ESIG
Large Loads
Task Force

# Resource Adequacy with Large Loads

## PLANNING FOR FLEXIBILITY TO ACCELERATE INTEGRATION

A Report by the
Energy Systems Integration Group's
Large Loads Task Force

**April 2026**

ESIG
ENERGY SYSTEMS
INTEGRATION GROUP

ESIG
ENERGY SYSTEMS
INTEGRATION GROUP

## About ESIG

The Energy Systems Integration Group is a nonprofit organization that marshals the expertise of the electricity industry's technical community to support grid transformation and energy systems integration and operation. More information is available at https://www.esig.energy.

## ESIG Publications Available Online

This report is available at https://www.esig.energy/working-groups/large-loads. All ESIG publications can be found at https://www.esig.energy/resources/reports-and-briefs/.

## Get in Touch

To learn more about the topics discussed in this report or for more information about the Energy Systems Integration Group, please send an email to info@esig.energy.

© 2026 Energy Systems Integration Group

# Resource Adequacy with Large Loads: Planning for Flexibility to Accelerate Integration

**A Report by the Energy Systems Integration Group's Large Loads Task Force**

## Prepared By

**Aaron Schwartz**, Telos Energy

**Derek Stenclik**, Telos Energy

## Project Managers

**James Okullo**, Energy Systems Integration Group

**Lisa Schwartz**, Lawrence Berkeley National Laboratory

## Project Team Members

**Mark Ahlstrom**, Energy Systems Integration Group

**Margaret Albright**, Bonneville Power Administration

**Eric Baran**, Western Interconnection Regional Advisory Body

**Conleigh Byers**, Harvard University

**Jim Caldwell**

**Juan Pablo Carvallo**, Lawrence Berkeley National Laboratory

**Anthony DeAugustino**, Xcel Energy

**Surya Dhulipala**, National Laboratory of the Rockies

**Michael Dzurak**, Arizona Public Service

**Erik Ela**, Energy Systems Integration Group

**Frank Felder**, Independent Electricity Consultants

**Maren Ihlemann**, EPRI

**Monica Jha**, Vistra Corp

**Sean Jones**, Tesla

**Nikhil Kumar**, GridLab

**Debra Lew**, Energy Systems Integration Group

**Thair Shaker Mahmoud**, Electric Reliability Council of Texas

**Trieu Mai**, Energy Systems Integration Group

**Subir Majumder**, Texas A&M University

**Roxana Melendez**, United States Department of Energy

**Varsha Menon**, United States Department of Energy

**Parag Mitra**, EPRI

**Nihal Mohan**, NextEra Energy Transmission

**Zachary Morales**, Northeast Power Coordinating Council

**Katie Mulvaney**, Colorado Public Utilities Commission

**Joshua Novacheck**, NextEra Energy

**Fernando Palma**, United States Department of Energy

**Alejandro Palomino**, Energy Strategies

**Russ Philbrick**, Polaris Systems Optimization

**Fatemeh Rajaei Najafabadi**, Imperial College London

**Abraham Silverman**, Johns Hopkins University

**Anna Sommer**, Energy Futures Group

**Paul Spitsen**, United States Department of Energy

**Priya Sreedharan**, GridLab

**Aaron Stoll**, Grid United

**Ben Wheeler**, Verrus

**John D. Wilson**, Grid Strategies

**Matthew Zapotocky**, Western Electricity Coordinating Council

**Zhi Zhou**, Argonne National Laboratory

## Disclaimer

This report was produced by a project team made up of diverse members with diverse viewpoints and levels of participation. Specific statements may not necessarily represent a consensus among all participants or the views of participants' employers.

## Acknowledgment

This work was supported by the U.S. Department of Energy under Lawrence Berkeley National Laboratory Contract No. DE-AC02-05CH11231.

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, or The Regents of the University of California. Ernest Orlando Lawrence Berkeley National Laboratory is an equal opportunity employer.

## Suggested Citation

Energy Systems Integration Group. 2026. *Resource Adequacy with Large Loads: Planning for Flexibility to Accelerate Integration*. A report by the Large Loads Task Force. https://www.esig.energy/reports-briefs/large-loads-resource-adequacy.

### PHOTOS

Cover: © iStockphoto/alacatr

p. x: © iStockphoto/Gerville

p. xii: © iStockphoto/monsitj

p. xiii: © iStockphoto/claffra

p. 5: © iStockphoto/VichienPetchmai

p. 10: © iStockphoto/YoungNH

p. 12: © iStockphoto/Jacques Delespee

p. 15: © iStockphoto/tigerstrawberry

p. 21: © iStockphoto/SweetBunFactory

p. 32: © iStockphoto/Hugo Kurk

p. 42: © iStockphoto/alacatr

## Preface to ESIG Large Loads Task Force Reports

This report is one of 11 reports by the ESIG Large Loads Task Force, which was formed to assist the power industry in addressing new challenges introduced by the rapid proliferation of large electronic loads such as data centers, as well as other large loads including manufacturing, electric vehicle fleets, and hydrogen production. The titles of the reports are as follows:

- Grid Integration of Large Loads: Introduction to the Large Loads Task Force, Data Needs, and Flexibility Overview
- Forecasting for Large Loads: Current Practices and Recommendations
- Interconnection Processes for Large Loads: Current Practices and Recommendations
- Large Load Performance Requirements: Current Practices and Recommendations
- Large Loads: Behaviors, Capabilities, and Limitations
- Reliability Impacts of Large, Power Electronics–Interfaced Loads
- Large Load Disturbance Events
- Large Load Modeling for Dynamic Studies: Current Practices and Recommendations
- Transmission Planning with Large Loads: Current Practices and Recommendations
- Resource Adequacy with Large Loads: Planning for Flexibility to Accelerate Integration
- Wholesale Market Design and Operations for Systems with Large Loads: Current Practices and Recommendations

# Contents

# Abbreviations

| | |
|---|---|
| AI | Artificial intelligence |
| BYONG | Bring your own new generation |
| CHILLS | Conditional High-Impact Large Load Service |
| ELCC | Effective load-carrying capability |
| ERCOT | Electric Reliability Council of Texas |
| FERC | Federal Energy Regulatory Commission |
| GPU | Graphical processing unit |
| HILLGA | High-Impact Large Load Generation Agreement |
| ISO | Independent system operator |
| LOLE | Loss-of-load expectation |
| LSE | Load-serving entity |
| PRM | Planning reserve margin |
| SPP | Southwest Power Pool |

# Executive Summary

Electricity demand in the United States is growing at a pace not seen in decades, driven primarily by a new large loads such as data centers, industrial electrification, hydrogen production, and electric vehicle fleets. These loads are unprecedented in size, with some hyperscale data center loads approaching or surpassing 1 GW at a single point of interconnection, and they are emerging during a time of tightening supply conditions caused by fossil plant retirements and supply chain bottlenecks for new generation.

This convergence of surging demand and constrained supply puts significant pressure on the resource adequacy of the bulk power system. Traditional planning frameworks, designed for incremental load growth and predictable resource mixes, are being tested by customers that can drive multi-gigawatt shifts in demand on short timelines. The risks are clear: if load growth outpaces the addition of new resources, system reliability could be eroded; conversely, delaying interconnections to preserve reliability threatens the economic viability of new developments and the nation's competitiveness in critical industries like artificial intelligence (AI) and advanced manufacturing.

**By treating large loads' flexibility as a planned capacity resource considered up front in planning, utilities and regional grid operators can transform a potential source of resource adequacy risk into a system asset.**

To manage these risks, the industry can move beyond short-lead-time measures developed outside of long-term planning—such as curtailment programs—and assess large load impacts up front in the electricity system planning process. This includes evaluating the latent flexibility potential of large data center loads, which can be operated through several pathways such as intelligent workload management, thermal control strategies, and on-site generation. Utilities, grid operators, and large loads—including hyperscale data center developers—from across the United States and globally have already called on the industry to align on a flexibility framework as a key tool for accelerating time to power.*

By treating this flexibility as a planned capacity resource considered up front, utilities and regional grid operators can transform a potential source of resource adequacy risk into a system asset. Integrating large load flexibility into modeling for long-term resource adequacy allows planners to flatten the peaks that drive capacity requirements, reducing the need for new firm generation, while increasing the utilization of existing capacity and thereby increasing system efficiency and reducing electricity system costs.

## Managing Resource Adequacy Risks in a Tight Supply Environment

Unlike the incremental, population- and economy-driven load growth of the past, today's large load development is characterized by discrete, binary events: a single data center or industrial facility can add hundreds of megawatts of demand in a stepwise shift. Forecasting these loads is exceptionally difficult, as data center developers often pursue multiple sites simultaneously to maximize interconnection success, clogging interconnection queues with many requests that may never materialize.

---

* https://dcflex.epri.com/flex-mosaic

Furthermore, the operating characteristics of these loads are diverse and rapidly evolving. The hour-to-hour operations of large loads are unique, even within a single class of large loads. For example, a data center's impact on the grid varies significantly depending on whether it is used for AI training, inference, or cloud storage—distinctions that are often invisible to the host utility or other grid operator. Without granular data on these end uses, planners may be challenged to model the potential coincidence of these loads with electricity system peaks or assess the ability of these loads to provide demand flexibility.

The forecasting uncertainties are compounded by supply-side resources that are increasingly capacity-constrained. New large loads are seeking to interconnect just as aging fossil generation is retiring and the deployment of new resources faces headwinds. Supply chain bottlenecks for critical equipment, such as gas turbines and transformers, along with lengthy generator interconnection queues, have limited the ability to add firm capacity in the near term.

This imbalance creates a growing tension between the utility's "obligation to serve" and the physical imperative to maintain system reliability. Navigating this environment requires finding solutions to integrate these loads without compromising resource adequacy.

## Current Approaches to Mitigating Near-Term Resource Adequacy Risks

Faced with the magnitude of new large loads and the urgency of near-term interconnection requests, utilities and regional grid operators are deploying a range of short-term measures to preserve reliability. Broadly, these approaches fall into three categories: interconnection moratoriums, load-curtailment programs, and "bring your own new generation" (BYONG) arrangements.

### Interconnection Moratoriums

In regions where system conditions have become too constrained to reliably interconnect new load, some utilities have issued interconnection moratoriums on new data centers. These pauses allow utility planning departments time to study the reliability impact of pending service requests and identify necessary mitigations. While effective at preserving near-term reliability, moratoriums are blunt instruments that can undermine broader economic development goals.

### Load-Curtailment Programs

Curtailment-based reliability programs allow new large loads to connect under the condition that their net output may be reduced during system emergencies or periods of high risk to preserve resource adequacy. Some curtailment-based reliability programs, such as SPP's Conditional High-Impact Large Load Service (CHILLS), allow large loads to interconnect on an accelerated timeline in exchange for being subject to curtailment during system emergencies. Others, such as those required by Texas Senate Bill 6, do not provide accelerated interconnection and instead impose mandatory requirements for large loads. These programs ensure that new large loads do not degrade reliability for existing customers and place operational risk onto these new loads. That potentially makes these programs less attractive than firm service pathways or those that provide large loads with greater agency in how and when they provide flexibility services.

### Bring-Your-Own-New-Generation Arrangements

Another approach involves BYONG arrangements, such as under SPP's High-Impact Large Load Generation Agreement (HILLGA), or PJM's January 2026 Critical Issue Fast Path Proposal, which allow large loads to offset their grid impact through self-supplied capacity. By committing to rely on on-site or contracted generation resources, large load customers can bypass certain capacity constraints and interconnect more quickly.

### Limitations of Interim Measures

While these mechanisms provide valuable tools for managing near-term risks, integrating them into long-term planning processes and frameworks is necessary to support the most cost-effective outcomes for long-term resource adequacy. Without integration, these tools can lead to suboptimal outcomes: piecemeal procurement of on-site resources may not align with broader electricity system needs, and curtailment risk can undermine the business case for large load development. Furthermore, without careful design, these programs can distort investment signals, potentially reducing the incentives for new firm capacity to enter the market. Moving from these short-lead-time solutions to a robust, long-term resource adequacy strategy requires integrating large load flexibility up front into system planning.

## A Framework for Implementation: The Six-Step Process

This report outlines a standardized six-step process to effectively integrate large load flexibility into long-term electricity system planning—specifically, for resource capacity (Figure ES-1). The framework leverages established planning tools, including capacity expansion and probabilistic resource adequacy models, to systematically evaluate how load flexibility affects resource adequacy and investment needs. By isolating the incremental impact of flexibility, the framework offers a transparent, data-driven foundation for decision-making.

### Step 1: Characterize Loads with Greater Fidelity

Planners need to move beyond generic load representations by segmenting large loads based on key operational features such as size, ramp rates, and flexibility potential. In addition, distinguishing between end uses, such as AI

FIGURE ES-1

Six-Step Process to Plan for Large Load Flexibility Up Front in the Capacity Planning Process

| Step 1 | Step 2 | Step 3 | Step 4 | Step 5 | Step 6 |
|---|---|---|---|---|---|
| Characterize loads with greater fidelity | Improve large load forecasts | Quantify capacity requirements of large loads without flexibility | Quantify capacity contributions of large load flexibility | Calculate the avoided infrastructure investments resulting from large load flexibility | Develop regulatory mechanisms that account for the value of large load flexibility |

Source: Energy Systems Integration Group.

training versus inference or hydrogen production, enables planning models to capture critical differences between loads such as weather dependence and coincidence with system peaks.

### Step 2: Improve Large Load Forecasts

Planners can develop "large-load layers" that separate these loads from other loads and employ scenario-based forecasting approaches to address uncertainty. Planners can also ground large load forecasts using evidence beyond interconnection requests and include high- and low-growth "bookend" scenarios to capture the full range of potential load-growth realization outcomes.

### Step 3: Quantify Capacity Requirements of Large Loads Without Flexibility

This step establishes a baseline by calculating the total amount of accredited capacity required to meet the resource adequacy criterion if new large loads were to come online as fully firm, inflexible demand. This "no-flexibility" scenario can be compared against scenarios that consider large load flexibility (Step 5) to compute its incremental value. The no-flexibility scenario can also be compared against a "business as usual" scenario that assumes no new large loads come online, allowing planners to attribute specific capacity requirements and costs directly to inflexible new large loads.

### Step 4: Quantify Capacity Contributions of Large Load Flexibility

Using probabilistic methods like effective load-carrying capability (ELCC), planners can accredit large load flexibility as a capacity resource, allowing it to be treated comparably to generation resources in capacity expansion modeling. This accreditation quantifies the firm capacity value of flexibility and can reflect how that value may vary based on parameters such as response duration, frequency of calls for flexibility, and load availability. Alternatively, the capacity contributions of large loads can be represented through a reduction in the planning reserve margin (PRM) requirement, whereby only the firm (non-flexible) portion of large load demand counts toward the PRM requirement. This approach does not require capacity accreditation of flexible large loads and instead socializes their flexibility benefits across all customers.

### Step 5: Calculate the Avoided Infrastructure Investments Resulting from Large Load Flexibility

By incorporating the accredited flexibility (or reduced PRM) into capacity expansion models, planners can determine the type and amount of supply- and demand-side resources that can be avoided or deferred. This step yields quantifiable metrics on system cost savings, providing an economic foundation on which to design flexibility incentives.

### Step 6: Develop Regulatory Mechanisms That Account for the Value of Large Load Flexibility

Finally, the quantitative insights from the previous steps are translated into regulatory mechanisms. Avoided costs and deferred investments can inform the design of large load tariffs, interconnection agreements, and market rules that fairly allocate costs and reward demand flexibility, aligning the economic interests of large customers with system adequacy goals.

## Methods and Data Needs for Effective Modeling

Accurately representing large loads within resource adequacy and capacity expansion models depends on integrating detailed, segment-specific data into planning workflows. Critical to this effort are 8,760-hour chronological load profiles that capture the coincidence of large load demand with system peaks. Where relevant, it is important that these profiles reflect weather dependence, particularly for cooling-intensive loads like data centers. Large load forecasts can also specify zonal locations to evaluate local resource capacity and transmission constraints, as well as ramp rates to capture how quickly demand will scale after the large load is energized.

In addition, planners need detailed specifications on on-site generation and storage assets, including generator size and run-hour limitations, to assess their potential reliability contributions. Given the high uncertainty around large load development, modelers can employ “bookend” scenarios that span both high and low load-growth outcomes and assign probabilities of realization to individual large load projects to better capture uncertainty.

Effectively accrediting large load flexibility as a capacity resource—for example, to support the six-step framework described in this report—requires additional data. Planning models can vary parameters such as the size of the flexibility response (MW), frequency of deployment, and duration of each flexibility event to test different combinations and assess how they affect demand flexibility accreditation. Models can account for event-specific limits, such as maximum depth of flexibility response, as well as horizon-wide constraints like annual energy budgets or total call limits. And to ensure operational realism, models ideally would incorporate response performance rates, representing the probability that flexibility will materialize when called, and account for weather-driven variations in flexibility availability.

## Recommendations

Operationalizing the six-step framework requires coordinated action across the industry. Table ES-1 (p. xiv) summarizes targeted recommendations for stakeholders to support proactive, flexibility-driven large load planning.

To ensure a reliable and affordable grid in the face of unprecedented load growth, the electric power industry can develop demand flexibility solutions that are harmonized with long-term planning and market design processes. Incorporating large load flexibility up front in the resource adequacy assessment process is one such solution essential for ensuring the timely interconnection of large loads while managing costs and system-wide reliability impacts. By quantifying the flexibility capabilities of large loads and potential capacity savings through the proposed six-step framework, planners can support win-win solutions for key stakeholders.

TABLE ES-1

## Summary of Recommendations

| Stakeholder | # | Recommendation |
|---|---|---|
| **Resource adequacy modelers** | A1 | Incorporate multiple load-growth scenarios |
| | A2 | Ensure close coordination with demand forecasters |
| | A3 | Evaluate the incremental impact of large load flexibility |
| **Resource planners at vertically integrated utilities** | B1 | Plan and procure large load flexibility up front |
| | B2 | Model multiple planning scenarios to isolate the capacity requirements attributable to large loads |
| | B3 | Incorporate flexible-load accreditation (e.g., ELCC) in integrated resource plans |
| | B4 | Solicit flexibility in negotiated tariffs and exchange faster interconnection for flexibility |
| **Independent system operators' and regional transmission organizations' market operators and planners** | C1 | Clarify the treatment of on-site generation and "bring your own new generation" (BYONG) options in capacity markets |
| | C2 | Develop clear accreditation rules and participation pathways for large load flexibility in capacity markets |
| | C3 | Incorporate flexibility options into interconnection and capacity market processes |
| | C4 | Design non-voluntary load curtailment programs carefully to avoid distorting capacity investment signals |
| **Data centers and other large loads** | D1 | Plan for flexibility early in the large load design and development process |
| | D2 | Ensure grid operator visibility through robust control architecture and telemetry configurations |
| | D3 | In bilateral contracts with utilities, provide firm flexibility commitments with clear limits |
| | D4 | Consider utility contracts that monetize avoided capacity and support faster interconnection |
| | D5 | Work with utilities and regional grid operators to provide representative, detailed operational data to support effective resource adequacy modeling |
| **State utility regulators** | E1 | Isolate the costs of capacity attributed to new large loads to ensure that cost allocation and retail rate structures are fair to all utility customers |
| | E2 | Use results from integrated resource planning modeling that assess incremental large load impacts—with and without flexibility—to inform cost allocation and tariff design |
| | E3 | Assess whether traditional/existing demand response programs provide appropriate incentives for desired flexibility responses |
| | E4 | Set goal posts around the types of flexibility that are needed rather than prescribe specific flexibility technologies |
| | E5 | Incorporate flexibility incentives and/or requirements in large load tariffs and rate designs |
| | E6 | Leverage utility resource adequacy analysis to set a cap on the amount of large load flexibility needed |

# Introduction

Electricity demand in the United States is growing at a pace not seen in decades, driven primarily by new large loads such as data centers, industrial electrification, hydrogen production, and vehicle fleet electrification.[1] These loads are reshaping power system planning. This increased demand growth is driving new investment in generation, transmission, and distribution assets and will require modernizing aging grid infrastructure. Managing this growth reliably and affordably poses new and complex challenges for regional transmission operators and regional transmission organizations, independent system operators (ISOs), utilities, and regulators tasked with maintaining an affordable grid.

Today's large loads are challenging compared to historical demand growth due to several factors. These include their unprecedented size, which in some cases approaches or surpasses 1 GW at a single point of interconnection; the high volume of large load interconnection requests; uncertainty surrounding the likelihood of large load materialization; and their distinct demand profiles. These challenges are compounded by rapidly evolving policies, regulations, and energy resource mixes. Further, supply chain bottlenecks—especially for gas turbines, transformers, and other equipment—limit the ability to add new firm capacity in the near term. Together, these factors create a tight supply environment just as new large loads are emerging at record scale.

These challenges are putting pressure on the resource adequacy of the country's power grids. Traditional frameworks for ensuring resource adequacy presume incremental load growth and a predictable mix of generation. But these frameworks are being tested by a class of customers that can drive near-term, multi-gigawatt shifts in demand. If load growth outpaces the rate of generation additions, system reliability could erode. Alternatively, grid operators may delay interconnections to maintain reliability, which would have adverse effects for new customers seeking to quickly energize. To maintain reliability, affordability, and fairness across existing and prospective customers in this environment will require innovations in power system planning and operations. Technological, policy, and regulatory innovations that tap into the demand flexibility potential of new large loads present an especially valuable opportunity for achieving power system goals in a way that meets the needs of utilities, large load customers, and other utility customers.

