Important Disclosures

Independent Third-Party Research Report

Important Disclosure

This research report was prepared by an independent third-party provider and is made available by StartEngine Crowdfunding, Inc. ("StartEngine") for informational purposes only. StartEngine and its affiliates, including StartEngine Primary, LLC, a registered broker-dealer, have not independently verified the information contained herein and do not guarantee its accuracy or completeness.

The views expressed in this report are those of the independent provider and do not necessarily reflect the views of StartEngine, its affiliates, or any of their employees.

This report does not constitute an offer to sell or a solicitation of an offer to buy any securities, nor does it constitute a recommendation to buy, sell, or hold any securities. Please carefully review the risks, conflicts of interest, and disclosures provided in the **Important Disclosures** section located at the end of this document.

Important Disclosures

Independent Third-Party Research

Important Disclosures

This research report was prepared by an independent third-party research provider without input, review, or influence by StartEngine Crowdfunding, Inc. ("StartEngine") or its affiliates and is distributed from StartEngine for informational purposes only. StartEngine has not independently verified the information contained herein and does not guarantee its accuracy or completeness. This report reflects the views of the independent provider and does not necessarily represent the views of StartEngine or its affiliates.

StartEngine does not endorse or make any representations regarding the conclusions or recommendations contained in this report. This report does not constitute a recommendation or solicitation by StartEngine to buy, sell, or hold any securities.

Investing in private companies involves a high degree of risk, including the loss of your entire investment. Past performance does not guarantee future returns.

This report is made available to all individuals on startengine.com and may be distributed to individuals who have expressed interest in specific companies or industries.

Please carefully review the company-specific risks, conflicts of interests, and disclosures below.

Company Specific Disclosures:

Conflict of Interest: StartEngine or its affiliates sell membership interests of:

- Series 17-1, a Series of StartEngine Private LLC which owns Cerebras Systems, Inc. Series C-1 Common Shares;
- Series 17-2, a Series of StartEngine Private LLC which owns Cerebras Systems, Inc. Series C-1 Common Shares;
- Series 17-3, a Series of StartEngine Private LLC which owns Cerebras Systems, Inc. Series C-1 Common Shares;
- Series 17-4, a Series of StartEngine Private LLC which owns Cerebras Systems, Inc. Series C-1 Common Shares; and
- Series 2-1, a Series of StartEngine Private Funds LLC which owns Cerebras Systems, Inc. Series C-1 Common Shares

Security Ownership: StartEngine or its affiliates owns Common Shares of Cerebras Systems, Inc.

For further questions regarding this report or its distribution, please contact StartEngine at contact@startengine.com representative.

SACRA

EQUITY RESEARCH

UPDATED
10/01/2025

Cerebras

TEAM

Jan-Erik Asplund
Co-Founder
jan@sacra.com

Marcelo Ballve
Head of Research
marcelo@sacra.com

DISCLAIMERS

This report is for information purposes only and is not to be used or considered as an offer or the solicitation of an offer to sell or to buy or subscribe for securities or other financial instruments. Nothing in this report constitutes investment, legal, accounting or tax advice or a representation that any investment or strategy is suitable or appropriate to your individual circumstances or otherwise constitutes a personal trade recommendation to you.

This research report has been prepared solely by Sacra and should not be considered a product of any person or entity that makes such report available, if any.

Information and opinions presented in the sections of the report were obtained or derived from sources Sacra believes are reliable, but Sacra makes no representation as to their accuracy or completeness. Past performance should not be taken as an indication or guarantee of future performance, and no representation or warranty, express or implied, is made regarding future performance. Information, opinions and estimates contained in this report reflect a determination at its original date of publication by Sacra and are subject to change without notice.

Sacra accepts no liability for loss arising from the use of the material presented in this report, except that this exclusion of liability does not apply to the extent that liability arises under specific statutes or regulations applicable to Sacra. Sacra may have issued, and may in the future issue, other reports that are inconsistent with, and reach different conclusions from, the information presented in this report. Those reports reflect different assumptions, views and analytical methods of the analysts who prepared them and Sacra is under no obligation to ensure that such other reports are brought to the attention of any recipient of this report.

