Stargate Project Launches with $500 Billion Commitment—OpenAI, SoftBank, Oracle, and MGX Deploy $100 Billion Immediately for Texas Data Centers

The largest AI infrastructure bet in history was announced with $500 billion in commitments. Seven months later, Bloomberg reported no funds had actually been raised.

The Announcement: Half a Trillion Dollars at the White House

On January 21, 2025, President Trump stood alongside three of technology's most prominent figures—OpenAI CEO Sam Altman, SoftBank CEO Masayoshi Son, and Oracle Chairman Larry Ellison—to announce the Stargate Project. The joint venture commits $500 billion over four years to build AI data center infrastructure across the United States, with

$100 billion deploying immediately for facilities starting in Abilene, Texas.

The numbers are staggering by any measure. For context, $500 billion exceeds the GDP of Sweden. It's roughly equivalent to the entire market capitalization of Meta. The initial $100 billion commitment alone would rank as the largest single infrastructure investment in American tech history.

SoftBank leads the financial side of Stargate LLC, with chairman Masayoshi Son serving as the entity's chairman. OpenAI holds operational leadership. Oracle and MGX (the Abu Dhabi sovereign wealth fund's technology arm) round out the initial equity funders. Technology partners include Microsoft, Nvidia, Oracle, and Arm.

The first two data centers in Abilene, Texas, have one facility targeted for completion by end of 2025. Five additional U.S. sites were announced subsequently, bringing total investment projections toward $400 billion across the expanded footprint. Total planned power capacity: nearly 7 gigawatts. Estimated job creation: 25,000 positions.

President Trump pledged emergency declarations to expedite energy infrastructure for the project—a recognition that power, not capital, may prove the binding constraint.

Why This Matters Beyond the Headlines

The Stargate announcement signals a fundamental shift in how AI infrastructure gets built and financed. For the past decade, hyperscalers—Amazon, Microsoft, Google—have dominated data center construction through internal capital expenditure. Stargate represents an alternative model: a consortium approach pooling capital across technology operators, financial sponsors, and sovereign wealth funds.

This matters because AI compute demand is outpacing any single company's ability to build fast enough.

Consider the trajectory. OpenAI's training runs have increased compute requirements by roughly 10x annually. GPT-4's training consumed an estimated 25,000 Nvidia A100 GPUs for months. The models after GPT-5 will require clusters that don't yet exist. Stargate is the explicit acknowledgment that even Microsoft's $50 billion annual capex can't keep pace

alone.

The geographic concentration in Texas reflects three interrelated calculations. First, Texas operates its own electrical grid (ERCOT), enabling faster permitting for power connections than interconnected grids requiring multi-state coordination. Second, Texas land costs roughly 60-70% less per acre than California or Virginia equivalents for industrial-scale parcels. Third, Texas has the most favorable data center tax incentives of any major state—no corporate income tax, and property tax abatements available for qualifying facilities.

Who wins from Stargate's construction? Nvidia is the obvious beneficiary—7 gigawatts of AI compute at current power densities translates to hundreds of thousands of GPUs. Arm benefits from Nvidia's continued dominance in AI training chips. Oracle gains a massive anchor tenant for its cloud infrastructure ambitions.

Who loses? Potentially Microsoft, which finds its exclusive partnership with OpenAI diluted. Oracle's inclusion as both equity funder and technology partner suggests OpenAI is diversifying its infrastructure dependencies. Microsoft's $13 billion investment in OpenAI bought preferential access, not permanent exclusivity.

The Money Question: Elon Musk vs. Everyone Else

Within hours of the announcement, Elon Musk posted on X: "They don't actually have the money." The comment wasn't idle sniping. Musk runs xAI, OpenAI's most aggressive competitor, and operates his own data center buildout in Memphis consuming over 100,000 Nvidia GPUs.

Musk's skepticism gained credibility when Bloomberg reported in August 2025 that no funds had been raised for Stargate due to market uncertainty and trade policy concerns. The announcement was ambitious. The execution faced reality.

Here's what the financing actually looks like as of mid-2025. SoftBank's first $10 billion will be borrowed from Mizuho and other lenders—announced in April 2025. JPMorgan lent $2.3 billion to OpenAI partners specifically for the Abilene sites in May 2025. That's $12.3 billion in confirmed debt financing against a $100 billion "immediate" commitment.

The gap between announcement and execution isn't unusual for megaprojects. What's unusual is announcing before financing is secured. SoftBank's Masayoshi Son has a history of bold proclamations—the original $100 billion Vision Fund announcement followed similar patterns of ambitious targets followed by gradual capital formation.

