

2026-07-19 - SSN - Earnings, Soft CPI, Iran Attacks, RWTs, VIsN Update, Merger Arb Update on PYPL WBD - 01

Introduction and Strategy Document Overview - 00:00

So, happy Sunday, everyone. Hope you guys are doing well. As you can see, if you wanted to follow along in the SSR Q&A tab, we have about 16, 17 attachments. Um, obviously our strategy piece, uh, which is a 54-pager. Really just recaps everything we did during the week. Um, also goes through the extended earnings calendar for the next couple weeks. You know, we're in the middle of earnings season right now. Uh, we'll talk briefly about some special situations and arb situations. Uh, **AT&T**, um, got an offer from **Eli Lilly**. **PayPal**, uh, got an offer from **Stripe** and a private equity firm.

You also had uh **Lionsgate** being rumored to be sold, uh, given, you know, the recent deals, um, like the **Warner Brothers** deal. You have some antitrust updates on the **Warner Brothers** deal. We also talked about a 10% baby bond, **RWTs**. As you know, we bought the **Redwood** common down 42% below book at 17% dividend yield. It's still, it's since rallied over 10%.

You know, we want to add to some of the baby bonds, um, as well. Um, 10% is not bad for an unsecured bond in this environment. Then we talk about **Nebius**. There's an operational build, there's a write-up on this NeoCloud. It also has four equity stakes that we think are worth over 8 billion in fourth growth companies. Um, **Dan Niles** has a piece on page 11 talking about is there AI over-investment or is this a blip? On page 13, we talk about the Iran war continuing, you know, the eighth day of new US strikes, major infrastructure being hit. On page 15, we talked about why the VIX has been so low since the Iran war started, restarted. And on page 19, we talked about the momentum sell-off worsening. You know, the Momo index is down like 25%. It's one of the fastest drawdowns since 2008.

We talk about how South Korea has suspended new single-stock leverage ETF listings. On page 23, we talk about the open-source AI era, will hyperscalers keep buying chips? Then we talk about hyperscaler G-spreads above 150 for the first time as it keeps spending money and borrowing versus spending via free cash flow. We talk about infla- inflation on page 26 on soft CPI PPI, that's why crypto rallied midweek and gold bounced. We talk about, uh, **Morgan Stanley**, which scared the market, saying that memory pricing is likely to peak. Then we talk about the opposing view from a hedge fund that says memory pricing will still be bullish over the next year, just so you can see both views. On page 32, there's a **JP Morgan** preview on **Meta** earnings.

On page 32, there's also discussion of the **TSM** double beat, but the stock still ended up down high single digits for the week. Page 33, we talk about the market preparing for the **SpaceX** unlock. Page 34, we talk about **ASML** boosting its net sales guidance but ending up down. We talk about the massive **IBM** miss on page 35. **Shopify** catalyst on page 37. Follow-up from a prior week. We talk about the **UNH** earnings beat, the **AVT** earnings beat on page 38. Page 39, we talk about why crypto bounced. Page 39, we also talk about the outflows from gold-backed ETF. And on page 40, we talk about higher real yields, uh, tightening their grip on precious metals. Page 41, we talk about S&P earnings momentum and how they're expected to increase by 24% this year, which is quite high, even though free cash flow growth will be lower.

Page 41, we talk about mortgage rates, um, they're highest since *August 2025*. Page 42, we talk about a negative, the negative catalyst for DGFX. Unfortunately, the reason why the stock is down is not just because of the AI sell-off. There's some regulatory uh, news in New York about um, building data centers. And we talk about why it isn't over for the company because it has some grand- grandfathered-in terms. Page 43, we talk about the Lucid bankruptcy halt and how the company refuted the halt, and now the stock's trading even higher than before the rumor came out. On page 44, we talk about **ASML** lifting its annual sales forecast for the second time this year but still selling off. Page 45, we talk about **Argus** starting **Sandisk** as a hold and why that resulted in that temporary stock sell-off, but we actually think short-term, **Sandisk** and **Micron** are both buys.

Page 46, we talk about **Stripe** and **Advent** bidding for **PayPal** and how that there could be a topping bid. Page 47, we talk about the **Vision** expert call we had and the summary model. We don't think there's a tremendous amount of upside because I found out there's quite a lot of customer concentration, but it's still, you know, relatively cheap to fundamentals. On page 50, we talked about, you know, we quote **Stena Research**, just a follow-up piece on the DeepSeek 2.0 and not so fast. Page 52, we talk about **Netflix** falling the most in four years on slowing earnings. **Netflix** actually could be a buy this

year. And then 54, we talk about Europe's reliance on, on US falling while reliance on China growing, which is actually quite dangerous for the entire manufacturing footprint in Europe.

Economic Calendar - 05:07

So let's quickly just jump into the economic calendar on page one. It's a very light economic calendar. The most important events for next week are the leading index, initial jobless claims, continuing claims, S&P US manufacturing PMI, service PMI, and new home sales, okay? The leading index, um, you know, is supposed to be frankly zero. Uh, so no surprises there. The market's kind of pricing in two, two and a half percent GDP growth for this year, lower than what was originally expected. We have initial jobless claims on the 23rd. Um, we have the service PMI and manufacturing PMI on the 24th.

And uh, on the, in terms of the earnings calendar, that's a lot more exciting, right? So last week, we had the banks report earnings. We had **JP Morgan**, **Bank of America**, **Goldman Sachs**, **Wells Fargo**, um, **Citi**. Then we had **ASML** on Wednesday, **J&J**, **Morgan Stanley**, **BlackRock**, **BNY**, **PNC**, **United**, **MT Bank**. On Thursday, we had **TSM**. Remember, **ASML**, **TSM**, most of the banks beat earnings, but you know, they still, the market still didn't rally, uh, because of this AI momentum fears and the new strikes on Iran. Then on Friday, you had **Volvo**, **Truist**, **Travelers**, uh, and **Fifth Third** report.

Earnings Calendar Review - 05:52

Now, on Monday, we have a big day. We have **Ryanair**, **Steel Dynamics**, **AGNC** mortgage REIT reporting pre-market. We'll see what their book value looks like. After the close, we have **Domino's**, **Zions Bank**, and **Bank Financial**. On Tuesday, we have uh, **Novartis**, **Charles Schwab**, **Danaher**, **Capital One**, **3M**, **Northrop Grumman**, **General Motors**, and **MSCI** reporting. And with respect to the home builders, we have **DR Horton**. On Wednesday, we have a big day. We have **Tesla**, **Philip Morris**, **GE Vernova**. Um, **IBM**, yeah, **IBM** pre-announced and the stock was down, you know, the most in many, many years. Um, we'll talk about that last week. **Texas Instruments**, uh, is also on the 22nd. Uh, **ServiceNow**, where we have a position is also on *July 22nd*. And then we have **Moody's**, where we have a position as well. Uh, and **CME Group**, where we don't have a position, but the stock has been really decimated.

Um, we'll see what **Southwest** has to say about oil prices on the airline side. And the most important companies to report next week are on *July 23rd*, um, **Alphabet** and **Intel**. So we'll see how the market's viewing hyperscalers and tech infrastructure. On the defense side, we have **Raytheon** reporting, medical devices, we have **Thermo Fisher Scientific**. Telco, we have **T-Mobile**. We also have the German tech giant **SAP** reporting on the 23rd, **Honeywell**, which just did a spin-off, **BMP**, the French bank, **Lockheed Martin** on defense, and **Comcast**, cable company, which is related to VSEN.

On Friday, we have **Charter**, which is also one of the biggest customers of VSEN, **American Express**, **Verizon**, uh, **SLB** on the energy services side, **Tenet**, and **Lamb Weston**. And then the following week is another massive week on the tech side.

Special Situations and M&A Updates - 08:21

So, in terms of the special situations update for this week, shares of **atai Beckley**, ticker **ATAI**, traded up 33% on Thursday after the company announced that they've entered a definitive agreement with **Eli Lilly**. Under the terms of the agreement, **Lilly** will acquire all outstanding shares of the company for 6.75 in cash plus up to 2.50 per share in the form of a CVR, contingent value right. Shares of **Electronic Arts** also moved higher, um, after positive press reports continued to fuel investor sentiment on approvals, uh, phase one clearance in the Euro zone. **Oak Eagle Acquisition** extended its tender deadline for its senior notes from *July 15th* to *July 30th*. **DNE** also announced last week that they made um, several regulatory filings overnight in Virginia, North Carolina, South Carolina. Um, so the deal is now expected to close in the second half of 2027. You know, it's a utility deal, so it'll take a long time to get approved.

Shares of **PayPal** traded up 20% on Wednesday following reports that the payment company **Stripe** and the PE firm **Advent International** jointly bid 60. uh, basically \$60 a share in cash. That's a 28% premium to the prior close, valuing **PayPal** um, at 53 billion or eight times forward EBITDA, which is quite low. According to the article, the firms made a proposal in early *May* and have yet to receive a response from **PayPal** two months later. were looking to advance discussions in the coming weeks, so there could be a topping bid, you know, much higher than \$60 a share.

The news follows a Bloomberg report from *February* that **Stripe** was considering an acquisition of all or parts of **PayPal**. Prior to that headline, **PayPal**'s shares had dropped 20% in *January* following an announcement of a CEO change and a 4Q profit miss. For **Stripe**, a deal with the company would combine its lead leading merchant-side payment platform with **PayPal**'s consumer-facing ecosystem, including Venmo. And I don't know how the trend, well, you know, given that it's an all-cash deal, you know, they would effectively delist **PayPal**. So it's not a back-door way for **Stripe** to IPO, although that would have been a good idea. Um, for **Advent**, the deal builds on its payment portfolio, including a recent 6 billion acquisition of a Canadian payments processor, Nuvei.

PayPal is also expected to generate about 6 billion of free cash flow next year, which is a 10% yield to the deal terms, which is quite cheap. The 8 times EBITDA multiple implied in the offer is roughly in line with **PayPal**'s trailing 12-month average multiple and a 1 to 1.5 discount to the average multiple over a three-year period, which is one of the reasons why people think that there could be a topping bid. In other news, **Lionsgate** is exploring a potential sale. **Warner Brothers** received EU FSR approval. NSA PSA secured shareholder approval and expects the deal to close on or about *July 22nd*, and **Re/Max** received um, session of its HSR. Now, **PayPal**'s, you know, chart you can see uh in the middle of the page, you know, it's rallied from, you know, a 40 level, but, you know, just given its free cash flow generation, I would say that the company, um, you know, could frankly receive a bid in the mid-60s.

Warner Brothers Antitrust Lawsuit Details - 12:09

Shares of **Warner Brothers** rallied about 2% on Tuesday after a group of 12 state attorney generals filed a suit to block the deal on antitrust grounds. Um, the strength uh, during the week was attributed to the fact that the complaint included very few surprises, so investors were already, um, accustomed to, you know, investors were expecting a worse outcome, uh, and that's why you saw some relief.

As expected, the complaint focused heavily on the combined company's market share in the distribution-wide theatrical films, particularly those with anticipated top-grossing, um, debuts. The key market share stat in the complaint was 30%, which is the percentage the states claim **P-Sky** and **Warner Brothers** would control of top-grossing films. After the merger, the top four US movie studios will control 93% of this market. The states argue that this control will lead to worse terms for movie theaters, i.e., worse revenue splits for companies like **AMC**, and less flexibility to discount tickets. It will also result in reduced output. Complaint references the drop in the number of films post the merger of **Disney**, uh, and **21st Century Fox**, which is true, and coordination amongst the four large movie companies.

Now, as compared to much of the pre-complaint speculation, we'd argue that the challenge focuses more on the overlap and basic cable channels and less on the labor market, which was one of the bigger risks. At \$27 a share, **Warner Brothers** shares are pricing in about a 75% odds of a deal completion, using a mid-teens standalone value. The next events include the judge assignment and likely hearing on the state's motion last week for a temporary restraining order.

We're also waiting for news from the European Commission, which should come on *July 22nd* on the Phase 1 deadline. The United Kingdom CMA, which is the most risky because they're socialists, on *August 7th*, phase one deadline, and an update from the UK's culture secretary on whether to intervene in the merger on public interest grounds. So you can see some details on the **Eli Lilly** deal on page three, um, on the **NextEra Dominion Energy** deal on page four, on the **PayPal Stripe** offer on page five with the press release link. **Lionsgate**, um, set to explore a sale via the content M&A wave. There's a link to that PR on page four. And then at the bottom of page four, um, you can see, you know, the biggest arb spreads for large caps only.

