

Semiconductors

Google Cloud Next: TPU v8 and rising of AI Hypercomputer

On Day 1 of Google Cloud Next 2026, held on April 22 in Las Vegas, delivered three key infrastructure updates: 1) Launched two 8th gen TPUs (TPU 8t and TPU 8i) with dedicated designs for training and inference, respectively, highlighting a clearer split between compute-intensive training and latency- and memory-sensitive inference workloads; 2) introduced a broader AI Hypercomputer architecture built around Virgo networking, tighter storage integration, and large-scale cluster orchestration, reinforcing that AI performance is increasingly defined at the system level rather than by standalone chip capability; 3) framed agentic AI as the next major demand driver, which raises the importance of memory bandwidth, interconnect scale, utilization, and power efficiency alongside raw silicon performance.

In our view, these updates reinforce a more system-led AI spending cycle. As bottlenecks shift from chip availability toward data movement, latency, and energy efficiency, the beneficiary set should broaden beyond accelerators to include memory suppliers, optical interconnect vendors, power infrastructure players, and storage-related participants. We expect Google-linked supply chain names to continue gaining traction. **Maintain BUY on Innolight (300308 CH, BUY) and InnoScience (2577 HK, BUY), both of which we view as key beneficiaries of Google's AI infrastructure expansion.**

- **For GPU vendors**, the message is not weaker demand, but a higher bar for differentiation. Google's continued support for Nvidia confirms leading GPUs remain central to frontier AI deployments, but future competition will depend less on compute power alone and more on the ability to deliver a complete compute fabric, including scale-up, scale-out, memory hierarchy, software, etc. TPU 8t's focus on large-scale training and TPU 8i's optimization for inference and agentic AI reinforce that vendors are now competing at the ecosystem level, not just the chip level.
- **For memory suppliers**, Google's architecture suggests memory is becoming a more central constraint in AI system design. TPU 8i was built to address the memory wall, with 288GB of HBM and 384MB of on-chip SRAM (vs. Groq LPU's 500 MB), and Google explicitly notes that SRAM capacity was sized around the KV cache needs of reasoning models. This implies the next phase of AI demand will depend increasingly on memory capacity, bandwidth, and lower-latency access, especially for long-context inference and multi-agent reasoning. That should remain supportive for HBM suppliers and suggests inference infrastructure may be more memory-intensive than previously assumed.
- **For optical interconnect**, Google's TPU v8 architecture reinforces that network scaling is becoming a primary system bottleneck. TPU 8t doubles interchip bandwidth vs. the prior gen and relies on Virgo to support large-scale training clusters, while TPU 8i adopts Boardfly topology to improve communication efficiency in reasoning and MoE workloads. Google also highlighted OCS as part of its system design, underscoring that optical is increasingly being used not just for bandwidth expansion but also for resiliency and hardware-level traffic optimization. With Virgo linking over 134k chips at very high non-blocking bandwidth, we see this as supportive of a structurally stronger outlook for optical interconnect.
- **Our broader takeaway**: Google is validating a more distributed AI value chain. The first leg of AI capex was accelerator-led, while the next looks more system-driven. For supply chain participants, that means the AI trade is widening, with incremental value increasingly accruing to those enabling system efficiency rather than just raw compute.

OUTPERFORM
(Maintain)

China Semiconductors Sector

Kevin ZHANG
(852) 3761 8727
kevinzhang@cmbi.com.hk

Aaron GUO
(852) 3916 3715
aaronguo@cmbi.com.hk

Figure 1: TPU 8t ASIC block diagram

| Feature | TPU 8t | TPU 8i |
|---------------------------|--|---------------------------------------|
| Primary workload | Large-scale pre-training | Sampling, serving, and reasoning |
| Network Topology | 3D torus | Boardfly |
| Specialized chip features | SparseCore (Embeddings) & LLM Decoder Engine | CAE (Collectives Acceleration Engine) |
| HBM Capacity | 216 GB | 288 GB |
| On-Chip SRAM (Vmem) | 128 MB | 384 MB |
| Peak FP4 PFLOPS | 12.6 | 10.1 |
| HBM Bandwidth | 6,528 GB/s | 8,601 GB/s (~1.3x of TPU 8t) |
| CPU Header | Arm Axion | Arm Axion |

