Data Center Flexibility: A Call to Action

Improving The Grid With A New Approach
To Data Center Development

Data Center Flexibility: A Call to Action

The objective of *Data Center Flexibility: A Call to Action* is to raise awareness of the untapped energy flexibility that data centers can offer the electric grid, laying the foundation to facilitate a more active and collaborative approach between project development and utility system planning. A shared recognition between each industry's emerging requirements will lead to more sustainable deployment of digital infrastructure in a manner that supports utility system planners at a critical moment in building the next generation of industrial demand domestically.

This white paper comes on the backdrop of opportunities identified at the "Future of Data Centers Summit: Defining The Path Forward For Data Center Innovation, Energy Management, & Sustainability." This event, hosted by Sidewalk Infrastructure Partners ("SIP") in June 2023, brought together dozens of leaders across energy, policy, academia, and technology.

Data Center Flexibility: A Call to Action builds on those discussions and is authored by SIP with the support of reviewers and collaborators from across the data center and energy ecosystems. It is intended to raise awareness and to serve as the first step towards workshops, publications, and other collaborative efforts to further develop and implement the concepts discussed in this paper. SIP invites interested stakeholders to learn more by visiting www.datacenterflexibility.com or emailing connect@datacenterflexibility.com.

SIP is a holding company that builds innovative technology-enabled infrastructure companies and projects that deliver positive social and environmental impact in partnership with the public sector. Its mission is to make infrastructure more efficient, sustainable, and inclusive, rooted in the belief that technology will help to solve some of the world's most pressing social and environmental challenges.

Acknowledgments

While the ideas and principles outlined incorporate the perspectives of hundreds of conversations across several industries, SIP would especially like to thank several contributors who lent their time and expertise in reviewing this paper, including: Anand Ramesh, Arman Shehabi, Chris Hickman, Dr. David Kathan, John Forrester, Mark Dyson, Matt Langley, Max Schulze, Nelson Abramson, and Dr. Stephen D. Comello.

Contributors were not asked to endorse this paper, nor are they responsible for any errors or omissions contained herein. All views and opinions expressed are those of SIP only and do not necessarily reflect the views and opinions of the contributors or their affiliates, employers, or investors.

Published March 2024

Executive Summary

Utilities are facing an inflection point in the evolution of the electric grid. They are managing a near-term surge in large-scale industrial loads, such as onshoring manufacturing and the tripling of US data center capacity over the remainder of the decade, the latter accelerated by the growth of artificial intelligence ("AI"). At the same time, these same utilities are undergoing longer-term transitions to renewable and decentralized energy sources, as access to reliable and affordable power becomes ever more critical with the adoption of electric vehicles and the electrification of buildings and other industrial processes.

Today's project development practices have put utilities in a bind between accommodating these demand side interconnection requests and continuing to reliably and affordably provide clean power to their customers, necessitating a re-evaluation of interconnection standards. The near-term solution has been pushing delivery of power to new large loads far into the future. However, not building these projects domestically in the near-term has significant economic development, innovation, and national security implications, and thus we must coalesce around new engagement strategies.

Data centers, expected to eclipse $100 billion of annual capex spend to bring online approximately 40,000 megawatts ("MWs") of new capacity requirements by 2030, are one of the largest drivers of these new load requirements.[1] Building out a strong digital backbone historically has enabled the United States to be the de facto global leader in innovation. Maintaining and growing this infrastructure – now with the emergence of AI and its massive power requirements – will be critical to cementing its leadership position.

Utility planners do not need to sacrifice power availability, reliability, sustainability and affordability when faced with serving new data centers. In fact, data centers themselves can be positive additions

---

[1] The Economist; Data centres improved greatly in energy efficiency as they grew massively larger; But can this continue into the age of AI?

to the grid if innovations in technology, policy, and business models are harnessed in their design, development, and operation.

Specifically, as energy storage and other distributed energy resources (“DERs”) become more cost effective and reliable as backup power sources to replace diesel generators, and as certain AI workloads that are interruptible increase as a percentage of total compute capacity, data centers can provide large scale flexibility to the grid that can both help utilities manage peak-demand challenges and also provide other valuable grid services. The benefits of this approach include improving resource utilization using existing transmission and/or generation, managing peak demand more sustainably and affordably, and supporting offtake for curtailed and/or additional renewable energy generation resources – all of which will significantly improve a utility’s load factor, the measure of system efficiency for the electric industry that underpins affordability for ratepayers.

In essence, data center projects can (1) provide operational flexibility to meet both customer compute demands and grid services needs, and (2) serve as strategic asset deployments for utilities. By acting as symbiotic developments with the grid that can bring new and evolving energy infrastructure technologies online over time, data center projects can help meet objectives defined in utilities’ capital, resource, and grid planning processes.

This white paper is meant as a call to action and an invitation for collaboration. Realizing the benefits of data center flexibility will require partnerships and co-creation across the utility sector, large users of compute, project developers, and other electric grid stakeholders with a collective approach to evaluating how project development can support energy infrastructure both locally and more broadly. By raising awareness of what is possible today, we will enable more productive discussions and planning at the project execution level.

