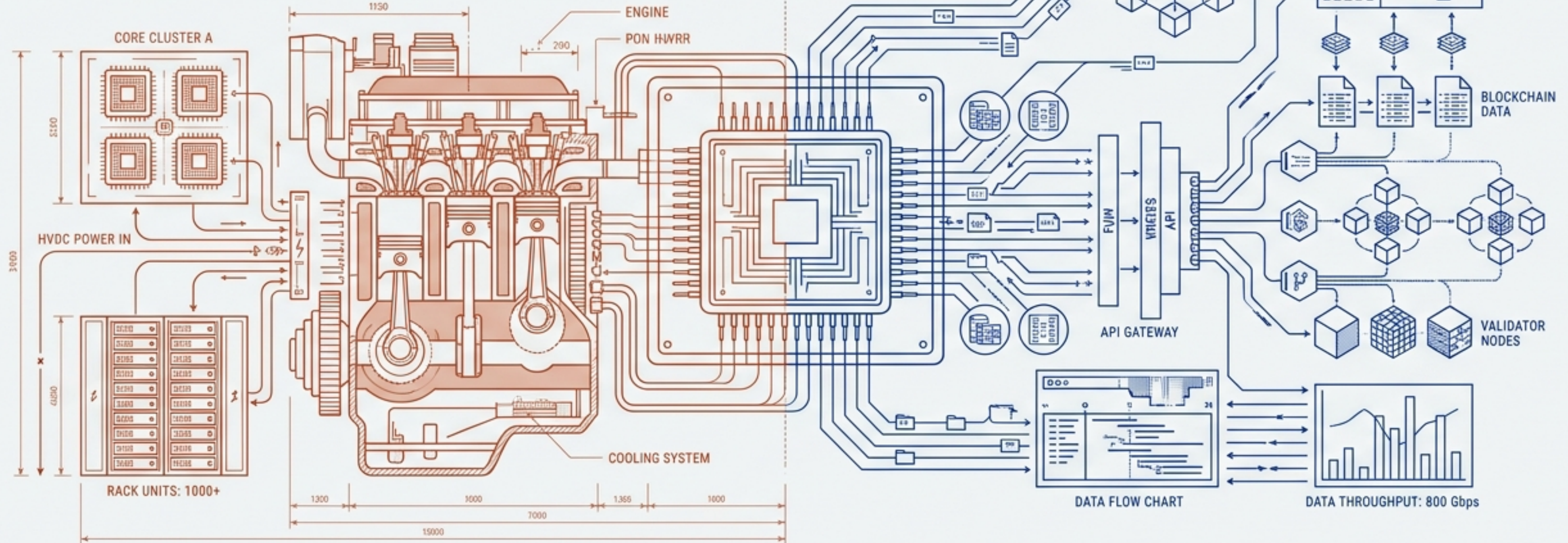


The Sovereign AI Factory

Reliance Intelligence Commercial Strategy:
Architecture of a Durable Revenue Engine



The Mandate: Convert compute into cash faster than it depreciates.

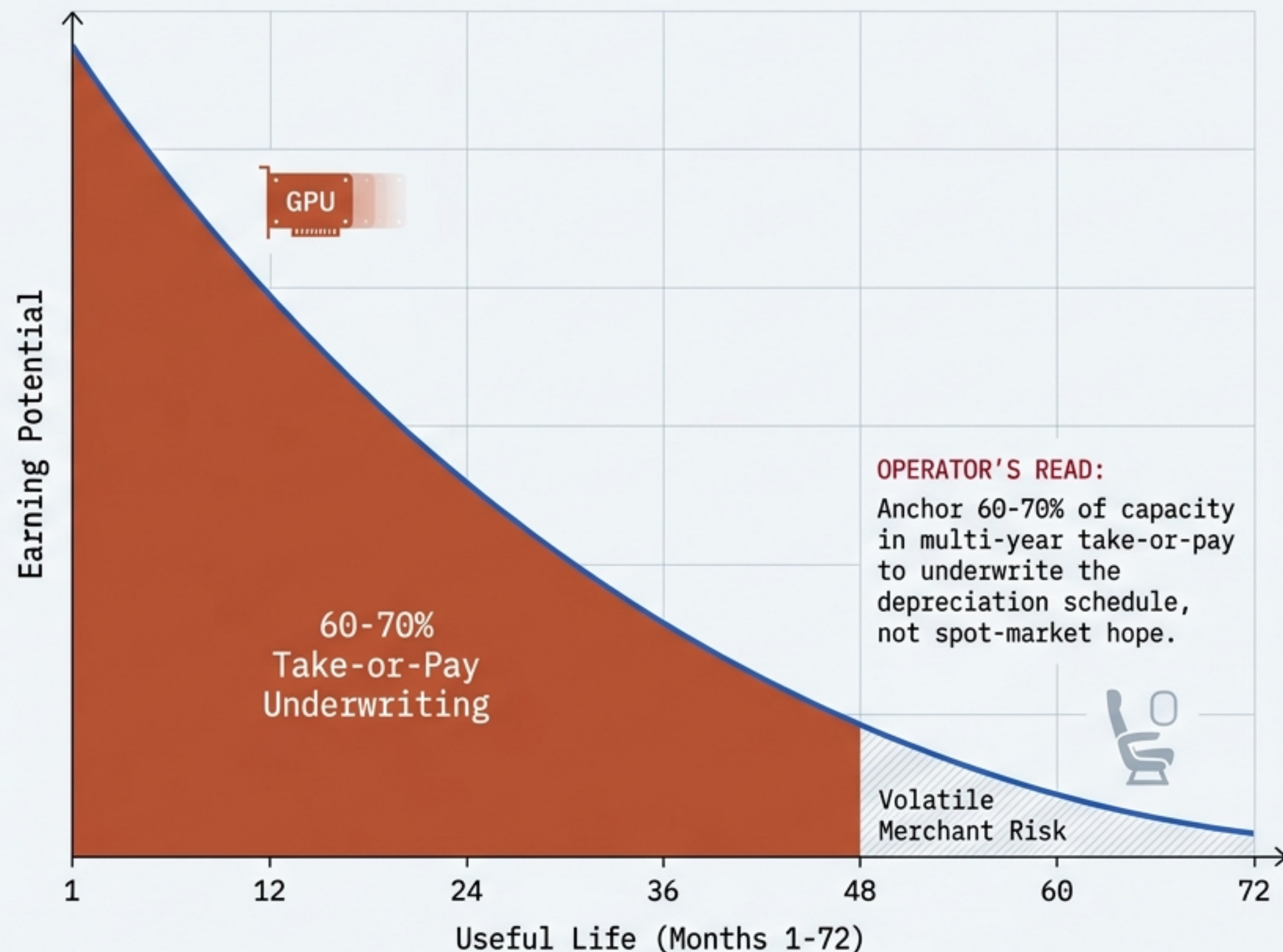
Reliance is building 120 MW of AI compute at Jamnagar, scaling past 200,000 H100-equivalents.

GPU infrastructure is a wasting asset. Every idle GPU-hour is perishable inventory.

The Five Operating Principles

1. Sell certainty first.
2. Price in tokens, cost in GPU-hours.
3. Captive demand is the launch, not the plan.
4. Billing is a product.
5. Sovereignty is the moat.