Source: Company data, CMBIGM

Figure 2: TPU 8t rack

| Architecture | Ironwood 2025 | TPU 8t 2026 |
|--|---------------|-------------|
| Pod size | 9,216 | 9,600 |
| FP4 EFlops/pod | 42.5 | 121 |
| Bidirectional scale-up bandwidth (Tb/s per chip) | 9.6 | 19.2 |
| Scale-out networking bandwidth (Gb/s per chip) | 100 | 400 |

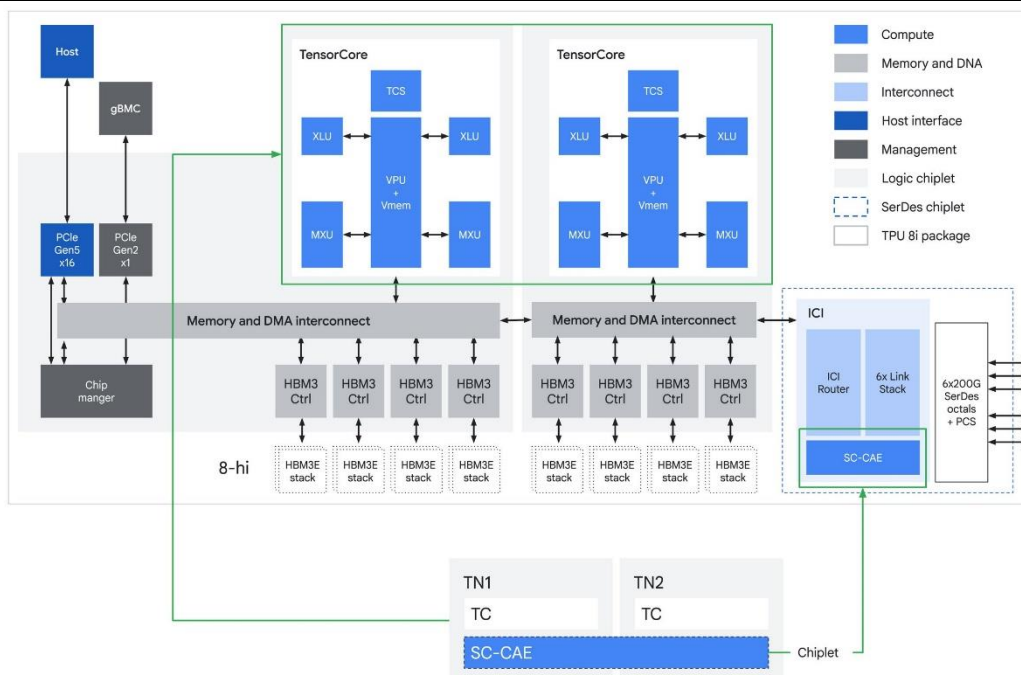
Source: Company data, CMBIGM

Figure 3: TPU 8i rack

| Architecture | Ironwood 2025 | TPU 8i 2026 |
|--|---------------|-------------|
| Pod size | 256 | 1,152 |
| FP8 EFlops/pod | 1.2 | 11.6 |
| Total HBM capacity/pod (TB) | 49.2 | 331.8 |
| Bidirectional scale-up bandwidth (TB/s per chip) | 9.6 | 19.2 |