Introduction

The United States has embarked on a $10 trillion energy transition over the next decade that will include replacing existing fossil generation assets with decarbonized and decentralized electricity generation, specifically wind, solar, and batteries in the near term, to electrify our economy more sustainably.[2] Driven by supportive federal and state policies and decreasing costs, these intermittent renewable resources have almost quadrupled their share of US power capacity from 6% in 2012 to 22% in 2022 and are expected to pass 50% by 2035.[3] Such growth is causing a paradigm shift for utilities to maintain reliable operations. Furthermore, most clean energy targets were established in an era of relatively flat load growth, but three recent trends have caused a dramatic increase in forecasted electrical demand.

The first trend is a continued and rapid increase in data center buildout, as evidenced by the construction of approximately 20,000 MWs of data center capacity over the last two decades. Building capabilities in artificial intelligence is the next extension of this digital backbone of innovation, but this will require ever more compute capacity, and thus power. As a result, the power footprint for US data centers alone is poised to triple from 2022 levels to 60,000 MWs by the end of the decade, representing over $100 billion of annual investment.[4] Demand historically was built in certain markets, such as Northern Virginia where data centers amounted to over 20% of the grid's power load,[5] but are rapidly clustering in new markets and consuming significant power resources in those areas as existing markets face limited power availability for development.[6] In aggregate, go forward data center load could equal up to 7.5% of the nation's projected electricity demand.[7]

The second trend is that in the face of increasing geopolitical turmoil, such as the Russian invasion of Ukraine and ongoing tariff uncertainty with China, the US has passed ambitious legislation (e.g., the CHIPS and Science Act; the Inflation Reduction Act) to bring more manufacturing onshore to bolster the country's domestic supply chain. Since 2021, there have been approximately $465 billion worth of semiconductor, EV, and battery factory projects announced, which are projected to consume significant amounts of power.[8]

The third trend is the acceleration of the overall electrification of the economy, with the transition of passenger and commercial vehicle fleets to electric as a key example. Electric vehicle sales increased more than tenfold over the past decade, reaching 1.4 million in 2023 (or over 9% of vehicle sales).[9] Future growth, underpinned by state and federal incentives targeting 50% sales penetration of EVs by 2030, will continue to require scaled access to power.[10] [11] [12] Another driver is

---

[2] McKinsey; The Net Zero Transition: What it could cost, what it could bring
[3] BloombergNEF; New Energy Outlook 2022
[4] The Economist; Data centres improved greatly in energy efficiency as they grew massively larger; But can this continue into the age of AI?
[5] Prince William Times; Region's sluggish solar can't match surging data center demand
[6] Washington Post; Amid explosive demand, America is running out of power
[7] Boston Consulting Group; The Impact of GenAI on Electricity
[8] Bloomberg; AI Needs So Much Power That Old Coal Plants Are Sticking Around
[9] BloombergNEF; Electric Vehicles Database
[10] National Conference of State Legislatures; State Policies Promoting Hybrid and Electric Vehicles
[11] Internal Revenue Service; Credits for new clean vehicles purchased in 2023 or after
[12] Federal Register; Executive Order 14037: Strengthening American Leadership in Clean Cars and Trucks

the electrification of heating, with sales of commercial and residential electric heat pumps outpacing traditional gas furnaces.[13] Together, these additions in commercial, residential and transportation power demand are expected to add 280 terawatt-hours ("TWh") of annual consumption to the grid by 2030.[14]

The core question now is not only whether or not the country can decarbonize, but whether the US can even power these economy-wide projects at the scale envisioned. Private and public sector leaders are increasingly questioning the ability of existing grid infrastructure and current development practices to meet this aggregate demand. Day-to-day execution of these large-scale projects falls to state and local constituents who must balance practical solutions with decarbonization initiatives. Forgoing or delaying these project development opportunities will have significant negative economic development, innovation, and national security implications.

There are several options that stakeholders can consider in meeting these power demand requirements at the transmission level, each with its own set of challenges from an availability, reliability, sustainability, and affordability perspective. Three of these generation focused options are: (1) delaying the retirement of thermal generation plants (e.g., coal), (2) building additional new natural gas plants, and (3) building large-scale transmission to tie in new utility-scale intermittent renewable generation to the grid.

Figure A

Bloomberg January 25, 2024

> "We do need way more energy in the world than we thought we needed before. We still don't appreciate the energy needs of this technology."

Sam Altman
Chief Executive Officer, Open AI

The Washington Post March 7, 2024

> "When you look at the numbers, it is staggering. It makes you scratch your head and wonder how we ended up in this situation. How were the projections that far off? This has created a challenge like we have never seen before."

Jason Shaw
Chairman, Georgia Public Service Commission

---

[13] Canary Media, Heat pumps outsold gas furnaces again last year – and the gap is growing
[14] National Renewable Energy Laboratory; Electrification Futures Studies

However, another option also exists that is rarely evaluated but which can be an important solution for utilities: industrial load flexibility.

Data centers are one of the largest drivers of industrial demand growth. Historically, data centers have been built to draw continuous power from the grid and with under-utilized energy infrastructure (e.g., backup diesel generators that rarely run). With other energy technologies (e.g., battery energy storage) becoming increasingly cost-effective and reliable as a backup power source, and the ability to have flexibility in the dispatchability of certain AI workloads (e.g., large language training models), intelligently built and operated data centers can provide large-scale flexibility to the grid that has several benefits to utilities, such as improving resource utilization using existing transmission and/or generation, managing peak demand more sustainably and affordably, and supporting offtake for curtailed and/or additional renewable energy generation resources.

