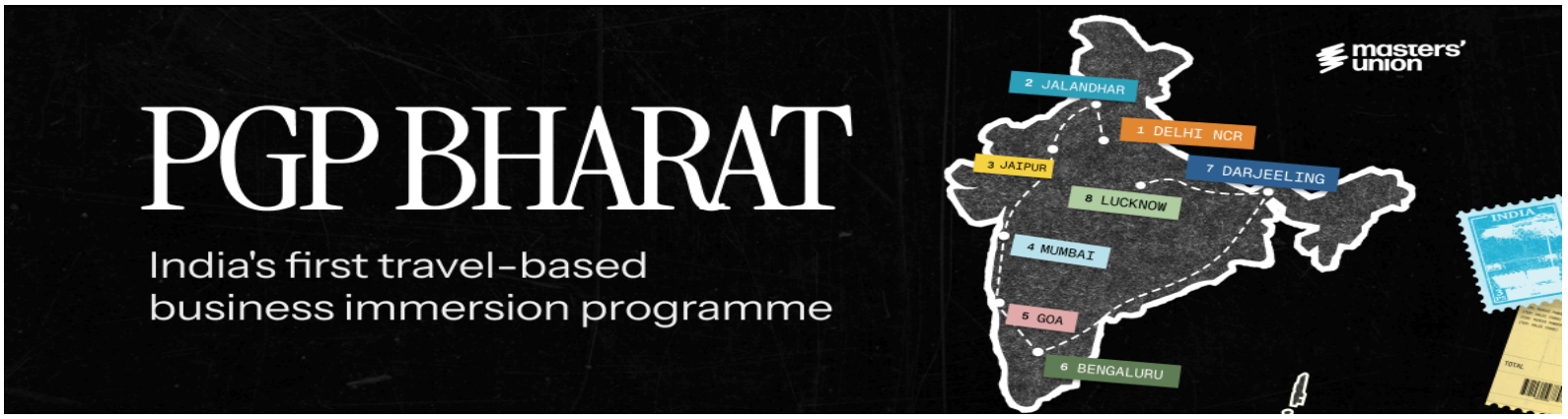


1. Why does this immersion matter?

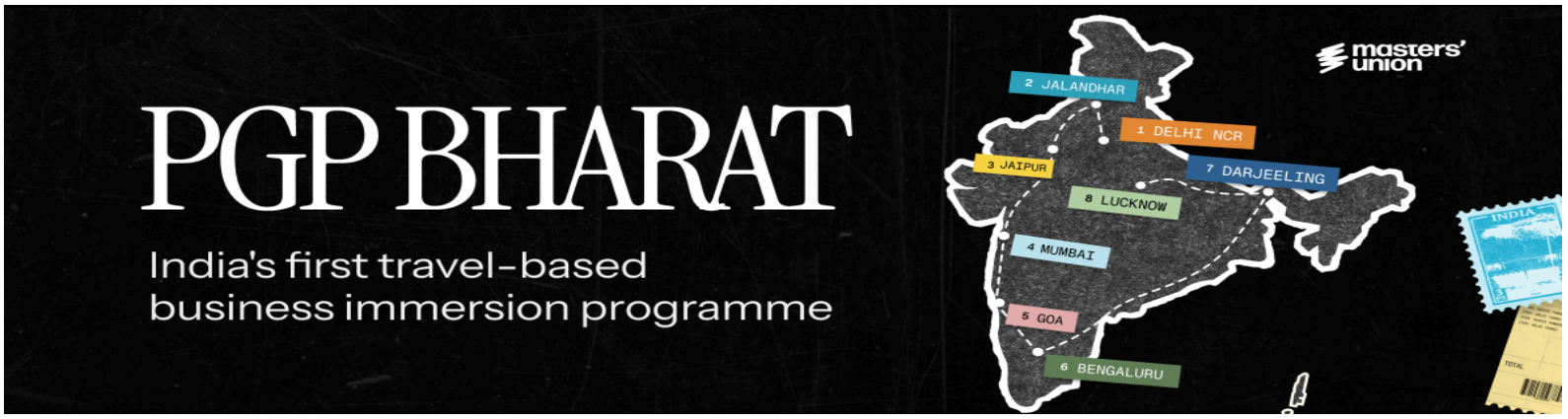
Every AI product you've used this year, whether it's a chatbot, an image generator, or a recommendation engine, needed somewhere to actually run. Not metaphorically. Physically. Racks of GPUs, pulling megawatts of power, cooled around the clock, sitting inside a building somewhere. Most MBA students think about AI as a product or a strategy question. Very few have stood inside the building where the compute actually lives.

