Utility Planning Best Practices

Data Center Load Considerations

Prepared for
Ameren Missouri
1901 Chouteau Ave.
St. Louis, MO 63103

Prepared by
CRA
Charles River Associates
1201 F Street NW, Suite 800
Washington, DC 20004

Date: 09/16/2025

Disclaimer

Charles River Associates and its authors make no representation or warranty as to the accuracy or completeness of the material contained in this document and shall have, and accept, no liability for any statements, opinions, information or matters (expressed or implied) arising out of, contained in or derived from this document or any omissions from this document, or any other written or oral communication transmitted or made available to any other party in relation to the subject matter of this document.

TABLE OF CONTENTS

1. INTRODUCTION

The rapid growth of data centers—driven by AI model training, hyperscale cloud expansion, and digital infrastructure—is fundamentally reshaping traditional load forecast approaches for many utilities across the U.S. In response, utilities are beginning to adapt their Integrated Resource Plans (IRPs) to account for this emergent, high-intensity demand profile and to ensure long-term system reliability and capital investment alignment. However, the pace and scale of this demand surge introduce significant uncertainty into planning processes. The lack of historical load analogs, coupled with the geographically concentrated nature of data center interconnection requests, introduces additional forecast uncertainty and raises critical questions around capacity procurement, transmission planning, and regulatory coordination. CRA has reviewed publicly available IRPs, regulatory filings, and utility practices to identify how data center load is being incorporated and what risk mitigation strategies are emerging as best practices.

2. CHARACTERIZING DATA CENTER LOAD

2.1. Segmenting Load

Data centers differ from other large loads due to their continuous, high-density energy consumption, differences in load flexibility, and demand for very high levels of reliability. Their need for uninterrupted operation drives uptime requirements ranging from 99.9% (“three 9s) to as high as 99.999% (“five 9s”). To properly consider these characteristic differences, it can be valuable to segment data centers apart from other large industrial loads for the purpose of forecasting.

For example, **Arizona Public Service (APS)** segments different loads by class as part of its load forecast procedure. It includes a specific segment which represents data centers. The APS load forecast is developed from several class-level analyses: Residential, Commercial and Industrial <3 MW, Commercial and Industrial >3 MW, EV Charging, Irrigation and Street Light Customers, and Commercial and Industrial – Extra High Load Factor (XHLF).[1] The final category is reserved for loads in excess of 5 MW and with a load factor of at least 92%, such as data centers. For the purposes of the load forecast, this class’s demand forecasts are prepared individually, and developed based on input provided by customer account managers.

Also, as part of its external adjustment for large loads, **Georgia Power Company (GPC)** segments large loads which are to be evaluated by its load forecast adjustment model. These loads are divided into six categories: cryptocurrency, data center, warehouse, battery manufacturing, chemicals, and solar. This allows them to make targeted adjustments to the modelling input ranges based on the known characteristics of data centers, and historic and current probability data gleaned from their pipeline and past data center landscape.[2]

[1] Arizona Public Service 2023 Integrated Resource Plan, Main Document

[2] Georgia Power Company 2025 Integrated Resource Plan, Technical Appendix

Distinguishing large data centers from other large industrial customers enhances the accuracy of electricity load forecasting by enabling more precise load characterization. This differentiation also helps utilities prioritize grid upgrades, as data centers typically feature high load densities, operate continuously 24/7 with minimal demand variation, and maintain consistently high load factors. Moreover, data centers are far from uniform. They can vary significantly in size, energy demand, ramp-up timelines, and sustainability requirements—including preferences for low- or zero-carbon energy sources. Sub-segmenting them allows for more accurate forecasting and better-aligned infrastructure planning.[3]

The three main types of data centers are Private or "Enterprise" data centers, Colocation data centers, and Hyperscale or Big Tech data centers. Private or "Enterprise" data centers are the smallest type, used only to support their owner's core business, like for example medical records storage for hospitals. Enterprise data centers are typically under 10 MW. Colocation data centers can vary drastically in size, typically ranging from 10-100 MW, but they can also be larger, particularly if specializing in AI-related computing. These data centers are characterized by their business model of leasing IT space to end users. These data centers "co-locate" many different customers' IT infrastructure under one roof, providing the building, cooling, fiber optic bandwidth, and security. Finally, hyperscale data centers are the largest type of data center. These data centers serve "Big Tech" brands like Microsoft, Google, Amazon Web Services, and Meta and are typically very large (from 500+MW up to several gigawatts). Hyperscale data centers have historically been differentiated by their sustainability commitments, each targeting 100% carbon-free or clean electricity.

Utilities like **Dominion Energy in Virginia (DEV)** are already segmenting data centers by size and load profile—hyperscale, AI/High-Performance Computing, and colocation—to tailor load forecasts and infrastructure planning to develop dedicated infrastructure plans to support high density AI workloads.[4] **GPC** is working with developers to model load ramp-up curves based on cooling technologies and compute intensity. it segments customers for rate design and, similar to Dominion, infrastructure planning.[5]

The next section outlines how U.S. utilities forecast data center load, including bottom-up approaches that leverage segmentation to inform and refine load projections.

2.2. Forecasting Approaches

Significant expansion of data center development has, in many cases, outpaced utilities' ability to coordinate a strategized approach to forecasting associated load. As a result, forecasting practices vary widely across the industry. Some utilities limit their projections to contracted or under-construction facilities within their service territory, effectively excluding speculative or early-stage projects.[6] This approach helps mitigate over-forecasting risk but may underrepresent future capacity needs. Other utilities—particularly those in regions where data center development is still nascent— rely on bespoke deterministic methods,

[3] 2024 United States Data Center Energy Usage Report, Berkeley Lab (LBNL), December 2024.

