US Stock Express

Daniel Yue

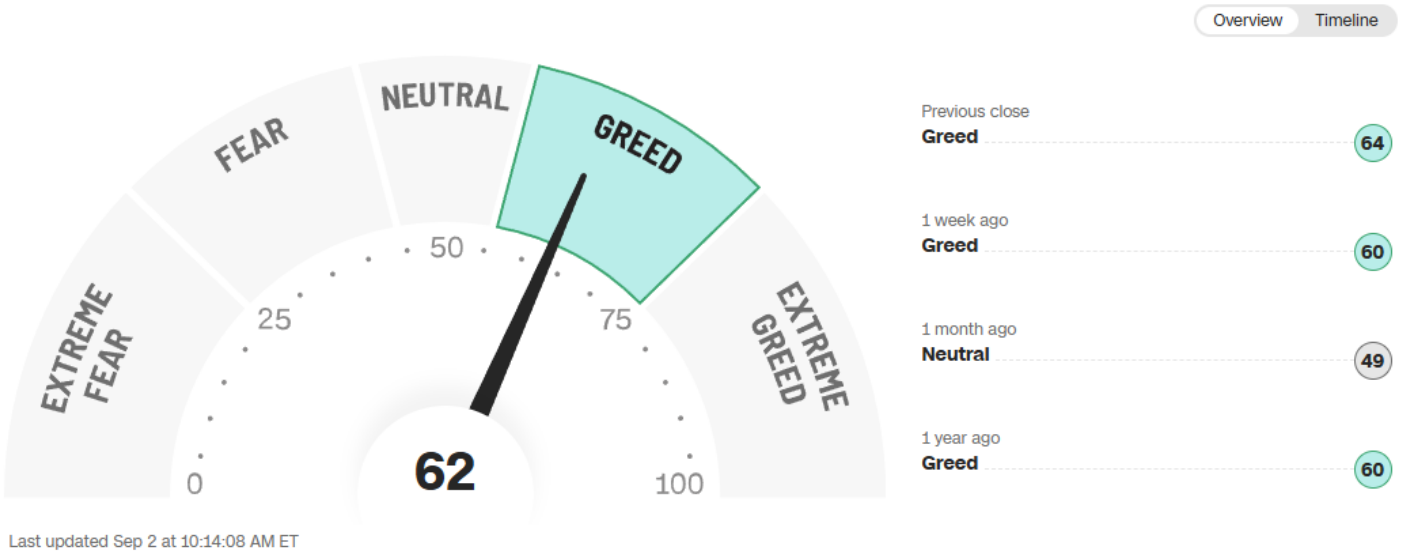
Email: info@ihandbook.org www.ihandbook.org

NVDA fell, BABA rose. Which would you choose?

Risk disclosure: Price can go up and down at any moment, use free money to trade and bear the risk according to your own capital;
Never trade with money that has a deadline for withdrawal.
All suggestions are for reference only, even AI cannot be 100% reliable, final decision still lies upon investors.
Copy trading cannot replicate another trader's background or psychological state.

Fear & Greed Index

What emotion is driving the market now?
Learn more about the index



Last updated Sep 2 at 10:14:08 AM ET

North East West South is *NEWS*

Rapidus, a Japanese startup foundry, recently announced the logic density of its 2HP 2nm process, which achieves 237.3 million transistors per square millimeter. This data is slightly higher than TSMC's N2 node's 236.2 million and significantly surpasses Intel's 18A node's approximately 184.2 million.

The European Defence Agency stated that EU military spending will reach a new high of €381 billion in 2025, with defense against Russia being the primary reason for the increased spending.

Tech tycoon Elon Musk expressed support for an anti-immigration protest in Japan. While the number of foreign-born residents in Japan is relatively low, political parties advocating "Japanese first" performed well in recent elections.

During his visit to Tokyo, Indian Prime Minister Narendra Modi stated that India and Japan will "shape the Asian century." Japanese Prime Minister Shigenori Ishiba pledged a 10 trillion yen investment in India.

Gold price went to $3509 record high in expectation of rate cut of the Fed.