That's what this visit is. Yotta's NM1 facility in Navi Mumbai is Asia's largest Tier IV data centre, certified by the Uptime Institute for Gold Operational Sustainability. It's also the physical home of Shakti Cloud, the GPU platform that Yotta is betting will make it one of India's key AI infrastructure providers as the country tries to build compute capacity it doesn't currently have enough of.

The story worth paying attention to here isn't just "big data centre, lots of GPUs". It's a company making an enormous, expensive bet years ahead of confirmed demand in a business where the government is simultaneously your biggest customer and the reason margins might get squeezed to nothing. Yotta currently operates over 10,000 GPUs at its NM1 data centre and is on track to scale its AI compute capacity to more than 80,000 next-generation GPUs by FY27-28 across its two campuses. That's roughly an 8x jump in AI compute capacity within about two years, built on infrastructure that takes years to plan and construct.

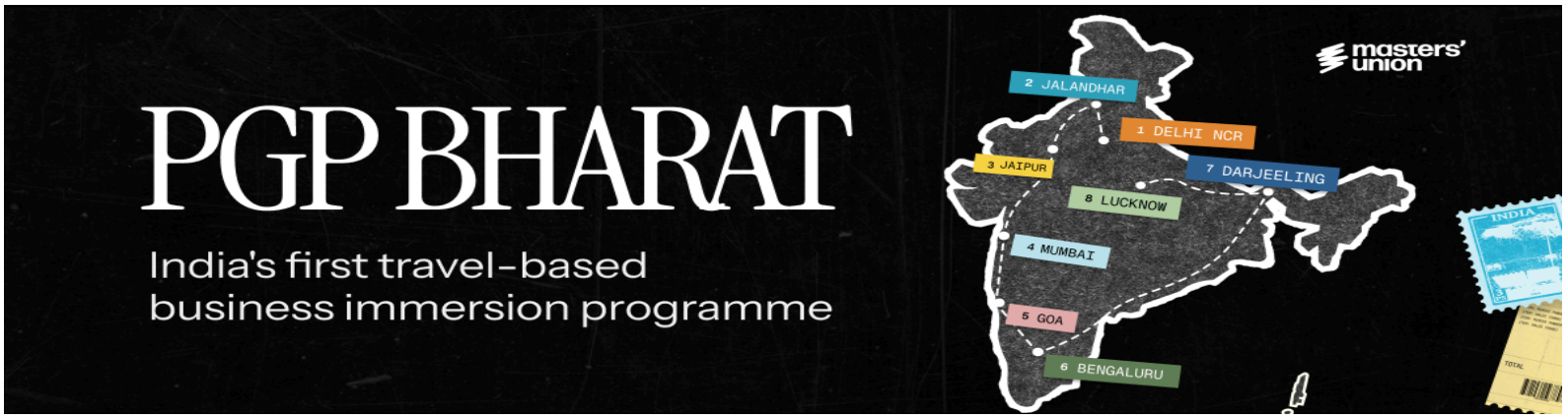
This pre-read gives you a way to think about that bet before you walk the floor, not a script to recite once you're there.

