Thematic Insights

Power hungry AI – investment implications in the era of energy transition

Natalia Luna
Senior Thematic Investment Analyst, Global Research

At a glance

- The growth in AI and associated data centre expansion is set to increase power demand. This has significant implications in the era of energy transition.
- We explore options for power provision including behind the fence locations at nuclear and gas plants alongside efforts to improve grid connectivity and efficiency.
- Emissions will increase as a result of data centre expansion. Big tech will likely use some non-renewable resources but we expect them to continue investing in renewables.
- The AI revolution is thirsty for energy. We see numerous opportunities including quality names in areas like energy efficiency and provision of clean energy infrastructure.

Understanding the implications of AI growth

We have undertaken deep research and analysis on the investment implications of AI growth. The work has involved eleven analysts across our research platform and brought together experts on different industries and themes.

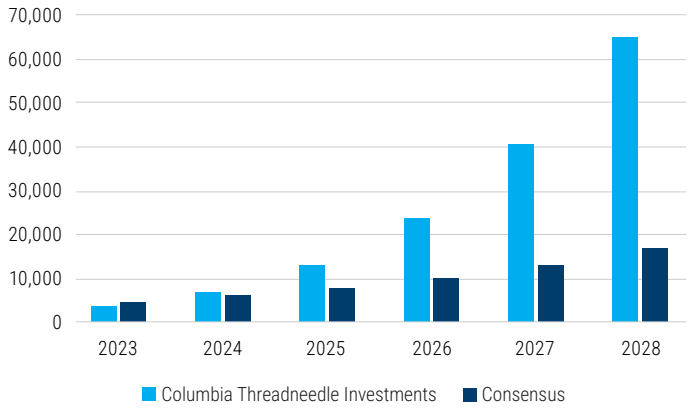
In this research we assess whether our semiconductor demand forecast through to 2028 can be met with an equal amount of data centre capacity. Within that framework we analysed the ability of the grid to meet power demands as well as considering non-grid options. We also explored the ability of industrial, machinery, and equipment suppliers to support growth of the AI ecosystem and looked at the impact of all this expansion on copper demand. And lastly, the impact on carbon emissions.

Here we focus on the implications of AI growth on power and the energy transition in the US. We find that meeting the high demands of AI will prove challenging but conclude with some confidence that between grid and non-grid options, these power demands can be met. However, we believe that some specific markets will not be able to generate adequate power, which will push AI demand to markets where power availability is greatest.

We anticipate higher demand than the consensus

The starting point of this analysis is our forecast for AI compute demand. Our predictions are materially higher than consensus and are driven by our FLOP/GPU estimates that see the AI Accelerator market potentially growing at 50% (compound annual growth rate) to $500B by 2028. As a result, our AI data centre energy demand forecast is also materially higher than consensus.

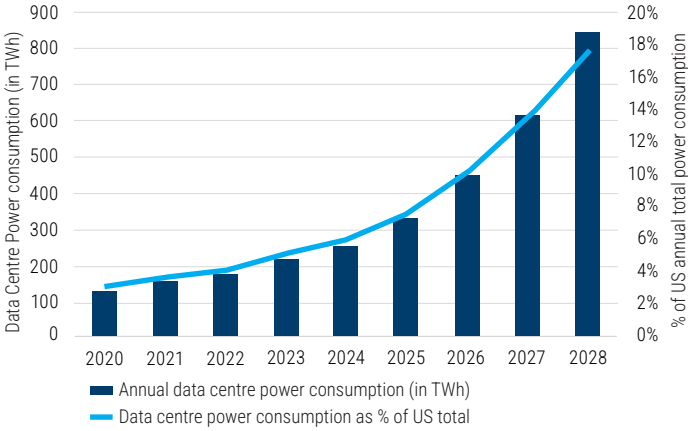
Figure 1: AI Data Centre Energy Demand (in MW)



Source: Columbia Threadneedle Investments, September 2024

Our data centre power consumption estimates (AI and non-AI) represent a greater proportion of total US annual power consumption, up to ~17.5% in 2028.