Redwood (RWTs) Baby Bond - 15:02

On page five, you can see the **Redwood** baby bond yielding about 10%, which we're still adding to the income portfolio. It was a new issue at the end of *May*. Um, so, you know, as of Friday's close at about 24.90, it's about 9.8%. Um, it's already announced two dividends. The first dividend is payable on *August 14th*. Um, so that's going to be, you know, above 60 cents. And then there's another 60 cent dividend payable on *November 13th*. So as long as you buy it before the X-date on *August 14th*, you should get both dividends. So that's about \$1.20 per share.

Nebius Deep Dive - 15:38

On the following page, on page six, we have an operational build-up by GPU, revenue, **Microsoft** contract, and **Meta** contract, which are the two biggest customers for the **Nebius** NeoCloud. Um, then you have the **Nebius** uh, bare metal revenues based on the type of GPU, the A100s, H100s, H200s, B200s, GB200s, B300s, Vera Rubin, and the Vera Rubin Ultra. The the the Vaman has been delayed, so that comes in, you know, closer to 2029. You have the number of GPUs in the installed base, the number of GPUs added per year, and the proportion of the year of total CAPEX, and what that means in terms of the, you know, the blended price per share. Um, and what the cloud revenues are going to be.

So what is **Nebius**? So, you know, **Nebius** is a NeoCloud, it could be undervalued if you include the stakes they have in growth businesses. Um, so, **Nebius** is a technology group formed out of the divested international assets of **Yandex**, which was a Russian e-commerce business and internet business. It's emerged as one of the most compelling pure-play AI infrastructure investments, um, and it also is net debt neutral. Uh, so it has much less, you know, debt to EBITDA than, you know, a company like **CoreWeave**. While the stock has experienced significant volatility after jumping 200% over the past year, now it's down 40% from that high. Wall Street remains bullish with a with a strong consensus buy rating between 244 and 265. You know, based on, you know, Friday's close of 177, you know, remember, Wall Street's always wrong, but, you know, on average, that's a 41% um, price target increase.

If you look at um, you know, reasons for the the company to perform, you know, it's a pure-play GPU as a service. Um, you know, **Nebius** isn't building uh, AI software. It's basically building the physical and virtual factories that power it. Um, the company functions as a specialized high-performance cloud provider, tailored entirely for AI machine learning workloads, because major hyperscalers like **AWS** or **Microsoft Azure** face massive capacity backlogs, AI developers are migrating heavily to nimbler GPU-as-a-service providers.

Nebius provides direct access to high-end **Nvidia** clusters with highly optimized infrastructure, allowing them to capture massive tailwinds from global the global compute shortage, which is still an issue. Number two, it's a sovereign AI company. So with respect to geopolitics and data uh, residency laws, um, they're creating a huge push for sovereign AI. The demand from European governments, which are far behind, and corporations to keep their AI workloads hosted inside Europe and under local legal jurisdiction is also helpful tailwind. The company's headquartered in Amsterdam, so it's strategically positioned itself as a premium European alternative to US cloud giants. The financial metrics show a business scaling at a very fast pace. In the first quarter earnings report, **Nebius** reported a massive sequential revenue jump to to 399 million, which was much better than Wall Street estimates. Um, you know, one of the bare theses has been, well, the company's spending heavily on CapEx like every hyperscaler or NeoCloud, it has two billion earmarked for expansion. However, the company recently de-risked its balance sheet by securing a 775 million asset-backed debt facility with a low interest rate directly tied to its past contracted customer cash flows. You know, they have customers in the Netherlands, for example. This allows them to aggressively scale their data centers without continuously uh, diluting current shareholders through new equity issuance like it was doing before. Because **Nebius** functions like a like a venture company locked inside a large cap stock wrapper, its traditional valuation multiples might look optically expensive. It trades at a high trailing P/E ratio, but its EBITDA is expected to grow five 5x from here.

If you look at, you know, its hyper-growth trajectory, global data center expansion plans, including targeting 1 gigawatt of power capacity and a massive data center build in Vineland, New Jersey, um, and long-term margin leverage suggests that cash flow will grow very quickly. Based on this tailored model, institutional research analysts flag the stock as inherently undervalued relative to its long-term earnings potential, and major institutional players like **Orbis**, **Allan Gray**, and **Two Sigma** have built out their positions. Uh, those are two large hedge funds in the name. So if you're looking, obviously to get exposure to **Nebius**, you have to be comfortable. It's a it's a high-beta risky asset. The stock has a beta over two, which means that it moves twice as fast as the general market. Beta is basically the slope coefficient on a regression of a stock's price over um, the **S&P 500**. Um,

you know, the stock recently pulled back roughly 20% in a single week due to broader, you know, AI sector profit-taking, so be very careful. The ultimate success of the investment depends on their ability to execute their massive data center buildout efficiently before global GPU supply completely catches up with demand. We believe that recent commentary from the AI ecosystem underscores strong AI demand driven by increasing adoption trends based on dynamics, including the adoption of agentic AI and also more cheap Chinese models and, you know, cheaper compute, uh, or cheaper models actually mean a higher demand for compute, which is, I know it's counterintuitive.

For example, in the semi space, **Micron** at its recent 3Q FY 2026 results guided next quarter revenue up 15% above consensus and commented that it expects memory supply demand imbalance to persist beyond 2027 with supply gradually improving in 2028, 2028. The **Meta** pivot to selling excess compute is also only incremental. Low-cost Chinese AI models may be incrementally negative, we're on page eight, for **Anthropic** and **OpenAI**'s margins, but they also mean that demand for compute should rise even faster because Chinese models are, you know, 95% cheaper than **Anthropic**. Not to mention that the Kimi.ai, the thing that everyone was scared about last week, the new Claude competitor, just today has restricted new users as they're running low on compute. If you look at their uh, Twitter handle, kimi.ai or @kimimoonshot, they basically said that Kimi's K3 model has received far more love than they expected, and its GPUs are feeling it. Over the past 48 hours, demand has pushed close to the limits for its current capacity. To protect the experience of existing subscribers, they're temporarily pausing new subscriptions and prioritizing compute for current members. Existing subscribed users are not effective. They're adding capacity as fast as they can, which means that they have to buy GPUs and memory, and will reopen new subscription spots in batches, which means that, you know, they're probably lagging behind.

Additionally, um, we highlight an increase in enterprise adoption of AI as companies deploy agentic AI with the aim of streamlining workflows, improving efficiency, and reducing costs. As such, we believe that **Nebius** is well-positioned to benefit from these tailwinds in the coming years, given number one, its solid execution in ramping power capacity. Number two, AI platform approach that combines hardware and software capabilities. And number three, relatively higher reliability of its solutions versus peers. First, recent data points from the AI supply chain point to a robust demand outlook given accelerating AI adoption. For example, **Qualcomm** at its recent investor day stated that it expects agentic AI to drive a step-function change in compute demand with up to 1,000x tokens per query and 40x token growth by 2030.

Additionally, **Micron** at its recent results highlighted that it signed 16 take-or-pay strategic customer agreements or SCAs across DC consumer and automotive end markets. Importantly, many of these agreements contain both floor and ceiling prices, and the cumulative amount of committed revenue from these agreements represents over 100 billion over 5 years at floor prices. As such, we believe that AI demand is a structural tailwind for the AI ecosystem in coming periods, and within the AI infrastructure space, **Nebius** screens favorably given its strong execution in ramping power and competitive positioning. Further, we highlight that investors remain focused on the execution of NeoClouds and ramping up capacity as demand continues to outstrip supply. And this vein, we note that **Nebius** has upgraded its contracted power target in successive quarters and now targets over 4 gigawatts of contracted power by 2026. At its recent earnings, the company also confirmed it's fully delivered the first **Meta** contract in first quarter of 26 with its **Microsoft** ramp on track. Additionally, we attended um, well a peer attended a company's inflection event in San Francisco where the company detailed a portfolio-based approach to capacity planning, which we see is helping de-risk targets. As such, we note that the power target for any given site is based on grid power levels rather than alternative energy source potential from hydrocarbons, which are treated as an upside. More broadly, the company also continues to utilize lease capacity where appropriate and maintains geographically diversified footprint across Europe, the US, and the Middle East. It's one of the leaders in Europe. And furthermore, we note that recent commentary suggests an acceleration in uptake of AI by enterprises, which bodes well for **Nebius**.

For example, we highlight that **Nebius** is seeing 3 to 5x over-subscription in demand for each GPU deployed outside of hyperscaler customers. Our confidence is underpinned by **Nebius**'s AI platform approach where it provides AI startups and enterprises with a broad suite of offerings across both hardware and software. Additionally, we highlight that its platform offers relatively better reliability versus many alternative compute providers, with one enterprise customer indicating reliability is levels up to 5x higher than competing solutions.

Finally, while some customers and investors have queried whether recent trends of token maxing could be a headwind for **Nebius** revenues, we note that the company's business model has always focused on improving the cost dynamics with falling cost per tokens, with new model releases in the past having actually led to further uptake of AI compute and revenue acceleration for **Nebius**, something we believe that will continue happening. Further, we continue to believe that AI platforms represent a viable solution for enterprises aiming to scale their business. In this vein, we highlight that the company recently launched its AI Cloud 3.6 with improved capabilities and **Nebius**'s first AI agent, **Nebius Echo**, for managing cloud infrastructure and using natural language. Changes to estimates. Recently, consensus increased top-line estimates by 10 to 61% across 2027 to 2030 to reflect positive views on **Nebius**'s core AI business given the backdrop of strong adoption trends by enterprises. Note, **Nebius** signed deals with multiple enterprises in the first quarter, including

Monday.com and Revolut in Europe. Additionally, we believe that **Nebius** plans to have over 4 gigawatts of contracted power supplied by the end of this year, giving the company flexibility to sign additional multi-year contracts similar to **Microsoft** and **Meta** later.

Further, gross margin estimates for 26 to 28 remain unchanged but are slightly higher in 29 and 30 to reflect a more favorable mix and higher top line. Additionally, our EBIT estimate is unchanged for 2026 but grows um, changes between -310% and 49% across fiscal 27 and fiscal 30 because 27 is a growth and investment year and 30 is a sharp growth year. Um, finally, EBIT estimates basically have grown 0 to 65% over the forecast period, uh, back-end loaded. **Nebius** currently has a market cap of 45 billion. The stock is down 40% from local highs. Over the last two weeks, sell-side estimates have only gotten higher with revenue uh in 2030 now closer to 44 billion.

The company also has some stakes in some big equity growth stories like **ClickHouse**, **AV-Ride**, which is partnering with **Uber**, **Toloka**, and **TripleTen**, which could be worth 8 billion alone, and the business is net debt neutral, which means its cash level is almost the same as its debt level. **ClickHouse** is an open-source database management system. **AV-Ride** is a leading developer in autonomous cars and robotics. **Toloka** provides data labeling for AI models, similar to what Scale AI does. **TripleTen** provides online education and training for tech AI-related jobs. But **Nebius** owns about 28% of **ClickHouse**, which recently raised a round in *January 2026* at a 15 billion valuation.

So, you know, if you quickly do that math, before further dilution, you know, 0.28 times 15 is, you know, 4.2 billion, you know, just in this, um, in this, uh, **ClickHouse** company. Then the company owns basically, uh, 83% of **AV-Ride**, which is an autonomous driving company, um, and that was valued recently at 4 billion, you know, so that's that stake was worth roughly 3.32 billion, and I don't think sell-side analysts really focus on this enough. **Nebius** also owns 50% of **Toloka**, which was, uh, in *May of 2025*, was given a 250, uh, sorry, 500 million valuation. So that's roughly worth 250 million. So already you add 4.2 billion to 3.3 billion, that gives you 7.5 billion plus 250, so you're at 7.75 billion. And then **Nebius** owns 100% of **TripleTen**, which is probably worth around 250 to 500 million because it recently did about 54 million in revenue up 87% year-over-year, and just given the growth, it'll get a high rev multiple. That's about 8 billion in value, okay?