[4] 2024 Dominion Energy Virginia Integrated Resource Plan, Virginia Electric and Power Company, October 2024.

[5] Georgia Power Company 2025 Integrated Resource Plan, January 2025.

[6] For example, Dominion Energy South Carolina and AEP Indiana Michigan Power

often applying arbitrary derating factors or “haircuts” to data center projects in the pipeline.[7] While this approach attempts to account for uncertainty, it lacks transparency and consistency, making it difficult to benchmark across jurisdictions. This methodological fragmentation reflects the broader challenge: the absence of a standardized framework for modeling high-density, high-variability loads, which complicates long-term resource planning.

The following subsections highlight key data center load forecasting approaches identified by CRA.

2.2.1. Top-Down Adjustment of Commercial Load Class

In this approach, utilities assume that data centers are a subset of the commercial customer class and apply macro-level adjustments to reflect anticipated growth. Growth assumptions are often based on historic commercial trends, regional economic development forecasts, or assumptions about technology sector expansion.

This method is easier to implement in regions with less data center activity lacking detailed project-level data or containing several projects in early-stage of development. However, it is less precise and may obscure data center-specific impacts and understate emerging demand.

2.2.2. Bottom-Up Deterministic Forecasting

This approach relies on detailed, site-specific information—such as interconnection requests, public filings, permitting activity, and direct engagement with developers—to forecast expected load from known data center projects. Utilities use this data to pinpoint when and where facilities will come online and, through segmentation, model their expected electrical demand (MW) and energy usage (MWh) profiles. The resulting data center load projection is then added, as a peak and energy adjustment, to the traditional, in most cases econometric, load forecast on other customer types.

This method enables highly accurate, location-specific, near-term forecasts, particularly in high-growth regions with active siting and permitting. However, it offers limited visibility into long-term trends or unannounced developments, underscoring the need for complementary forecasting strategies.

Duke Energy Carolinas (DEC) employs a bottom-up methodology to enhance its long-term load forecasting by explicitly incorporating anticipated demand from large-scale economic development initiatives, referred to as “Mega-projects.” These include hyperscale data centers and advanced manufacturing facilities, such as those associated with semiconductor fabrication and electric vehicle production.

These loads are modeled independently from the base econometric forecast to avoid conflating macroeconomic trends with discrete, high-impact developments. To qualify for inclusion, a project must meet two primary criteria: (1) a minimum load threshold of 20 MW, consistent with the statistical forecast error margin, and (2) a demonstrated level of maturity and planning certainty sufficient to justify inclusion in the forecast.

Even when these criteria are met, **DEC** applies two successive discount factors—typically ranging from 30% to 60%—to account for project uncertainty and to mitigate the risk of

[7] For example, Entergy Louisiana

double-counting demand already embedded in macroeconomic indicators.[8] Projects are continuously monitored and evaluated through a structured qualification process that assesses the likelihood of site activation or cancellation.[9]

The other utility that includes a bottom-up approach as part of its data center load forecasting process is **DEV**. It identifies its eight largest or fastest-growing data center customers, modeling each individually. The customers selected represent a significant majority of the utility's Data Center load. All other customers are combined into a ninth segment. This segmentation allows forecasting each data center customer's load using statistical methods and public and confidential customer information (see section 2.2.4).[10]

Xcel Energy's (Xcel's) forecasting of data center load involves a combination of analyzing interconnection requests and assessing project viability. Given the speculative nature of many interconnection requests, Xcel employs a conservative approach by assuming a lower realization rate of proposed projects.[11]

2.2.3. Stakeholder-Informed Forecasting

This methodology builds on the bottom-up approach by incorporating project-specific intelligence while extending beyond publicly disclosed developments. Utilities work directly with data center developers, industry associations, state energy offices, and economic development agencies to access semi-public or confidential data under non-disclosure agreements (NDAs). This data may include anticipated load profiles, deployment timelines, site clustering behavior, and infrastructure co-location plans. By integrating these insights, utilities can construct more granular and forward-looking forecasts that reflect near-term market dynamics.

Ongoing stakeholder engagement is central to this approach, enabling iterative updates that align forecasts with evolving project pipelines and regional development trends. This methodology enhances the credibility and temporal relevance of forecasts, particularly in high-growth zones where traditional queue-based methods lag market activity. However, its effectiveness hinges on the strength of inter-institutional relationships and the willingness of counterparties to share sensitive information. Even with robust collaboration, data gaps and project attrition risks remain, limiting the completeness and certainty of the resulting forecasts.

Currently, **GPC** actively engages with data center developers to understand their specific power requirements, timelines, and potential challenges. This collaboration includes discussions on site selection, infrastructure needs, and energy consumption patterns. GPC uses this engagement as a key input for its Load Realization Model, specifically considering existing relationships and the progress of customer discussions in determining the probability of a large load that has decided to locate in Georgia will select GPC as its electric service provider (see section 2.2.4).

---

[8] A similar methodology has been adopted by ERCOT in their April 2025 demand forecast where data center load forecasts are adjusted down to reflect actual load relative to expected (49.8%) based on 2022-2024 period and Officer Letter load further adjusted to 55.4% based on recent experience.