The Perishable Inventory Curve



Market Context: The Structural Right to Win

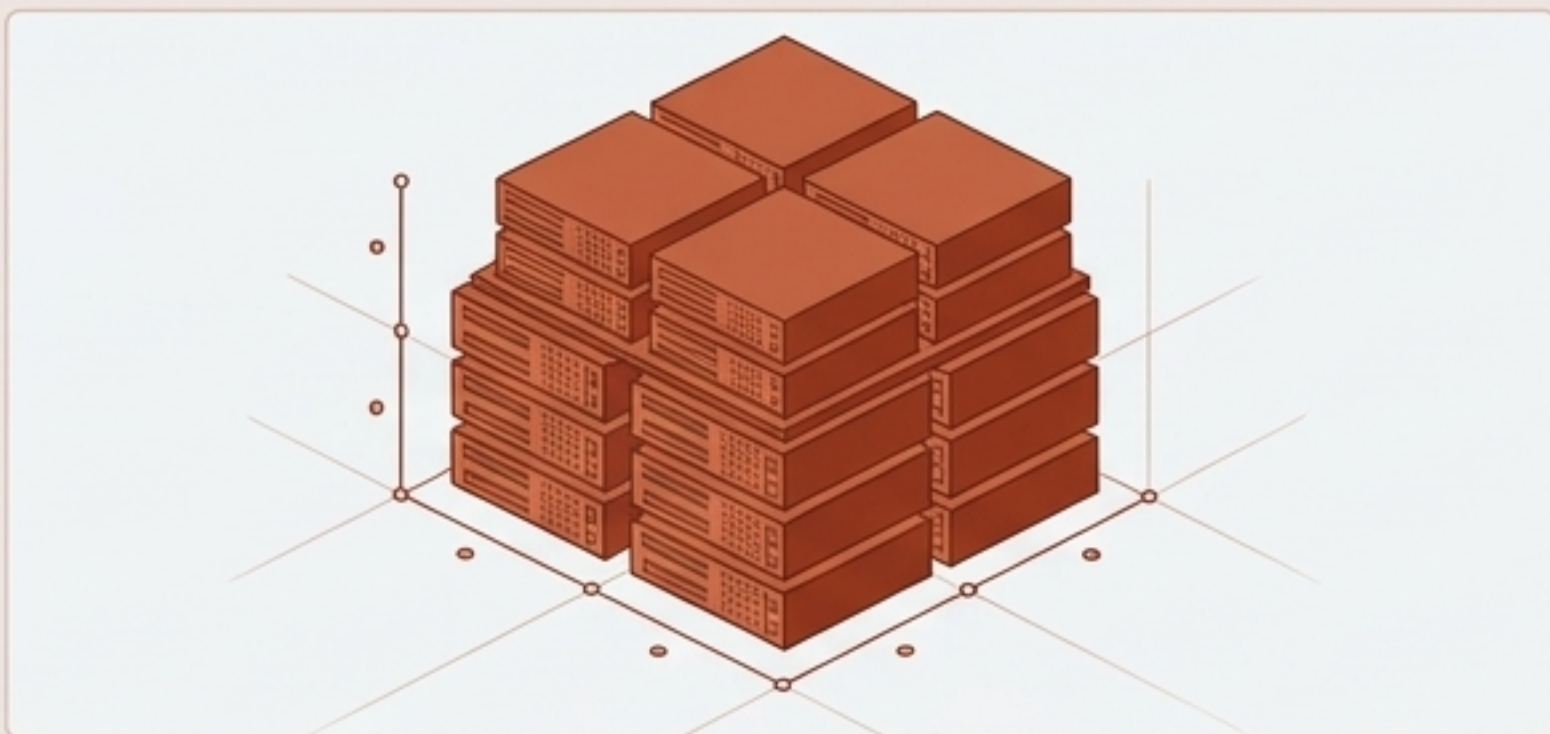
OPERATOR'S READ: Sovereignty and Edge are the true moats. Well-capitalized competitors can buy GPUs, but they cannot easily replicate a national telecom's regulatory footprint and Indic-language model quality on Indian soil.

The Competitive Anatomy Matrix

	Hyperscalers (AWS/Azure)	Domestic Neoclouds	Telecom-Adjacent	Subsidized (IndiaAI)	Reliance Intelligence
	Wins on brand/ecosystem. Fails on Rupee pricing and edge latency.	Wins on price. Fails on scale economics and captive power.	Wins on relationships. Fails on purpose-built AI stack.	Wins on pilot pricing. Fails on enterprise SLAs.	Wins via Kutch renewable power, sovereign data residency, Jio edge latency, and 500M+ consumer/enterprise distribution.
Price/Cost Floor	✓ Wins on Rupee pricing and ₹1.0.M+	✗ Fails on ₹3.000n price	✓ Wins on telecom- relationships	↑ Wins on pilot pricing	✓ Wins via Kutch renewable power
India-Sovereignty & DPDP Compliance	✗ Fails on ECRT DPDP agender and DPDP Compliance	✗ Fails on scale economics and ECNE	✗ Fails on purpose- built AI stack	✗ Fails on enterprise exterpt SLAs	✓ Jio sovereign data residency, 500M+ consumer/enterprise
Latency/Edge Access	✓ Winss on Rupee pricing and edge latency	✗ Fails on scale economics and captive power	✓ Wins on purpose- purpose-built AI stacks	✗ Wins on pricing Fails on enterprise SLAs	✓ Wins via Kutch residency, Jio edge latency, and latency
Enterprise Distribution	✓ Wins on brand/ Fails on bae/sem/ rupee ecosystem	✗ Wins on price Fails on economics and captive power	✓ Wins on relationships Fails on purpose- built AI stack	✗ Wins on pilot pricing Fails on enterprise SLAs	↑ 500M+ setan, and 500M+ consumer/ enterpr distribution

The Commercial Engine: Two Wrappers, One Asset Base

Capacity Business



- **What is sold:** GPU-hours (Take-or-pay clusters)
- **Utilization Risk:** Held by the customer
- **Revenue Quality:** Ratable (ASC 606), supported by RPO/Backlog. Underwrites debt
- **Margin Profile:** 30-45% (Stable floor)

Token Business



- **What is sold:** API Inference (Per 1M Tokens)
- **Utilization Risk:** Held by the platform
- **Revenue Quality:** Usage-based, highly sensitive to efficiency and rate-card deflation
- **Margin Profile:** 50-70%+ (Yield management premium)

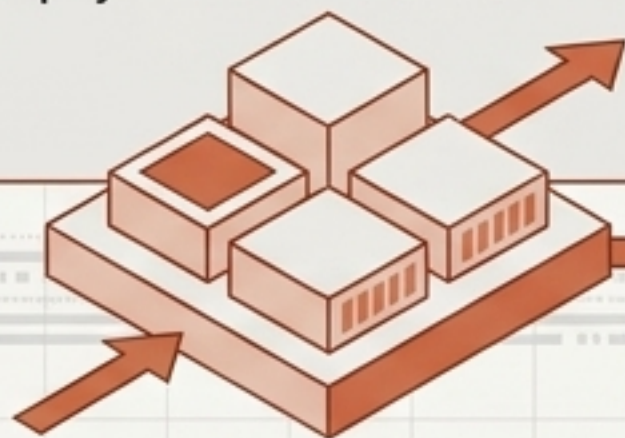
OPERATOR'S READ

The strategic difference is who owns the utilization risk. Capacity contracts transfer risk to the buyer to fund the build; Token APIs retain the risk but charge a massive premium for traffic density.

Go-To-Market Sequencing: From Certainty to Flywheel

Wave 1: Certainty & Utilization (Launch)

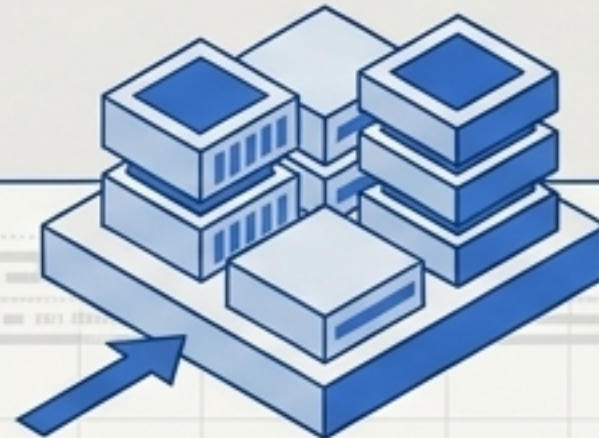
- **SKUs:** Reserved GPU Capacity, Token-as-a-Service.
- **Motion:** Captive Reliance demand (JioBrain, Retail) sets the utilization floor. Anchor external tenants (CEO-to-CEO motion) lock multi-year take-or-pay.



Captive & Anchor Demand

Wave 2: Margin & Differentiation (+2 Quarters)

- **SKUs:** Fine-Tuning, Private Enterprise Endpoints, Telecom Edge Inference.
- **Motion:** Target BFSI and Govt with 30-50% sovereignty premiums. Time-boxed 30-45 day sales cycles.



Specialized Sovereign
Endpoints & Edge

Wave 3: The Ecosystem Flywheel (+12 Months)

- **SKUs:** Agent Runtime, Marketplace Revenue Share.
- **Motion:** 10+ models on marketplace beta. 80/20 standard revenue split, improving to 70/30 for premium telecom edge bundles.