All rights reserved. All material presented in this report, unless specifically indicated otherwise is under copyright to Sacra. Sacra reserves any and all intellectual property rights in the report. All trademarks, service marks and logos used in this report are trademarks or service marks or registered trademarks or service marks of Sacra. Any modification, copying, displaying, distributing, transmitting, publishing, licensing, creating derivative works from, or selling any report is strictly prohibited. None of the material, nor its content, nor any copy of it, may be altered in any way, transmitted to, copied or distributed to any other party, without the prior express written permission of Sacra. Any unauthorized duplication, redistribution or disclosure of this report will result in prosecution.

Cerebras

Visit Website

AI chip hardware for training large-scale models in scientific and enterprise use cases

#ai #ai-chips

Details

HEADQUARTERS

Sunnyvale, CA

CEO

None

SACRA

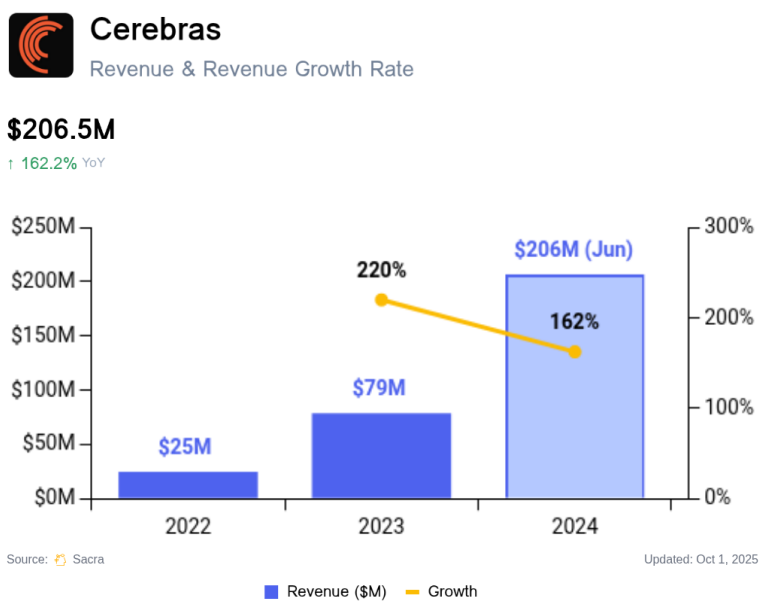
REVENUE
$206,480,000
2024

VALUATION
$8,100,000,000
2025

GROWTH RATE (Y/Y)
162%
2024

FUNDING
$1,910,000,000
2025

Revenue



Sacra estimates Cerebras will reach $272M in revenue for 2024, based on their H1'24 revenue of $136.4M, representing approximately 245% year-over-year growth.

The company maintains significant customer concentration with their primary enterprise client G42, which accounts for 87% of revenue, stemming from a strategic partnership that includes a commitment to purchase $1.43B of computing systems and services.

Cerebras' revenue model centers on their hardware and software offerings: sales of their wafer-scale chip systems, their CSoft software platform, and associated services. Their AI supercomputer systems command premium pricing due to superior performance claims, including 10x faster training time-to-solution and over 10x faster output generation speeds compared to GPU-based alternatives.

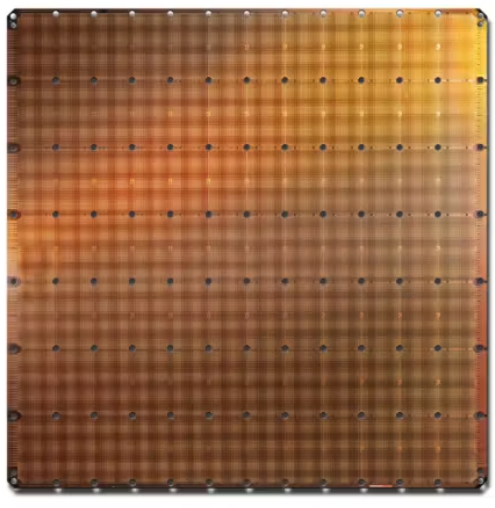
The company's gross margins have shown improvement from 11.7% in 2022 to 33.5% in 2023, though they experienced some compression in early 2024 (41.1%) due to volume-based discounts offered to G42.

Valuation

In September 2025, Cerebras raised a $1.1B Series G co-led by Fidelity and Atreides at an $8.1B post-money valuation. The company continues to plan to go public after CFIUS-related delays postponed its planned 2025 IPO.

