Power surge: AI, renewable energy, and the future of electricity

Goldman Sachs Exchanges

Carly Davenport, US Utilities Research Analyst, Goldman Sachs Research

Alberto Gandolfi, Head, European Utilities Research team, Goldman Sachs Research

Brian Singer, Global Head, GS SUSTAIN, Goldman Sachs Research

Date of recording: June 3, 2024

Allison Nathan: The demand for power, driven in part by the rise of generative artificial intelligence is set to grow to levels not seen in a generation. So, where will this additional power come from?

Carly Davenport: We think it's really going to be an all of the above approach because we're coming off of a time of no power demand growth. And now seeing this acceleration from a number of different areas. And so, we think that this is going to be a mix largely of new natural gas fire power generation, as well as new renewable generation, particularly on the solar side, coupled with battery stores that we've seen a lot of demand for in recent years.

Allison Nathan: I'm Allison Nathan and this is Goldman

Sachs Exchanges.

[MUSIC INTRO]

Joining me to discuss the global implications of generative AI, data centers, and the impending boom in power demand are my colleagues in Goldman Sachs Research, Carly Davenport, Alberto Gandolfi, and Brian Singer. Carly and Brian are joining me in the New York studio. And Alberto is joining us remotely from Milan.

All right, so we're all here. Let's jump in. If generative AI, which we've heard so much about, becomes a huge part of the technological landscape as so many people are promising, what impact will that have on power demand overall? Brian, why don't you start and, Carly, I'm sure you'll want to chime in.

Brian Singer: Yeah, I mean we've called this generational growth with the pun intended because the extent to which we're seeing multiple drivers of power demand growth led by data centers as you'll hear from Carly and Alberto, is driving a level of generation growth in the US, Europe, and potentially the world not seen in decades. And the level of

complexity that's seen when we're looking at autonomous vehicles, when we're looking at the ChatGPT searches, and a whole host of other potential innovations that could become transformative.

The implications are not only for higher power demand, but potentially for greater greenhouse gas emissions. But at the same time, AI and data centers bring a lot of benefits.

Allison Nathan: So, I was amazed to learn that a ChatGPT search uses 10 times more power than a Google search. Brian, why is that?

Brian Singer: So, one of the reasons is that if we're trying to actually get the level of complexity that a ChatGPT search is looking for, where they are going to be looking across not only the internet, but for all of the models that have been trained to give ChatGPT and equivalent type AI databases more knowledge, that needs to come from and use a lot more energy. And so, it becomes a lot more energy intensive when you're going through the operations that are required to try to increase the efficacy and accuracy of the types of searches that are happening that are requiring artificial intelligence versus those that would just be a

normal search.

Allison Nathan: Carly, do you have thoughts?

Carly Davenport: Yeah. And to Brian's point, as we think about the power demand implications, right, this is what we think is a generational opportunity. Power demand in the US has been essentially flat for the past decade. And as we look forward to 2030, we're now expecting to see about a 2.5 percent power demand CAGR in the US. And data centers are an important part of that. As we think about breaking that down, it's about 90 basis points of the growth of that 2.4 percent. And so, we do see it as an important shift and one that data centers is key to.

Brian Singer: I mean, Allison, to give you a sense, we talked about what data center to power demand had been kind of flattish 2015 to 2019. And then you had the pandemic that took place. And so, I think some of this inflection we may have already seen in the data center aspect of growth, power demand growth globally, may have gotten lost in the pandemic in '21 and '22 as we have recovered. And that's another reason why we think we're only starting to have that conversation and see some of

that impact now.

So, our projection is that data center power demand will rise 160 percent by 2030, relative to 2023 levels. And to put that into context, if that 160 percent growth from power demand coming from data centers, if that growth alone were its own unique country, it would be a top 10 global power consuming country.

Allison Nathan: But AI and data center demand, they're not the only thing driving power demand right now. So, talk to us about some of the other drivers increasing the level of demand for power.