The following sections walk through current data center practices today, how data centers can evolve to be more flexible, what benefits a flexible data center can provide, and how to begin unlocking this opportunity.

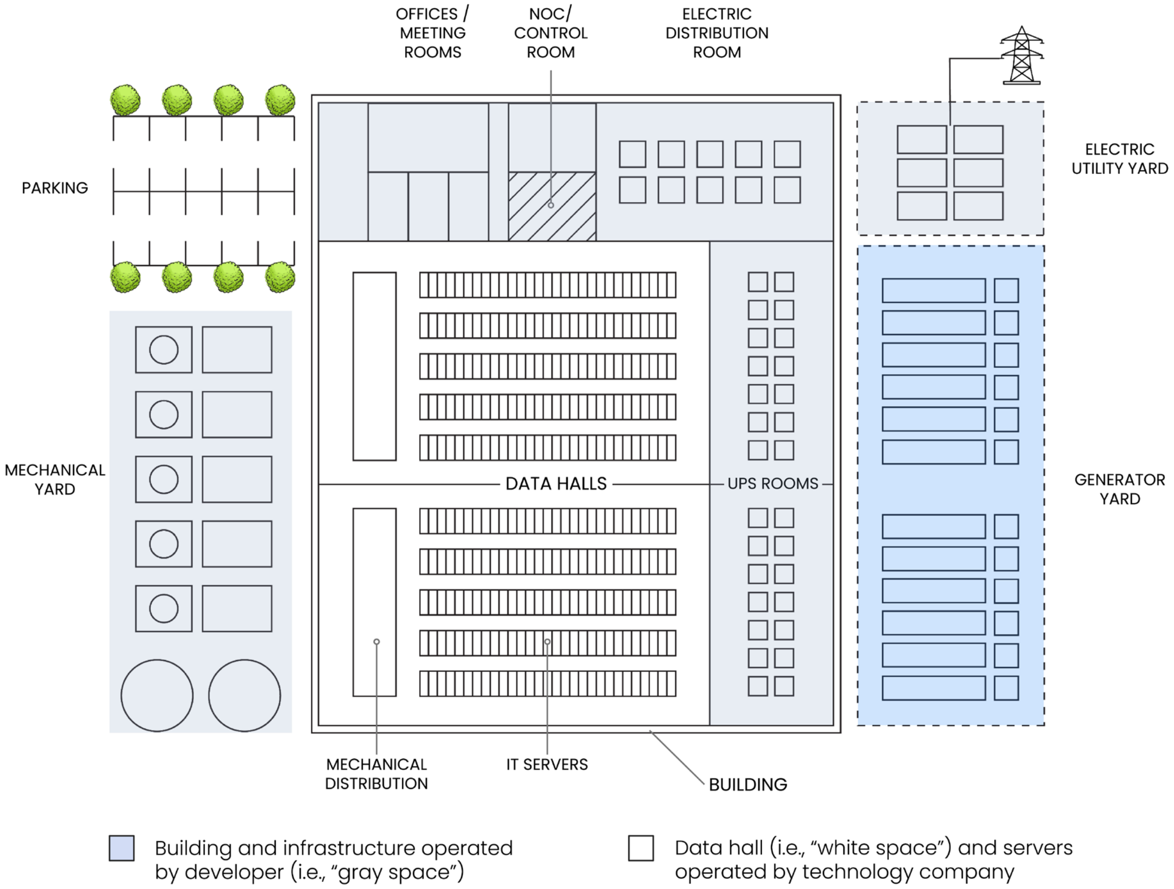
Status quo practices have overwhelmed the grid

Data centers are built today with backup diesel generators that are underutilized and polluting, and the industry's massive growth has led to significant power constraints in core markets that will be exacerbated as demand continues to grow.

A data center is a specialized warehouse that hosts IT infrastructure (i.e., racks of servers), providing primarily two services: electricity, to provide continuous power to the servers, and cooling, in order to keep the servers from overheating. Generally, the IT hardware is either owned by an enterprise that has an abundance of compute requirements and the sophistication to manage its own IT infrastructure, or a cloud services provider that manages IT infrastructure on behalf of other enterprises. The organizations that require the most compute needs, such as Amazon, Apple, Google, Meta, Microsoft, and Oracle, are often referred to as "hyperscalers."

In an increasingly digital world, enterprises rely on and value uptime in order to continuously run their operations, which is why data centers have been historically built as depicted in *Figure B*. Locationally, facilities have been built in areas with strong grid stability, low electric rates, and low natural disaster risk. Additionally, due to the benefits of clustering data centers around each other for technical and policy reasons, a significant amount of capacity has

Figure B



aggregated around several core markets. This has been driven by factors such as the availability of a skilled labor force, a desire for low latency (the amount of time for data to travel to the end user), and data gravity (the efficiency in having large data sets and processes located close to each other).

To provide backup in the event of an extended grid outage, data centers traditionally have been built with diesel generators sufficient to run the servers so that no loss of operations is experienced by the end user. However, since data centers are generally built in locations with high grid availability, these diesel generators are rarely used.

