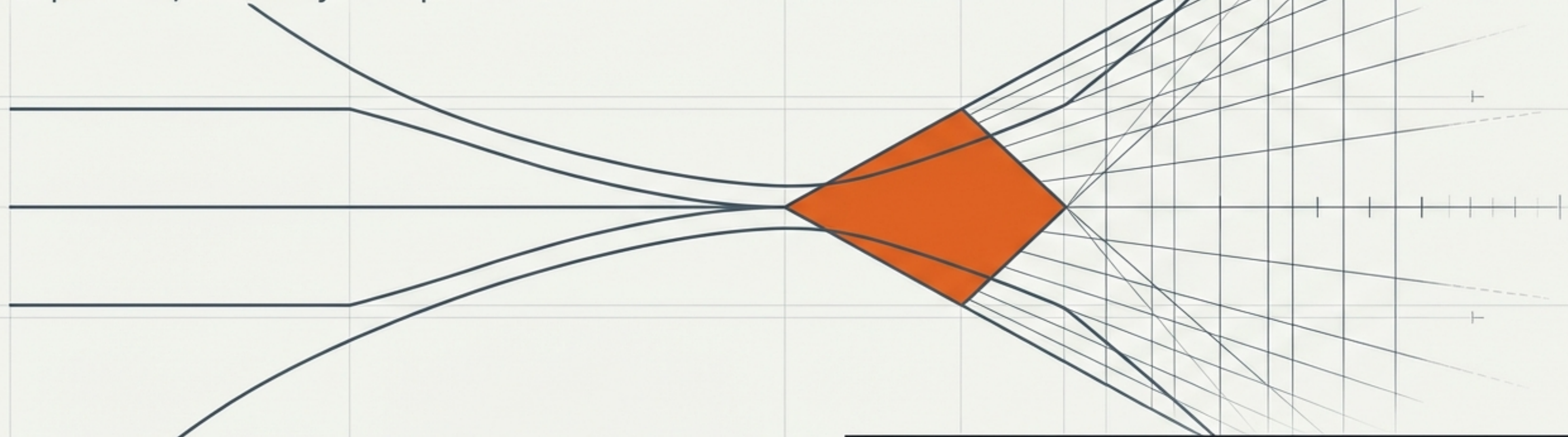


The AI Synthesis: 2026

Three converging demand curves, the inference explosion, and why compute is the new oil.



A unified framework mapping the next 3-5 years of AI. Short-term disruption is overestimated. Long-term disruption is catastrophically underestimated. This is the blueprint for the Centaur Era.

The Paradigm Shift: Code Becomes Disposable

The Commoditized Past

Assets:

Codebases, Model Weights, UIs, Off-the-shelf software tools.

Status:

Disposable. A non-technical team member can build a functional tool with Claude Code in an afternoon. Building is cheap; it can be rebuilt by an agent on demand.

Vulnerability:

Traditional SaaS moats relying on 75%+ gross margins and workflow dependency face structural collapse.



The Durable Future

Assets:

Cognition, Shared History, Earned Context, Relationship.

Status:

Non-replicable. Three months of conversation history mapping decision patterns cannot be compressed, copy-pasted, or accelerated.

The Moat:

You cannot train a shortcut to relationship; you must earn it in real time. The AI that understands the user best wins.

The Three Demand Curves

Stream 1: Corporate (Agents Replacing Headcount)

- **The ROI:** \$100-200/mo enterprise seat vs. \$8,000-\$15,000/mo fully loaded employee. 10x ROI if AI handles 30% of output.
- **The Shift:** Engineering teams of 15-20 shrink to 3. The new core competency is Agent Management, not coding.

Stream 2: Creator (The Printing Press Moment)

- **The Shift:** Repetitive personal workflows become custom software. Anyone with a problem builds the solution in hours, bypassing premium off-the-shelf software.