Brian Singer: Sure, so above and beyond AI, and first, I think, just take a step back even within data centers to say there's been an inflection in data center power demand growth over the last three years that wasn't seen between 2015 and 2020. Because let's think about it. Data demand is not new. Data demand's been growing for decades. So, why is this only now being talked about?

Largely it's been because the data demand growth that we were seeing in the last decade was offset by efficiency gains

within data centers. In other words, the energy intensity of data centers was moving down to completely offset the increase in demand.

This was being driven by two key forces. The first is the shift from traditional data centers, which are more energy intensive, into cloud data centers and hyperscale data centers which are less energy intensive. The second is that within cloud and hyperscale data centers, they themselves were getting less intensive. And so, you know, the result was this has only now become an issue because those efficiency gains have started to decelerate.

We're not entirely through that shift into the cloud and hyperscale. But we think we're moving into the later innings of that. And we've now seen a lot of consolidation within hyperscalers and within the cloud as well.

So, the result is even before we talk about AI, data center demand for power demand had started to increase 2020 to 2023. And we think that's going to continue in addition to AI.

Then we look more broadly within what's driving electricity

demand and the work that Carly and Alberto have done, and you have electrification, whether it's vehicles or whether it's residential. And then in certain areas, particularly in the United States, you're seeing also the industrialization that's partly driven by policy. Partly driven by just our broader reshoring and industrial growth. So, all of those are major contributing factors that are driving increased power demand.

Allison Nathan: But I often hear the phrase don't bet against the engineers because we often in history have said, look, we're coming to the end of these efficiency gains and things are going to feel a lot different. So, why is this time different then?

Brian Singer: I think it's super fair. And it's why there could be a big range, whether it's around our estimates or other estimates for what data center power demand could bring. We've seen efficiency gains in that last decade of more than 10 percent per year in terms of a reduction of energy intensity. It's come down a lot. We do project that it's going to move a little bit higher for just the reason that you're asking about, because it's hard to say, well, it's going to be zero and we think about all the innovation

that's been brought to us from technology land.

And even if we look at AI, and we look at some of the recent, newer generation AI high powered servers, we've seen an 85 percent reduction in energy intensity between the current generation relative to two generations ago.

Now, if customers respond to those innovations by just buying more and more and more, then that could potentially offset the impact of that innovation. So, whether we're going to see technology users be budget constrained, or whether we're going to see them be demand constrained, or whether there is no constraint and they're going to just have voracious budgets and appetite for more and more AI-related servers and consume everything they can, that could lead to a very wide range of outcomes in terms of what data center power demand could be.

Allison Nathan: So, Carly, if the engineers can't deliver the amount of power that we need because of all the factors that Brian just laid out, where will this additional power be sourced from?

Carly Davenport: Yeah, we think it's really going to be an

all of the above approach because we're coming off of a time of no power demand growth. And now seeing this acceleration from, as Brian said, a number of different areas. And so, we think that this is going to be a mix largely of new natural gas fire power generation, as well as new renewable generation, particularly on the solar side, coupled with battery stores that we've seen a lot of demand for in recent years.

And so, if we look at our own estimates, we think it will be about 60 percent natural gas, 40 percent renewables. And the reason we chose that split was really to try to take into consideration two things. One is the reliability needs of data centers. You know, as Brian mentioned, there's a lot of training that needs to be done with these AI-driven models. And so, they need to be running essentially all the time. And so, natural gas is a dispatchable form of generation. You can turn it up and turn it down to meet the needs of demand.

Renewables, on the other hand, are more intermittent in nature. So, solar only works when the sun is shining. Wind only works when the wind is blowing. And so, that creates more of an issue on the reliability side. But a lot of these

companies building data centers, especially the hyperscalers of the world, they have commitments to green electricity consumption. So, we don't think that these companies are going to abandon their green aspirations and goals in order to meet the data center demand.

And so, we do think renewables are going to be an important part of this equation going forward.

Allison Nathan: And so, what have we observed already taking place? I mean, how much investment is flowing in? How much capacity is being built?

Carly Davenport: Yeah. So, we think that you'll need to see about 50 GW of new power generation capacity in the US in order to support the data center growth that we see. And based on our own cost assumptions, we think that points to about $50 billion of investment just on the power generation side.