Ecosystem Marketplace
& Revenue Share

Portfolio Economics: The 7-SKU ROI Profile

OPERATOR'S READ

Private Enterprise Endpoints offer the highest risk-adjusted return. They command a 30-56% sovereignty premium that hyperscalers cannot easily compete away in India.

The 7-SKU ROI Heatmap

SKU	Gross Margin Potential	Capital Intensity	Payback Period	Risk Profile
Reserved Capacity	30-45% GM	Very High Cap	3.5-4.5 yrs	Zero util risk
Token-as-a-Service	50-70% GM	High Cap	2.5-3.5 yrs	High commoditization risk
Private Endpoints	55-70% GM	High Cap	2.5-3 yrs	Best risk-adjusted ROI
Fine-Tuning	60-70% GM	Medium Cap	2-3 yrs	Lumpy demand
Edge Inference	40-60% GM	Medium Cap	3-4 yrs	Premium real, unproven volume
Agent Runtime	70%+ GM	Low Cap	<2 yrs	Decouples from token commoditization
Marketplace	20-30% GM	Minimal Cap	Fastest ROI	Cold-start problem

Enterprise Packaging: The Margin Engine

OPERATOR'S READ

Success requires deal desk discipline. Every non-standard deal carries a margin floor tied directly to the cost per token of the underlying hardware.

Tier 1: Launch

- ✓ Token allotment + shared endpoints
- ✓ 99.5% SLA
- ✓ Target: Mid-market's first workload

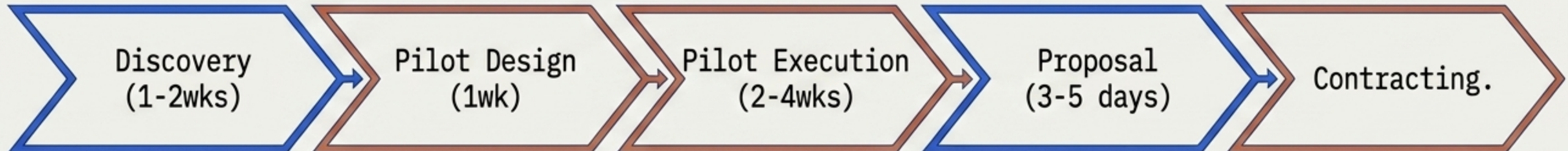
Tier 2: Scale

- ✓ Large token pool + fine-tuning credits + provisioned throughput
- ✓ 99.9% SLA + CSM
- ✓ Target: Growth-stage

Tier 3: Sovereign

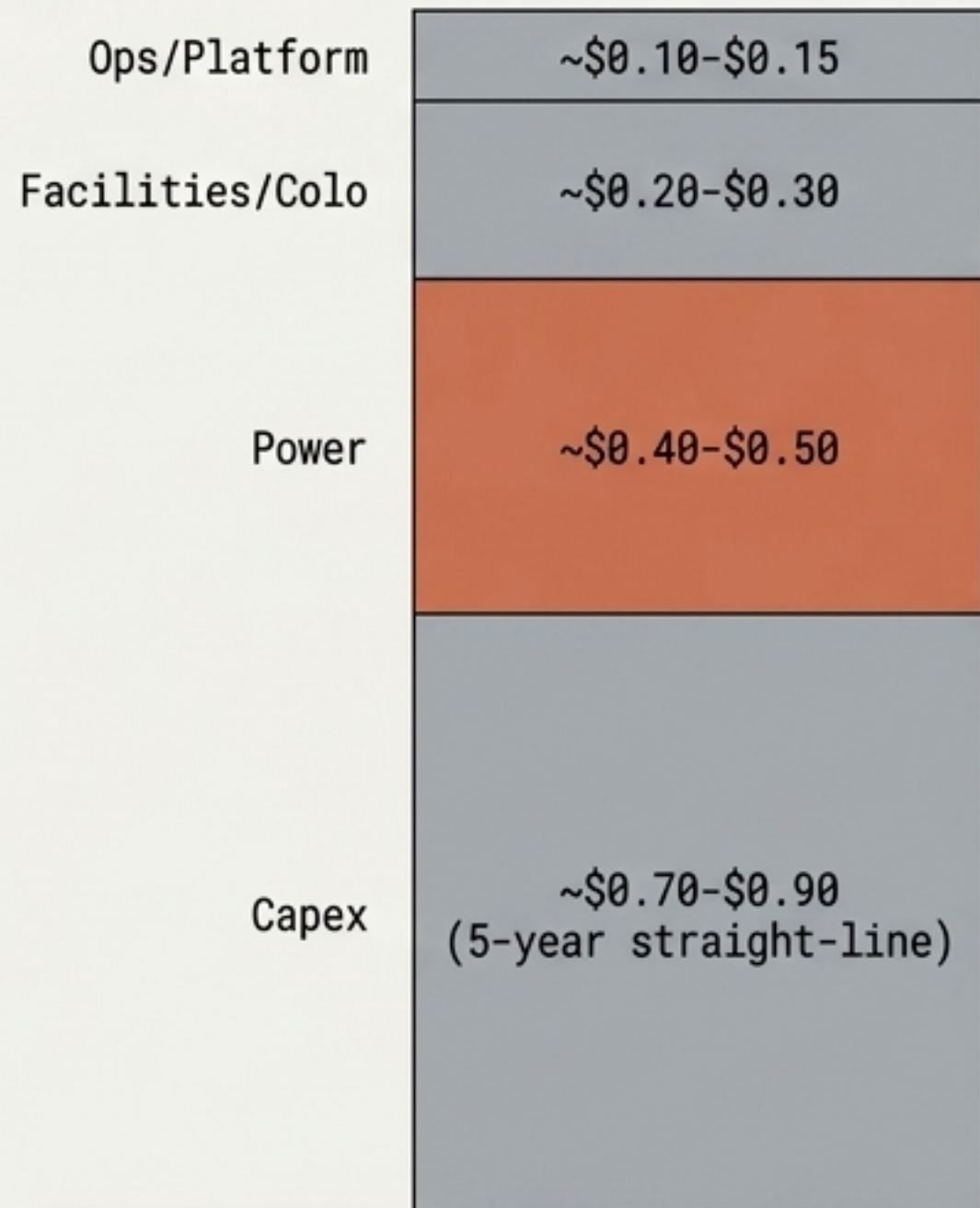
- ✓ Private single-tenant + data-residency guarantee + dedicated capacity
- ✓ 99.95% SLA + TAM
- ✓ Target: BFSI/Govt

The Land-and-Expand Cycle (30-45 days):

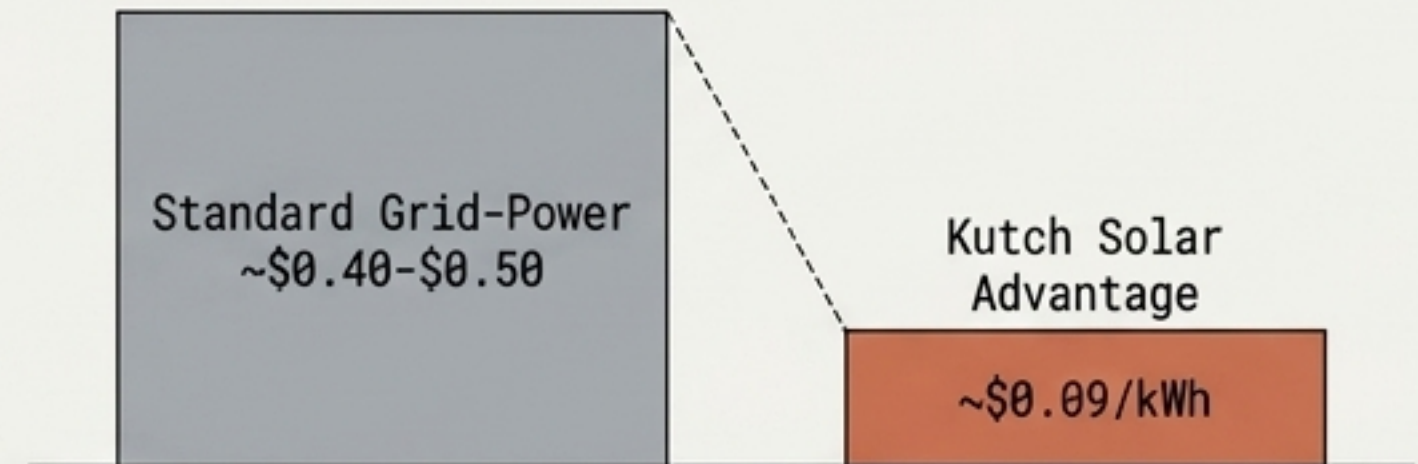


The Cost Floor: Hardware Depreciation and the Power Moat

The Unit Economic Cost Stack (\$1.40-\$1.80/GPU-hour)



