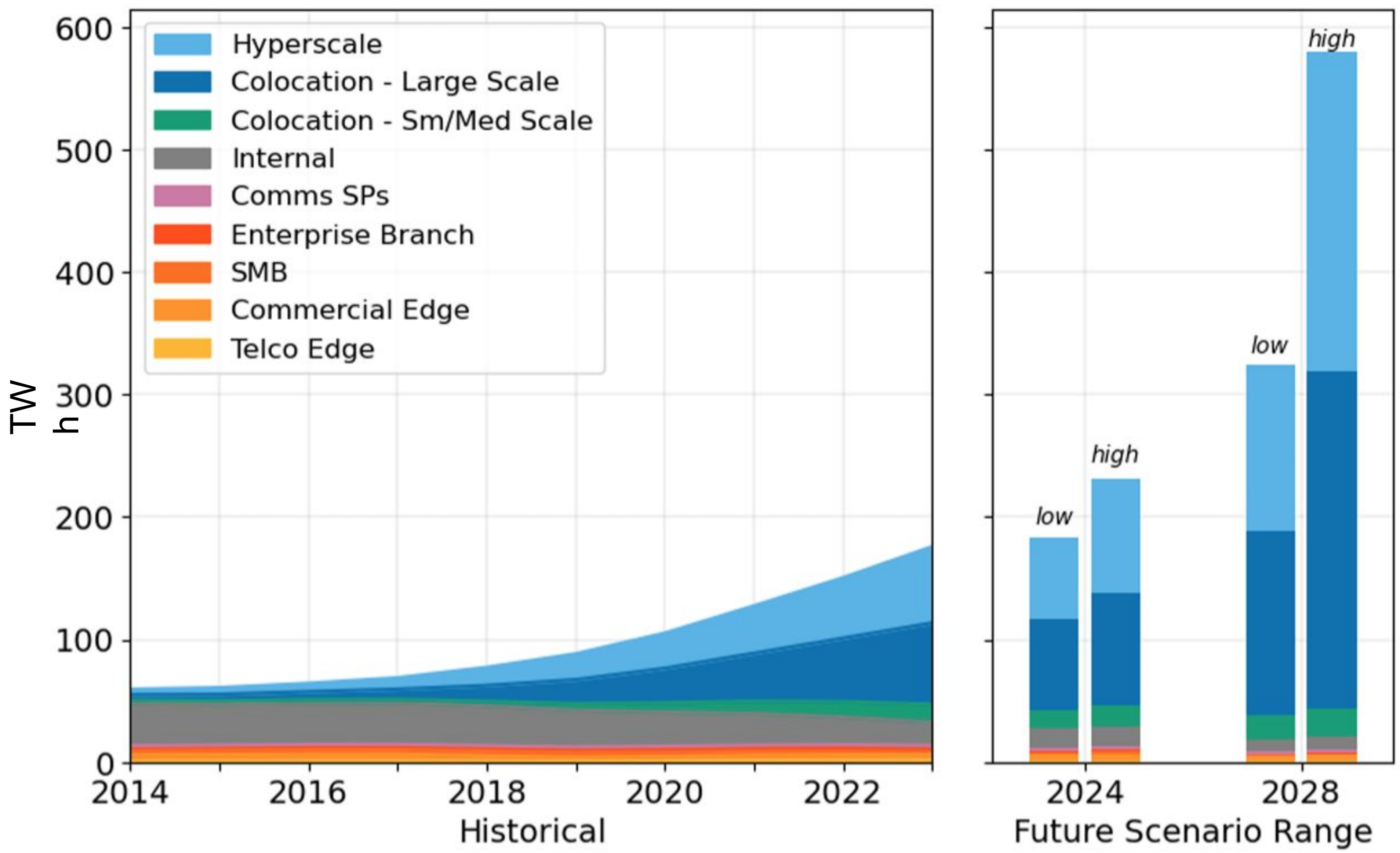


Source: Shehabi, A., et al. 2024 United States Data Center Energy Usage Report. Lawrence Berkeley National Laboratory, Berkeley, California. LBNL-2001637

AI data centers lead US electricity load growth

- Data centers could account for up to 44% of US electricity load growth through 2028
- AI workloads are projected to represent 50% to 70% of data center demand by 2030
- Hyperscale and large-scale colocation data centers account for the vast majority of growth

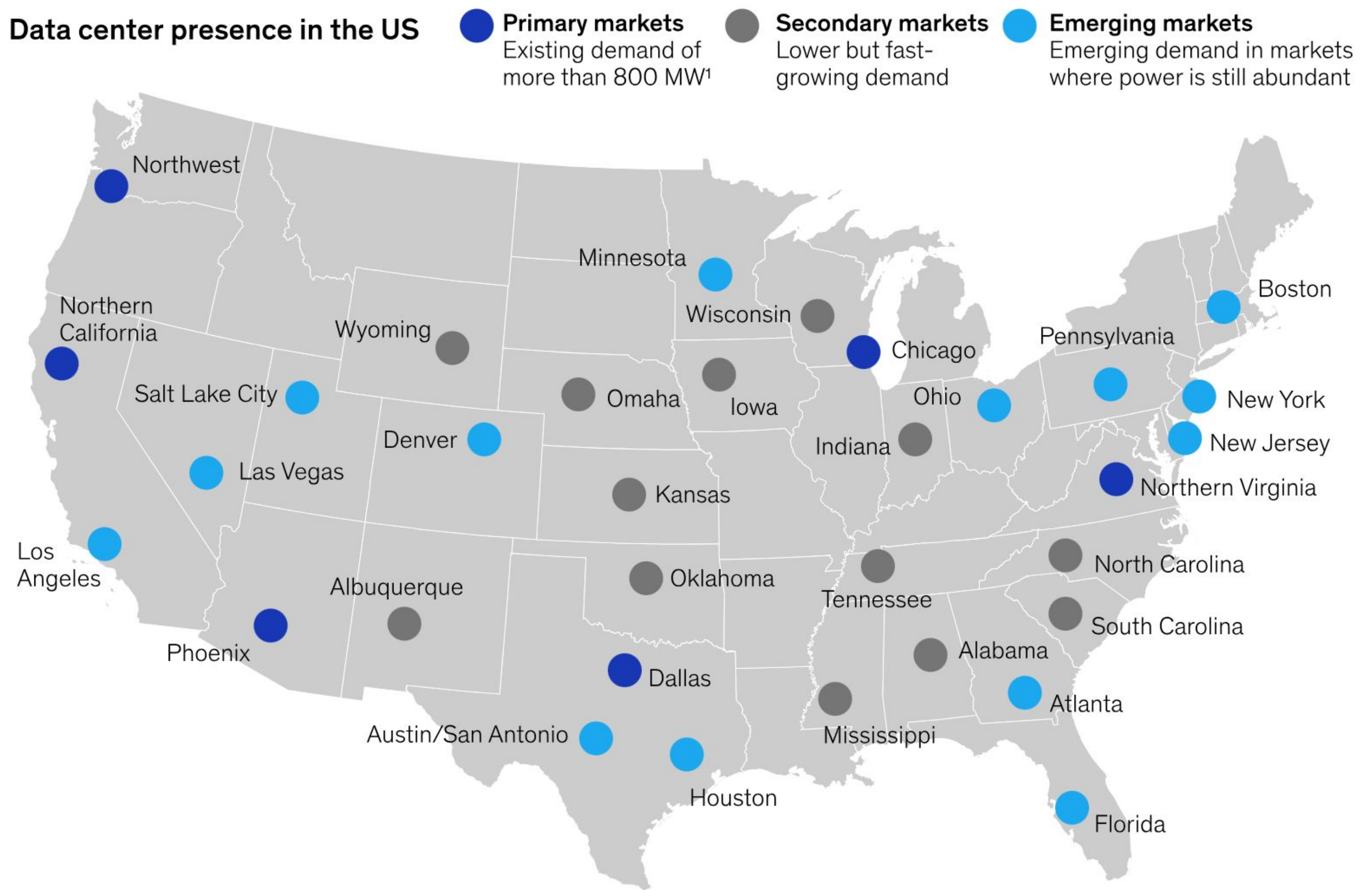
US data center electricity use by space type, 2014-2028



Source: Shehabi, A., et al. 2024 United States Data Center Energy Usage Report. Lawrence Berkeley National Laboratory, Berkeley, California. LBNL-2001637

AI data centers are siting in more remote locations

- Higher shares of AI training workloads enables siting in more remote locations
- More remote siting enables greater ability to co-locate with on-site generation
- Remote siting may also enable greater ability to permit and operate backup generators



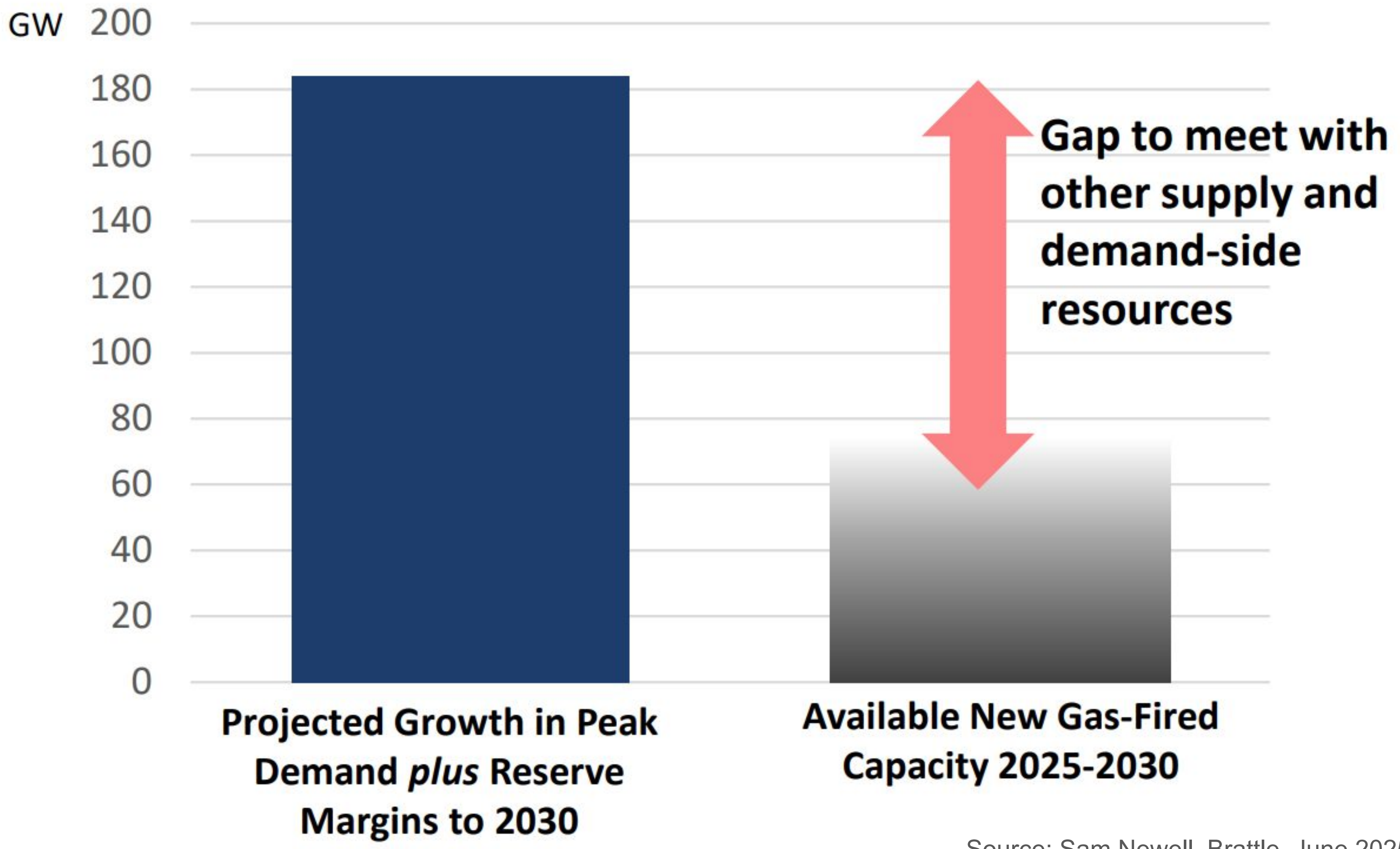
Source: Datacenters.com; S&P Global Market Intelligence
451 Research; McKinsey Data Center Demand model