Figure 2: Data Centre Power Consumption (in TWh)



Source: Columbia Threadneedle Investments, September 2024

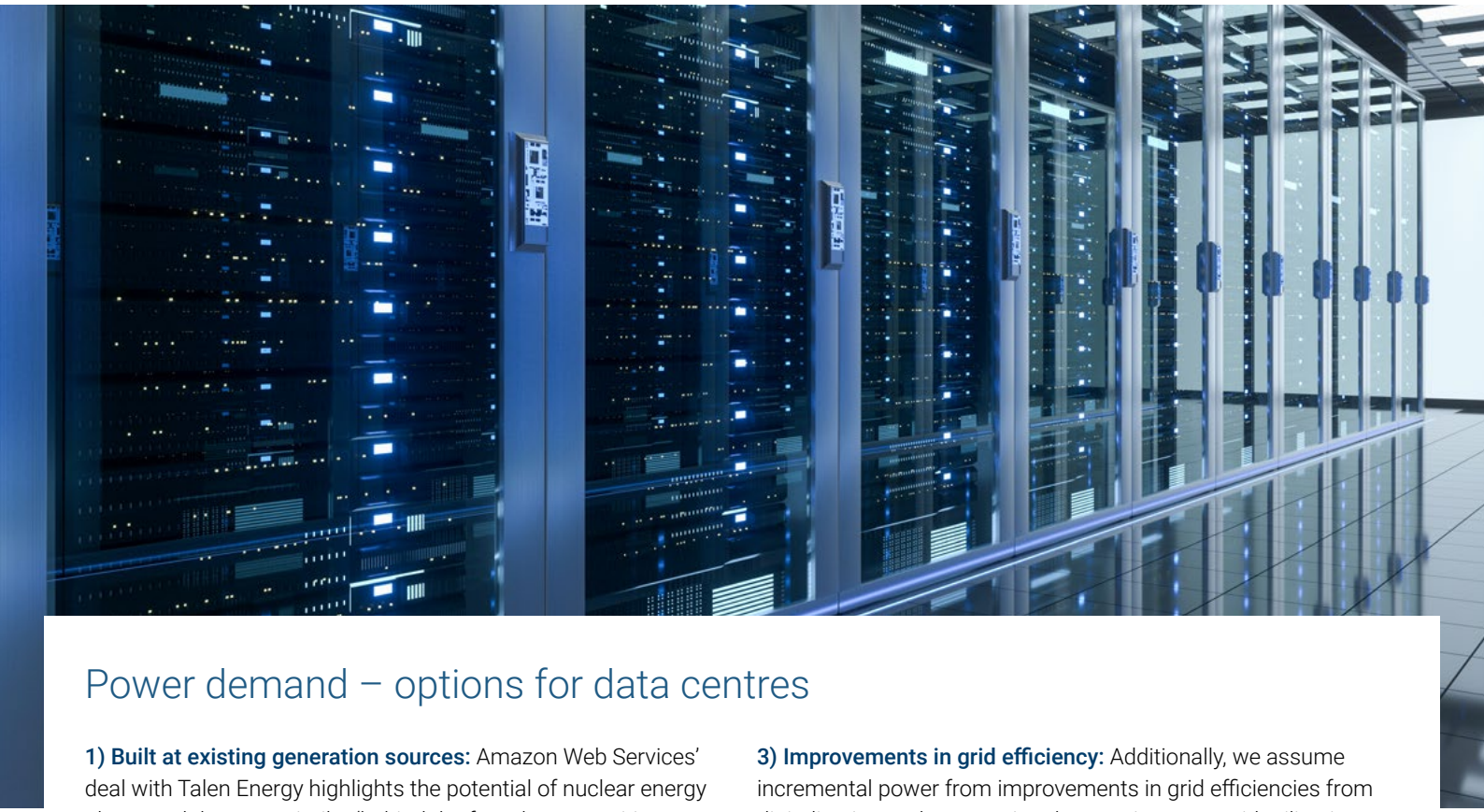
While US power demand has been mostly flat for the past decade, our AI forecasts point to an increase in coming years. In fact, US data centre energy demand could increase from 211 TWh in 2023 to over 800 TWh in 2028, representing 2.8% growth annually.

When we think how the data centre power demand could be satisfied, we consider that siting of facilities has multiple considerations:

- Intersection of sufficient power, fiber and transmission connection points, as well as water resources
- Proximity to population centres
- State and local support.

Jargon busters

AI Accelerator – a deep learning processing unit designed for AI workloads.
FLOP – Floating point operations per second – a measure of computer performance.
GPU – Graphical processing unit – initially designed for graphics and image processing. Used in training of AI learning.