When the scale of projects were in the tens of MWs each, and thus small relative to the local utility's generation and transmission system, building data centers in geographic clusters was both preferred and acceptable. Now, forecasted data center projects are growing increasingly large in size, with the aggregate size requirements of certain campuses as large as hundreds of – or even more than a thousand – MWs for a planned development over a several year period. Such projects can require the same amount of power as hundreds of thousands of homes, but in a footprint as small as hundreds of acres. The result of this explosive, high-density demand growth is significant power constraints in core markets, generally coupled with an abundance of underutilized diesel generator energy infrastructure.

Areas across the country, from Northern Virginia to Silicon Valley and many in between, have had to put temporary pauses on new data center construction in order to more proactively plan and manage go-forward energy demand. Continuing to build data centers with status quo practices will exacerbate the problem and lead to billions of dollars of inefficient capital expenditures in seldomly used diesel generators that additionally face increasing concerns from local communities regarding their environmental impacts.

New technology enables a different approach to data centers

Advances in battery technology and the ability for certain workloads — such as AI training — to be interruptible enable flexible data center use cases that can benefit the grid.

It is highly unlikely that the demand for new data centers will decrease, and as a result there must be a more proactive approach to understanding how the next decade can support the corresponding energy infrastructure on the backdrop of the energy transition. Fortunately, advances in technology can enable more bi-directional interactivity between data center projects and utilities (i.e., the data center both consuming power from and providing services to the utility), which in turn would allow the data centers to be assets to the grid. Broadly, this falls into two categories, as outlined below.

First is the actual data center energy infrastructure. At a high level, instead of building diesel generators for backup, other technologies can provide backup to the IT infrastructure and run in the event of grid outages. Today, lithium ion battery energy storage can provide sufficient backup to the IT infrastructure in certain locations relative to historical grid outage data, and can also provide grid services in peak demand periods. Battery technology has significantly declined in cost over the past several years, and prices are on pace to decline another two-fold by 2035, thus becoming increasingly cost competitive compared to diesel generators without their negative emissions implications.[15] Utility scale batteries have proven effective in alleviating peak demand constraints in several markets, such as CAISO and PJM, and will continue to play an important role as many grid mixes shift towards more intermittent renewable generation sources.[16] Furthermore, new technologies, such as other long-duration energy storage options and nuclear micro and small modular reactors, may be commercialized later this decade and could ultimately be able to serve longer-duration backup and grid support use cases. Beyond their use in short-duration uninterruptible power supplies (“UPS”), however, battery storage technology has not been widely adopted in data centers to date given historical cost and reliability requirements.

---

15. Bloomberg New Energy Finance; 2023 Lithium-Ion Battery Price Survey
16. Wall Street Journal; Giant Batteries Helped the US Power Grid Eke Through Summer

Second is the underlying IT infrastructure, and the workloads that are run on the servers. Some workloads require high availability in the form of uninterrupted power flows. However, significant portions of certain workloads (e.g., AI training) do not have the same level of stringent requirements, and can be interrupted.

AI compute begins with the "training" phase, during which a model is given a vast amount of information and patterns to learn. During training, the model analyzes data and adjusts its internal structure to make accurate predictions. This takes significant computational processing power, and thus energy. As an example of one such model, OpenAI's GPT-3 required as much processing power as the 20 most powerful supercomputers combined,[17] with the training of its successor, GPT-4, costing over 100x that of GPT-3 due to increased compute and thus energy requirements.[18] Separately, Google estimates that 40% of its AI energy use is driven by training.[19]

Once fully trained, a model enters the "inference" phase, where it uses its learned knowledge to analyze new, unseen data and make predictions based on what it has learned. Outputs from these models can be text, images (which can be 60x more energy intensive to generate than text),[20] and video (which will take even more energy to generate).[21]

As AI becomes more prevalent in a variety of day to day tasks – such as integration into search engines or users communicating directly with AI agents – and gains mass adoption across specific enterprise use cases, the number of models and thus communications across models will grow exponentially.[22] The result will dramatically increase AI energy use, even in the near-term. One such estimate suggests that AI could use as much electricity as an entire small industrialized country by 2027.[23]

However, training can be accomplished in a more flexible manner. The model owner could theoretically schedule the provisioning of power to run the training model in accordance with certain agreed upon planning windows with the utility in order to have it to be treated as a large-scale, flexible asset to the grid. Relatedly, companies have also been testing moving workloads around to different data centers geographically depending on grid conditions.[24] This has many benefits and could provide significant relief to the grid in a more sustainable manner at a meaningful scale.

Harnessing both of these flexibility levers (dispatchable batteries and IT load) requires advances in the way that data centers are being built today from a design perspective, a closer dialogue with customers on uptime requirements related to the underlying IT, and understanding of a utility's grid challenges.

---

[17] Forbes; Artificial Intelligence Electricity Use Is In The Crosshairs
[18] Data Center Dynamics; Generative AI & the future of data centers: Part I - The Models
[19] Columbia Climate School; AI's Growing Carbon Footprint
[20] Luccioni, Jernite, Strubell; Power Hungry Processing: Watts Driving the Cost of AI Deployment?
[21] Wired; OpenAI's Sora Turns AI Prompts Into Photorealistic Videos
[22] Forbes; AI Agents And The Era Of The Intelligent Interface
[23] New York Times; A.I. Could Soon Need as Much Electricity as an Entire Country
[24] Bloomberg; AI Is Exploding Data Center Energy Use. A Google-Created Technique May Help

Data center flexibility can benefit the utility

Data centers can be the connector that brings new energy infrastructure technologies to market throughout their development timelines in order to help utilities meet their resource and capital planning objectives.