Load growth is colliding with resource constraints

- Transformer order lead times have risen to 2-5 years—up from less than one year in 2020—while costs have surged by 80% (NIAC 2024)
- Lead times for high-voltage circuit breakers reached 151 weeks in late 2023, marking a 130% year-over-year increase (Wood Mackenzie 2024)
- Interconnection wait times have grown significantly, with some utilities reporting delays up to 7 to 10 years (Li et al. 2024; Saul 2024; WECC 2024)

Credit: IEEE

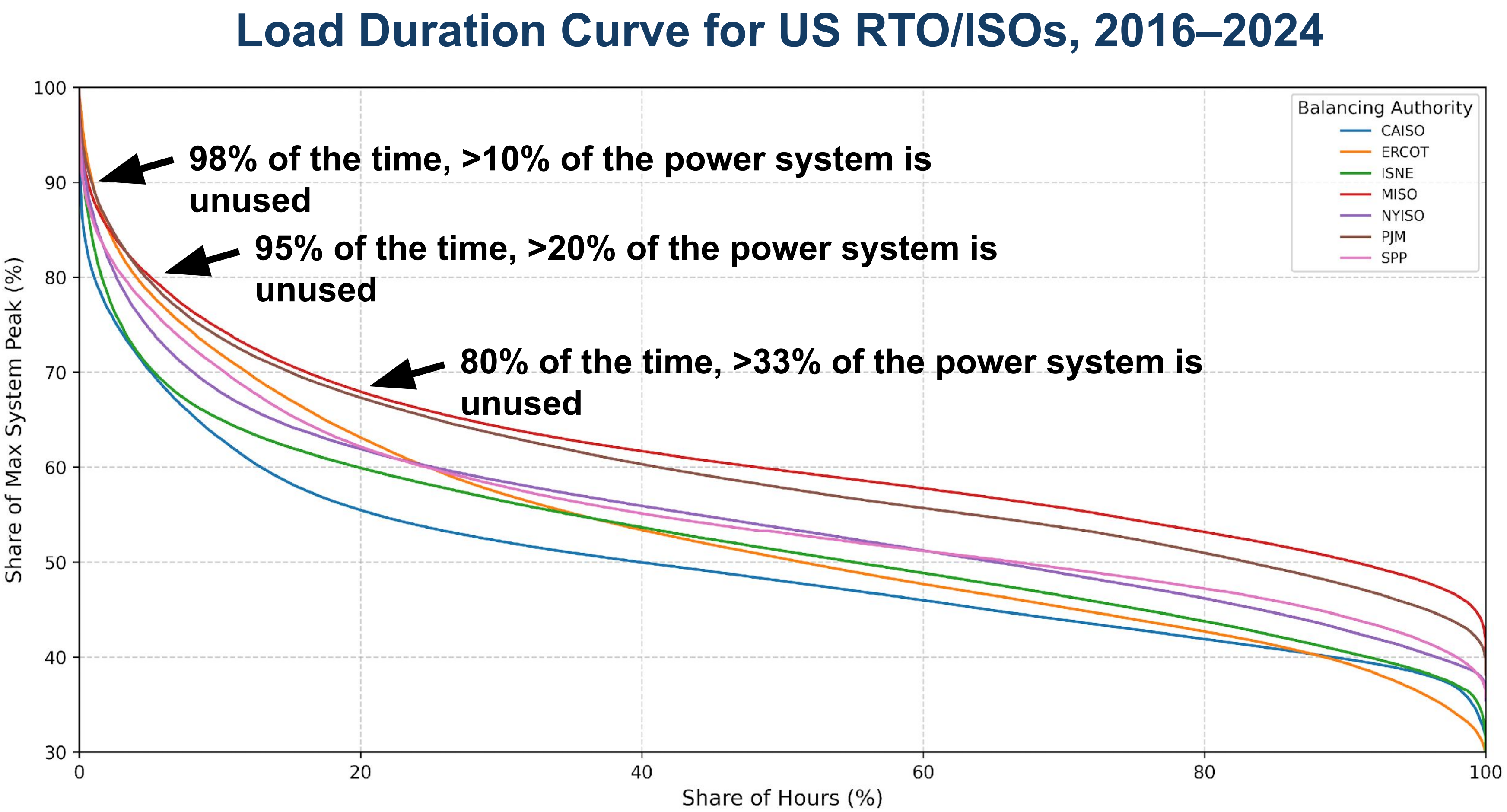
Bridging the supply-side gap



Source: Sam Newell, Brattle, June 2025

Headroom exists outside extreme peak hours

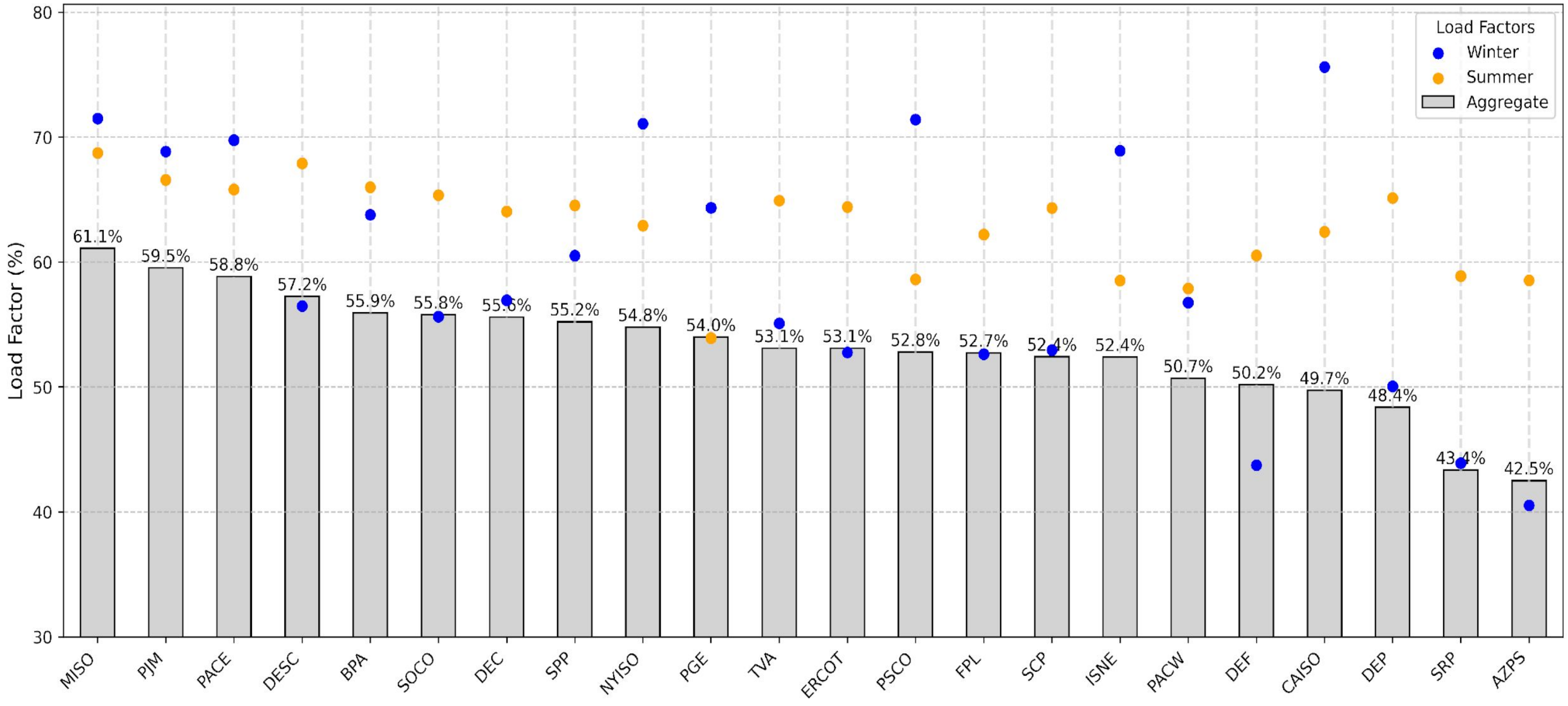
- The load duration curve illustrates system utilization by ranking demand from highest to lowest over a given period
- More than 10% of the system is built to serve 35 hours/year of extreme peak load (avg.)