The Kutch Advantage: Structural Cost Savings



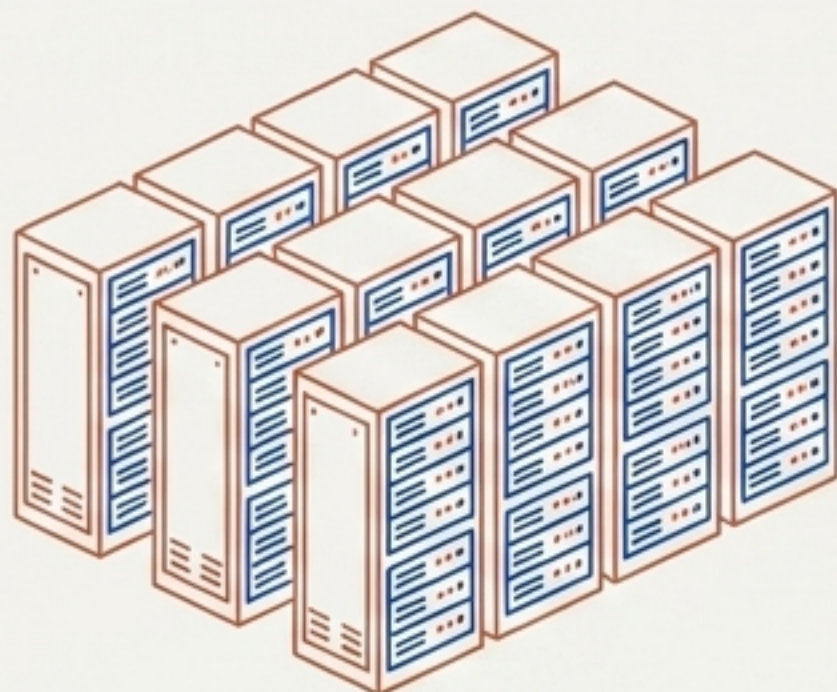
- **Capex:** ~\$40-48k/GPU (5-year straight-line depreciation). Half to two-thirds of total cost.
- **Power:** 700W-1,000W per chip. 1.3 PUE.
- **Facilities:** \$135/kW-month.
- **Financing:** Break-even sits in the high-\$1s before any profit.
- **The Kutch Advantage:** Captive solar transfer rates (\$0.09/kWh) structurally lower the COGS floor versus grid-blended hyperscalers.

OPERATOR'S READ

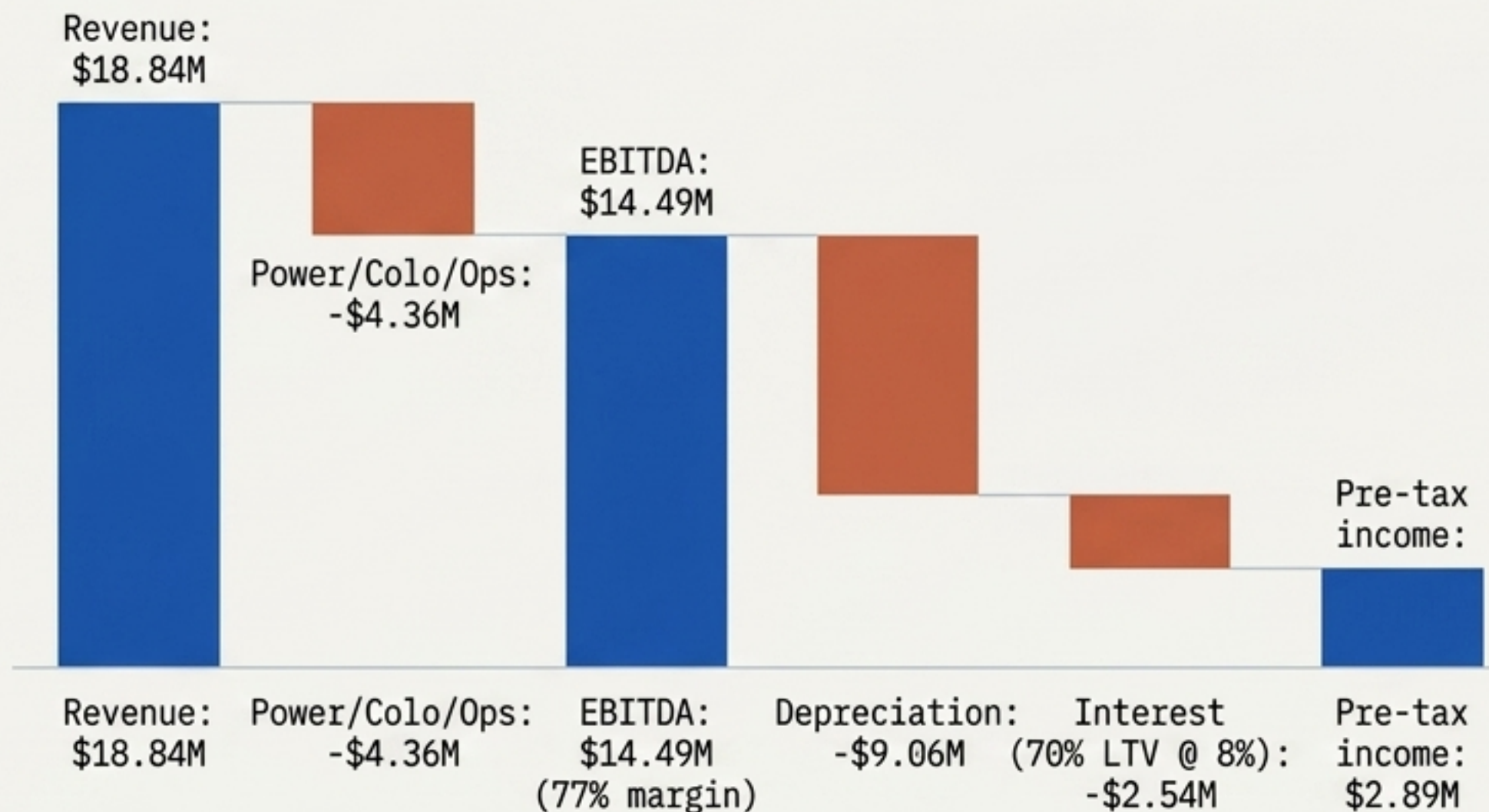
Do not use the power advantage to win price wars on spot compute. Bank the margin and direct the advantage toward sovereignty-premium SKUs.

Physical Fleet Economics: The 1,024-GPU Capacity Cluster

The Asset



- 128 servers x 8 H100s = 1,024 GPUs
- \$45.3M all-in Capex
- Assumes \$2.10 take-or-pay rate