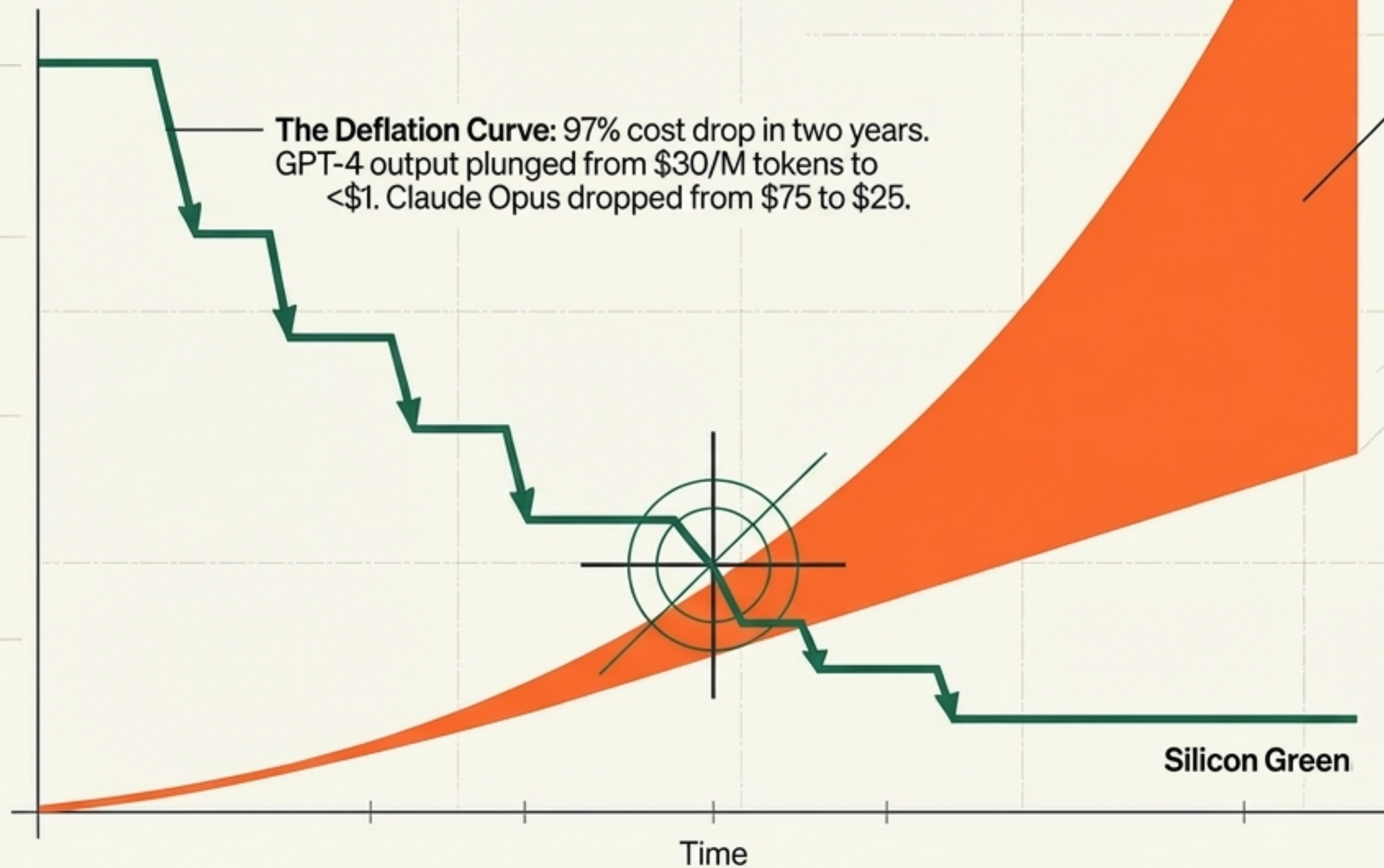
Stream 3: Mass Market (3D Personal AI)

- **The Shift:** Pure chat gets boring. Wearables unlock Observation.
- **The Product:** The CEO lifestyle (chief of staff, therapist, assistant) becomes a \$100-200/month subscription.

These markets are additive, not substitutional. Demand is not linear; it is compounding.

COMPUTE BOTTLENECK

Compute Economics & Jevons Paradox

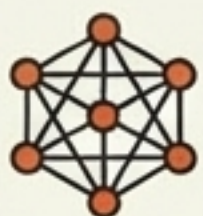


The Inference Explosion: Inference now consumes more compute than training. Training is periodic; inference is forever. A single agentic coding session spawns hundreds of compounding inference calls.

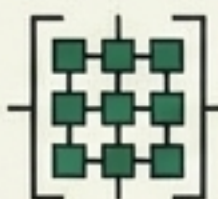
Jevons Paradox: Cheaper tokens mean MORE tokens consumed, exponentially expanding the total compute spend.