**Traditional frameworks for ensuring resource adequacy presume incremental load growth and a predictable mix of generation. But these frameworks are being tested by a class of customers that can drive near-term, multi-gigawatt shifts in demand.**

To identify challenges and develop harmonized best practices for integrating large loads reliably and efficiently, the Energy Systems Integration Group (ESIG) launched the Large Loads Task Force, which convened a multi-stakeholder group of ISOs, utilities, data center operators, regulators, and researchers. The Large Loads Task Force consists of seven specialized project teams, each focusing on a key aspect of large load integration, from interconnection and operations to planning and market design.

---

[1] Nationally, there is no single definition of what constitutes a "large load." The ESIG Large Loads Task Force uses the following definition: "Any commercial or industrial individual load facility or aggregation of load facilities at a single site behind one or more point(s) of interconnection that can pose reliability risks to the BPS [bulk power system] due to its demand, operational characteristics, or other factors" (NERC, 2025a, 1).

This report, developed by the resource adequacy project team, focuses on how to ensure resource adequacy and adapt the planning, operational, and market frameworks that underpin reliable electricity service in an era of rapid, uncertain, and unprecedented large load growth. In particular, flexible operations for data centers are a key tool for ensuring resource adequacy, as is a framework to consider flexibility as part of utility and ISO planning and market design processes.

## Resource Adequacy Challenges Stemming from Large Loads

Resource adequacy is the ability of the bulk power system to maintain reliable operation under uncertainty in both generator availability and load. Resource adequacy analysis evaluates whether enough capacity is available to meet demand across seasons, days, and hours, accounting for outages, variable energy resources, and forecast errors. These analyses produce quantitative metrics that can help planners characterize the system's unique risks in terms of the frequency, duration, and severity of emergency events driven by capacity shortfalls, for example, loss-of-load probability or expected unserved energy. These metrics help guide investment decisions about the types and quantities of resources needed to achieve an acceptably low level of reliability.

Among the various types of large loads, data centers provide the most immediate and transformative challenge for resource adequacy (see Box 1). Fueled by the rapid expansion of artificial intelligence (AI) and cloud computing, data center electricity use has more than doubled in recent years, from roughly 1.9% to 4.4% of total U.S. electricity consumption between 2018 and 2023 according to Lawrence Berkeley National Laboratory (Shehabi

BOX 1

### What Do New Large Loads Mean for Resource Adequacy Planning?

A central concept in resource adequacy is the **planning reserve margin (PRM) requirement**, the expectation that utilities or load-serving entities (LSEs) procure sufficient accredited resources to meet their customers' expected peak demand, plus a buffer—the PRM. This buffer helps ensure that the grid maintains sufficient capacity to operate reliably even if forecast error, extreme weather, or unexpected system conditions lead to higher-than-anticipated demand or generator underperformance. This capacity can be provided by supply-side resources such as new generation or battery energy storage systems or by new demand-side resources such as distributed energy resources and demand-side load flexibility.

When a new large load connects to the grid as a firm customer (a customer with contractual guarantees for uninterruptable service), system demand increases and the utility or LSE that serves the large load may be required to procure additional accredited capacity. Because of the PRM requirement, however, new firm loads require LSEs to secure accredited capacity in excess of the contracted firm load. For example, in a region with a 15% PRM, a new load that drives a 100 MW increase in system peak demand would require 115 MW of new accredited capacity.

The regulatory and financial obligation to procure this capacity generally falls on the host utility or LSE. In restructured markets, where regulated utilities do not own generation assets, LSEs typically have a capacity obligation to procure sufficient accredited capacity to cover the PRM requirement for their share of load. In other regions, all utilities remain vertically integrated and often meet their PRM requirement through integrated resource planning and procurement (Biewald et al., 2024).

In today's supply- and transmission-constrained environment, contracting or building new generation capacity to meet PRM requirements—and delivering it to new large loads—is increasingly difficult and costly, underscoring the value of flexible solutions that can reduce the firm generation capacity required to serve these loads.

et al., 2024). The same study projects data center electricity consumption to reach 6.7% to 12.0% by 2028. This surge has already begun to tighten current and prospective PRMs in several regions (Figure 1). As load growth forecasts increase faster than supply-side resource deployment, the amount of generation capacity projected to be available in excess of forecasted load is diminishing. Utilities have been forced to reconsider near-term generation and storage capacity additions, defer retirements, and explore new mechanisms for managing load.

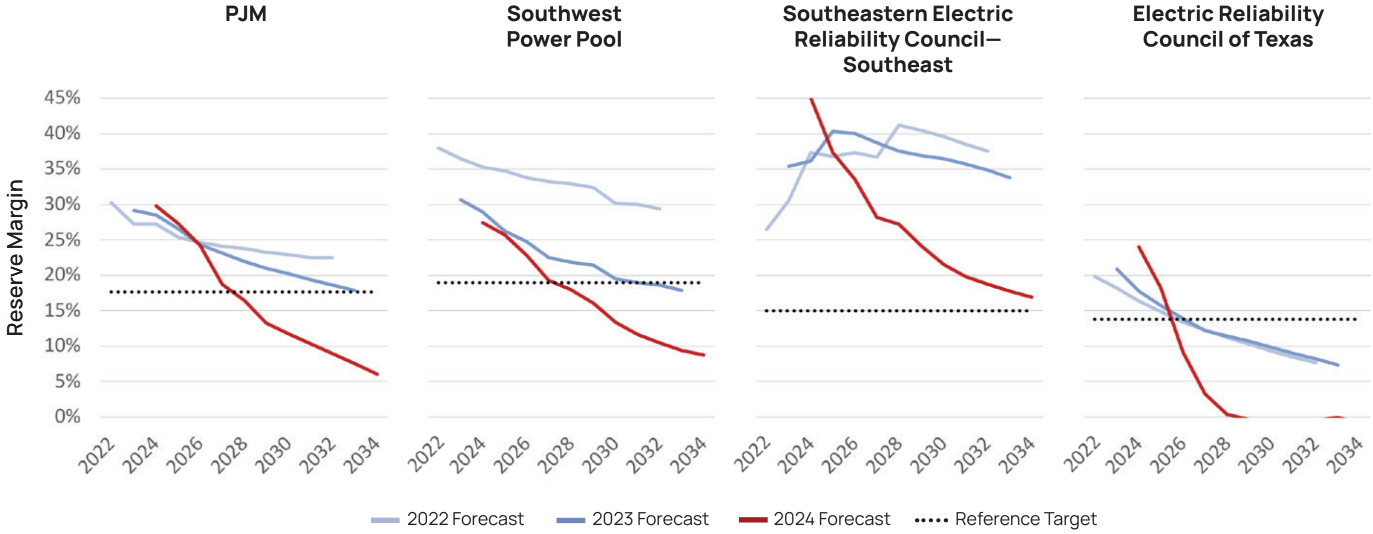
Recent market outcomes underscore the need for, and value of, flexibility to complement traditional capacity resources. Capacity clearing prices in PJM's 2027/2028 Base Residual Auction hit record highs for the third auction in a row (Howland, 2025b). The capital cost of new gas turbines has increased considerably (Besuner, 2025). Lead times for turbine delivery can extend up to seven years, constraining near-term capacity options. While longer-term resources come online, flexibility programs can be implemented more quickly, providing near-term adequacy relief.

**While longer-term resources come online, flexibility programs can be implemented more quickly, providing near-term adequacy relief.**

One key aspect that informs whether and how quickly these new data center loads can connect to the grid is whether the system has sufficient supply-side and demand-side resources to support the operation of these and all other loads under extreme or unexpected grid conditions. Even before today's large load–driven demand growth, resource adequacy planning was challenged by an evolving resource mix, including greater shares of variable energy resources, the growing frequency and magnitude of severe weather events, and rising end-use electrification. The ESIG Redefining Resource Adequacy Task Force has written about the needs and

**FIGURE 1**

**Tightening Reserve Margins in 2022, 2023, and 2024 Forecasts**



**Reserve margins for select regions, prior to resource additions, are tightening, as shown in this comparison of 2022, 2023, and 2024 forecasts for several system operators. Despite near-term increases compared to 2022 and 2023 forecasts in PJM, the Southeastern Electric Reliability Council (SERC)-Southeast, and the Electric Reliability Council of Texas (ERCOT), 2024 reserve margin forecasts in all regions quickly fell to levels substantially below historical projections, including falling below reserve margin targets within three years in PJM, SPP, and ERCOT.**

Note: Reserve margin represents "Existing-Certain and Net Firm Transfers Reserve Margin" prior to new resource additions.

Source: Energy Systems Integration Group; data from North American Electric Reliability Corporation, "Electricity Supply and Demand," accessed 2025, https://www.nerc.com/programs/reliability-assessment-performance-analysis/electricity-supply-demand.

opportunities for evolving resource adequacy planning approaches in this environment, including a growing recognition of the role of the demand side (ESIG, 2021).[2]

As the resource adequacy construct continues to evolve, it must further account for the unique characteristics of, and uncertainty surrounding, large loads.

## Diverse and Rapidly Changing Operating Characteristics

The probabilistic resource adequacy models that underlie modern resource adequacy analyses require detailed load data that reflect hour-by-hour changes in system demand and operational representations of the supply and demand resources that can be used to meet that demand. Because of the size of today's large loads, their operations can have a material impact on system load shapes. However, it can be challenging to effectively represent large loads' diverse and changing operating characteristics in resource adequacy models.

Data center demand depends heavily on the type of workload. For example, a data center's servers can be used for AI training, AI inference, cloud storage, or traditional enterprise computing. Yet these end uses are often not visible to host ISOs and utilities. While "data center" or "large load" are used to describe facilities that can host any of these workloads, representative hourly profiles for computing and cooling requirements for each workload type may be different. Furthermore, a high degree of uncertainty can exist around seasonal and daily operational modes, even within a given workload type. This poor visibility combined with high uncertainty makes effective resource adequacy planning challenging. For example, a data center's cooling needs may affect the degree to which its power consumption is impacted by temperature, while the degree to which high data center consumption coincides with the system's peak load will affect the likelihood that it will contribute to system stress periods. Planners today are challenged in adequately capturing these types of operational characteristics.

These loads' operating characteristics can also evolve. For example, a data center that hosts AI training workloads one year may host AI inference workloads the next. New hardware can also cause a data center's demand profile to change over time. For example, the graphical processing units (GPUs) used by AI data centers must be replaced or repurposed every few years, and newer units may have different efficiencies and operating characteristics that can affect a data center's demand profile.

Data are also lacking around operational capabilities that could underpin a data center's ability to provide flexibility, such as temporal workload shifting or switching to on-site generation or storage. While technological innovation and policy and regulatory developments are creating new opportunities for load flexibility, these responsive load capabilities remain largely untested at scale and require granular operational data to accurately model their impact on system adequacy.

## Forecasting Challenges

Despite their growing importance, large loads, especially data centers, remain exceptionally difficult to forecast in the medium and long term.[3] First, developers often pursue multiple sites simultaneously to maximize their success at getting interconnection at least at one grid point. They may submit interconnection requests across multiple utility footprints and RTOs/ISOs, all while maintaining confidentiality around their ultimate build-out decisions to preserve optionality and protect investment planning decisions from competitors. This means that load interconnection queues are commonly inflated with duplicative interconnection requests, even from the same developer. Many developers may not even have a data center tenant in mind but instead are anticipating a customer later in the development process. Such practices inject uncertainty into the load forecast development process and risk utilities setting capacity requirements based on inflated forecasts of future energy demand. In such a situation, ratepayers may be left to foot the bill for stranded infrastructure costs built to serve a new load that did not materialize.

---

[2] https://www.esig.energy/working-groups/system-ops-market-design/redefining-resource-adequacy/

[3] While the load forecasting considerations of large loads is discussed in detail in a companion ESIG report *Forecasting for Large Loads: Current Practices and Recommendations* (https://www.esig.energy/reports-briefs/forecasting-for-large-loads/), the present report considers aspects of large load forecasting on resource adequacy and adjustments to the load forecast to facilitate probabilistic resource adequacy assessment (ESIG, 2025a).

Second, new data centers can drive discrete, stepwise shifts in system demand and resource needs, adding tens, hundreds, or even thousands of megawatts of new demand at a time. While some large loads may energize at their full contracted interconnection capacity, many data center sites ramp up operations over several years as buildings, computing clusters, and cooling systems are commissioned in stages. However, these "ramp rates" may not be known to utilities, which may instead assume a facility's full contracted interconnection capacity for load forecasting and capacity planning purposes, potentially creating a disconnect between forecasted and actual demand at a data center site.

Finally, the rapid pace of technological innovation—including more efficient computational models, data center designs, and workload orchestration software—means that future energy consumption patterns may be different from today's. This results in a high degree of planning uncertainty, with wide margins between "high" and "low" load growth scenarios. While ISOs and utilities take a variety of approaches to account for these sources of uncertainty in forecasting and resource adequacy planning, including contract-based, probabilistic, and econometric approaches, no approach is perfect (Daymark Energy Advisors, 2025).

### Tight Supply Conditions

The increase in forecasted energy demand driven by large loads comes at a time of generation capacity constraints. Many aging coal, gas, nuclear, and oil plants have retired or are scheduled for retirement within the next decade. At the same time, gas turbines are in short supply, with lead times stretching several years due to global supply chain limitations (Cohen, Fitch, and Shwisberg, 2025). Meanwhile, the addition of renewable resources, including paired storage projects, faces headwinds due to backlogged interconnection queues and changes to federal tax incentives. Together, these factors limit both the ability and incentives for bringing new supply to market that can support resource adequacy.

## Increased Resource Adequacy Risks for Large Load Developers, Consumers, Utilities, and Markets

When rapid, hard-to-predict data center growth occurs in a constrained supply environment, the grid faces heightened resource adequacy risks. If not managed proactively, this imbalance threatens to impact reliability, increase ratepayer costs, and erode customer confidence. System operators and planners are hesitant to authorize interconnections until risk is appropriately managed. This has broad implications:

- **For large loads,** unpredictable or inadequate reliability can threaten server uptime guarantees (such as maintaining "five nines," or 99.999%, server availability), strand investments, and delay large load project timelines.
- **For consumers and ratepayers,** higher capacity procurement costs and infrastructure build-outs could translate to higher electricity rates, depending on how costs are recovered and allocated.

---

> **When rapid, hard-to-predict data center growth occurs in a constrained supply environment, the grid faces heightened resource adequacy risks. If not managed proactively, this imbalance threatens to impact reliability, increase ratepayer costs, and erode customer confidence.**

- **For ISOs, utilities, and markets,** failure to manage these risks could undermine public trust in institutions and processes at a time when they are already under scrutiny for rate increases. It has also introduced *political* decisions in ISO processes that are typically driven by state and federal regulatory bodies (NEDC, 2026). Undermining public trust could deter large load investments in some jurisdictions and drive them to others.

Underlying these challenges is a foundational tenet of electricity regulation: the utility's or ISO's obligation to serve. These entities are required to plan and operate their electric power system to provide universal access to reliable electricity to meet the needs of existing and new customers. The interpretation of this obligation varies by jurisdiction. For regulated utilities, for example, some state public utility commissions may prioritize serving existing customers, even if that requires limiting or placing additional conditions on new interconnections. Other commissions maintain that utilities must accommodate all new loads, regardless of their size or timing, provided technical requirements are met. In both cases, the outcome is the same, even if timelines for and the volume of large load interconnections look different: utilities and ISOs must find ways to integrate large loads while upholding system reliability.

## Solution: Evaluation of Resource Adequacy Impacts of Large Loads Up Front in Planning Processes

Resource adequacy analysis provides a quantitative foundation for decision-making around resource additions and retirements. It can also inform the design of demand-side flexibility and co-located generation and storage systems that extend beyond traditional supply-side generation and accelerate the power system's ability to interconnect and serve large loads reliably. However, these benefits can only be achieved if the system impacts of large loads, and the full set of demand- and supply-side solutions available to mitigate them, can be effectively evaluated before interconnection in resource adequacy modeling.

**The industry needs a consistent, transparent framework for assessing large load impacts up front in planning processes, including how the large load affects system adequacy, how load flexibility can be accredited, and how capacity obligations can be managed.**

Large load development's scale, speed, forecast uncertainty, and operational uncertainty strain current planning tools and processes used by utilities and ISOs. This results in inconsistent approaches across the country and limits the ability of large load developers to accurately anticipate reliability requirements and interconnection timelines. The industry needs a consistent, transparent framework for assessing large load impacts up front in planning processes, including how the large load affects system adequacy, how load flexibility can be accredited, and how capacity obligations can be managed.

Drawing from the broad and diverse expertise of the ESIG project team participants, this report offers recommendations on how resource adequacy analyses can reflect large loads and how these analyses can be used to inform load flexibility and other mitigation options. Several short-lead-time solutions that utilities and ISOs are implementing today are introduced for mitigating near-term, large load-driven resource adequacy risk. The need to integrate these options more deliberately into long-term planning is highlighted. We then describe a standardized playbook for integrating large loads into resource adequacy assessments to help utilities and ISOs make more efficient, data-driven decisions about capacity needs. A clear framework is provided to accelerate interconnection of large loads and ensure reliable electricity for all types of loads.

# Approaches to Mitigate Near-Term Resource Adequacy Risks

Power system planners can implement short-lead-time solutions to mitigate or address near-term adequacy risks. The magnitude of new large loads in load growth forecasts and the volume of near-term large load interconnection requests has driven utilities and ISOs to pursue both new and traditional short lead-time strategies for managing this growth while maintaining near-term resource adequacy. Broadly, these solutions fall into three categories: (1) interconnection moratoriums, (2) load curtailment programs, and (3) bring-your-own-(new)-generation (BYOG or BYONG) arrangements.

Today, these solutions are often embedded exclusively within a utility's or ISO's interconnection process rather than within long-term planning. In the interconnection context, these solutions are evaluated based on their ability to provide immediate reliability protection through targeted modifications to planning, operations, and interconnection processes, without requiring a redesign of utility planning frameworks or market structures. However, these solutions need to also be incorporated into long-term resource planning frameworks and complemented by improvements in how we measure, model, and assess the resource adequacy impacts of large loads.

## Near-Term Solutions

### An Interconnection Moratorium

If system conditions are too constrained to reliably interconnect new large loads, a moratorium on new large load interconnections is a potential short-lead-time solution for ensuring continued system adequacy. In such cases, utilities, ISOs, or states might limit or pause the interconnection of new large loads to preserve reliability for existing customers. For example, in March 2023 American Electric Power issued a moratorium on new service requests from data center customers in central Ohio to give its transmission planning group time to study the impact of pending large load service requests (PUC of Ohio, 2025).

While protecting reliability, interconnection moratoriums raise difficult legal and regulatory questions related to the "obligation to serve." Moratoriums could also undermine economic development goals, which includes advancing domestic investment in emerging industries like AI and advanced manufacturing. But the alternative—a load being added to the system without sufficient resources to supply it—is even more worrisome. While utilities and ISOs may have an obligation to serve any load, they also have an obligation to meet reliability needs of customers, both existing and new. For example, in a complaint to the Federal Energy Regulatory Commission (FERC), PJM's independent market monitor argued that PJM should not be allowed to interconnect new large data center loads unless they can be served reliably in terms of both transmission and capacity adequacy, noting the "basic responsibility of PJM to maintain a reliable grid" (Howland, 2025a).

### Load-Curtailment Programs

Many grid operators are instead designing curtailment-based reliability programs. These allow new large loads to connect on the condition that they may be curtailed during system emergencies or periods of high risk to preserve resource adequacy. Though largely an operational tool, this solution maintains resource adequacy requirements in the near term. Load curtailments may be voluntary or involuntary, and provide accelerated interconnection in exchange. Table 1 (p. 8) gives details on the Southwest Power Pool's (SPP's) Conditional High-Impact Large Load Service (CHILLS) and frameworks emerging under PJM's Critical Issue Fast Path Proposal, as well as provisions laid out in Texas Senate Bill 6 (SB 6) that affect the Electric Reliability Council of Texas's (ERCOT's) load curtailment requirements. These programs take different approaches but share a key feature: they authorize grid operators to curtail large load demand ahead of rolling blackouts to protect existing customer classes, regardless of the large load's price sensitivity and willingness to pay for electricity. In addition to programs differing in whether they are voluntary or involuntary, programs also vary with respect to incentives, such as accelerated interconnection or reduced capacity obligations, and opt-out mechanisms.

Curtailment-based programs can offer benefits for the power system and participating loads, including potentially accelerating interconnection of new large loads. These programs ensure that new large loads do not degrade reliability for existing customers, and provide operators with tools to maintain resource adequacy.

TABLE 1

## Overview of Proposed and Emerging ISO Curtailment-Based Reliability Programs

| | Southwest Power Pool Conditional High-Impact Large Load Service (CHILLS) | PJM January 2026 Critical Issue Fast Path Proposal | Texas SB 6 (2025) |
|---|---|---|---|
| **Which large loads qualify** | Program is available to "high-impact large loads" (HILLs), which are commercial or industrial loads with peak demand of greater than or equal to 10 MW if 69 kV or below, or peak demand of greater than or equal to 50 MW if greater than 69 kV. | The PJM board defines large loads as individual load additions at or above 50 MW at a single point of interconnection. | SB 6 applies to load with peak demand of 75 MW or greater, though the Public Utility Commission of Texas may elect to set a lower threshold. |
| **Role for curtailment** | In exchange for interruptible transmission service until upgrades are in place, CHILLS loads receive faster interconnection. | "Connect and manage" loads that do not participate under the voluntary "bring your own new generation" (BYONG) framework may be curtailed by PJM for a "limited number of hours during the year." | SB 6 requires mandatory curtailment or deployment of back-up generation for large loads with an on-site back-up generator. |
| **When curtailment may occur** | These loads are subject to curtailment during emergency or other unforeseen conditions that threaten system reliability. | Curtailment will occur prior to the deployment of pre-emergency demand response. Curtailment events are expected to occur "infrequently, for limited durations, and only when necessary to prevent broader system impacts." | Large loads may be required to deploy back-up generation or curtail load with reasonable notice before or during an emergency alert, based on a threshold to be developed by ERCOT. |
| **Implications for interconnection** | CHILLS is a temporary solution that provides for accelerated interconnection, with participants expected to transition to firm service within five years. | Generation sponsored by large loads under BYONG and with a contractual commitment can qualify for a new expedited interconnection track. | SB 6 does not impact interconnection timelines. |
| **Implications for markets and resource adequacy** | CHILLS loads are considered "non-market registered demand response for purposes of [resource adequacy] requirements." | The PJM board does not propose that "connect and manage" loads be removed from the capacity market. | ERCOT has no forward capacity market or capacity requirements; SB 6 uses administrative curtailment authority and demand response procurement rather than capacity obligations. |

Source: Energy Systems Integration Group, summarizing information from the Southwest Power Pool's Conditional High-Impact Large Load Service (CHILLS) (https://spp.org/markets-operations/high-impact-large-load-hill-integration/), frameworks emerging under PJM's Critical Issue Fast Path Proposal (https://insidelines.pjm.com/pjm-board-fast-tracks-effort-to-reliably-serve-large-loads/), and provisions laid out in Texas Senate Bill 6 (SB 6) that affect the Electric Reliability Council of Texas's load curtailment requirements (https://capitol.texas.gov/BillLookup/History.aspx?LegSess=89R&Bill=SB6).