The company has now raised about $1.91 in total funding across multiple rounds. Key investors include Benchmark and Foundation Capital.

Product



Cerebras WSE
1.2 Trillion transistors
46,225 mm² silicon

Largest GPU
21.1 Billion transistors
815 mm² silicon

Founded in 2016, Cerebras Systems is an AI hardware manufacturer and computing services provider, specializing in wafer-scale technology for artificial intelligence workloads. The company's flagship product is the CS-3 system, powered by the Wafer Scale Engine 3 (WSE-3). The WSE-3 is the largest and fastest chip ever created, measuring 46,000 square millimeters (about the size of a dinner plate) and built on TSMC's 5nm process.

Traditional AI chips, like those from NVIDIA or Groq, are typically the size of a postage stamp and are clustered together to achieve high performance. In contrast, the WSE-3's design offers a few critical advantages for AI workloads:

Data Locality: By keeping all computations on a single, large chip, Cerebras dramatically reduces the need for data movement. In traditional GPU clusters, data must constantly move between chips, consuming time and energy. The WSE-3 keeps data local, resulting in significantly faster processing and lower power consumption.

Memory Bandwidth: The WSE-3 has vast on-chip memory with unprecedented bandwidth. This allows for faster data access and reduces bottlenecks commonly experienced in distributed GPU systems.

Simplified Scaling: While GPU clusters require complex software to distribute workloads across multiple chips, Cerebras's single-chip approach simplifies this process. This makes it easier to scale up to larger AI models without the overhead of managing distributed computing resources.

These advantages make Cerebras systems particularly well-suited for training large AI models and handling complex, data-intensive workloads. For instance, training a 175 billion parameter model on 4,000 GPUs might require 20,000 lines of code for distribution, while Cerebras can accomplish this with just 565 lines of code in one day.

Product-market fit

Overall, Cerebras positions itself for customers who have outgrown smaller-scale AI solutions and need supercomputer-level performance. Their systems are designed for organizations training large models or working with vast datasets, typically those spending millions on AI compute annually. In addition to hardware, Cerebras offers professional services, including assistance with data preparation, model design, training oversight, and optimization. This service component accounts for between a quarter and a third of new customer engagements.

The company has found product-market fit across several sectors.

In the government and research space, Cerebras has run the table on supercomputer labs, including Argonne National Laboratory, Lawrence Livermore National Laboratory, and Sandia National Laboratory, among others. Governments worldwide are also investing in Cerebras technology for sovereign clouds to meet domestic AI requirements.

In the private sector, Cerebras serves customers in pharmaceuticals and life sciences, such as GlaxoSmithKline, which uses the technology for pioneering work in epigenomics, while Mayo Clinic is using Cerebras systems to enhance medical diagnostics and personalized medicine.

In the energy sector, TotalEnergies, a $100 billion French company, is a notable customer. While not specifically named, Cerebras also works with companies in the finance space on complex modeling and risk assessment.

Business Model

Cerebras Systems has a few separate revenue streams:

1. Hardware Sales: The primary revenue source comes from selling CS-3 systems, powered by the Wafer Scale Engine 3 (WSE-3). These systems are priced in the millions of dollars, targeting high-end customers with substantial AI computing needs.

2. Professional Services: Accounting for 25-33% of new customer engagements, this includes data preparation, model design, training oversight, and optimization services.

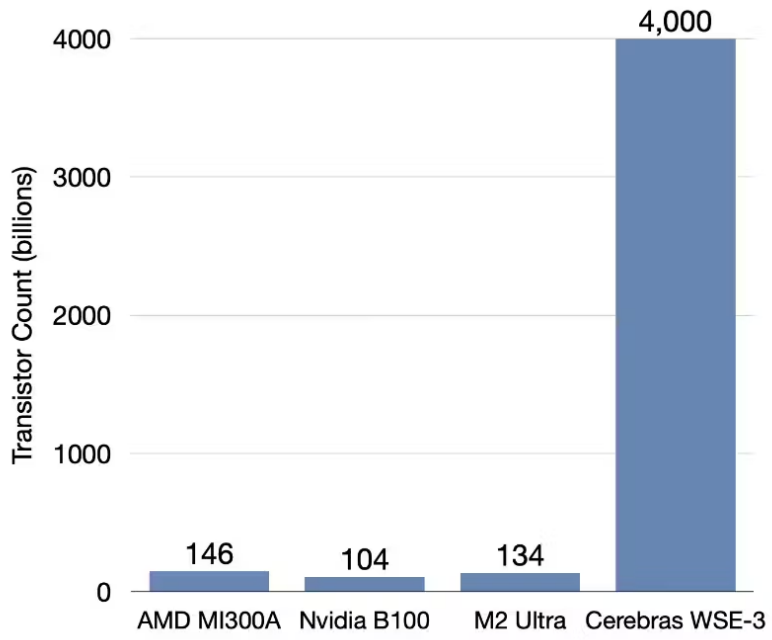
3. Cloud Services: Cerebras offers cloud-based access to its systems through partnerships and its AI Model Studio, providing a recurring revenue stream.

