

MAINCODE BUILDS AN AI FACTORY FOR AUSTRALIA WITH AMD

CASE STUDY

Maincode's upcoming \$30M MC-2 AI factory uses AMD Instinct™ MI355X GPUs to scale sovereign, cost-efficient AI systems for Australian enterprises



Maincode is an Australian AI research lab and product company. The team builds AI models and systems for banks, insurers, and other organizations that need tight control over data, risk, and cost. It focuses on AI that customers can run in their own environments, under their own rules, with economics that work when pilots give way to real workloads.

“We know we can pick up the phone, speak to someone at AMD, and be confident they’ll deliver.”

Dave Lemphers, CEO & Founder, Maincode

To support this vision, Maincode is investing thirty million dollars in MC-2, a next-generation AI factory built on AMD Instinct™ MI355X GPUs. MC-2 will give Maincode a central hub for training and hosting AI, operating as an AI factory rather than a traditional compute cluster. Customers will then run their AI systems on their own hardware, often on premises behind their firewalls, using MCX, Maincode's upcoming operating system for AI that will link to MC-2.

EVOLVING FROM PUBLIC CHAT TO GOVERNED AI SYSTEMS

Most Maincode customers begin with simple chat experiments using public LLM APIs. The trouble comes when they try to plug those into the governed workflows on their own infrastructure. Regulated workloads require deterministic, auditable behavior, but, as CEO Dave Lemphers puts it, “when you rely on a generative model to classify real documents, you get different responses each time,” turning variability into significant risk for a bank or insurer.

“In almost any scenario, AMD lets us afford as many as twice the GPUs.”

Fabian Waschkowski, Research Scientist, Maincode

Maincode's answer is to keep the language interface flexible while making the backend predictable. Models handle natural-language input, then process it against strict business rules running on MC-2. Lemphers says, “MCX is designed to be an operating system for AI. MCX will package Maincode's research outputs into a governed distribution that customers can run under their own compliance rules. It will be a mini version of our cluster, from AMD Instinct GPUs up through a custom software stack, that lets companies run AI safely and with full compliance.”

INDUSTRY

Data Center

CHALLENGES

Banks and insurers need compliant AI they could customize and run on their own infrastructure, not public LLM APIs that could produce variable answers and governance risk

SOLUTION

Using AMD Instinct™ MI355X GPUs, Maincode is building its MC-2 AI factory—and MCX, an upcoming AI operating system that enables customers to deploy custom governed models on their own infrastructure

RESULTS

MC-2 on AMD Instinct MI355X GPUs delivers stable performance, up to 2x more usable FLOPs per dollar vs their prior comparable NVIDIA platform, and positions Maincode as a leader in Australian AI

AMD TECHNOLOGY AT A GLANCE

AMD Instinct™ MI355X GPUs



Maincode is building its AI on AMD Instinct™ GPUs to help Australian customers run AI in-house under their own compliance rules

TIGHTER MODELS, FASTER RESULTS, BETTER ECONOMICS WITH AMD INSTINCT GPUs

“With smaller models, you control how many parameters go in and what the architecture looks like,” researcher Lukas Wesemann says. “They become really cost-efficient and orders of magnitude faster at the specific thing they are designed to do.”

For such workloads, memory capacity and bandwidth matter as much as raw FLOPs. Higher HBM capacity enables architectures that would otherwise require complex tensor parallelism. Built on 4th Gen AMD CDNA™ architecture, AMD Instinct MI355X GPUs offer up to 288 GB of HBM3E and 8 TB per second of on-package bandwidth. “With AMD Instinct GPUs, we can push much more onto a single GPU and rely less on parallelism strategies,” researcher Fabian Waschkowski says. Focused small models combined with high-capacity accelerators provide Maincode a simpler training path, stronger performance per watt, and a tighter match between model design and hardware.

“With AMD Instinct GPUs, we can push much more onto a single GPU.”

