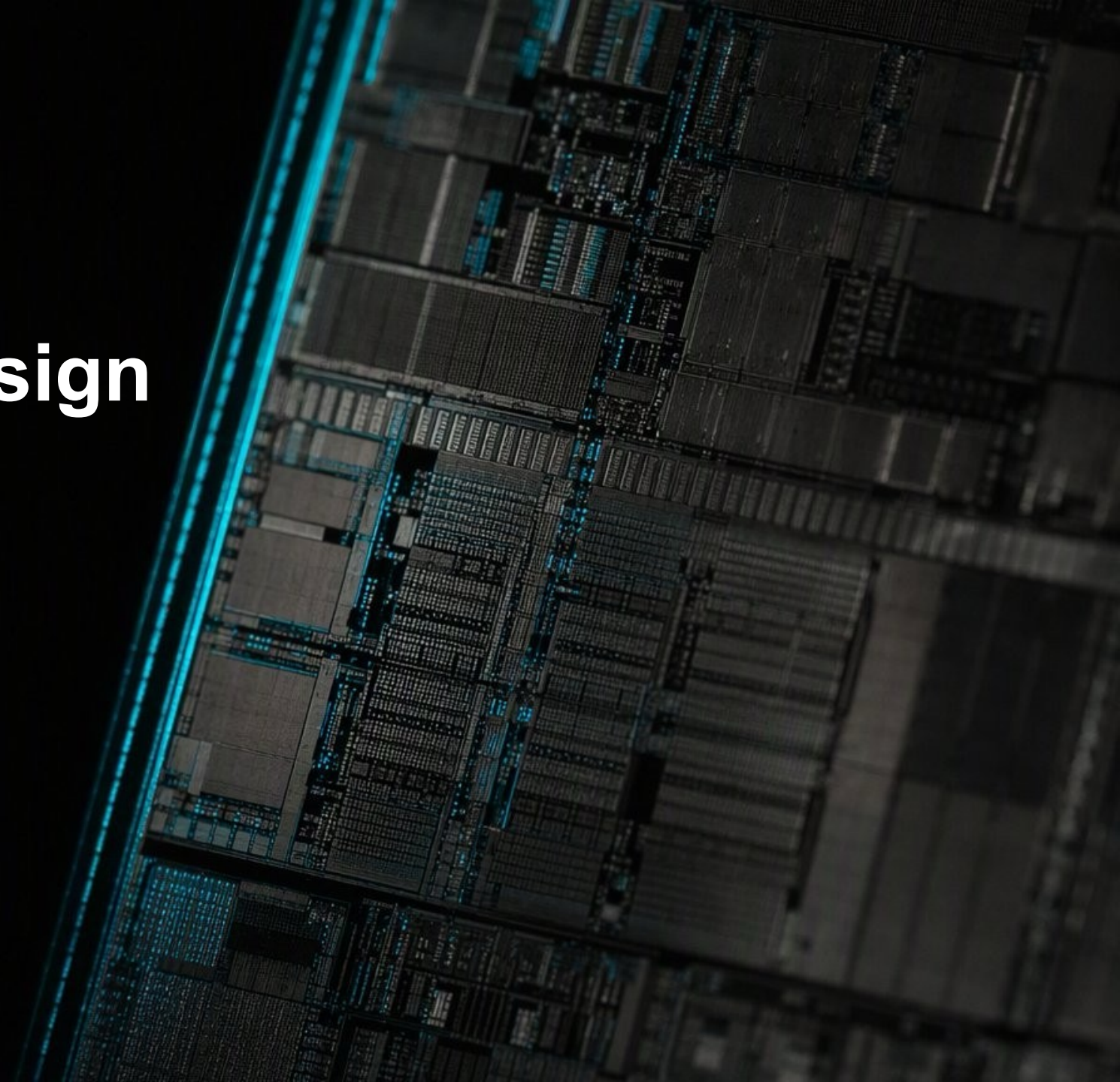
AI-Driven Chip Design and the Future of Computing