Power demand – options for data centres

1) Built at existing generation sources: Amazon Web Services' deal with Talen Energy highlights the potential of nuclear energy plants and there are similar 'behind the fence' opportunities at gas plants. In total, these opportunities could potentially satisfy the power needs through 2027 based on our current demand forecast. Customer willingness to take outage risk and adherence to stated decarbonisaton goals will be items to watch.

2) Power needs can be satisfied with projects coming out of interconnection queues: Transmission interconnection challenges have been in focus given the increasing number of projects in queues (waiting for transmission equipment or upgrades before a project can connect to system). There was ~2,600 GW of power capacity seeking grid interconnection at the end of 2023 (exceeding the total installed US power plant fleet capacity). Interestingly, 95% of this capacity comes from renewable energy projects.

We see encouraging regulatory efforts to improve the grid connection process. For example, the Federal Energy Regulatory Commission (FERC) Order 2023 will require that projects are considered in batches or clusters. They will impose penalties for transmission providers that don't complete studies on time. Moreover, the Department of Energy (DOE) issued a roadmap in April 2024 aimed at improving interconnection processes and creating new fast-track options for interconnection.

Using relatively conservative assumptions, we see additional power that could become available through capacity expansion over the next five years.

3) Improvements in grid efficiency: Additionally, we assume incremental power from improvements in grid efficiencies from digitalisation and automation that can increase grid utilisation. Therefore, our analysis indicates that existing generation (behind the fence) sources could offer a workaround to transmission interconnection challenges. But even if these opportunities do not materialise, we estimate that there will be sufficient power through 2028.

More power could become available if:

- Interconnection processes improve and more capacity is added
- More grid capacity is available to data centres vs other end-users
- Technology drives efficiency gains.

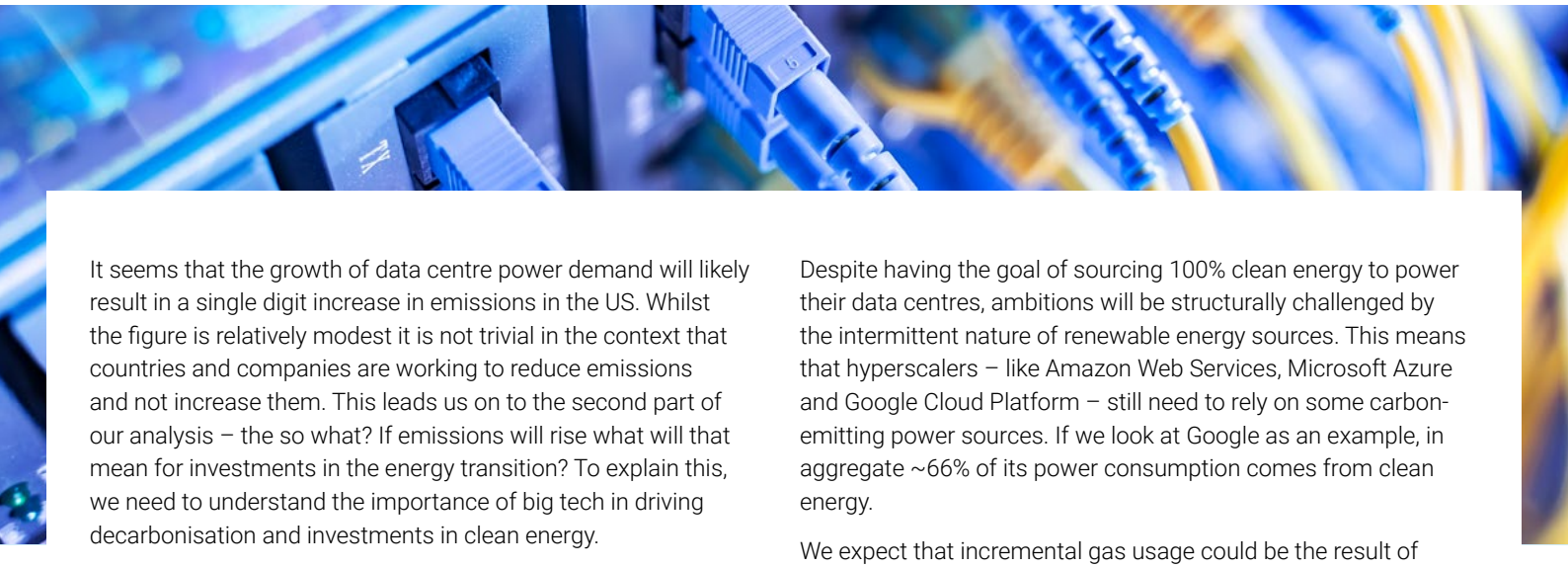
What would be the impact on carbon emissions?