[9] 2023 Duke Carolinas Resource Plan, Appendix D

[10] Dominion Energy Virginia Data Center Demand Forecasting Process

[11] Xcel Energy Investor Presentation

In addition, since its 2023 IRP Update, **GPC** has been submitting quarterly large load economic development reports to the Georgia Public Service Commission (PSC). These reports detail the pipeline of large load projects, including data centers, tracking their progression from initial interest to committed service agreements and construction milestones. As of Q1 2025, the long-term large load pipeline had grown to approximately 52 GW, with 8.3 GW from committed customers.[12]

Using this approach, **DEV** supplements its statistical forecasting models with direct input from data center developers, including construction schedules, anticipated power requirements, and operational parameters. This developer-provided intelligence is used to calibrate and refine model outputs, ensuring alignment with real-world deployment expectations. DEV also monitors project progression through key milestones—such as substation engineering Letters of Authorization (LOAs), construction LOAs, and executed Electrical Service Agreements (ESAs)—to assess project viability and mitigate the risk of overbuilding infrastructure for speculative or stalled developments. This continuous feedback loop enables DEV to dynamically update its forecasts and synchronize infrastructure investments with verified demand signals.

Based on inquiries from data center developers and perceived increases in the potential for new loads from hyperscaler data centers, the **Northern Indiana Public Service Company (NIPSCO)** developed a load reference scenario that includes two to three data center projects, or up to 2,600 MW of new load.[13]

APS, as mentioned above, also relies on customer account data to develop data centers' individual load forecasts and apply derating factors to projects in earlier stages of development. APS notes that it "would be unlikely to find reliable independent causal variables to substitute for this method." though they also acknowledge that the customer class is relatively new, and their forecasting capabilities will evolve over time.[14]

While developer engagement is a valuable input to load forecasting, overreliance on it presents challenges due to inconsistent commitment levels and limited data transparency. Developers often withhold critical details—such as whether a project is soliciting service from multiple utilities—raising the risk of double counting and complicating territorial planning. The lack of end-user disclosure further undermines segmentation-based forecasting approaches that rely on understanding load characteristics by sector or application. Increasingly, utilities are also seeing speculative proposals without confirmed end users, reducing data quality and project viability. These information asymmetries introduce significant uncertainty into both short-term infrastructure planning and long-term resource adequacy analysis

2.2.4. Probabilistic Modeling

This forecasting methodology employs probabilistic modeling to simulate a distribution of potential data center build-out scenarios over time. Rather than relying on point forecasts, the model executes thousands of Monte Carlo simulations, varying key parameters such as site commissioning dates, facility size, geographic location, and operational load profiles. This stochastic approach enables planners to quantify uncertainty, assess tail

[12] GPC Large Load Economic Development Report, Q1 2025 PD

[13] Report on NIPSCO's 2024 Integrated Resource Plan, April 2025.

[14] Arizona Public Service 2023 Integrated Resource Plan

risks, and evaluate system flexibility requirements under a range of plausible futures. It is frequently integrated with resource adequacy studies and reliability risk assessments to inform long-term capacity planning.

This technique is increasingly adopted by utilities operating in high-growth, mature data center markets where deterministic methods fall short in capturing volatility and scale. Its primary value lies in its ability to quantify risk and support robust decision-making under uncertainty. However, it is highly data-intensive, computationally demanding, and sensitive to input assumptions. As such, rigorous validation, transparent documentation, and stakeholder alignment are critical to ensure credibility and regulatory acceptance.

GPC is utilizing probabilistic modeling as an external adjustment to its base forecast methodology. Starting with its 2023 IRP Update, GPC incorporated its Load Realization Model because the current level of growth would not otherwise be captured in the historical trends which underly the baseline forecast and would not provide a clear view of large load demand.[15] The Load Realization Model is a probabilistic model that uses Monte Carlo simulations to calculate load percentiles based on the pipeline project's commercial operation date (COD), year-by-year ramp-up trajectory, and the total announced load, as well as the project's class, commercial or industrial, and segment. GPC models the binary probability of each the project selecting the state of Georgia, selecting GPC as its electric service provider, and successfully reaching its COD. If a modeled project satisfies those requirements, it is included in the portfolio, but GPC further considers the probability of the project's metered load varying from its announced load, and the probability and impact of project delays to the ramp-up timing and COD date. GPC uses a Monte Carlo Simulation to run 100,000 repeated sampling iterations, and uses the median, or P50 value, in its most recent RFP.[16] [17]

The organic load growth threshold for industrial load growth is 45 MW, according to GPC's 2025 IRP.[18] Compare this to 7,300 MW of committed large load customers, customers which have signed a Request for Electric Service from Georgia Power.[19] This external adjustment is not unprecedented, as GPC has made external adjustments to its load forecasts in the past for non-data center large loads which represented novel industries for the state, including for example in the 2016 and 2019 IRPs to account for the addition of a new large LNG facility.[20]

2.3. Scenario-Based Planning

Utilities construct multiple load growth scenarios—typically baseline, moderate, and high—to reflect varying trajectories of data center expansion. These scenarios incorporate assumptions around AI-driven compute demand, hyperscaler investment cycles, and regional policy or incentive structures. The resulting load profiles are integrated into

[15] Georgia Power Company's 2023 Integrated Resource Plan Update Docket No. 55378

[16] GPC 2025 IRP Technical Appendix

[17] GPC previously reported the p95 value in its 2023 IRP Update

[18] Georgia Power Company's 2025 Integrated Resource Plan Docket No. 56002

[19] GPC 2025 IRP

[20] Georgia Power Company's 2025 IRP – Technical Appendix Volume 1, Section 1 Load and Energy Forecast

capacity expansion and resource adequacy models to evaluate system resilience and infrastructure needs under uncertainty.

Developing credible scenarios requires clearly defined assumptions, rigorous documentation, and coordination with key stakeholders, including developers, regulators, and economic development agencies. This collaborative approach ensures scenario inputs reflect both market intelligence and planning realism.