Ralph Wittig
Corporate Fellow, AMD
November 9, 2026

ICCAD 2026 · KEYNOTE



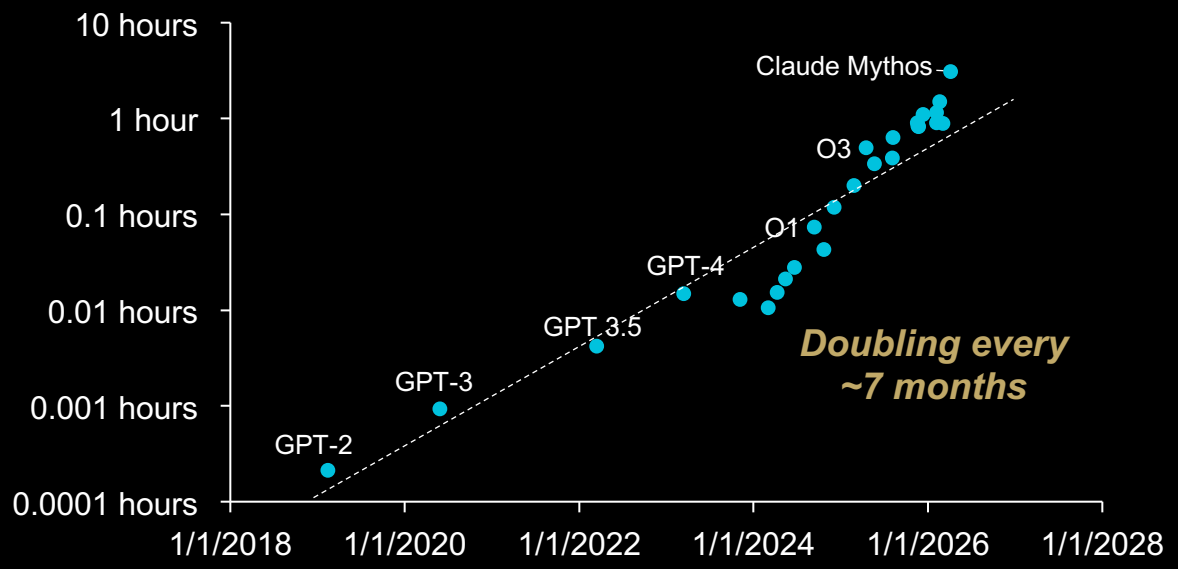
Ongoing Trends in AI

More Work Becomes Possible

Frontier Capabilities doubling every 7 months

Length of task completed with 80% Success

Source: METR

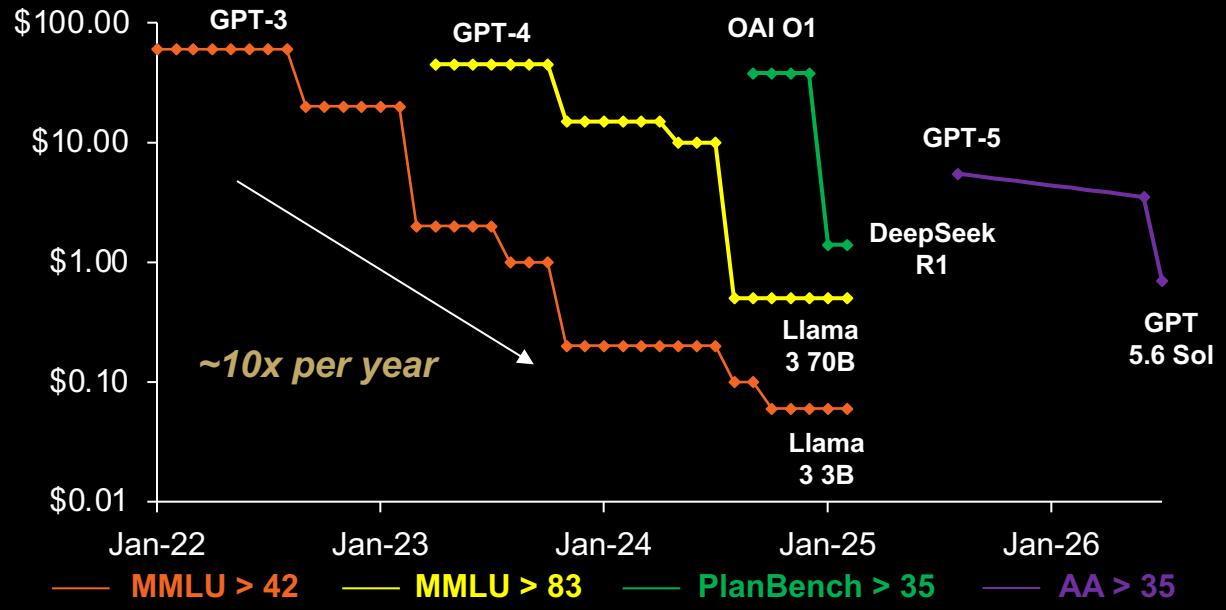