YOTTA



2. Company snapshot

#	Detail	Information
1	Organisation Name	Yotta Data Services Private Limited
2	Industry/Sector	Cloud infrastructure, data centres, AI compute
3	Ownership	A 100% subsidiary of the Hiranandani Group
4	NM1 Location	Panvel, inside the 600-acre Hiranandani Fortune City, Navi Mumbai
5	NM1 Facility Size	820,000 square feet across 16 floors, housing up to 7,200 racks
6	NM1 IT Power Capacity	36 MW, part of a campus scalable to 1 GW
7	Certification	Uptime Institute Tier IV, Gold Operational Sustainability certified; world's second-largest Uptime Institute-certified Tier-IV facility
8	GPU Compute (current)	Over 10,000 GPUs operating at NM1
9	Flagship Platform	Shakti Cloud (GPU-as-a-Service, AI-HPC compute)
10	Other Campuses	Greater Noida (D1), GIFT City Gujarat (G1)
11	Website	yotta.com



3. Why is Yotta spending on 70,000 more GPUs before it knows demand will show up?

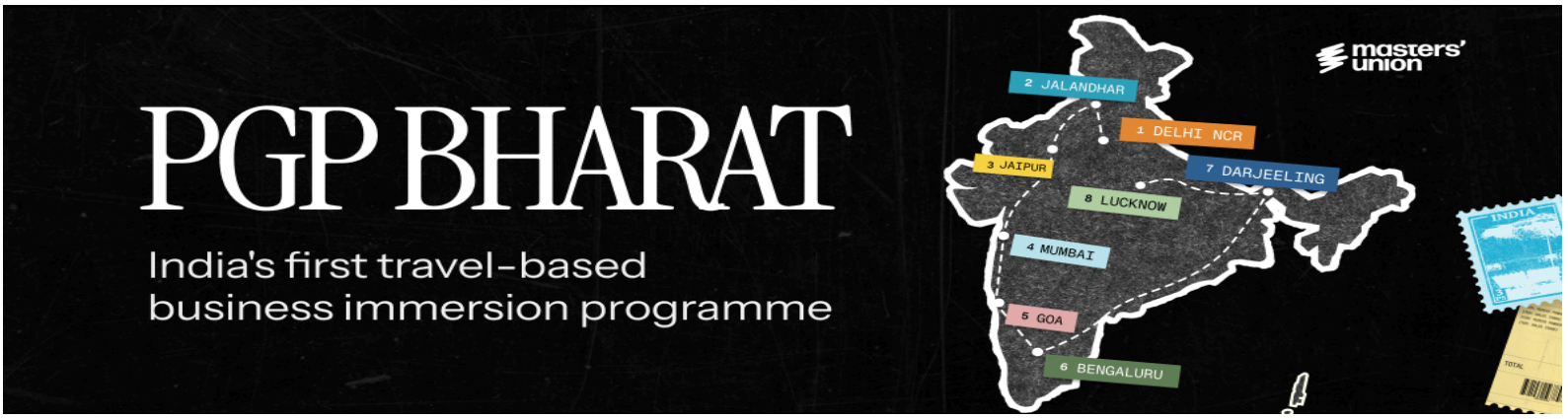
The 3 Strategies for Capacity Planning

Every capacity decision in operations comes down to a simple, uncomfortable question: do you build ahead of demand, behind it, or exactly alongside it? This is one of the oldest and most reliable frameworks in operations management, described in Chase, Jacobs and Aquilano's *Operations and Supply Chain Management* and in Nahmias' *Production and Operations Analysis*, and it comes down to three strategies.

- A **lag strategy** means you wait until demand is confirmed before you add capacity. It's the safest option for your cash, but you risk turning customers away while you catch up, and a competitor can fill that gap before you do.
- A **match strategy** means you add capacity in small, closely tracked increments as demand grows. It's balanced, but it only works if your demand forecasting is genuinely good, which is hard in a market moving as fast as AI compute.
- A **lead strategy** means you build capacity well ahead of confirmed demand, betting that growth will show up and fill it. It's expensive and risky if you're wrong, but if you're right, you own your market position before anyone else gets theirs.