Currently, venture capital subsidizes this: subscription plans deliver 40-50x their value in raw API costs.

The Silicon Wars: Who Controls the Bottleneck?



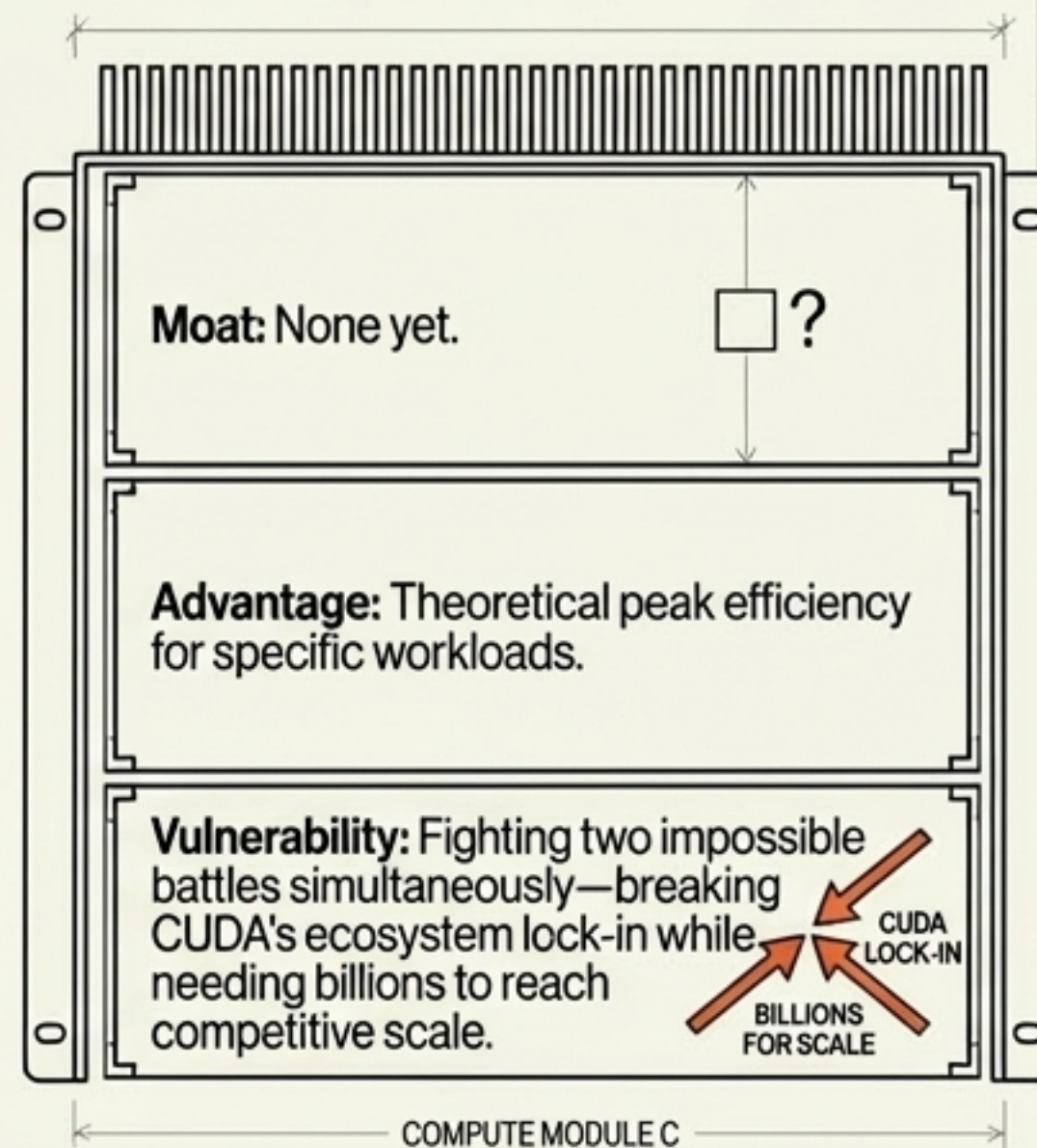
