Bubble or Goldmine? Analysis of Recent Tech Investments in Data Centres

Introduction

Last year, we discussed how surging demand for data storage and processing - driven by streaming, e-commerce, and AI—propelled record investments in server infrastructure worldwide. There existed a large FOMO surrounding Data Centres and AI. Currently, with increasing investor pressure amid large Capex requirements and slow materialisation of profits from data centre investments, we see companies increasingly adjusting their strategies. Notably, two of the largest tech companies – Meta [NASDAQ: META] and Microsoft [NASDAQ: MSFT] – seem to be adopting divergent strategies concerning their data centre investments.

Microsoft's Pullback

Microsoft has halted and delayed various data centre expansions in several regions, including Indonesia, the UK, Australia, and multiple US states (Illinois, North Dakota, Wisconsin). This wave of cancellations and postponements, initially estimated at around 200 megawatts (MW) of capacity, has since ballooned to a reported 2 gigawatts (GW) of projects across the US and Europe.

Although Microsoft publicly reaffirms its commitment to spend $80bn on AI-oriented infrastructure in 2025, the deceleration of data centre investments remains an issue worth examining. With no official statement from Microsoft's management, their bearish outlook on data centre demand matching current supply may be one explanation. This aligns with the broader concern of massive data centre expansions in the last year not aligning with demand forecasts. An early indicator may be the downward trend in pricing on the public GPU rental market. In the last twelve months, the price of renting two of Nvidia's [NASDAQ: NVDA] consecutive chips – H100 and H200 – have both fallen by over 50%. While natural for chips to lose value as newer models are rolled out, a historical 20%-30% fall in value is much lower than the observed 50% drop, indicating a weaker demand-side factor in the current pricing. The oversupply thesis seems to have traction, especially with additional circumstantial evidence of data centres being built without underlying hosting contracts. While indicative for the broader industry, Microsoft most likely projected demand excessively due to their relationship with OpenAI - originally a major customer for Microsoft's data centres due to large LLM training workloads. Microsoft's press release from January 2025, covers how the terms of its partnership with OpenAI have changed, including the addition of a right of first refusal (ROFR) for Microsoft, regarding hosting incremental training workloads of OpenAI. In conjunction a clause was added allowing Open AI to build its own capacity. Hence, it is clear why Microsoft is reducing its data centre capacity. Furthermore, despite the seemingly large scaling back, Microsoft seems to have fulfilled all its medium-term capacity needs. This cutback is only on long-term projects that would far exceed the company's current requirements. is taking a step back to reevaluate the long-term demand growth potential for data centres before committing to long-term projects far exceeding current capacity needs.

Meta's Acceleration

Meanwhile, Meta is moving in the opposite direction unveiling plans to spend $65bn on new data centres and AI infrastructure in 2025. Meta has sought to raise as much as $35bn to fund a "mega data centre" buildout in the US with Apollo Global Management [NYSE: APO] poised to lead the financing package. While bold, these investments underscore Meta's desire to remain at the forefront of AI and social platform infrastructure. Furthering this trajectory Meta is planning to add a full gigawatt of computing capacity while scouting for sites for a multi-hundred-billion-dollar AI campus. While speculative, this potential $35bn deal with Apollo and other projects marks a significantly different direction in terms of CapEx compared to Microsoft.

Origin of the Divergence

The fraying relationship with OpenAI is a significant factor behind Microsoft's cautious approach to further data centre expansion. In addition, data centre ROIs are declining and returns from AI are expected to take years, creating concern among investors. In the last 2-3 years, Microsoft's CapEx has grown while growth has decelerated, exacerbating investor pressure. On the flip side, Meta seems to be betting on exponential growth in AI-driven social media, metaverse, and generative AI services soaking up future excess capacity. They are competing in the "Data Centre Measuring Competition" with xAI and OpenAI, for who will be able to create Artificial General Intelligence (AI possessing human-level cognitive abilities). This arms race accounts for the increased investment in data centres. Whether this will be sustained in the long run is a mystery. In the past, Meta caved to investor pressures after investment in the Metaverse consumed the company's profitability. Therefore, Meta might align its spending with Microsoft's more conservative approach, if their bet doesn't pan out.

Big Tech Partnering with Asset Managers

In the past private equity (PE) firms have funded large tech deals such as a $5bn financing for Databricks Inc. led by Blackstone or an $11bn joint venture (JV) of Apollo and Intel [NASDAQ: INTC]. If Meta's investment with Apollo finalizes, it would be one of the largest team-ups between big tech and asset managers in the data centre sector. This collaboration would follow this broad trend of asset managers and tech firms combining their capabilities in data centre and AI investments such as Intel's and Apollo's $11bn JV in a high-volume chip-manufacturing facility in Ireland. Similarly, Microsoft, Global Infrastructure Partners, Blackrock [NYSE: BLK], and MGX, launched a partnership (since scaled down due to Microsoft's step-back on data centres) to mobilise almost $100bn into data centre capacity expansion projects in the USA. Big tech firms seem to be looking to increase their flexibility and deleverage their investments in data centres, by inviting PE partners into the mix. A clear advantage of joining forces is the scale at which investments can be conducted. More parties can raise more capital, leading to larger projects. For asset managers, on the other hand, providing credit or equity for these investments opens exposure to a quickly growing market providing outsized returns in the coming years.