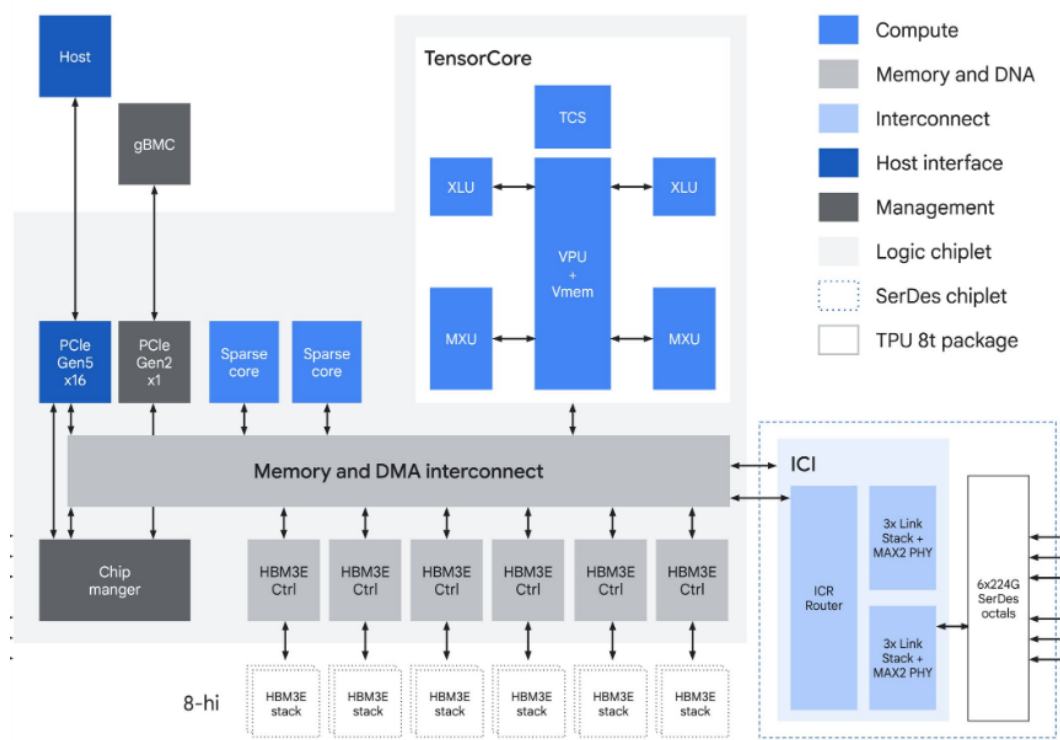
Source: Company data, CMBIGM

Figure 4: TPU 8i ASIC block diagram



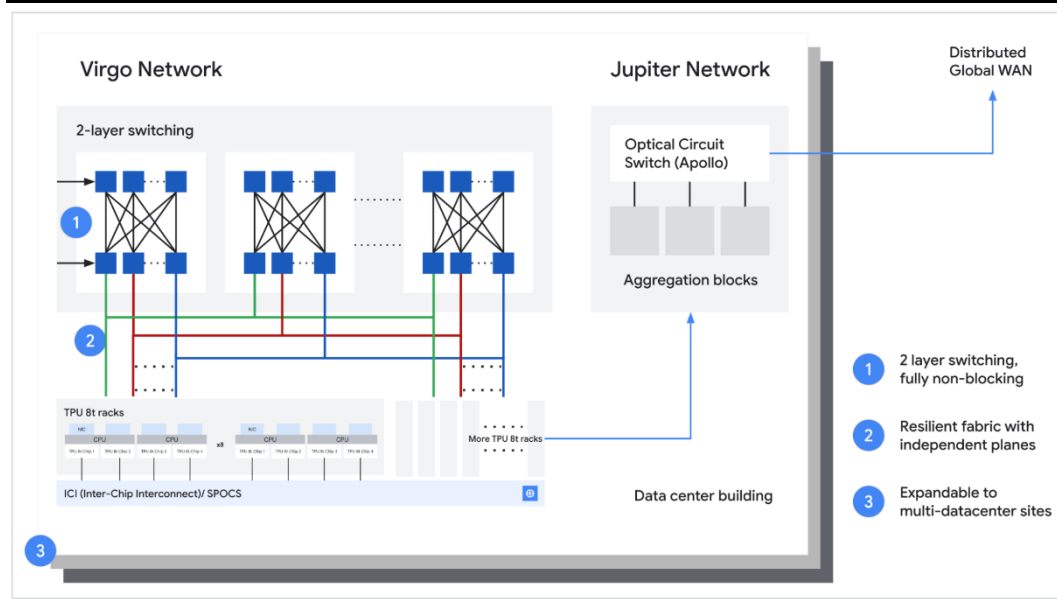
Source: Company data, CMBIGM

Figure 5: TPU 8t ASIC block diagram



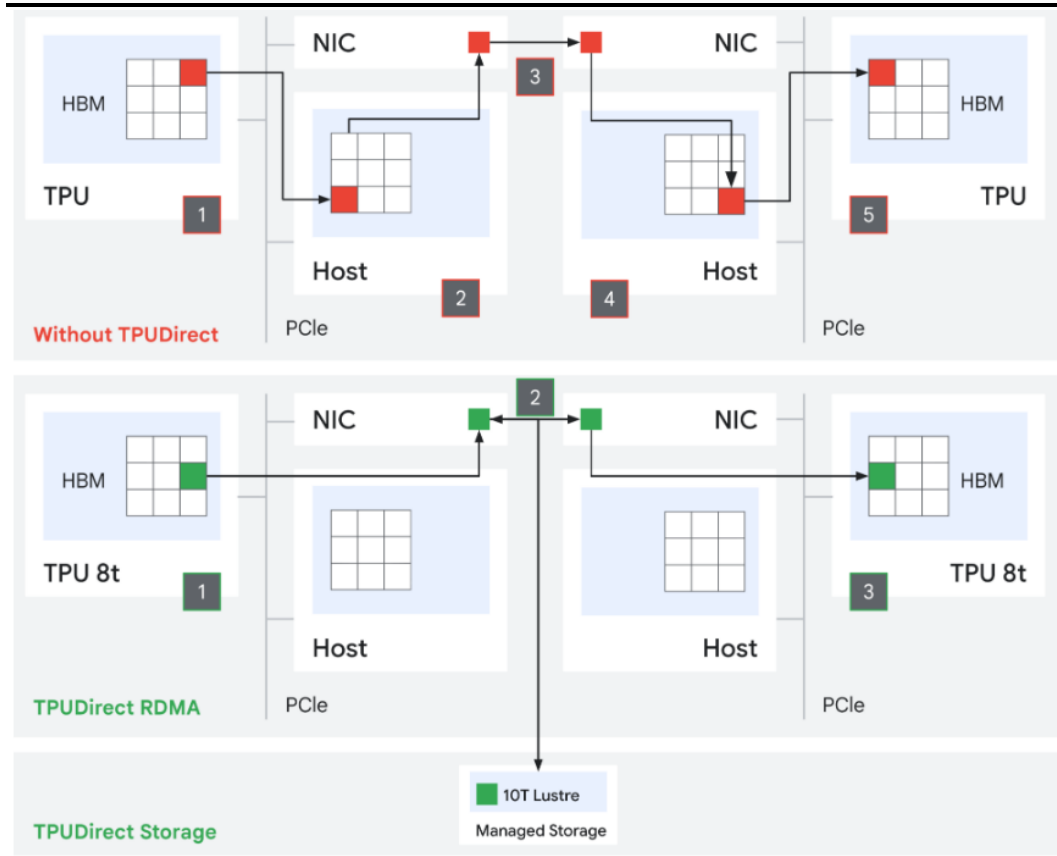
Source: Company data, CMBIGM

Figure 6: TPU 8t rack level connectivity to Virgo fabric



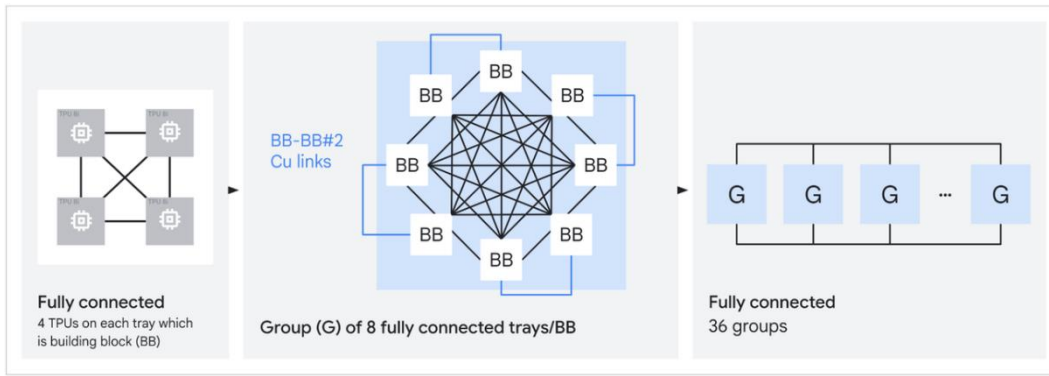
Source: Company data, CMBIGM

Figure 7: TPUDirect RDMA & Storage



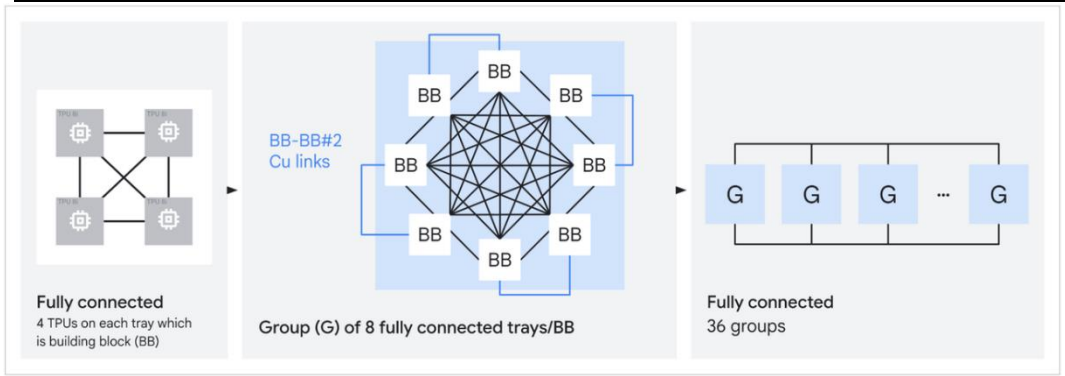
Source: Company data, CMBIGM

Figure 8: TPU 8i hierarchical Boardfly topology



Source: Company data, CMBIGM

Figure 9: Maximum seven-hop ICI network diameter via OCS on TPU 8i pod



Source: Company data, CMBIGM

Disclosures & Disclaimers

Analyst Certification

The research analyst who is primary responsible for the content of this research report, in whole or in part, certifies that with respect to the securities or issuer that the analyst covered in this report: (1) all of the views expressed accurately reflect his or her personal views about the subject securities or issuer; and (2) no part of his or her compensation was, is, or will be, directly or indirectly, related to the specific views expressed by that analyst in this report.