Fabian Waschkowski, Research Scientist, Maincode

That strategy also impacts how Maincode evaluates infrastructure. The team looks at cost per token, throughput per dollar, and usable FLOPs for a set budget. This system-level evaluation differs from traditional HPC benchmarking. They ran what Lemphers calls an exploded modeling exercise across accelerators, nodes and rack design, power, cooling, networking, and support. “It is all about how many FLOPs we can get for a given amount of money,” Waschkowski says.

Lemphers adds, “MC-2 would cost nearly twice as much all-in had we gone with a comparable NVIDIA-based solution.” Or as Waschkowski puts it, “If the whole package gives you a 2x advantage in usable FLOPs for the same money, then the real question is whether you want a certain number of GPUs or roughly twice as many for the same budget. In almost any scenario, AMD lets us afford as many as twice the GPUs.”

LOOKING FOR A PARTNER, NOT JUST A HARDWARE SUPPLIER

Early hardware experiences did not match the depth or pace Maincode needed. “We found a lot of legacy HPC thinking and enterprise optimization,” Lemphers says. “We need technical depth and agility, not just another GPU configuration.” HPC approaches are built for batch compute rather than AI factories that require hardware, software, and model co-design. “Working with NVIDIA was very commercial,” he continues. “It was basically just get in line, collect your box. There was no strategy for us and no real strategic relationship there.”

“When we engaged with AMD, what they brought to the table quickly felt like co-design and co-partnership.”

Dave Lemphers, CEO & Founder, Maincode

“When we engaged with AMD, what they brought to the table quickly felt like co-design and co-partnership,” Lemphers says. AMD matched the engineering speed and depth Maincode required for frontier AI research. The AMD team worked with Maincode to port its stack to the AMD Accelerator Cloud, validate performance on AMD Instinct™ MI300X GPUs, and think through what MC-2 would later look like on AMD Instinct™ MI355X.



Maincode’s AI factory runs on AMD to deliver more throughput, density, and efficiency per rack.



Maincode quickly ported its code using AMD ROCm™ and validated its confidence in the trajectory of AMD.

PUTTING AMD INSTINCT AND AMD ROCM TO THE TEST

Before committing MC-2 to AMD Instinct™ MI355X GPUs, Maincode wanted to prove the platform end to end. “We initially trained MC1 models on our NVIDIA cluster, and then we ported our training infrastructure and libraries over to AMD Accelerator Cloud,” Waschkowski says. Moving to AMD was a strategic choice for architectural flexibility and future scale. “AMD ROCm libraries let PyTorch target AMD Instinct GPUs,” he adds, “so everything basically worked out of the box.” AMD ROCm provides an open ecosystem that supports custom kernels and experimental architectures.

Lemphers sees AMD ROCm and the broader software ecosystem as equally strategic. “We saw the ground AMD made up in a very short time relative to its competition,” he says/. “For us, the rate of change and velocity AMD demonstrates matter more in the long term than the current feature surface area.”

“The rate of change and velocity AMD demonstrates matter more in the long term.”

Dave Lemphers, CEO & Founder, Maincode

“Training over weeks at a time, we saw that AMD GPUs were really stable and performing well, which made bringing in a cluster of AMD Instinct™ MI355X GPUs an easy decision,” Waschkowski says.

For Lemphers, reliability includes people as much as hardware. “When something goes wrong, hardware won’t save you. We know we can pick up the phone, speak to someone at AMD, and be confident they’ll deliver. That’s the biggest part. From the earliest planning conversations, AMD demonstrated complete commitment to delivering MC-2 on time.”

TCO AND SCALING MC-2 ON AMD INSTINCT GPUS

“Selecting your GPU is just a tiny part of the process,” Lemphers says. “You must consider what kind of chassis it is going into, power, cooling, and density. There is so much more to the story.”

Those economics matter even more when they look ahead. “It ended up being a 1.5 to 1.8 times TCO advantage delivering MC-2 on AMD. That’s considerable when you think about scaling to an MC-2 in every city,” he says. This cost profile will enable MC-2 style deployments across Australian regions. AMD Instinct™ MI355X GPUs deliver high GPU density, strong power efficiency, and a liquid-cooled design that help Maincode fit more accelerators into each rack and do more work within the same space and power envelope. This maps directly to Maincode’s focus on FLOPs per dollar and throughput per rack over any single benchmark score.