The flexibility of both the energy and IT infrastructure of a data center development is an asset that can be considered as part of a utility's capital, resource, and grid planning processes, such as an Integrated Resource Plan ("IRP"). An IRP takes a multi-decade view on forecasted demand growth and generation mix, utilizing a bottoms-up build of potential loads to develop a model to a variety of objectives. Shorter-duration 4-6 hour lithium ion batteries are valuable assets today that are often included in planning, but face hurdles in reaching operations in some markets given a myriad of development and economic considerations. Further, as renewable penetration increases and more storage is brought onto the grid, utilities will increasingly have to solve for more hours on their Load Duration Curve ("LDC"), as defined on the next page. New technologies that are not yet proven at scale, such as longer duration batteries, will be required in order to meet availability, reliability, sustainability, and affordability resource requirements.

As shown in *Figure C*, a data center can be a natural connector to bring this energy infrastructure to market now and over time.

Figure C

Provide input into energy resource requirements and system needs

Incorporate new DER technology behind the meter for trial use cases and at scale as commercialized

Multi-decade planning tool for utilities that assesses forecasted load demand and associated energy resource requirements.

Data centers can serve as strategic asset deployments for utilities, acting as symbiotic developments that can bring evolving energy infrastructure technologies online that can collaboratively serve both the data centers and the utilities to meet the objectives defined in their IRPs.

Commercialization timelines for emerging technologies – such as long-duration energy storage and nuclear fission and fusion – line up with IRP requirements.

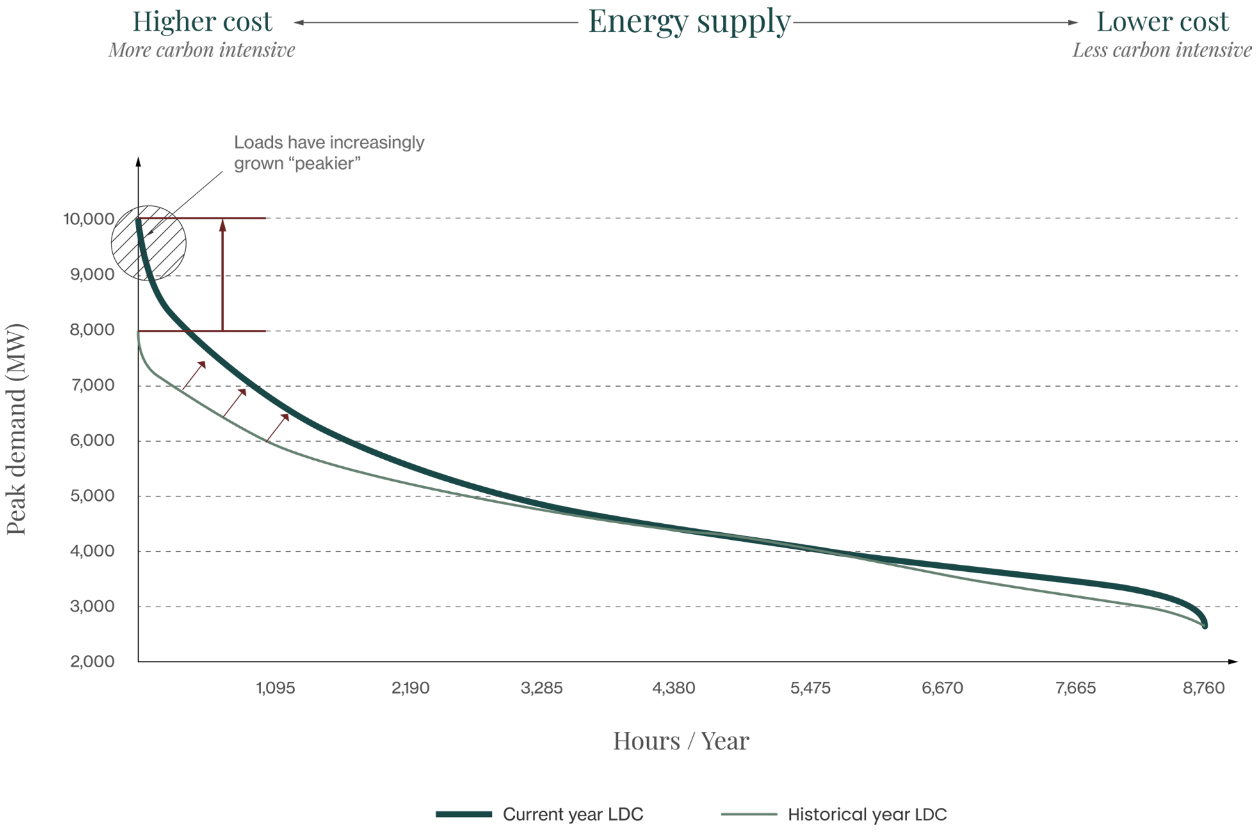
A phased data center development can be a pathway to incorporating valuable energy assets at scale into a utility's plan, by virtue of having an existing demand source (i.e., the data center) provide flexibility resources versus having to implement new independent resources. A well-structured project development plan can incorporate a myriad of assets "behind the meter" that can meet data center availability requirements as well as create value through providing a variety of services to the grid. Given the scale of these projects, the benefits can be substantial.