Besides, the analyst confirms that neither the analyst nor his/her associates (as defined in the code of conduct issued by The Hong Kong Securities and Futures Commission) (1) have dealt in or traded in the stock(s) covered in this research report within 30 calendar days prior to the date of issue of this report; (2) will deal in or trade in the stock(s) covered in this research report 3 business days after the date of issue of this report; (3) serve as an officer of any of the Hong Kong listed companies covered in this report; and (4) have any financial interests in the Hong Kong listed companies covered in this report.

CMBIGM Ratings

BUY : Stock with potential return of over 15% over next 12 months
HOLD : Stock with potential return of +15% to -10% over next 12 months
SELL : Stock with potential loss of over 10% over next 12 months
NOT RATED : Stock is not rated by CMBIGM

OUTPERFORM : Industry expected to outperform the relevant broad market benchmark over next 12 months
MARKET-PERFORM : Industry expected to perform in-line with the relevant broad market benchmark over next 12 months
UNDERPERFORM : Industry expected to underperform the relevant broad market benchmark over next 12 months

CMB International Global Markets Limited

Address: 45/F, Champion Tower, 3 Garden Road, Hong Kong, Tel: (852) 3900 0888 Fax: (852) 3900 0800

CMB International Global Markets Limited ("CMBIGM") is a wholly owned subsidiary of CMB International Capital Corporation Limited (a wholly owned subsidiary of China Merchants Bank)

Important Disclosures

There are risks involved in transacting in any securities. The information contained in this report may not be suitable for the purposes of all investors. CMBIGM does not provide individually tailored investment advice. This report has been prepared without regard to the individual investment objectives, financial position or special requirements. Past performance has no indication of future performance, and actual events may differ materially from that which is contained in the report. The value of, and returns from, any investments are uncertain and are not guaranteed and may fluctuate as a result of their dependence on the performance of underlying assets or other variable market factors. CMBIGM recommends that investors should independently evaluate particular investments and strategies, and encourages investors to consult with a professional financial advisor in order to make their own investment decisions.