DataBricks just raised about 3 billion at 188 billion valuation, which is approximately 27x their 6.9 billion ARR, which is insane. **ClickHouse** is a similar business. They raised a round in *January 2026*. The company was reportedly doing about 250 million ARR but was growing 200% a year, which means that they raised at a 60 times multiple. Now that we're six months past that round, I don't know what ARR is, but if they're growing 200 to 250% a year, I wouldn't be surprised if they're close to a 500 million ARR with a good shot at 1 billion ARR next year, um, which means that the valuation could be, you know, much higher next year. We all know that the robotaxi wars are coming with **Waymo** currently in the lead, **Tesla** behind them, and **Uber** with their multiple partners in third place. The numbers are still small for all of these companies, but that is going to change, um, drastically over the next three to five years.

It's very hard to know where **AV-Ride** valuation goes from here and how many robotaxis they'll be able to deploy with the **Uber** partnership they have. So **Uber** invested about 375 million in their last round, which is quite a big deal. However, if **AV-Ride** becomes a leading provider on the **Uber** platform, it is possible that they could have at some point over 100,000 robotaxis deployed, you know, past 2030, in which case the company could be valued at 10 times the last round in a bull case, right? Depending on revenues, margins, CAPEX, depreciation. It's worth noting that deploying 100,000 plus robotaxis will still require a lot of dilution. So **Nebius**, you know, does have a 83% stake, but that could get watered down as the company grows. Um, you could see the model on, you know, below our discussion on page 10. And basically that what that does is it takes, it shows you 2023 to 2030 revenues, gross profit, adjusted EBIT, um, and net income. And then it shows you the segment revenues, um, by **Nebius** AI infrastructure division, the bare metal, the AI cloud, ancillary business, total segment revenues on page 10. And then it shows you the cash flow statement where, believe it or not, the company is supposed to do over 22 billion of operating cash flow um, over the next four years. Um, **Wolfe Research**, um, sees **Nebius** doing about 34 billion in revenue and 21 billion in EBITDA by 2030. And consensus has actually grown in the last two weeks to 44 billion in revenue with the ramp beginning in the fourth quarter of this year as their massive data center in Vineland, New Jersey comes online and it unlocks more contracted revenue. The path is backed by about 46 billion of signed **Microsoft** and **Meta** deals, including 27 billion of **Meta** agreement with 12 billion fixed and 15 billion of optional capacity that **Nebius** can either fill for **Meta** or redirect to higher paying customers, protecting the downside while preserving spot market upside. The margin story comes from a cloud pipeline that grew 3.5x quarter over quarter, excluding hyperscalers, a rising mix of inference and software, and the 643 million EGAN AI acquisition, which together should sustain growth gross margins in the high 60s. You can see the **Wolfe Research** project, uh, projections on page 11.

A Brief Intermission - 34:24

Now, right before we get into the **Dan Niles** discussion on AI overinvestment, um, I'm just going to get a glass of water and then be right back.

Dan Niles on AI Overinvestment - 35:46

Okay, so every great industrial revolution has had overinvestment due to potential for riches for the last company standing. Eventually, this industrial revolution does result in a bust. AI is no different, but he thinks that the current speed bump is not the beginning of an ultimate bust, at least this year. On *Saturday, June 20th*, Niles laid out a near-term concern on the AI speed bump due to, one, token minimization, two, cheaper Chinese models, three, competitive low-cost open-source LLM models, four, rising semiconductor costs, and five, impact on Q3.

On *Monday, June 22nd*, in hindsight, it turned out to be a short-term top for the semiconductor index and the momentum trade. Token maximization in *March* turned into token minimization in *June* with the most extreme example being the **Coinbase** CEO on *June 26th* posting on how they cut their token spend by nearly 50% by largely routing AI queries to cheaper models. In the near term, the question becomes, can token usage by the other 99% of firms go up fast enough to offset the top 1% of firms like **Coinbase** cutting their AI bills?

Numerous cheaper LLMs have been introduced recently. Last week, the introduction of the Moonshot Kimi K3 China-based challenge, the performance of the most US, most advanced US models like Claude. However, we learned today that that model doesn't have enough compute. So it takes that risk off the table temporarily. Profitable AI-native revenue growth that is readily positive is what is needed to keep the whole ecosystem functioning.

Google is reporting this Wednesday. It will provide the first major data point on the trade-offs between cheaper tokens and more token production. **Google's** cloud platform has seen revenue growth year-over-year accelerate from 34% in Q3 of 2024 to 48% in Q4 and 63% in the first quarter of 2025. The growth rate should accelerate further in the second quarter due to token maximization. I think that Niles was referring to 2026, not 25. While he expects forward revenue and EPS to move higher post-results due to their core business, comments on GCP growth are likely to drive um, the stock's reaction. I also think that the market is going to focus very closely on margins.

Now, **Google** is his favorite consumer AI play given that they have a complete AI stack. He is not sure they are immune from the leading 1% of companies cutting their AI bills. **Uber**, for example, blew their entire AI budget in a single quarter.

Intel, which reports on Thursday, you know, by the way, **Uber** has um, **Google** Cloud as their primary supplier. **Intel** reports on Thursday, provides multiple ways to win at the AI infrastructure layer, which he is more bullish on than the increasingly commoditized LLM model layer. Number one, agentic AI is driving a surge in demand for server CPUs, which are the new bottleneck. Number two, their advanced packaging has already attracted several hyperscalers. And number three, their foundry business supported by the US government continues to improve and attract new customers. As for the AI speed bump stocks, the **Morgan Stanley** Momentum or Momo index, which subtracts the long index from the short index, fell 28% since *June 22nd* through *July 16th* in just 24 days with a slight reprieve of 1.6% on Friday. Historical corrections of over 10% since 1996 have averaged 20% from peak to trough, but have taken 48 days on average to bottom. But the rally of the Momo index of 40% from *March 30th* to *June 22nd* was also much sharper than historical precedents.

Market Update and Geopolitical Situation in Iran - 40:13

As I'm talking, you know, futures have bounced a little bit. **Nasdaq's** up 0.53% as the Korean sell-off is not as bad as people thought. **S&P 500** is up 0.15%. Russell was down -25, but now it's flat. This is a mid-mid-uh, hour futures update. So in his opinion, investing is about risk versus reward, and on *June 20th*, it was not a good risk versus reward with increasing examples of token minimization. From a technical basis, Momo is still not fully oversold with an RSI only reaching 35 on *July 16th*, and on average, it bottoms at 31. But the risk versus reward is much more favorable today with sentiment having fallen further following, number one, the negative pre-announcement by **IBM**, which declined 26% last week, despite prior claims of being an AI beneficiary. Number two, the 3% drop in **ASML** and the 8% decline in **TSM** last week, despite positive earnings. And number three, the 10% decline in both the semiconductor index and the momentum index just last week, on top of a decline in the prior week.

Finally, he believes that the advent of agentic AI, which arguably started on *January 30th of 2026* with the formalization of open claw, requires 10 to 100x more tokens versus chat-based AI. As a result, he believes it's prudent to start adding back some exposure in AI-related infrastructure names as we have been last week. He believes this needs to be balanced with prudence at the public cloud layer given the near-term focus on AI bills needing to be controlled by the top 1% of corporations. On page 12, you can see the decline in **TSM** despite the double beat in earnings, which we'll talk about in a moment. Attacks on Iran continue into its eighth day. The US actually proposed its own tolls in the Strait of Hormuz, which would have been much higher than Iran's tolls, and then canceled them. By Friday, we saw six days of US strikes in a row, eight by Sunday, with the US focusing on taking out IRGC infrastructure that could be used to damage vessels. They basically blew up about six bridges. Um, they attacked several um, water uh, desalinization plants, power stations. Uh, it's becoming quite a merciless uh, set of attacks on both sides with Iran also attacking the US Air Force base in Kuwait.

Brent crude rallied to about \$86 a barrel and **WTI** rallied to about \$81 a barrel on Friday. And as of now, **WTI** has rallied to about \$84 a barrel, and Brent is actually over \$90 a barrel. So basically, Trump, you know, it's now, he said like 45 times that we've won the war, um, in the last 100 plus days. Um, and obviously, it's still going on. Um, but he's he also says that they have reinstated the Iranian blockade, so there should be much fewer vessels going through the strait, which brings back the risk of inflation, unfortunately, which was looking to abate last week with the low CPI and low PPI. He also called the United States the guardian of the Hormuz Strait, you know, um, and then on the following page, he talks about how oil is flowing like never before. It's frankly the opposite. Oil is not flowing like it ever has before because the strait is effectively now closed again. The UAE is actually looking to build a new port and a container terminal on the country's east coast to bypass the Strait of Hormuz entirely. Now details include that Dubai's **DP World** is in talks to develop a new multi-purpose port in the coastal area of Fujairah. The new project would allow containers to enter and leave the country without having to pass through the strait. Shipments would then be moved on the trucks overland to Dubai, Abu Dhabi, and neighboring Gulf countries, and the Iran war is going to result in many new pipelines and ports being built to avoid this hazardous strait. There were some peace talks in Rome on *July 15th*, which were not fruitful. US, Israeli, and Lebanese officials concluded two days of talks in Rome, described as productive by senior US official. The parties agreed on a structure and guidelines for a regional pilot process expected to be implemented in the next coming days, according to Axios reporter Barak Ravid.

VIX Analysis: Why It's Low Despite Geopolitical Risk - 45:03

Now, let's quickly discuss why the VIX has been so low since the Iran war restarted. Now, the war came back on *July 6th*. The fear gauge went down. That's not complacency. It's the third natural experiment for the framework and the cleanest one yet. On *July 6th* and 7th, Iran fired on three commercial vessels transiting the Strait of Hormuz. The US answered with strikes on more than 80 targets the following day, then a 140 target overnight assault, and then more. President Trump declared the truce over. The naval blockade of Iranian ports was reinstated on *July 14th*, with the United States now styling itself the guardian of the strait and charging shippers 20% of the cargo value for passage. Iran then struck tankers in Omani waters, fired at US facilities across five Gulf states, and as of this writing, hit a water desalinization plant in Kuwait, while American munitions fall on Iranian bridges. By any sober reading, the 2026 Iran war has restarted and is only escalating.

The CBOE volatility index of the **S&P 500** closed the first Friday of the restarted war at about 15.02, down 5% on the day, while the **S&P 500** finished within a whisker of its record close. Through two weeks of tanker attacks, 100 target strike nights and reinstating blockade, the VIX has traded in a small band between 15 and 17.5. It closed *July 16th* at 16.73. The financial commentary has one word for this: complacency. The market we are told is asleep. The fear gauge is broken or suppressed or both. We've spent five months arguing the opposite, and the restart is now the third consecutive natural experiment to come back with the same answer. The VIX is not low because the market is ignoring the war. The VIX is low because the market has learned in real time through two prior tests, exactly where this war's risk actually lives. And it's pricing that risk with precision in the places the headline gauge does not measure.

Two wars, one gauge. The VIX price round one blind and round two with a map. If you look, um, you know, round one versus round two. If you start with a comparison that frames everything else, this is the same conflict, the same choke point, the same belligerents, many of the same weapons run twice, five months apart. When the war began on *February 28th*, the VIX repriced violently. It jumped 23% to 26.43 within days, pushed towards 29 intraday in the first week, and held

close from the mid-to-high 20s for two weeks, peaking near 31 in late *March*. The market was pricing in an unknown. Nobody knew whether a closed Strait of Hormuz meant 1973-style asymmetric repricing of everything, whether the inflation shock would unanchor rates, whether the dispersion architecture would hold under a genuine supply-like catastrophe. Now, there's several several reasons why oil prices didn't go ballistic. Number one, you had China slow down its buying because it has, you know, had over a billion barrels of reserves. Number two, you had Japan and the US also leaking out reserves. US reserves are now all the way at 1980 levels, so uh, it's a temporary reprieve. And then you also had several countries that, like Vietnam, couldn't import as much oil as they thought anyway, so you had demand destruction which resulted in lower demand and lower prices.

The restart is the same war with an answer key. Since *July 6th*, the VIX has not closed above 17.5. At the identical stage of the episode, tanker attack strikes blockade, the index gauge is trading roughly 10 volatility points below its round one path. What changed between *March* and *July* is not the war. It is the market's information about the transmission mechanism. Round one tested in live conditions and at full scale whether an identifiable sector-constrained geopolitical shock gets routed through the broad equity index or around it. The answer came back unambiguous. OVX, or oil volatility, to 126, while the VIX never sustained above the low 30s. Energy freight and rates carrying premium, the S&P absorbing the most consequential oil disruption in a generation through sector rotation rather than index-level panic. The dispersion mechanism did not merely survive its hardest test, it was seen surviving by every options desk and alligator in the market. Round two is being priced by a market that watched round one. That's not only complacency, that's Bayesian updating.