Yotta's GPU scale-up is a lead strategy. The first cluster of NVIDIA H100 GPUs, more than 4,000 units, arrived at NM1 in March 2024. Since then, capacity has kept climbing: Yotta committed to deploying a cluster of 16,384 GPUs at NM1, and separately, Yotta deployed DDN infrastructure to power 8,000 NVIDIA B200 GPUs as part of its AI cloud buildout. All of this is happening while the company is still filling out its current 10,000-plus GPU base, on the way to the 80,000-plus target by FY27-28.

Here's what makes this genuinely a bet rather than a safe expansion. GPUs are expensive, power-hungry, and depreciate as newer chips get released. If Indian enterprises' and the government's demand for AI compute doesn't grow as fast as Yotta expects, a large share of that capacity sits idle, burning capital and power costs without generating matching revenue. If demand does grow that fast, Yotta becomes one of the few providers in India with enough scale to serve it, which matters enormously in a business where switching infrastructure providers mid-project is disruptive and expensive for a customer.



The question worth carrying onto the floor: what does Yotta's leadership actually believe has to be true about India's AI compute demand over the next two to three years for this bet to pay off, and what are they seeing that makes them confident enough to build ahead of it?

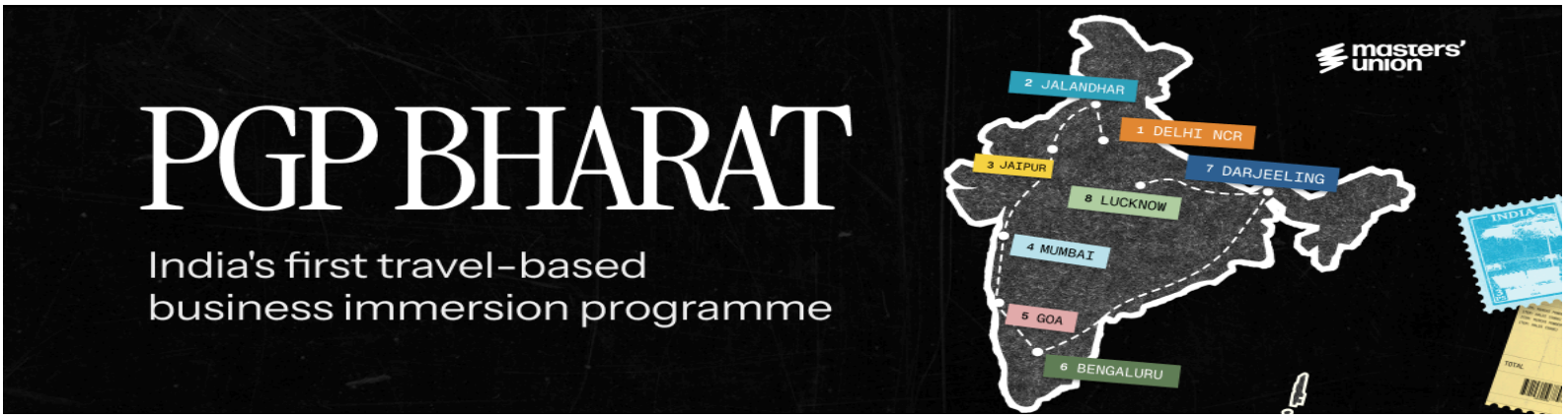
4. Why does a data centre run its own power substation instead of just buying electricity like everyone else?

Most businesses buy the inputs they need from someone else who specialises in making them. A restaurant doesn't farm its own vegetables. A car company usually doesn't mine its own steel. The decision of what to build in-house versus what to buy from the market is one of the oldest questions in economics, first framed clearly by Ronald Coase in his 1937 paper "The Nature of the Firm" and later extended by Oliver Williamson's work on transaction cost economics in the 1970s and 80s. The core idea: you bring something in-house when the cost, risk, or coordination problems of relying on an outside supplier outweigh the cost of doing it yourself.

