

Cerebras Systems Analyst Quote Sheet

Karl Freund, Principal Analyst, Cambrian AI Research

"Cerebras is the company [whose] architecture is skating to where the puck is going: huge AI."

"One of the largest challenges of using large clusters to solve AI problems is the complexity and time required to set up, configure and then optimize them for a specific neural network... With Weight Streaming, Cerebras is removing all the complexity we have to face today around building and efficiently using enormous clusters – moving the industry forward in what I think will be a transformational journey."

Mike Orme, Consultant Analyst, Thematic Intelligence team, GlobalData

"Saving time is crucial in competitive AI model training and inferencing. By doing everything on a single piece of silicon within a closely knit mesh of processors, fast memories, and compilers, Cerebras' CS-2 shows up the core weakness of the Nvidia approach of using lots of different types of separate chips that are working in parallel to do the same job."

The CS-2 is a formidable feat in silicon engineering. CS-2 technology is a speed fiend in comparison to a GPU cluster, and offers a very big bang per watt too. There are several other well-funded contenders with wafer scale offerings, but Cerebras is the clear pack leader for now."

Dan Olds, Chief Research Officer, Intersect360 Research

"Cerebras' ability to bring large language models to the masses with cost-efficient, easy access opens up an exciting new era in AI. It gives organizations that can't spend tens of millions an easy and inexpensive on-ramp to major league NLP. It will be interesting to see the new applications and discoveries CS-2 customers make as they train GPT-3 and GPT-J class models on massive datasets."

Mike Orme, Consultant Analyst, Thematic Intelligence team, GlobalData

"Nvidia faces mounting competition from left field as a new breed of lavishly funded specialist AI companies such as Cerebras, Graphcore, and SambaNova design AI chips from scratch, rather than adapting GPU technology."

"GSK, TotalEnergies and Argonne Labs, for example, have used Cerebras Wafer Scale Engines to power heavy-duty, real-world applications."

Linley Gwennap, President and Principal Analyst, The Linley Group

"The wafer-scale approach is unique and clearly better for big models than much smaller GPUs. By coupling the WSE with the new MemoryX and SwarmX technology, Cerebras has created what should be the industry's best solution for training very large neural networks."

Contact Us

info@cerebras.net
www.cerebras.net/press-kit