More Work Becomes Economical

Cost of Intelligence dropping ~10x every year

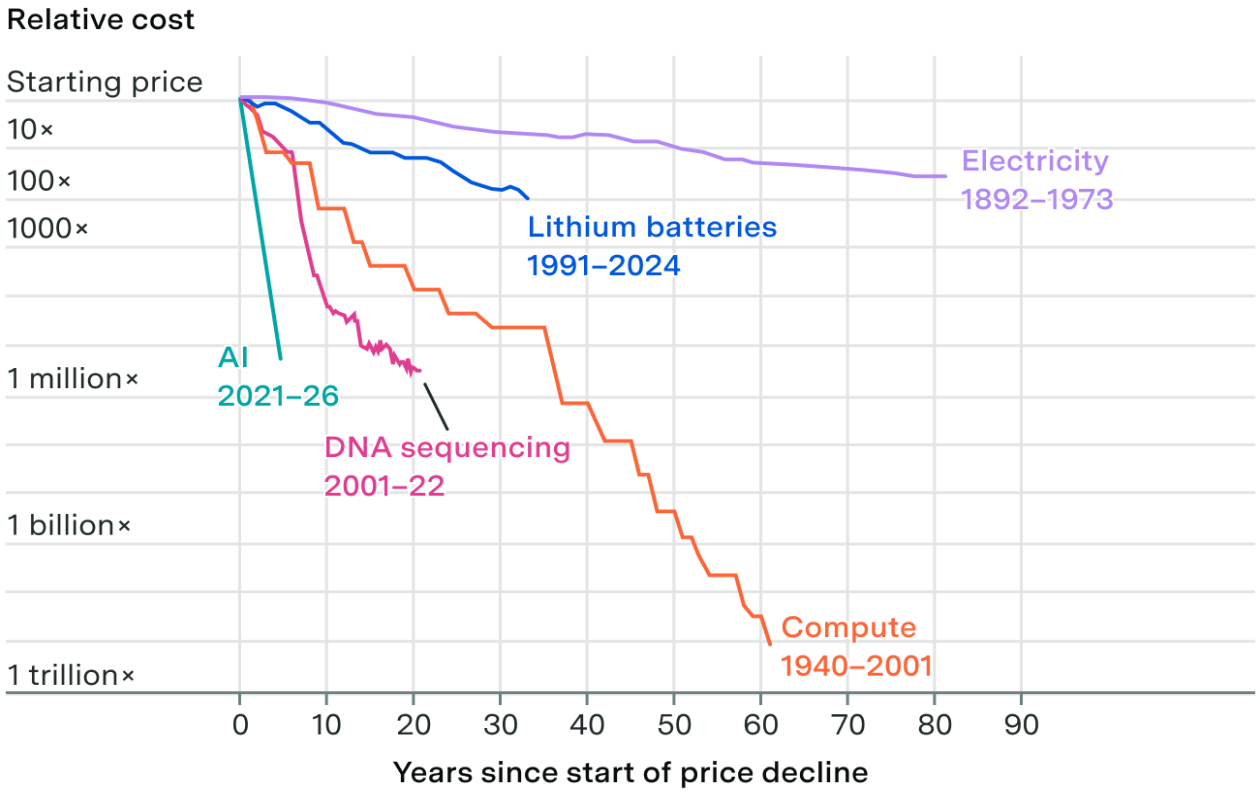
Cost per task

$ per million token (input + output) / 2



AMD
together we advance_

The price of artificial thought may have fallen faster than for any other transformative technology in history