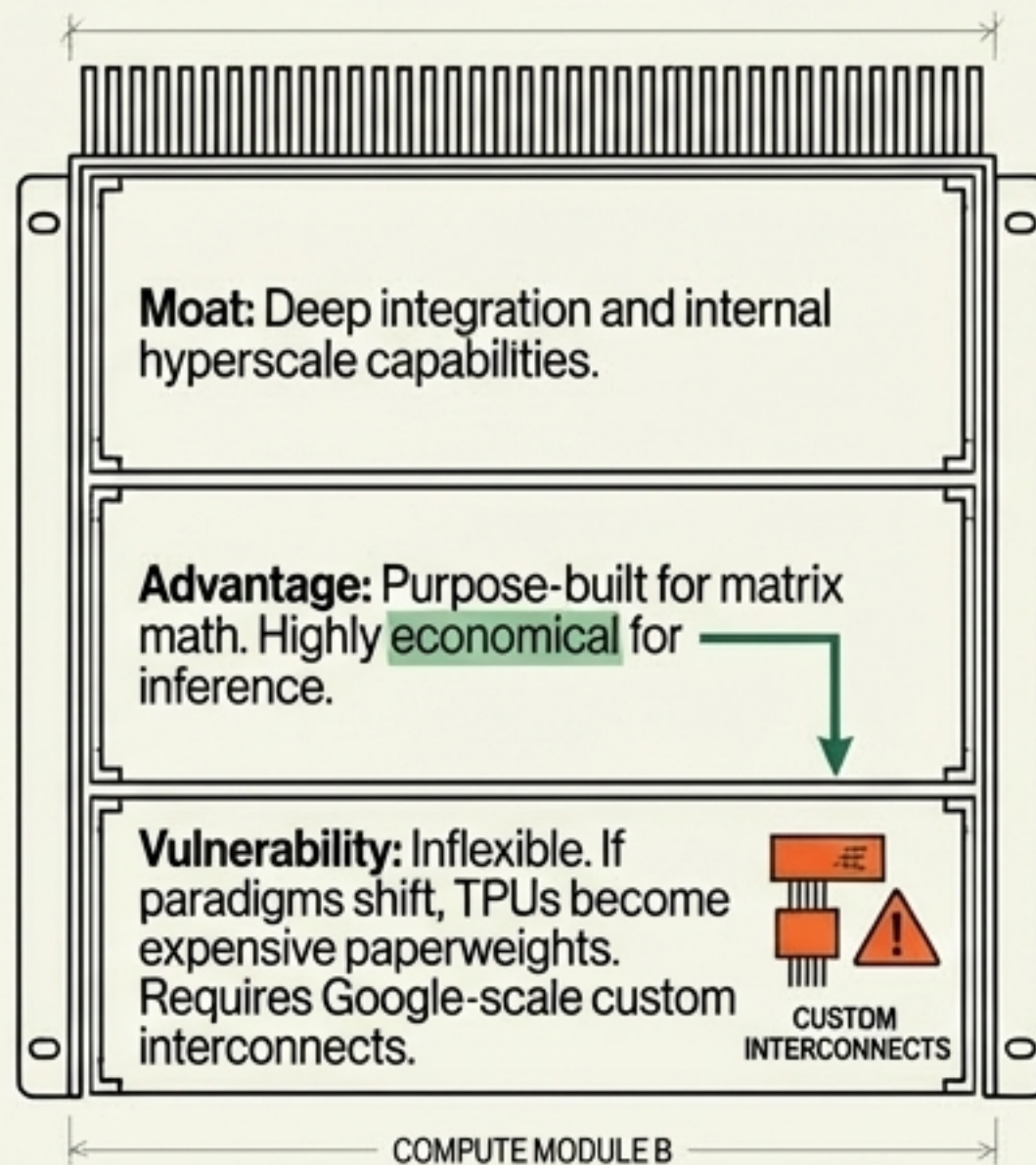
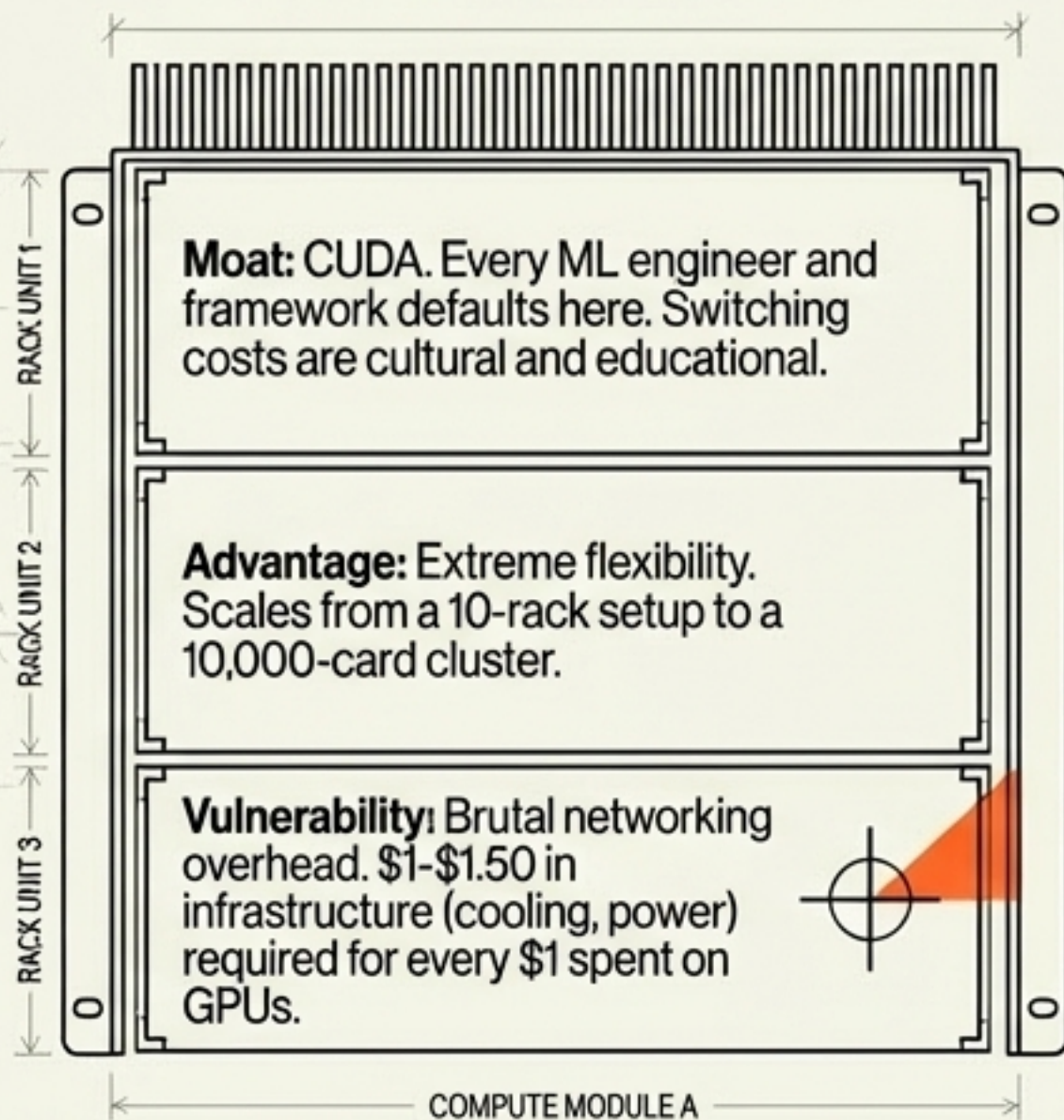
Nvidia:
The Default Ecosystem