Now, when you're building new generation, right, you also need to build the supporting infrastructure, the supply chain, all of these new pieces need to be built out.

And so, the actual investment needed to facilitate the growth overall is likely much higher. And as we look at our own utilities coverage, you know, we expect cumulative investment over the next four years to be about 40 percent higher than investment over the last four years. And that's a function, one, of the data centers, but also a confluence of other factors driving growth and needs for investment in the grid.

Allison Nathan: So, we haven't really seen it yet, but we are forecasting a big jump in the years to come?

Carly Davenport: Yeah, we saw a big step up in utility investment in 2023 versus 2022 levels. Not much of that was yet driven by this data center dynamic. But we expect that as we look forward, more of that will materialize. And it's going to be really regional. So, there are certain regions that have already seen it. And there are certain regions that are becoming hotbeds for data center growth that we'd expect to see that acceleration in the next few years.

Allison Nathan: If we think about the entities that are now embarking, potentially, on this major investment cycle to meet this generational demand, we don't think of

regulated utilities as the most nimblest participants in the market. So, is it reasonable to assume that they are going to be able to adjust, invest, and deliver this capacity?

Carly Davenport: Yeah, we think they're going to be a key component of it. But there will be other players involved. There are independent power producers out there that can help meet the need. You could see some of these hyper scalers and tech companies participate in some of the investment opportunity. But because of the scale of the utilities, especially when we think about actually building the grid infrastructure, the wires, they are, we think, an important player in making sure that that gets built out.

It is going to ultimately be garnered by the regulators. These are regulated businesses. And so, the utilities have to go to their state regulatory commissions and get these investment plans approved. And so, we're going to need to see a really big focus from the utilities on making sure that bills stay affordable for residential customers. Making sure that there are protections in place to make sure that the investment to support data centers is not being borne on the back of the residential consumer. And we've started to see some of that play through the regulatory process thus

far. But it's going to be a key piece to watch to determine if in fact the utilities are able to invest what is needed to facilitate this growth.

Allison Nathan: But if you think about their revenues as being somewhat capped by the regulators, where does the funding come from? How do they find the sources of funding to get this done?

Carly Davenport: Yeah. I mean, these are not free cash flow generative companies. Right? There's a lot of debt capacity that they're going to need to add to make this happen. There's a lot of equity issuance that is going to need to happen to facilitate this growth.

But they also can get some investment actually from the data center customers. Like if a utility needs to just build one transmission substation or one transmission line that's only going to support the data center, there’s are a lot of constructs in place across a number of states that would allow the utility to just go to get the capital from the customer to make sure then that that cost is not socialized across the rest of their customer base. And it's only actually borne by the data center customer.

So, we think that there are a number of special contracts and rates and ways that they can structure this to make it such that the data center customers actually are contributing and making sure that they're paying their fair share.

Allison Nathan: And are we seeing utilities and some of the other power producers presenting plans for development, hiring more people? I mean, it seems like there are going to be more people required to lay all this infrastructure as well. Where are we in that cycle?

Carly Davenport: Yeah. We've started in a number of regions, especially in places like the southeast and Georgia. We've seen utilities massively step up those plans for their expectations of power demand and also the infrastructure that they need to build it.

There are a number of plans that we'd expect to get revised as we think about 2025/26/27 that we think that if they haven't already reflected that growth, we'll start to see those numbers getting revised higher.

But this is really only something that we've started to talk about really starting in late last year. It came about very quickly. And so, not all of the utilities, not all of the regional grid operators have yet sort of taken that step change higher in their forecast. We've started to see it come through this year, but it's still something that's evolving.

Allison Nathan: Brian, what do you think?

Brian Singer: If we look at what the implications are for all of this growth across the supply chain, it's really significant because it's not just as Carly said the data center aspect, you have regulated utilities that Carly and Alberto focus on that are going to have the opportunity to invest capital because of their fleet transformation. And then also because of the grid hardening and grid reliability investments that are going to be needed, in part because of more volatile rising temperatures.