US power systems operate at 53% avg. load factor

- Load factor is the ratio of average demand to peak demand and is an indication of system utilization
- Aggregate load factors range between 43% to 61%, with an average and median value of 53%
- Winter load factors were notably lower than summer (59% vs. 63% average)

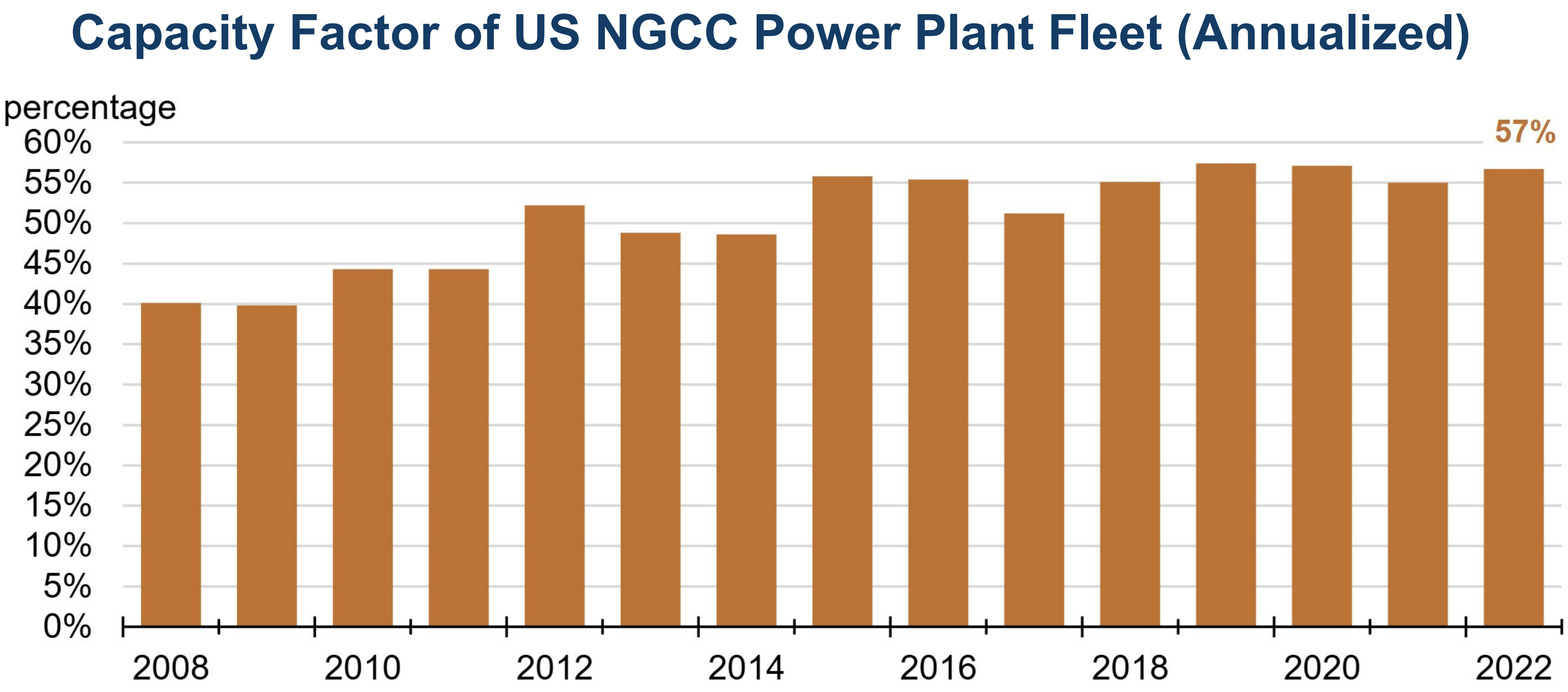
Load Factor by Balancing Authority and Season, 2016–2024



US NGCC fleet operates at <60% capacity factor

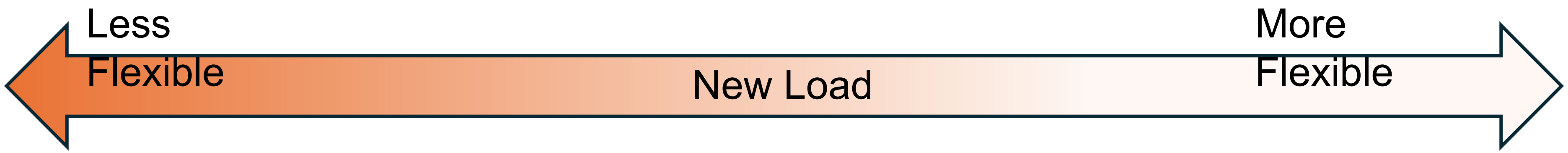
- NGCC plants run well below full potential, with a fleetwide capacity factor of 57% in 2022 – despite being among the most efficient thermal resources

- Even the most efficient NGCC units are underutilized – new plants average 64% utilization despite cutting-edge turbines and low marginal costs



Source: EIA Form EIA-860M, Monthly Electric Power Industry Report

Implications for system planning and interconnection



Inflexible Loads

- More burden on system peaks
- More likely to trigger need for additional generation and transmission capacity
- More likely to require longer interconnection

Flexible Loads

- Less burden on system peaks
- Less likely to trigger need for additional generation and transmission capacity
- Potential for accelerated grid interconnection

Types of load flexibility

On-site power and storage
Utilizing co-located storage, renewables, or other generators

Temporal flexibility
Scheduling computational loads before or after periods of high system stress

Spatial flexibility
Distributing workloads across one or multiple data centers in different geographic locations

Reduced operations
Planning for reduced workload during defined periods

Data center uptime: perception vs. reality

- Duke Energy, 2024 testimony to NC Utilities Commission:
  - *"[Data centers] are operating at a **consistent load factor 365 days a year**, seven days a week, 24 hours."*
- Energy Futures Group, 2025:
  - *"A large load data center, since it is constantly active, will have a high **load factor of 90-100%**."*
- Resources for the Future, 2024:
  - *"It's important to note that these data centers have a very high load factor. Basically, they run very consistently. Let's say it has a **90 percent load factor**.*
- ARES Wealth Management Solutions, 2024:
  - *"Data center capacity calculated assuming a **90% load factor**"*

Data center uptime: perception vs. reality

- **Load factor ≠ Utilization rate**
  - If a customer's peak load is only 85% of its rated capacity, and its load factor is 90%, its capacity utilization rate is 77%
- **Load factor ≠ Server uptime**
  - Load factors for *servers* get conflated with load factors or even utilization rates
- **Server uptime ≠ "Five Nines"**
  - Servers operate far below 99.999% utilization, which is a customer-facing uptime guarantee
- **Variance in non-IT load is not well understood**
  - Constant, high server utilization leaves significant room for seasonal load variability
- **Data is very limited**
  - *"The lack of primary performance and utilization data indicates that much greater transparency is needed around data centers. Very few companies report actual data center electricity use and virtually none report it in context of IT characteristics such as compute capacities, average system configurations, and workload types."*

Terminology

Power Systems Planning

Term	Definition	Formula	Typical Use
Load Factor	Avg vs. peak actual load	Avg load ÷ peak load	Load forecasting
Nameplate capacity	Max rated grid power draw (i.e. interconnect limit)	Based on electrical gear ratings	Interconnection agreement, System planning, Utility equipment sizing
Utilization Rate	Avg load vs. nameplate capacity ("capacity utilization rate")	Avg load ÷ nameplate	Load forecasting, System planning

Data Centers / Compute

Term	Definition	Formula	Typical Use
Load Factor	Equivalent to Power Systems; Potential distinction between full facility vs. server load factor		
Nameplate capacity	Max internal load supported by UPS/gens	UPS + generator stack (max design load)	Facility design/eng and redundancy (2N); often oversized for expansion
Power Usage Effectiveness (PUE)	Total facility energy vs. IT energy	Total Energy / IT Energy	Energy efficiency benchmark (1.1 - 1.5 typical)
Utilization Rate	Server workload vs. max compute capacity ("compute utilization")	Actual server load ÷ max server capacity	IT resource efficiency
Infrastructure Usage Effectiveness (IUE)	Ratio of actual facility power usage to max facility capacity	Avg Total Facility Power ÷ Max Facility Power	Measure of how effectively infra is utilized

Data center uptime: perception vs. reality

Question: When an electric utility cites X gigawatts of AI data center load requests, how much utilization should regulators anticipate?

Source	Figure	Paper
Shehabi et al. 2025 (LBNL)	50%	*2024 United States Data Center Energy Usage Report*
Guidi et al. 2024	75%	*Environmental Burden of United States Data Centers in the Artificial Intelligence Era*
Patel et al. 2024 (SemiAnalysis)	80%	*AI Datacenter Energy Dilemma – Race for AI Datacenter Space*
Knittel et al. 2025 (MIT)	80%	*Flexible Data Centers and the Grid: Lower Costs, Higher Emissions?*
Netherlands Central Bureau of Statistics	"Unknown"	Cited in *How Data Centers Have Come to Matter*

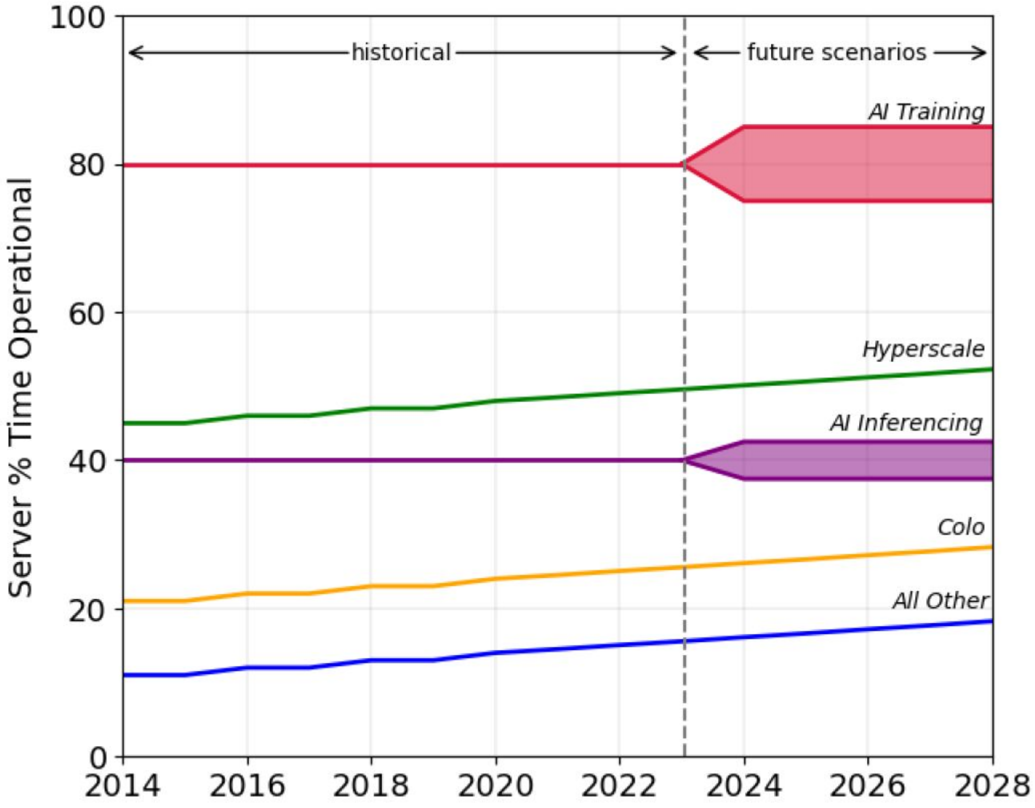


Figure Source: Shehabi, A., et al. 2024 United States Data Center Energy Usage Report. Lawrence Berkeley National Laboratory, Berkeley, California. LBNL-2001637

Potential values of flexibility

1 Faster Time-to-Power

- Dominion Energy: 7+ year wait time
- Centerpoint Energy: 700% increase in data center queue
- APS, ERCOT, others developing priority interconnection for flexible loads

2 Increased Interconnection Capacity

- Utilities and txm service providers limit load capacity based on worst-case load study results
- Flexible loads can be interconnected at higher capacities

3 Mandate Compliance

- Grid reliability and price affordability are suffering around the country.
- Demand response mandates are coming: e.g., legislation already proposed in TX, VA.