Figure D shows a representative LDC for a typical utility, plotting the electricity load for every hour of the year from peak demand to lowest demand. The LDC is a critical analytical tool that ultimately drives infrastructure and resource decisions for utilities. The figures included in this section are based on actual data from a utility that are rounded to both anonymize the utility and also to simplify the example.

As the economy has continued to electrify, central station intermittent renewables have been added, and the nation has experienced more extreme weather events, LDCs have become increasingly "peakier," meaning the far-left hand side of the LDC has increased upwards (compared to a flatter LDC across the majority of the hours of a year). To accommodate

Figure D

Representative load duration curves



these peak demand events, utilities have had to build new fossil based peaker plants, and now are considering delaying the retirement of existing fossil based plants as well being required to build significant additional transmission. This has resulted in a significant underutilization of infrastructure throughout large parts of the year (referred to in the electric industry as a “low load factor”), which increases the overall system cost to utilities and their customers.

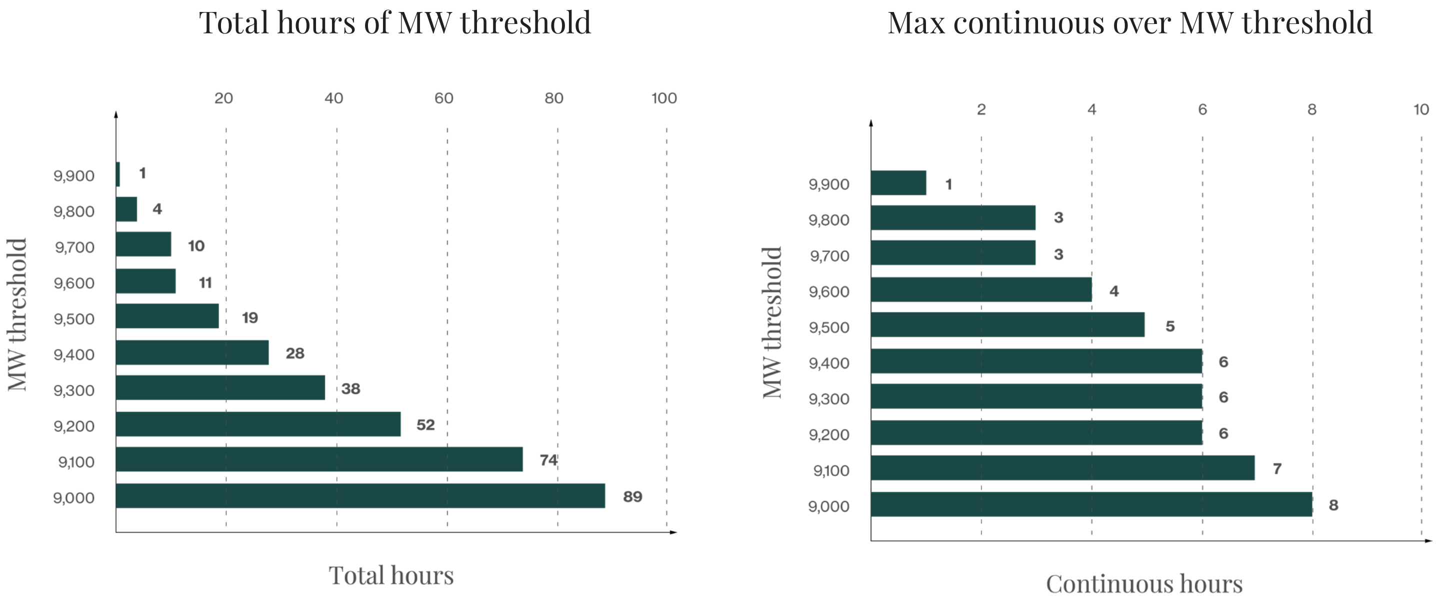
Understanding the LDCs at a more granular level can help better inform stakeholders on the utility constraints that they are attempting to solve for in their planning. As mentioned above, many of these constraints occur in solving for the peak demand hours. Absent flexibility at the load level in the grid, these investment decisions for the peak hours will often include overbuilding infrastructure and/or relying on fossil based peaker plants.

For example, as shown in *Figure E*, which is likewise representative for many utilities, the utility only exceeds 96% of its peak load for 11 hours (0.1% of the year), with the maximum continuous hours over that peak being four hours. This means that this peak load in excess of 9,600 MWs could be solved with large-scale four-hour battery energy storage rather than a new generation plant or transmission infrastructure. Further, the utility only exceeds 90% of its peak load for 89 hours (< 1% of the year), with the maximum continuous hours over that peak being eight hours.

Designing a data center project with flexibility has the potential to have a dramatic impact in the ability of utilities to help manage their peak demand. In the example below, the utility is assumed to be managing a data center queue of 1,000 MWs. In *Figure F*, the “Status Quo Scenario” assumes that the data centers are brought online with no flexibility, thus increasing peak demand requirements by 1,000 MWs, or a 10% increase in existing peak load.