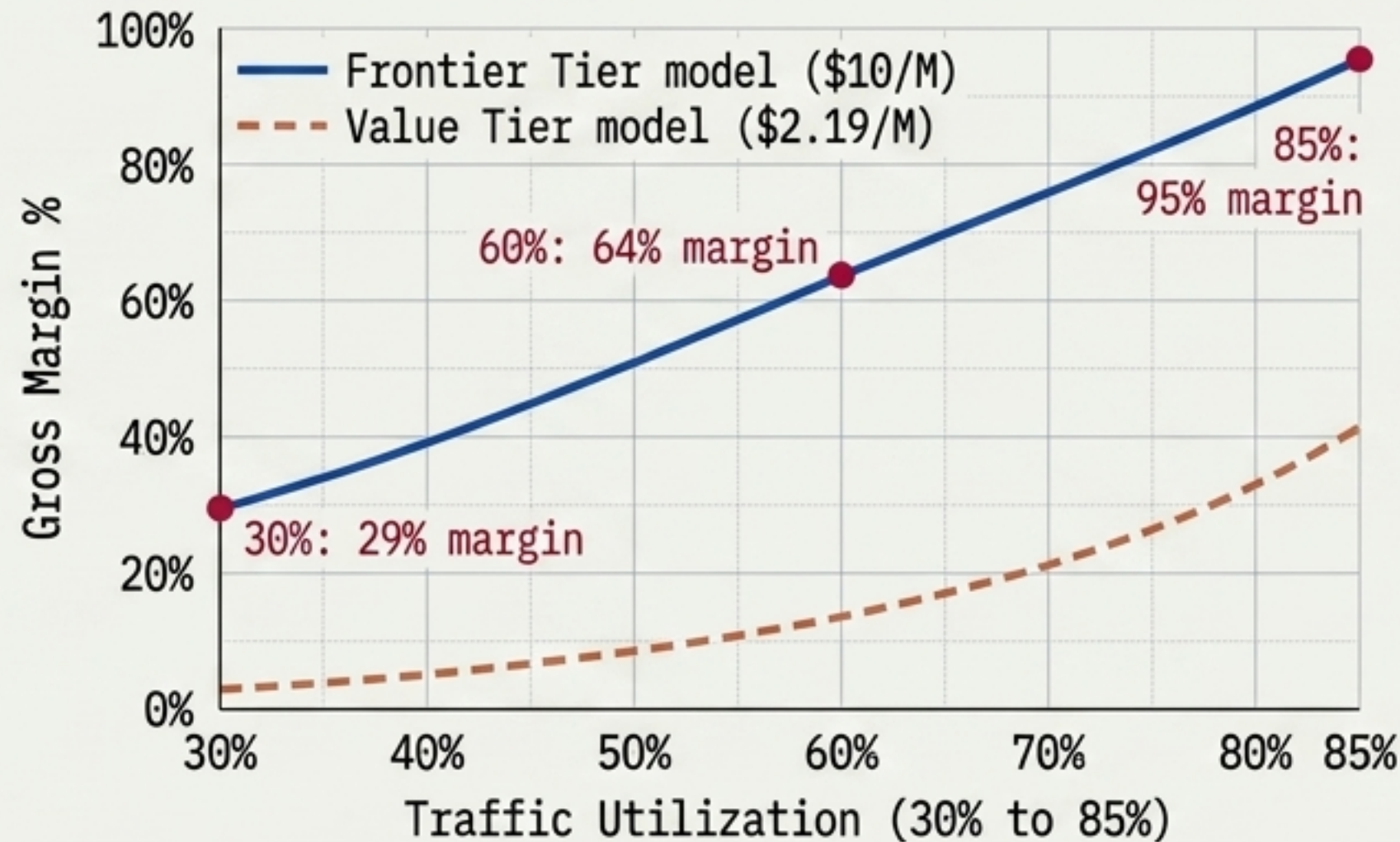
OPERATOR'S READ: The business is highly leveraged. At \$2.10, the cluster pays back in ~3.8 years. If the achieved price drops to \$1.70, pre-tax income falls below breakeven after financing. Backlog, not GPU count, is the real asset.

Token Economics: Utilization is the Margin

The Asset

- One 8xH100 node
- All figures
- Charged internally at \$16.80/hr transfer price
- 70B-class model

The Traffic-Density Margin Multiplier



The Math

- 30% utilization: 29% margin (\$1.56 cost per 1M output)
- 60% utilization: 64% margin (\$0.78 cost). Node generates ~\$47/hr revenue.
- 85% utilization: 95% margin (\$0.55 cost).

OPERATOR'S READ

The token wrapper can dramatically out-earn the capacity wrapper on identical silicon. The reward for operating complex multi-tenant routing and batching is an achieved price nearly 3x the take-or-pay rate.

The Reality of Usage-Based Margin

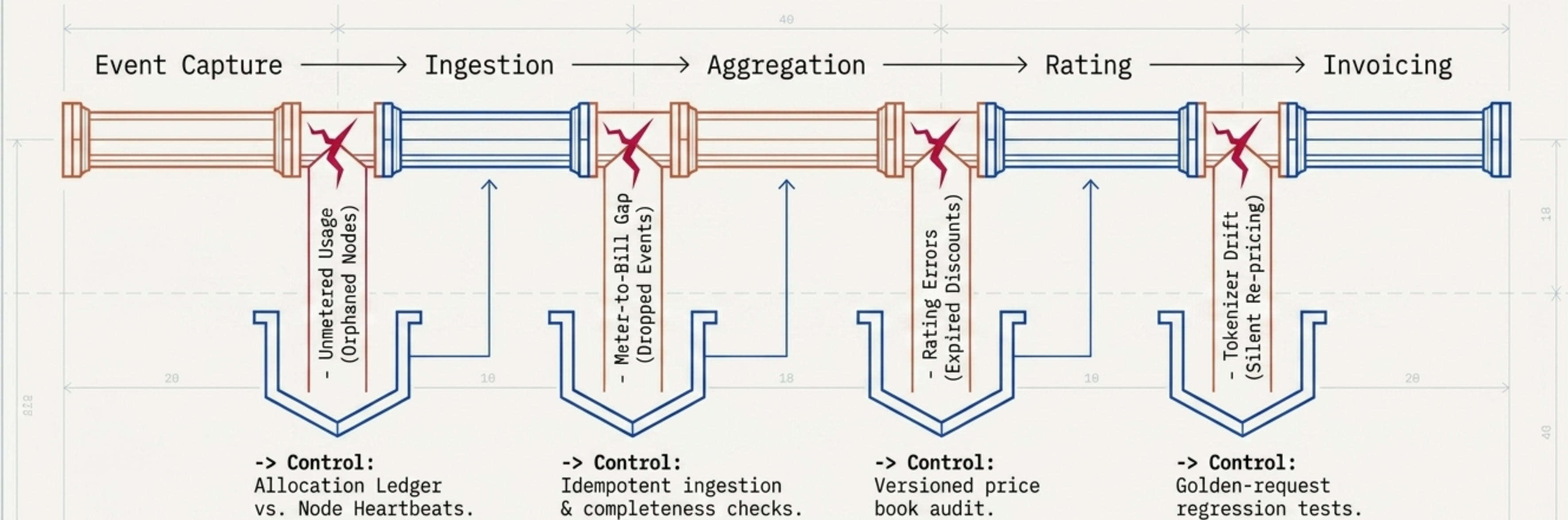
1-3%

Margin is manufactured in pricing, but it is protected in operations.

1-3% of revenue routinely vanishes between the network switch and the invoice in un-managed consumption businesses. At \$100M+ ARR, that is a leadership-level sum.

Billing is not back-office plumbing. It is a first-class product with an SLA. A recovered dollar of leakage carries zero incremental COGS.

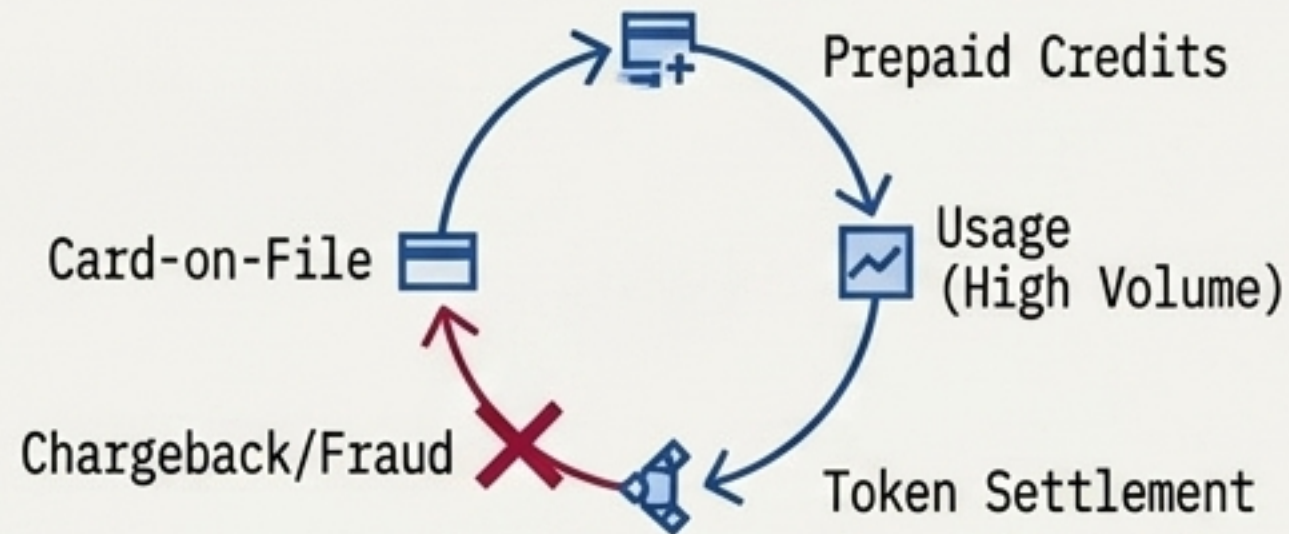
Revenue Assurance: Plugging the Meter-to-Cash Pipeline



THE BACKBONE CONTROL: Automated Daily Three-Way Match (Intent Ledger vs. Telemetry vs. Invoiced Output). Treat discrepancies like trading-desk breaks.