4 Flexibility Revenues

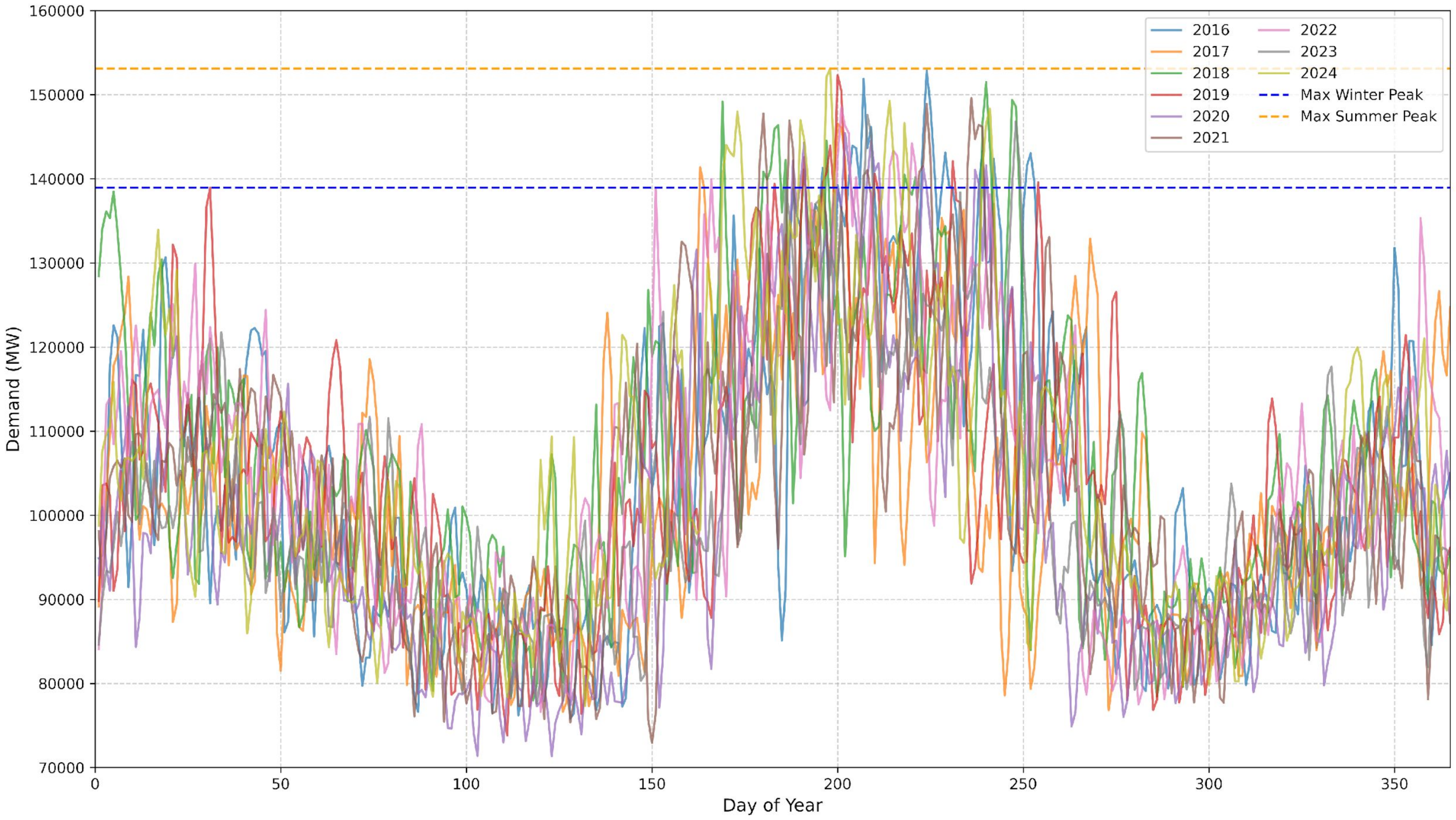
- As peak demand soars, regional power systems are looking for shock absorbers.
- Rising flexibility revenues in 2025 could become material.

Analysis

Method: Identify Seasonal Peak Thresholds

- Balancing authorities develop resource expansion plans to support different peak loads in winter and summer
- To account for variation, we identified the max winter and summer peak observed for each balancing authority
- These thresholds serve as the upper limits for system demand during their respective seasons

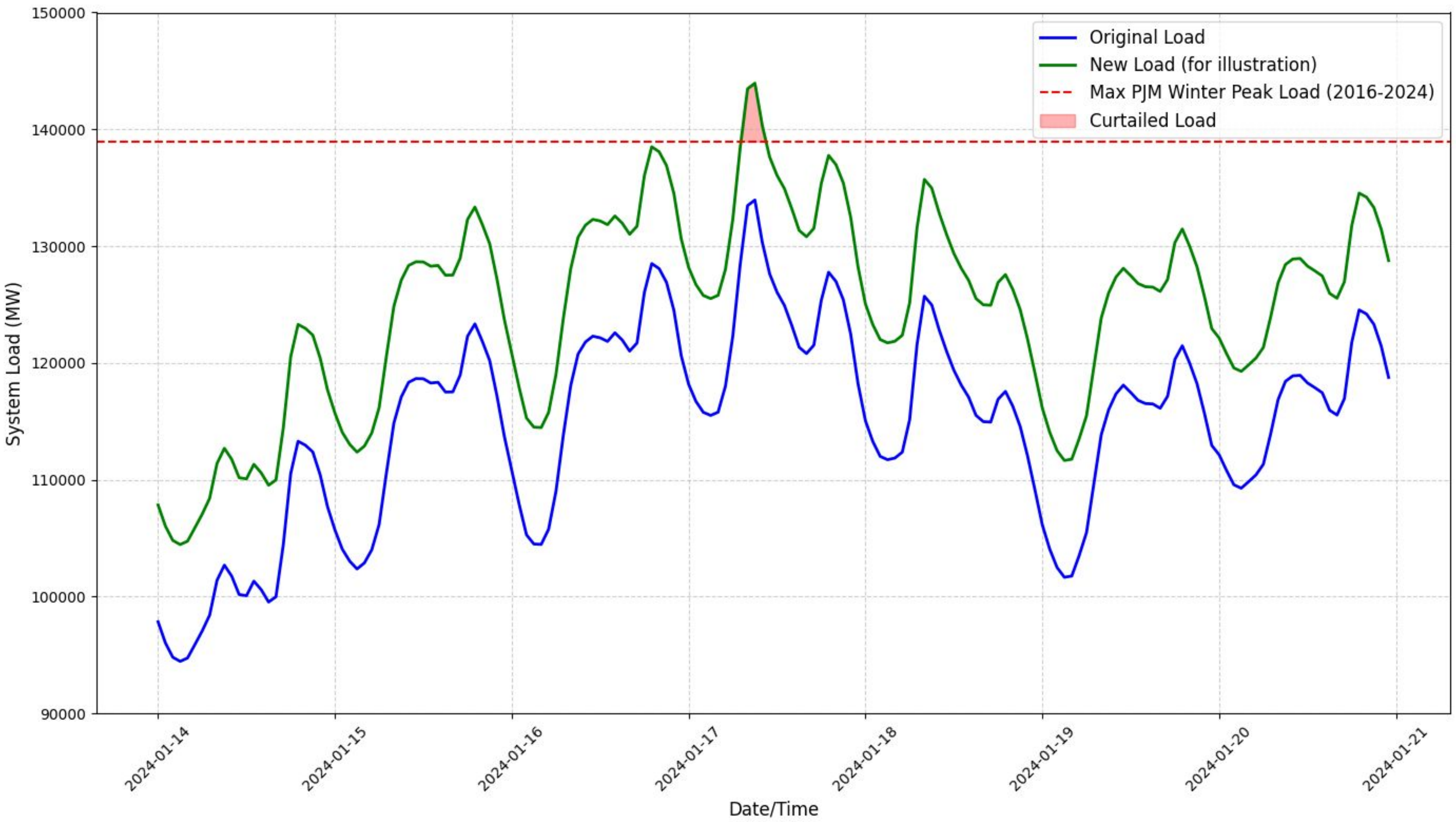
Daily and Seasonal Maximum Demand in PJM (2016-2024), MW



Method: Calculate Load Additions for Curtailment Limits

- New, constant load was added in all hours
- Curtailment was calculated as the difference between the new load and the seasonal peak threshold in each hour, summed across all hours in a year
- The curtailment rate for each load increment was defined as the total annual curtailed MWh divided by the new load's max potential annual consumption

Illustrative Load Curtailment in PJM (Jan. 14–21, 2024)