So, it just creates, I think, a very unique environment not seen in a long, long time for regulated utilities from an investment and rate base, and ultimately, earnings perspective.

Carly Davenport: And on top of what Brian just said, all of these new factors driving investment, we also have an existing power grid that's really old.

Allison Nathan: Right.

Carly Davenport: And so, there's a lot of investment that's just needed to enhance, upgrade beyond useful life of the existing asset base. So, now add all of these things on top of it and it really is a confluence of factors, you know, that as Brian mentioned earlier, we think is really a generational opportunity.

Allison Nathan: We're already just entering summer and we're hearing about blackouts all the time.

Alberto, let me turn to you. Obviously, we're facing a step change in power demand in the United States. Are you seeing a similar change taking place in Europe?

Alberto Gandolfi: Very much so. We are perhaps lagging the United States a little bit in terms of time. It's hard to say if it's six months or three years to be fair. But the inflection point is in sight. And our analysis suggests

actually potentially even stronger demand growth in Europe vis-à-vis the United States.

Why is that? Well, first of all, it's not just about data centers and artificial intelligence. It's also about electrification. And Europe has the most ambitious electrification plan in the world. It's called REPowerEU. It was conceived at the height of the energy crisis. It's essentially a plan to produce roughly 70 percent of electricity from renewables by the end of the decade. It's a very pervasive plan to electrify mobility, to electrify heating, even industrial manufacturing processes.

That, per se, is worth roughly 30 - 35 percent increase in power consumption. Whenever completed, the REPowerEU electrification plan is a plan to 2030. We think it's three to five years behind. So, let's call it a 2033 - '35 plan. But that would inflict significantly and decisively power demand.

On top of which, the spark that ignites power demand inflection is actually data centers and AI. We are anecdotally seeing and hearing from the corporates we cover, demands on data centers in different regions. And the numbers are really, really big.

So, if we put together data center AI-driven demand with the electrification process, we see the potential for 40 to 50 percent increase in power demand in Europe over the coming 10 years. This has extremely important implications in industry. You know, the top line as far as volume is concerned has been declining for 15 years. We are about to see potentially 10 years of substantial growth. I mean, we are hearing from corporates that as soon as some of the constraints on carbon Brian mentioned gets resolved, power demand could start actually to grow very quickly in the low- to mid-single digits.

Allison Nathan: And how are the power suppliers in Europe positioned to deal with that increase in demand?

Alberto Gandolfi: That's fascinating. It's fascinating for two reasons. Number one, if you see what happened in the past 12 months, most developers of power capacity, which is really in Europe renewables, so most renewable developers have been lowering investments. Have been lowering capex. Why? Well, largely to reflect the increasing interest rates. And largely to reflect some, "pressure" from capital markets which have been asking for higher returns.

So, the capital allocation has shifted away from chasing top line growth to actually chasing, optimizing, maximizing returns. And the early indications actually seem to suggest that this is working.

So, we are in a moment where we may be able to see power demands positively inflecting. And by quite, quite large number. When actually, the supply is not really catching up. It's actually been adjusted downward. We are seeing the big European oil majors scaling back the effort in renewables. This is actually going to create some issues three years down the line. But for utilities, it's most likely very positive because it implies higher returns.

Now, there's a second level of constraint in the infrastructure that we can elaborate if you want. It's power grids, particularly power distribution. That is where data centers will be connected. It's the medium voltage grid that is actually not ready for this data. By the European Union suggests that the European grid is more than 40 years old. Like, it was built as a state-of-the-art infrastructure 20 - 25 years ago. But it's like a beautiful car that you have not looked at for about two decades. And now the need for

modernization and upgrade is very urgent and very big.

Allison Nathan: So, what is it going to take to modernize the European grid and get it to a place where it needs to be to meet this data center and other demand?

Alberto Gandolfi: The data centers actually accelerates the urgency for these investments to take place as soon as possible. On our estimate, Europe needs about 800 billion euros of investments in transmission and distribution over the coming 10 years. And when you actually look at the single year investments, by the end of the decade, this is roughly implying that capex needs to double versus current levels. So, there's a doubling in the investment need in transmission distribution.