All prices adjusted for inflation. Compute series stops merely for lack of newer data.
Sources: Electricity: Hersh (1999), Figure 2.1, and EEI (1995), Table 63. Compute: Nordhaus (2007), Appendix Table 2. Batteries: Fouquet (2026) via Our World in Data. DNA sequencing: National Human Genome Research Institute. AI line is 47%/quarter (12.7x/year), per estimates reported here for 2023–26; the extrapolation of that rate back to 2021 is broadly validated by Appenzeller (2024)'s finding that per-token prices for models clearing a certain capability threshold fell 10x/year in 2021–24.

EPOCH AI | CC-BY

epoch.ai

The trend towards cheap and capable digital intelligence is **historic!**

The Internet made all the worlds data available to anyone

AI can do the same thing for all the worlds expertise and critical thinking

We are seeing the first implications of the "AI Revolution"

AMD
together we advance_

Almost Nobody Has Seen What AI Can Really Do

> The pace you have seen is one data point. The frontier is moving faster than most of this room has had a chance to see.

~6B
OF THE WORLD

Almost no contact with these systems.

~1–2B
CASUAL USERS

Free chatbots, occasionally. A better search box.

~20M
PROFESSIONALS IN CODE & MATH

Frontier tools daily. Work that took weeks finishes from one prompt.

~5,000
INSIDE THE AI LABS

Agent teams run projects for weeks — faster than people can review them.

https://x.com/karpathy/status/2109361546505966046

Software disruption happened first. Hardware disruption is coming soon.

AMD
together we advance_

Chip Design Is Tailor-Made for AI

AI is strong at work that is digital and automatically checkable. Chip design is both — and the world cares about the result.

DIGITAL

The artifact is a file

RTL, netlists, testbenches, firmware, GDSII. Nothing physical happens until the mask.

CHECKABLE

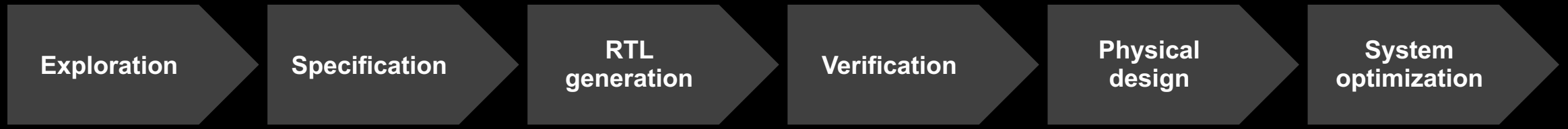
Correctness can be tested by machine

Simulation, formal, emulation, coverage, equivalence, timing and signoff.

These AI Trends are driven by Silicon innovation — and future Silicon Innovation will be enabled by these AI Trends

AMD
together we advance_

AI Throughout Chip Design



Exploration	Specification	RTL generation	Verification	Physical design	System optimization
Search vast architecture spaces	Draft and check specs from intent	Generate RTL and HLS from specs	Write tests, triage failures	Tune flows, predict PPA	Co-tune hardware, software, models

AI-ASSISTED PRESENT

Copilots suggest code, tests and fixes; engineers drive every step

AGENTIC FUTURE

Agents plan, call EDA tools, check results and iterate across stages

We believe that AI Agents will be applied wherever there is **robust and actionable feedback**