Whether the second war is actually being priced. If the framework is right, if the post-COVID architecture routes identifiable shocks to where their exposure lives, then the restart premium should be visible, large, and located outside the VIX. And it's all three. If you look at the second war being priced in, most of the volatility is actually in oil, with a 39% spike in oil volatility, a 13% spike in the price of oil, a 9% spike in the VIX and **Nasdaq** fall, and that's also because of the AI sell-off. A 4% increase in the two-year Treasury yield, a 6% increase in S&P fall, a 0.3% decline in the **S&P 500** and a 3% decline in Treasury fall. Since, and that's because of lower CPI, PPI. Since the restart, the crude oil volatility index has repriced roughly 39%, jumping 18% in a single session on *July 7th* to 47.59 that day on towards 60 mid-month, while the VIX ticked up to 16 on the same session, almost nothing. The oil vault equity vault ratio ran roughly 3x, the same asymmetric signature the first war stamped on the tape at four times. **WTI** posted a 9.4% single-day increase on *July 13th*, its largest since *April*, settling at 78.14. I mean, today as we're talking is at 80, almost 84.

The rates complex is carrying its share. The two-year Treasury yield pushed to a 15-month high near 4.23% as the market prices the inflation transmission of a reclosed Hormuz into Federal Reserve policy. The market is still worried about a surprise July hike because of this. So it's not a cut story anymore, a hike risk story is what the market is priced in. The MOVE index firmed, Korea's KOSPI, an economy with an acute imported energy exposure, fell 7.3% in a single session on the weekend, um, while S&P futures slipped half a percent but they've still since rebounded. Korea was obviously also down because of margin calls, because of the ban on single stock leveraged AI ETFs and just a general correction in the AI names post the listing of **SK Hynix** in America. And the **Nasdaq** 100 volatility sits at 27.3 against a VIX of 16. Single name and concentrated index volatility is trading at nearly double the broad index measure.

The last pair deserves a moment because it connects to the single most direct piece of evidence the framework has ever been handed. One-month implied correlation, the option market's own estimate of how much stocks will move together, is sitting at a two-year low. With the single-stock implied volatility running well above index level, implied correlation is the dispersion mechanism expressed as a number. When the market prices individual stocks to move a lot while the index moves a little, it is saying in the actual prices of actual options that the stress will be absorbed name by name, sector by sector, but not index-wide. The mechanism, um, so this is actually quite interesting. So what the market is doing is basically saying that we're going to absorb the shock through a sector rotation while the overall index volatility will not be very high. The war premium is enormous, it's simply not in the VIX, it's in the oil VIX because VIX measures the one thing this war twice now has demonstrably failed to threaten, broad correlated index level equity risk. The vol surface is not shaped shaped like complacency yet. There's a second independent reason to reject the complacency reading, and that's the shape of the volatility curve itself.

Complacency has a shape, flat and cheap everywhere. A mark, a market that has genuinely stopped worrying about risk sells the entire surface of the vol curve, near-dated vol, far-dated vol, tails, everything. What the current service actually shows is nothing like that. One-day implied volatility closed the week of the restart under 10. Spot VIX was at 15, uh, but the front VIX future traded at 17.5, the three-month traded at 18.57, which is a steep firm contango, paying up nearly nine

vol points from the one day to the three-month vol point. And the SBOE skew index has been pinned between 144 and 146 through the entire episode, deep in the very elevated zone above 140, the same tail bid reading that defined the *June's* war-end divergence when the VIX fell 8.4% on the peace news and skew rose in the same session. If you read those three coordinates together, and we'll move on to the next point in a moment, the market is selling the immediate center because the dispersion mechanism has twice demonstrated uh, it absorbs this specific war's shocks without index-level volatility. It's paying up for the later-dated volatility because people are worried that volatility will rise because of the escalation path, the CPI transmission, and the Fed's reaction function are generally unresolved. And it's keeping crash protection expensive because the post-COVID market never sells the tails anymore, war or peace, a structural feature that we've documented since the volatility surplus got its name. Relaxed center, bid curve, bid tail. This this is not a market that's super complacent or asleep. It's a market that's pricing in a precise, differentiated risk map or a rotation. And the map, um, you know, is something that we'll continue to talk about.

Final Thoughts on the VIX and Falsifiable Edges - 56:15

The framework owes readers um, its falsifiable edges. So here's a reading in the current tape that doesn't flatter the thesis. 20-day realized volatility in the **S&P 500** has roughly doubled since early *June*, from under nine to now the mid-teens, brushing 18% at one point. With spot VIX in the mid-teens, the implied to realized ratio has compressed towards one, and on at least once session, implied printed a hair below realized. The structural at-the-money premium, the IV to RV ratio, has persistently been in the 1.3 to 1.5 range and is now closely to one. Two things about that. First, it is exactly what the term structure says it is. The market is treating the war-driven churn in realized volatility as transient, dispersion-absorbed noise rather than the beginning of a correlated regime or a big risk-off, which is why the premium has migrated out the curve to the three-month point rather than vanishing. Second, it is the fragile point in the whole configuration and it's worth naming plainly. An index selling one-day volatility under 10 while realizing in the teens is the market leaving itself no cushion against the one transmission channel this war has genuinely uh, seen as systemic, the inflation to rates channel. Oil at 78 is absorbable. Oil at 110 with a two-year yield breaking out and a Fed forced from hold to hike is not a dispersion story, it's a correlation story. And correlation is the one thing the current surface is priced for the absence of. So that's the main risk, that oil goes to 110. That is the falsifiable marker, safe for the record. The framework says that the VIX stays in the teens while OVX, rates volatility, and single-name volatility carry this war unless and until the rates channel activates. A sustained VIX break above 25 that arrives through oil-driven inflation repricing rather than equity equity-native catalysts would mean that the dispersion mechanism has found its limit. If you look at the two-year yield, not the tanker count, that's what's going to tell you where the market's going to go.

So there's a nice table that summarizes this on page 18, and we'll close this discussion and move on. The war's onset tested whether a generational supply shock routes through the index gauge. OVX went to 126 and the VIX never sustained above the low 30s. The war's end tested when the unwind would occur. OVX collapsed by more than half while the VIX modestly ends and skew rose on the very session the peace arrived. The tail refusing to relax on good news. And now the restart of the war has tested what a market does when the same shock arrives a second time. It prices the mechanism rather than the headlines, sells the center, a twice-watched hold, pays for the curve and the tail, and posts implied correlation at a two-year low while the oil complex repriced 39% in seven days. Onset, unwind, and restart, three different phases, three different directional shocks, one identical structure signature. The probability that this pattern is an artifact, that the VIX is merely suppressed, and the suppression coincidentally reproduces the framework's predicted geometry through a spike, a collapse and a respike gets smaller with every test. Suppression is a level story. It cannot explain a shape that keeps reorganizing itself in exactly the same asymmetric way across opposite sign events. The war restarted 11 days ago. Oil volatility repriced by a fifth. Crude had its biggest day since *April*. The two-year yield hit a 15-month high. Korean equities fell 7% in a single day. Single-name volatility is trading at twice the index. Crash protection is pinned at a very high elevated level. The volatility curve is paying up nine points from tomorrow to three months from now. And the VIX is still at 16 because the VIX is the gauge of one risk the war has now failed twice to produce. The fear is not missing, it is felt correctly.

The Momentum Sell-Off and South Korea's ETF Suspension - 60:31

So, now if we jump to the next page, we'll look at the momentum sell-off. So this has been a very sector-specific sell-off. Um, we've had a historic sell-off for momentum in *July*, according to **Goldman** and **Morgan Stanley**. The **Morgan Stanley** Momo index was down 28%. The **Goldman Sachs** high beta momentum index is down 24% month to date through the first half of *July*, the worst performance since *April of 2009* after Lehman Brothers crashed. The **Morgan Stanley** tech momentum index 17-day rate of change was down 35%, the worst ever recorded through its 27-year history. You can see that spike down in purple. You can see the **Goldman Sachs**, um, monthly 27% decline also in the table on page 19.

What made matters worse, um, last week was that South Korea suspended new listings of single-stock leverage ETFs. So the company, the country will temporarily halt these new listings, according to Financial Services Commission. Authorities will temporarily suspend all single-stock leverage products and raise minimum cash deposits required for investing in domestic and overseas single-stock leverage products to 30 million won from 10 million won. So it's 3x how much money you need to buy these leverage products, which is a good thing because of over-speculation. But this resulted in the Korean index selling off quite dramatically. The past month has been the craziest 30-day period in the history of the Korean stock market. Actual vol has now topped to the highest levels in Korea since the great financial crisis. The Korean stock market, the KOSPI VIX was at 81, which is basically where our VIX was in 2008. That's how crazy the vol is. If you look at this red graph in the middle of page 20, you can see that their decline was so fast that their VIX is the same level it was back in 2008.

You can look at **SK Hynix** now down 40% from its all-time high. Um, and, you know, the market, you could argue that earnings are overinflated, which is why the market is so low. But the Korean, you know, PE multiple is based on how much earnings have grown at **SK Hynix** and **Samsung**, is at only basically 5.89 times.

Open Source AI Era and Hyperscaler Chip Demand - 62:51

Gavin Baker, who is a very successful tech portfolio manager, is basically said last week that the mega bull case for AI infrastructure would be if market share actually shifted away from certain frontier labs with 90% inference margins towards cheaper models, because the more and more cheaper models are available, the more you will have demand for compute. Basically, what he's saying is it would increase the ROI on AI spend for end customers by increasing intelligence per dollar, which would then drive incremental token demand. Margin dollars would effectively get redistributed from frontier labs to AI infrastructure providers and infrastructure winners would be those with the lowest per-token cost, and winners at the model layer would be those with the highest token efficiency. So, you know, he's talking about NeoClouds like **Nebius**. He's talking about very efficient, you know, Chinese models that are being adopted by **Microsoft**. So investor Gavin Baker posted on X that the risk-reward now appears attractive again for cheap stocks possessing durable competitive advantages expected to exceed earnings expectations over the next 6 to 12 quarters.

So in terms of the open source era, will hyperscalers keep buying chips? Semiconductor investors have had a run of unsettling news. Hyperscalers are announcing far fewer data centers and chip stocks that rode the AI story have been sliding. On top of that, on *July 16th*, Moonshot released Kimi K3, a 2.8 trillion parameter AI model that placed level with the top US models in a blind arena, um, like the new Claude model. The new model is scheduled for release, the Kimi model, on *July 27th*. A wave of open models promising frontier level performance at a fraction of the price has led to a simple worry. Would hyperscalers keep buying chips at this scale? Has the market really moved past the semiconductor theme? This article works through the numbers in the public view, how open models make money, what frontier developer's margins actually look like, what happens to hyperscaler profits when workloads move to cheaper models. The goal is to answer one question: does the switch to open models make hyperscalers buy fewer chips, and where does that compute and memory demand go from here?

So number one, how open models make money and how fast they're spreading. A token is the smallest unit of text an AI model processes, and the model fees are charged per token. Frontier models are the top closed models such as **OpenAI's** GPT or **Anthropic's** Claude, sold only through a paid API. Open models are models whose weight files are published so anyone can download and serve them. They are so-called open-weight models, and that's what **Microsoft** is doing with these Chinese open-weight models. They're basically tweaking them and just using them and only paying for the compute. Open model developers earn money differently from the frontier. Because the weights are public, a company that hosts the model itself owes the developer nothing. Even when the developer sells its own paid API, competing hosts sell the same model cap what it can charge. The economics of that hosting business are real. Inference specialist reports serving

70 billion plus models profitably at 0.9 to \$2 per million tokens. And together AI said that its annual bookings passed 1.15 billion. Serving margins exist in the open camp. What is zero on the self-hosting path is any mandatory premium for the model itself.

On capability, EpocAI's measurements put the best open models about four months behind the frontier models, which is quite unnerving. The Kimi K3 model, which was released on *July 16th*, placed above the previous frontier generation in blind area rankings right away, and level with the new, levels with the new frontier models. The problem, and the reason why this is not a big worry, is that Kimi basically ran out of compute today. And it'll take many, many quarters for it to grow its compute. Usage data shows the shift as well. OpenRouter is a routing platform that lets developers call hundreds of models through one API. So where its traffic goes is a wide proxy for model market share. In this platform, the share of calls sent to models priced under \$1 per million tokens rose from 18% in *January 2026* to 41% in *June*, which means more and more big US companies are using cheaper Chinese models that are open source than US models. The share of tokens processed by Chinese open models such as DeepSeek, Qwen, and Kimi, went from under 2% in late 2024 and now to about 61% in *May* and *June of 2026*. Counting all open source weight models, the latest weekly snapshots are about 69%. So volume is moving to cheaper, mostly open models. The arithmetic of frontier margins we'll discuss in a moment, but I do need to get a little bit more water because we've been talking for a while.