The two companies that have recently come out with an upgrade in the investment plan, particularly one in Germany, has essentially just done that. Therefore, corroborating our thesis that investments and growth rates need to double.

In renewables, we estimate even greater investments, close to 850 billion euros. As I said, precisely in a moment where

companies are starting to adjust downward their investment profile. So, power grids in our opinion require a very strong step up in investments, but doable because companies have been working on it for quite some time. Renewables, at some point there will have to be a rush to invest because clearly there's a widening gap between the supply and demand.

Allison Nathan: So, it sounds like the European power industry is facing a lot of the same challenges as the US industry. If we take a step back, we've talked about conventional sources of energy such as natural gas and renewable sources of energy as a part of the solution here. But in this conversation, we do often hear again about nuclear energy and the extent to which it can be part of the solution.

So, Carly, talk to us a little bit about that in the US.

Carly Davenport: Yeah, I think in the US we don't underwrite any incremental nuclear power plants being built by an investor-owned utility. The US has not historically had the best track record of building nuclear plants on time or on budget. And so, for an investor-owned

utility to undertake that would be, I think, quite risky from an equity price perspective.

So, we don't think any new assets will be built. But we do expect to see data centers, and they already have, look to strike deals with companies that operate unregulated nuclear plants because nuclear power basically solves exactly what the data centers are looking for. It's zero carbon power and it's reliable. And so, we've already seen some deals kind of come through. And there could be some innovations in the future related to small, modular nuclear reactors. We think that's probably a 2030s type event. But it is a focus of some of these hyperscalers to invest in that technology and actually build that out themselves.

Allison Nathan: When you say unregulated nuclear capacity, sorry, what exactly is that? Because we all think it is the most regulated area of the power complex.

Carly Davenport: Yeah. So, yes, it is regulated still by the Nuclear Regulatory Ccommission and there are very tight specs and things that they need to be built. But I mean that they are not regulated by a state utility commission. And so, they don't necessarily have to direct their power to

the grid.

So, if you operate an unregulated nuclear plant, there would be not really anything that would preclude you from striking a direct contract with a data center customer to just sell your power from your nuclear plant directly to the data center. And it doesn't have to go to the grid. And so, that's really what we mean. We wouldn't expect to see any plant that is under a state utility commission regulation that is powering the grid, we wouldn't expect to see that reallocated away to directly power a data center. But we have already seen deals come about for these plants that are not necessarily serving the grid or don't have to serve the grid. And they can just serve a data center directly.

Brian Singer: I think one of the key points is that for the technology companies, as has been said, are among the most forward and front footed in trying to mitigate the impact of their carbon footprint, both their own and downstream, despite the expansions that are happening.

And given the extent of what we're talking about from a power generation growth and power demand growth with data centers, they are likely to take and appear to taking

an all in approach. Supporting renewables via power purchase agreements. Supporting the development of some of the small, modular reactor nuclear technology as Carly discussed.

And then also, looking at some other technologies for carbon emissions, abatement reduction or removal, like direct air capture or carbon capture and storage.

Allison Nathan: Right. So, if we think about all the constraints that the industry is facing in terms of delivering this increase in power that we see coming, which ones concern you the most, Carly?

Carly Davenport: Yeah, I think there are two that have been top of mind. I think the first is just the permitting and logistical constraints of actually getting both generation and transmission projects built. It can take, you know, the average permitting period for a transmission project is usually around four years. And then you're building a generation project, to actually get that connected to the grid, there can be constraints around the timing that it takes the grid operator to conduct their studies to make sure that connecting that project to the grid won't cause a

reliability issue.

Transmission can be a constraint in getting connected to a gird. The interconnection queues across the country can range anywhere from, on average, 40 to 70 months. And so, it can be an incredibly long process actually to get your projects built and connected to the grid. So, we think that is one key constraint as we think about it takes you probably two to three years to build a data center, you could be in the queue to get your power connected potentially longer than that in some cases. We do think there's some existing supply length in the system to meet demand in the interim. But as we think about the long-term constraints, we think that is certainly one to keep in mind and it's an area that regulators at both the state and a federal level have been focused on trying to alleviate.