Xcel Energy uses scenario-based planning for its load forecasting, developing a base case, with low and high sensitives. The base case in the 2024-2040 Upper Midwest Resource Plan uses econometric analysis to develop jurisdictional sales forecasts based on customer type.[21] In order to reflect data center demand, Xcel makes forecast adjustments for several categories including Large C&I, i.e. data centers.[22] While the exact methodology of the adjustment is not shared, Xcel notes that it uses a discrete adjustment in place of the prior uncertainty modeling method with Monte Carlo simulations, and provides some detail on the assumptions made regarding data centers included in the forecast. The base model includes one presumably big tech site, with several smaller additions. The high sensitivity assumes data center customer expansions on top of the base case, and the low sensitivity maintains the big tech site and some smaller additions, at a more conservative magnitude.

DEV also develops high, medium, and low data center load scenarios to include in its IRP. The process starts by deriving the high scenario, for this, **DEV** employs nine customer segments and data-driven approach to forecast electricity demand from data centers, which represent a rapidly growing share of its load. For each segment, Dominion develops a high-case forecast using three statistical techniques: (1) linear regression of peak demand; (2) polynomial regression of demand; and (3) linear regression of energy sales to peak demand, resulting in 27 different forecasts (i.e., three models for each customer segment). After, DEV selects the appropriate demand model for each segment based on customer-provided intelligence.[23] Finally, DEV uses historical monthly metered data to create the forecasted demand values by month within each year.[24] For the low scenario it uses historical metered data to develop six different statistical models of the overall industry, then these six models are averaged to develop the “low” forecast. Finally, DEV takes an average of the by-customer segment (“high”) and aggregate (“low”) forecasts to derive the “medium” scenario which is submitted to PJM.

Besides considering the set of data centers projects included in its reference scenario, **NIPSCO** also developed an emerging high load sensitivity to incorporate up to six potential data center projects for a total of 8,600 MW.[25] Although, as stated by the utility, “Such load additions are not attributable to a specific customer(s) but represent NIPSCO’s attempt to reasonable [*sic*] estimate total load additions that may come to fruition under various future

---

[21] Including Residential, Small Commercial and Industrial, and Large Commercial and Industrial.

[22] Other forecast adjustment categories include energy efficiency, BTM solar, and Beneficial Electrification

[23] DEV’s description of this process provides example of the information received by customers. If none of the three models aligned with customer intelligence as to future business growth, then an adjusted growth curve is used.

[24] DEV adjusts the initial MWh forecast by applying the historical industry average load factor to the selected model to derive the MWh forecast. The final step taken was the removal of retail choice MWh.

[25] Report on NIPSCO’s 2024 Integrated Resource Plan, April 2025.

states of the world.", it allows NIPSCO to capture uncertainty in its IRP and evaluate system needs under different load growth trajectories.

2.4. Location-specific Considerations

It may also be valuable to incorporate location-specific information into load forecasting and planning, especially in constrained transmission zones or load pockets. Such an approach would help ensure that load forecasting accurately reflects the unique characteristics and limitations of the local grid, which may factor into data center siting decisions, project timelines and potential delays, and ultimate probability of reaching COD. While utilities do have a responsibility to serve customers, data centers are driven by speed to market and likely to self-select out if their desired location requires significant upgrades in order to be viable. On the other hand, locations with sufficient capacity to support data centers quickly may drive increased loads in a specific location due to the cluster effect and need to be reflected by a steeper load forecast. In addition, location-informed forecasting can help prevent double counting if two or more developers are evaluating nearby sites and there may only be capacity to support one of them. Ultimately, forecasting that is informed by location-specific inputs will help ensure the most realistic results and prevent over- or under-estimation of the demand.

There is limited evidence available in public filings for load forecasts which integrate full-scale location considerations with specific regard to data centers, but location or jurisdiction-based modeling is somewhat common. Xcel Minnesota partially incorporates location-specific considerations, as they recognize a redacted "especially promising location" in their service territory, which is very likely to be developed into a data center. They include development of this site in all three data center outlooks, base, low and high.

Facilitating development of large projects through land availability and structured integration processes has also driven data center load growth. For example, **Alliant Energy** utilities in Iowa and Wisconsin are well-positioned for data center development due to their strong energy infrastructure, competitive economic environment, and strategic partnerships. The utility has invested heavily in reliable transmission and distribution systems, including collaboration with ITC Midwest, to ensure consistent and resilient power delivery. Projects like the $750 million QTS data center in Cedar Rapids highlight Alliant's ability to support large-scale developments with shovel-ready sites and utility coordination. In addition to infrastructure, **Alliant Energy** offers economic incentives and cost predictability that attract large commercial customers through its "individual customer rate", or ICR, construct. .[26] The utility's five-year rate freeze in areas like Cedar Rapids provides pricing stability, while economic development programs further enhance project feasibility. Alliant's shift toward a more sustainable energy mix also appeals to data center operators seeking to meet clean energy goals.

To sum up, the diversity of data center development across regions has led to a wide spectrum of forecasting methodologies, each with trade-offs in precision, scalability, and applicability. In mature markets—such as Northern Virginia or Atlanta—utilities have access to detailed project-level data, including interconnection status, load profiles, and construction timelines. Allowing the combination of deterministic baselines with probabilistic overlays that account for project maturity, developer credibility, and regional market signals. All these enable more granular,

[26] Alliant Energy Year-End 2024 Earnings

scenario-based modeling that can differentiate between speculative and committed load, and assess impacts on peak demand, capacity accreditation, and transmission constraints.