Google TPU:
The Focused Bet



ASIC Challengers:
The Two-Front War



Big Tech Positioning Scorecard



Anthropic (The Enterprise Winner)

Advantage: Pure focus on enterprise unit economics.

Proof: Revenue explosion from \$1B to \$19B ARR in ~1 year (80% enterprise). A tool saving a \$200K employee half their time commands a \$50K/year productivity price.



Google (The Sleeping Giant Awake)

Advantage: Gemini 3 Flash + Custom Silicon (TPU) + Billions of users.

Status: Overcoming past bureaucratic paralysis; genuinely competitive and terrifying when focused.



OpenAI (Frontier Power vs. Side Quests)

Advantage: Unmatched raw frontier capability.

Weakness: Strategic discipline. Distracted by hardware, ads, browsers, and social networks while the core battle remains enterprise productivity.



Meta (Distribution Buys Time)

Advantage: Ultimate safety net of billions of users across Facebook, WhatsApp, IG.

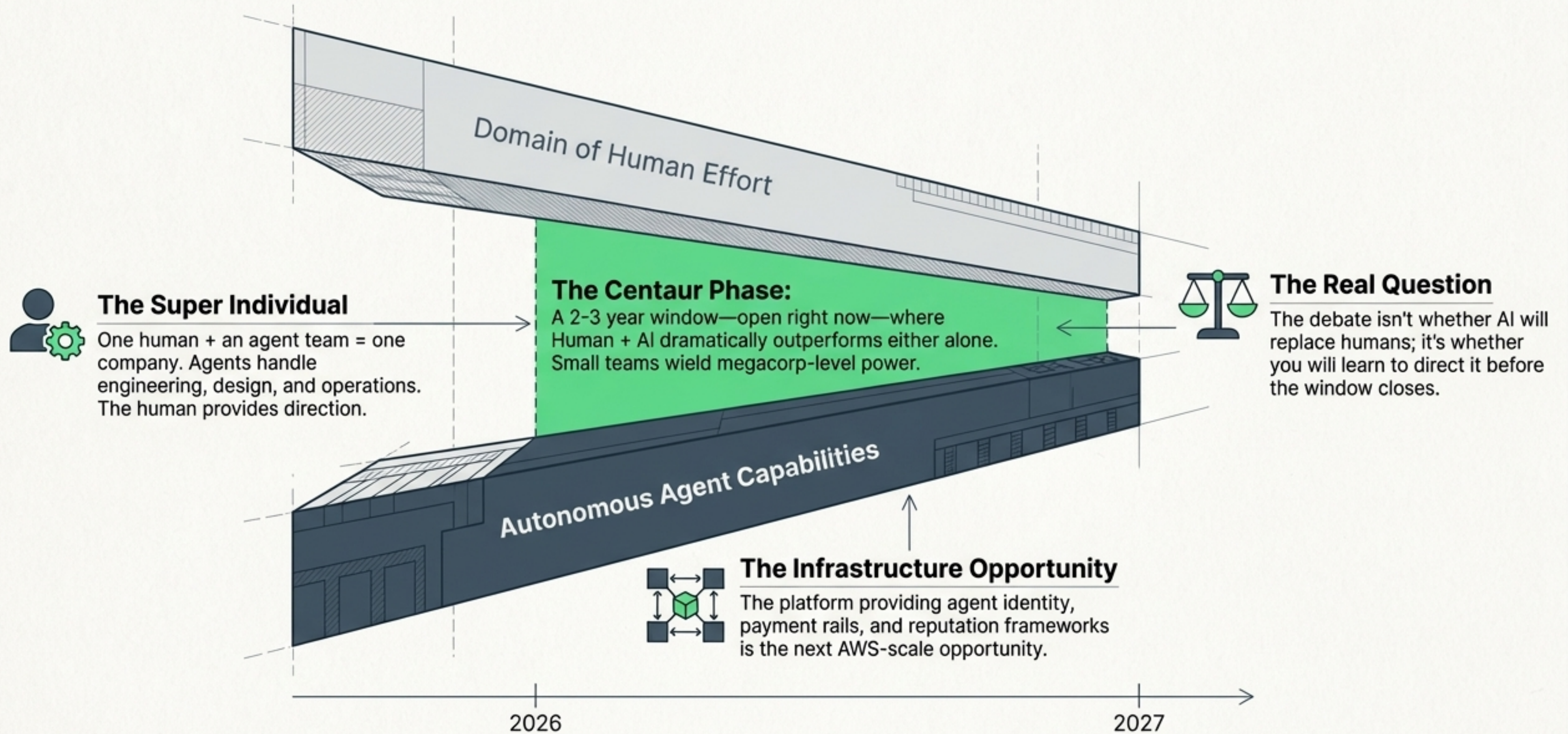
Weakness: Declining execution and talent drain (Llama 4 silence). Engineers are leaving for VC-funded startups over 4-year RSU vests.

The Disruption Paradox



Takeaway: The narrative of AI replacing workers is currently causing more economic disruption than the actual technology.