AMD
together we advance_

AI is driving AI Model Development

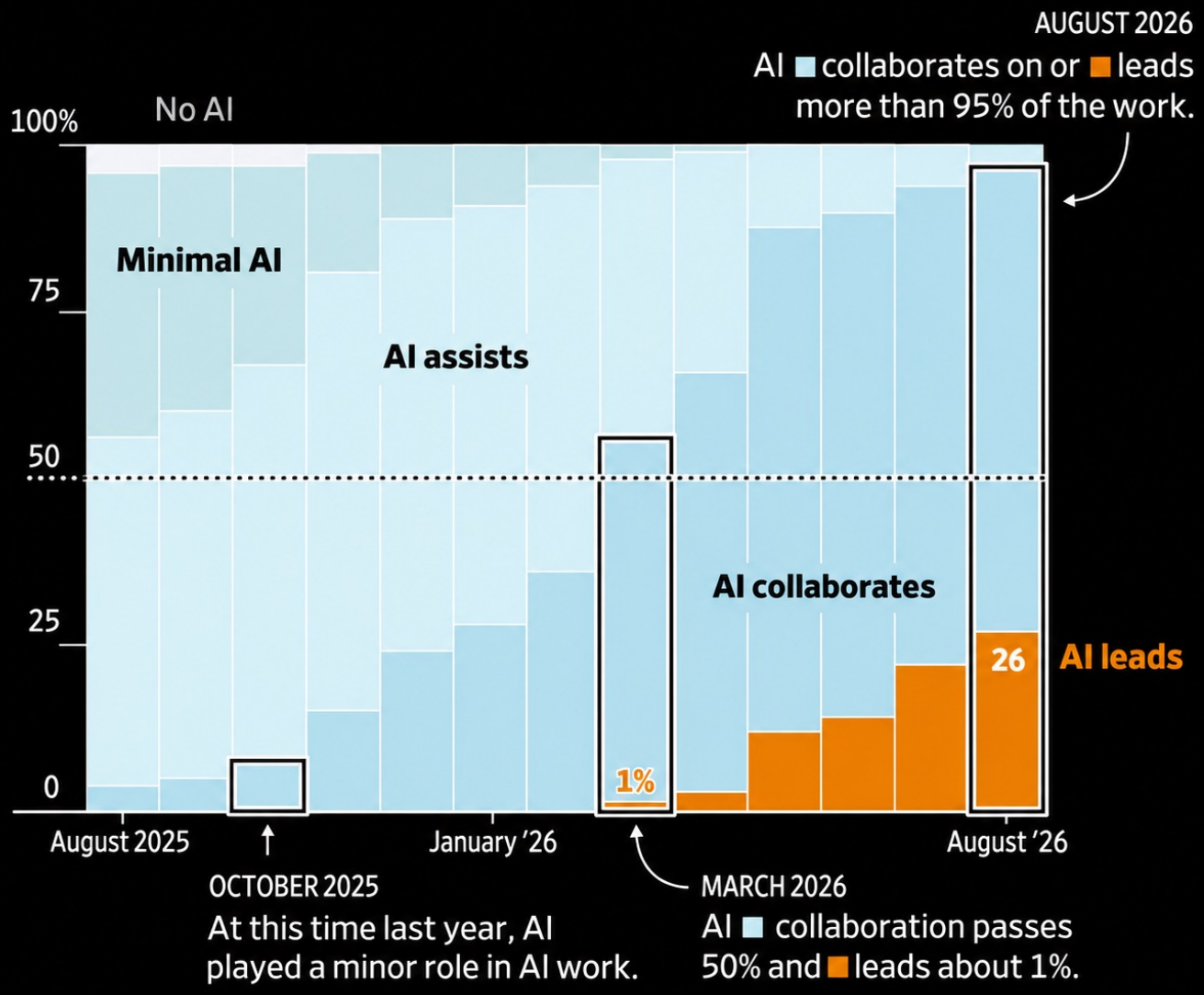


Chart : WSJ - ai-anthropic-rsi-recursive-self-improvement
Data :Anthropic R&D Automation Index v2026.07