However, proposed curtailment approaches in most regions are not informed by resource adequacy analysis to determine the likely amount of curtailment needed. This places uncapped curtailment risk onto large load customers, without annual or daily curtailment limits. As more loads connect in an area under curtailment-based agreements, the cumulative risk of curtailment for these loads may grow, creating uncertainty that can undermine the business case for new large load development.

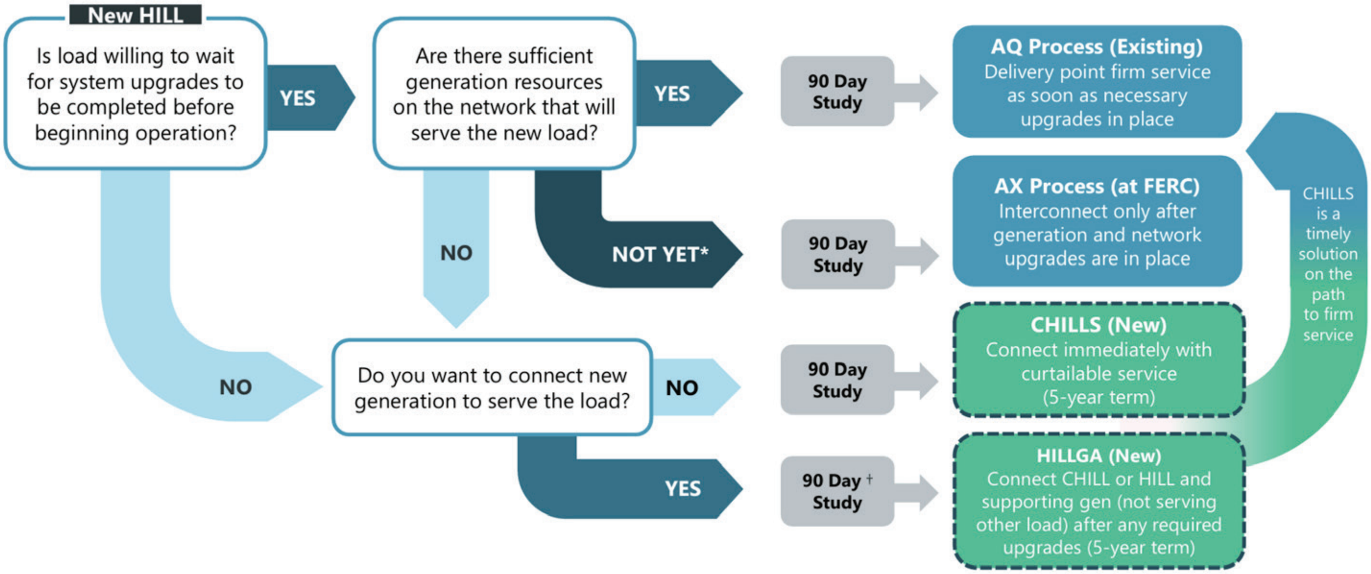
While interconnection moratoriums would restrict development by decree, uncapped curtailment risk could restrict development indirectly. If not thoughtfully designed, curtailment-based programs could distort resource adequacy signals or reduce transparency in capacity markets. For example, curtailment-based programs could reduce utility capacity obligations by lowering the amount of firm capacity required to serve curtailable large loads. This could lead to reduced incentives for new capacity to enter the market, such as through a lower demand curve driving reduced capacity prices. While this outcome may be desirable if large loads participate in curtailment-based reliability programs indefinitely, it could slow down the procurement of additional generation and transmission that may be needed to support uninterrupted service in the long run.

## Bring-Your-Own-New-Generation Arrangements

A third class of short-lead-time solutions for near-term resource adequacy is bring-your-own-new generation

FIGURE 2

**Proposed Paths for the Interconnection of Load and Supporting Generation in the Southwest Power Pool, as of July 2025**



Load may pursue multiple paths (AQ, AX, CHILL) simultaneously.

* "**Not Yet**:" The utility has pending generation with rights (GIA), or planned generation.

† HILLGA for "Common Bus" and "Local Area" to be completed in 90 days. HILLGA for lager "Deliverability Zone" requires additional study.

**SPP has proposed new pathways for accelerated interconnection of "high-impact large loads" (HILLs), which include a specific expedited interconnection pathway for HILLs with new contracted generation under a bring-your-own-new-generation arrangement. This pathway, called the High-Impact Large Load Generation Agreement (HILLGA) process, is shown on the bottom-right corner of the flowchart. HILLGA is relevant for HILLs, which do not want to wait for system upgrades before beginning operations and which can bring their own new generation, as shown by following the flowchart from the upper-left-hand corner.**

Notes: CHILLS = Conditional High-Impact Large Load Service; FERC = Federal Energy Regulatory Commission; GIA = generator interconnection agreement; HILL = high-impact large loadH; HILLGA = High-Impact Large Load Generation Agreement.

Source: Southwest Power Pool, "Large Load Stakeholder Engagement Forum," July 1, 2025 (virtual meeting), https://spp.org/spp-documents-filings/?id=540500. "Large Load Stakeholder Engagement Forum Meeting Materials 20250701," Item 01.

(BYONG). These arrangements allow large loads to interconnect more quickly if they offset their grid impact through self-supplied generation capacity. In exchange for reduced capacity obligations or accelerated interconnection, large load customers commit to relying on on-site or contracted resources. These programs differ in design: some require generation to be located at or adjacent to the point of interconnection, while others permit contractual ownership in other parts of the system; some allow contracts with existing generation, while others require new incremental capacity. For example, SPP proposed a HILLGA process that allows large loads and supporting generation to be studied together and enables expedited interconnection, as laid out in Figure 2 (p. 9). The HILLGA process was approved by FERC in January 2026 (SPP, 2026a).

The BYONG approach provides large loads with autonomy and flexibility, allowing them to manage their reliability needs independently of grid constraints. However, this approach can also lead to suboptimal system planning if individual large loads procure resources that do not align with broader system needs. For example, a set of self-supply resources optimized for a single load's peak demand may not contribute meaningfully during regional capacity shortfalls and may not provide grid services where they're most needed. Utility and ISO accreditation rules for BYONG resources also introduce risk. If accreditation of self-supply resources changes over time due to broader changes occurring across the power system, a load would need to continually add new generating resources to fulfill its capacity obligation. Finally, siting constraints for co-located generation may push large loads to locate in remote areas, and the added distance could increase latency or operational inefficiency.

BYONG can also increase resource adequacy risk if existing generation leaves the market to contract directly with large loads. Under this scenario, all or a portion of the generation capacity may not be made available again to the market, depending on service agreements as a BYONG resource. FERC's order on large loads co-located with generating facilities in PJM addresses this issue for ISO regions by directing PJM to prohibit existing capacity from leaving the grid to serve co-located load until all required network upgrades are in place and by requiring the existing generator to pay their full cost (FERC, 2025a). While these requirements can help mitigate network reliability

risk and incentivize existing generators to stay in the market, generators could still choose to leave the market, reducing reserve margins and increasing resource adequacy risk. Opportunities for new generation to contract directly with large loads as BYONG resources can similarly risk exacerbating resource adequacy challenges by reducing the amount of new generation participating in electricity markets.[4] Without careful planning, BYONG could exacerbate such procurement challenges in a tight-supply environment.

## Lessons from Near-Term Solutions to Support Long-Term System Planning and Capture the Value of Load Flexibility

Each of these options—interconnection moratoriums, curtailment, and BYONG—can play a role in addressing near-term resource adequacy concerns. But none provide a comprehensive regulatory framework, provide proactive system planning rooted in least-cost procurement or market efficiency, or are fully incorporated in regional resource adequacy frameworks.

Further, these options do not fully recognize the potential value of large load flexibility, which may include not just curtailment or on-site generation, but also load shifting, geographical redistribution, and/or coordination with system needs over seasonal and daily peaks. The differences between these reactive, near-term resource adequacy options and proactive, long-term, purpose-built flexibility programs are four-fold. The near-term options:

- Are not integrated with the long-term generation capacity planning process, solely meeting individual large load needs rather than simultaneously solving both large load and system-wide capacity needs
- Are one-size-fits-all approaches to curtailment and BYONG, whereas tailored flexibility programs can be designed for a wider spectrum of options in the size, duration, and timing of flexibility needs and the associated range of benefits
- Do not follow an economic signal, where well-designed flexibility programs can follow both short- and long-term price signals
- Are designed, in most cases, to be reactive, short-term solutions to incorporate the large loads as soon as possible until transmission and generation additions can “catch up” but are not necessarily intended to be permanent solutions

> **While near-term solutions might be viewed as stopgap measures, they provide an essential real-world “testing ground” for longer-term and more sophisticated, resource adequacy analysis-informed flexibility programs. Rather than purely temporary, near-term programs are a foundational starting point.**

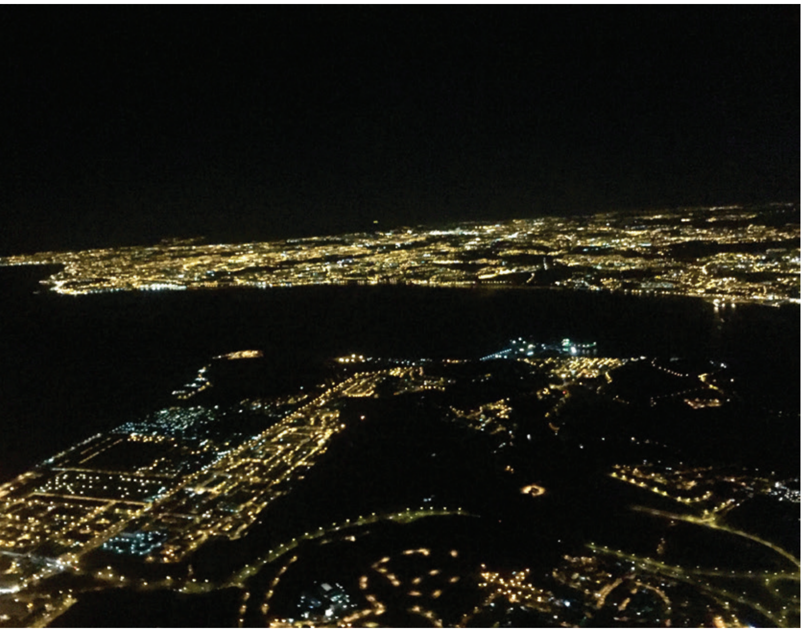
While near-term solutions might be viewed as stopgap measures, they provide an essential real-world “testing ground” for longer-term and more sophisticated resource adequacy analysis-informed flexibility programs. These early implementations can offer critical data on large load performance and the reliability of on-site resources during actual grid stress events. Rather than seeing these near-term programs as purely temporary, they should be viewed as a foundational starting point.

Ultimately, the goal of utility regulators and system planners is not to rely indefinitely on short-term solutions, but rather to use them as interim measures while developing proactive solutions that are integrated with long-term planning. Longer-term solutions can plan for, procure, and compensate large load flexibility, like any other reliability resource. To do so effectively will require adapting resource adequacy modeling methods and data inputs to support the integration of large loads.

---

[4] This would exacerbate existing capacity constraints. In its 2027/2028 Base Residual Auction, for example, PJM fell short of its 20% installed reserve margin procurement target.

# Methods and Data Needs for Effective Modeling of Large Loads



Accurately representing large loads within utility planning models requires new levels of data granularity and methodological rigor. To account for the high uncertainty inherent in today's large load forecasts, planners can model large loads as separate load components. Detailed, segment-specific data—covering load size, location, hourly operating characteristics, on-site generation specifications, and realization probabilities—can be collected and integrated into modeling workflows. This can help planners isolate the effects of large loads and more precisely model their behavior and system impacts in both probabilistic and deterministic (i.e., scenario-based) analyses.

## End-Use Load Layers

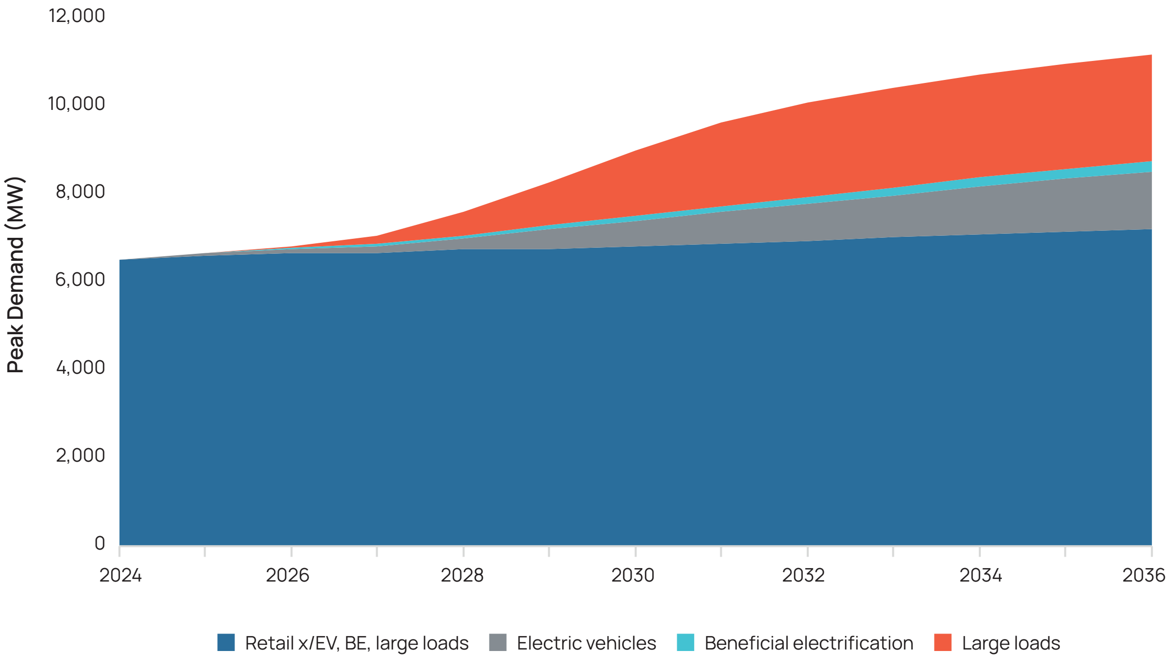
Traditional load profiles in resource adequacy models are insufficient for large loads. Load profiles used as inputs to capacity expansion and resource adequacy models traditionally represent zonal system demand as a single aggregate hourly load shape. This load profile is either aggregated at the regional level or split across planning zones, and is often composed of forecasts of each customer class (residential, commercial, and industrial). While these forecasts may be developed using methodologies that account for specific end uses, the final load profile used in probabilistic resource adequacy models often does not include this specific end-use segmentation. While this traditional approach has been sufficient for capturing gradual, system-wide growth, it is no longer adequate in an era of large, concentrated, and often uncertain load additions. Treating all demand growth as uniform fails to capture the operational characteristics of large loads with distinct chronological hourly load profiles, seasonality, and interconnection timelines.

In many cases resource adequacy planners also apply temperature-load relationships across the entire aggregate load profile to create multiple weather years of load data necessary for probabilistic resource adequacy analysis. But this overstates the weather dependence of the future load, which may contain a large amount of data center and other industrial demand that is less influenced by temperature. The lack of end-use segmentation in load profiles also limits planners' ability to isolate the specific impacts of large loads and evaluate targeted solutions, such as flexibility.

Instead, planners can develop load profiles split into distinct components, explicitly separating large loads from the baseline forecast (Figure 3, p. 13). Creating a "large load layer" allows for scenario development and proper representation in resource adequacy models and provides greater transparency to planners, regulators, and stakeholders by identifying the portion of future demand growth attributable to specific industrial or commercial large load developments. The large load layer

FIGURE 3

## Peak Demand Load Forecast Separated into Components



**This peak demand load forecast from the Public Service Company of Colorado (PSCo) (also known as Xcel Energy) includes load layers for beneficial electrification (BE), electric vehicles (EV), and large loads. Load forecasts need to be separated into key load forecast components, including a large load layer to allow for scenario development and proper representation in resource adequacy models.**

Source: Public Service Company of Colorado, "2024 Just Transition Solicitation, Section 2.2 Xcel Energy, Load Forecast 2024–2036." https://xcelnew.my.salesforce.com/sfc/p/#1U0000011ttV/a/8b000003N3eQ/QodCDDNR0HUF75Rb_nOkwu_MzfeKzPduYws13FjugGs.

can be further segmented by specific end uses like data centers (potentially further segmenting by data center type), oil and gas extraction, advanced manufacturing, and vehicle fleet electrification. This approach also allows large loads to be excluded or adjusted in sensitivity scenarios, enabling targeted analyses of how their presence or absence affects system capacity needs and risk exposure.

## Large Load Scenario Development

The inherent uncertainty associated with large load development means that no forecast will be perfect. To mitigate uncertainty, planners can evaluate a range of future demand scenarios to reflect the full spectrum of potential outcomes. These scenarios would include not just business-as-usual and high-growth projections, but also cases that assume a reduction or delay in data center load growth. Evaluating the full range of outcomes helps planners identify low-regrets strategies of investments that remain valuable even if large load growth materializes differently from what is expected, while being prepared to scale up investments if demand follows a higher trajectory.

While developing scenarios is typically reserved for long-term resource planning (i.e., integrated resource plans), it is important that the underlying probabilistic resource adequacy analysis also includes similar large load scenarios. This is because the changing load profile will inherently change the size, frequency, duration, and timing of resource adequacy risk. These scenarios will also inform whether changes are required for total capacity obligations in the PRM requirement or in resource accreditation.

## Large Load Data Needs

Effective modeling of large loads in probabilistic resource adequacy analysis requires that resource adequacy workflows integrate new kinds of data. Broadly, these data needs fall into six categories, detailed in Table 2 and described in more detail below.

## Size

The capacity of large loads (in MW) is one of the most direct indicators of their system impact. Planners need to collect information on the total non-coincident peak demand and on the ramp period (how quickly a load is expected to scale from initial interconnection to full operation, measured in MW per year). Large load size also drives the modeling need to treat these loads discretely as individual data center campuses (if over a certain MW threshold), rather than folding them into aggregated demand growth forecasts. Clustering or aggregation methods may be used to reduce computational burden while preserving essential load characteristics where multiple loads are expected within a single planning region.

## Location

Locational granularity is also important. While some resource adequacy studies operate under a “copper sheet” assumption (treating the system as a single, unconstrained electrical region), most utilities and system operators employ zonal or transfer-limited models that reflect underlying transmission constraints and regional capacity deliverability. Therefore, large load forecasts need, at a minimum, to specify the zones or subregions in which new loads are expected to materialize. This zonal resolution helps planners evaluate how large load additions will affect local capacity requirements, where new resources may be needed, and how transmission limitations could influence capacity requirements. Although transmission planners may need precise nodal locations for transmission interconnection studies and production cost simulations, resource adequacy analyses can typically operate effectively at a broader zonal level. As a result, the required locational data in resource

TABLE 2

### Data Needs for Modeling Large Loads in Resource Adequacy Analysis

| Data Need | Essential | Preferred |
|---|---|---|
| **Size** | • Nameplate capacity (specified in interconnection agreement) | • Annual ramp rate |
| **Location** | • Zonal assignments | • Nodal location |
| **Hourly operating characteristics** | • 8,760 chronological hourly load profiles with *aggregated* large load layer<br>• Capacity utilization rate<br>• Load factor | • 8,760 chronological hourly load profiles with disaggregated *individual* large load profiles |
| **Weather dependence** | • Temperature dependence of load reflected in hourly load profiles | • Multiple weather years of load data |
| **Load and resource configuration specifications** | • Asset type<br>• On-site generator size<br>• Run-hour or emissions limitations<br>• Availability duration [for battery energy storage or fuel-limited resources] | • Operational costs<br>• Detailed operational characteristics<br>• Forced outage rates |
| **Load forecast range** | • Load growth forecasts<br>• Probabilities of realization for individual projects<br>• Large load layers by end use | • Stochastic modeling of load-growth outcomes<br>• Reference load-growth forecast that excludes large load growth |

Source: Energy Systems Integration Group.

adequacy assessment is less granular than in other planning processes. Providing this degree of locational and temporal detail ensures that large load forecasts can meaningfully inform both system-wide and regional reliability planning.

## Hourly Operating Characteristics

A critical factor for resource adequacy is the load profile, particularly a load's coincidence with system peak. To capture the operational dynamics of large loads, planners can collect hourly (8,760) chronological demand profiles, measured in MW, segmented by end use, and ideally reflective of weather dependencies. For data centers, for example, cooling loads may show strong temperature correlations, while industrial or hydrogen loads may operate more uniformly. This level of detail allows resource adequacy models to evaluate how large loads contribute to or relieve stress during peak and scarcity periods, and is required to accurately compute probabilistic resource adequacy criteria and metrics such as ELCC.