The Arithmetic of Frontier Margins - 01:08:44

Okay, so the arithmetic of frontier margins. While volume moves at this speed, where has the money gone?

If you start with frontier mart, uh, the biggest model developer's margins, we can start there. So public reports put **Anthropic's** gross margin, which is revenue minus the cost of serving, at -94% in 2024, improving to positive mid-60% by 2026, with a 36% operating margin before model training costs.

OpenAI's inference cost reportedly quadrupled in 2025, and its adjusted gross margin slipped from 40% to 33% as it lost market share to **Anthropic**.

If you add training costs, the picture changes. Reporting based on **OpenAI's** 2024 investor projections puts its 2026 loss at \$14 billion. **Anthropic** told investors that it expects its first quarter operating profit in the second quarter of 2026 to be about 560 million at a roughly 5% net margin, while stopping short of promising a profitable full year.

The accounting basis are not identical but for scale, **Anthropic's** 36% before training, a research estimate, sits near **AWS's** 35 to 38% operating margin, or 37.7% for the first quarter of 2026. While the margin including training is about 5% even in the first profitable quarter, far below the hyperscalers' 30s. What matters more than the margin rate is the premium that's actually being paid.

As of the late, as of late *May 2026*, **Anthropic** processed about 11% of tokens on open router, but took 42% of estimated model spend on the platform. Benchmark gaps have narrowed to a few points, yet a 25x price gap, 25 to \$30 per million tokens for Opus versus under \$1 for budget open models, keeps getting paid.

The reason is that companies do not pay for benchmark points. They pay for an agent that doesn't fall apart in the middle of a 50-step job, a reliability where failure is expensive, and to avoid revalidating their whole workload on a new model.

So if you look at the table in the graph below, for very complex projects, money stays with the frontier, even though it's 11% of tokens and that seems low, it's 42% of spend. So the real risk that open models pose to frontier developers is not a falling margin rate. It is the loss of the mid-tier volume. Open model hosts are now offering APIs that are drop-in compatible with **OpenAI's** and **Anthropic's**. So switching models can be done in a one-line change.

They work well and uh, and they work well with defined requirements such as document classification, extraction, summarization, and simple code. What stays on the frontier is the hardest work where a capac, where the capacity gap is worth money. Prices there are holding, but as the middle drains away, frontier revenue growth depends more and more on how fast that top tier actually expands.

And that top tier, right, is going to be funded by the biggest VCs, the biggest companies, and even the biggest companies like **Microsoft** are diversifying to the open-source Chinese models, as Satya Nadella said a couple weeks ago.

Impact of Open Models on Hyperscaler Revenue - 01:12:39

So how does the switch to open models change hyperscaler revenue? Hyperscaler AI revenue comes through three accounts. One, renting compute to frontier developers. Number two, charging cloud customers for model usage, and number three, selling their own products with models inside, like **Microsoft** Copilot.

The open model switch changes the customer account first. So let's look at a case. An enterprise, let's say pays \$25 for a job with 1 million tokens on a frontier API like Claude or **OpenAI**, and follows the money before and after the switch. Before, the payment reaches the hyperscaler indirectly. The \$25 first becomes the revenue for a developer such as **OpenAI** or **Anthropic**. At the reported 65% gross margin, the developer spends \$8.75 per million tokens on the serving cost. That is buying cloud compute and keeps \$16.25 to fund payroll and training and capex. The hyperscaler then books 8.75 of revenue. At a 35% infrastructure operating margin, that's \$3.06 of profit and \$5.69 of server operations. For simplicity, this model treats the developer's entire serving cost as hyperscaler compute purchases, assuming that after the switch, the company deploys directly on the same hyperscaler's cloud and prices every tier on the same basis of 1 million output tokens.

When the same company moves to an open model, the pass-through disappears. The weights are free, so nothing is owed to a developer, and everything the customer pays to run the model in the cloud is infrastructure revenue for the hyperscaler. Structurally, this is better revenue. Compute rental that is used to ride on someone else's model becomes a direct sale to the hyperscaler's own customer. What comes with it is a price collapse. The same million token price falls from 25 to under \$1 for million tokens.

If you look at money flow, you know, before, you had an enterprise customer paying \$25 for a 1 million token job. The frontier developer would keep \$16 of that. The hyperscaler would take 8.75. It would spend 5.70 on operations and profit \$3 of that first 25 spend. Now, in an open model, a direct model, the enterprise customer pays \$1 for the same 1 million token job. But it reuses part of the \$24 it saved. So the full payment is \$1 plus the 24 times the re-use rate. Now, people forget about this. So it's not going to be a complete collapse. All the money that is saved by the enterprise customer, okay, they're going to look at that and they're going to re-spend that on additional compute, because lower costs for compute means higher demand.

So the hyperscaler revenue will now equal the full repayment. Profit of 35% and the server operations of 65% multiplied by the \$1 plus the 24 times the reuse rate. So at 100% reuse rate, the revenue is 25, the profit is 8.75, and the server ops is 16, uh, is 16.25. So what ends up happening is at 100% reuse rate, the hyperscaler actually earns more money, which, which I don't think people really get.

So that \$1 is not a discounted version of the \$25 model. It is the price of a small model whose active parameters and therefore compute per token are a 10th to a 50th of the frontiers. Open serving prices form where competing hosts add a margin over the cost. Hence the profitable 0.9 to \$2 price points. So in both paths, the compute the hyperscaler provides equals about 65% of the revenue it books. So if you look below, what a million tokens cost by model tier, you can see that the developer share, you know, completely goes away with the budget open model, okay? But in that scenario, you reinvest all of that you save as the enterprise user, and that money goes to the hyperscaler.

If the switch is to a top open model, there is little to weigh. So basically Kimmy 3, this this Chinese model that's supposed to be just as good as **Anthropic**, the payment of 9.75 is larger than the frontier's 8.75. So the hyperscaler revenue and server operations rise the moment the switch happens regardless of any reuse. The case where the outcome can go either way is the trade down to a full budget model. So the base case continues on the \$1 path, the least favorable one, favorable one for the hyperscaler. Under the assumptions above, the size of the customer account after the switch is set by the reuse rate. How much of the saved \$24 per million tokens goes back into tokens? If none of it does, the hyperscaler revenue is one and the profit is 35 cents, far below the \$3 below, before. If 32% is reused, the profit matches. So what this is telling you is for a breakeven at the hyperscaler, you only need a 32% reinvestment rate, okay, not 100% reinvestment rate. So basically, if the enterprise customer spends 25 per million tokens on Claude and spends \$1 for a cheap Chinese commodity model, and then reinvests 1/3 of the \$24 it saved, that's actually a breakeven profit for the hyperscaler, okay?

If all of it's reused, revenue is 25, profits 8.75, server operations are 16.25, which is 2.9 times the level before the switch, which means that demand for GPUs and memory would actually be much higher. So the Chinese models actually mean more demand for compute, as long as the reuse rate is high, and the reinvestment rate is high, and more demand for memory.

So if you look at the diagram on the next page, if you look at the same \$25 split by reuse rate, the 32% reuse rate is the hyperscaler's break even. If companies respond just a third of what they save, the hyperscaler earns just as much as before. So where are companies actually sitting? Reports keep coming in of annual AI budgets being used up in a single quarter and token usage is growing five to seven times a year. That is not responding savings, that is just growing the budget itself. Reality sits to the right of the rightmost bar on the chart. In that zone, the 16.25 premium that is going into the frontier developer disappears and the hyperscaler profit and server operations fill the space. The other two accounts point the same way. In the own product account, swapping in an open-source model removes the license fee from the cost line and improves the margins. The developer rental accounts carries a risk. The mid-tier volume loss squeezes the frontier revenue. Developers could slow their training compute purchases, but multi-year rental contracts with **OpenAI** and **Anthropic** sit in the backlog. So this account is not one that breaks so quickly.

So a hyperscaler whose profits grow with this switch faces the next question, raise, hold, or cut compute and memory investment. And the truth of the matter is they likely don't cut. So the whole point of this piece is basically to say that these Chinese models are not going to result in less demand for compute.

Soft CPI/PPI Data and Federal Reserve Commentary - 01:21:01

Now, we are seeing some risks because to pay for all this compute, you do need to borrow money. So you have seen the hyperscaler G-spread or the risk spread on bonds for borrowing money has gone up and it's been peaking in the middle of *July*.

Now, we're going to take a break from AI for a moment and discuss the soft CPI and PPI trends last week, which were very positive for crypto and resulted in a small bounce in gold before the Iran war restarted in a bigger way.

So last week's whirlwind of Federal Reserve commentary and inflation data created a few market gyrations, but ultimately left interest rates lower by the end of the week. In a need for rates less, rate hikes less urgent. Evidence that inflation may finally be back on the decline is welcome news for investors, even as it arrives with a reminder that the newly opaque Federal Reserve may limit the extent to which we can enjoy it. We think that the current economic backdrop continues to support a Fed on hold, neither hiking nor cutting rates, which we argue is the best outcome for risk assets in the near term. The Federal Reserve will only be compelled to hike rates if inflation becomes a more powerful enemy and the economy is running uncomfortably hot. If the Fed is cutting interest rates anytime soon, it will be because the enemy is weak growth and a softening labor market. It's important for investors to know their enemy when thinking about the forward rate path and find ways to diversify around various risks.

We learned this week that what inflation takes away from investors, it can give back by moderating and allowing rates to fall and equity markets to multiples to expand. Both consumer and producer price inflation for *June* came in significantly softer than expected, permitting investors to breathe at least a temporary sigh of relief. If it wasn't for the war, the market would have ended up higher.

The Impact of Geopolitical Events on Inflation and Markets - 01:23:21

Now, if, and this is a big if, given geopolitical events this week, energy prices remain comfortably below their peaks from earlier this year, inflation could continue to moderate in coming months, but that's a big if. Traditionally, the three main sources of inflation are energy prices, housing, and labor. Right now, neither housing nor labor look like major risks. Energy remains a major wild card. As commodity traders become more sensitive to the risks of a prolonged conflict in the Middle East, futures markets may begin to reflect a higher for longer environment for oil prices.

Should the conflict de-escalate once more, however, energy should continue to exert a downward pull on annual inflation. You can see *December*, uh, *September* and *December* Brent crude futures, which had bottomed at the beginning of the *July*. Uh, they have spiked again. If they rise above 100, we're kind of back in shit territory. But as long as they stay, um, below 90 and go back into the 70s, and there's another peace treaty, the market can rally.

Markets of course continue to view inflation risks through the prism of interest rates. Upside inflation surprises tend to push rates higher, as investors demand a higher yield for fixed coupon payments and B, anticipate higher and tighter policy via higher interest rates by the Federal Reserve. A central bank's reaction function matters a great deal for how markets respond to inflation news. History is rife with examples of how central banks, primarily but not exclusively in emerging markets, can either ignore or miscalculate the risks of higher inflation and see their currencies and several bonds sell off sharply. Stocks reactions have been more case-dependent. On one hand, higher inflation can help profit margins. On the other hand, too much inflation can lead to financial instability and economic collapse, neither of which is good for business. More credible central banks, which have nearly always included the Federal Reserve, tend to get the benefit of the doubt from investors on handling inflation risks. Communicating and executing policy calmly and competently can keep investors' inflation expectations low, and those low expectations can hopefully become a self-fulfilling prophecy. One look at the five-year, five-year forward break-even rate of inflation, essentially the inflation risk premium priced into the Treasury market, tells you that at no time in the past five years have inflation expectations become unanchored from the Fed's 2% target, despite the fact that the target's never been reached.

Enter Kevin Warsh, the new Fed chair who has vowed to bring inflation back down to target to help preserve the institution's credibility. Warsh testified before Congress last week for the first time and chose to become less hawkish. His comments revealed little about his plans for monetary policy, however, or even his framework for making decisions. Chair Warsh's colleague on the Federal Board of Governors, Chris Waller, has been on the opposite side of the transparency spectrum lately. Waller gave a speech last Monday in which he strongly hinted at a willingness to vote for a rate hike on *July 29th*, which is the next Fed meeting, if inflation data, specifically core inflation, worsened. His comments sent the odds of a *July* hike to as high as 45% intraday on Monday before CPI and PPI brought it back down. You can see those in the blue and black graphs in the middle of page 28.