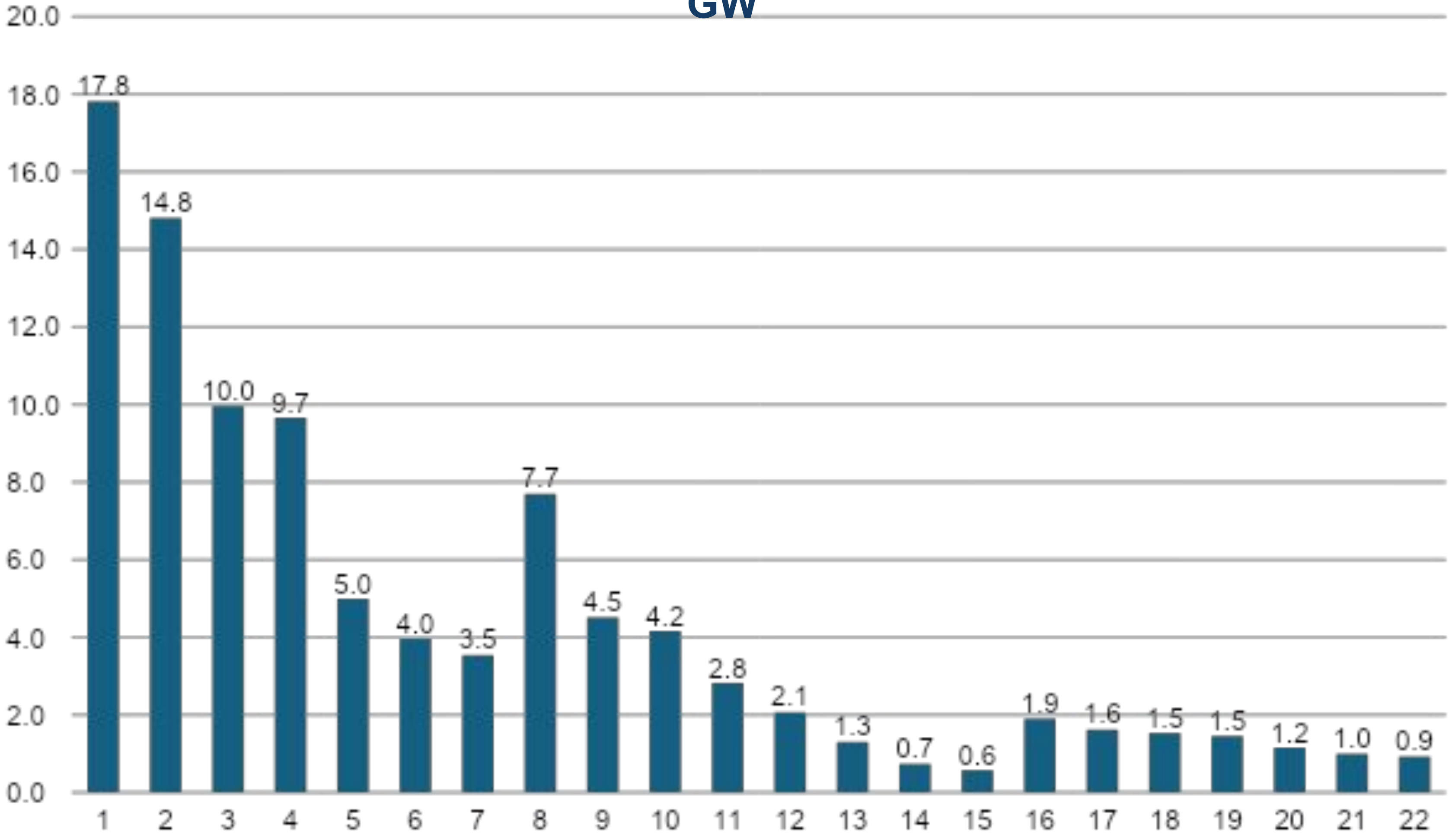
Results: Curtailment-Enabled Headroom

- Headroom across the 22 analyzed balancing authorities is between 76 to 215 GW, depending on the applicable load curtailment limit

- 76 GW of headroom is available at an expected load curtailment rate of 0.25%

- This headroom increases to 98 GW at 0.5% curtailment, 126 GW at 1.0%, and 215 GW at 5.0% curtailment

Headroom Enabled by 0.5% Load Curtailment by Balancing Authority, GW



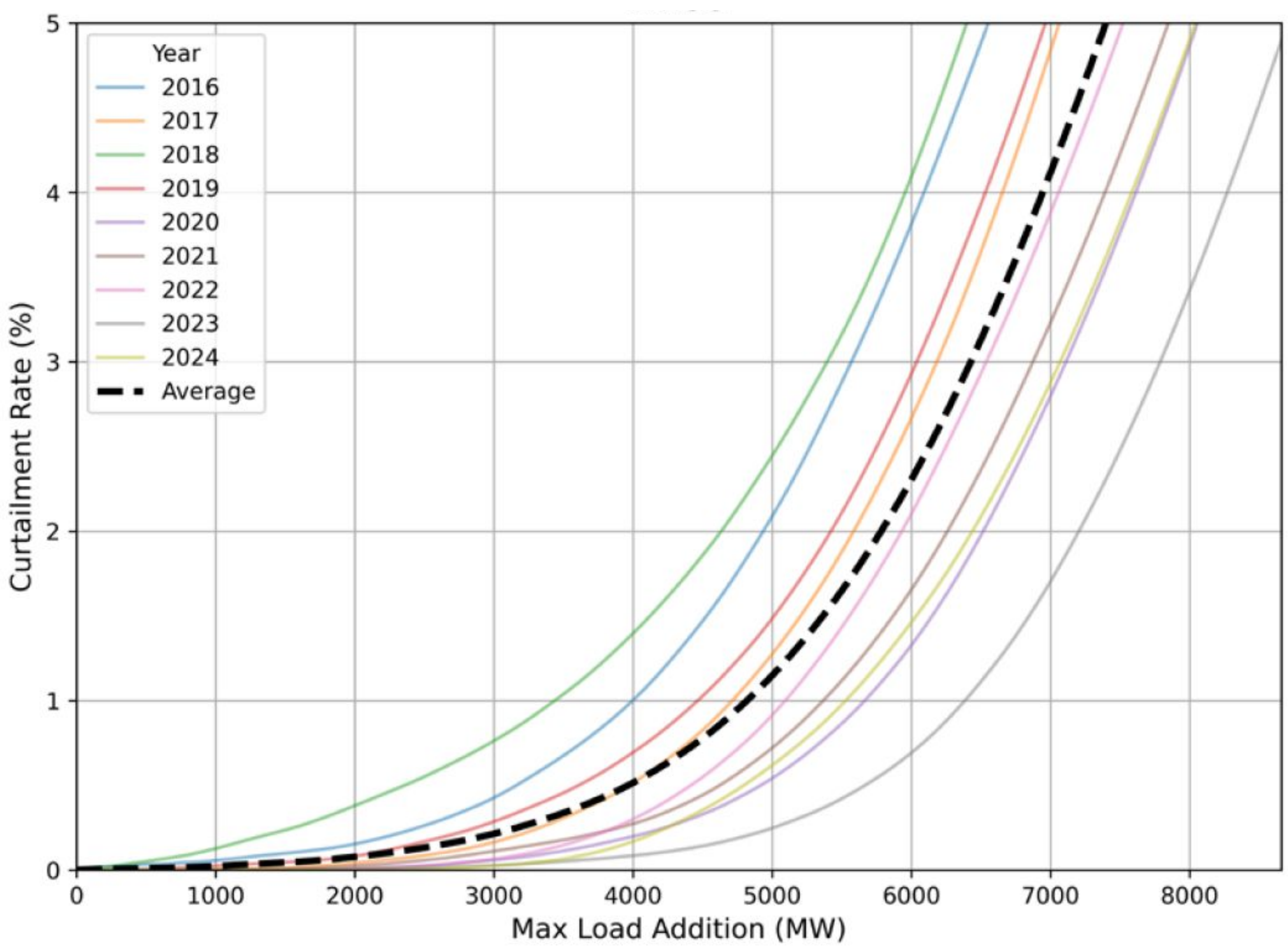
Results: Curtailment-Enabled Headroom

- To visualize the relationship between load additions and curtailment, curtailment was calculated for small incremental load additions

- This plot captures year-by-year variability in demand patterns, including the effects of extreme weather and economic conditions

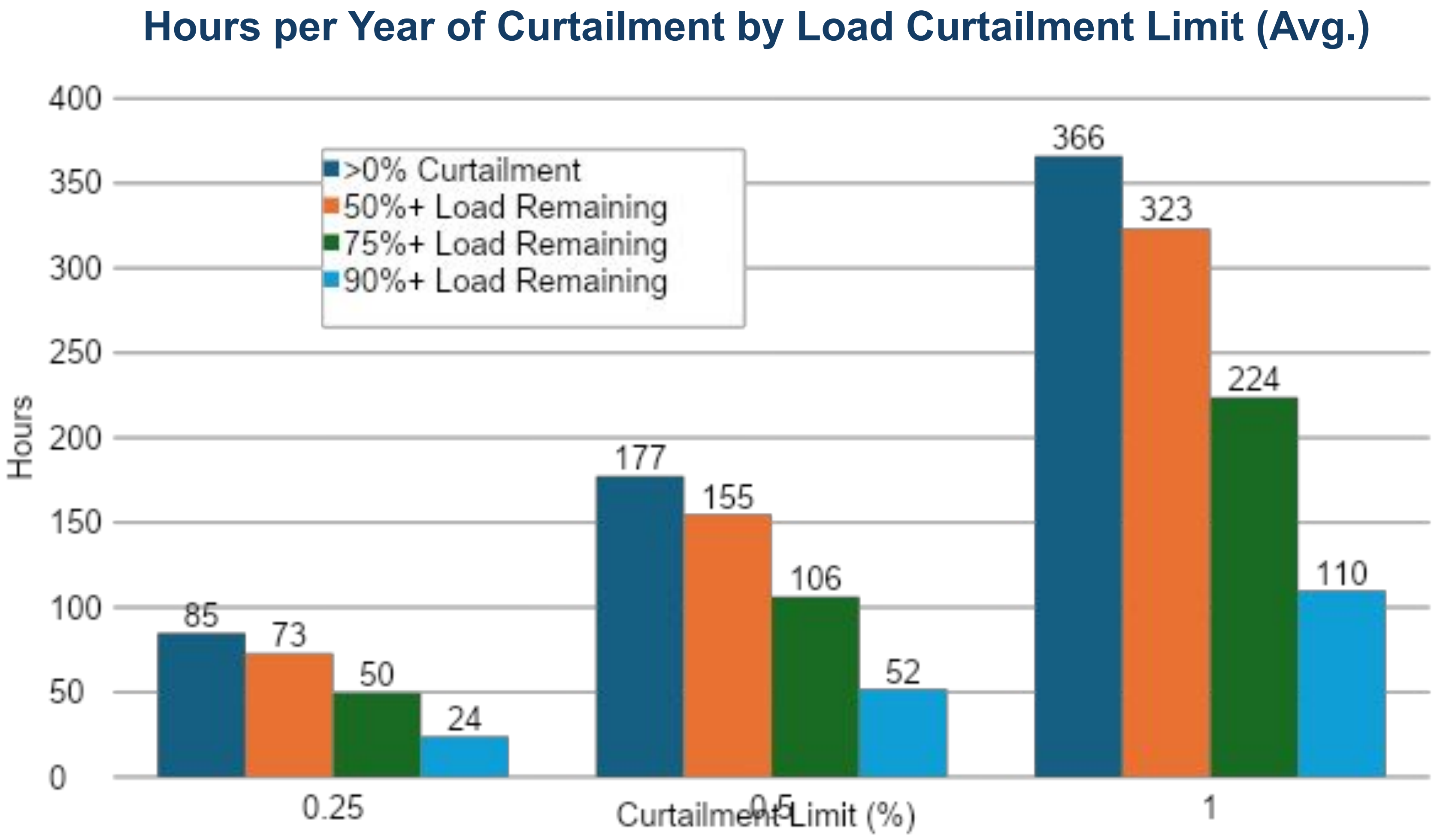
- A similar plot is presented for each balancing authority in Appendix A of the report