Conversely, in emerging or low-penetration markets, utilities often lack visibility into the data center development pipeline. In these contexts, simplified deterministic approaches—such as applying uniform derating factors to interconnection queue volumes or only accounting for potential load requests—are more feasible.

3. PLANNING FOR LOAD GROWTH

3.1. Shorter IRP Cycles

Integrated Resource Plan requirements vary by state, but in general, where required at all, they are mandated in multiannual cycles which vary from 2 to 5 years. In some jurisdictions, these full IRPs may be complemented by mandatory annual IRP updates, but in others IRP Updates are only required under specific circumstances or not required at all. These cycles are not necessarily equipped to handle the rapid growth in data center energy demand that the US has been experiencing since the release of ChatGPT in late 2022. As a result, IRPs are evolving to account for the temporal mismatch between resource planning cycles and data center project developments.

Utilities are increasing the cadence of forecast revisions—moving from biennial or triennial to annual or even quarterly updates. This allows planners to respond more dynamically to near-term changes in load drivers, especially large industrial or digital infrastructure developments.

For example, **GPC,** which has been experiencing extraordinary data center growth due its historical positioning as a data center hub pre-ChatGPT, elected in 2023 to provide an optional mid-cycle update to address the rapid growth it was experiencing which had not been reflected in the previous year's IRP. The Georgia Public Service Commission only requires that utilities file an amendment to the IRP if it "anticipates submitting an application for a certificate to construct or purchase a supply-side or demand-side capacity resource, which was not previously approved as part of the IRP, or it finds that other conditions such as an increase in its projected load forecast warrant an amendment." GPC's 2022 IRP forecasted flat load growth, not including any external adjustment to its baseline forecast which resulted in less than 400 MW anticipated load growth between winter 2023/34 and winter 2030/31. But by 2023, Georgia was experiencing rapid economic growth, which resulted in a load growth result 17x greater than was previously forecasted, 6,600 MW by winter 2030/31, and elected to file an IRP Update.

Since its 2023 IRP Update, GPC has also filed quarterly Large Load Economic Development Reports with the Georgia Public Service Commission to continue to foster a constructive relationship with regulators and reflect the dynamic market conditions it is operating under.

Similarly, **DEC** issued a Fall 2023 IRP update following its Spring 2023 filing to account for a surge in new economic development commitments—including manufacturers, electric transportation, data centers, and advanced cloud computing—which collectively added approximately 2 GW of incremental peak load. By Q4 2024, DEC announced an additional 2 GW of new data center agreements, rendering the most recent IRP load assumptions potentially outdated and underscoring the need for more agile, iterative planning processes.[27]

3.2. Accelerated Procurement of Flexible and Dispatchable Resources

In response to unprecedented load growth, utilities are rapidly evolving their procurement strategies to ensure system reliability, affordability, and alignment with long-term clean energy goals. A common approach is the use of all-source requests for proposals (RFPs), which allow utilities to competitively evaluate a range of technologies—including solar, storage, natural gas, and hybrid systems—based on cost, availability, and system need. Many utilities are also accelerating the addition of dispatchable resources such as natural gas peakers or reciprocating engines to address near-term reliability concerns, often selecting units that are hydrogen-capable or convertible in the future.

Both, **GPC** and **DEV** have issued all-source RFPs to meet capacity needs by early 2030s. In its outstanding RFP, GPC is open to various resource types, including renewable energy, storage, and traditional generation sources and seeks to procure ~9 GW multiple resource types to address 2029–2031 capacity needs. Similarly, DEV issues annual RFPs targeting utility-scale solar, storage, and hybrid projects. The 2023 RFP sought over 1 GW of new resources, while the 2024 RFP seeks 650 MW.[28]

Alternatively, **Entergy Louisiana** is pursuing an expedited route to dispatchable generation, filing directly with the commission in Louisiana for expedited Certificates of Public Convenience and Necessity (CPCN) for two natural gas-fueled power plants, one 1,500 MW facility in northeast Louisiana to serve a new $5 Billion Meta data center, and another 754-MW separate but related plant in southern Louisiana. While this mechanism goes around the traditional IRP process in the state, Entergy argues it is necessary to meet the immediate and substantial energy needs of data centers, which will facilitate, they say, much needed economic development in the State.

Also, **DEC** is planning the additions of 3.6 GW of new gas-fired capacity, including combustion turbines and combined-cycle units by 2035, and **Xcel** is adding gas plants that are convertible to green hydrogen for future decarbonization.

3.3. Other Supply Option Strategies

In addition to issuing RFPs and fast-tracking procurement of dispatchable resources, utilities are extending the lives of existing assets—particularly coal and gas units—through operational modifications or delayed retirements to ensure sufficient firm capacity while newer resources are developed. While some utilities are investing in utility-owned renewables and battery storage projects to maintain control over system costs and performance.

Hybrid projects, especially solar plus storage, are increasingly favored for their ability to provide both energy and capacity value. Several utilities are also exploring or initiating development of emerging technologies such as small modular nuclear reactors[29] and long-

[27] Q3 2024 Duke Energy Corp Earnings Call

[28] 2024 Dominion Energy Virginia RFP

[29] Dominion Energy Virginia and TVA.

duration energy storage to meet decarbonization goals and provide clean firm capacity. Overall, the prevailing strategy emphasizes flexibility and scalability—phasing in new resources based on actual load realization and aligning procurement with location-specific growth such as data center clusters or industrial development zones.