Given that the growth of AI will be very power intensive we also analysed the impact on carbon emissions and the implications for investments in the energy transition.

To quantify the carbon impact, we have taken our estimates on power supply to meet our data centre demand forecasts and run two scenarios:

1) In a worst-case scenario of power being sourced from fossil fuels, we would see an increase of 7% in US emissions by 2028

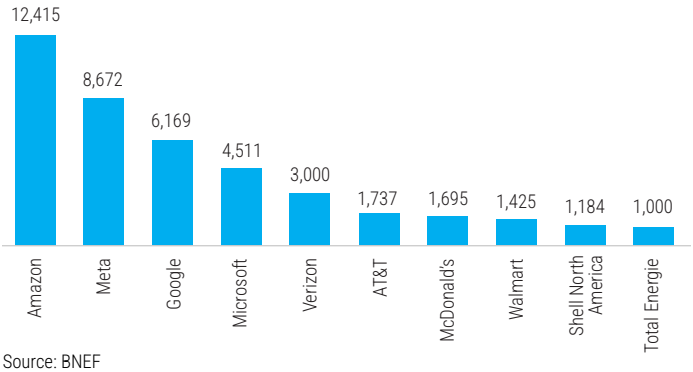
2) More realistically, if a meaningful portion of the power comes from clean energy sources the growth in US emissions could be around 2%.



It seems that the growth of data centre power demand will likely result in a single digit increase in emissions in the US. Whilst the figure is relatively modest it is not trivial in the context that countries and companies are working to reduce emissions and not increase them. This leads us on to the second part of our analysis – the so what? If emissions will rise what will that mean for investments in the energy transition? To explain this, we need to understand the importance of big tech in driving decarbonisation and investments in clean energy.

All the big tech companies have ambitious emission reduction targets. As AI grows and related emissions increase, these targets will be under pressure. However, the important point here is that linked to these targets to mitigate emissions all big tech companies have made commitments around using 100% clean energy to power their data centres. Some have committed to run 24/7 on carbon-free energy on every grid (Google). Others (Amazon, Microsoft) have committed to sourcing 100% of energy from renewables i.e to have power purchase agreements (PPA) for green energy contracted for 100% of electricity consumed. This commitment is what has been driving large investments in renewables.

Figure 3: 2010-2022 cumulative contracted capacity (MW)



Source: BNEF

As per the chart, big tech are the largest corporate buyers of renewable energy contracts, so they have been critical in financing the development of the sector and will continue to play a key role. Recently, for example, we have seen related announcements from Amazon[1] and Microsoft[2].

Despite having the goal of sourcing 100% clean energy to power their data centres, ambitions will be structurally challenged by the intermittent nature of renewable energy sources. This means that hyperscalers – like Amazon Web Services, Microsoft Azure and Google Cloud Platform – still need to rely on some carbon-emitting power sources. If we look at Google as an example, in aggregate ~66% of its power consumption comes from clean energy.

We expect that incremental gas usage could be the result of increased energy demand as AI grows. And should increased gas demand align with our data centre forecasts then we could see an 8% rise in usage by 2028.

Nonetheless, we don't believe that higher power needs will derail the industry's commitment to invest in clean energy.

Even if in practice data centres remain physically connected and reliant on their grid connections, a focus on lower carbon power through PPAs will allow them to maintain the green attributes of having the clean energy contracted. They will be able to count these as carbon credits which in turn aligns with their decarbonisation goals. Thus, despite its limitations, renewable energy sources will still be part of the power solutions for hyperscalers

We thus still expect big tech to remain a key driver of the clean energy revolution, and this could expand to include storage which in turn provides flexibility to the grid. This will support the growth potential of storage developers as well as manufactures of clean tech equipment.