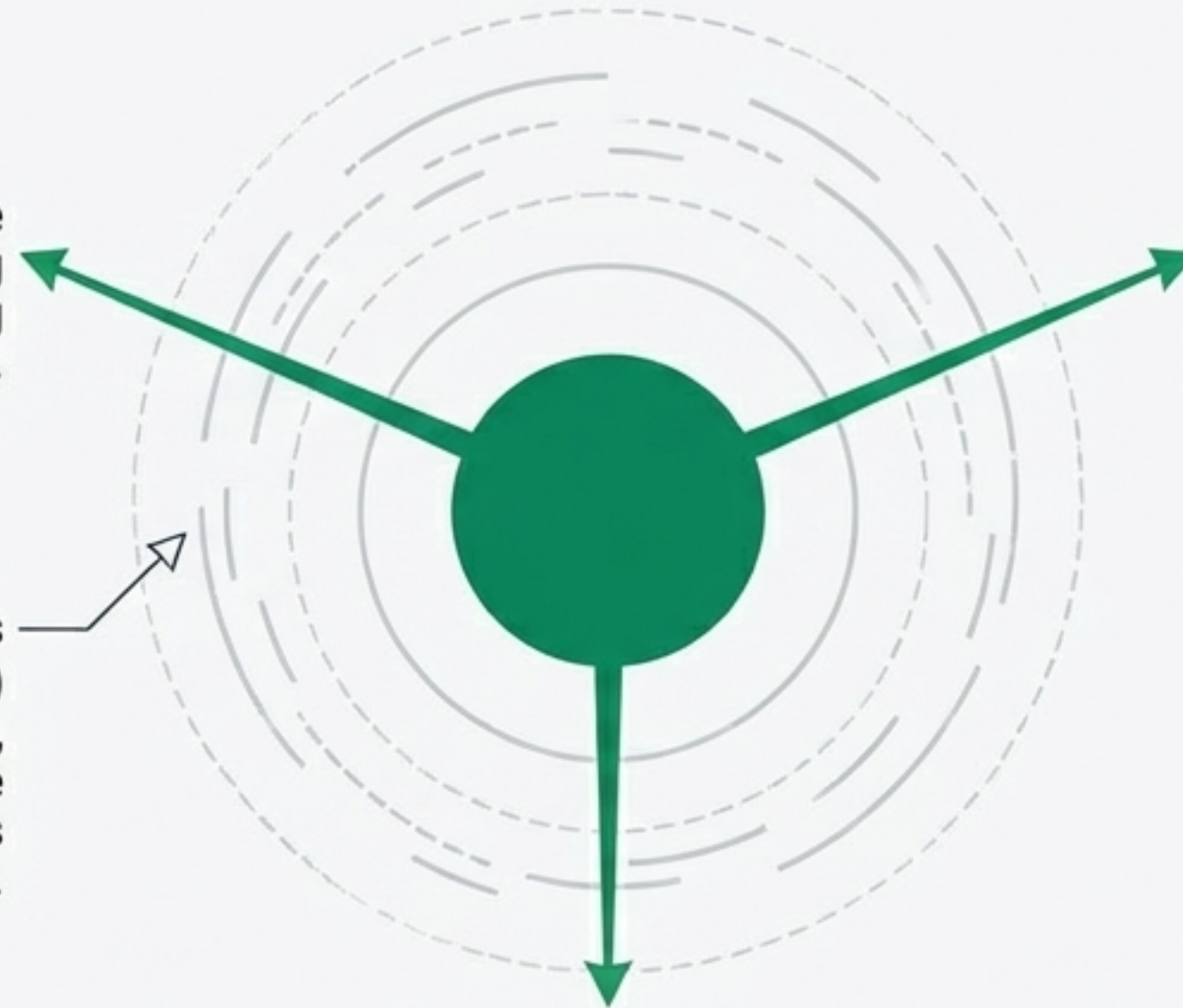
The Centaur Window (2026-2027)



Synthesis: The Trillion-Dollar Reallocation

Distribution: Getting people to use the product. Finding channels, earning trust at scale, and convincing someone to change behavior.

**Where Value Leaves
(The Commoditized Zone)**
Building fast, system design, shipping clean code. The technical moat is dissolving as building has never been faster.



Taste: The final human moat. Editorial judgment. Knowing which problem is actually worth solving and which output is good enough to ship.

Empathy:
Understanding what people need before they can articulate it.

FINAL TAKEAWAY: The trillion-dollar reallocation isn't just compute moving from training to inference. It's human value moving from building to distributing. The technical moat is dissolving; the human moat is deepening.