## Weather Dependence

Where relevant, the 8,760 profiles need to reflect weather sensitivity in load behavior, particularly for end uses with significant cooling requirements, such as data centers or advanced manufacturing. Higher ambient temperatures often increase cooling loads precisely when system margins are tightest, compounding adequacy risks. Incorporating this dependence enables planners to capture correlated risks between high load and low resource availability that together create periods of grid stress. When specific load layers for large loads are incorporated, this allows the application of large load–specific temperature dependence responses to that portion of the load. Modeling temperature response generically—that is, identical to the typical weather-dependent response of a customer class—can risk over- or under-stating a large load's unique temperature dependence. Ideally, multiple weather years of hourly load data will be available to inform weather dependence across a wide range of weather conditions. These data should also be coincident and consistent with other load, renewables, and outage data.

It is important to note that capturing the weather dependence of large loads, particularly for data center demand, requires improvements to modeling techniques and requires multiple years of historical, empirical data for utility and ISO load forecasters and researchers. These improvements in data availability and modeling techniques should be viewed as an emerging need and an area of future research. At a minimum, however, it is important that resource adequacy planners do not apply the same historical temperature-load relationships observed in the past to reflect the aggregate demand of the system in the future.

## Load and Resource Configuration Specifications

Many large load facilities, particularly data centers and industrial campuses, include on-site generation or storage systems that can influence system adequacy. Modeling data center load as the net of on-site generation should be avoided, as there will be significant differences between the risk profile and probabilistic characterization of the underlying data center load and those of the on-site generation. As discussed above, off-site capacity can also be contracted by large loads, whether centralized generation and storage resources or local distributed energy resources. To represent these assets accurately, planners need detailed information on generator size (MW), operational and cost characteristics, and run-hour or emissions limitations. This allows them to quantify the extent to which on-site generation can be relied upon as a resource that can support large load flexibility and to assess its potential capacity accreditation. Collecting these data may require formalized or mandatory data requests of large load operators.

## Load Forecast Range

Forecasting large load growth requires accounting for significant uncertainty in realization (the likelihood that proposed projects will materialize and when and where they will come online). Essential data include probabilities of realization for individual projects, regional or utility-specific load growth forecasts, and, where possible, end-use breakdowns that distinguish between data centers, industrial electrification, hydrogen production, and other load types.

Methodologically, planners can apply a “bookend” forecasting approach, developing both high- and low-growth scenarios to capture the plausible range of large load additions. Probabilistic sampling techniques can further refine these scenarios by weighting projects based on realization probabilities, enabling stochastic modeling of load-growth outcomes. This approach can be anchored by a reference forecast that excludes large load additions entirely to serve as the baseline for isolating their incremental effects on capacity requirements and system costs. These approaches are novel to resource adequacy planning and are discussed further in the ESIG Large Loads Task Force report *Forecasting for Large Loads: Current Practices and Recommendations*.[5]

Collectively, these data and modeling improvements enable planners to move from generic load representations toward a segmented, probabilistic approach that captures both the uncertainty and operational diversity of large load growth. This foundation can allow planners to begin incorporating large load flexibility and responsiveness into long-term planning, a preferred pathway for providing system value in terms of lowered costs, enhanced resource adequacy, and reduced emissions.

---

[5] https://www.esig.energy/reports-briefs/forecasting-for-large-loads

# Preferred Path

## Longer-Term Solutions and Purpose-Built Flexibility Programs Planned Up Front in System Planning Processes

While data centers and other large loads can strain existing grid resources, they also possess unique capabilities that, if planned for intentionally, can enhance resource adequacy. This can occur by incentivizing (or requiring) additional capacity on the system but also by reducing the corresponding load during critical periods. A promising approach to managing these challenges is to develop purpose-built flexibility programs for large loads designed and integrated from the outset into long-term planning frameworks.

### The Role of Large Load Flexibility in Resource Adequacy

Large load flexibility refers to the ability of a large load to voluntarily decrease or increase its power consumption from the grid—either through changes in consumption or through the use of on-site generation or storage or nearby contracted off-site generation or storage capacity—with little to no impacts to its normal operation. Flexibility can be in response to market signals, rate structure, or direction from a utility or regioal grid operator. In other words, load reductions implemented regardless of the load's operational impact can be considered a load curtailment, while load reductions coordinated to limit disruption and operational impact can be considered load flexibility.

Flexibility can be deployed over a range of time frames, from seconds to hours to days, and can support both operational needs (e.g., balancing, frequency regulation, and contingency response) and planning objectives (e.g., reducing long-term capacity requirements).[6] Large load flexibility can improve resource adequacy and unlock faster, more cost-effective investment pathways.

**If large load flexibility can be realized at a scale to balance load growth with the pace of new generation interconnection, the question of how to integrate these loads while maintaining resource adequacy fundamentally changes.**

Data centers in particular have latent flexibility potential through workload management, thermal control strategies, and on-site generation and storage. Purpose-built programs that recognize and compensate for this flexibility can provide meaningful reliability benefits or minimize reliability impacts while enabling faster, lower-cost integration of new load.

If large load flexibility can be realized at a scale to balance load growth with the pace of new generation interconnection, the question of how to integrate these loads while maintaining resource adequacy fundamentally changes. Traditional utility planning has long grappled with the peak demand dilemma, the reality that a handful of extreme hours on the hottest or coldest days of the year determines the entire system's capacity requirement. Planners have historically needed to procure enough firm capacity to meet these rare peaks while fully aware that much of that infrastructure would remain underutilized for the vast majority of the year (Norris et al., 2025).

The load duration curve, which arranges hourly demand from highest to lowest, illustrates the challenge of the peak demand dilemma (Figure 4). While simplified, a

---

[6] The EPRI DCFlex project has drafted a standard with five tiers of load flexibility services that span both operational and planning objectives (as discussed in the presentation "The StarFLEX™ standard," EPRI, by Anuja Ratnayake to the ESIG Large Loads Workshop, October 30, 2025).

demonstrative current-day system's load duration curve (the left pane) shows that a system's capacity requirements are driven by only a few high-demand hours (150 GW in this example, plus a PRM), leaving substantial surplus generation and transmission capacity idle during non-peak conditions. If large loads are added to the system without flexibility (middle pane), the capacity requirement increases based on how much this load contributes to the coincident high-demand hours. In this example, that number is 40 GW. This yields a large increase in the capacity requirement and maintains the same amount of surplus capacity (in MW) during the remaining hours of the year.

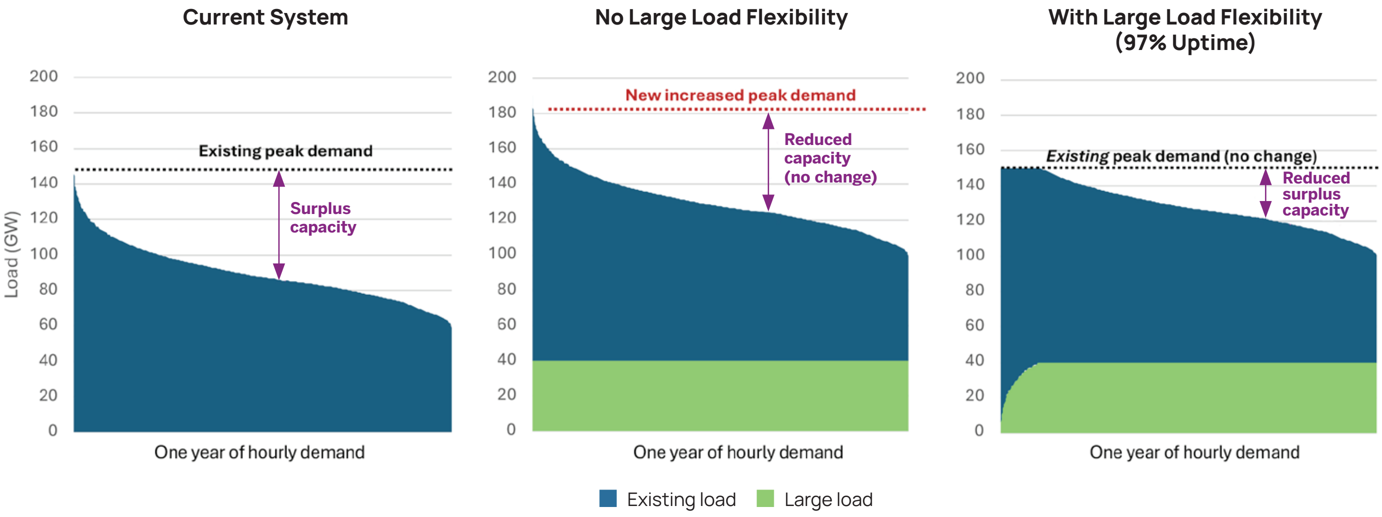
This traditional narrative can be inverted by large loads with flexibility (Figure 4, right pane). If these loads can temporarily reduce demand, shift workloads, or rely on on-site generation for just a few critical hours each year, they can help flatten the system's peak and increase the overall utilization of existing generation capacity. In the example in Figure 4, with a 3% annual load reduction, the total capacity obligation remains unchanged from the existing system, despite 40 GW of nominal demand being added across the majority of hours in the year. This approach adds a large amount of demand without new capacity obligations and reduces the amount of surplus generation that is required solely for rare events, improving the efficiency of the system.

> **If large loads can temporarily reduce demand, shift workloads, or rely on on-site generation for just a few critical hours each year, they can help flatten the system's peak and increase the overall utilization of existing generation capacity.**

At the same time, flexibility-backed load growth throughout the rest of the year from data centers and industrial electrification can stimulate new generation investment. This investment can occur either through direct procurement via power purchase agreements or indirectly by increasing wholesale energy prices that generators receive to remain in the market or incentivize new development. Under some cost allocation arrangements, flexibility-backed load growth may improve asset utilization and reduce average system costs.

FIGURE 4

## Load Duration Curves



**Illustrations of a load duration curve and capacity obligations without large loads (left), with large loads but no flexibility (middle), and with flexible large loads (right).**

Source: Energy Systems Integration Group.

> **When designed and operated with flexibility in mind, large loads can become an integral part of a balanced and efficient resource adequacy framework.**

In this way, large loads are not necessarily a threat to resource adequacy. Instead, when designed and operated with flexibility in mind, they can become an integral part of a balanced and efficient resource adequacy framework. Integrating large load flexibility in resource adequacy and resource planning can simultaneously enhance reliability, improve capacity utilization, and reduce system costs, transforming a planning challenge into an opportunity to strengthen the modern grid.

## Large Load Flexibility Options

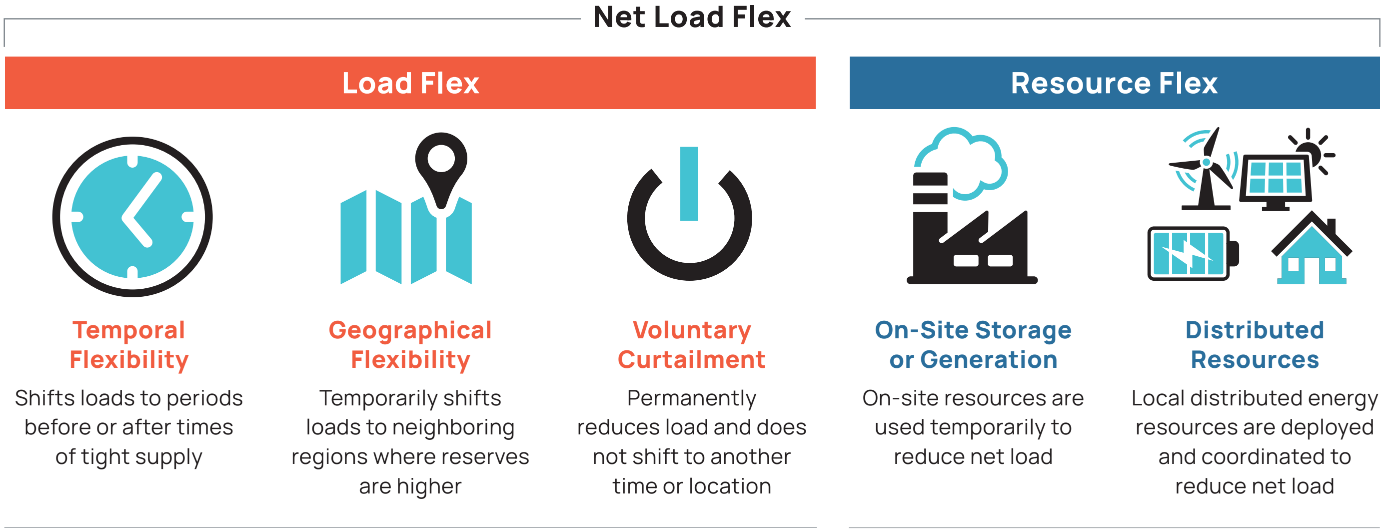
As shown in Figure 5, large load flexibility can take several distinct forms, each offering different mechanisms for reducing strain on the power system and supporting resource adequacy. Broadly, these flexibility options fall into five categories: temporal demand flexibility, geographical demand flexibility, voluntary curtailment, on-site generation and storage, and contracts with new, nearby available distributed energy resources or local capacity resources. Notably, none of these options actually call for involuntary load curtailments, but rather shift large load demand to other regions or times, or serve the load from on-site or nearby resources. And these options are not mutually exclusive. A large load can combine multiple approaches to get an aggregate, net-load response of the large load, generation, storage, and distributed assets and can be achieved independently or through a third-party service level agreement.

### Temporal Flexibility

Large loads can reduce or reschedule demand over short time periods to help meet peak or emergency conditions. For example, compute workloads such as AI training, non-urgent batch processing, or crypto operations can be paused or shifted to off-peak periods.[7] EPRI's DCFlex–Salt River Project pilot demonstrated that a data center could achieve a 25% reduction in compute-cluster power usage over three hours during a simulated peak-demand event,

FIGURE 5

Large Load Flexibility Options



**Net Load Flex**

**Load Flex**

**Temporal Flexibility**
Shifts loads to periods before or after times of tight supply

**Geographical Flexibility**
Temporarily shifts loads to neighboring regions where reserves are higher

**Voluntary Curtailment**
Permanently reduces load and does not shift to another time or location

**Resource Flex**

**On-Site Storage or Generation**
On-site resources are used temporarily to reduce net load

**Distributed Resources**
Local distributed energy resources are deployed and coordinated to reduce net load

**Five options for large load flexibility include temporal flexibility, geographical flexibility, voluntary curtailment, use of on-site storage or generation, and use of new, nearby available distributed energy resources or local capacity resources.**

Source: Energy Systems Integration Group.

---

[7] See the ESIG Large Loads Task Force report *Large Loads: Behaviors, Capabilities, and Limitations* (https://www.esig.energy/reports-briefs/large-loads-behaviors-capabilities).

validating the feasibility of short-term load response without significant business disruption (Warren, 2025). This approach is most analogous to conventional demand response programs and applies to large loads generally, as well as data centers specifically.

### Geographical Flexibility

Multi-site data center operators with distributed workloads can sometimes shift computing tasks between regions in response to grid conditions.[8] While such actions may increase latency or operating cost, they can provide system-level flexibility by reducing workloads in grid-strained regions. The pilot program underway between Google and Duke Energy as part of EPRI's DCFlex initiative combines both geographical and temporal workload shifting, exploring how coordinated demand management across facilities can enhance regional reliability (Ratnayake, 2025). This approach is most applicable to data center demand and may not be applicable to other types of large loads that cannot shift workload or output to other regions.

### Voluntary Curtailment

Unlike temporal and geographical flexibility, an additional load flexibility option voluntarily curtails load without the expectation of shifting the demand to another time period or region. In this option, load is simply reduced without a "payback," or corresponding increase. While this may not be an option for all large loads, voluntary curtailment is highly applicable to certain use cases, such as for price-responsive large loads like crypto mining, for which voluntary or market-based curtailment is more economical than continued operation during periods of high power prices.

### On-Site Generation and Storage

Many large facilities can operate independently of the grid during supply-constrained periods using on-site generators or battery energy storage systems. These can be in the form of large, utility-scale resources at the same (or nearby) location or as smaller distributed resources within the large load. Data centers are already designed with the capability to maintain uptime requirements in case of grid disruptions or disturbances, though the diesel-fired or other fossil-based back-up generation commonly used is often constrained for emergency use only due to air quality regulations. However, the role of on-site resources can be extended beyond emergency-only back-up capabilities to providing voluntary resource adequacy support with updated data center designs and upfront planning. For example, a study by the National Laboratory of the Rockies, in collaboration with Verrus, simulated a 70 MW grid-interactive data center using operational data, and showed that battery storage–driven demand response could provide substantial capacity relief without compromising service continuity (Vaidhynathan et al., 2025).

### Contracted Distributed Energy Resources

A fifth and emerging source of flexibility for large loads comes from partnerships with contracted local resources, including distributed energy resources located in the capacity zone as well as supply-side capacity resources. For example, large load customers are increasingly considering collaborating with nearby residential and commercial customers to deploy behind-the-meter batteries that can be used to help meet local and system-level adequacy needs (Giacobone, 2025). In this model, a large load may co-develop or finance batteries at customer sites, allowing the host customer to use the resource for their own benefit most of the time while retaining the ability to schedule or aggregate those resources during critical periods when the large load wants or needs to reduce local net energy consumption. When dispatched by the large load, these distributed energy resources can offset its net demand, effectively contributing to capacity during tight system conditions.

## Load Flexibility Benefits for Long-Term Planning

Large load flexibility needs to be planned for and procured proactively in utility or regional grid operator planning processes, and early in the design of data center campuses. Long-term planning processes typically evaluate capacity needs over a 5- to 20-year horizon and differ across restructured and vertically integrated markets, but

---

[8] This is when data center service level agreements with their customers allow. Certain service level agreements require data to be located in a particular location, limiting the capability for workload shifting.

both can integrate large load flexibility in meaningful ways. Across the various options, flexibility can be explicitly modeled as a demand-side capacity resource, reducing the need for new firm generation and better aligning load management programs with broader-capacity portfolios.

Integrating flexibility into long-term planning models and capacity market frameworks provides several key benefits, as it:

- **Reduces total capacity requirements.** Traditional capacity planning methodologies assume that each megawatt of new firm load requires an equivalent amount of firm generation plus a PRM. If large loads can reduce or shift consumption during resource adequacy risk hours, their effective firm demand will be lower than their nameplate capacity rating. This can meaningfully reduce the total capacity additions required to maintain reliability and meet PRM requirements, thereby lowering system costs.
- **Better leverages behind-the-meter generation.** Many large loads are already equipped with on-site generation for back-up or emergency use. These assets include diesel, gas, or battery storage, and represent an untapped source of resource adequacy support. If operated under defined conditions during system stress, on-site generation can substitute for or defer grid-connected capacity additions while reducing the reliance on older, inefficient, and more polluting peaker plants.
- **Enhances reliability during scarcity events.** A diverse portfolio of capacity and flexibility resources creates a more resilient grid. In the current environment, where supply chain constraints, transmission bottlenecks, and rising capital costs limit the pace of new generation, flexibility offers a faster—and often more affordable—solution for ensuring reliable operations during system-stress events.

In contrast to reactive, near-term solutions, purpose-built flexibility programs planned up front as part of capacity planning offer a preferred path forward for capturing the value of flexibility for resource adequacy.

## Opportunity to Consider Flexibility *Up Front* in Planning

Power system planners, in both restructured and non-restructured markets, have long relied on demand-side flexibility from commercial and industrial customers as a resource adequacy tool. Programs that compensate customers for temporarily reducing or shifting consumption during critical grid conditions are now well-established components of many resource adequacy

TABLE 3

## Value of Electricity and Price-Driven Flexibility Responsiveness for Data Center Categories

| Category | Ballpark Power Capacity | Ballpark Consumption Value* | Flexibility | Back-Up Generation as Grid Resource** |
|---|---|---|---|---|
| **AI training** | 100–1,000 MW[1] | > > $10,000 ($/MWh) | Curtailments may not be economic at nearly any cost | Can use for emergency grid support |
| **AI requests (inference)** | 10–50 MW[2] | > $10,000 ($/MWh) | Curtailment unlikely to be economic | Can use for emergency grid support |
| **Conventional data centers** | 10–100 MW | ~ $1,000–$10,000 ($/MWh) | Evidence of load shifting during high prices, emergencies | Can use for emergency grid support |
| **Cryptocurrency processing** | 10–500 MW[3] | ~ $250–$1,000 ($/MWh) | Relatively flexible | Can use for emergency grid support |

* Electric energy is priced up to ~$5,000/MWh during emergencies, while the social costs of grid emergencies (such as rotating outages) can exceed $10,000/MWh.

** Utilization of emergency back-up generators for grid support depends on emissions and permitting.

[1] Patel et al., "AI Datacenter Energy Dilemma – Race for AI Datacenter Space, SemiAnalysis," March 13, 2025, https://newsletter.semianalysis.com/p/ai-datacenter-energy-dilemma-race

[2] Metrobloqs, "Urban AI: Why Inference Workloads Demand Edge Data Centers," accessed June 22, 2025, https://www.metrobloks.com/post/urban-ai-why-inference-workloads-demand-edge-data-centers

[3] EIA, "Tracking Electricity Consumption From U.S. Cryptocurrency Mining Operations," February 1, 2024, https://www.eia.gov/todayinenergy/detail.php?id=61364

Source: A. Levitt and S. Shparber, "The Challenges and Opportunities of Onsite Resources in Accelerating Data Center Interconnection and Reliability," Brattle and Mintz webinar, June 24, 2025, https://www.brattle.com/insights-events/publications/the-challenges-and-opportunities-of-on-site-resources-in-accelerating-data-center-interconnection-and-reliability/.