4. Subscription Models: The company is exploring subscription-based offerings, placing hardware on customer sites under a license agreement.

Hardware sales likely constitute the majority of revenue, followed by professional services. Cloud and subscription services are likely growing segments.

While gross margin hasn't been publicly reported, AI chipmaker AMD reported gross margins of 47% last quaarter.

Competition



AI Accelerator Hardware

The AI accelerator hardware market is dominated by NVIDIA's GPUs, which hold approximately 85% market share. NVIDIA's H100 GPU is the current industry standard for AI training and inference, offering high performance and widespread software support through CUDA.

AMD and Intel are NVIDIA's primary competitors in this space, with AMD's MI300 series and Intel's Gaudi2 chips aiming to capture market share.

The key characteristic that Cerebras is differentiating itself around in this category is simplicity of scaling.

While GPU-based systems require complex networking and software to distribute workloads across multiple chips, Cerebras' single-chip approach eliminates much of this complexity. This can potentially reduce the time and expertise required to set up and manage large-scale AI training infrastructure.

Cloud AI Services and Specialized AI Chips

Major cloud providers like Google, Amazon, and Microsoft are developing their own specialized AI chips to reduce dependence on NVIDIA and offer differentiated AI services. Google's Tensor Processing Units (TPUs), Amazon's Trainium and Inferentia, and Microsoft's Azure Maia AI Accelerator represent significant investments in this area.

Cerebras competes in this category by offering its technology as a cloud service, allowing customers to access WSE-powered systems without the need for on-premises hardware. The company has partnered with Cirrascale Cloud Services to make its CS-1 and CS-2 systems available in the cloud.

This approach allows Cerebras to target customers who need high-performance AI training capabilities but may not have the resources or desire to manage their own hardware.

Cerebras' advantage here lies in its ability to offer a unique hardware architecture that isn't available from major cloud providers. The company claims its systems can handle larger models and provide faster training times than cloud-based GPU clusters, potentially attracting customers with demanding AI workloads.

AI Chip Startups

A number of well-funded startups are developing novel AI chip architectures, including Graphcore, SambaNova Systems, and Groq.

These companies are exploring various approaches to AI acceleration, such as Intelligence Processing Units (IPUs) and tensor streaming processors.

In this category, Cerebras stands out for its wafer-scale approach and focus on very large language models. While other startups are developing chips that compete more directly with traditional GPUs, Cerebras has carved out a niche for training massive AI models that may be impractical on other hardware.

Cerebras' recent partnership with Qualcomm for AI inference represents a strategic move to address the full AI workflow.

By optimizing its training output for Qualcomm's Cloud AI 100 inference processor, Cerebras aims to offer a more complete solution that spans from model development to deployment. This collaboration could give Cerebras an edge over other AI chip startups that focus primarily on training or inference.

TAM Expansion

The AI hardware market is projected to reach nearly $250B by 2030, and companies are showing willingness to spend hundreds of millions or even billions on supercomputer-scale AI clusters.

With its unique wafer-scale technology providing substantial speedups over alternatives, Cerebras's upside hinges on becoming the architecture of choice for the high end of the market: multi-billion dollar AI supercomputers used by hyperscalers, governments, and the largest enterprises.

In addition to this core tailwind from the growth of the AI market, Cerebras has several opportunities to expand its overall TAM.

Expanding from Training to Inference

While Cerebras initially focused on training, where its huge memory capacity and bandwidth shine, the company is expanding to also accelerate inference. This opens up the broader market of deploying trained models into production to generate text, images, and other AI-powered experiences.

