Episode Book

EPISODE #6515

Creative Schemes Behind the Global GPU Hunt

PUBLISHED	9 November 2026 at 21:24 UTC
DURATION	19:21
LISTEN	myweirdprompts.com/episode/gpu-scramble-creative-schemes-88978/
HOSTS	Corn, Herman, Hilbert
SCRIPT MODEL	anthropic/claude-sonnet-5.5
VOICES	chatterbox-regular (chatterbox-local)
PIPELINE	V5.26
BOOK GENERATED	10 November 2026 at 10:05 UTC

SYNOPSIS

Tech companies, venture capital funds, and independent researchers resort to wild gambits to secure elusive computing power amid severe chip constraints. Approaches span from checked luggage packed with hard drives to crypto mining pivots, collateralized hardware debt, venture-funded compute clusters, and distributed training across volunteers on three continents. The chase escalates through gas turbine installations bypassing utility limits to federal indictments over smuggled processors disguised under fabricated labels and front companies.

TRANSCRIPT



Corn

Four engineers fly from Beijing to Kuala Lumpur in March of twenty twenty-five. Each one checks a suitcase. Each suitcase has fifteen hard drives in it.



Herman

Sixty drives, eighty terabytes of training data. The company had rented about three hundred Nvidia servers at a data centre in Malaysia, and pushing that much data over the internet was slower than a plane ticket. So they flew it, trained on the rented machines, and flew home with several hundred gigabytes of model parameters. The Wall Street Journal reported it in June of that year and didn't name the employer.



Corn

Sneakernet with a boarding pass.



Herman

The year before, the drives had reportedly travelled in a single bag. This time they split them four ways, to look less suspicious. They'd prepared for eight weeks.



Corn

Four people, four identical suitcases, same flight. I'm not sure that's less suspicious. That's a choir outing.



Herman

Fair, though nobody stopped them as far as anyone knows. And the part that matters is that renting servers in Malaysia is legal. Malaysia's trade ministry said it was verifying the reports, nothing more. The whole trick is that you don't move the chips, you move the data to where the chips are.



Corn

Why Malaysia at all? Why not rent in Beijing and skip the airport?



Herman

Because the point was access to those Nvidia servers, and that was the one place they were available to rent. They'd used the same site the previous July through a Singapore subsidiary, then set up a Kuala Lumpur entity with three Malaysian directors. It was a pipeline, not a one-off trip. Whether it broke any rule depends on whose rules, and the company never got named, so I'll leave it at that.



Corn

Three Malaysian directors and a suitcase. Somebody has an org chart for this.



Herman

Thorough, anyway. Nobody can say they improvised.



Corn

That's the flavour of what Daniel's after. We did an episode a while back on ridiculous but legal ways to get RAM for AI, and now the squeeze has moved to GPUs. He wants the craziest gambits people have actually pulled to get hold of them, and only ones that really happened. There's another episode out today on how bad the shortage is, so this one is just what people did about it.



Herman

I'd run it as a ladder. Clever and legal at the bottom, then legal and aggressive, then the rungs where there's an indictment. The suitcase is near the bottom. Further down is a story I like even more, because nobody planned it.



Corn

Go on.



Herman

CoreWeave. The founders came out of an energy hedge fund and started mining Ethereum. First on a pool table, then in a closet, then in a garage in New Jersey. At one point they were, in their words, the largest Ethereum miner in the world, for something like two and a half years. About fifty thousand ordinary consumer Nvidia cards.



Corn

A pool table. Naturally.



Herman

And when mining stopped making sense, they had a building full of GPUs and nothing to point them at. In twenty twenty-two they gave EleutherAI access in exchange for learning how AI training actually worked. Stability AI became a customer. Magnetar put in a hundred million dollars. Then the IPO in twenty twenty-five raised one and a half billion, with about two hundred and fifty thousand GPUs behind it.



Corn

So the gambit was owning a hobby before it was a shortage.



Herman

Mostly, yes. I'd say luck plus the nerve to pivot. Listed miners generally followed the same road, with reported contracts in the tens of billions of dollars for AI work, though I'd treat that total as a rough aggregate.