Power is exactly this kind of decision for a data centre. NM1 doesn't just plug into the grid and hope for the best. It runs N+1 redundancy on upstream transformers, a 4+1 setup, distributes power through its own 33kV feeders and 33/11kV substations, runs N+2 redundancy on downstream transformers with a 30+2 setup, and operates its own power distribution infrastructure under its own operational licence.

That's a lot of engineering and regulatory effort to avoid depending entirely on the external grid. The reasoning is straightforward once you think about what a power failure actually costs a Tier IV data centre: not just lost revenue, but broken trust with enterprise and government customers who chose Yotta specifically because it promised close to zero downtime. A generic utility connection can't guarantee that level of reliability on its own, so Yotta absorbed the cost and complexity of building redundancy in-house rather than depending on someone else's infrastructure.

The question to test on the floor: which parts of NM1's operations follow this same logic of "we'll control it ourselves because outside reliability isn't good enough," and which parts does Yotta happily outsource because someone else already does it better or cheaper?



5. Why does one weak link matter more than a hundred strong ones?

A data centre isn't one system; it's a chain of dependent systems: power arrives, gets converted and distributed, cooling keeps the hardware at a safe temperature, and only then can the actual compute, the GPUs, do useful work. The Theory of Constraints, developed by Eliyahu Goldratt in his 1984 book *The Goal* and now a standard part of operations management courses alongside capacity planning, makes a simple but easy-to-forget point: the output of any chained system is limited by its single weakest link, not by the average strength of the whole chain. Adding capacity anywhere except at that weak link doesn't actually help.

This matters enormously for a facility like NM1. You can have thousands of idle GPUs sitting in racks, but if the power supply, the cooling system, or the network fabric connecting the GPUs to each other can't keep up, the extra GPUs don't translate into extra usable compute. In AI workloads specifically, GPUs often need to talk to each other constantly during training, so a slow network link between racks can become the bottleneck just as easily as raw power capacity.

This is also why "how many GPUs do you have" is an incomplete question to ask on the floor. The better question is what Yotta considers its actual constraint right now: is it GPU supply from NVIDIA, is it power availability, is it cooling capacity, or is it something less visible like skilled engineers to manage GPU clusters at this scale? Whatever that answer is tells you more about where the business is actually headed than the headline GPU count does.

6. Why is the same government programme both Yotta's biggest opportunity and its biggest margin threat?

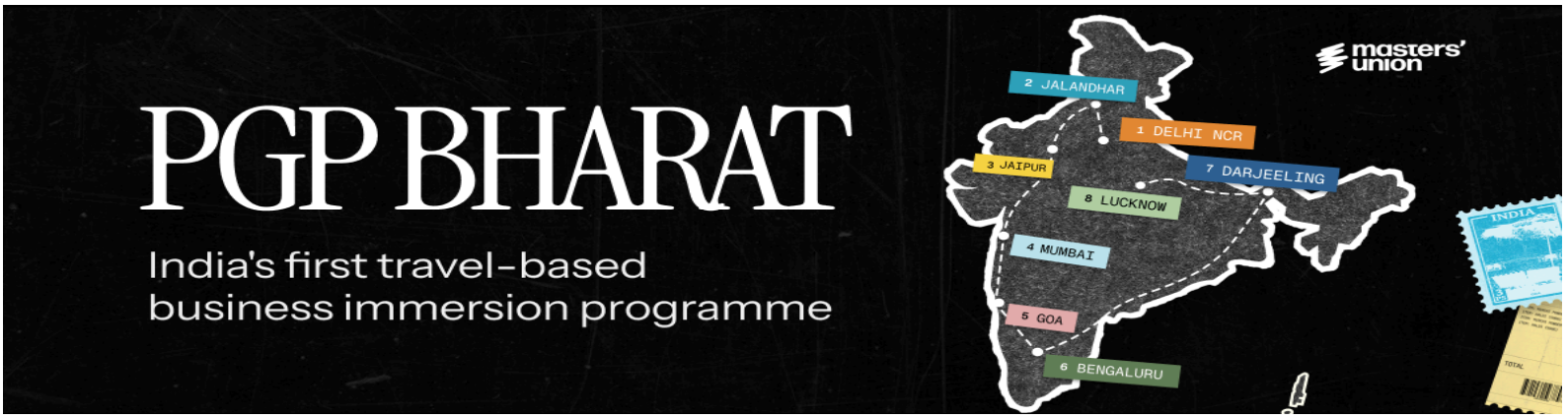
An industry's competitive intensity shapes what a smart strategy actually looks like inside it. Applied here: the IndiaAI Mission is a Rs 10,300 crore government programme to build sovereign AI infrastructure, and it is simultaneously the reason Yotta has a guaranteed customer and the reason its pricing power is under pressure.