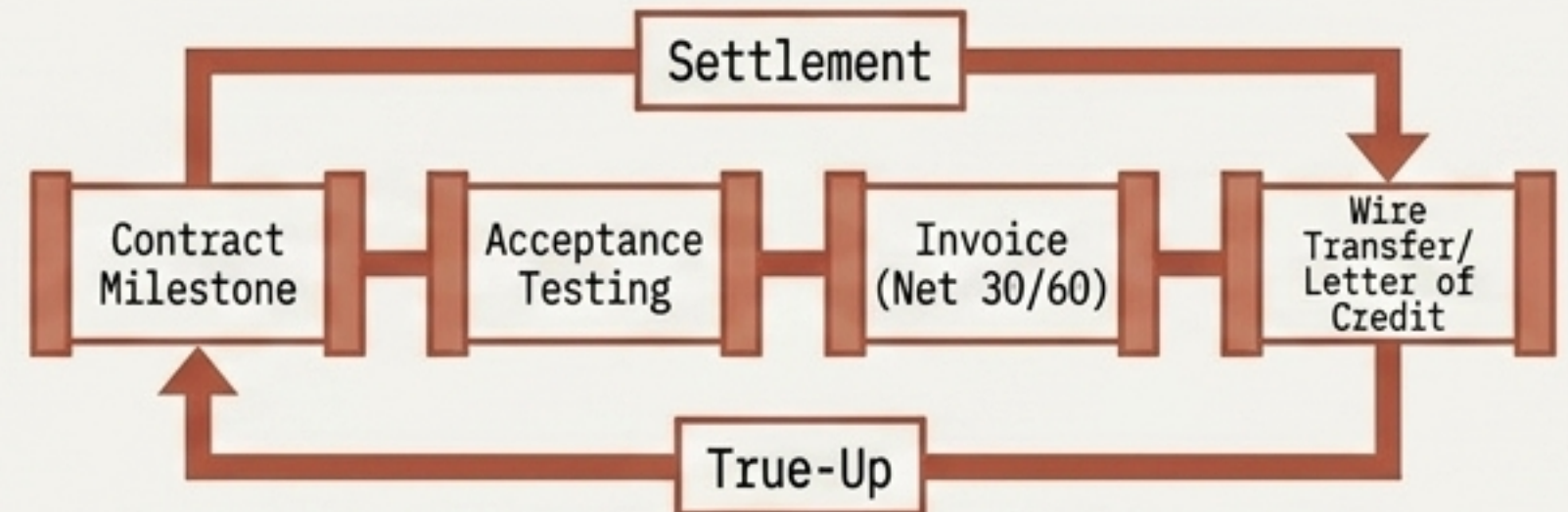
Cash Settlement & ASC 606 Concentration

Token Settlement (Consumer/SaaS Profile)



- High volume, card-on-file, prepaid credits (deferred liability).
- Risk: Stolen-card GPU abuse (crypto mining) where we eat the chargeback and the burned power.

Capacity Settlement (Industrial Leasing Profile)



- Invoices on net-30/60. Milestone billing tied to acceptance testing.
- Risk: Massive counterparty concentration.
- Note: CoreWeave's \$99.4B backlog is highly concentrated.

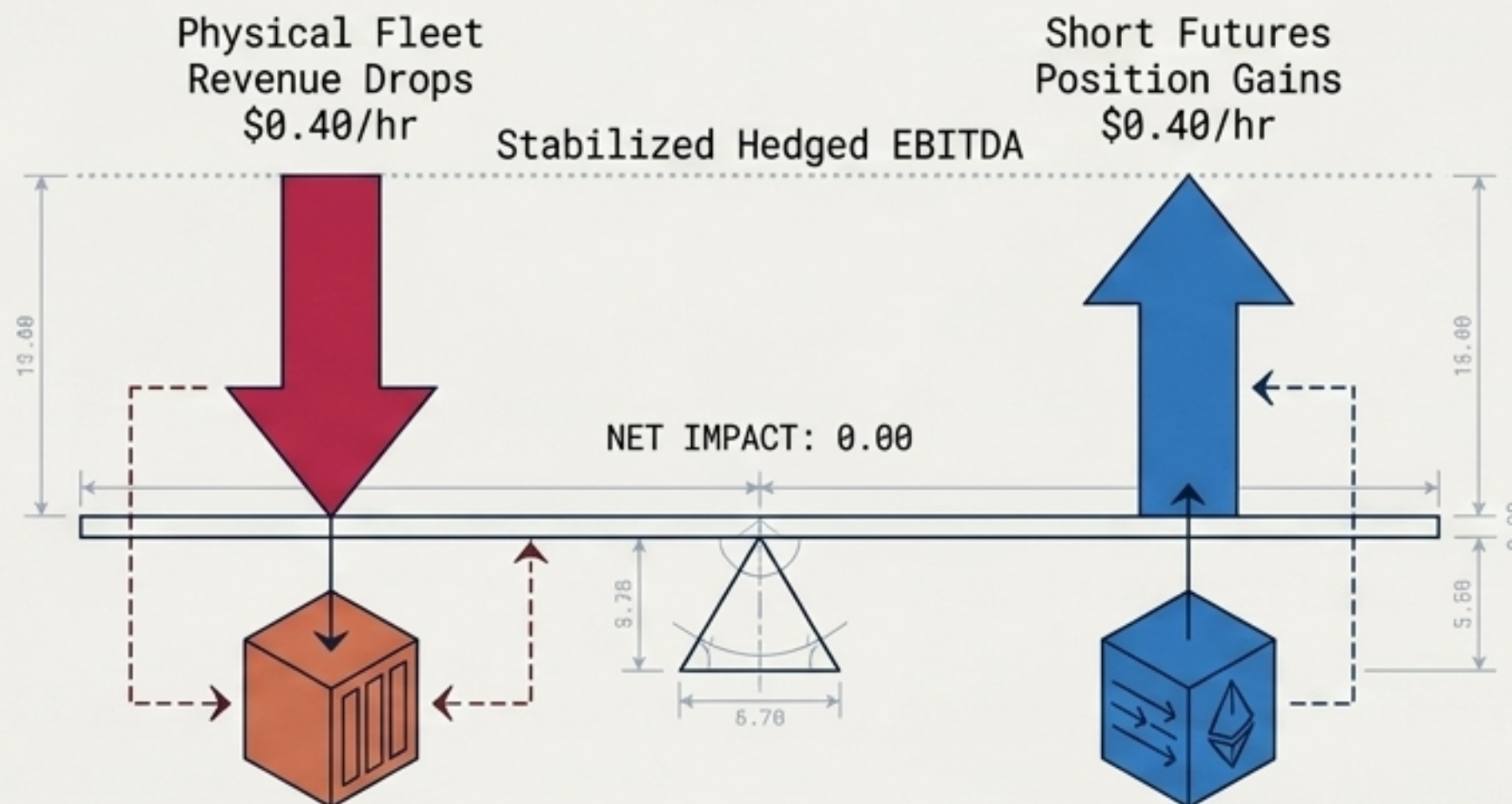
OPERATOR'S READ: Receivables risk in the capacity business is single-name credit risk. It requires credit-committee discipline, not standard accounts-receivable routine.

Risk Management: Hedging the Merchant Tail

The Hedging Seesaw

The Risk:

When next-gen scales, H100 rental prices will sag faster than our 5-year depreciation schedule.



The Math (1,024 GPUs @ 85% Util):

Spot falls from \$2.10 to \$1.70.