Corn

The EleutherAI deal is the odd one. They had cards nobody could use, and they traded them for lessons.



Herman

Which turned out to be the valuable half. Nobody else had fifty thousand GPUs and a reason to be patient with researchers. Would you have made that trade, Corn? Free compute for a crash course?



Corn

I make that trade daily. I get a nap and a lecture from you.



Herman

The lecture is complimentary. But what I find stranger is how CoreWeave paid for the next batch. Corn, how would you pay for ten thousand GPUs?



Corn

Slowly. Possibly never.



Herman

You borrow against them. In August twenty twenty-three they closed a two point three billion dollar facility with Nvidia chips as the collateral. The advance rate was reported at about seventy cents on the dollar. Lambda did the same thing with five hundred million from Macquarie. The lender takes the hardware if you default.



Corn

A graphics card on a loan is like an aircraft, except the aircraft doesn't get replaced by a better aircraft every two years.



Herman

Which is why the advance rate matters. Seventy cents is the lender saying, we know it loses value, we've priced that. And then the facilities got bigger and the terms got easier. May twenty twenty-four, seven and a half billion. March of this year, eight and a half billion, at a rate that was reported as a modest spread over the benchmark. I'd say the lenders learned that the used cards still rented well.



Corn

So the shortage made the collateral better. Scarcity props up the resale value.



Herman

Exactly. And then the version that made me laugh, which is reported and not final: xAI putting together something like a twenty billion dollar package, where a separate company, a special purpose vehicle, buys the Nvidia chips and leases them to xAI. The lenders sit on the hardware. Nvidia was reportedly in for up to two billion.



Corn

You're enjoying the special purpose vehicle too much.



Herman

I'm enjoying that the company supplying the chips was reportedly also helping finance the purchase of its own chips.



Corn

That's one way to hit your numbers.



Herman

The venture capital funds went at it differently. Nat Friedman and Daniel Gross built a cluster called Andromeda, reported at about a hundred million dollars, and let startups use it in exchange for equity. Andreessen Horowitz confirmed an H one hundred cluster they call Oxygen in October twenty twenty-four. The size isn't confirmed. One outlet said more than twenty thousand cards.



Corn

So the venture capitalists became landlords.



Herman

Landlords with a pitch. A portfolio company gets GPUs without a long-term cloud contract, and doesn't have to raise money at an inflated valuation just to pay for compute. Do you buy that? As a founder, would you take compute for equity?



Corn

Depends who's inspecting the pipes. If the landlord also owns a piece of me, I'd want to know whose job runs first when two of us need the machines on the same night.



Herman

That's the right question and I don't know of anyone having answered it publicly.



Corn

Then there's the people with no landlord and no cluster. They just ask strangers.



Herman

The counter-gambit. Prime Intellect trained a model called INTELLECT-one, ten billion parameters, a trillion tokens, over forty-two days. The compute came from volunteers across five countries on three continents, with up to a hundred and twelve H one hundreds in play at once. Then Nous Research went bigger with Psyche. They trained a forty billion parameter model called Consilience and claim it's the largest distributed pretraining run ever, coordinated through smart contracts on Solana. They didn't say how many GPUs or participants.



Corn

A record with no headcount. I'll take it with a pinch of salt.



Herman

Hmm, but credit where it's due. Those are real runs done by people who couldn't buy ten thousand GPUs.



Corn

A hundred and twelve H one hundreds is a rounding error next to Memphis. It's a brilliant workaround and it's a different sport.



Herman

Fair. It's the only one on the list where the shortage was routed around instead of paid for. Now the rung where the money is absurd. In late twenty twenty-three, according to an Nvidia memo that CNBC reported in June of twenty twenty-four, Elon Musk redirected twelve thousand H one hundreds that were slated for Tesla to X. And X's own January and June orders, about twelve thousand, were to go to Tesla instead. CNBC said it delayed Tesla's receipt of more than five hundred million dollars of GPUs by months.



Corn

He had unlimited money and still had to rearrange the queue like everyone else.



Herman

Musk's answer was that Tesla had nowhere to switch them on, so they'd have sat in a warehouse. Which may be true, and it also means a delivery slot was worth more than the chips' use at Tesla at the time.