Gold's Performance Amidst Geopolitical Uncertainty - 01:27:10

Barring a prolonged return to oil prices above \$100 a barrel and gasoline above \$4 a gallon, it's already getting close there because of crack spreads, which we can discuss another time, inflation may remain quiet enough for the Fed to keep policy on hold for the remainder of the year. If action must be taken in either direction, both hikes and cuts were floated as possibilities in the *June* meeting minutes and is unlikely to be good news for equity markets. Hikes would mean high inflation and volatile bond markets while cuts would only come amidst a serious macroeconomic deterioration. The pricing out of hikes, which would show up as a drop in the blue line above, would be positive for just about every asset class.

Before we wrap up, it's worth spending a moment on gold. The yellow rock or the barbaric relic seemed a tailor-made for a year like 2026 in which geopolitical uncertainty and inflation have both spiked unexpectedly. Yet gold has been the worst-performing asset in the last couple months. Gold has seen a particularly sharp correction, just like crypto, as real interest rates, yields on tips, or interest rates excluding inflation drifted higher. As the graph above shows, gold investors can take solace in the fact that their position is up still 19% over the previous year, and part of the decline has been the government forcing less speculation on gold in China. But gold's price has fallen below \$4,000 an ounce after briefly spiking above \$5,400 an ounce in *January*.

You could argue that gold is suffering from a surplus of Fed credibility. Despite elevated inflation, the Fed has been able to keep inflation expectations low, as we showed above, while real interest rates have actually risen.

Market Sentiment, Fed Transparency, and the July FOMC Meeting - 01:28:33

There's simply less of a need to hedge against inflation risks using gold than there might have been with a wobblier or less credible central bank. The Fed's greatest asset, its credibility, is gold's worst enemy.

No one likes a week of soft inflation better than us, but if we are permitted one nitpick, it's that the markets are still wary of a surprise hike from the Fed this summer. The *July* FOMC meeting is less than two weeks away. There isn't any more inflation data coming before then. And yet, there is still a non-trivial chance of a quarter point hike priced into futures. We think that the remaining odds of a hike derive from markets simply not knowing what the chair is thinking or what he'll do. This lack of transparency has contributed to the market's diminished understanding of the Fed's reaction function to incoming data. Indeed, this may be a feature not a bug of the of the Warsh administration. It will take some time for investors, not to mention investment strategists, to adjust to this new reality. Inflation surprises still represent a serious

risk to asset prices, but this week provided a view into disinflation's ability to lift both the floor and the ceiling for portfolio values for the remainder of the year. That outcome is still possible if geopolitical risks stay at a simmer and the Fed starts a steady course.

Market Wrap-Up and the AI Infrastructure Trade - 01:29:47

Lastly, we would be remiss if we didn't note the lack of a rally in equity markets this week despite cooler inflation and falling expectations for rate hikes. Equity market weakness has instead come from a sharp unwind in the dominant and large-weighted AI infrastructure trade. Stocks of AI infrastructure producers have been negatively correlated to inflation risk this year, performing the best in *April* and *May* as prices and rates weighed on the rest of the market, yet struggling in *June* and *July* as inflation risks have abated somewhat.

HBM Memory Pricing and Its Impact on Chipmakers - 01:30:31

Now, **Morgan Stanley** says memory pricing is likely to peak, okay? But on the following page, we present the opposing view on memory pricing from **BlueFin Research Partners**.

So HBM looming price increases and potential impact on **Nvidia**, **AMD**, and **Broadcom** GPUs and XPU's. The ongoing price negotiations with the memory suppliers suggest that HBM4 pricing could double or even triple in 2027. We believe **SK Hynix's** recent press release highlighting delays in HBM4 production ramp in order to redirect resources to conventional DRAM was in part a veiled threat to **Nvidia** that HBM margins need to come in line with DRAM or else. Table one highlights the HBM4 content per GPU or XPU platform including the present COGS as well as the COGS if ASPs were doubled or tripled. **Nvidia** Rubin R200 utilizes 288 gigabytes of HBM4 RAM with the present COGS approaching \$5,000. Based on **Nvidia's** typical 75% gross margin, the HBM4 content would increase from roughly 19,000 to 38,000 if prices doubled or 58k if ASPs tripled. In *April*, supply chain reads suggested that the VR, the Vera Rubin 200 rack was approximately \$6 million. We are now hearing that this rack is roughly \$8 million due to memory price increases and could exceed \$9 million if HBM4 prices triple. **AMD's** MI450 has 50% more HBM4 content than the R200 at 432 gigabytes and **AMD** does not mark up the memory nearly high as high as **Nvidia**.

Nevertheless, their solution costs will increase significantly within anticipation of the HBM4 price increases. Some industry sources believe the shelving of the **Google** Pumafish, the initial 2-nanometer ASIC designed with **Broadcom**, could be due to HBM4 pricing. Pumafish was slated to have 576 gigabytes of HBM4 RAM, putting COGS at an astronomical level approaching 20,000 if prices are doubled and 30,000 if prices are tripled. The anticipated price increases may have changed the economics unfavorably with Pumafish, while doubling the HBM4 may provide a significant performance boost, it may not make economic sense if it's going to result in a very expensive XPU. You can see the **Nvidia** R200, the **AMD** MI450 and the **Google** Pumafish, and how much HBM4 memory they actually consume. 288 gigabytes for **Nvidia**, 432 gigabytes, which is insane for **AMD**, and 576 gigabytes for **Google Pumafish**.

Meta Pre-Earnings Analysis - 01:33:26

GPM comments on **Meta** pre-earnings. Um, you know, the main risk here is how is **Meta** going to monetize its huge capex investment and are they going to show much higher capex? One of the things we're worried about is that now that the stock is up 20% over the past two weeks, the risk is that while **Meta** said that they're going to rent out excess compute, and that's why the stock bounced, they might still increase capex spending for the second quarter, and that I think would scare investors.

If you look at the 12-month free cash flow of **Oracle**, **Meta**, **Alphabet**, **Microsoft**, and **Amazon**, you know, it's basically, you know, now gone from 300 billion of free cash flow to essentially negative.

TSM Earnings Report - 01:34:19

TSM, you know, had a double beat, but US ADRs were lower on high capex investment. Um, it was a really stellar earnings release. EPS was 437 versus 337 for the quarter. Revenue was 40 billion versus 39.7 billion. Gross margin was 67.7%, operating margin was 60%, net profit was 55%. So basically, they had a great quarter, but because the company is investing so much in capex, 100 billion in Arizona, um, raises its capex guidance from 60 billion to 64 billion due to

generative AI, um, that kind of scared investors because ideally you want a company to grow its earnings, but then, you know, as it slows down, you want it to give money back to the investors. But in this case, um, you know, that's not really happening.

So for those of you who are not familiar, **Taiwan Semiconductor's TSM** revenue growth by platform is about 66% HPC for AI, 22% for smartphones, 5% for IoT, 4% for automotive, 2% for others, and 1% for DCE. You can see how much the stock sold off at the beginning when the earnings came out. The Taiwan stocks fell into technical correction actually last week. I mean, Korea was already in a correction, now it's basically in a full sell-off, you know, being down over 20%.

Um, the decline in Taiwan's Taichung basically took its decline from *June* to nearly 11%. The market had been caught off guard by the Chinese AI pioneer Moonshot, but now I think people are more comfortable that it's not a very big threat.

SpaceX Upcoming Unlock and Trading Strategy - 01:36:19

So let's just jump, given the amount of time. The market is preparing for the **SpaceX** unlock. The next unlock is going to be in *August*. Uh, so what we are thinking, what we have is we are short **SpaceX** versus long **SATS**, or which is now **ECHO**, E C H O. And so I think the, at least last week, **SpaceX** stock was down more than **SATS**, so that spread should be compressing.

ASML Sales and Market Reaction - 01:36:47

ASML also had really good sales for the year. They make the um, lithography machines for etching semiconductors, and they saw net sales of 43 to 45 billion, um, which was higher than the estimate of 39 billion, but the stock still didn't do well. Um, the gross margin they saw was 55 to 57, which is fantastic, beating an estimates of 52 and a half, I thought was fantastic. Um, second quarter revenues were up from 9.3 from 8.85 billion to 9.3 billion. So overall, the company had a good quarter, but because of the AI fear, um, the market sold it off, even though the CEO said, we're planning to add 30% to our 2026 low North America EUV capacity. The market still doesn't believe that it that this is a growth business, which is quite, you know, quite interesting. Markets absolutely very, very skeptical of these names.

IBM Earnings Miss and Discretionary IT Spending - 01:37:52

Um, now, **IBM** had a massive miss, um, which we can talk about. Discretionary IT spending is worsening and will likely be the main theme across most software companies when they report results. What's happening is that more and more companies are spending on AI and they have less money to spend on regular software. The CEO, Arvind Krishna, said the company had expected supply chain issues to weigh on the results, but he said that the company failed to predict that its customers would end up shifting away from their spending away from **IBM** products to servers, storage, and memory purchases to hedge against further price increases. You know, one of the main issues is that, you know, **IBM** has been very slow to innovate. Um, you know, **Oracle** has a big problem as well. **Oracle's** problem is that they have too much exposure to **OpenAI**, and that's a much different problem. You can see when **IBM** missed on its guidance, the stock fell over 20% pre-market and basically stayed down.

Um, so they talked about, you know, what played out was worse than our expectations driven by a shortfall in our Z performance in the associated software stack, primarily in transaction processing. In the last few weeks of *June*, they saw clients shift their quarterly capex spend towards server, storage, and memory, uh, to secure supply-constrained infrastructure ahead of the expected price increases. This dynamic impacted client buying patterns. While we anticipated some supply chain-related impact on our expectations, we didn't anticipate the magnitude of the capex re-prioritization. In addition, clients were distracted with rapidly evolving industry-wide cybersecurity, uh, changes during the quarter.

Shopify Catalysts - 01:39:38

We talk about some **Shopify** catalysts such as exponential AI traffic, higher margin financial scaling, and up market enterprise adoption. If you want to read it, this is just an update from our **Shopify** work a couple weeks ago.

UnitedHealth Group (UNH) Earnings Beat - 01:39:50

UNH basically beat, um, on earnings. It raised its outlook and reported quarterly profit well ahead of Wall Street news. I think **UNH** is reporting expected profit between 19 and a half and \$20 a share. It's a little bit better than their guidance last quarter, you know, after the stock sold off. Um, but you guys can read this full piece.

Abbott Labs Earnings and Stock Performance - 01:40:12

Abbott Labs stock was up 11% after they posted strong results after many quarters of lackluster results. The revenue was up 13% year-over-year to 12.6 billion, that exceeded analyst expectations of 12.54 billion. Overall, the company did quite well across medical devices, uh, diagnostics, that established pharma and nutrition.

Crypto Rally and Gold ETF Outflows - 01:40:37

Um, you know, the reason why crypto rallied very briefly was that, you know, the chance of a Fed hike fell to about 8% last week. The largest gold-backed ETF, **GLD**, also recorded a \$14 billion outflow since, uh, since *March*, but you are seeing more central banks add gold and you're seeing investors actually restart to buy gold now that it's under 4,000, which is quite interesting. There is, um, 9.6 billion of outflows across Bitcoin ETFs, but there's 14 billion of outflows across gold. It's funny how everyone is rushing to buy gold at 5,600, but no one wants to buy it under 4,000, which is what you're supposed to do.

Precious Metals, Real Yields, and Market Futures - 01:42:33

So higher real yields, um, are one of the reasons why, uh, precious metals prices have been so weak. Uh, just give me one minute to get some more water here.

So higher real interest rates resulted in silver basically selling off much faster than gold. As I'm talking, S&P 500 futures are now back to flat after a brief rally with Asia opening and you and, uh, US Tech **Nasdaq** futures are also flat. And the small cap Russell is now back to negative. Crude oil is still around 84 for WTI and 90 for Brent. If you look at bond prices, the US 10-year Treasury yield is actually up 2.8 basis points to 4.57%, which means mortgage rates are going to make another new high. US 30-year is at 5.07 and the US 5-year is at 4.28. We're even seeing a spike in overnight SOFR, which is the benchmark interest rate for pricing commercial real estate and securitization tranches. So that's frankly not a good thing. You know, you really just need this war to be over, and I don't really see it de-escalating immediately.