> **In contrast to reactive, near-term solutions, purpose-built flexibility programs planned up front as part of capacity planning offer a preferred path forward for capturing the value of flexibility for resource adequacy.**

frameworks, although their capacity accreditation and eligibility guidelines are being reassessed across many ISOs (ESIG, 2025b). In traditional demand-side management programs, however, commercial and industrial participants negotiate flexibility after interconnection, enrolling voluntarily or through negotiated tariff agreements in exchange for incentive payments. Upon receiving a signal from the utility or ISO, they can shed or shift load and receive compensation typically denominated in $/MWh of curtailed energy.

However, this post-interconnection approach is unlikely to be sufficient for certain large loads. Hyperscale data centers in particular have uptime and service reliability commitments that carry economic consequences that dwarf conventional demand response payments, and these data centers can be deployed significantly faster than traditional supply-side resources. Recent analysis by The Brattle Group, summarized in Table 3, highlights the high economic value of continuous operation for various data center types, estimating the value of computation for AI training and inference exceeding $10,000 per MW. This exceeds the value of lost load (VoLL) used by most utilities and markets to determine how much loads are compensated for shedding or shifting load. It means that curtailing power to a data center after it is already built and operational, using only real-time pricing mechanisms, would require very high prices that would likely exceed the highest wholesale price caps. In addition, under most interconnection agreements, the customer typically has an inalienable right to continue consuming once the load is interconnected, regardless of end use. For example, the Midcontinent Independent System Operator (MISO) recently received FERC approval to raise its VoLL to $10,000 for load-shedding events. ERCOT's energy-only

**If flexibility is valued up front as an avoided capital cost ($/MW) rather than as an after-the-fact operational response to real-time prices ($/MWh), this will increase optionality for large loads to participate in flexibility programs and ultimately make participation more likely.**

market features up to a $5,000/MWh day-ahead and $2,000/MWh real-time system-wide offer cap (FERC, 2025b; ERCOT, 2025).

Instead, planners and policymakers can think differently. If flexibility is valued up front as an avoided capital cost ($/MW) rather than as an after-the-fact operational response to real-time prices ($/MWh), this will increase optionality for large loads to participate in flexibility programs and ultimately make participation more likely. Incorporating flexibility early in the planning process creates new opportunities for planners to align incentives and avoid costly overbuilds while providing large load developers with an incentive to develop their sites proactively with flexibility in mind.

## Accelerated Interconnection in Exchange for Flexibility

One such opportunity is for utilities and grid operators to offer accelerated interconnection in exchange for flexibility, as is currently being proposed in the SPP CHILL service (SPP, 2026b). Large loads could receive faster interconnection approvals or streamlined study processes if they commit to defined levels of dispatchable flexibility, supported by verifiable on-site generation or load-management capabilities. This arrangement aligns system needs (timely access to capacity) with developer interests (speed to market). While all large loads place a high value on reliability and uptime, making a tradeoff between speed to interconnect and uptime may be necessary, namely when large loads are seeking to interconnect before the grid has been able to add the new resources or transmission needed to serve them.

For example, a large load wants to interconnect in two years, but the grid has insufficient capacity to meet resource adequacy requirements, which would require four years to develop, construct, and interconnect new capacity. One option is that the large load waits an additional two years to interconnect, essentially curtailing 100% of its load over that two-year time period (17,520 hours of load curtailment). An alternative would be to accept a modest amount of required load curtailment, either through shifted workloads or the use of on-site generation (e.g., 2 to 5% of total energy needs), and interconnect on time with fewer than 500 hours of load flexibility or curtailment controlled by the grid operator.

Tariff-based incentives for flexible design are another possibility. Large load customers could be presented with transparent capacity and network upgrade costs prior to interconnection and offered lower long-term rates or reduced capacity charges if they agree to flexible operating conditions. In essence, flexibility becomes a tool to offset the need for incremental firm capacity additions, providing both cost savings and reliability assurance. These mechanisms shift the incentive structure offered by utilities from short-term operational payments to long-term capital savings—where the benefits can reach billions of dollars in avoided infrastructure investment and deferred generation procurement. However, most investor-owned utilities do not have an incentive to do this if they earn a guaranteed rate of return on those capital expenditures. Large load owners could then shop for utility rates, uptime guarantees, and flexibility requirements that best fit their needs before choosing a location to interconnect and before the data center supply is developed.

## Proactive Flexibility in Vertically Integrated and Restructured Market Frameworks

Both non-restructured and restructured markets can accommodate upfront flexibility planning. In non-restructured markets, flexibility can be explicitly evaluated in integrated resource plans (if available) and procurement processes, either as a load-modifying or supply-side capacity resource. In restructured markets, capacity market and reliability construct designs can allow large loads to count verified flexibility toward either their ndividual or the utility's aggregate capacity obligations, or to sell that flexibility value back into the market. In both contexts, resource adequacy analysis is the key analytical tool that can support the accounting and integration of large load flexibility.

## A Six-Step Process for Bringing Flexibility Up Front in Planning

To effectively integrate large load flexibility into long-term planning, both grid operators and large load developers need access to clear, actionable planning information. For system planners at ISOs and utilities, this means understanding how large load flexibility alters capacity requirements, resource procurement strategies, and reliability risk. For large load customers, it means knowing how flexibility can reduce long-term costs, accelerate interconnection timelines, and shape contractual arrangements with capacity providers or flexibility aggregators.

The six-step process summarized in Figure 6 provides a structured and replicable method to help quantify flexibility needs. The process allows planners to identify the specific quantity, location, and type of load flexibility required to maintain reliability under different load-growth scenarios and translate those needs into avoided capacity investments. The six-step process leverages well-established planning tools, including capacity expansion and resource adequacy models, to systematically evaluate how large load flexibility affects system reliability and investment needs. By isolating the incremental impact of flexibility on key planning metrics, this process can provide the economic foundation for new contract structures and tariffs that equitably allocate costs and benefits. It also offers a transparent, data-driven foundation for decision-making that can be applied consistently across regions and market structures.

> **This process allows planners to identify the specific quantity, location, and type of load flexibility required to maintain reliability under different load-growth scenarios.**

### Step 1: Characterize the Loads with Greater Fidelity

Addressing flexibility proactively in power system planning starts with clearly and comprehensively characterizing large loads.[9] These loads represent a wide range of technologies, industries, and operational profiles—from hyperscale data centers to industrial electrification and hydrogen production. Each of these load types can behave very differently from both conventional demand and from one another. Even within data center loads, there is a wide range of hourly load patterns and opportunities for flexibility. For example, data centers vary based on end use, including cloud computing applications, where compute needs fluctuate based on end-user demand; AI training, which follows its own, less predictable load profile; AI inference models to support queries; and crypto mining, which typically operates at high load factors but is also highly flexible. In addition, multi-tenant data centers may have multiple use cases operating simultaneously at a single facility. Capturing these differences is essential to accurately represent their effects on the power system and to assess their potential to support resource adequacy through flexible operations.

FIGURE 6

Six-Step Process for Planning for Large Load Flexibility

| Step 1 | Step 2 | Step 3 | Step 4 | Step 5 | Step 6 |
|---|---|---|---|---|---|
| Characterize loads with greater fidelity | Improve large load forecasts | Quantify capacity requirements of large loads without flexibility | Quantify capacity contributions of large load flexibility | Calculate the avoided infrastructure investments resulting from large load flexibility | Develop regulatory mechanisms that account for the value of large load flexibility |

Source: Energy Systems Integration Group.

---

[9] See the ESIG Large Loads Task Force reports *Grid Integration of Large Loads: Introduction to the Large Loads Task Force, Data Needs, and Flexibility Overview* (https://www.esig.energy/reports-briefs/large-load-task-force-introduction) and *Large Loads: Behaviors, Capabilities, and Limitations* (https://www.esig.energy/reports-briefs/large-loads-behaviors-capabilities).

Planners may segment large loads according to their impact on system adequacy. This type of segmentation serves three purposes. First, it enables model parameterization. As with generation resources, different types of large loads have distinct behaviors that can be reflected in model inputs, such as hourly load shapes, coincidence with system peaks, and flexibility response capabilities. Accurate segmentation allows planners to generalize load classes while still capturing critical distinctions that matter for grid reliability. The diversity and novelty of many large loads entering interconnection queues today makes this particularly important.

Second, segmentation supports targeted data collection. By classifying large loads by characteristics relevant to resource adequacy, planners can identify and prioritize the operational data needed to evaluate their effects. This includes not only typical metrics like annual energy use or peak demand, but also more specialized data such as ramp rates, availability windows for flexibility, and on-site generation configurations.

Several organizations have already developed frameworks that offer a starting point for segmenting large loads and facilitating standardized collection of key data. For example, EPRI's DCFlex initiative classifies data centers based on their size, workload type, ownership model, and reliability requirements—all factors that influence their ability to provide flexibility (EPRI, 2025). Similarly, NERC has proposed grouping large loads into broad categories such as data centers and other computational loads, industrial loads, and hydrogen production facilities (NERC, 2025a). These high-level categories can help guide modeling, regulatory treatment, and reliability assessments. But the key to effective reliability planning is ensuring that core resource adequacy characteristics are clearly parameterized within each segment.

More information on large load categories and potential examples can be found in the ESIG report *Grid Integration of Large Loads: Introduction to the Large Loads Task Force, Data Needs, and Flexibility Overview*, which illustrates the variety of large load types and subtypes (as shown in Figure 7, p. 26), each of which can feature distinct operating behaviors and modeling needs.[10]

Planners can develop a segmentation framework that informs forecasting, resource adequacy modeling, and market design by characterizing large loads based on how they may influence resource adequacy—including size, load profile, price responsiveness, flexibility potential, and on-site generation. This foundational step enables utilities, system operators, and regulators to treat large loads with the same analytical rigor historically applied to generation resources.

**Outcome:** A classification of large loads to improve data inputs (Step 2) and parameters for modeling large load flexibility to differentiate potential resource adequacy requirements and contributions (Step 4).

## Step 2: Improve Large Load Forecasts

Large loads challenge traditional load forecasting methodologies because of their size, unique operational characteristics, and business dynamics.[11] First, forecasters must recognize the inherent uncertainty associated with large load development. Developers frequently pursue multiple sites simultaneously and may cancel or delay projects based on economic or regulatory conditions. Unlike population-driven demand, large loads tend to come online in discrete, project-specific increments, with substantial step changes in demand that are orders of magnitude larger in impact on location, timing, and scale compared to traditional load growth. To mitigate some uncertainty, planners can develop a range of future demand forecasts to reflect the full spectrum of potential outcomes.

To the extent possible, forecasts for resource adequacy analyses—and ultimately resource procurement—need to utilize contractual agreements rather than rely solely on interconnection queue requests. While load interconnection requests are a visible and easily accessible data source, they are a poor proxy for actual load realization. Projects often hold multiple queue positions for optionality, and a significant portion of queued requests never materialize. Signed, long-term power contracts or executed interconnection agreements provide a stronger signal of intent to build, so forecasts can be bifurcated to include a portion that is based on long-term contracts and another

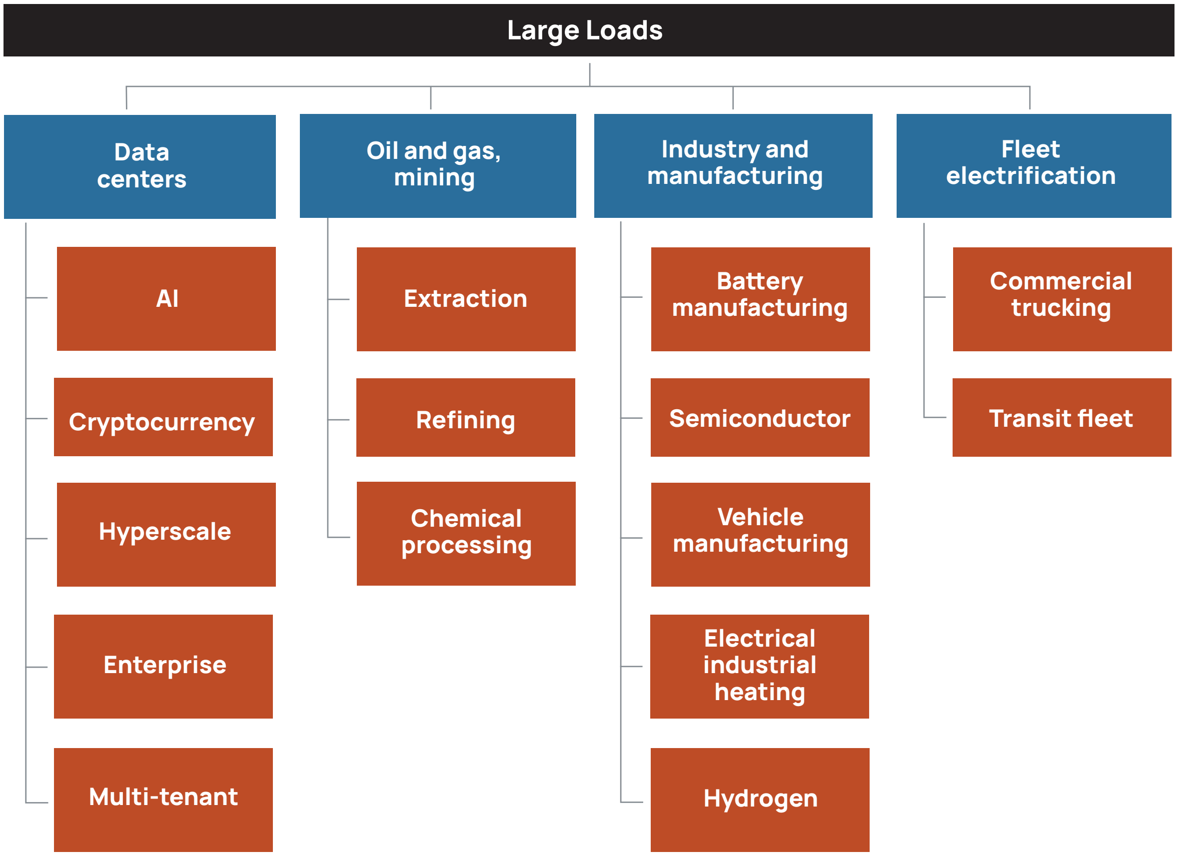
---

[10] https://www.esig.energy/reports-briefs/large-load-task-force-introduction

[11] https://www.esig.energy/reports-briefs/forecasting-for-large-loads

FIGURE 7

### Large Load Segmentation Across Subtypes and End Uses



Source: Energy System Integration Group.

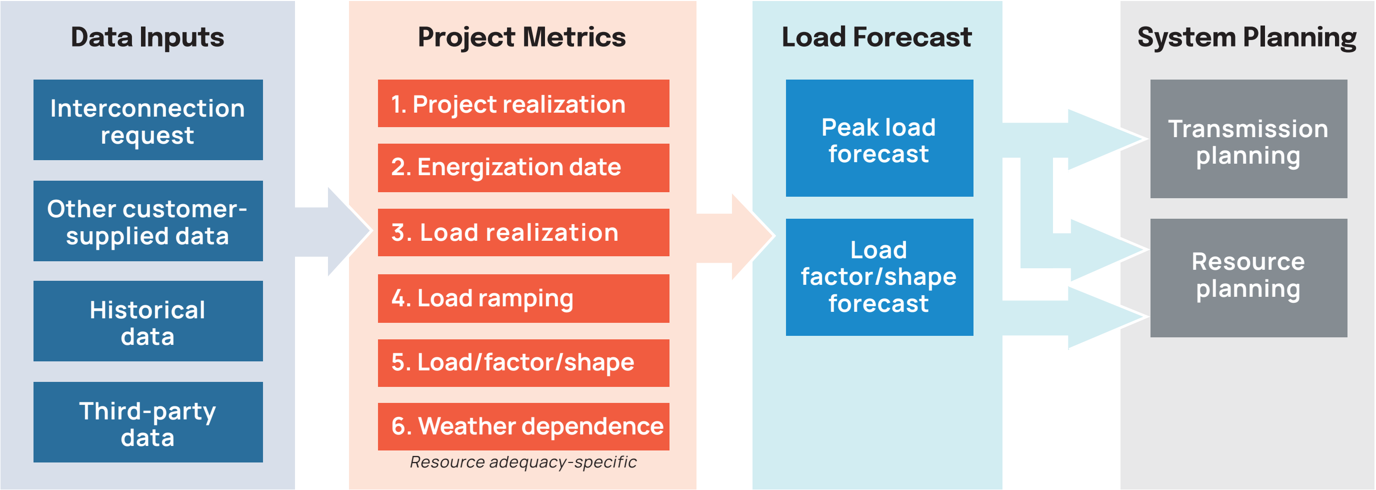
that is more uncertain. Other indicators such as active permitting, confirmed equipment orders, and ongoing construction activity can also be used to assign probabilistic weights to different projects in the forecast, improving accuracy and confidence (ESIG, 2025a).

Forecasts will also need to be specific in terms of both peak demand and annual energy use (Figure 8, p. 27, load forecast). In many cases, large load interconnection requests include a request of maximum offtake but do not provide useful annual load factor assumptions. To be useful in resource adequacy applications, large load additions can be represented with the key inputs of both their expected annual energy consumption (MWh) and their contribution to seasonal or system peaks (MW). Currently, many load forecasters and resource adequacy modelers are forced to choose an assumed load factor for large loads despite little or no historical and publicly available information to support those assumptions. Typically, these assumptions err on the conservative side, using a high load factor (i.e., assuming baseload, or flat consumption profiles). But this assumption fails to acknowledge that large load demand, including data centers, fluctuates at hourly, daily, and seasonal scales.

To be useful for resource adequacy analysis, the demand profiles for large loads should also include 8,760-hour chronological profiles to capture diurnal and seasonal variation in load behavior, ideally over multiple weather years. When possible, these profiles should be weather-dependent to reflect cooling or other environmental sensitivities common to data centers and industrial facilities. This “weather dependence” forecast metric, specific to resource adequacy assessment, can be included alongside the five core metrics that large load forecasts can use to characterize and evaluate new loads

FIGURE 8

## Large Load Forecasting Process



**Large load forecasting process from inputs and metrics to chronological 8,760 demand profiles.**

Source: See the Large Loads Task Force report *Forecasting for Large Loads: Current Practices and Recommendations* at https://www.esig.energy/reports-briefs/forecasting-for-large-loads/.

(Figure 8, forecast metrics), as described in the in the ESIG Large Loads Task Force report *Forecasting for Large Loads: Current Practices and Recommendations.*[12]

**Outcome:** Large load forecasts that include necessary data for resource adequacy analysis, including seasonal peak demand (MW), annual energy (GWh), hourly 8,760 chronological demand profiles, and regional breakdown of load additions.

## Step 3: Quantify Capacity Requirements of Large Loads Without Flexibility

Before planners can assess the value of large load flexibility, they need to first establish a clear baseline: the total amount of new capacity that would be required if large loads were to come online as firm, inflexible demand. This step provides the point of comparison needed to identify avoided or deferred infrastructure builds, costs, and associated emissions in later steps. The baseline capacity requirements associated with new large loads will depend on both the unique operating characteristics and the load profiles of the loads themselves, as identified in Step 1 and Step 2, and the assumptions planners make about system reliability and risk tolerance. Figure 9 (p. 28) provides a visual overview of Step 3, showing how new large loads without flexibility require new accredited capacity to meet increased capacity requirements.

Despite the new challenges posed by the size and uncertainty of large loads, planners can still rely on the same foundational resource adequacy planning principles that have guided power system reliability planning for decades. These include the PRM, possibility of fuel disruptions, renewable variability, and forecast error. The PRM ensures that sufficient firm capacity is available to reliably meet load, even under stressed system conditions.[13]

To determine how much capacity is required and how much each resource contributes to meeting the PRM requirement, planners use probabilistic resource adequacy models. These models simulate thousands of potential

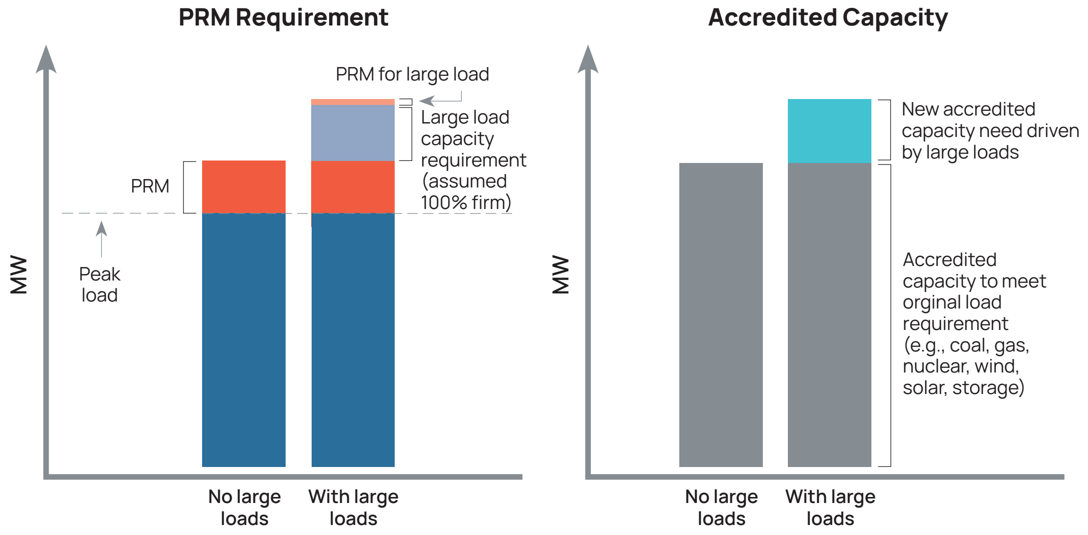
---

[12] https://www.esig.energy/reports-briefs/forecasting-for-large-loads/

[13] Some grid planners and researchers are developing alternatives to resource planning that is based on the PRM and capacity accreditation, topics that are covered in other ESIG task forces including the Redefining Resource Adequacy Task Force. This report does not cover potential alternative capacity-planning frameworks, as it is intended to help integrate large loads into many existing utility planning frameworks. The principles described here for large loads would broadly apply to other planning frameworks as well.