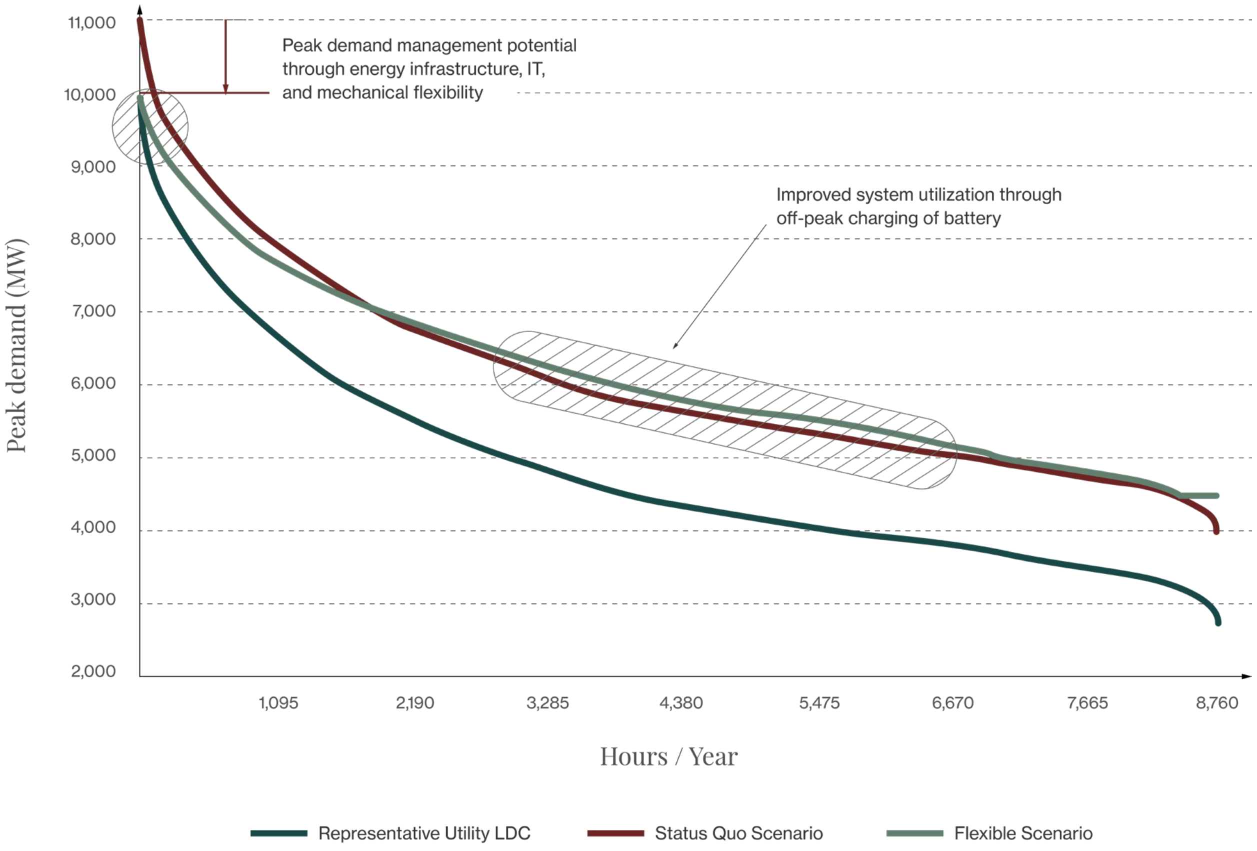
The “Flexible Scenario” incorporates both energy infrastructure (specifically battery energy storage)

Figure E


Total hours of MW threshold

Max continuous over MW threshold

Figure F Comparison of pro forma LDCs to representative utility LDCs



and IT load flexibility to better utilize existing utility infrastructure and help manage peak issues. In this scenario, the data centers can work collaboratively with the utility to define certain periods of the year where AI training workloads and mechanical loads can be flexible for prescriptive periods of time during peak periods. Given that certain workloads are not constrained by the same limitations in operation as shorter-duration batteries, the blending of both resources can be a valuable resource in helping solve for longer-duration flexibility needs. For example, by requiring power for 99% of the year (instead of 99.999%), flexible IT could help solve for the 89 hours (the 1% of the year) when the system is over 90% peak capacity and would otherwise need a resource for up to eight continuous hours.

Relative to adding 1,000 MWs in the “Status Quo Scenario,” the “Flexible Scenario” peak demand would only increase modestly relative to the Representative Utility LDC or could even decrease depending on the amount of flexibility offered to the utility.

In this example, the utility would be able to leverage several flexibility tools to help alleviate certain constraints, and importantly, the load factor (the measure of efficiency or utilization for energy infrastructure) would increase by over 10%, driving significant additional benefits for the utility. Note that this analysis does not yet assume the data center incorporates emerging technologies (e.g., other long-duration energy storage) throughout the data center's development and lifecycle, which could further improve its capabilities and usefulness for grid services.

While the specific impacts will vary based on the utility's infrastructure and generation mix, the benefits include more affordable peak demand management, increased system utilization (load factor), and thus a more reliable, sustainable and affordable grid for consumers.