Besides all-source and renewables RFPs,[30] **GPC** has requested coal and gas plant retirement deferrals and is evaluating gas conversions to maintain dispatchable capacity. The utility is also expanding its use of hybrid solar-plus-storage projects to enhance flexibility and system reliability, expand implementation of battery energy storage systems (BESS), distributed energy resource (DER) and demand response (DR) programs, as well as build simple cycle combustion turbines (CTs) at existing generation sites.

3.4. Expanded Demand Side Options

Traditionally, data centers have not been considered flexible loads and have largely remained outside traditional demand response or load curtailment programs. However, large-scale operators are increasingly exploring Geographic Load Balancing (GLB), where computational workloads are dynamically shifted across multiple sites based on regional grid conditions, energy prices, or carbon intensity. While GLB offers potential for grid-interactive flexibility, its effectiveness is constrained when multiple regions face simultaneous stress, limiting redispatch options.[31]

Emerging AI-driven workload orchestration tools are enhancing the ability of data centers—particularly those supporting cloud and AI applications—to modulate power demand in near real time. These tools enable more responsive load management, offering utilities new avenues for integrating data centers into grid flexibility strategies.[32]

Some operators are already optimizing workloads for latency and IT efficiency, which may incidentally align with grid needs and ease the transition toward more active participation in demand-side programs. Realizing this potential requires deeper coordination between data center operators, utilities, and ISOs. Initiatives like EPRI's DCFlex[33] —a collaborative effort involving over 40 stakeholders including Duke Energy, PG&E, and New York Power Authority (NYPA)— are actively exploring how data centers can support grid reliability and operational flexibility.

Additionally, some data centers may be exploring partial colocation of backup generation assets. In these configurations, facilities offset grid demand using on-site generation rather than curtailing load, providing a form of dispatchable capacity that could be integrated into broader grid planning frameworks.

---

30 GPC will still proceed with the All-Source RFP outlined in the 2022 IRP and anticipates procuring ~9 GW of additional capacity to meet system requirements by winter 2030/31 between the All-Source RFP, ~1 GW of retirement extensions, 380 MW of uprates, and ~1 GW via a Renewable RFP.

31 Opportunities and Challenges for Data Center Demand Response, working paper

32 Emerald AI LFLTF Background, March 2025.

33 Electric Power Research Institute (EPRI) Data Center Flexible Load (DCFlex) Initiative Overview

4. ADDRESSING INFRASTRUCTURE AND INTERCONNECTION CHALLENGES

4.1. Coordination with Transmission Planners

Transmission planning is a critical constraint in accommodating large, energy-intensive loads like data centers, primarily due to the mismatch in development timelines. While transmission projects typically require 3 to 6 years to permit and construct, data centers can be operational in under 18 months.[34] This temporal disconnect means that much of the transmission infrastructure currently under development was likely scoped before the recent surge in AI-driven data center demand, creating potential bottlenecks in delivering capacity where it's now urgently needed. CRA has seen that data centers frequently site along existing transmission corridors, but the rapid clustering of high-density loads in these areas can strain grid infrastructure that was not designed with such demand in mind. The disconnect between fast-paced data center development and the much longer timelines for transmission expansion also means that projected load growth may outpace the grid's ability to deliver adequate capacity, potentially creating localized reliability and congestion challenges.

Coordination with transmission planners will help better inform the transmission need and planning process as load forecasts evolve, and in the reverse, can inform project success modeling inputs if it is clear on the transmission side that projects will not be able to come online as quickly as they would like, and may fall off due to the long wait times for transmission.

In organized regional markets, utilities collaborate closely with Regional Transmission Organizations (RTOs) to ensure that new resource additions and forecasted load growth are integrated into long-term transmission planning. For example, MISO's Long-Range Transmission Planning (LRTP) initiative supports the development of multi-value transmission projects designed to accommodate evolving system needs, including large-scale load growth and renewable integration.[35] Examples of other RTOs are highlighted in the table below.

RTO	Planning Initiative	Key Features	Load Integration
MISO	Long-Range Transmission Planning (LRTP)	Supports multi-value transmission projects, accommodates AI-driven data center growth and renewable integration.	Integrates data center growth scenarios into multi-value transmission projects.
PJM	Load Analysis Subcommittee, Planning Committee, and Regional Transmission	Enhances long-term load forecasting, projects significant growth in energy demand, incorporates stakeholder input.	Structured processes for large load adjustments. Load growth is factored into baseline reliability studies and market efficiency analyses

34 LLTF April Meeting & Technical Workshop Presentations

35 MISO Long Range Transmission Planning (LRTP)

	Expansion Plan (RTEP)[36]		
SPP	Integrated Transmission Planning (ITP)[37]	Scenario-based modeling of high-impact loads, stakeholder-driven approach, informs regional transmission expansion.	Models speculative data center developments to ensure system reliability under various future conditions.

Local transmission projects are increasingly being scoped and fast-tracked to support high-growth zones, particularly data center clusters and industrial development corridors. Utilities often conduct injection studies to evaluate the feasibility of new load interconnections and identify necessary system upgrades. **DEV**, for example, collaborates closely with data center developers to align transmission expansion timelines with anticipated load deployment. These localized upgrades are subsequently submitted as supplemental projects through PJM's Transmission Expansion Advisory Committee (TEAC) process for regional coordination and review.[38]

DEC conducts transmission assessments on an annual basis as part of its comprehensive planning process. This annual assessment evaluates transmission needs using seasonal peak load forecasts, contingency studies, and power flow modeling over a 10-year horizon.[39] In addition to the annual assessments, Duke Energy Carolinas engages in a triennial (every three years) strategic planning process known as the Multi-Value Strategic Transmission (MVST) planning process. The MVST process incorporates scenario-based analyses to account for different possible futures.[40]