FIGURE 9

## Quantifying Capacity Requirements of Large Loads Without Flexibility: A Baseline



**Visual overview of the load and capacity requirement with and without large loads (left) and the accredited capacity necessary to meet the new requirements (right). When large load additions are assumed to be 100% firm, they increase firm capacity requirements equal to the nameplate capacity of the new large loads, plus an additional capacity buffer based on the planning reserve margin (PRM) requirement (left). Additional accredited capacity is then required to meet the increased capacity need (right).**

Source: Energy Systems Integration Group.

system conditions, spanning multiple weather years and generator outage combinations, to estimate the likelihood of shortfalls and identify the total firm capacity needed to meet a specified reliability criterion, often expressed as a 1-day-in-10-year loss-of-load expectation (LOLE). The required amount of firm capacity to achieve this criterion can be translated into a PRM by representing it as a percentage of the expected peak demand. These models also calculate capacity accreditation values, such as ELCC, which indicate the portion of a resource's nameplate capacity that can be counted toward meeting the PRM requirement.

Resource adequacy models are used to determine the amount of firm capacity needed to meet the reliability criterion and the amount each resource can contribute based on their respective ELCCs. However, resource adequacy models do not provide insight into the least-cost mix of resources that can meet these requirements or the financial or environmental implications of different portfolios. To address these issues, planners turn to capacity expansion models, which take the PRM and resource ELCCs as inputs and identify cost-optimal portfolios of new resources that meet reliability, policy, and operational constraints. These models not only specify the mix and timing of new capacity additions (in megawatts by resource type) but also estimate the total portfolio cost (net present value) and emissions of the portfolio over the planning horizon.

To isolate the capacity requirements attributable to inflexible large loads, planners can run multiple planning scenarios. These include a business-as-usual case without new large load growth, and one or more large load growth scenarios consistent with the forecasts developed in Step 2. By comparing results across these scenarios—each meeting the same reliability standard—planners can clearly attribute any increase in required firm capacity to the addition of large loads. This approach is consistent with how utilities allocate capacity obligations to LSEs today, and it can support transparent, equitable cost assignment for new large load customers.

This step provides the foundation for understanding the system impact of large loads under firm, inflexible lassumptions.

**Outcome:** Capacity requirements (MW) attributed to large loads, by forecast year, assuming no large load flexibility.

### Step 4: Quantify Capacity Contributions of Large Load Flexibility

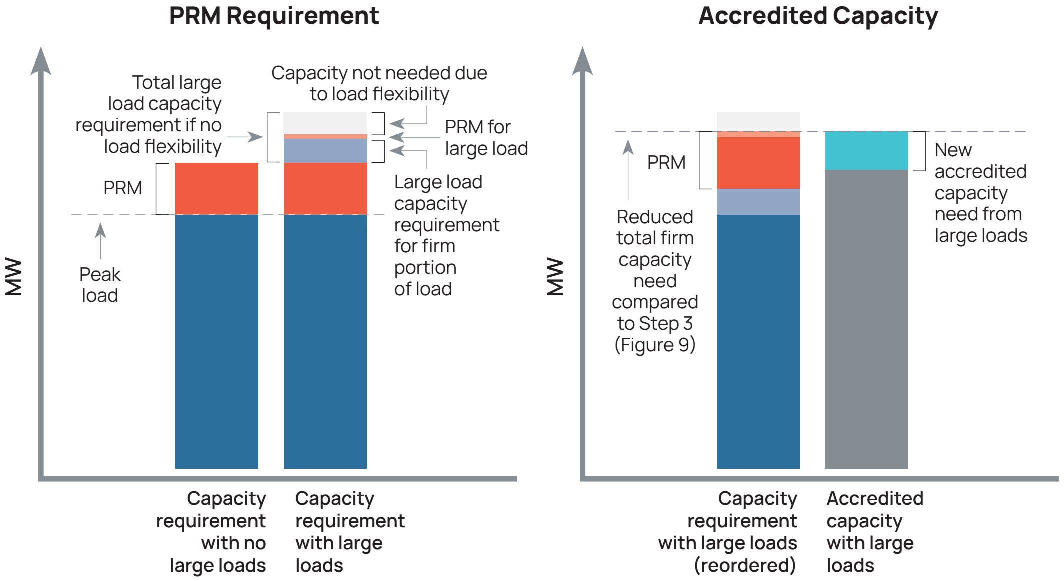
Once planners have established the baseline capacity requirements for large loads under inflexible conditions, the next step is to quantify the capacity contributions that can be provided by large load flexibility. This analysis uses resource adequacy models to measure how much firm capacity the system can avoid procuring because flexible loads are able to adjust their demand, rely on on-site resources, or otherwise reduce stress on the grid during critical periods.

> Once planners have established the baseline capacity requirements for large loads under inflexible conditions, the next step is to quantify the capacity contributions that large load flexibility can provide. Planners determine how much firm capacity the system can avoid procuring because flexible loads are able to adjust their demand, rely on on-site resources, or otherwise reduce stress on the grid during critical periods.

Large load flexibility can be represented in two primary ways within a resource adequacy framework, depending on whether its reliability benefits are treated as shared across the system or directly credited to the load providing them. Both approaches will result in equivalent system reliability but distribute the benefits of flexibility differently: the first distributes those benefits equally across load while the second recognizes the specific contributions of flexible loads and accredits those specific resources accordingly, making it cheaper for the load to interconnect.

FIGURE 10

**Option A for Quantifying Capacity Contributions of Large Load Flexibility: Reduce the PRM Requirement**



- Only the firm portion of large loads counts toward the PRM.
- The PRM requirement is effectively reduced.
- Large load flexibility is not accredited.
- Flexibility benefits are socialized across the system's load.

**Overview of the load and capacity requirement with and without large loads (left) and the accredited capacity necessary to meet the new requirements following the addition of large loads (right). The left side shows how, when only a portion of large load additions are assumed to be firm, they increase capacity requirements equal to the firm capacity of the new large loads, plus an additional capacity buffer as a percentage of the firm large load capacity based on the planning reserve margin (PRM) requirement. The right side shows how additional accredited capacity is then required to meet the increased capacity need. Since only firm capacity of new large loads contributes to the capacity requirement, large load flexibility cannot count as an accredited capacity resource.**

Source: Energy Systems Integration Group.

### Treating Reliability Benefits as Shared Across the System as a "Load Modifier"

In the first approach, shown in Figure 10 (p. 29), for flexible large loads, only the firm (non-flexible) portion of the large load demand is counted when calculating the PRM requirement. This effectively reduces the overall reserve margin needed to meet the reliability criterion (the right bar in "PRM Requirement") and lowers total system capacity requirements (the right bar in "Accredited Capacity")—both relative to the baseline shown in Figure 9 (p. 28). Because the benefit of flexibility is distributed across all customers through lower generation and transmission costs, this approach spreads the benefits of large load flexibility proportionally across all customers. This improves adequacy system-wide without directly accrediting the flexible load itself.

### Crediting Reliability Benefits Directly to the Load Providing Them as a "Competitive Resource"

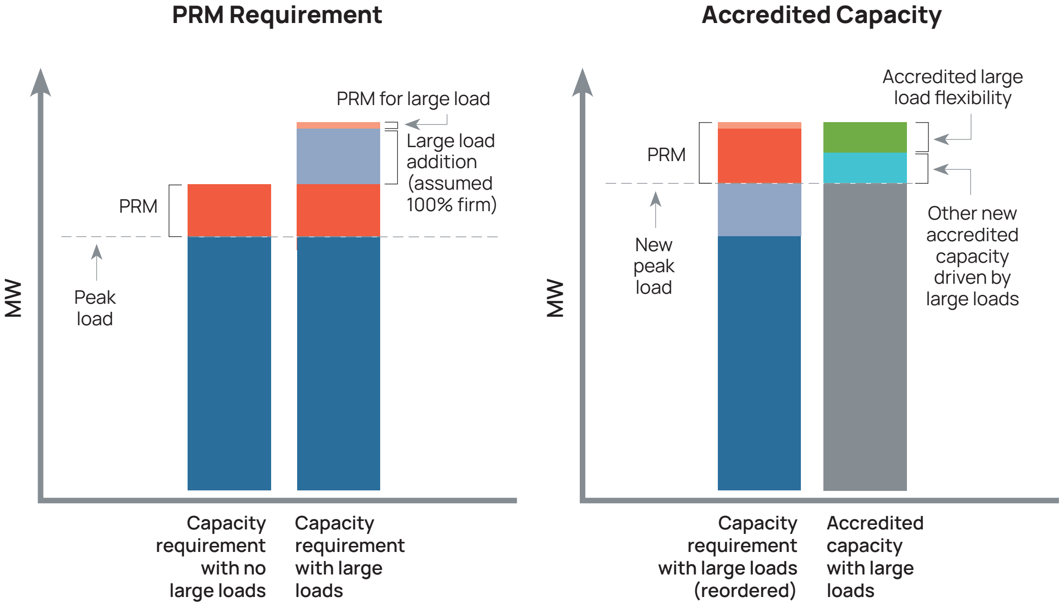
Alternatively, planners can directly assign capacity accreditation to flexible large loads using a "demand-side ELCC" or equivalent method, as shown in Figure 11. In this framework, all large load demand is included in the PRM calculation, but the flexible portion is assigned an ELCC based on its expected performance during system stress events. The total PRM remains unchanged, but this flexibility is treated as a capacity resource with its own accredited contribution to the PRM requirement.

This approach internalizes the value of flexibility to the large load resources themselves, allowing the large load customer to offset part of its own capacity obligation or receive explicit compensation for its contribution to reliability.

---

FIGURE 11

**Option B for Quantifying Capacity Contributions of Large Load Flexibility: Accredit Large Load Flexibility**



- All large load counts toward the PRM.
- The PRM requirement remains unchanged.
- Flexibility is accredited and has capacity commitments.
- Flexibility benefits are internalized to the large loads.

**Overview of the load and capacity requirement with and without large loads (left) and the accredited capacity necessary to meet the new requirements (right). The left side shows the same baseline as given in Figure 9 (p. 28), namely, how when large load additions are assumed to be 100% firm, they increase firm capacity requirements equal to the nameplate capacity of the new large loads, plus an additional capacity buffer as a percentage of the large loads' nameplate capacity based on the planning reserve margin (PRM) requirement. The right side shows how additional accredited capacity, which can include large load flexibility, is then required to meet the increased capacity need.**

Source: Energy Systems Integration Group.

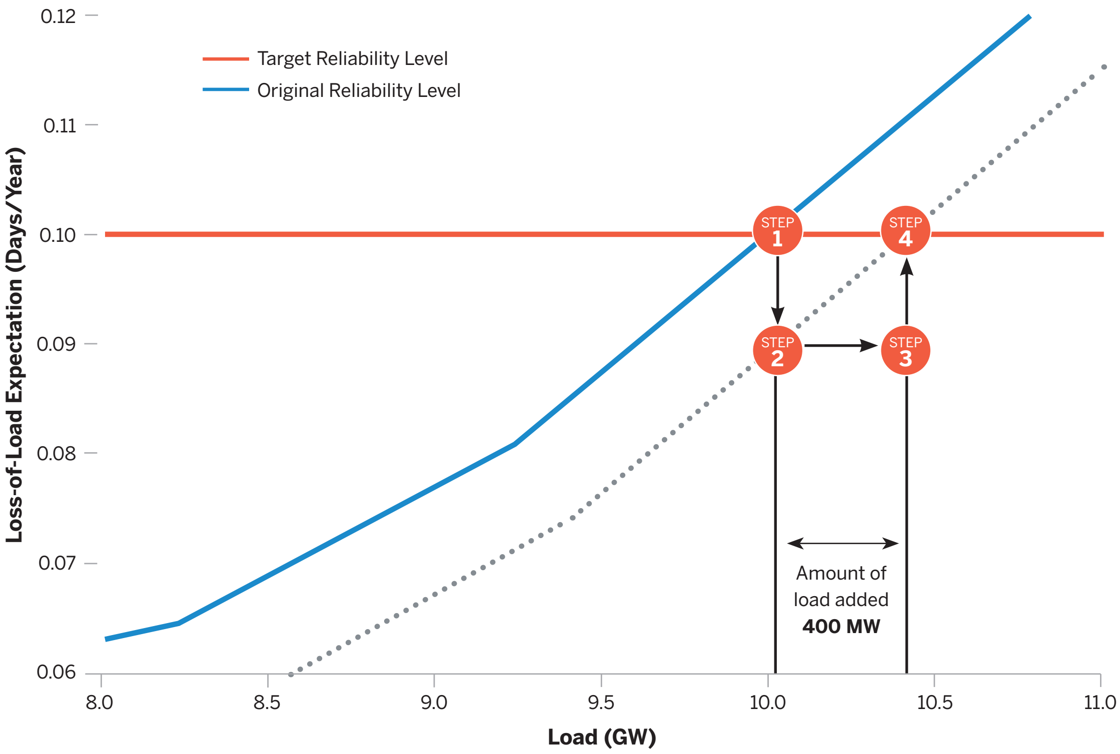
### Quantifying the Capacity Contribution of Large Load Flexibility

To compute the capacity contribution of large load flexibility, planners can apply the same methods used to accredit generation resources within existing resource adequacy frameworks. This "demand-side ELCC" approach is not new. Several utilities and consultants have applied similar methods to evaluate capacity contributions from demand-side resources in integrated resource plans, and recent studies have developed ELCC-based accreditation for large loads, including data centers and industrial facilities (E3, 2024; Cox, Schwartz, and Stenclik, 2025). Using established probabilistic models and methods ensures that flexibility is evaluated under the same reliability framework as generation resources, allowing direct comparison and consistent treatment in planning and procurement processes.

The process for computing demand-side ELCC typically involves four key steps, which directly apply for large load flexibility accreditation, as illustrated in Figure 12. First, planners calibrate a baseline case that meets the system's target reliability criterion, such as a 1-day-in-10-year LOLE, for a chosen future planning year, assuming that large load demand is fully firm and inflexible. Second, they re-run the model allowing flexibility from large loads, such as voluntary curtailment or on-site generation dispatch, which will reduce the system's LOLE below the target level. Third, planners incrementally add firm load back into the model until the LOLE returns to the target reliability

FIGURE 12

The Four Steps in the ELCC Methodology



**Load flexibility can be calculated with the same method used to calculate effective load-carrying capabilities (ELCCs) for generating resources. The system is first brought to the reliability criterion (e.g., 1-day-in-10-year loss-of-load expectation (LOLE)) (Step 1). In Step 2, a large load's flexibility is added to the system, including constraints on response size, duration, number of calls per month or year, etc. This load flexibility reduces LOLE, making the system more reliable. The modeler then adds a fixed amount of load to the model in Step 3 until the original LOLE criterion is reached (Step 4). The difference between the amount of load added relative to the amount of load flexibility added for a given large load is its ELCC.**

Source: E. Ibanez and M. Milligan, "Comparing Resource Adequacy Metrics" (preprint, National Renewable Energy Laboratory, 2014), https://www.nrel.gov/docs/fy14osti/62847.pdf.

criterion. Finally, the ELCC of the large load flexibility is calculated as the ratio of the added firm load (in MW) to the amount of flexible load modeled in the second step. This result quantifies how much firm capacity the flexible load is effectively providing to the system.

In quantifying capacity contributions, planners will need to characterize flexibility across a range of parameters to capture its operational limits and system value. These could include the size of the flexible load (in MW), the frequency of deployment (how often flexibility can be called upon in a year), and the duration of each event (hours per call). These dimensions define a capacity accreditation matrix of large load flexibility, as shown in Figure 13. This allows planners to test different combinations of parameters and identify how incremental flexibility affects total system reliability. For example, a short-duration flexibility resource available for frequent dispatch may provide different adequacy value from a long-duration flexibility resource that can be used only a few times per year. By conducting sensitivity analyses across these parameters, planners can capture diminishing returns or saturation effects, ensuring that flexibility is credited appropriately without overstating its contribution. In addition, the capacity accreditation matrix of large load flexibility offers a clear starting point for negotiations around flexibility capabilities and requirements between large load developers, utilities, grid operators, and other stakeholders.

FIGURE 13

## Capacity Accreditation Matrix of Large Load Availability

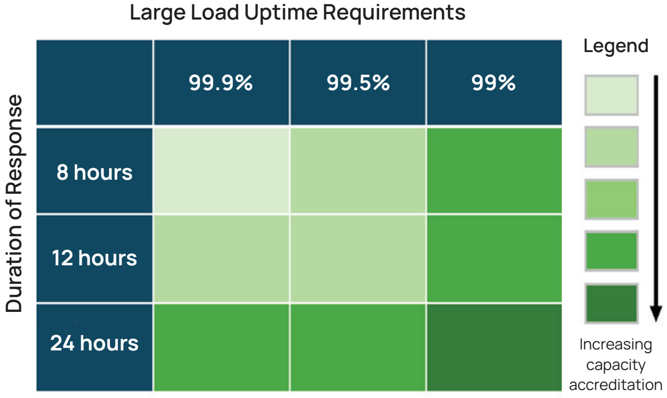


**Illustration of a capacity accreditation matrix of large load flexibility based on availability, either as an uptime requirement (the percentage of total annual energy the load is unable or unwilling to be flexible) or maximum duration of flexibility response.**

Source: Energy Systems Integration Group.

For this step, it is generally useful to assume zero cost for flexibility to determine its avoided-cost value or "break-even" benefits. This assumption isolates the physical capacity value of flexibility before incorporating cost and policy considerations. In later steps, the avoided capacity cost (dollars per megawatt) can be compared with the actual cost of enabling or contracting flexibility

and help inform the development of appropriate tariffs, contractual agreements, and financial incentives.

This provides a quantitative measure of how flexible large loads contribute to maintaining system reliability. It defines the firm capacity value of flexibility in physical terms and establishes a foundation for calculating avoided capacity additions, deferred infrastructure, and system cost savings.

**Outcome:** Capacity accreditation (MW and %) or contributions from large load flexibility.

### Step 5: Calculate the Avoided Infrastructure Investments Resulting from Large Load Flexibility

This step builds directly on the modeling approach established in Step 3 by repeating the capacity expansion analysis with changes. In Step 5, the analysis incorporates the flexibility potential of large loads, as quantified in Step 4 via large load–specific capacity accreditations. This is most relevant for vertically integrated utilities doing integrated resource planning or for ISOs in their scenario-based, long-term transmission planning analyses.

Flexibility can drive key resource adequacy analysis outcomes through a demand-side ELCC (treating flexible load as an accredited capacity resource) or by applying a lower PRM where only the firm portion of large load demand is counted toward system requirements. In either case, this modeling reveals how much infrastructure can be avoided or deferred when flexibility is incorporated proactively into planning.

As in Step 3, this analysis calculates total capacity additions (MW), portfolio net present value, and associated emissions and fuel consumption over the planning horizon. However, the inclusion of flexibility will likely change the least-cost portfolio composition. By comparing results from this step to those generated in the baseline "no-flexibility" scenario, planners can isolate and evaluate the full system benefits that large load flexibility provides. See Box 2 (p. 34) for an example.

This modeling reveals how much infrastructure can be avoided or deferred when large loads' flexibility is incorporated proactively into planning.

An important benefit is the reduction of capacity needs. When large loads are flexible, the total amount of new firm capacity required to maintain reliability is reduced. This can result in a smaller resource build-out over the planning horizon. In addition to reducing the absolute quantity of new capacity, flexibility can also defer the need for investments. Even when some new generation or storage remains necessary to support large load growth, flexibility can reduce or delay the point at which those resources are needed—a delay that can provide valuable time for newer, lower-cost, or lower-emissions technologies to become available and can help spread capital costs across a longer planning horizon. These deferred investment timelines can be translated into concrete, accelerated interconnection timelines that are especially attractive for AI data center loads, where speed-to-power is essential for maintaining a competitive business position.

As a result of these avoided or deferred capacity investments, and potentially lower production costs, the inclusion of large load flexibility can drive reduced system costs.

As a result of these avoided or deferred capacity investments, and potentially lower production costs, the inclusion of large load flexibility can drive reduced system costs. Because capacity expansion models are designed to select least-cost portfolios that meet reliability and policy constraints, when flexibility is included in the model, this will often lead to portfolios that rely less on expensive or inefficient peaking resources and instead optimize for lower total cost. The difference in net present value between the flexibility and no-flexibility cases represents a real, monetizable benefit. These avoided costs can form the basis for new tariff structures, incentives, or interconnection terms that reflect the value of flexibility and fairly allocate costs between flexible and inflexible loads (see Step 6).

This modeling process can be repeated across the matrix of flexibility options described in Step 4. Varying the magnitude, frequency, and duration of flexible load

BOX 2

## Quantifying the Value of Data Center Flexibility in Nevada

A recent study leveraging data from NV Energy's 2024 integrated resource plan demonstrates how incorporating data center flexibility into resource adequacy analysis can materially reduce capacity needs and total system costs. Data center load growth was a key focus of NV Energy's 2024 Integrated Resource Plan, and data center demand is expected to grow from roughly 5% of peak load in 2025 to nearly 18% by 2030.

Using the ELCC modeling framework, the study modeled flexible data center load as a demand-side capacity resource. Scenarios tested 1 to 2 GW of aggregated flexibility with availability levels consistent with 99.5% to 99.9% uptime. One GW of data center flexibility was able to defer 600 to 850 MW of new capacity needs. Across all scenarios, flexibility yielded $300 to $400 million in net-present-value savings over the planning horizon. As shown in Figure 14, this flexibility allowed the capacity expansion model to select fewer capacity resources such as new gas turbines and instead rely on lower-cost energy resources such as solar-plus-storage, driving system-wide cost savings.