The funding structure reveals the project's real nature: a rolling commitment that will scale up or down based on market conditions, AI demand trajectories, and each partner's financial position.

SoftBank carries roughly $180 billion in debt against $100+ billion in investment assets. Oracle's free cash flow runs approximately $15 billion annually. OpenAI remains unprofitable, generating approximately $4 billion in annual revenue against costs that exceed that figure. MGX, backed by Abu Dhabi's sovereign wealth, represents the most reliable capital source.

The $500 billion figure should be understood as a ceiling, not a floor. If AI demand continues on current trajectories and financing markets remain favorable, the full commitment is achievable. If either condition changes, the project scales down. Both Musk's skepticism and the partners' optimism can be simultaneously correct depending on which future materializes.

Technical Architecture: What 7 Gigawatts Actually Means

Power consumption has become the primary constraint on AI infrastructure scaling. The 7-gigawatt capacity planned across Stargate sites deserves unpacking because it reveals the technical architecture required for next-generation AI training.

One gigawatt equals approximately 1,000 megawatts. A typical modern data center runs 50-100 megawatts. A hyperscale facility might reach 200-300 megawatts. Stargate's Texas facilities alone are targeting capacities that would rank among the largest data centers ever built.

Current AI training clusters—like Microsoft's Azure facilities running OpenAI workloads—operate at roughly 100-150 watts per GPU for inference, but 400-700 watts per GPU for training workloads. Nvidia's H100 GPUs draw approximately 700 watts at peak.

The forthcoming Blackwell architecture draws even more.

At 7 gigawatts, Stargate could theoretically power 10 million H100-equivalent GPUs simultaneously—roughly 50x the current global installed base of AI training capacity.

Cooling represents the other critical technical challenge. Traditional data centers use air cooling, which works up to approximately 30 kilowatts per rack. AI training racks now exceed 100 kilowatts per rack. This requires liquid cooling infrastructure—either direct-to-chip liquid cooling or full immersion cooling.

The Abilene sites are reportedly designing for liquid-cooled racks from day one. This isn't optional for AI training at scale—it's a basic requirement. Air-cooled facilities cannot support the power densities that modern AI training requires.

Power sourcing for 7 gigawatts presents its own challenges. For reference, the entire ERCOT grid serves approximately 85 gigawatts of peak demand. Stargate's full buildout would represent roughly 8% of Texas's entire grid capacity. The emergency declarations Trump pledged would likely involve expediting transmission line construction and potentially new generation sources.

Natural gas peaker plants can be built in 18-24 months. Solar plus battery installations can come online faster but with lower capacity factors. Nuclear—the preferred option for continuous baseload power—takes 7-10 years minimum under current regulatory frameworks. The power question may ultimately constrain Stargate more than capital availability.

What Most Coverage Gets Wrong

The prevailing narrative frames Stargate as either transformative (the administration's view) or vaporware (Musk's characterization). Both miss the more interesting middle ground.

Wrong Take #1: This is primarily about AI competition with China

The announcement framing emphasized national security and competing with global AI powers. While accurate as a political positioning, it obscures the commercial reality.

Stargate's immediate purpose is providing OpenAI with training infrastructure independent of Microsoft's exclusive control. The national security framing secures regulatory cooperation and political cover. The commercial motivation drives the actual investment.

Wrong Take #2: The $500 billion number matters

It doesn't. What matters is the first $20-30 billion actually deployed and whether it generates returns sufficient to attract the next tranche. Infrastructure investments of this scale happen in phases, each dependent on the success of prior phases. Amazon Web Services didn't announce $500 billion in 2006—it built iteratively based on demonstrated demand. Stargate will follow the same pattern regardless of announced totals.

Wrong Take #3: This represents government industrial policy

The federal government isn't investing capital. It's providing regulatory accommodation—expedited permitting, emergency declarations for power infrastructure, potentially favorable treatment under export control regimes. This is closer to the public-private partnerships that built American railroads and highways than to direct government investment in specific technologies.

What's underhyped: The Oracle inclusion

Oracle's participation as both equity investor and technology partner signals something important about the cloud infrastructure landscape. Oracle Cloud Infrastructure has grown rapidly but remains a distant fourth behind AWS, Azure, and Google Cloud. Stargate gives Oracle a pathway to hosting the most important AI workloads in the world—OpenAI's training runs.

Microsoft's Azure has hosted OpenAI workloads exclusively under their partnership agreement. Oracle's inclusion suggests that agreement is evolving. For enterprise CTOs evaluating cloud strategy, this signals Oracle as a credible option for AI workloads in ways it wasn't six months ago.