Here's the tension. The government's Request for Empanelment invited applications from agencies to provide AI compute and cloud services, aiming to make 10,000 GPUs available to support India's AI mission, and a total of 19 bidders submitted proposals, with 10, including Yotta, CtrlS, Jio Platforms, E2E Networks, and Tata Communications, qualifying. By mid-2026, the mission had empanelled more than 38,000 GPUs, well past its original target. That's a lot of well-funded competitors chasing the same government contracts, all offering broadly similar hardware from the same handful of chipmakers.

The pricing reflects that competition. The lowest winning bids came in at ₹115.85 per GPU-hour for standard GPUs and ₹150 per GPU-hour for H100-class GPUs, with an additional government subsidy of up to 40% bringing effective end-user rates to roughly ₹67 and ₹92 per GPU-hour. That's a government-set price ceiling that every empanelled provider, including Yotta, has to operate under for this segment of demand, regardless of their own cost structure.

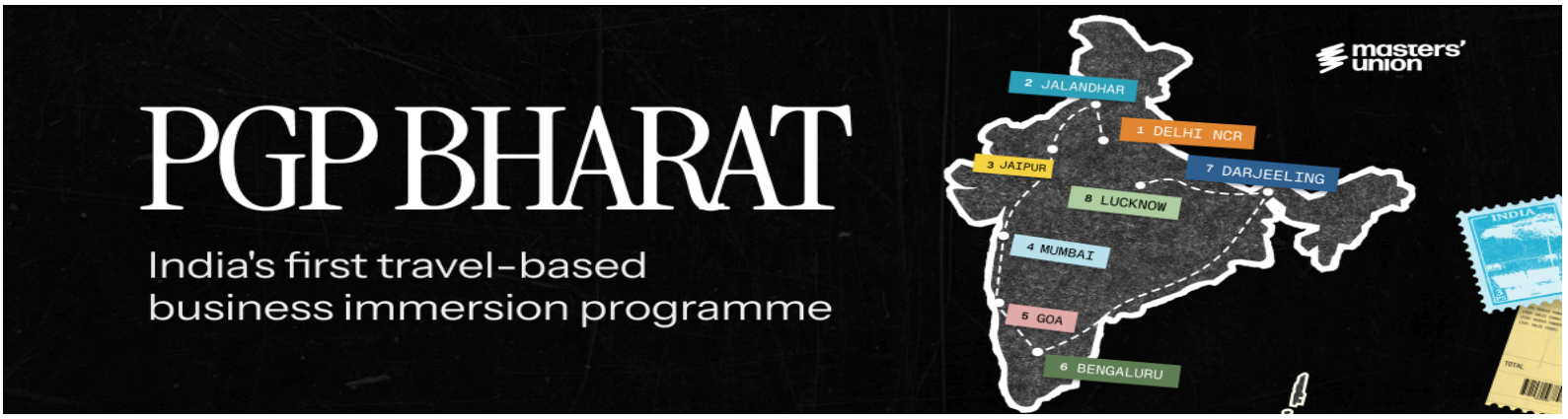
So Yotta gets guaranteed large-scale demand through the IndiaAI Mission, which justifies building capacity ahead of the open market. But that same demand comes at government-negotiated rates that leave little room for differentiation on price. Meanwhile, not every empanelled provider has executed at the same pace, with Jio Platforms and CtrlS Datacenters facing delays in deploying their GPUs while Yotta, NextGen, and E2E Networks made significant headway in installing and commissioning theirs. That execution gap, not the contract itself, is where Yotta's actual competitive advantage seems to be showing up.