In terms of other clean energy sources, nuclear and small modular reactors (SMR) are good power solutions for data centres as they can provide baseload clean power[3], as highlighted by the recent deal between Constellation Energy and Microsoft. However, we believe they are more likely a potential solution to meet needs beyond the next decade given the time lag for scale, government approvals and construction. We plan to present an analysis of the nuclear potential in a future article.

[1] AWS plans to invest €15.7 billion in Spain, supporting 17,500 jobs annually in local businesses (aboutamazon.eu)
[2] Brookfield and Microsoft Collaborating to Deliver Over 10.5 GW of New Renewable Power Capacity Globally | Brookfield Renewable Partners
[3] https://www.world-nuclear-news.org/articles/constellation-to-restart-three-mile-island-unit-powering-microsoft

Conclusion – investment implications

We believe that renewables will play a role in meeting the power demand for data centres but they will not be the only source of power. Higher energy requirements will result in the utilisation of a combination of sources, on site back-up generation, and grid connectivity.

Select regulated utilities are seeing pockets of increased commercial load, which is largely data centre driven. This trend should provide two key benefits:

- Increased load spreads costs over a larger base and can benefit customer bills. This dynamic could potentially smooth future regulatory outcomes
- Incremental capital opportunities will be additive to rate base growth.

Additionally, power-levered names could also benefit from these trends. All else held equal, increased power demand would be supportive of near term power prices. Also, from our discussions with companies, tech customers seem willing to pay a premium to reinforce their green credentials.

Moreover, our outlook for incremental power via interconnection queue projects would require new pipelines. Natural gas midstream companies will benefit as they construct and operate these natural gas pipelines and related infrastructure.

And despite pressure on carbon emissions from AI power intensity, we don't expect firms to be derailed from their commitments on clean energy investment. We believe investments for renewables and storage will continue to grow this decade and spread into nuclear and SMRs during the next decade; seemingly demand for clean technologies such as CCS (carbon capture and storage) could accelerate.

With regards to the risks for investments in this space it is clear from discussions with companies that the lack of grid capacity represents a real challenge – one likely to be exacerbated by AI and data centre growth.

Permitting remains the major issue to accelerate investments in transmission due to the lack of progress at federal and state level from regulatory bodies to overcome regional issues. With this backdrop, companies emphasise that digitalisation of the grid is becoming very important as infrastructure needs to get smarter and more intelligent. The likes of smart meters, sensors, and software automation need to be better deployed. This aligns with our view on the importance of grid efficiencies as part of power solutions which in turn mean opportunities for electrical equipment companies.

View from the desk

As discussed in this thematic piece, the AI revolution will be thirsty for energy. And as investors, this thirst offers up a range of interesting opportunities. One angle to this is the pursuit of companies that can help decarbonise or make physical infrastructure more energy efficient. Schneider Electric is a world leader in this space, using smart power distribution systems with energy management software to ensure that buildings optimise their energy usage. Studies by Schneider Electric show that implementing their digital building and power management solutions in existing office buildings can cut operational carbon emissions by up to 42%, with a payback period of less than three years. Schneider applies its cutting-edge technology to data centres, designing and implementing carbon reduction strategies for them, whilst leveraging data and digitalisation to drive operational efficiency.

The boom in AI applications will lean heavily on grids across the world, and so the degree to which these grids are themselves powered through renewable energy sources is a key variable in the environmental impact of AI. We are invested in SSE, a leader in developing, owning and operating renewable energy sources in the UK and Ireland, encompassing onshore and offshore wind, as well as hydroelectricity. SSE is currently working on the world's largest offshore wind farm project, a 3.6GW facility in the North Sea. SSE is targeting a tripling of its renewable installed capacity by 2031 and is a vital part of the UK's net zero energy transition.

Where there is growth, there is opportunity. We as a team are focused on investing client money in businesses that are sustainably enabling the technological opportunities offered up by AI, and the energy space is one of the exciting ways of doing so. Such businesses should see enduring demand for their products and services, and this demand should help drive top-line growth and share price performance for our clients.

Nick Henderson
Portfolio Manager, Global Equities

The mention of any specific shares should not be taken as a recommendation to deal.

Contact us

columbiathreadneedle.com

Follow us on LinkedIn

To find out more visit **columbiathreadneedle.com**

COLUMBIA THREADNEEDLE INVESTMENTS

CTEA7178051.1- | WF241364 (10/24)