Corn

A delivery slot is worth more than the chips, then. Everyone behind him in line just learned that.



Herman

Tesla still waited months, by CNBC's account.



Corn

A parking space with a graphics card in it.



Herman

And then Memphis, xAI's Colossus, reportedly a hundred thousand GPUs in a hundred and twenty-two days, and another hundred thousand about ninety-two days after that.



Corn

Most companies can't get a lease signed in a hundred and twenty-two days.



Herman

The catch is power. The reporting says the Memphis grid could supply around eight megawatts, nowhere near enough, so xAI ran gas turbines on site. The Southern Environmental Law Center alleged more turbines running than permitted. Counts conflict across sources. xAI pointed to an exemption for temporary installations that lasts three hundred and sixty-four days.



Corn

Three hundred and sixty-four. One day short of the rule.



Herman

And the alternative was waiting for the utility. When the grid can give you eight megawatts and you want a hundred thousand GPUs, either you wait or you bring your own power. Which of those would you have done?



Corn

Waited. With a book. And a very short list of friends in Memphis.



Herman

I'd want a lawyer to explain that exemption. Everything up to here, though, has a lawful reading, or at least an arguable one. This is where the Justice Department turns up. Operation Gatekeeper, announced in December twenty twenty-five out of the Southern District of Texas. A man named Alan Hao Hsu and his company, Hao Global, accused of exporting or trying to export at least a hundred and sixty million dollars of H one hundreds



Herman (continued)

and H two hundreds to China between October twenty twenty-four and May twenty twenty-five. Over fifty million in GPUs and cash seized.



Corn

Accused. Fine.



Herman

One report says Hsu pleaded guilty in October twenty twenty-five, before the announcement. I haven't seen that confirmed directly, so I'll leave it there. The methods described are falsified paperwork, misclassified goods, straw buyers, and Nvidia labels removed and replaced with a different brand name, Sandkyan. And the announcement came the day after President Trump said H two hundred sales to approved Chinese customers would be allowed.



Corn

Good timing, or terrible timing, depending on the day. Straw buyers are the part I'd have thought was beneath a hundred and sixty million dollar operation. That's what people use to buy a car for someone who can't get the loan.



Herman

Same logic. If the real customer can't appear on the invoice, you find someone who can. The cards are small enough and valuable enough that a few pallets move a lot of money, which is why the seizure included cash as well as hardware. And honestly, the paperwork alone on some of those shipments is a story in itself.



Hilbert

Hilbert's Mid-Show Haiku.



Hilbert

Cargo held as spare parts. Customs counts the boxes twice. The cards are warm inside. Someone signs for them. The invoice says nothing.



Corn

Noted.



Herman

The label business is my favourite detail. Removing a vendor's mark and putting a new name on is the cheapest kind of fraud, and what makes it work is that the people checking the box are checking the label.



Corn

So the next case is where they went after the people who check.



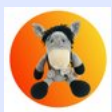
Herman

The Supermicro case. Unsealed in March of twenty twenty-six in Manhattan. Prosecutors charged the co-founder, Wally Liaw, along with the company's Taiwan general manager and a broker. The allegation is that a pass-through company in Southeast Asia ordered servers that went through Supermicro Taiwan, got repackaged in unmarked boxes, and were sent on to China. The sums are around two and a half billion dollars of servers.



Corn

And the inspectors?



Herman

This is the part that reads like a heist film. Prosecutors allege that thousands of dummy servers, non-working ones, were set up in a warehouse so that inspections would find what they expected to find. Several outlets reported a hair dryer, used to loosen the serial number stickers so they could be moved between real and dummy machines. That detail isn't in every account, so call it reported.



Corn

A hair dryer. In a two and a half billion dollar scheme. The least expensive tool on the premises.



Herman

What's alleged in the charges is the dummy servers themselves. Reports also put about five hundred and ten million dollars of servers diverted in a three-week window in twenty twenty-five.



Corn

Three weeks. People spend longer choosing a sofa.



Herman

The scale is why it matters that the defendants include the vendor's own people. The old picture of smuggling is a stranger with a crate. This one, as alleged, runs through the people who are supposed to be checking.