🧠 AI Chip Comparison: Alibaba, Huawei, Cambricon vs Nvidia

Chip Model	Process Node	Transistor Count	Peak FLOPs (FP16)	DRAM / Memory	Key Differences vs NVIDIA H100
Alibaba Hanguang 800	~12nm (est.)	Not disclosed	~78 TFLOPs (inference)	On-chip SRAM (no external DRAM)	Optimized for inference; lacks training capability; used in Alibaba's data centers [1]
Cambricon MLU370	7nm	Not disclosed	~200 TFLOPs (INT8)	64GB HBM	Focused on inference; lower versatility than H100; used in edge and cloud AI [2]
Huawei Ascend 910B	7nm	~22 billion (est.)	~256 TFLOPs (FP16)	32GB HBM	Strong training chip; lacks CUDA-like ecosystem; limited global reach [3] [1]
Huawei Ascend 910C	6nm	Not disclosed	Claimed parity with H100	Likely 64GB HBM	Designed to rival H100; ecosystem still maturing; not widely benchmarked yet [3] [1]
NVIDIA H100	4nm (TSMC)	80 billion	1,979 TFLOPs (FP16)	80GB HBM3	Industry leader; unmatched software stack (CUDA, TensorRT); global adoption [3]

By Microsoft AI Copilot

🔍 Key Takeaways

- **Process Node**: Nvidia leads with a cutting-edge 4nm process, while Chinese chips range from 6nm to 12nm.
- **Transistor Count**: H100 packs 80 billion transistors — far ahead of competitors. Huawei's Ascend 910B is estimated at ~22 billion.
- **Performance (FLOPs)**: Nvidia's H100 delivers nearly 2 PFLOPs (FP16), dwarfing others. Huawei's 910C aims to match it, but real-world benchmarks are scarce.
- **Memory**: H100 uses 80GB HBM3; Huawei's chips use HBM2 or HBM2e. Alibaba's chip is inference-only and uses SRAM.
- **Ecosystem**: Nvidia's CUDA ecosystem is a major differentiator. Chinese chips are building their own stacks but face hurdles in developer adoption and global integration.

If Huawei's 910C truly rivals the H100 in performance, it would be a seismic shift — but Nvidia's software ecosystem and global reach still give it a major edge. Want to dive deeper into how these chips are being deployed in real-world AI models or data centers?

Alibaba said they are developing a new kind of semiconductor on their own and manufacture on their own without TSM, and they are going to replace NVDA. People outside China said they are still far away from replacing or even giving threats, but those inside China said they can stand on their own soon and no need to rely on TSM and NVDA. Who is correct?

I asked the Microsoft AI to make a comparison on the above items, how far that NVDA really is ahead of them? The above is just a comparison with H100, but they still need the H20, B30A and RTX6000D as said on The Express of Aug 22nd 2025.

For any questions investors can ask the Microsoft AI directly.

World Observation

Day 1288
Russia/Ukraine Conflict

Riding the AI Waves

(13) AMD

AMD started to fall since October 2024, and TSLA started to fall since December 2024, I told investors not to buy them during the fall. But I started to buy on the bottom of April 7th. TSLA will have Full Self Driving (FSD) version 14 come out in Sep, and version 15 will come out in Dec which means it is matured and in mid-2026, FSD will come to commercial use. So if there is a fall recently due to the news of developing of semiconductor of BABA replacing NVDA, its time to buy NVDA, TSLA and TSM.

The CEO of AMD is Lisa Su, or called Mother Su, sha is born in southern Taiwan. Jensen Huang of NVDA is her uncle, anyway of the same family tree even not so close. That is why NVDA is developing together with AMD. Gradually the island of Taiwan is becoming a myth of semiconductors. The question is that would AMD and NVDA rise together and fall together? Mind that the relation between these two companies is somewhat in a case of competitor, similar to that of zero sum game. That is in case of market share, NVDA is growing and eating the market share of AMD. So if they grow together, better buy NVDA than AMD.

The market has risen a lot since April 7th, and people are waiting for a correction before the Employment Report this Friday. The writer apologizes for stopped issuing The Express for several days due to personal problem on internet and desktop computer at the same time, but everything is solved now. Anyway, stocks purchase from March to August should locked up for 5 years and with effective from September, everything can start from ground zero.