Silver briefly traded, uh, at fresh eight-month lows last week before tripling losses, extending a remarkable sharp reversal from this year's highs. The metal is now roughly down 20, 21% year-to-date, dramatically underperforming gold, which is only down 7%. The divergence speaks to silver's dual identity. It behaves both as a precious metal and an industrial input, leaving it more exposed when the tape sours. But it's also that there's less Asia speculation in silver. The move in rates has been become increasingly difficult for the metals complex to ignore. Real yields have pushed higher as the market digests the latest rise in oil prices and Chairman Kevin Warsh's message of patience, raising the opportunity cost of holding non-yielding assets.

Gold has held up relatively well thanks to continued structural support from central bank demand, like the Chinese central bank aggressively buying, the People's Bank of China. Yet silver has enjoyed no such cushion. Its higher volatility has instead amplified the adjustment, turning what began as a broad precious metals pullback into a clear case of silver underperformance. On page 41, you can see the S&P 500 earnings are now expected to increase 24% this year. We've never seen this type of earnings growth, uh, outside of a post-recessionary rebound. An unprecedented boom fueled by massive EPS gains in Big Tech. Remember that free cash flow gains have been much slower than these EPS gains this year because of all the capex spending.

US Mortgage Rates and Housing Market Trends - 01:45:27

Mortgage rates in the US reached 6.55% last week, which is the highest since *August of 2025*. US mortgage rates rose to the highest level in more than a year as a flare-up in the tensions in the Middle East renewed concern about concerns about inflation. The average for the 30-year fixed loan was about 6.5%, 6.55% up from 6.49% a week earlier, **Freddie Mac** said in a statement on Thursday. Rates haven't been this high since last summer. Mortgage rates have been climbing for the second straight week, tracking an increase in Treasury yields as the peace deal brokered between the US and Iran about a

month ago broke down. The war has weighed on the housing market with mortgage rates climbing by more than half a percentage point since the conflict started in late *February*.

Sellers outnumbered buyers by almost half a million in *June*. Listings are lingering longer, according to **Redfin**. Even so, the median sale price hit a fresh record high in *June*, propped by wealthier buyers. There's a pool of higher-income buyers who are purchasing seven-figure homes, according to Chen Zhao, **Redfin's** head of economics. But a lot of first-time and average move-up buyers are priced out of the market as mortgage rates are around 6.5%, making monthly payments hard. Contracts to buy previously owned homes dropped 5.4% last month from *May*, declining in all four major US regions, according to the National Association of Realtors.

DGXX and the New York Data Center Moratorium - 01:47:05

Now, um, we'll talk about China's trade surplus later. Um, negative catalysts for **DGXX** in New York, and why it isn't over. So update on **DGXX**. So last week, Governor Kathy Hochul of New York signed a one-year executive order freezing air permits for new hyperscale data centers, 50 megawatts plus. So where does this leave **DGXX** uh, roughly 140 megawatts of power capacity in the state?

So here's a clear balance breakdown. So the CEO basically said on *July 9th* that they basically tackled the New York permitting concerns that have weighed on investors. He noted that **DGXX's** sites have been operating for years and framed the current crypto mining setups as tier one data center operations. His key line is that we are grandfathered in in terms of upgrading into AI data centers, so we are carved out of any issues banning New York State builds. We're just upgrading to tier three.

The statewide moratorium is also not a blocker. Hochul's executive order signed last week pauses new large-scale data center air permits for one year. It gives the state time to develop unified standards amidst concerns over grid strain and electricity costs. Because **DGXX's** North Tonawanda facilities are existing, long-running operations, the statewide pause on fresh permits does not appear to directly block upgrades. The CEO's grandfathered framing aligns with how these moratoriums are usually written. They target new development, not vesting existing assets with prior permits and operating history.

This is where it gets complicated. A mid-*June 2026*, North Tonawanda extended its local moratorium for another 12 months. The law explicitly covers new data processing centers and expansion of existing facilities. Data processing centers is defined broad, is defined broadly to include AI processing, server farms, and high-density compute. Expansion includes increases in processing capacity, intensity of use, and infrastructure. The extension was passed in part to study proposals to convert the existing crypto mining site into AI data data centers, exactly what **DGXX** is evaluating.

Existing operations can continue. A meaningful upgrades that increase capacity or change infrastructure like require local approvals are now paused. There is a hardship exception process through the Common Council, although it's not automatic. A successful shift to modern AI infrastructure, especially with quiet, quieter liquid or glycol cooling systems, could actually help address the core community complaint over time. So how will this play out? In the next 6 to 12 months, existing operations will continue. The seat light order is largely a non-issue for **DGXX**. The local moratorium creates a temporary pause on formal approvals for significant expansions. **DGXX** will likely explore whether its planned upgrades trigger expansion rules or pursue dialogue with the city or evaluate the hardship exception path. In the medium term, the resolution is problem, probable. Many data center conversions navigate the local opposition through better technology, quieter, more efficient systems, economic development arguments, and phased approaches. Once North Tonawanda completes its zoning updates, the path should become clear. 140 megawatts of controllable power remains a strategic, valuable asset in a power-constrained AI world.

The upside case would be successful conversion, even partial, to the AI colocation or GPU services from, you know, obviously the the crypto space. The bottom line is the CEO's public comments hold up well on the state level. The new Hochul order targets new large projects and do not appear to block upgrades to existing projects. And the local North Tonawanda moratorium is more directly relevant because they explicitly includes expansions and was extended partly to study this exact conversion. It's not a clean all-clear, but it's also not an outright ban that some bears have portrayed on **DGXX**.

Lucid Bankruptcy Rumors and Stock Rebound - 01:51:31

Lucid also had a bankruptcy halt last week, and the company refuted the bankruptcy and said it has liquidity until next year. That resulted in a bunch of degens buying the stock, which traded all the way down, you know, to, you know, 270, 280, and then now rallied all the way back to 736 pre, uh, after the close on Friday.

ASML Sales Forecast Lift - 01:51:54

ASML also lifted its annual sales forecast for the first time, you know, with 43 to 45 billion of sales versus, you know, 36 to 40 billion. Um, they saw net sales of 43 to 45 billion and gross margins of 54 to 56%. So, you know, their numbers were pretty good. It's just the stock didn't move. Um, they're planning to add 30% to their 2026 low NA EUV capacity so they can produce help Bjensen produce even more chips. Um, one of the reasons why there was an AI selloff last week was that **Argus** cut **SanDisk** to hold. Um, but you know, that's frankly not a big deal.

Stripe and Advent Bid for PayPal - 01:52:38

Stripe and **Advent** basically bid for **PayPal** at 60.50, um, 60.50 per share, or about 28% more than **PayPal's** closing price on Tuesday. The bid is backed by 50 billion in bank financing, which is highly levered. **PayPal** was founded in the 1990s by Elon Musk and a bunch of the **PayPal** mafia owners. Um, it's still it's since struggled with monetizing its payments, but it's done a pretty good job. Uh, it's just it's not growing as fast. So they that's why **Stripe** and **Advent** are trying to buy them. **PayPal** was planning to cut about 20% of its workforce in the next two to three years. The planned reductions represent the part of the new CEO Enrique Lores' effort to reduce costs and turn around the payments company. Um, **Stripe**, which was founded by Patrick and John Collison, has become one of the industry's most coveted players. And if they think that if they can buy this company, they can find some synergies and increase the cash flow. ner thinks, you know, the stock's worth 70.

VISN Asset Sale Analysis - 01:53:46

So **VISN**, we had an expert call on this name, and what we did was we mapped out all the asset sales. You might have to zoom in to read everything. We laid out the enterprise value in 2023, 2024, 2025, and post the two asset sales and then the third asset sale they're going to do. Frankly, gets us to um, about 32% upside on the stub and 19.5% upside on the stock, because half the stock is cash. We have to strip out the cash, and if you strip out the full cash and the dividend, um, it's about 45% upside just on that business's value. Um, and that business, you know, frankly only trades at three, four times EBITDA, it should trade at six to seven times EBITDA. We did it on a cash flow basis because a lot of these big cable companies like **Charter**, **Comcast**, **Altice**, they need these upgrades because they don't have fiber, and they have these old coaxial cables in the US. So if in order to to to transmit, you know, analog to digital signals and have multiple signals per cable, they need this infrastructure, okay, that is installed by this company. And what they've done is they've sold the majority of of, uh, their businesses except for the upgrade technology and licensing fees. So we really just need to see whether the government will approve the deal or not.

Um, you guys can read the expert call with Carlos Treves, who's the VP of product and growth at **Zayo Communications**, which is an industry that buys this compute. Um, so he talked about the pros and cons and the major risks down to page 49.

China AI Risk Analysis - 01:55:30

On page, um, 50 is a similar piece to what we shared earlier, um, from Stemma Research, which basically talks about how, you know, this whole China AI risk, uh, is overblown. Um, but you know, I would just let you let you read through it because I've already gone through some of these points. Um, and then there's a really cool graph which shows the cost of a million tokens, and you can see how like Fable 5 is 57.4, and if you look at, you know, all the Chinese models, right, they're much, much cheaper.

Netflix Stock Performance and Strategy - 01:56:24

The preliminary UMich sentiment index was actually quite good last week, which is one of the reasons why the market bounced before selling off again. Um, **Netflix** shares fell the most in four years. They were down 12.5% on Friday, despite heavier content spending and expansion into podcasts, gaming, and short-term video. I just think they're trying to do too much, um, and that's why the business frankly isn't growing. Um, **Netflix** announced a flurry of new deals with popular social media personalities like Alan, the Allen Chicken Show, and Nick, uh, Giovanni. Um, the amount of people, the amount of time people spent on **Netflix** grew 2% in the first quarter, which isn't impressive. A slight improvement over last year, but not a big improvement in subscribers or in time spent, which is likely why the stock sold off. But I think it'll be a buy again.

Europe's Reliance on US and China - 01:57:13

Um, we can talk about, you guys can read this piece about Europe's reliance on the US following while increased, uh, reliance on China growth. Um, but just given how long the call is going to be, um, let's just get to some of the questions.

Q&A Section - 01:57:24

Thoughts on Nokia - 01:57:29

So **Nokia** is back down in the tens. Um, of all the AI names, it's probably the least interesting. Um, it's not going to be one of the first companies to rally on the way back up. It's not as correlated. But it's one that we can spend a little bit of time on. Um, basically, orders for **Nokia** AI and cloud are supposed to be strong. They raised its full-year, uh, guidance, just like all these other companies. Um, I'll be the **Nokia** changes and estimates to, uh, to the growth channel. So you guys can just everyone can just see the same, same data and understand the question. And thank you, Bitcoin Tina. **Google** earnings are Wednesday, um, in the p.m., uh, *July 22nd*. Um, not Thursday a.m. So I pasted **Nokia's** estimates. Um, so 30% of **Nokia's** revenue comes from, uh, optical networks and IP networks business units, which are expect, which are more exposed to, um, AI infrastructure. Just put this in the growth tab as well.

In this whole segment, the optical networks and IP network segment is supposed to grow at 10 to 12% a year, and they believe that **Nokia** can grow the sales significantly. At the same time, due to operational gearing in the in the NI business unit coupled with synergies from the, uh, **Infinera** deal, the earnings of the company could grow about 14% this year. If **Nokia** can deliver on its earnings growth, uh, we believe that the stock could rerate with its earnings multiple rising in the mid-teens. So **J.P. Morgan** thinks the stock's worth about 18 euros, which would be a 29 times multiple, which just seems very, very high. The peg ratio of one should be an appropriate metric to value the stock. So if the because the company's growing so quickly, um, they're saying that if you use a peg ratio instead of a P/E ratio, **Nokia** actually looks interesting. So it could be one to spend more time with. I'll just copy this and um, share, you can't always trust Claude because they oftentimes they make mistakes with the numbers. Um, but you know, you can often times just test it out and see, you know, what they've said about other companies and how it's played out. Um, but here, I'll just paste this in as well, the **J.P. Morgan** thesis. Um, so the company in 2027, right, it's supposed to do adjusted EBITDA of about 4 billion. And the current enterprise value of **Nokia**, let's just check, 4 billion EBITDA and the enterprise value is 48 billion. So, yeah, not especially cheap. So 48 divided by 4. So 12 times. Um, and then on earnings, they're supposed to be 41 cents. So 10 divided by 0.41, so 24 times earnings. Um, it could just be a momentum reversion type name. Um, but I would get, if I were to do the momentum trade in a very big way, you know, size up **Nebius**, size up **Oracle**, size up, um, **Micron** and **SK Hynix** and **Nokia** and all these names. Um, on **Oracle**, I'd be more careful. I wouldn't add that because of **OpenAI** right away. Um, but if you're to add all these names, you really need interest rates to stop going up and you need the Iran war to stop. But in the meantime, you know, **Nokia** could be one to do, uh, some more work on. So let's just go back to the Q&A tab.