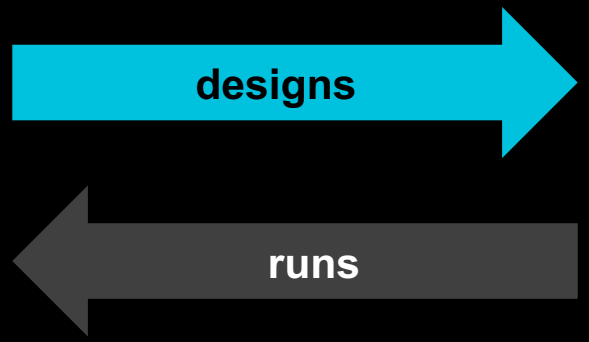
AMD
together we advance_

AI Designs Chips. Chips Run AI.

AI MODELS & AGENTS

The designer

Explore · specify · generate · verify



SILICON & SYSTEMS

The engine

Latency · energy · memory · cost per token

What happens when AI is both the workload we design for and a participant in designing the next generation of computing systems?

> What matters is not either box. It is how fast the loop turns – latency of a single synthesis run can set that pace.

AMD
together we advance_

What This Talk Covers

AI as a designer

Why AI can be effective at chip design, and what that already looks like in practice.

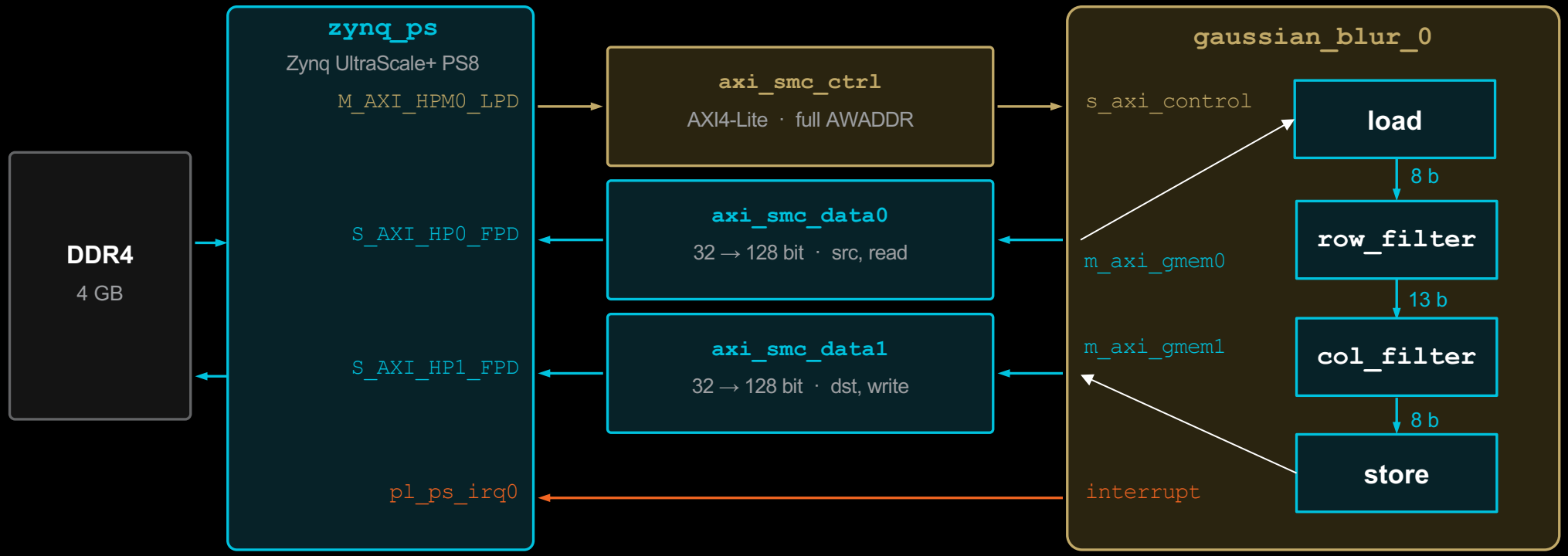
What AI asks of the hardware

The demands AI workloads place on silicon, and what this community should do next.