And then I think the second one, we sort of talked about it before, is just this affordability constraint, particularly for the regulated utilities, to make sure that they are able to keep customer bills in a healthy place and to make sure that the investment is for the data center growth is not borne by the residential customer because if they can't prove that to the regulators, then it's going to be really

difficult for them to get these investment plans approved.

Allison Nathan: Right. Alberto?

Alberto Gandolfi: Well, for me, I would equal Carly on permitting. But I would add two other points. I think equipment, availability of equipment worries me a bit because, obviously, now we are seeing big corporates announcing capex increases in power grids. And they sign big framework agreements with the suppliers, with the cap goods companies. So, they will obtain their equipment.

But companies that are a bit late in other regions to this trend may actually find themselves in a queue where they have to wait for, you know, two, three, four, five years for transformers or substations. So, that could be a bottleneck.

And last, raw materials. You know, we know that if you read the research from our colleagues at Goldman Sachs that copper could become a constraint. The amount of copper in an electrified economy has to go up. The amount of copper to connect these new facilities will have to go up. So, those three, I think, permitting, equipment, and raw

materials are the key risks and the key bottlenecks that we continue to monitor.

Allison Nathan: Carly, Brian, Alberto, thanks so much for joining us.

Brian Singer: Thank you.

Carly Davenport: Thank you.

Alberto Gandolfi: Thank you.

Allison Nathan: This episode of Goldman Sachs Exchanges was recorded on Monday, June 3rd, 2024. I'm your host, Allison Nathan.

If you enjoyed this show, we hope you subscribe on your platform of choice and tune in next week for another episode. And if you want more insights from Goldman Sachs, make sure to visit GS.com and sign up for Briefings, a weekly newsletter from Goldman Sachs about trends shaping markets, industries, and the global economy.

Thank you for listening.

The opinions and views expressed in this program may not necessarily reflect the institutional views of Goldman Sachs or its affiliates. This program should not be copied, distributed, published, or reproduced, in whole or in part, or disclosed by any recipient to any other person without the express written consent of Goldman Sachs. Each name of a third-party organization mentioned in this program is the property of the company to which it relates, is used here strictly for informational and identification purposes only and is not used to imply any ownership or license rights between any such company and Goldman Sachs. The content of this program does not constitute a recommendation from any Goldman Sachs entity to the recipient and is provided for informational purposes only. Goldman Sachs is not providing any financial, economic, legal, investment, accounting, or tax advice through this program or to its recipient. Certain information contained in this program constitutes “forward-looking statements,” and there is no guarantee that these results will be achieved. Goldman Sachs has no obligation to provide updates or changes to the information in this program. Past performance does not guarantee future results, which may vary. Neither Goldman Sachs nor any of its affiliates

makes any representation or warranty, express or implied, as to the accuracy or completeness of the statements or any information contained in this program and any liability therefore (including in respect of direct, indirect, or consequential loss or damage) is expressly disclaimed.

This transcript should not be copied, distributed, published, or reproduced, in whole or in part, or disclosed by any recipient to any other person. The information contained in this transcript does not constitute a recommendation from any Goldman Sachs entity to the recipient. Neither Goldman Sachs nor any of its affiliates makes any representation or warranty, express or implied, as to the accuracy or completeness of the statements or any information contained in this transcript and any liability therefor (including in respect of direct, indirect, or consequential loss or damage) are expressly disclaimed. The views expressed in this transcript are not necessarily those of Goldman Sachs, and Goldman Sachs is not providing any financial, economic, legal, accounting, or tax advice or recommendations in this transcript. In addition, the receipt of this transcript by any recipient is not to be taken as constituting the giving of investment advice by Goldman Sachs to that recipient, nor to constitute such

person a client of any Goldman Sachs entity. This transcript is provided in conjunction with the associated video/audio content for convenience. The content of this transcript may differ from the associated video/audio, please consult the original content as the definitive source. Goldman Sachs is not responsible for any errors in the transcript.