The question worth asking on the floor: if the government sets the price for a meaningful chunk of demand, where does Yotta expect to make its real margin, on government contracts, on enterprise customers paying market rates, or on some other layer of the business entirely?



7. Glossary

Term	Meaning
Capacity Strategy (Lead, Lag, Match)	A framework for deciding whether to build capacity ahead of demand (lead), behind confirmed demand (lag), or in small increments alongside demand (match).
Tier IV (Uptime Institute)	The highest data centre reliability rating, requiring fully redundant systems so that no single equipment failure causes downtime.
N+1 / N+2 Redundancy	A design principle where a system has one (N+1) or two (N+2) more units than the minimum needed, so equipment can fail or be maintained without stopping operations.
GPU-as-a-Service	A business model where customers rent access to GPU computing power by the hour, rather than buying and maintaining their own hardware.
Transaction Cost Economics	A framework, from Coase and Williamson, for deciding whether a business should build a capability in-house or buy it from an outside supplier, based on which option has lower total cost and risk.
Theory of Constraints	A framework stating that the output of any linked system is limited by its single weakest link, and that improving anything else doesn't increase overall output.
Sovereign Cloud	Cloud infrastructure that keeps data, processing, and control physically located within a country's own borders and legal jurisdiction.
IndiaAI Mission	A Government of India programme that subsidises AI compute access for startups, researchers, and government entities, delivered through empanelled private data centre providers.
MeitY	Ministry of Electronics and Information Technology, the government body running the IndiaAI Mission.