Curtailment Rate vs. Load Addition, MW – NYISO



Results: Annual Hours of Curtailment

- A large majority of curtailment hours retain most of the new load
- 88% of hours during which load curtailment is required retain at least half of the new load
- 60% of the hours retain at least 75% of the load, and 29% retain at least 90% of the load

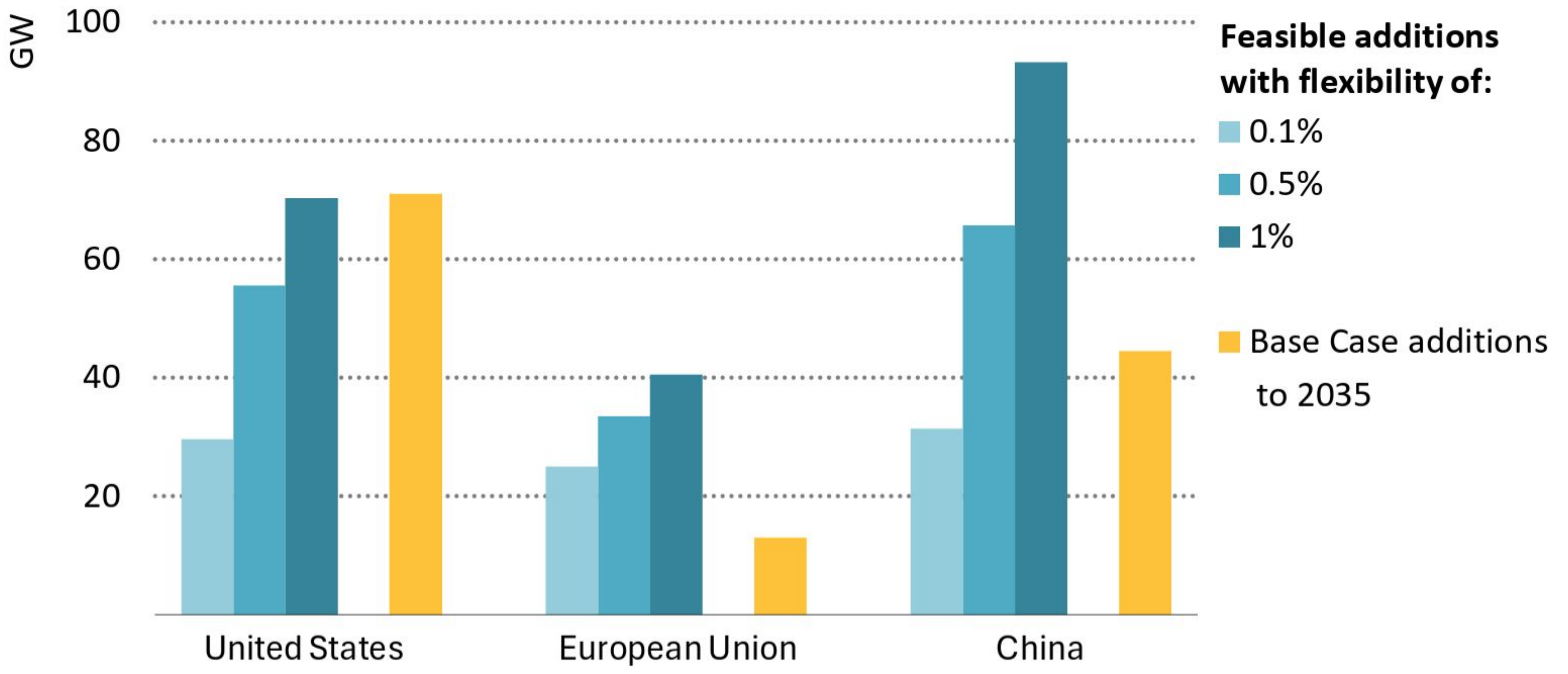


IEA Finds Similar Headroom for US, EU, China

- IEA 2025: "*if data centres are flexible for 0.1-1% of the time, there is enough room in current electricity systems to integrate all new data centre capacities to 2035.*"

- Note: subject to confirmation, the difference in IEA's US results are likely due to using aggregate national demand figures, rather than BA-specific figures

Data Center Capacity Additions to 2035 with Flexibility (IEA 2025)

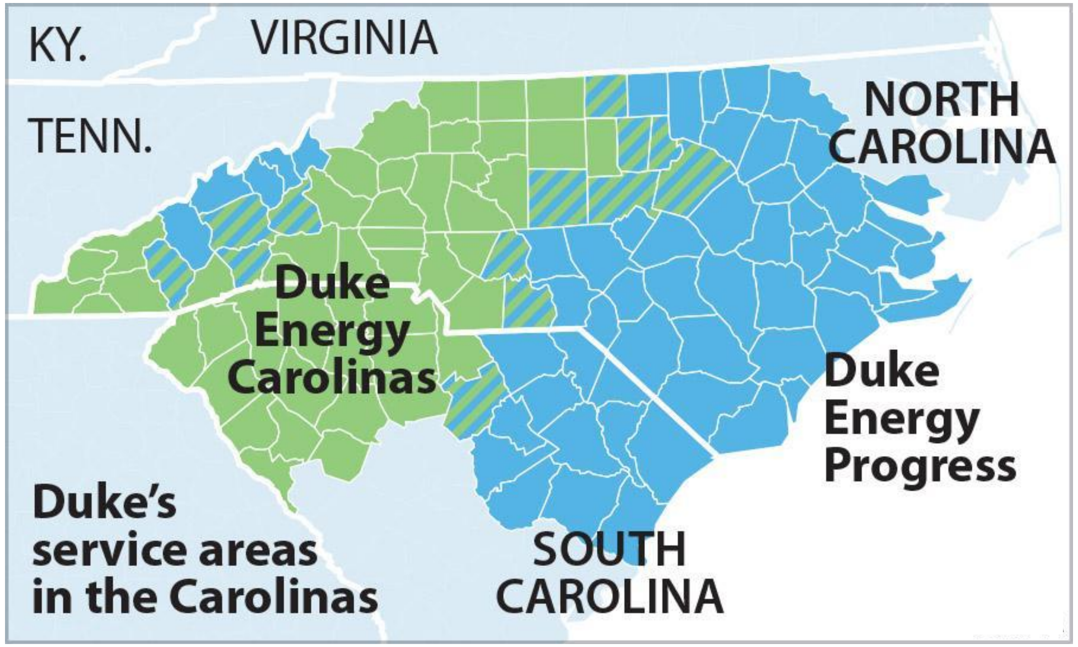


Source: IEA (2025), Energy and AI, IEA, Paris https://www.iea.org/reports/energy-and-ai, Licence: CC BY 4.0

Illustrative Simulation of Load Additions – DEC/DEP

Preliminary results for illustrative purposes only - Not for citation

- **Model:** GRACE EMS (DukeU)
- **Balancing Authorities:** Duke Energy Carolinas (DEC) and Duke Energy Progress (DEP)
- **Combined System Peak**: 37.4 GW (Feb. 2025 record)
- **Load and Weather Year**: 2022
- **Resolution**: Hourly
- **Scenarios**:
  1. Existing load (2022)
  2. 4,100 MW new load
  3. 8,000 MW new load



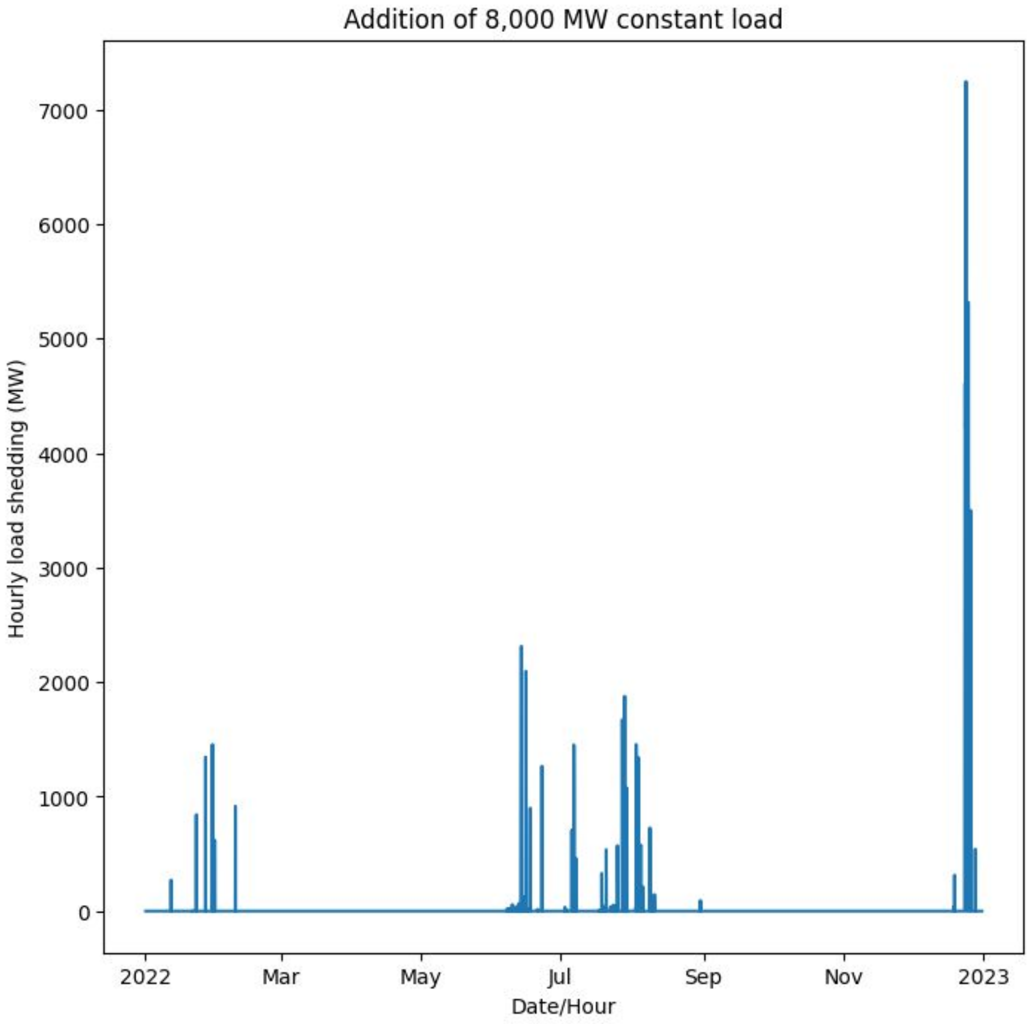
Illustrative Simulation of Load Additions – DEC/DEP