Comparison of Nebius and CoreWeave - 02:02:37

Thank you, Bitcoin Tina. David is saying, why **Nebius** over **CoreWeave**? The reason why we like **Nebius** over **CoreWeave** is while **CoreWeave** is also likely a buy, it just has too much debt. You know, if you look at **Nebius**, it basically has no net debt. Debt minus cash is basically almost flat or zero. So the company is basically growing very quickly, it doesn't have a lot of debt. Whereas if you look at **CoreWeave**, it has about 40 billion of debt. They're actually raising some more debt this week, apparently, I've heard. Um, so while both of them look cheap, um, **CoreWeave** is also burning a boatload of cash, and it's supposed to, you know, it's burning a lot of cash and borrowing a lot of debt, and it's being more aggressive with

its growth. I think **Nebius** is also growing quickly, but it's growing quickly without borrowing a lot of money. Um, and its growth rates, I think, are like three to four times higher than **CoreWeave**. So you could frankly buy both, uh, if you believe in some sort of AI recovery. It just **Nebius** I just think has no net debt, so and it's growing faster, so I think it's a little bit more interesting at the moment.

Model Reuse in Frontier vs. Open Models - 02:03:49

Why is the reuse different from a frontier versus an open model? Why do the need to reuse one version not the other? Well, basically, Bitcoin Tina, let's just say you had 100 million to spend, okay? And that's all you had to spend in your budget. You spend that 100 million. Um, 25 is the cost of these high-end models, let's say, right? Um, so you can basically buy, you know, 1 million tokens, 4 million times. Now, let's say we are using a model that costs 25 times less. So that means you can buy 1 million tokens 100 million times, okay? So you can do the same calculations with a cheaper model, let's say give or take. Okay, you can do 25 times more calculations. Okay. What that means is that, you know, you could either do 25 times more calculations or you could just do the same amount of calculations and, um, realize that you have a lot of cash left over in your AI budget. And then if you spend basically a third of that, um, on the on hyper scalers, they will earn the same amount as they did if you just bought the compute using your 100 million in total. So basically, what what we're saying is that of the 100 million, um, if you spend on the same amount of compute, instead of, instead of a 100 million, you basically would only spend 4 million. So that would leave you with 96 million. So then if you take 96 million and divide it by three and you spend 32 more million plus the four, you spend 36, uh, million instead of 100 million to get the same output. Um, so if you're spending that much money to get the same output and your output is profitable, it makes sense for you to spend a little bit more to get that type of return.

So, all we're trying to say here is that, you know how **Uber** spent all of its, uh, money on AI compute in the first quarter? Now imagine if they only spent 5% of the money in the first quarter, but all of that AI spend was profitable, then they could spend more. But that, the fact that they're spending more on compute, on cheaper compute, that money they spend still has to go to the hyperscalers on the cloud spend. So if they end up spending more on tokens in general, right, and recycling that capital, that ends up resulting in more, um, and more hyperscaler usage because, you know, the GPU content and the store, all of that stuff and the memory usage is the same as if you spend on a Chinese model or if you spend on a frontier model, on a good Chinese model or a frontier model. But if you if you spend the same amount of money on a commoditized model, you get way more demand and that ends up being more, way more compute and way more hyperscaler revenue.

It's just like if you were buying bananas and if you, you know, you were buying **Chiquita** bananas for like 20 cents a pop, are you were buying like, I don't know, bananas made out of gold for \$100 a pop. And you had a family of 100 people. And you could only afford to buy three of the gold bananas and you spent all your money on those and you made some profit on them. What if I then you found out you could buy these **Chiquita** bananas and feed everyone in your family, um, with the same budget. And in fact, you could actually only spend, you know, to feed everyone spending 75% less. You would spend as much money as possible to make everyone in your family happy. And frankly, you might even buy more bananas because the incremental utility of buying those bananas for your for your family might be higher. So you would probably buy way more bananas in in overall spend. And those bananas would require way more, uh, fertilizer, way more water, and the entire ecosystem would have much more demand because that money that you're spending is not going to be captured by that high margin banana, the company that makes the high margin banana, it would be more spread out into the commodity banana space, which which would use more resources. And in this case, those resources are more compute. And it ends up translating into more cloud demand as well and the hyperscalers make more profit. Um, the reuse argument is in is also a Jevons's paradox argument, which is the same argument that if you have more output coming from engineers and the cost per engineer output falls with AI, you actually should hire more engineers, because the output per hour per engineer is much higher, which means the marginal cost of the engineer falls. So if AI makes engineers cheaper, instead of firing your engineers, you should just keep and hire more of the better engineers to get more output. And so yes, it absolutely is a Jevons's paradox argument.

Kimmy's Compute Sourcing - 02:09:35

Um, so going into Gerardo's, uh, question, which is compute sourcing for Kimmy. Is Kimmy's compute capacity coming from the Chinese hyperscalers like **Alibaba** or from US hyperscalers and cloud providers? Well, Kimmy's compute capacity for now is coming from **Alibaba**, which should be very good for **Alibaba's** demand. But on a global scale, right, if, one, Kimmy is not the only Chinese model. **Microsoft** is not using Kimmy, right? **Microsoft** is using repurposed Chinese models, um, and is basically, is basically recreating them in the US, um, and those, and is basically buying US compute to power the Chinese open source models. So while Kimmy's compute capacity is coming from Chinese hyperscalers like **Alibaba**, and it shouldn't mean, you know, more demand. But by the way, even those Chinese hyperscalers are buying **Nvidia** GPUs. Um, they're just buying them from Singapore. And that's why like, if you look at **Nvidia's** financial statements, they're selling all these GPUs to Singapore. Singapore doesn't have that many data centers. You know, those GPUs are going to to other places in Asia, probably China, and the government is just allowing it to happen. But anyway, um, yes, obviously, uh, they're buying from **Alibaba**, but the Chinese models in the US that **Microsoft** is using that are being repurposed, right, they are using US infrastructure. And so those are all the commodity models.

OpenAI is definitely losing market share, which is why I'm a little bit more worried about **Oracle**. Um, but these are all really good questions.

Outlook on Defense Stocks and Joby Aviation (JOBY) - 02:11:30

Thoughts on defense stocks like **RST**, **LMT**, and **ITA**? What do you think of **JOBY** at \$7 or below? I actually like the big defense companies more than **JOBY**. Now, I like **AVAV** more than **JOBY**. Um, just **Joby Aviation** itself, um, I know that it sold off a lot, like with all the Momo stocks, but of all the Momo stocks, you know, I don't know why you'd buy the **JOBY** chart because it's burning like basically three-quarters of a billion a year. It will basically run out of all its excess cash over the next three years. It's not forecast to generate positive EBITDA till like past 2030. So, you know, there are a lot of other companies in the growth space that I'd buy, I'd rather buy than **JOBY**.

Impact of Data Center Delays on AI Stocks - 02:12:22

Um, Yojo is asking if data centers aren't going to come on as expected, what's the impact of **Anthropic** AI stocks in general? Data centers are delayed. It's not that they're not going to come on as expected. It's just that the number of data centers that is going to be built, um, 70% of the data centers that were supposed to be built in 2027 have been delayed because of power, because of just simply not having the equipment they need in time. Um, so I think what that means is it might actually just extend the cycle. Instead of data center demand going absolutely parabolic and all these data centers being built in 2027 and then, you know, the market, you know, slowing down in 2028 because fewer data centers are built. What's actually happening is data centers are, you're probably going to see more data center, um, builds ramp in 2028 much more than they did in 2027 because of those delays. So the whole data center cycle is going to get pushed out, and eventually the over-ordering of memory, the over-ordering of GPUs, um, will start to become more prevalent. And, you know, there may have been over-ordering already and a lot of those GPUs might be put into place in 2028. But the pace of data center builds will likely be higher in 2028 than it is in 2027. It's just very difficult to time and to know how much memory did they over-order, because, you know, **Nvidia** has been trying to lock up all this capacity because the Vera Rubin racks require so much more memory. So, you know, **Morgan Stanley** is basically saying that they think memory is going to, prices are going to peak this year, but I think that they're wrong, um, simply because all these data centers aren't even going to be built. You know, if, you know, demand could actually be much higher going into 2028. Um, so again, I think what we've seen, I think there will be an infrastructure overbuild. I don't think we're there yet because of these capacity constraints. And, you know, are some of these companies cheap? Um, they've sold off a lot, so I think they're tradable from from a, uh, you know, from the recent sell-off. Um, optically they may look cheap, but earnings could also and margins could also be inflated. So I'm looking at this as a trading opportunity. I'm not looking at it as a buy and hold. Um, but if you see interest rates peak, if you see the war go away, if you see people not as worried about these Chinese models because they're running out of compute, um, it's a backdrop that does set up for a bounce in the AI names. Um, again, I'm not a tech expert, I'm not an AI expert. Uh, I prefer value and special situations and asymmetric type plays. But if you if I was a betting man and, you know, just the speed of the momentum decline, there should be some sort of a bounce, uh, and it doesn't look like the AI trade is over.

Closing Remarks: The World Cup and Market Outlook - 02:15:50

Especially if you believe that people are going to still continue to build the generative AI, the compute required for that is just much, much higher. Um, so my, it's my educated guess. Now again, this may not happen because the Iran war could escalate, interest rates could rise, and obviously these companies need to issue a lot of bonds, right, to fund their capex. So if interest rates rise, it could, you know, unravel the whole thing. But, you know, we're only a couple months now, I mean it's almost *August*. Midterm elections are in *November*, they're already going to start campaigning. I don't understand why, uh, this war should continue, um, much longer. But again, I've been, you know, we've been proven wrong. Uh, the war stops and then escalates again, stops, escalates again. I'm frankly really worried about what happens after midterm elections, after *November*, because at that point Trump will have really nothing to fear. Um, but in the meantime, I do think that the war should stop, there should be some sort of a bounce. Um, interest rates should peak and there should be some sort of a resurgence in risk assets. Um, but, um, we'll stop there. I hope you guys have enjoyed the call. It's been like, you know, a three-hour call. And, um, you know, kudos to Spain for winning the World Cup. Um, this is a very, very, um, surprising World Cup with, you know, how well, you know, um, the US did, which was great, how well the UK did, you know, UK beating Norway, that was a big upset. Um, and now Spain being Argentina. Sorry, Messi. Um, very, very exciting World Cup. I hope you guys were able to watch, were able to enjoy the fact that the United States hosted it. I think it really helped with foreign relations. You know, you saw the Japanese, you know, loving the barbecue in Texas and everyone getting along. Uh, it shows the world that, you know, governments are really the problem. It's not the people. You know, people all over the world, um, are great. Um, it's really the governments that are causing all this pain globally. Uh, but the World Cup this year was absolutely fantastic. Um, I hope you guys all saw the game and I hope you guys are, uh, ready for the trading week tomorrow.

Final Market Outlook - 02:24:20

Um, you know, it looks like we're, we're going to see some, some choppiness. Um, you know, rates are higher, oil is higher. You know, **Nasdaq** is slightly, uh, bouncing after the sell off last week, but, you know, frankly, who knows what will happen by the morning. I am a, you know, if the war does go away, I'm optimistic. I would be adding precious metals, I'd even be adding a little bit of crypto, I'd be adding AI names, I would be adding income names. I would just be adding in general. Um, I do have quite a lot of cash right now. Um, but let's see how it plays out. Um, there's no reason to rush. Um, patience is key, and we'll talk to you guys, uh, this week. Recording will be up shortly.

Date of Transcript: 2024-06-25

Filename: ssn_transcript_v6 - 2026-01-26 - @DA

Model (may not be accurate): Gemini

Prompt: ssn_transcript_v6 - 2026-01-26 - @DA

QA Check: Completed