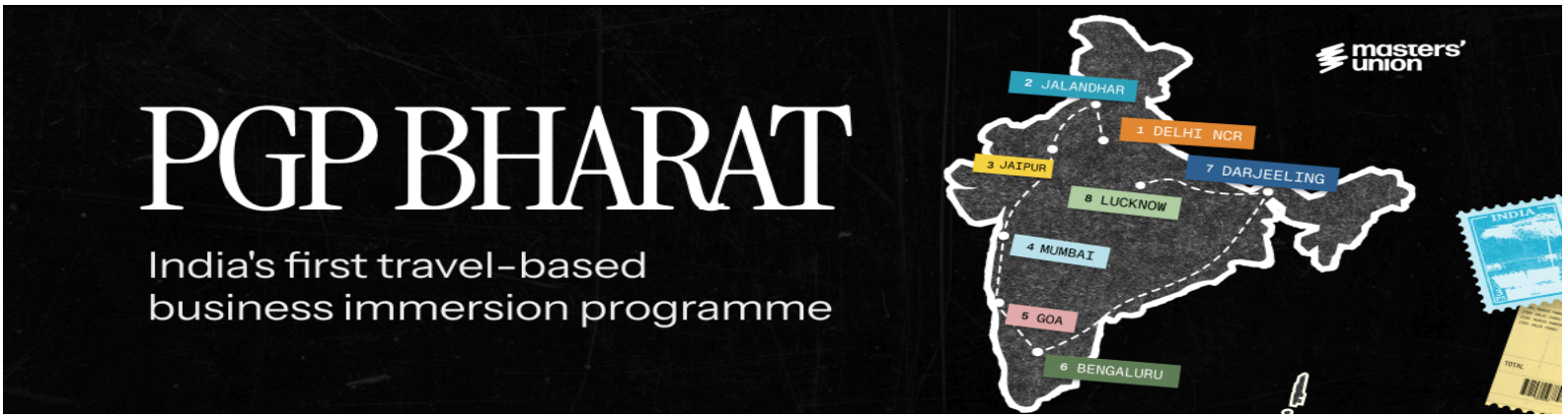


8. Guiding Questions

1. If demand for AI compute in India grows slower than expected, what does Yotta do with tens of thousands of unused GPUs?
2. How much of Yotta's GPU capacity is committed to IndiaAI Mission government contracts versus open enterprise demand, and how does that mix affect margins?
3. As more competitors get empanelled and GPU prices get pushed down by government bidding, what stops this business from becoming a commodity?
4. Yotta scaling from roughly 10,000 to 80,000+ GPUs is a huge jump in a short window. What has to be true about hiring, training, and organisational capacity for a company to operate at that scale, not just the hardware?
5. If power, cooling, and GPU supply are all potential bottlenecks, which one is Yotta actually most worried about right now, and how would that change the next capacity decision they make?

9. Resources to go through before Immersion:

1. Will Sunil Gupta's Yotta successfully build India's own sovereign cloud?
2. Meet Yotta Data Services, an Indian startup taking big AI bets
3. Yotta Built India's First Adapted AI Infrastructure With Shakti Cloud
4. Yotta raises $150 million at $3.9 billion valuation to fund AI infrastructure expansion - The Economic Times



10. Sources

1. Yotta Data Services. "Yotta Data Services Collaborates with NVIDIA to Catalyze India's AI Transformation." yotta.com/press-releases
2. Yotta Colocation. "Inside Yotta NM1: Asia's Largest Tier IV Data Center." colocation.yotta.com/blog/inside-yotta-nm1-asias-largest-tier-iv-data-center
3. Baxtel. "Yotta Navi Mumbai NM1 Data Center." baxtel.com/data-center/yotta-navi-mumbai-nm1
4. CXOToday. "Yotta receives India's first-ever consignment of NVIDIA H100 GPUs at its NM1 data center park." cxotoday.com, March 19, 2024
5. PR Newswire. "Yotta Data Services Receives Frost & Sullivan's 2026 Indian Company of the Year Recognition for Leadership in Sovereign AI Infrastructure." June 23, 2026
6. Yotta Data Services. "Yotta Empaneled in India AI Mission to Accelerate AI Adoption with Advanced GPU & AI Cloud Services." yotta.com/press-releases, February 17, 2025
7. IndiaAI.gov.in. "MeitY Accelerates Empanelment of Agencies to Provide AI Compute and Services for the IndiaAI Mission." indiaai.gov.in
8. Economic Times, via CBInsights. "IndiaAI plan moves forward with over 17,000 GPUs successfully installed."
9. Ecorpit. "GPU cloud pricing in India 2026: H100 to B200 compared." ecorpit.com
10. Business Wire. "DDN and Yotta Power Sovereign AI for India with Multi-Million-Dollar EXAScaler Deployment." businesswire.com, October 15, 2025
11. NVIDIA. "India Fuels Its AI Mission With NVIDIA." blogs.nvidia.com
12. NVIDIA Customer Stories. "Yotta Built India's Sovereign AI Infrastructure with Shakti Cloud." nvidia.com/en-gb/customer-stories
13. Coase, R. (1937). "The Nature of the Firm." *Economica*, 4(16), 386-405.
14. Goldratt, E. (1984). *The Goal: A Process of Ongoing Improvement*. North River Press.
15. Chase, R., Jacobs, F.R., Aquilano, N. *Operations and Supply Chain Management*.
16. Nahmias, S. *Production and Operations Analysis*.