This report or any information contained herein, have been prepared by the CMBIGM, solely for the purpose of supplying information to the clients of CMBIGM or its affiliate(s) to whom it is distributed. This report is not and should not be construed as an offer or solicitation to buy or sell any security or any interest in securities or enter into any transaction. Neither CMBIGM nor any of its affiliates, shareholders, agents, consultants, directors, officers or employees shall be liable for any loss, damage or expense whatsoever, whether direct or consequential, incurred in relying on the information contained in this report. Anyone making use of the information contained in this report does so entirely at their own risk.

The information and contents contained in this report are based on the analyses and interpretations of information believed to be publicly available and reliable. CMBIGM has exerted every effort in its capacity to ensure, but not to guarantee, their accuracy, completeness, timeliness or correctness. CMBIGM provides the information, advices and forecasts on an "AS IS" basis. The information and contents are subject to change without notice. CMBIGM may issue other publications having information and/ or conclusions different from this report. These publications reflect different assumption, point-of-view and analytical methods when compiling. CMBIGM may make investment decisions or take proprietary positions that are inconsistent with the recommendations or views in this report.

CMBIGM may have a position, make markets or act as principal or engage in transactions in securities of companies referred to in this report for itself and/or on behalf of its clients from time to time. Investors should assume that CMBIGM does or seeks to have investment banking or other business relationships with the companies in this report. As a result, recipients should be aware that CMBIGM may have a conflict of interest that could affect the objectivity of this report and CMBIGM will not assume any responsibility in respect thereof. This report is for the use of intended recipients only and this publication, may not be reproduced, reprinted, sold, redistributed or published in whole or in part for any purpose without prior written consent of CMBIGM.

Additional information on recommended securities is available upon request.

For recipients of this document in the United Kingdom

This report has been provided only to persons (I) falling within Article 19(5) of the Financial Services and Markets Act 2000 (Financial Promotion) Order 2005 (as amended from time to time) ("The Order") or (II) are persons falling within Article 49(2) (a) to (d) ("High Net Worth Companies, Unincorporated Associations, etc.") of the Order, and may not be provided to any other person without the prior written consent of CMBIGM.

For recipients of this document in the United States

CMBIGM is not a registered broker-dealer in the United States. As a result, CMBIGM is not subject to U.S. rules regarding the preparation of research reports and the independence of research analysts. The research analyst who is primary responsible for the content of this research report is not registered or qualified as a research analyst with the Financial Industry Regulatory Authority ("FINRA"). The analyst is not subject to applicable restrictions under FINRA Rules intended to ensure that the analyst is not affected by potential conflicts of interest that could bear upon the reliability of the research report. This report is intended for distribution in the United States solely to "major US institutional investors", as defined in Rule 15a-6 under the US, Securities Exchange Act of 1934, as amended, and may not be furnished to any other person in the United States. Each major US institutional investor that receives a copy of this report by its acceptance hereof represents and agrees that it shall not distribute or provide this report to any other person. Any U.S. recipient of this report wishing to effect any transaction to buy or sell securities based on the information provided in this report should do so only through a U.S.-registered broker-dealer.

For recipients of this document in Singapore

This report is distributed in Singapore by CMBI (Singapore) Pte. Limited (CMBISG) (Company Regn. No. 201731928D), an Exempt Financial Adviser as defined in the Financial Advisers Act (Cap. 110) of Singapore and regulated by the Monetary Authority of Singapore. CMBISG may distribute reports produced by its respective foreign entities, affiliates or other foreign research houses pursuant to an arrangement under Regulation 32C of the Financial Advisers Regulations. Where the report is distributed in Singapore to a person who is not an Accredited Investor, Expert Investor or an Institutional Investor, as defined in the Securities and Futures Act (Cap. 289) of Singapore, CMBISG accepts legal responsibility for the contents of the report to such persons only to the extent required by law. Singapore recipients should contact CMBISG at +65 6350 4400 for matters arising from, or in connection with the report.