GPC conducts annual transmission assessments as part of its comprehensive planning framework, covering both near-term and long-term horizons in alignment with regulatory requirements.[41] In addition to its internal assessments, GPC participates in the Internal Joint ITS Planning Process—a coordinated effort among the four co-owners of Georgia's Integrated Transmission System (ITS)—to plan and prioritize high-voltage grid upgrades.[42] GPC also engages in the Southeastern Regional Transmission Planning (SERTP) process, a collaborative multi-utility initiative that planning forum that facilitates regional coordination. Through SERTP, stakeholders gain access to transmission models, expansion plans, and supporting data, enabling a collaborative approach to addressing regional transmission needs.[43]

36 PJM Load Analysis Subcommittee (LAS), Load Forecast Development , Regional Transmission Expansion Plan (RTEP)

37 SPP Integrated Transmission Planning (ITP)

38 PJM's Transmission Expansion Advisory Committee

39 2023 DEC and DEP IRP – Attachment L

40 Duke Energy Carolinas proposed changes to Multi-Value Strategic Transmission Planning

41 Guideline for Planning the Southern Company Electric Transmission System

42 Georgia's Integrated Transmission System (ITS)

43 Southeastern Regional Transmission Planning (SERTP)

4.2. Load Queue Management

Effective load queue management enables utilities to organize, prioritize, and evaluate interconnection requests. It also provides critical insight into the scale, timing, and geographic concentration of future loads. By identifying project readiness, grouping similar applications, and applying standardized review procedures, utilities can reduce delays, alleviate processing bottlenecks, and ensure fair access for all load customers. This is especially important in high-growth regions, where the volume of hyperscale data center requests has rapidly expanded.

Project pipeline visibility supports scenario-based planning, allowing utilities to assess the grid impacts of pending projects and prioritize investments in substations, feeders, and transmission infrastructure. Transparent queue management fosters stronger utility-developer relationships and informs engagement with Regional Transmission Organizations (RTOs/ISOs) to ensure broader system reliability.

Some utilities adopt clustering or "first-ready, first-served" approaches to encourage project discipline and reduce speculative applications. These practices not only improve administrative efficiency but also contribute to the orderly integration of large-scale data center loads into the grid.

DEV has implemented a structured Large Load Interconnection Process specifically designed to manage high volumes of applications from hyperscale developers. Through this process, DEV engages with developers early, assesses project readiness using standardized criteria, and conducts detailed feasibility studies in collaboration with PJM Interconnection. Queue transparency is maintained through published timelines and status updates, while the utility also provides guidance on grid-constrained zones and potential siting alternatives. As a result, DEV has been able to manage over 10 GW of active and proposed data center load while prioritizing system reliability and minimizing unanticipated infrastructure strain.[44] Their queue management process now serves as a model for other utilities adapting to rapid load growth in high-demand corridors.

DEC has implemented a large load interconnection process to manage the integration of substantial non-residential electric loads into its transmission and distribution systems. This process generally applies to new or expanding loads of 10 MW or greater, though specific thresholds may vary depending on regional or voltage-level considerations. Consistent with practices at DEV, DEC initiates early engagement with customers to define project scope, timelines, and potential system impacts. The interconnection process involves a series of technical studies—Feasibility, System Impact, and Facilities Studies—mirroring the structure of generator interconnection procedures. Customers are typically responsible for the costs associated with these studies and any direct interconnection costs.[45]

GPC actively monitors and reports on the scale and status of its large load project pipeline—including total, committed, and under-construction projects. This information is made publicly available through quarterly filings submitted to the Georgia Public Service Commission (PSC), offering transparency into the utility's strategy for managing substantial new load additions. In addition to high-level pipeline metrics, these reports detail the transmission and distribution infrastructure planning necessary to support new large

[44] Dominion Energy Virginia Data Center Request Process

[45] Duke Energy Carolinas Large Load Interconnection Procedure (LLIP)

customers, ensuring that system investments align with projected in-service dates. The filings also evaluate the economic impact of these projects, including job creation and capital investment, while assessing potential effects on existing customer rates.[46]

Xcel Energy has streamlined its interconnection process to facilitate the integration of large-scale data centers. The company employs hosting capacity analyses to identify optimal locations for new connections and to assess the grid's ability to accommodate additional load. This approach helps in reducing the complexity and time required for interconnection studies.[47]

To address the varied approaches to large load integration nationwide, NERC has launched the Large Load Task Force (LLTF) to assess the reliability impacts of emerging large loads on the bulk power system (BPS). The LLTF will begin by identifying the unique characteristics and risks of these loads, then prioritize and validate them. It will also pinpoint gaps in current planning and operational practices, recommending enhancements to help transmission planners and operators maintain grid stability—particularly in terms of voltage and dynamic performance.[48]

5. RISK MITIGATION STRATEGIES

Utilities can implement a range of mitigation strategies to reduce the risks associated with large load additions. These strategies address planning uncertainty, infrastructure adequacy, financial exposure, and system reliability.

Some of these strategies—such as enhanced forecasting and scenario planning, interconnection and queue management reform, and transmission planning integration—have already been addressed in earlier sections. The following subsections introduce additional strategies intended to expand the spectrum of viable approaches.

5.1. Flexible and Phased Infrastructure Investment

Flexible and phased infrastructure investment is a strategic approach utilities are using to manage uncertainty and reduce financial risk associated with large load additions. Rather than committing to large, capital-intensive infrastructure upgrades based on speculative forecasts, utilities are designing systems that can expand in stages as actual demand materializes. This includes deploying modular resources like small gas turbines, reciprocating engines, and battery energy storage systems (BESS) that can be added incrementally. These resources are quicker to install and can serve as bridge solutions while longer-term generation or grid infrastructure is developed. Utilities also invest in distributed energy resources and demand-side programs that can defer traditional grid upgrades by targeting local reliability and capacity constraints.

