

What optical module does Groq use



Overview

Groq, Inc. is an American (AI) company that builds an (ASIC). The architecture was originally introduced as a Tensor Streaming Processor (TSP) but was later rebranded as a Language Processing Unit (LPU) following the widespread adoption of large language models after the breakthrough of. The company also develops related and software to accelerate AI inference performance.

Key Takeaways

Groq leverages high-speed optical modules, including 400G, 800G, and emerging 1.6T solutions, to enable low-latency, high-bandwidth connectivity in AI and data center environments.

Optical Modules Overview

Optical modules are critical components for high-speed data transmission in data centers, telecommunications, and AI clusters. They convert electrical signals into optical signals and vice versa, enabling low-latency, high-throughput interconnects. The market for optical modules is rapidly expanding, driven by hyperscale data centers, 5G/6G deployment, and AI workloads, with projections reaching USD 42 billion by 2033 for general optics and USD 13.

36 billion for 1.6T modules by 2033. Key types include pluggable modules, embedded modules, and on-board modules, with pluggable modules currently dominating adoption.



High-

Speed Modules for AI and Data CentersThe transition from 400G to 800G and 1.6T optical modules is accelerating, particularly in hyperscale AI clusters. 800G modules are experiencing 60% year-over-year shipment growth, while 1.

6T modules are entering initial production, supporting the next generation of AI workloads. These modules are essential for co-packaged optics (CPO), which integrate optical transceivers directly with processors, reducing latency and increasing bandwidth density—an approach highly relevant for Groq's AI accelerators. **Groq's Likely Use of Optical Modules**Groq's AI accelerators are designed for high-performance, low-latency computation, which requires dense, high-bandwidth interconnects.

Optical modules, particularly CPO and linear pluggable modules, enable Groq systems to scale efficiently across multiple nodes in AI clusters. By leveraging 400G and 800G modules, Groq can support massive data throughput for training and inference workloads, while emerging 1.6T modules will further enhance interconnect performance in future deployments.

Market and Technology Trends

Silicon photonics and indium phosphide (InP) components are driving higher data rates and lower power consumption per bit, critical for AI-scale deployments. AI-driven demand is creating a new growth segment for optical modules, with 30-35% annual growth projected in 2025-2026 for AI cluster connectivity. Geopolitical and supply chain factors influence module availability and pricing, affecting deployment strategies for companies like Groq.

ConclusionGroq's high-performance AI systems rely on advanced optical modules to achieve the bandwidth and low-latency connectivity required for modern AI workloads. The adoption of 400G, 800G, and emerging 1.6T modules, along with co-packaged optics, positions Groq to efficiently scale AI clusters while keeping power consumption and latency optimized, aligning with broader trends in hyperscale data center and AI infrastructure growth.

Article Content

Groq, and the Hardware of AI – Intuitively and Exhaustively Explained

Before we talk about Groq, we'll break down what AI fundamentally is, and explore some of the key components of computer hardware used to run AI models. Namely; CPUs, GPUs, and TPUs.

What Is the Groq Chip and How It Works: LPU Architecture Explained

The Groq chip is not a GPU — it is an entirely new class of processor called a Language Processing Unit (LPU), designed from first principles for one specific task: running large language

You Need to Know About Groq

Groq is the maker of the Language Processing Unit (LPU), a chip optimized for fast AI inference. The Groq LPU is a single-core processor designed for LLMs, interconnected with a fast

Nvidia builds out LPU chip team following \$20bn Groq acquihire ...

Groq has yet to use co-packaged optics for its LPUs, but Nvidia has integrated CPO into its networking switches. This role is likely to ensure the LPU works with Nvidia's broader CPO switch

Groq LPU: not a replacement for GPU+HBM, but a low-latency

At GTC 2026, Groq/LPU will likely be presented as an LPX-style pluggable module within a broader inference system stack. It addresses the productization gap in real-time inference without

AI Inference Accelerator | NVIDIA Groq 3 LPX

NVIDIA Groq 3 LPX is the inference accelerator for NVIDIA Vera Rubin, designed to meet the low-latency and large-context demands of agentic systems. Vera Rubin and LPX unite the extreme

Groq is fast, low cost inference.

Groq has created immense savings and reduced so much overhead for us. We've been able to keep costs for our main offerings incredibly low, helping keep our

Can Groq Really Take On Nvidia?

The Groq LPU is a single core processor designed for LLMs, interconnected with a fast switchless routing fabric using 288 QSFP28 optical cables.

A closer look at Nvidia's Groq-powered LPX rack systems

We're told the chip is based on Groq's second-gen LPU tech with a handful of last-minute tweaks made just before tapping out at Samsung's fabs. The chip doesn't use Nvidia's proprietary

NVIDIA Groq 3 LPX: Everything we know

NVIDIA's Groq 3 LPX pairs 256 LP30 LPUs with Vera Rubin NVL72, delivering up to 35x higher inference throughput per megawatt.

Groq, and the Hardware of AI – Intuitively and

This article discusses Groq, a new approach to computer hardware that's revolutionizing the way AI is applied to real world problems. Before we talk

Infrastructure for Scalable AI Reasoning | NVIDIA Vera Rubin Platform

NVIDIA Vera Rubin powers agentic AI and reasoning models at scale, eliminating bottlenecks in communication, coordination, and memory for efficient multi-step workflows.

Deep Dive 4: Groq (Private) – The Technologist

What does Groq do? Groq designs and manufactures specialized, high-speed AI inference chips that act like turbo-charged calculators for robots, enabling artificial intelligence

Inside the LPU: Deconstructing Groq's Speed

Discover how Groq's Language Processing Units (LPUs) achieve breakthrough AI inference speeds with 4 key architectural innovations: SRAM-centric design for instant weight

A Capella: A Unique Compilation | PDF | Wellness

This document contains a long list of words beginning with "ab-" or "ac-". It does not appear to be a coherent text, but rather a collection of unrelated terms.

Why is Everybody Talking about Groq?

Plus, at that scale, networking all of those chips starts adding bottlenecks -- currently, Groq's optical interconnect only scales up to 264 chips, not 10,000. The thing that still perplexes me

What is Groq? The AI chip maker explained (2026) | eesel AI

A deep dive into Groq, the company behind the ultra-fast LPU inference chips. Learn how Groq's technology works, its benefits, limitations, and what it means for AI.

NVIDIA Groq 3 LPU Explained: How the Non-GPU

NVIDIA's Groq 3 LPU delivers 35x more inference throughput per megawatt than GPUs with 150 TB/s SRAM bandwidth. Full architecture

Contact Us

Continue your research online:

Website: <https://ensaa.es>

Technical inquiry: <https://ensaa.es/contact.html>

This document is for preliminary research. Verify specifications against current standards, product documentation and project requirements.