Preliminary results for illustrative purposes only - Not for citation

- **Assumptions:**
  - Addition of a constant 4,100 MW and 8,000 MW new load distributed across DEC/DEP

- **Findings:**
  - 4,100 MW**:** negligible curtailment
  - 8,000 MW: 0.3% of annual energy curtailment

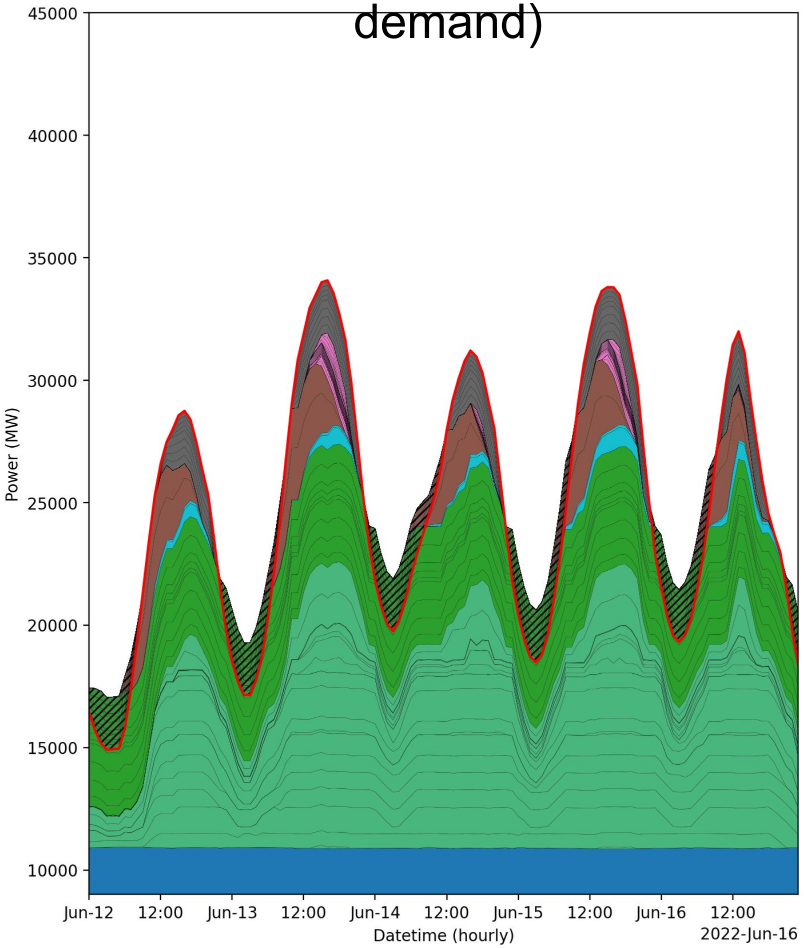
- **Next steps**:
  - Incorporate historical generator deratings for more accurate curtailment ratios
  - Identify “hot-spots” or locations with high/low curtailment, by incorporating network constraints



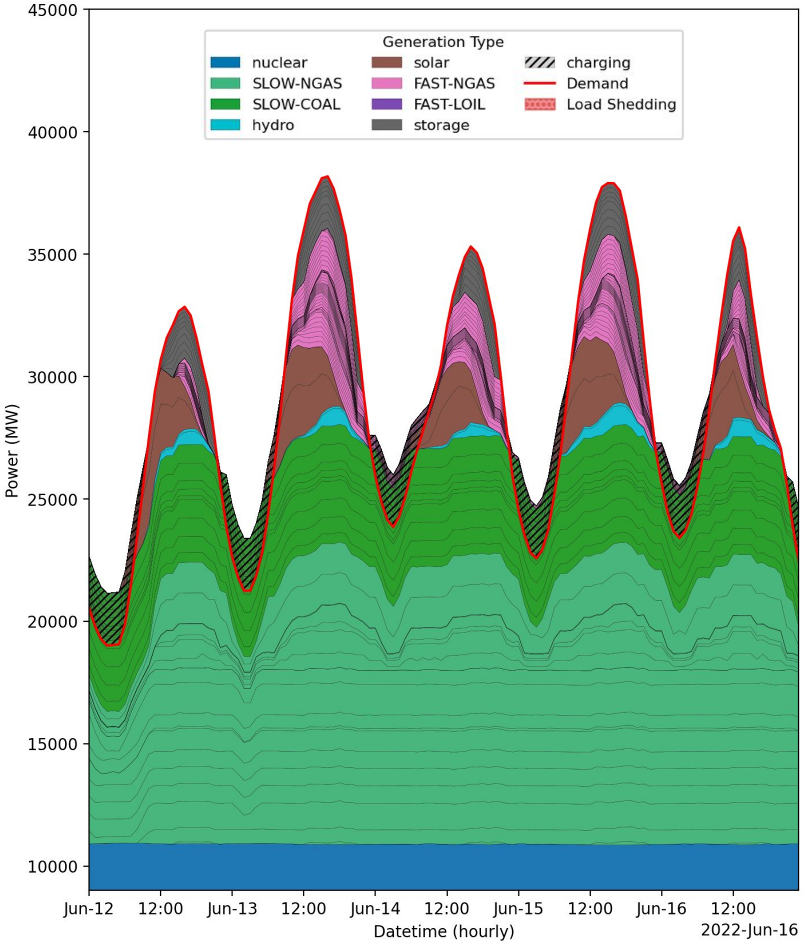
Illustrative Simulation of Load Additions – ~~DEC~~/DEP

Preliminary results for illustrative purposes only - Not for citation

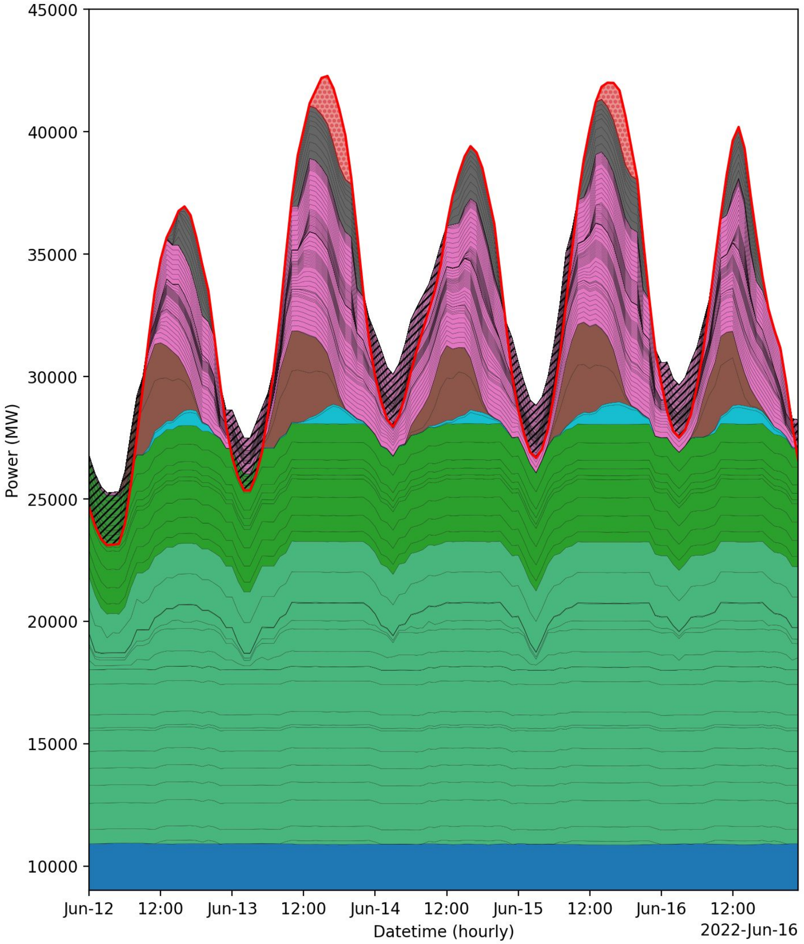
Baseline (2022 historical demand)



4,100 MW additions



8,000 MW additions



Recent developments

Recent developments

- **NYISO flexible load forecast**
  - New Jersey legislation
  - MA Governor’s legislation
  - PA Office of Consumer Advocate
  - Alberta - Beacon AI commitment
  - Texas Senate Bill 6
  - SPP initiative
  - PJM colocation docket
  - Google flexibility contracts

Over 2,000 MW of large demand facilities are expected to be served in New York within the next decade. Approximately 1,200 MW of these large loads are cryptocurrency mining and hydrogen production facilities, and the RNA base case assumes these loads would be flexible during peak conditions to reflect their characteristics…”

Source: NYISO Reliability Needs Assessment 2024

Recent developments

- NYISO flexible load forecast
- **New Jersey legislation**
- MA Governor's legislation
- PA Office of Consumer Advocate
- Alberta - Beacon AI commitment
- Texas Senate Bill 6
- SPP initiative
- PJM colocation docket
- Google flexibility contracts

"The Board of Public Utilities may approve utility rate filings that relax the requirements of this subsection if a large load data center commits to providing sufficient operational flexibility or commits to bringing additional sources of energy and capacity online to meet its load, such that these requirements are not necessary to protect ratepayers."

Source: New Jersey A5462, https://legiscan.com/NJ/text/A5462/2024

Recent developments

- NYISO flexible load forecast
- New Jersey legislation
- **MA Governor's legislation**
- PA Office of Consumer Advocate
- Alberta - Beacon AI commitment
- Texas Senate Bill 6
- SPP initiative
- PJM colocation docket
- Google flexibility contracts

> "Each electric company shall offer a comprehensive flexible interconnection program designed to enable the efficient connection of new customer loads and to maximize the deployment of distributed energy resources, while minimizing associated electric infrastructure costs."