Open AI and Stargate

On January 21, 2025, Open AI, developer of Chat GPT, announced its new project Stargate. Lead by Open AI responsible for the operational part and co-lead by Softbank Group [TYO:9984] who has financial responsibility. The project is a fund dedicated solely to creating and funding infrastructure projects for data centres to support the training and development of Artificial Intelligence. The fund is aiming to raise $500bn over the next four years with $100bn being deployed immediately. The infrastructure built will be solely for the use of Open AI. Whilst being Open AI's project, the Japanese holding company Softbank is inducing a serious investment into the fund, as well. At the unveiling of the new project at the White House, Softbank's chair Masayoshi Son was present. The announcement was led by Donald Trump who used the project as a sign of a win over China, despite no government funding being involved. Among the other investors are Oracle [NYSE: ORCL], MGX, Abu Dhabi State AI fund, as well as other private stakeholders not shared with the public. Notably, there was no involvement from Microsoft as they are pursuing their own data centre projects. Upon request, Microsoft's CEO shared that while they support the project, they decided to focus on their personal infrastructure strategy.

Musk Dispute

After the announcement at the White House, a post by Elon Musk titled "They don't actually have the money", indicated the promised $100bn was not secured. This was one of the few instances where Trump's largest supporter

(by donation) was opposing the White House. Surely at least partly motivated by his ongoing dispute with Open AI. The dispute concerned Open AI's profit structure and has driven apart the co-founder and the AI company. Following the tweet, Altman responded publicly stating Musk was lying, and inviting him to a visit to the first data centre site. Said site is in Abilene, Texas, and is currently developed by Crusoe, which in return will be leased by Oracle for the use of Stargate. Although construction has already started in June 2024, the project is now part of Stargate, a strategic move by Oracle, who started the venture first and then reached out to Open AI to secure it as a customer. Crusoe is known as a so-called "neocloud" company, which are web service companies, mainly focused on AI topics. They tend not to develop their chips, rather buying them externally, predominantly from Nvidia.

DeepSeek

Hangzhou DeepSeek Artificial Intelligence Co. Ltd. is a Chinese AI lab that has sparked headlines worldwide with its new models R1 and V3. Facing major GPU restrictions placed in 2021 and 2023 by Biden Deepseek developed its AI to optimize the software rather than the hardware. This led to a relatively inexpensive creation of a comparable program to all its competitors roughly costing only a supposed $6mn. Must be disclosed however that this figure is most likely minimized (a private company in China thus practically no information is known). The announcement of the efficiency of the training of the model in January disrupted tech stocks, leading to a drop in industry stock prices triggered by the expectation that now fewer chips would be needed in the AI sector. There have been debates about whether the data about the computing power needed was reliable. Furthermore, Open AI claimed to have evidence that the Chinese start-up engaged in "distillation", a process where developers train their language models by using the output of other language models, thus potentially breaching intellectual property rights. Industry participants, specifically from the data centre industry claim that DeepSeek would not harm their demand, but rather, higher efficiency and more open-source orientation would help the development of the industry. Additionally, DeepSeek is not an outlier. Several US universities, such as Stanford have published working papers on models that they have trained for single-digit dollar amounts in under a day. These developments showcase that AI is, as expected, advancing further and further changing the investment landscape as it progresses. Whether data centre developers should be worried about DeepSeek or whether they can adapt to the fast-changing demand structures will become apparent as the market matures.

Conclusion

"There are three ways to make a living in this business: be first, be smarter, or cheat." (Margin Call)

Success is highly based on timing and skill. AWS [NASDAQ: AMZN] is partly successful due to its prior position as a pioneer in the cloud computing space. Microsoft's current presence in the operating system space is a direct result of their MS-DOS (predecessor of Windows) system pioneering the market with new technologies. Google [NASDAQ: GOOGL] similarly dominated the search engine industry through its PageRank algorithm. While being smarter is often much more strenuous than simply being first, it leads to long-term dominance. Open AI was truly the first firm to effectively develop AI. Amazon, Google, and Microsoft are now attempting to catch up through massive capital expenditures in 2025. Amazon is planning to increase CapEx from $83bn (2024) to $100bn focusing on AI/Cloud infrastructure (data centres, GPUs, R&D). Amazon specifically AWS underwent a stellar 2024 with a 19% year-on-year (YoY) revenue growth leading to a $39.8bn net income in 2024 (solely AWS). Likewise, Microsoft, Meta, and Google all experienced an outstanding year in 2024 with large revenue growth around the board. They are expecting their CapEx for 2025 to be $80bn, $65bn, and $75bn respectively. The paradox of growth in the AI space is more CapEx leads to a better product creating more sales. This necessitates more AI infrastructure thus CapEx. As the market is still in its infancy and new customers grow on trees this cycle is bound to continue. Partly this investment can be attributed to this cycle. Like the advent of the internet, AI represents a

once-in-a-lifetime opportunity to create a legacy product. If a firm can get ahead and create a new staple product such as the iPhone. While the margins of the AI space of 50%-60% currently are lower than traditional SaaS 70&-85% margins, high room for growth and margin improvement drive this investment. These firms are betting that high initial investments will lead to long-term high margins. Like last year Google, Microsoft, Meta, and Amazon collectively do not want to miss out on this future cash cow. They are also financially in a unique position to be able to afford these high capital expenditures due to their exceptional 2024. They can sacrifice short-term profits for long-term success. A large question mark remains surrounding Deepseek. If what they claim is true it does pose the question if all this capital deployed by these tech giants is effective. US AI might simply be too inefficient to compete in the long run.

TAGS: Data Centres, AI, Tech