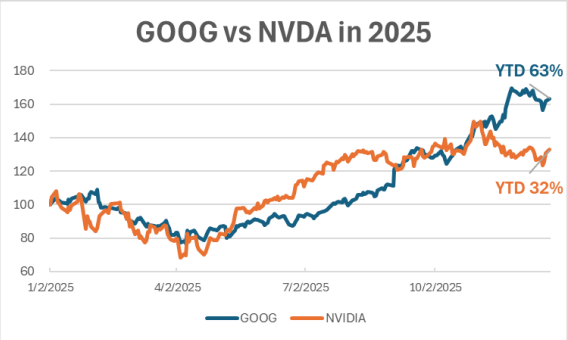
PeaQ News
Vol. 5 No. 22

GOOGLE v NVIDIA: THE AI COMPUTE BATTLE



- The explosive growth of generative AI has turned computing hardware into a strategic battleground. Nvidia's GPUs remain the industry standard, but Google's Tensor Processing Units (TPUs) are increasingly emerging as a credible alternative for large-scale cloud and inference workloads.

TPUs Move Beyond Google

- Originally developed for internal use, TPUs are now being commercialised more aggressively via Google Cloud:
  - AI-native firms such as Anthropic have committed to large TPU deployments.
  - Reports that Meta is evaluating TPUs highlight a growing desire among hyperscalers to reduce reliance on Nvidia.
  - This marks a shift from experimental use to meaningful external adoption.
- Beyond being a Google optimisation tool, TPUs are becoming a competitive offering in the AI infrastructure market.

The TPU advantage

- The key attraction of TPUs lies in economics and efficiency:
  - Strong performance-per-dollar, particularly for inference workloads.
  - Better power efficiency, increasingly important as AI data centre energy costs surge.
  - Lower total cost of ownership for large-scale, always-on AI systems.
- As AI usage scales, cost and energy efficiency are becoming decisive factors in hardware choice.

Software, the Next Battleground

- Google recognises that hardware alone is not enough to steal market share from NVIDIA.
- CUDA, NVIDIA's proprietary software, is the de-facto standard for programming GPUs. Nowadays, PyTorch is the dominant AI framework used by AI developers and it integrates seamlessly with CUDA.
- PyTorch is not well integrated with Google's software stack. Thus, in order to compete better with NVIDIA, Google is actively working to improve TPU compatibility with PyTorch.
- Broader software support would significantly lower adoption barriers.
- The competition is increasingly shifting from silicon performance to developer experience.

Market Implications

- AI infrastructure is likely to move toward hybrid compute stacks, mixing GPUs, TPUs and other accelerators.
- Nvidia may face incremental pricing pressure, particularly in cloud inference.
- Google's TPU push positions it as both a cloud provider and as a strategic AI infrastructure player.

Conclusion

- NVIDIA's GPUs remain central to today's AI ecosystem, but as AI deployment scales, economics and energy efficiency are becoming increasingly decisive. This shift favours Google, whose TPUs are emerging as a credible, cost-effective alternative for large-scale cloud and inference workloads. As software support improves and adoption broadens, TPUs are positioned to capture a growing share of AI infrastructure spend.
- PeaQ recommends an overweight position in Alphabet to capture the upside from this structural shift toward more cost- and energy-efficient AI compute.

22nd December 2025

For more information about our services, please contact:

André Chung Shui
Managing Director
+230 52512216
ac@peaqadvisors.com

Ishrat Cheeroo
Senior Investment Manager
+230 55012900
ic@peaqadvisors.com

Keshav Maghoo
Investment Analyst
+230 4687900
km@peaqadvisors.com

Nashini Ramye
Portfolio Analyst
+230 4687900
nr@peaqadvisors.com

Source: FED, ECB, BOE, BoM, Bloomberg, SEMDEX, Statistics Mauritius, MSCI, CNBC, Reuters