What Technical Leaders Should Do Now

For CTOs, senior engineers, and tech founders reading this, Stargate creates several actionable considerations.

Evaluate GPU supply chain implications

Stargate will consume massive quantities of Nvidia hardware over the next four years. If your organization depends on acquiring Nvidia GPUs or cloud instances running them, expect continued supply constraints and elevated pricing. The inference market—which uses GPUs less intensively than training—may see better availability as hyperscalers prioritize training clusters.

Consider AMD's MI300X series and Intel's Gaudi accelerators as alternatives for inference workloads. They're not drop-in replacements, but they're increasingly viable for production inference at 60-70% of Nvidia's price per performance.

Watch Oracle Cloud Infrastructure

If Stargate succeeds, OCI becomes the de facto home for cutting-edge AI training. Oracle has historically offered aggressive pricing to win enterprise workloads. Organizations with substantial cloud spend should evaluate OCI's AI capabilities—the Stargate announcement indicates Oracle is betting heavily on this market.

Plan for power as a constraint on AI adoption

The power intensity of modern AI creates real constraints on deployment options. Organizations running AI workloads on-premises face the same physics—you cannot run modern GPU clusters without adequate power and cooling infrastructure.

If your data center was built more than five years ago, it probably cannot support modern AI training densities without significant retrofitting.

Cloud deployment avoids this problem by shifting it to the hyperscaler. Hybrid approaches—cloud for training, edge for inference—may prove optimal for organizations

with existing data center investments.

Consider the regulatory vector

Stargate's emergency declaration pathway suggests AI infrastructure will receive preferential regulatory treatment under the current administration. Organizations planning AI infrastructure investments should factor this into timing decisions. The permitting environment for data centers, power connections, and related infrastructure is likely more favorable now than it will be under future administrations with different priorities.

Where This Goes in 12 Months

By January 2026, we'll know whether Stargate is real or vaporware. Specific predictions:

The first Abilene facility will be operational but at reduced capacity from initial targets. First-generation megaprojects consistently face delays—expect 60-70% of planned capacity online by year-end 2025.

Total deployed capital will reach $25-40 billion, not $100 billion. This still represents the largest AI infrastructure investment in history and provides sufficient capacity for OpenAI's near-term training needs. The gap between announced and deployed capital will generate continued skepticism without materially affecting the project's strategic value.

Microsoft's relationship with OpenAI will formalize a transition from exclusive to preferential. Microsoft will remain OpenAI's primary cloud partner for inference and API serving, but training workloads will increasingly run on Stargate infrastructure. The equity relationship will evolve accordingly.

Power constraints will emerge as the binding bottleneck. Capital is available. Chips are constrained but available. Land is available. The 7-gigawatt power requirement cannot be satisfied by existing infrastructure. Emergency declarations will help but cannot conjure new power plants into existence. The 2029 timeline depends more on power generation buildout than on any other factor.

At least one additional technology partner will join, likely from the semiconductor supply chain. TSMC or Samsung as a manufacturing partner, or a major interconnect provider like

Arista or Broadcom, would strengthen the consortium's ability to execute at the announced scale.

The Bigger Picture

Stargate represents a bet that AI training compute will remain valuable enough to justify unprecedented infrastructure investment. That bet requires believing several things simultaneously: that larger models continue improving, that training (not just inference) remains compute-bound, that the competitive dynamics of AI development justify this spending, and that the consortium partners can maintain cooperation across a four-year buildout.

Each of those beliefs has serious counterarguments. Scaling laws may plateau. Algorithmic improvements may reduce compute requirements. Regulatory intervention may constrain AI development. Partner interests may diverge as the market evolves.

What isn't in question is the scale of the ambition. Whether through Stargate or alternative approaches, someone will build the infrastructure to train the next generation of AI systems. The U.S. partners assembling this consortium are attempting to ensure that infrastructure gets built domestically, operates under American legal frameworks, and benefits American companies.

Elon Musk is correct that the partners don't have $500 billion sitting in bank accounts. Sam Altman and Masayoshi Son are correct that financing at this scale is achievable over four years if the underlying asset generates sufficient returns. The next twelve months will determine which version of reality materializes.

For technical leaders, Stargate's importance lies not in its announced scale but in what it signals: AI infrastructure has become strategic in ways that justify historically unprecedented investment, and the race to build it is now the central contest in technology.

WRITTEN BY

Artur Markus

Software engineer in Basel building AI infrastructure and agentic systems, with a background in pharmaceutical automation, GxP compliance, and validation. I help teams turn AI from a chatbot into a reliable operational layer.

Book a meeting