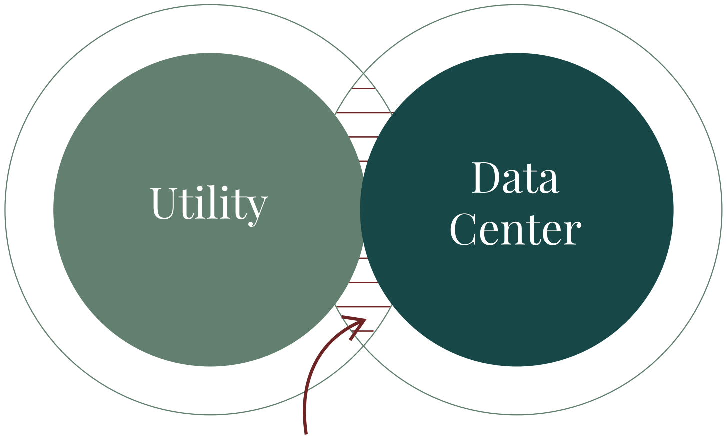
This requires a new approach to interconnection

Data centers, utilities, grid operators, and policymakers must collaborate in order to understand, plan, and ultimately deploy projects that benefit utilities and their stakeholders.

Historically, data center interconnections have been a one-way process whereby developers request access to power, and the utility works on power studies to arrive at a feasible power solution and associated timeline. This was acceptable in an era with abundant power resources and when data center capacity was not being implemented at the scale it is today. However, in the face of the increasing peak supply-demand imbalance and the significant volume and scale of new interconnection requests for a multitude of large-scale industrial use cases (including but not limited to data centers), the existing practice is no longer feasible and results in significantly delayed timelines, or outright denial, for new projects.

Moving forward, a more collaborative approach to planning is needed that creates a "win-win" for both data center developments and utilities. Utilities can more effectively enable access to power for a project in exchange for that project providing valuable grid services to help the utility manage its challenges. This collaborative process will improve the overall system instead of increasing constraints for the utility to

- IRP (or other strategic plan) that indicates key priorities and resourcing shortfalls
- Data sharing outlining core assumptions and resources underpinning IRP
- Prioritization of credible partners with willingness to invest in long-term infrastructure supportive of IRP



- Long-term investment horizon for project
- Phased development plan that incorporates emerging technology and solutions to IRP

Collaborative approach to project plans that can be a "win-win" for both parties manage. As outlined above, many of these issues can be solvable with flexibility and collaboration.

With each utility facing different challenges, it is important that both the utility and data center developer agree upfront on a project development plan and opportunity, and then work collaboratively to translate the plan into defined utility-facing products that ultimately can be implemented operationally and appropriately included in the mix of resources for a utility's capital, resource, and grid planning process. These products will take into account the various resources (i.e., the energy infrastructure and IT flexibility) in the near-, medium-, and long-term.

It will be important in this process that both parties also work collaboratively with regulators. Creating projects at this scale with this amount of flexibility and generation capability is a novel concept, and will require education and understanding to appropriately develop the proper structure for this type of resource to be considered in the interconnection process given the value it can create.

Additionally, as utilities plan for a multitude of interconnection requests, it is important that the utility and developer agree upfront on siting characteristics such that large-scale infrastructure planning can be synergistic with project development, particularly as the utility balances the buildout of other electrification initiatives (e.g., commercial charging depots).

Starting with a utility-first lens to solve energy challenges will enable a clearer set of goals for the data center with respect to flexibility requirements, which will allow a new collaborative approach that will break through the current backlog of projects helping to solve utility issues instead of creating them.

Conclusion and Recommendations

To summarize, data center projects can (1) provide operational flexibility to meet both customer compute demands and grid services needs, and (2) serve as strategic asset deployments for utilities. By acting as symbiotic developments with the grid that can bring new and evolving energy infrastructure technologies online over time, data center projects can help meet objectives defined in utilities' capital, resource, and grid planning processes.

In order to maximize the value of these benefits and enable access to power to build critical domestic infrastructure, all parties – data centers, utilities, grid operators, and policymakers – must collaborate to understand the value data centers can bring. Clearly identifying specific utility constraints, and understanding how data center flexibility can enable grid services, will allow more effective access to power in a way that benefits all parties and provides significant cost savings for energy consumers.

To realize the benefits of data center flexibility, SIP invites interested stakeholders to provide feedback so that we can collectively take additional action, starting with the below suggested next steps:

1. Creating channels to convene a diverse set of stakeholders – such as utilities, policy makers, and/or compute users and developers – through events, working groups, and other exchanges of information to brainstorm and ultimately implement flexibility initiatives;
2. Commissioning research on the potential grid impact of the suggested approach, such as through partnerships with universities, national labs, and other researchers;
3. Ensuring that potential federal funders, including the Advanced Research Projects Agency–Energy (ARPA-E), the Loan Programs Office (LPO) and other US Department of Energy programs, are aware of the benefits of data center flexibility;
4. Obtaining private sector commitments to invest capital into more flexible data center projects;
5. Implementing pilot projects and creating case studies on their results for dissemination across utilities, policymakers, and other key stakeholders.

It is a challenging – and exciting – time in our country's history. With the rapid technology advancements and increased capabilities that digital infrastructure and onshoring manufacturing provides, we are poised to take another significant step forward in global innovation and leadership. If we do not solve the power constraint problem, we will fall behind in this evolution. We look forward to engaging with stakeholders at this critical moment.