The flexibility required to achieve these savings was minimal, averaging between 3 and 42 hours per year across the participating fleet, or 1 to 13 hours for an individual data center when data centers were individually flexed to their maximum amount of flexible load. The NV Energy case study illustrates how demand-side ELCC modeling can quantify avoided-infrastructure benefits and serve as a replicable playbook for utilities and ISOs seeking to integrate large load flexibility into future planning and procurement frameworks.

FIGURE 14

### Comparison of Expansion Plan

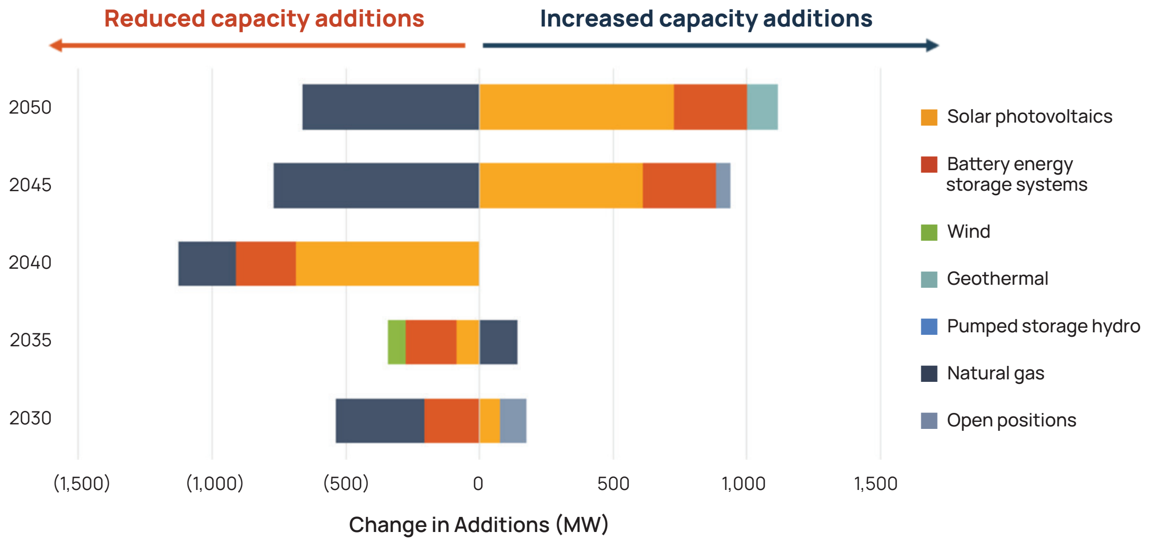


**Illustration of the change in the expansion plan resource mix when data center flexibility accreditation counts toward the capacity requirement compared to an expansion plan assuming non-flexible, fully firm data center loads. This change reflects fewer traditional capacity resources such as new gas turbines and a greater share of lower-cost energy resources such as solar-plus-storage, driving system-wide cost savings.**

Source: C. Cox, A. Schwartz, and D. Stenclik, *Bringing Data Center Flexibility into Resource Adequacy Planning: A Case Study of NV Energy*, (GridLab and Telos Energy, 2025), https://gridlab.org/datacenter_flex/.

contributions allows planners to assess how different operational profiles affect system needs and investment strategies. This enables utilities, regulators, and large load developers to compare different flexibility strategies on an apples-to-apples basis and select those that offer the greatest benefit per dollar or per megawatt of curtailment potential. This process also provides a quantitative framework for identifying diminishing returns—clarifying, for instance, when adding more flexible load delivers marginal benefits or when flexibility is most valuable in relation to system peak periods.

This step provides the system-wide metrics needed to justify and design flexibility-oriented planning and investment frameworks. It demonstrates how large load flexibility not only reduces the amount of infrastructure required but also improves the timing, cost, and environmental footprint of system development. These findings are essential for informing policy, procurement, and rate design.

**Outcome:** Quantification of accelerated timeline for interconnection and avoided costs and emissions attributed to large load flexibility.

## Step 6: Develop Regulatory Mechanisms That Account for the Value of Large Load Flexibility

The final step in bringing large load flexibility up front in long-term planning processes is to embed it within the current regulatory framework that ensures resource adequacy,[14] allocates costs fairly, and supports system reliability and utility and state goals. The outputs from Step 5, including quantified avoided capacity needs, deferred investments, system cost savings, and emissions reductions, provide the necessary analytical foundation to inform decision-making by utilities, regulators, and large load developers. This step ensures that the value of flexibility is not only modeled, but translated into practical mechanisms that govern interconnection, procurement, and rate design.

First and foremost, the information developed in Step 5 can inform utility negotiations with prospective large load customers. By quantifying how flexibility affects the timing and magnitude of required system investments, utilities can develop clear and evidence-based proposals for how quickly a load can interconnect and what capacity-related costs it may be expected to bear. For example, if a flexible large load reduces peak demand or delays the need for firm capacity compared to the no-flexibility baseline, the utility can offer an accelerated interconnection timeline, provided that flexibility is verifiably available under stress conditions. Conversely, inflexible loads that contribute to near-term adequacy challenges may be assigned higher costs or longer interconnection timelines to make sure that system reliability is not compromised as new capacity is procured.

These avoided or deferred capacity needs also support the development of new large load tariff designs grounded in the principle of cost causation. Flexible loads that demonstrably reduce their impact on system adequacy can be rewarded with lower capacity charges, reduced interconnection fees, or eligibility for performance-based incentives. In contrast, firm, inflexible loads that drive new capacity additions can be assigned a proportional share of those infrastructure costs, ensuring that ratepayers are protected from subsidizing unmitigated large load growth. This framework also provides regulators with a defensible basis for approving differentiated rate classes, such as new large load tariffs that account for the presence or absence of flexibility features in customer operations. Although new large load tariffs have been proposed and/or approved across the U.S. (see Figure 15, p. 36), none today explicitly incorporate flexibility. However, recent utility partnerships with Google to provide data center flexibility, and a commitment from Dominion Energy to propose a "high-load interruptible load tariff," suggest that flexibility provisions within large load tariffs may be on the horizon.

In jurisdictions with climate mandates or utility clean energy public policy targets, the avoided emissions quantified in Step 5 offer a powerful tool for aligning large load planning with broader policy goals.[15] By enabling faster interconnection of flexible loads, utilities can better

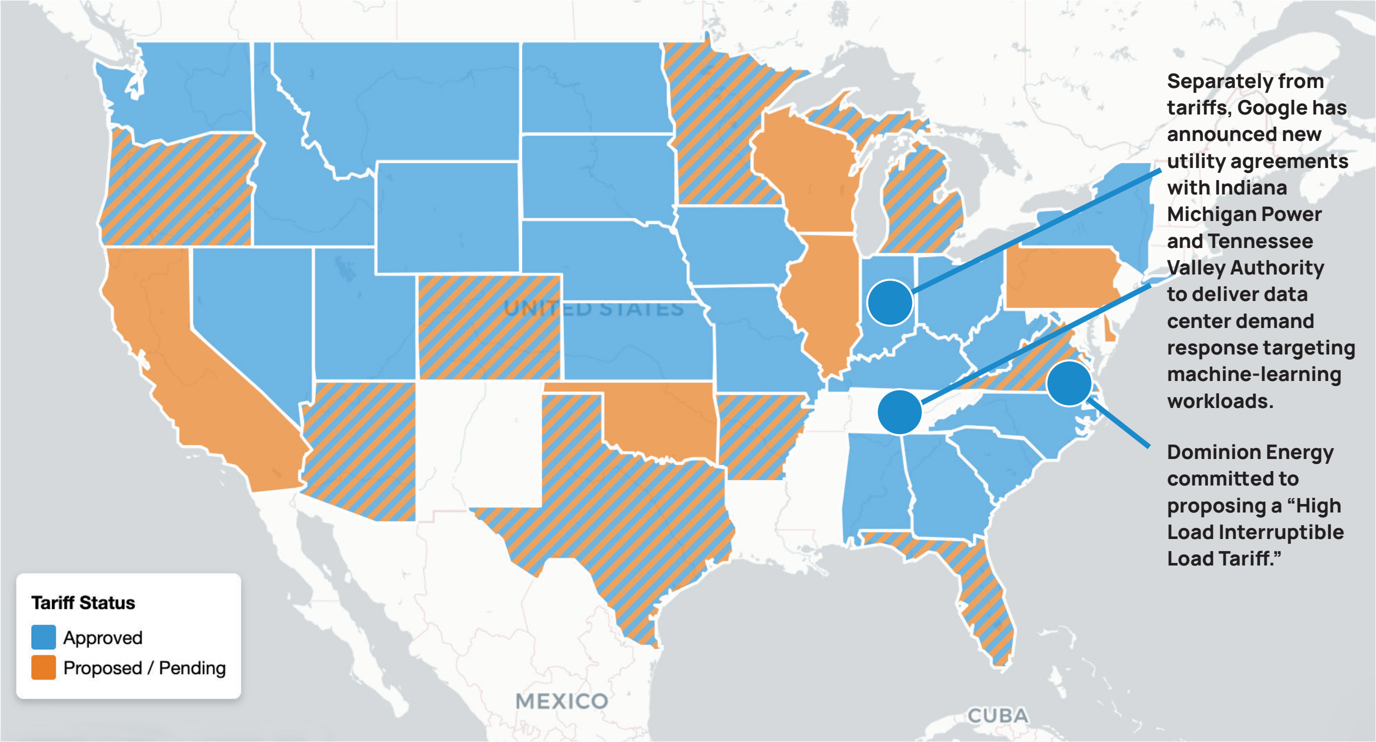
---

[14] Either the vertically utility integrated resource plans in the state regulatory context, or through the ISO capacity market requirements regulated by FERC.

[15] However, it is important to quantify not only the system-level emissions changes that result in the alternative resource mix, but also the direct emissions from the large load if it uses on-site generation and back-up diesel generation to meet flexibility needs.

FIGURE 15

## Map of Approved and Pending Large Load Tariffs, March 2026



Separately from tariffs, Google has announced new utility agreements with Indiana Michigan Power and Tennessee Valley Authority to deliver data center demand response targeting machine-learning workloads.

Dominion Energy committed to proposing a "High Load Interruptible Load Tariff."

**Map of approved and pending large load tariffs (as of March 2026), with additional examples of tariffs and utility agreements that explicitly incorporate flexibility. Approval status is based on the Smart Electric Power Alliance and NC Clean Energy Technology Center's analysis on March 31, 2026.**

Source: Smart Electric Power Alliance and NC Clean Energy Technology Center, "Database of Emerging Large-Load Tariffs (DELTa)" (2026), https://sepapower.org/large-load-tariffs-database/.

meet clean energy commitments without compromising reliability. Regulators can use these insights to prioritize load designs and interconnection pathways that are consistent with state policy, while utilities can integrate avoided emissions into the relevant resource planning efforts. Large load developers themselves can also benefit: the ability to track and report avoided Scope 2 emissions, those associated with electricity consumption, can support corporate sustainability goals and compliance with voluntary or mandatory reporting regimes.

To make these benefits actionable, utilities and regulators can work together to create a standardized menu of options that links interconnection timelines, rate structures, and flexibility requirements. For example, a load that commits to a defined number of MW of load curtailment or on-site generation during peak hours might receive expedited interconnection and reduced capacity charges, while a firm load without flexibility provisions would proceed on a standard schedule with full cost recovery. This standardized menu helps streamline negotiations, improves transparency, and incentivizes load designs that incorporate grid-supportive capabilities. It also supports scalability by reducing the need for one-off negotiations that can slow interconnection and strain utility and regulatory bandwidth.

**To make these benefits actionable, utilities and regulators can work together to create a standardized menu of options that links interconnection timelines, rate structures, and flexibility requirements.**

New service agreements, market rules, and tariffs that incorporate flexibility will also need to include clear accountability mechanisms to ensure that the promised flexibility is delivered when needed. These mechanisms may include performance tracking, audits, real-time telemetry requirements, and enforcement through incentives or penalties. Ensuring operational compliance is critical to maintaining the integrity of the resource adequacy framework and to building trust among stakeholders that flexible loads are as reliable as their modeled contributions suggest.

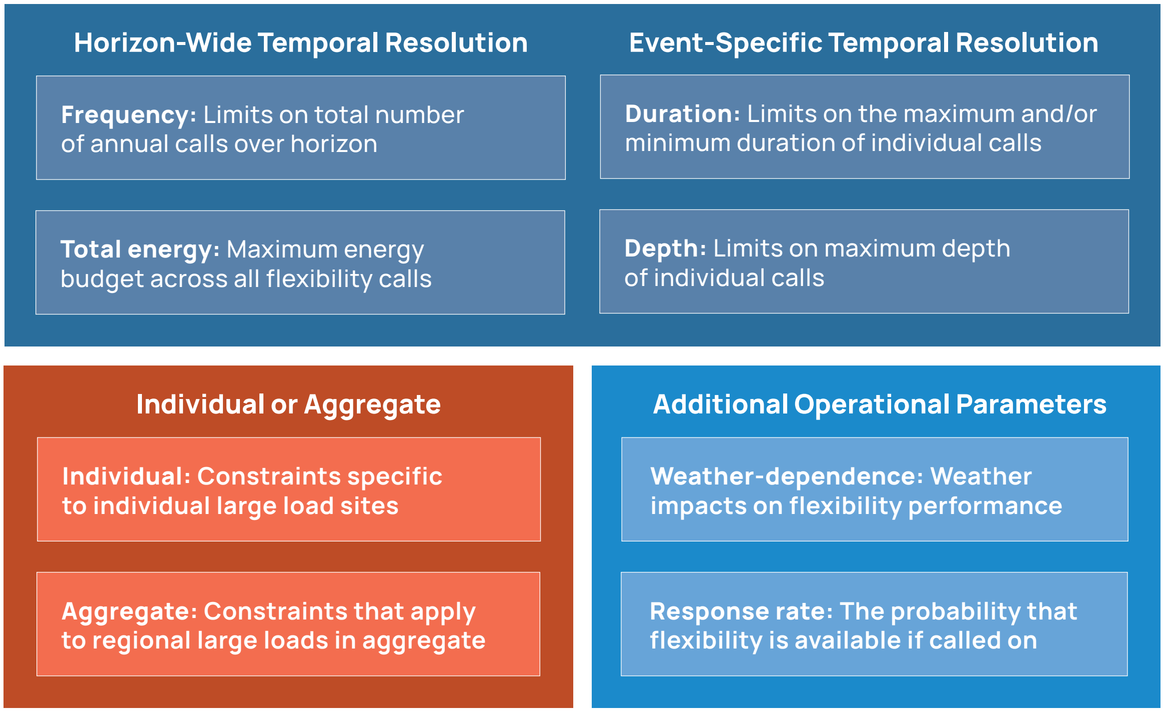
This final step ensures that the planning and modeling insights from the prior steps are fully integrated into the regulatory, commercial, and operational frameworks that govern large load development. In doing so, this step allows planners and policymakers to proactively manage reliability, accelerate interconnection timelines, reduce costs, and align large load growth with long-term system and climate objectives.

**Outcome:** Linkage of large load flexibility to capacity market design, regulations, and rate structures for large loads.

## Modeling Parameters for Large Load Flexibility

The six-step framework provides a replicable approach for quantifying the value of large load flexibility and integrating it into utility and market planning processes. While the framework is grounded in established planning tools such as resource adequacy and capacity expansion models, it also highlights the need for enhancements to current modeling practices, modeling methods, and data requirements for integrating large loads into long-term planning. Delivering on the full potential of large load flexibility will require methodological improvements, better input data, and more granular treatment of load behavior to ensure that flexibility is accurately valued and fully reflected in planning and procurement decisions. Figure 16 summarizes these flexibility-specific modeling parameters.

FIGURE 16
**Flexibility-Specific Modeling Parameters**



**Horizon-Wide Temporal Resolution**

**Frequency:** Limits on total number of annual calls over horizon

**Total energy:** Maximum energy budget across all flexibility calls

**Event-Specific Temporal Resolution**

**Duration:** Limits on the maximum and/or minimum duration of individual calls

**Depth:** Limits on maximum depth of individual calls

**Individual or Aggregate**

**Individual:** Constraints specific to individual large load sites

**Aggregate:** Constraints that apply to regional large loads in aggregate

**Additional Operational Parameters**

**Weather-dependence:** Weather impacts on flexibility performance

**Response rate:** The probability that flexibility is available if called on

**Key parameters and modeling configurations for representing large load flexibility in resource adequacy modeling.**

Source: Energy Systems Integration Group.

Planners can model the flexibility of large load facilities in aggregate or through a hybrid approach. The most suitable approach will depend on the number of large loads planning to interconnect within a planning jurisdiction, the type and quality of the data available to model individual sites, and the resolution and computational requirements of the resource adequacy modeling tools. In any case, modeling a large load flexibility response, whether of an individual data center or of an aggregated fleet of data centers, requires several event-specific and horizon-wide parameters.

### Limitations in the Depth and Duration of the Flexibility Response

Event-specific parameters include the depth and duration of the flexibility response. Regarding depth, large loads may only be willing or able to flex a portion of their demand. Similarly, the maximum duration for which large loads are able to provide flexibility services can be limited by contractual obligations, limits around data center workload shifting, limited-duration on-site generation, and other factors. Parameterizing resource adequacy models to reflect these limitations is essential to realistically and accurately model how flexibility may be able to support resource adequacy during specific system-stress events. This parameterization can be performed for individual large loads or through heuristics applied to the large load fleet in aggregate.

### Limits to the Number of Individual Flexibility Calls or Total Number of Flexibility Hours

In addition to event-specific constraints on flexibility behavior and performance, resource adequacy models can characterize additional limitations that span multiple events. There may be limits to the number of individual flexibility calls—or total number of flexibility hours—that large loads are willing to participate in over the course of a season or year. As an alternative to constraints on the total duration of flexibility response over a specific time horizon, a maximum energy budget can also allow resource adequacy models to limit flexibility. When modeling an individual large load in a resource adequacy model, using run-hour limits is a reasonable approach. However, if the modeling represents a portfolio of large loads in aggregate, an energy budget is the better approach because it allows the model to spread the portfolio's run-hour limits or call limits without individual load representation.

### Effective Performance for Large Load Flexibility During Resource Adequacy Events

Several additional operational parameters allow planners to more precisely model the resource adequacy value of large load flexibility. Like any dispatchable grid resource, large load flexibility may not show up when called on by the grid operator or may produce a smaller reduction in demand than anticipated. Because of the novel nature of new large load operations and flexibility schemes, the effective response performance rate for large load flexibility is uncertain but cannot be ignored. Just as supply-side resources have a forced outage rate that represents the probability of unanticipated unavailability, large load flexibility can be modeled with an expected response performance rate, ideally developed and refined over time as the industry gains experience and data surrounding flexibility performance.

### Flexibility-Specific Weather Impacts and Variations in Seasonal Performance

Resource adequacy models can incorporate flexibility-specific weather impacts and variations in seasonal performance. In data centers, for example, cooling loads may be higher during the summer months, creating more demand that could potentially be flexed. With a response performance rate, temperature and weather impact data on large load demand and performance can be closely monitored to better understand how these factors could affect flexibility.

# Recommendations

There are numerous ways in which power system planners and stakeholders can improve their processes, planning, and regulatory requirements pertaining to resource adequacy to better account for the needs, opportunities, challenges, and considerations of large loads. Here we consolidate these actions and divide them into five groups, so they can be tailored to meet the needs of (A) resource adequacy modelers, (B) planners at vertically integrated utilities, (C) planners and market development personnel at ISOs, (D) data center operators and developers, and (E) state and federal regulators.

## For Resource Adequacy Modelers

Many key findings and recommendations from this report rely, partially or wholly, on robust probabilistic resource adequacy modeling. This modeling simulates the power system across a wide range of uncertainties that can affect supply and demand including temperature; load variability; wind, solar, and hydro availability; and generator outages. Resource adequacy modeling must evolve to accurately assess system needs and reflect the unique characteristics of large loads, their chronological demand profiles, potential weather dependence, and characteristics of load flexibility. To provide accurate information to system planners, large load developers, and regulators, resource adequacy modelers can improve their modeling by:

### (A1) Incorporating multiple load-growth scenarios.

Probabilistic resource adequacy analysis can incorporate multiple load-growth scenarios, allowing planners to identify near- and long-term adequacy risks driven by large loads across a range of plausible uncertainties.

### (A2) Ensuring close coordination with demand forecasters.

Utility load forecasts are often developed by dedicated load forecasting teams or external consultants, and then used as inputs into grid planning models. Load forecasters must be aware of the requirements of resource adequacy models to most effectively characterize large loads, including modeling large load flexibility. Forecasts need to:

- Be separated into key load forecast components including a large load layer
- Be grounded in contractual agreements rather than interconnection queue requests
- Specify annual peak and energy demand
- Include 8,760-hour chronological profiles
- Capture large load–specific weather dependence

### (A3) Evaluating the incremental impact of large load flexibility.

By performing capacity expansion modeling with and without large load flexibility, planners can isolate and quantify flexibility's impact on total capacity needs, portfolio costs, and emissions.

## For Resource Planners at Vertically Integrated Utilities

Within vertically integrated utilities, planners have several opportunities to better incorporate and evaluate large loads as part of integrated resource planning and utility procurement processes. At vertically integrated utilities, people involved in the integrated resource planning process who are contracting for supply-side resources, development of new rates and contract negotiations with large loads, and the regulatory process can adjust their processes to include:

### (B1) Planning and procuring large load flexibility up front.

Long-term planning models used as part of integrated resource planning and utility procurements can incorporate flexibility options, allowing them to provide system value and compete against other capacity options.