Unhedged fleet loses \$3.05M.

A 35% short position gains \$1.07M.

Net result: ~35% of EBITDA erosion is neutralized.

The Hedging Seesaw

OPERATOR'S READ:

Never hedge the contracted core. Take-or-pay is already fixed; hedging it creates speculative risk. We only hedge the volatile spot/on-demand tail to de-risk debt financing.

Scenario Planning: Stress-Testing the P&L

Base Case

- 1.0x Price, 1.0x Coverage
- EBITDA: \$87.3M

Price Crash (Hedged)

- 0.85x Price
- EBITDA drops to \$69.0M (-\$18.3M)
- Insight: Survivable because the futures desk offset the spot crash.



[HEDGE ENGAGED]

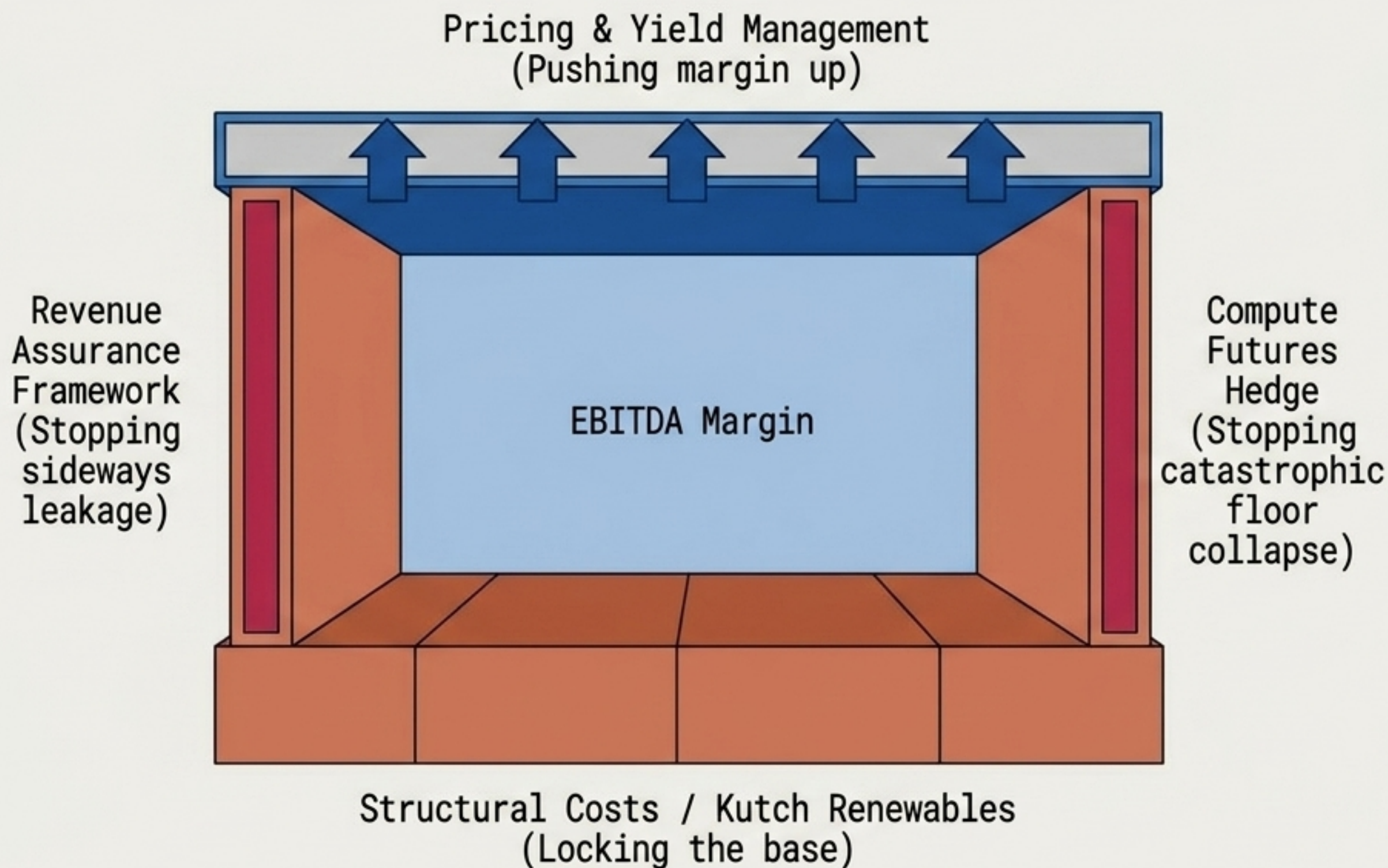
Demand Shock (Upside)

- 1.05x Price, 1.15x Coverage, 1.10x Fleet
- EBITDA: \$110.4M (+\$23M)
- Insight: Upside is highly asymmetric.

Cost Inflation

- 1.15x Opex (Power/Colo)
- EBITDA: \$80.1M (-\$7.2M)
- Insight: Must be offset by token load-shaping and utilization efficiency.

Synthesis: The Architecture of Margin Defense



Core Insight:

Margin is manufactured in pricing, but it is protected in operations.

A 56% EBITDA margin business rapidly becomes negative if depreciation and interest consume the core.

Without the 'Walls' of strict billing controls and financial hedges, the space between cost and price evaporates in a market downturn or operational drift.

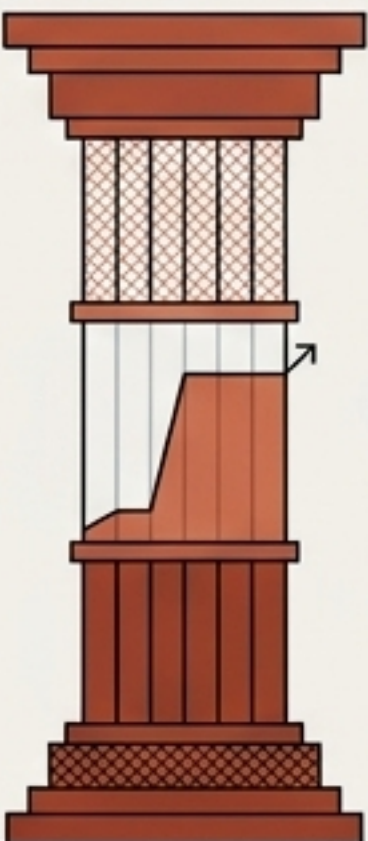
Organization Build: The Five Commercial Pillars

Target:

Scale to ~64 headcount by Year 1. Reach 60% ready-now succession depth by month 12 to eliminate single points of failure.

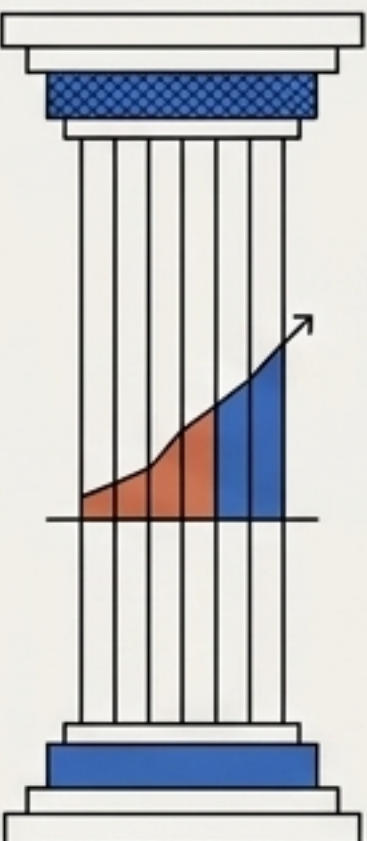
OSS/BSS & Billing Platform

Telecom-grade metering and settlement.
(Front-loaded hiring).