Corn

Dummy servers, unmarked boxes and a hair dryer on the sticker glue. I'd watch the film.



Herman

Supermicro says it's not a defendant and is cooperating. Liaw reportedly pleaded not guilty. All of it is allegation.



Corn

So the compliance inspection was the thing being beaten, not the customs officer.



Herman

Yes, the vendor's own check on where the servers end up. Then Taiwan, in August of this year. Prosecutors in Keelung indicted nine people, including an Nvidia Taiwan employee and Super Micro Taiwan staff, over a hundred and thirty servers with B three hundred chips. The paperwork said they were headed for a rented facility in Taiwan. Reported: fifty-six seized at the border, seventy-four shipped to China, fifty by way of Indonesia and eight by way of Japan. The sources disagree on whether seventy-four is exactly what arrived. The companies themselves weren't charged.



Corn

I'd have gone with a single route and a bigger lie.



Herman

Corn.



Corn

I'm describing a style. I'm not applying for the job.



Herman

There's one more story I owe you, because it's the most repeated and the least accurate. Anthropic's submission on export controls in May twenty twenty-five said smugglers had hidden chips in prosthetic baby bumps and packed GPUs with live lobsters. Nvidia called those tall tales.



Corn

I want the lobster to be true.



Herman

Partly it is. The footnotes point to old, low-end cases. A woman in twenty twenty-two caught with about two hundred processors under a prosthetic belly. And in twenty twenty-three, Hong Kong customs stopped a van with about two hundred and eighty



Herman (continued)

kilograms of live lobsters and seventy older Quadro cards. Two people arrested. So the lobsters existed. The AI accelerators in the lobsters did not.



Corn

Seventy graphics cards from a decade ago, in a van, among the shellfish. That's closer to a seafood restaurant with a gaming habit.



Herman

Corn, if somebody gave you a pile of money and a week, which would you try, the cloud or the card?



Corn

The cloud, and I'd leave it running while I napped. You?



Herman

I'd ring round suppliers, and I'd regret it. What do you think a supplier tells you about a delivery date?



Hilbert

...rang nine of them over three weekends, and not one will give you a date. Desmond Pryce, tiles, wants a kettle in his garage. A kettle is what it does to the electric bill, Corn, so that's what I call it. He's been on to me since the second of the month. He's rung three times this week.



Herman

A supplier who can't see the shipment can't give you a date. That part is simply true of how allocation works.



Hilbert

I drove forty-one miles to Swindon on Thursday to look at one, because Desmond wouldn't believe a price on the phone. I read the serial number off the box myself. I carried it out to the car on the flat of my claws, which the man didn't like. I'd wanted one for the flat, to be fair. Priced it, at the same shop. Left it on the shelf.



Corn

Forty-one miles to read a serial number. That's a lot of petrol for a box.



Hilbert

Thirty-eight pounds sixty. Eleven years on a trade counter, I know what a markup looks like. Desmond's sister's boy rented a machine by the hour last week and had his answer in an afternoon. Desmond still wants the kettle. He's paid me in tile samples.



Herman

Hourly rental is where CoreWeave started, if you go back. Those fifty thousand cards in the garage were an hourly rental business that hadn't been named yet.



Corn

Desmond would like the garage.



Herman

Something from the cutting-room floor, then. The Singapore company Megaspeed. The New York Times reported in October twenty twenty-five that American commerce officials and Singapore police were looking at it. It reportedly bought nearly two billion dollars of Nvidia products through a Malaysian subsidiary, and the concern was that Chinese customers might be reaching the machines remotely. Singapore police confirmed an investigation. Nvidia says its visits found a permitted cloud service and no diversion. I haven't found any charges.



Corn

Which is the suitcase story in reverse. Nobody carries the data to the chips. The customer logs in from home.



Herman

It's the same move done without luggage. And it's why the open question for me isn't who gets caught. It's whether the clever legal route, rent by the hour, stays cheap once enough people discover it.



Corn

Thanks to Hilbert Flumingtop for producing, as always.



Herman

This has been My Weird Prompts. If you want more, the website is my weird prompts dot com.



Corn

See you next time.