“It ended up being a 1.5 to 1.8 times TCO advantage delivering MC-2 on AMD”

Dave Lemphers, CEO & Founder, Maincode

BUILDING AN AUSTRALIAN AI FUTURE WITH AMD

MC-2 and MCX will contribute to an Australian-built AI capability for regulated sectors. “At the end of the day, customers need us to be ahead of them,” Lemphers says. “We can do that because of our technical relationship with AMD.”

That relationship has positioned Maincode as a leading AI research partner for AMD in Australia. Because MCX is built on AMD Instinct GPUs and the AMD ROCm software stack, Maincode has a solid, open foundation for the long term rather than a potential proprietary dead end.



Maincode is building a sovereign AI future for Australia with AMD.

Advanced security features, including hardware-based protections and support for secure multi-tenant GPU sharing, align with the need for strong isolation and compliance in finance, government, and other regulated sectors.

“AMD treats us as a first-class citizen in a way we did not experience with any other partner.”

Dave Lemphers, CEO & Founder, Maincode

For Maincode, MC-2 and MCX are also a statement about Australia’s AI future. “We are in the very early stages of what we believe is going to be a massive economic and industrial evolution in AI.

In AMD we found a partner who believes in Australian-made AI as strongly as we did,” Lemphers says. “AMD treats us as a first-class citizen in a way we did not experience with any other partner.” He concludes, “If you are a deep science research team, peeling back the layers and getting deep in it, you quickly find out AMD is the better solution.”



EXPLORE [AMD INSTINCT GPU](#)

Sign up to receive [AMD data center communications](#).

ABOUT MAINCODE

Maincode is an applied AI research lab and product company based in Melbourne, Australia. The team designs, trains, and operates custom, task-specific AI models for enterprises that need control over data, risk, and performance. Maincode provides end-to-end “AI model manufacturing” – from data engineering and evaluation to deployment on Australian infrastructure – and is building MC-2, a new AI factory, to power products such as its Matilda family of Australian-made language models. Maincode integrates data engineering, model design, long cycle training, and deployment governance into one research and production loop. For more information visit maincode.com.

ABOUT AMD

For more than 50 years AMD has driven innovation in high-performance computing, graphics, and visualization technologies. Billions of people, leading Fortune 500 businesses, and cutting-edge scientific research institutions around the world rely on AMD technology daily to improve how they live, work and play. AMD employees are focused on building leadership high-performance and adaptive products that push the boundaries of what is possible. For more information about how AMD is enabling today and inspiring tomorrow, visit the AMD (NASDAQ: AMD) [website](#), [blog](#), [LinkedIn](#), and [X](#) pages.

DISCLAIMERS

All performance and cost savings claims are provided by Maincode and have not been independently verified by AMD. Performance and cost benefits are impacted by a variety of variables. Results herein are specific to Maincode and may not be typical. GD-181

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. The information contained herein is subject to change and may be rendered inaccurate for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product releases, product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. Any computer system has risks of security vulnerabilities that cannot be completely prevented or mitigated. AMD assumes no obligation to update or otherwise correct or revise this information. However, AMD reserves the right to revise this information and to make changes from time to time to the content hereof without obligation of AMD to notify any person of such revisions or changes. GD-18.

COPYRIGHT NOTICE

©2026 Advanced Micro Devices, Inc. All rights reserved. AMD, the AMD Arrow logo, AMD CDNA, AMD Instinct, Infinity Fabric, ROCm, and combinations thereof are trademarks of Advanced Micro Devices, Inc. Other product names contained herein are for identification purposes only and may be trademarks of their respective owners. Certain AMD technologies may require third-party enablement or activation. Supported features may vary by operating system. Please confirm with the system manufacturer for specific features. No technology or product can be completely secure.