### (B2) Modeling multiple planning scenarios to isolate the capacity requirements attributable to large loads.

When planners model a business-as-usual case without new large load growth and one or more large load–growth scenarios, they can isolate the capacity additions that are attributed to large loads. These results can inform cost allocation of new resource needs attributable to large loads. If, in addition, planners also model scenarios with and without flexibility, these can also inform accelerated interconnection timelines for large loads that make flexibility commitments.

### (B3) Incorporating flexible-load accreditation (e.g., ELCC) in integrated resource plans.

By quantifying capacity contributions, planners can characterize flexibility across a range of parameters to capture its operational limits and system value. These could include the size of the flexible load (in MW), the frequency of deployment (how often flexibility can be called upon in a year), and the duration of each event (hours per call), each of which may drive different accreditation classes.

### (B4) Soliciting flexibility in negotiated tariffs and exchanging faster interconnection for flexibility.

Accelerated speed-to-power via faster interconnection timelines is a powerful incentive for large load types such as data centers. If modeling shows that flexibility can avoid or defer capacity additions that would otherwise be necessary to maintain the system's PRM requirements, resource planners can design flexibility requirements to allow flexible large loads to interconnect more quickly than inflexible loads that would require long-lead-time capacity additions.

## For ISO Market Operators and Planners

Many of the same considerations that apply to vertically integrated utility planners also extend to personnel at ISOs, although the wholesale market framework introduces additional considerations. In particular, ISOs are tasked with designing market rules and accreditation mechanisms that enable large loads to participate in resource adequacy programs while preserving efficient market price signals and investment incentives. ISO market operators and planners can adapt their processes to effectively integrate large loads into resource adequacy constructs by:

### (C1) Clarifying the treatment of on-site generation and "bring your own new generation" (BYONG) options in capacity markets.

ISOs need to clearly define how behind-the-meter or co-located generation associated with large loads will be treated for accreditation and performance obligations. If the on-site generation is dispatchable by the system operator and can reliably offset grid demand during scarcity events, it should be eligible for capacity accreditation similar to peaking units. However, strict performance and telemetry criteria must be in place to ensure that these resources can be relied on during system stress events and that their participation does not compromise overall transparency or reliability.

### (C2) Developing clear accreditation rules and participation pathways for large load flexibility in capacity markets.

ISOs will want to establish standardized, transparent rules that define how large load flexibility can be accredited as a firm capacity resource. This will include performance standards, testing protocols, and obligations under scarcity conditions. When properly defined, these rules will enable large loads to participate on equal footing with other capacity resources, holding firm performance commitments and contributing to reliability during tight system conditions.

### (C3) Incorporating flexibility options into interconnection and capacity market processes.

ISOs can provide meaningful incentives for flexibility by linking interconnection timelines and capacity obligations to flexible performance commitments. For example, if large loads are allowed to receive accelerated interconnection or reduced capacity requirements in exchange for proven, contractually binding flexibility, this can help align load growth with system needs while maintaining fairness and transparency in market participation.

### (C4) Designing non-voluntary load curtailment programs carefully to avoid distorting capacity investment signals.

In some regions, non-voluntary curtailment provisions are being used as a temporary reliability option until new generation and storage resources can come online. While these measures can protect near-term adequacy, ISOs must ensure that they do not inadvertently suppress energy prices or reduce apparent capacity demand, as this can artificially depress capacity market clearing prices and delay needed investment in new supply resources. Such programs should be limited in duration, transparent in design, and complemented by market mechanisms that signal the need for new capacity additions.

## For Data Centers and Other Large Loads

Data centers and other large loads are key stakeholders that can play a crucial role in directly or indirectly shaping how utilities and grid operators plan, value, and integrate large loads into their systems. They can support this important role by:

### (D1) Planning for flexibility early in the large load design and development process.

Different data center designs lend themselves to different levels of flexibility potential. For example, data centers with emergency-only back-up diesel generators may be constrained by the large load's back-up generator run-hour or emissions limits much more than those configured with multiple fuel options or on-site battery storage specifically designed to provide grid flexibility support alongside emergency back-up power. Other design and contracting decisions—such as pursuing off-site flexibility options through contracted resources or accreditation of load and resources together as an "energy park"—would ideally be planned for up front.

### (D2) Ensuring grid operator visibility through robust control architecture and telemetry configurations.

Establishing reliable communication channels between large loads and the system operator is important. Deploying real-time telemetry, status monitoring, and control interfaces helps establish reliable communication channels between the two. This will allow operators to observe load behavior and, where appropriate, dispatch flexibility. Robust visibility supports accurate forecasting, operational coordination during stress events, and the ability to verify performance when flexible load participation is credited in resource adequacy frameworks.

### (D3) In bilateral contracts with utilities, providing firm flexibility commitments with clear limits.

When negotiating service agreements, large loads can commit to predefined, measurable flexibility, such as by establishing a maximum number of calls per year, minimum notice times, duration limits, annual flexibility budgets, or other terms. These firm flexibility commitments can be incorporated into custom tariffs or bilateral contracts, enabling utilities to count this load as dispatchable in their resource adequacy planning and reducing their capacity requirements.

### (D4) Considering utility contracts that monetize avoided capacity and support faster interconnection.

Contracts can create win-win outcomes for large loads and utilities when they quantify and accurately compensate large loads for the value of flexibility they provide—or conversely, require the costs of new capacity required to

---

**Contracts can create win-win outcomes for large loads and utilities when they quantify and accurately compensate large loads for the value of flexibility they provide—or conversely, require the costs of new capacity required to meet inflexible large loads to be borne by those specific large load customers.**

meet inflexible large loads to be borne by those specific large load customers. Utility contracts can enable expedited timelines for large load customers that reduce the need for new firm capacity while providing utilities with certainty around the flexibility provisioned by large loads, allowing utilities to treat flexible loads as part of their resource stack.

**(D5) Working with utilities and regional grid operators to provide representative, detailed operational data to support effective resource adequacy modeling.**

When large loads can share data on hourly load profiles, temperature dependence, planned outages or maintenance periods, flexibility characteristics, on-site resource configurations, and other attributes, planners can model large loads more accurately. These data support segmentation, probabilistic analysis, and accurate capacity accreditation, and help to ensure that large loads are neither over- nor under-accounted for in resource adequacy assessments.

## For Regulators

Given that large loads drive significant new infrastructure and capacity needs, regulators play a central role in ensuring that planning, cost recovery, and rate structures remain fair and equitable. Regulators can achieve these goals by:

**(E1) Isolating the costs of capacity attributed to new large loads to ensure that cost allocation and retail rate structures are fair to all utility customers.**

By requiring utilities to model scenarios with and without new large loads, regulators can identify the incremental capacity and infrastructure investments driven by these loads. This analysis supports cost causation principles by ensuring that new large loads, rather than existing customers, bear a proportionate share of the costs they introduce. These findings can guide the design of targeted rate structures, flexibility incentives, or capacity contributions that align with the loads' system impacts.

### (E2) Using results from integrated resource planning modeling that assess incremental large load impacts–with and without flexibility–to inform cost allocation and tariff design.

Analyses performed as part of resource planning can help identify and isolate the incremental costs of capacity required to serve new large loads. However, procedural and informational barriers between integrated resource planning, tariff design, and rate case dockets mean that relevant insights from integrated resource planning modeling that can inform cost allocation in other proceedings may not be considered. Regulators can develop pathways for relevant data and analyses performed as part of integrated resource planning to enter the record for other relevant proceedings.

### (E3) Assessing whether traditional/existing demand response programs provide appropriate incentives for desired flexibility responses.

Regulators can review whether current demand response programs and frameworks are sufficient to induce the desired operational response from large loads. This review includes evaluating whether program structures, incentives, and dispatch protocols are aligned with the needs of large loads, and whether these can be expected to attract meaningful participation from large load categories such as data centers, which are unlikely to respond to traditional price-based demand response incentives. Where gaps exist, regulators may need to modernize program rules or establish new flexibility pathways better suited to these large load categories.

### (E4) Setting goal posts around the types of flexibility that are needed rather than prescribing specific flexibility technologies.

Rather than defining how large loads' flexibility must be delivered—for example, specifying specific load curtailments, behind-the-meter generation, or specific equipment—regulators can instead focus on establishing clear performance criteria such as response time, duration, frequency of calls, annual or seasonal flexibility budgets, and availability during system stress events. This outcome-based approach allows utilities and load customers to identify cost-effective and implementable solutions while ensuring that the flexibility procured aligns with resource adequacy needs. By establishing the goal posts rather than the path, regulators can enable a competitive and adaptive ecosystem of flexible technologies.

> **By establishing the goal posts rather than the path, regulators can enable a competitive and adaptive ecosystem of flexible technologies.**

### (E5) Incorporating flexibility incentives and/or requirements in large load tariffs and rate designs.

When approving or designing tariffs for large loads, regulators can ensure that flexibility is recognized and valued, either through incentive-based mechanisms (e.g., discounted rates or accelerated interconnection for flexible loads) or explicit requirements (e.g., minimum response capabilities or participation in flexible dispatch programs). This approach aligns the economic interests of large customers with system reliability goals and supports more efficient long-term planning and cost allocation.

### (E6) Leveraging utility resource adequacy analysis to set a cap on the amount of large load flexibility.

Resource adequacy studies can quantify how much flexible capacity is needed to maintain reliability under different load growth scenarios. Regulators can use these results to establish reasonable flexibility thresholds for large loads, striking a balance between full firm capacity and unlimited flexibility requirements. This will ensure that large loads contribute meaningfully to system adequacy without being subject to excessive flexibility obligations that may deter investment. A balanced approach provides predictability for developers while supporting efficient capacity planning and procurement.

# Large Load Flexibility in the Planning Process for Robust, Long-Term Solutions

The electric power industry stands at a pivotal moment. For the first time in decades, the pace of load growth is surpassing the addition of new generation and grid infrastructure. This reversal of long-standing trends poses a significant challenge for system planners and regulators. If left unaddressed, the mismatch between demand and available capacity could lead to widespread resource adequacy and reliability concerns. This outcome would be unacceptable for all stakeholders, including existing ratepayers, utilities, large load customers, and regulators. Without deliberate and coordinated action, widespread resource adequacy and reliability concerns remain a very real possibility.

Resource adequacy, and reliability more broadly, should never be assumed. Resource adequacy is the product of intentional, coordinated, and data-driven planning across multiple domains: utility and ISO system planning, market design, regulatory policy, and customer participation. Each plays a critical role in ensuring that the grid remains resilient.

Today, the industry is responding through a combination of reactive, short-term, and transitional measures designed to preserve near-term reliability. These include development moratoriums in constrained areas, non-voluntary load curtailment programs, and BYONG arrangements that allow large loads to offset their capacity obligations with on-site resources. These mechanisms are practical and necessary tools to manage immediate adequacy risks while new capacity additions and transmission upgrades catch up. However, they are inherently temporary solutions while long-term coordinated planning occurs at the system level.

**Flexibility, whether temporal, geographical, contractual, or physical, should be identified, quantified, and integrated as early as possible in the large load development lifecycle.**

The long-term solution lies in systematically incorporating large load flexibility into the planning process from the outset. Flexibility, whether temporal, geographical, contractual, or physical, should be identified, quantified, and integrated as early as possible in the large load development lifecycle. When designed and coordinated with broader utility and ISO planning efforts, this proactive approach can give both large load customers and system planners faster interconnection, more efficient capacity procurement, and more reliable system outcomes.

Achieving these outcomes depends on the continued advancement of probabilistic resource adequacy modeling. These tools must evolve to reflect the realities of modern load growth and system operation by incorporating large load layers explicitly within load forecasts, evaluating a wide range of large load scenarios and sensitivities, and improving large load data inputs. These data inputs include 8,760-hour chronological profiles, weather-dependent demand behavior, segmentation by end use, and defined flexibility options.

Through this modernization, planners can move beyond static assumptions and instead quantify the true capacity requirements and opportunities associated with large loads. The result is a well-planned, transparent, and equitable power system—one that maintains resource adequacy for all ratepayers, provides a clear path for new large load development, and supports the efficient transition to a reliable future grid.

# References

Besuner, P. 2025. *The New Reality of Power Generation: An Analysis of Increasing Gas Turbine Costs in the U.S.* GridLab, Energy Futures Group, and Halcyon. http://gridlab.org/portfolio-item/gas-tubine-cost-report/.

Biewald, B., D. Glick, S. Kwok, K. Takahashi, J. P. Carvallo, and L. Schwartz. 2024. *Best Practices in Integrated Resource Planning*. Synapse Energy Economics and Lawrence Berkeley National Laboratory. https://eta-publications.lbl.gov/sites/default/files/2024-12/irp_best_practices_2024_synapse_lbnl_24-061_0.pdf.

Cohen, J., T. Fitch, and L. Shwisberg. 2025. “Gas Turbine Supply Constraints Threaten Grid Reliability; More Affordable Near-Term Solutions Can Help.” Rocky Mountain Institute. https://rmi.org/gas-turbine-supply-constraints-threaten-grid-reliability-more-affordable-near-term-solutions-can-help/.

Cox, C., A. Schwartz, and D. Stenclik. 2025. *Bringing Data Center Flexibility into Resource Adequacy Planning: A Case Study of NV Energy.* GridLab and Telos Energy. https://gridlab.org/datacenter_flex/.

Daymark Energy Advisors. 2025. *Considerations for Fair and Efficient Large Load Integration.* North Carolina Utilities Commission. Docket No. E-100 Sub 208. October 14. https://starw1.ncuc.gov/NCUC/ViewFile.aspx?Id=6743c572-de60-47be-af18-f1b3065f3406.

E3 (Energy and Environmental Economics). 2024. “Appendix D: Value of Flexibility” in *Virginia Data Center Study: Electric Infrastructure and Customer Rate Impacts.* Virginia Joint Legislative Audit and Review Commission. https://jlarc.virginia.gov/pdfs/presentations/JLARC%20Virginia%20Data%20Center%20Study_FINAL_12-09-2024.pdf.

EPRI. 2025. “DCFlex: Grid Flexibility Needs and Data Center Characteristics.” White paper. https://www.epri.com/research/products/000000003002031504.

ERCOT (Electric Reliability Council of Texas). 2025. “Real-Time Market Operations.” Presentation. https://www.ercot.com/files/docs/2025/09/22/2025_09-Real-Time-Market-Operations.pdf.

ESIG (Energy Systems Integration Group). 2026a. *Grid Integration of Large Loads: Introduction to the Large Loads Task Force, Data Needs, and Flexibility Overview.* A report by the Large Loads Task Force. https://www.esig.energy/reports-briefs/large-load-task-force-introduction.

ESIG (Energy Systems Integration Group). 2026b. *Large Loads: Behaviors, Capabilities, and Limitations.* A report by the Large Loads Task Force. https://www.esig.energy/reports-briefs/large-loads-behaviors-capabilities.

ESIG (Energy Systems Integration Group). 2025a. *Forecasting for Large Loads: Current Practices and Recommendations.* A report by the Large Loads Task Force. https://www.esig.energy/reports-briefs/forecasting-for-large-loads/.

ESIG (Energy Systems Integration Group). 2025b. *Gaps, Barriers, and Solutions to Demand Response Participation in Wholesale Markets.* A report by the Distributed Energy Resources Working Group. https://www.esig.energy/demand-response-in-wholesale-markets.

ESIG (Energy Systems Integration Group). 2021. *Redefining Resource Adequacy for Modern Power Systems*. A report by the Redefining Resource Adequacy Task Force. https://www.esig.energy/reports-briefs/resource-adequacy-for-modern-power-systems.

FERC (Federal Energy Regulatory Commission). 2025a. “FERC Directs Nation’s Largest Grid Operator to Create New Rules to Embrace Innovation and Protect Consumers.” December 18, 2025. https://www.ferc.gov/news-events/news/ferc-directs-nations-largest-grid-operator-create-new-rules-embrace-innovation-and.

FERC (Federal Energy Regulatory Commission). 2025b. “Order Accepting Tariff Revisions.” Docket No. ER25-579-000. https://elibrary.ferc.gov/eLibrary/filelist?accession_num=20250404-3062.

Giacobone, B. 2025. “Can VPPs Unlock Grid Capacity for Data Centers?” *Latitude Media*, October 29, 2025. https://www.latitudemedia.com/news/can-vpps-unlock-grid-capacity-for-data-centers.

Howland, E. 2025a. “No More PJM Data Centers Unless They Can Be Reliably Served: Market Monitor.” *Utility Dive,* November 26, 2025. https://www.utilitydive.com/news/pjm-data-center-interconnection-market-monitor-ferc-complaint/806527/.

Howland, E. 2025b. “PJM Capacity Prices Hit Record High as Grid Operator Falls Short of Reliability Target.” *Utility Dive,* December 18, 2025. https://www.utilitydive.com/news/pjm-interconnection-capacity-auction-data-center/808264/.

Ibanez, E., and M. Milligan. 2014. “Comparing Resource Adequacy Metrics.” Preprint. National Renewable Energy Laboratory. https://docs.nlr.gov/docs/fy14osti/62847.pdf.

Levitt, A., and S. Shparber. 2025. “The Challenges and Opportunities of On-Site Resources in Accelerating Data Center Interconnection and Reliability.” Brattle and Mintz webinar, June 24, 2025. https://www.brattle.com/insights-events/publications/the-challenges-and-opportunities-of-on-site-resources-in-accelerating-data-center-interconnection-and-reliability/.

NEDC (National Energy Dominance Council). 2026. “Statement of Principles Regarding PJM.” https://www.energy.gov/documents/statement-principles-regarding-pjm.

NERC (North American Electric Reliability Corporation). 2025a. “Characteristics and Risks of Emerging Large Loads: Large Loads Task Force White Paper.” https://www.nerc.com/globalassets/who-we-are/standing-committees/rstc/whitepaper-characteristics-and-risks-of-emerging-large-loads.pdf.

NERC (North American Electric Reliability Corporation). 2025b. “Electricity Supply and Demand.” https://www.nerc.com/programs/reliability-assessment-performance-analysis/electricity-supply-demand.

Norris, T. H., T. Profeta, D. Patino-Echeverri, and A. Cowie-Haskell. 2025. “Rethinking Load Growth: Assessing the Potential for Integration of Large Flexible Loads in U.S. Power Systems.” Nicholas Institute for Energy, Environment and Sustainability, Duke University. https://nicholasinstitute.duke.edu/publications/rethinking-load-growth.

PSCo (Public Service Company of Colorado). 2025. “2024 Just Transition Solicitation, Section 2.2 Xcel Energy, Load Forecast 2024–2036.” https://xcelnew.my.salesforce.com/sfc/p/#1U0000011ttV/a/8b000003N3eQ/QodCDDNR0HUF75Rb_nOkwu_MzfeKzPduYws13FjugGs.

PUC of Ohio (Public Utilities Commission of Ohio). 2025. *In the Matter of the Application of Ohio Power Company for New Tariffs Related to Data Centers and Mobile Data Centers.* https://dis.puc.state.oh.us/ViewImage.aspx?CMID=A1001001A25G09B43531I00509.

Ratnayake, A. 2025. “Unlocking AI Potential with Data Center Flexibility.” *Energy Central,* June 12, 2025. https://www.energycentral.com/intelligent-utility/post/unlocking-ai-potential-with-data-center-flexibility-PtPoXIAuRMzs5Ff.

Shehabi, A., S. J. Smith, A. Hubbard, A. Newkirk, N. Lei, M. A. B. Siddik, B. Holecek, J. Koomey, E. Masanet, and D. Sartor. 2024. *2024 United States Data Center Energy Usage Report.* Lawrence Berkeley National Laboratory, Energy Analysis and Environmental Impacts Division. https://escholarship.org/uc/item/32d6m0d1.

Smart Electric Power Alliance and NC Clean Energy Technology Center. 2026. “Database of Emerging Large-Load Tariffs (DELTa).” https://sepapower.org/large-load-tariffs-database/.

SPP (Southwest Power Pool). 2026a. “FERC Approves SPP’s Large Load Connection Proposal.” *SPP blog,* January 15, 2026. https://www.spp.org/news-list/ferc-approves-spp-s-large-load-connection-proposal/.

SPP (Southwest Power Pool). 2026b. “SPP’s High Impact Large Load Interconnection Solutions: Powering Growth, Accelerating Opportunity.” https://spp.org/markets-operations/high-impact-large-load-hill-integration/.

SPP (Southwest Power Pool). 2025. Large Load Stakeholder Engagement Forum, July 1 (virtual meeting).

Vaidhynathan, D., K. Prabakar, G. Martin, A. Ramesh, B. Wheeler, C. Coco, J. Clidaras, M. Monsch, S. Kim, S. Talukdar, and S. Marti. 2025. *Vulcan Test Platform: Demonstrating the Data Center as a Flexible Grid Asset.* National Laboratory of the Rockies. https://docs.nlr.gov/docs/fy25osti/94844.pdf.

Warren, C. 2025. “EPRI’s DCFlex Initiative Explores How Large Flexible Loads Can Support Grid Reliability and AI-Driven Growth.” *EPRI Journal,* August 20, 2025. https://eprijournal.com/flexible-loads-resilient-grids/.

# Resource Adequacy with Large Loads: Planning for Flexibility to Accelerate Integration

**A Report by the Energy Systems Integration Group's Large Loads Task Force**

---

This report is available at https://www.esig.energy/reports-briefs/large-loads-resource-adequacy.

To learn more about the ESIG Large Loads Task Force and the recommendations in this report, please send an email to info@esig.energy.

The Energy Systems Integration Group is a nonprofit organization that marshals the expertise of the electricity industry's technical community to support grid transformation and energy systems integration and operation. More information is available at https://www.esig.energy.

ESIG
ENERGY SYSTEMS
INTEGRATION GROUP