As generative AI takes off, demand for cost-effective, high-performance inference will surge. Cerebras' wafer-scale engine could become a popular choice for deploying the largest models where real-time responsiveness matters, such as robotics, autonomous systems, and interactive applications. The inference market is projected to be even larger than the training market long-term.

Verticals

Cerebras has initially focused on the high-end AI research and development market, with some expansion into healthcare and life sciences, but there is substantial potential for expansion into other verticals that require massive computational power:

1. Financial Services: High-frequency trading, risk modeling, and fraud detection systems could benefit from the WSE's computational capabilities.

2. Climate Modeling: The company's systems could enhance climate simulations, improving our understanding of global weather patterns and climate change.

3. National Security and Defense: Government agencies could leverage Cerebras' technology for complex simulations, cryptography, and intelligence analysis.

Risks

Software compatibility: The AI ecosystem has standardized around Nvidia's CUDA platform. Cerebras' bespoke software stack may struggle to gain traction with developers accustomed to CUDA's mature tooling and libraries. Without a strong software moat, Cerebras risks being a marginal player.

Long-term business model: In the near-term, Cerebras can grow by selling high-margin hardware to enterprises for whom AI is mission-critical. But in the longer-term, it's unclear if being a specialized chip vendor is the best position in the value chain. The industry may re-centralize around a few hyperscale AI platforms who view chips as a cost center. Cerebras may need to forward-integrate into cloud services to capture sufficient value.

Niche market fit: Cerebras is building extremely large, expensive systems that only make sense for the most compute-intensive AI workloads. This limits their total addressable market to a small number of hyperscalers and deep-pocketed enterprises. If AI model growth slows or shifts towards efficiency over scale, Cerebras may be overly specialized for a shrinking niche.

Funding Rounds

Series G

Share Name	Issue Price	Issued At
Series G	$36.23	Sep 2025

Series F

Share Name	Issue Price	Issued At
Series F	$27.7448	Jun 2024
Series F-1	$14.66	Jun 2024

Series E

Share Name	Issue Price	Issued At
Series E	$18.3249	Nov 2019

Series D

Share Name	Issue Price	Issued At
Series D	$16.1458	Nov 2018

Series C

Share Name	Issue Price	Issued At
Series C	$8.9474	Aug 2017

Series B

Share Name	Issue Price	Issued At
Series B	$2.754493	Dec 2016

Series A

Share Name	Issue Price	Issued At
Series A	$0.85	May 2016

Figures sourced from the latest Certificate of Incorporation we have available.

DISCLAIMERS

This report is for information purposes only and is not to be used or considered as an offer or the solicitation of an offer to sell or to buy or subscribe for securities or other financial instruments. Nothing in this report constitutes investment, legal, accounting or tax advice or a representation that any investment or strategy is suitable or appropriate to your individual circumstances or otherwise constitutes a personal trade recommendation to you.

This research report has been prepared solely by Sacra and should not be considered a product of any person or entity that makes such report available, if any.

Information and opinions presented in the sections of the report were obtained or derived from sources Sacra believes are reliable, but Sacra makes no representation as to their accuracy or completeness. Past performance should not be taken as an indication or guarantee of future performance, and no representation or warranty, express or implied, is made regarding future performance. Information, opinions and estimates contained in this report reflect a determination at its original date of publication by Sacra and are subject to change without notice.

Sacra accepts no liability for loss arising from the use of the material presented in this report, except that this exclusion of liability does not apply to the extent that liability arises under specific statutes or regulations applicable to Sacra. Sacra may have issued, and may in the future issue, other reports that are inconsistent with, and reach different conclusions from, the information presented in this report. Those reports reflect different assumptions, views and analytical methods of the analysts who prepared them and Sacra is under no obligation to ensure that such other reports are brought to the attention of any recipient of this report.

All rights reserved. All material presented in this report, unless specifically indicated otherwise is under copyright to Sacra. Sacra reserves any and all intellectual property rights in the report. All trademarks, service marks and logos used in this report are trademarks or service marks or registered trademarks or service marks of Sacra. Any modification, copying, displaying, distributing, transmitting, publishing, licensing, creating derivative works from, or selling any report is strictly prohibited. None of the material, nor its content, nor any copy of it, may be altered in any way, transmitted to, copied or distributed to any other party, without the prior express written permission of Sacra. Any unauthorized duplication, redistribution or disclosure of this report will result in prosecution.

Published on Oct 01st, 2025