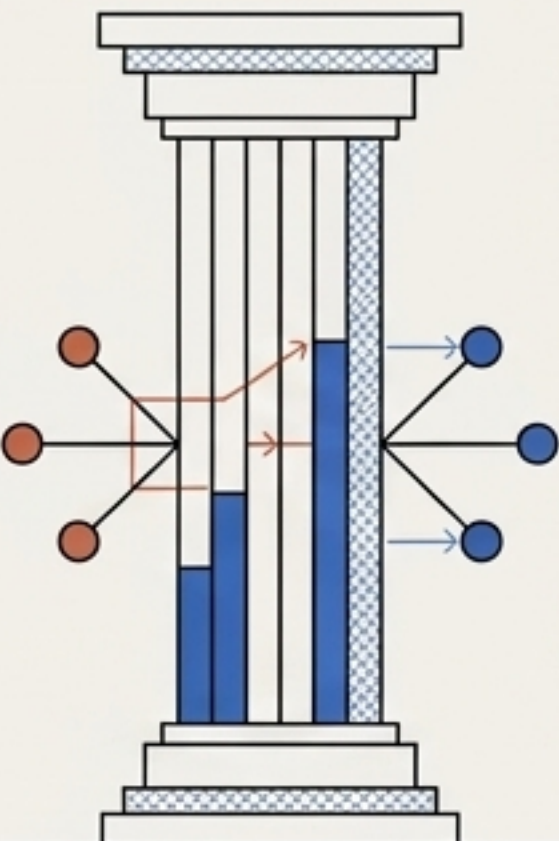
Enterprise Commercial

Deal desk, anchor tenant sales, pricing.



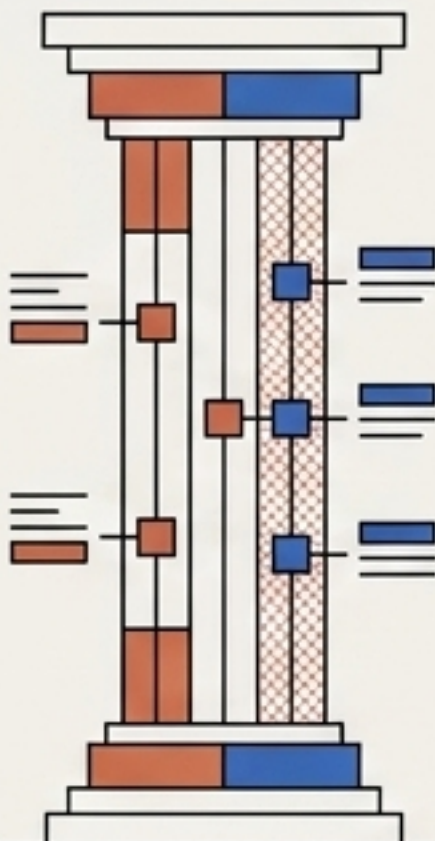
Marketplace & Ecosystem

ISVs, rev-share ops.



FinOps & Revenue Assurance

Leakage controls, utilization analytics.



Customer Success & Capacity Mgmt

Capacity ledger, net revenue retention (NRR).



The Commercial Cockpit: Live Mandate Scorecard

reliancecloud.live | System Status: **OPERATIONAL** ✓

On Track



EBITDA Margin: **69.4%**



SKU Deployment: **7 of 7**



Watch List



Meter-to-Cash: **97%**



Cost per 1M Tokens: **\$0.63**



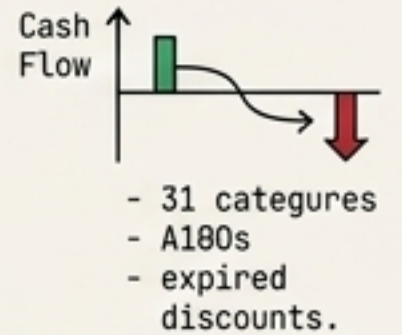
Off Track (The Problems)



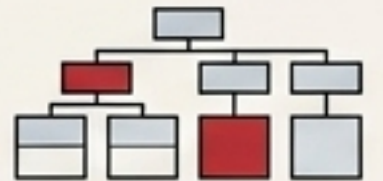
- **Committed Revenue Share: 68%** (Target $\geq 80\%$).
4 contracts approaching renewal.



- **Revenue Leakage: 0.88% / \$781K**
(Target $\leq 0.50\%$).
31 open disputes (orphaned A100s, expired discounts).



- **Succession Depth: 33%**
(Target $\geq 60\%$).

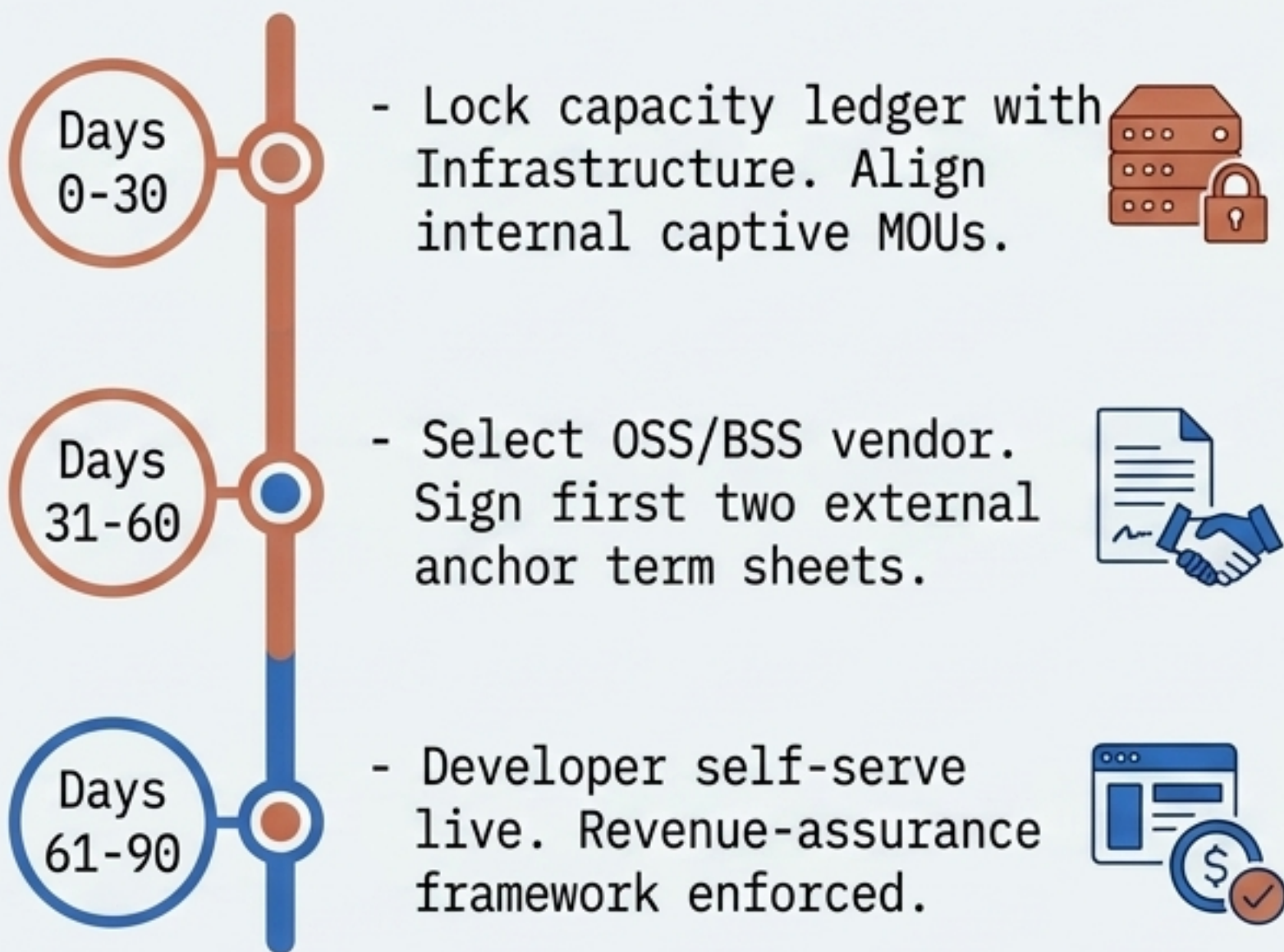


OPERATOR'S READ:

The business is earning very well today, but the **operational plumbing** is showing stress. Fixing **leakage** and securing **renewals** are the immediate fires.

Execution: The First 90 Days & Immediate Priorities

The Roadmap



Decisions Required for Tomorrow's Review

- ☐ 1. Confirm the physical commissioning curve with Infrastructure (gating mechanism for sales).
- ☐ 2. Approve the Top-8 external anchor pursuit list for immediate CEO-to-CEO outreach.
- ☐ 3. Approve the OSS/BSS build vs. buy timeline to ensure billing doesn't bottleneck launch.