Source: MA Energy Affordability, Independence, and Innovation Act

Recent developments

- NYISO flexible load forecast
- New Jersey legislation
- MA Governor’s legislation
- **PA Office of Consumer Advocate**
- Alberta - Beacon AI commitment
- Texas Senate Bill 6
- SPP initiative
- PJM colocation docket
- Google flexibility contracts

> “… large load tariff provisions that enable interruptible load and backup generation is appropriate so long as ratepayers are protected from stranded costs. Flexibility, optionality, and other creative methods to reduce costs and timelines should be… codified in the proposed model large load tariff so that they are transparent and documented.”

Source: https://www.puc.pa.gov/pcdocs/1875752.pdf

Recent developments

- NYISO flexible load forecast
- New Jersey legislation
- MA Governor’s legislation
- PA Office of Consumer Advocate
- **Alberta - Beacon AI commitment** (next slide)
- Texas Senate Bill 6
- SPP initiative
- PJM colocation docket
- Google flexibility contracts

Beacon AI commitment

"What we committed to with [Alberta Electric System Operator] is that… if we need to pull 400 MW off the grid because of peak usage… we can start the backup generators up and run it for 24 hours until the peak goes down. We'd be happy to do that.

The second thing is once we get to doing our own self-generation, we'll actually be helping contribute back to the grid, because if we're not using it at that moment, then the grid can have it.

We hope that that helps at least, not shelve, but at least address some of the concern around peak load usage because it's a legitimate thing. Not only Beacon, any data center operator, the last position they can be in is that they caused the brownout or some type of over usage on the grid."

– Josh Schertzer, CEO, Beacon AI, June 2025 (source)

Recent developments

- NYISO flexible load forecast
- New Jersey legislation
- MA Governor’s legislation
- PA Office of Consumer Advocate
- Alberta - Beacon AI commitment
- **Texas Senate Bill 6** (next slide)
- SPP initiative
- PJM colocation docket
- Google flexibility contracts

Recent developments

DIVE BRIEF

Texas law gives grid operator power to disconnect data centers during crisis

Senate Bill 6 pairs mandatory curtailment with a voluntary demand response procurement program.

Published June 25, 2025

By Brian Martucci

"After the independent organization deploys all available market services, except for frequency responsive services, the independent organization may direct the applicable electric utility or municipally owned utility to **require the large load customer to either deploy the customer's on-site backup generating facilities or curtail load**."

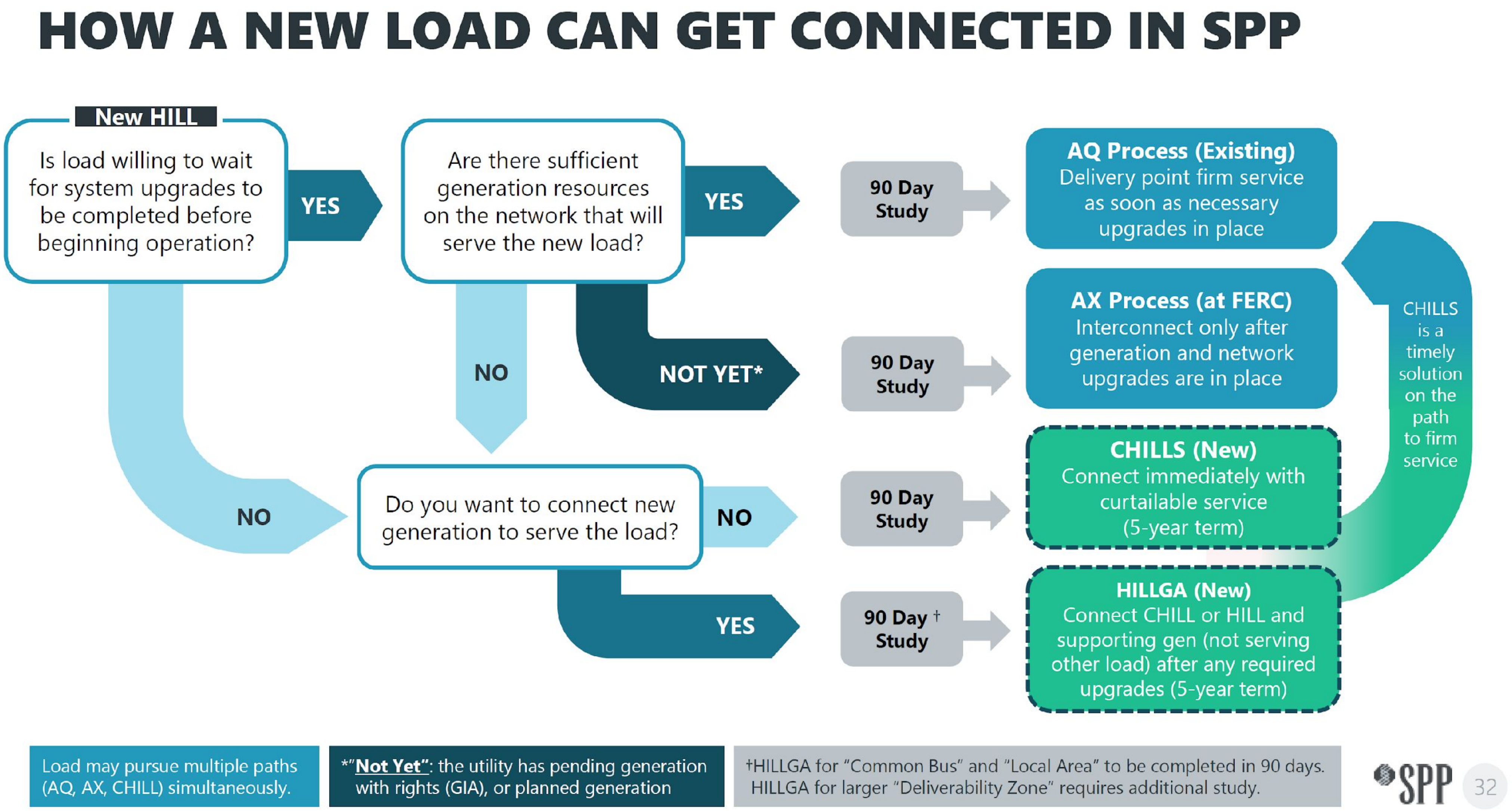
Source: TX Senate Bill 6

Recent developments

- NYISO flexible load forecast
- New Jersey legislation
- MA Governor’s legislation
- PA Office of Consumer Advocate
- Alberta - Beacon AI commitment
- Texas Senate Bill 6
- **SPP initiative (next slide)**
- PJM colocation docket
- Google flexibility contracts

Recent developments

CHILL = Conditional High Impact Large Load
HILLGA = High Impact Large Load Generation Assessment



Recent developments

- NYISO flexible load forecast
- New Jersey legislation
- MA Governor’s legislation
- PA Office of Consumer Advocate
- Alberta - Beacon AI commitment
- Texas Senate Bill 6
- SPP initiative
- **PJM colocation docket**
- Google flexibility contracts

Enabling non-firm loads - PJM colocation docket

PJM's proposed transmission service tiers (#1-3 pre-existing, #4-8 new) - Docket EL25-49

	Tier	Shorthand	Description	Network Load?	Generator retains CIRs?	PJM duty to serve	Load billing
Existing	1	**Classic Firm**	Separate POI, traditional NITS	Yes – separate interconnection	Yes (full)	Full obligation	NITS, energy, ancillary, capacity on full load
	2	**Shared-POI Firm**	Same POI but separate meters	Yes	Yes (full)	Full obligation	Same as Tier 1
	3	**Net-BTM**	Behind-the-Meter Generation (BTMG) netting	Yes, but only net of on-site gen	No (only net injections get CIRs)	Full obligation on net withdrawal	NITS/energy/ancillary/capacity on net withdrawal
Proposed	4	**Island-Only**	"Fully isolated" load behind gen, no grid back-up	No (unless customer elects & state law allows)	Only for output above load	None (unless load opts into NITS)	New rate design needed for ancillary support; otherwise no NITS/capacity charges
	5	**Island + Standby**	Same as Tier 4 but with PJM-approved back-up service	No (becomes network only during PJM-approved intervals)	Same as Tier 4	Only when PJM authorizes back-up	Similar to Tier 4; ancillary still charged
	6	**BYOG Firm**	"Bring-Your-Own-Generation" (BYOG) – load is Network and adds new gen	Yes	Yes (new or repowered gen)	Full obligation	Standard NITS/energy/ancillary/capacity
	7	**Curtailable Firm-Lite**	Non-capacity-backed interruptible service (transitional)	Yes (but commits to curtail before emergencies)	Yes	Reduced priority; curtailable	Pays capacity only on non-interruptible portion
	8	**Self-Curtail DR**	Network load that enrolls as Demand Response	Yes	Yes	Full obligation (but may curtail via DR)	Standard charges; DR revenues offset

Enabling non-firm loads - PJM colocation docket

"According to the Indicated Transmission Owners, however, all new load should be required to take front-of-the-meter service and become a [Network] customer paying for service based on gross load…

It is no surprise that they want to foreclose customers from any other option: large new [Network] customers mean extensive system upgrades, which the transmission owners get to build and add to their rate base.

Having this be the only option in PJM is unjust and unreasonable: it results in years of delay in connecting new load, induces massive and unnecessary overbuild of the system, imposes substantial additional costs on existing customers who are already struggling to pay their bills, and fails to send efficient signals about where load should be sited."

–Constellation Energy, 4/23/25, PJM Colocation Docket #EL25-49

Recent developments

- NYISO flexible load forecast
- New Jersey legislation
- MA Governor’s legislation
- PA Office of Consumer Advocate
- Alberta - Beacon AI commitment
- Texas Senate Bill 6
- SPP initiative
- PJM colocation docket
- **Google flexibility contracts**

Google announces new flexibility contracts

Why it matters:

- First-of-its-kind US agreements
- From operational flexibility to planning flexibility
- Inclusion of machine learning workloads
- Definitive, long-term contracts
- Time-bound flexibility

Reuters

Google agrees to curb power use for AI data centers to ease strain on US grid when demand surges

By Laila Kearney

August 4, 2025 12:48 PM EDT · Updated August 4, 2025

Aa

Duke NICHOLAS INSTITUTE *for* ENERGY, ENVIRONMENT & SUSTAINABILITY

GRACE Lab
A Grid that is Risk-Aware for Clean Electricity

NICHOLAS SCHOOL *of the* ENVIRONMENT

SIGN UP FOR EMAIL:
bit.ly/niemail

@NichInstitute

bit.ly/nilinkedin

@NichInstitute

@NichInstitute

@NicholasInstitute