On the transmission and distribution side, utilities are phasing upgrades to substations, feeders, and lines based on confirmed customer milestones, such as interconnection payments, permits, or construction progress. Some use mobile or temporary substations to provide near-term service while permanent solutions are underway. In substation design,

46 GPC's Large Load Economic Development Report Filing

47 Xcel Energy Hosting Capacity Analysis Process

48 LLTF April Meeting & Technical Workshop Presentations

utilities often install only a portion of the full buildout initially but plan for future additions by reserving space, installing expandable equipment, and standardizing design templates. These practices ensure that upgrades can be efficiently scaled as demand grows, without the need for major redesign or permitting delays.

Many utilities are also integrating “performance-triggered investment” strategies into their capital plans. These strategies set thresholds—such as a specified megawatt load or utilization rate—that must be met before advancing to the next phase of infrastructure deployment. This prevents overbuilding and aligns spending with tangible load growth. Additionally, utilities coordinate closely with large customers through pre-application reviews and load development discussions to sequence grid investments with customer readiness. When combined with flexible cost recovery mechanisms and tariff structures, phased infrastructure investment allows utilities to balance service reliability, financial prudence, and economic development objectives in an era of increasingly unpredictable large load growth.

Xcel Energy is applying phased infrastructure planning across its service territories in Minnesota and Colorado to support anticipated large load growth from electric vehicle adoption, hydrogen production, and industrial electrification. In its approach, Xcel prioritizes flexibility by combining non-wires alternatives, such as targeted demand response and distributed energy resources, with scalable grid infrastructure. For example, in Colorado, the utility employs a “least-regrets” planning framework to identify transmission projects that provide the most value across multiple load growth scenarios. It also uses modular equipment and standard substation templates to expedite future expansions. Xcel coordinates closely with regulators and regional stakeholders to ensure project timelines and infrastructure investments are adaptable, allowing the utility to meet emerging load without overcommitting capital or jeopardizing system reliability.

GPC has integrated flexible investment principles into its large load interconnection strategy, particularly in response to industrial and data center development in Metro Atlanta and western Georgia. GPC employs expandable substation designs, installing only the infrastructure needed to serve confirmed load while reserving space for future transformers and feeders. Transmission investments are phased and closely coordinated with economic development agencies, with project schedules triggered by customer readiness indicators such as signed interconnection agreements or verified construction timelines. GPC’s capital planning incorporates contingencies to adapt to changes in customer commitments, and the utility leverages its parent company Southern Company's (SOCO) Large Load Review Process to assess system impacts and identify modular investment options. This ensures grid upgrades are cost-effective and responsive to actual customer demand.

DEV has adopted a phased infrastructure investment strategy. it links transmission and substation buildouts to customer development milestones—only proceeding with upgrades once data center developers meet key thresholds such as permitting, financing, and construction progress. The utility utilizes modular substations and mobile transformers to deliver interim service while permanent infrastructure is built out. This staged approach enables DEV to align capital expenditures with actual load realization, reducing the risk of stranded assets. Additionally, the Long-Term Planning Scenarios incorporate these flexible build strategies to ensure scalability and cost control in grid expansion.

6. SUMMARY UTILITY BEST PRACTICES

Summary of Utilities Practices for Large Load Considerations

Utility	Load Characterization	Planning Approach	Load Integration Approach	Risk Mitigation Focus
Dominion Energy Virginia (DEV)	Data Center load sub-segmentation; statistical forecasting model supported by bottom-up, stakeholder-informed approach.	Load Scenario-Based Planning; transmission planning coordination with developers and PJM; Investing baseload emergent technologies (SMRs).	Active load queue management; Structured large load integration process	Phased infrastructure investment; Continue Issuing All-Source RFPs, modular substations and mobile transformers to deliver interim service.
Georgia Power Company (GPC)	Large load segmentation; probabilistic model supported by bottom-up, stakeholder-informed approach.	Periodic transmission planning assessment and coordination with regional transmission entities; Shorter IRP cycles; accelerated Procurement of Dispatchable Resources	Active load queue management; leverages SOCO's Large Load Review Process	Phased transmission infrastructure investment with schedules triggered by customer readiness indicators; Continue Issuing All-Source RFPs; coal and gas plants retirement extensions; expandable substation designs.
Duke Energy Carolinas (DEC)	Large Load segmentation; bottom-up approach.	Economic development-driven Scenario Planning; Periodic transmission planning	Structured large load interconnection process	New generation and transmission infrastructure investment; pursuing

		assessment and coordination with regional transmission entities; Shorter IRP cycles; accelerated procurement of gas-fired resources		"take-or-pay" contracts
Xcel Energy	Bottom-up with derating factors	Load Scenario-Based Planning; accelerated procurement of gas-fired resources	Streamlined large load interconnection process;[49] employs hosting capacity analyses	Investment in new generation, transmission, and distribution system upgrades; prioritizes flexibility by combining non-wires alternatives, with scalable grid infrastructure; strategic contractual agreements

7. CONCLUSION

The scale and speed of data center development are prompting a re-evaluation of traditional IRP approaches. Utilities that adapt with flexible, transparent, and risk-aware planning processes are better positioned to integrate data center loads while maintaining reliability and affordability.

[49] Xcel Energy Interconnection Guideline