AMD
together we advance_

Copilot Case Study: End-to-End FPGA Design

"In your opinion, what is a good example that shows how Vitis HLS can help design a simple and efficient IP that is laborious to do in RTL?"



92 lines of HLS C++ -> 11,839 lines of Verilog @ 460 MHZ
Separable, multiplier-free implementation with linebuffer + bitwidth optimization
Optimized memory allocation to avoid DDR bank conflicts
Interrupt-driven software driver
Deployed on Kria board

One shot prompt
~100x Token compression with HLS

AMD
together we advance_

From Copilots to Agents

COPILOT

Suggests Options

Unit of work: a suggestion

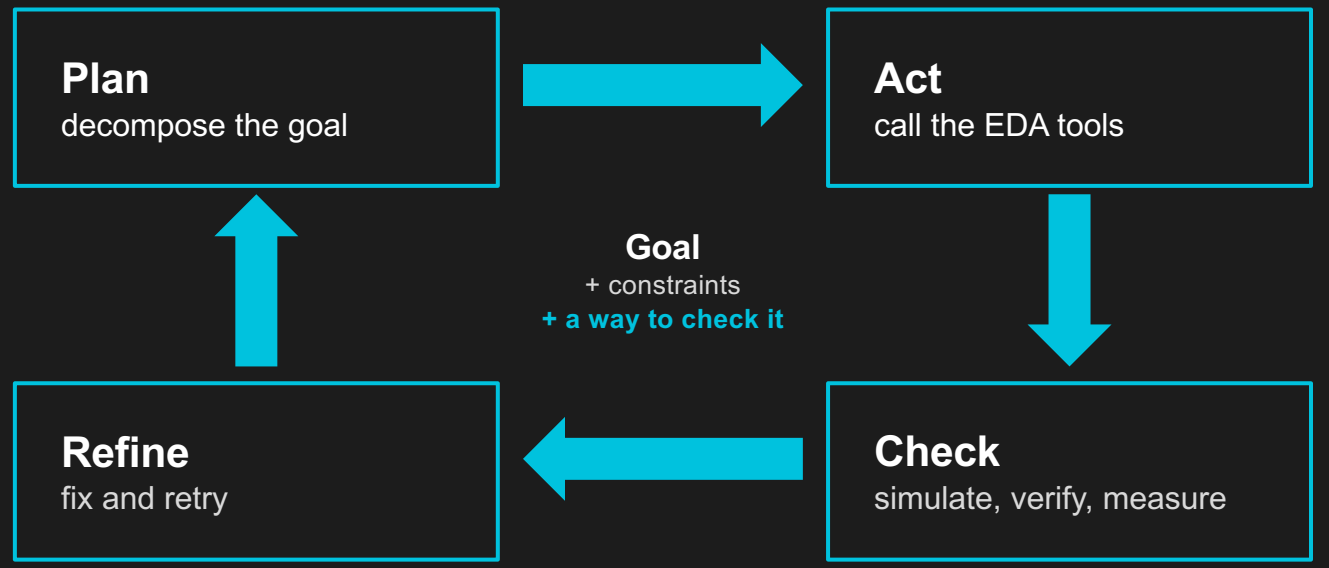
The engineer drives

1. Engineer prompts
2. Model suggests code or a fix
3. Engineer verifies and decides

AGENT

Plans, acts, checks, iterates

Unit of work: a verified delta



AMD
together we advance_

Getting the Most out of Agents

COPILOT AS ROBOT

We specify the how

"Use the profiler to measure this kernel."

AGENT AS GOAL-SEEKER

We specify the what

"Find the maximum performance this hardware can reach."

AMD
together we advance_

EDA Tools Provide a Robust Foundation

SIMULATION

Self-checking testbenches

Constrained random stimulus

FORMAL METHODS

Sequential equivalence

Provable assertions

EMULATION & PROTOTYPING

Software bringup

Hardware fuzzing

SIGNOFF

Timing, power, DRC, and LVS

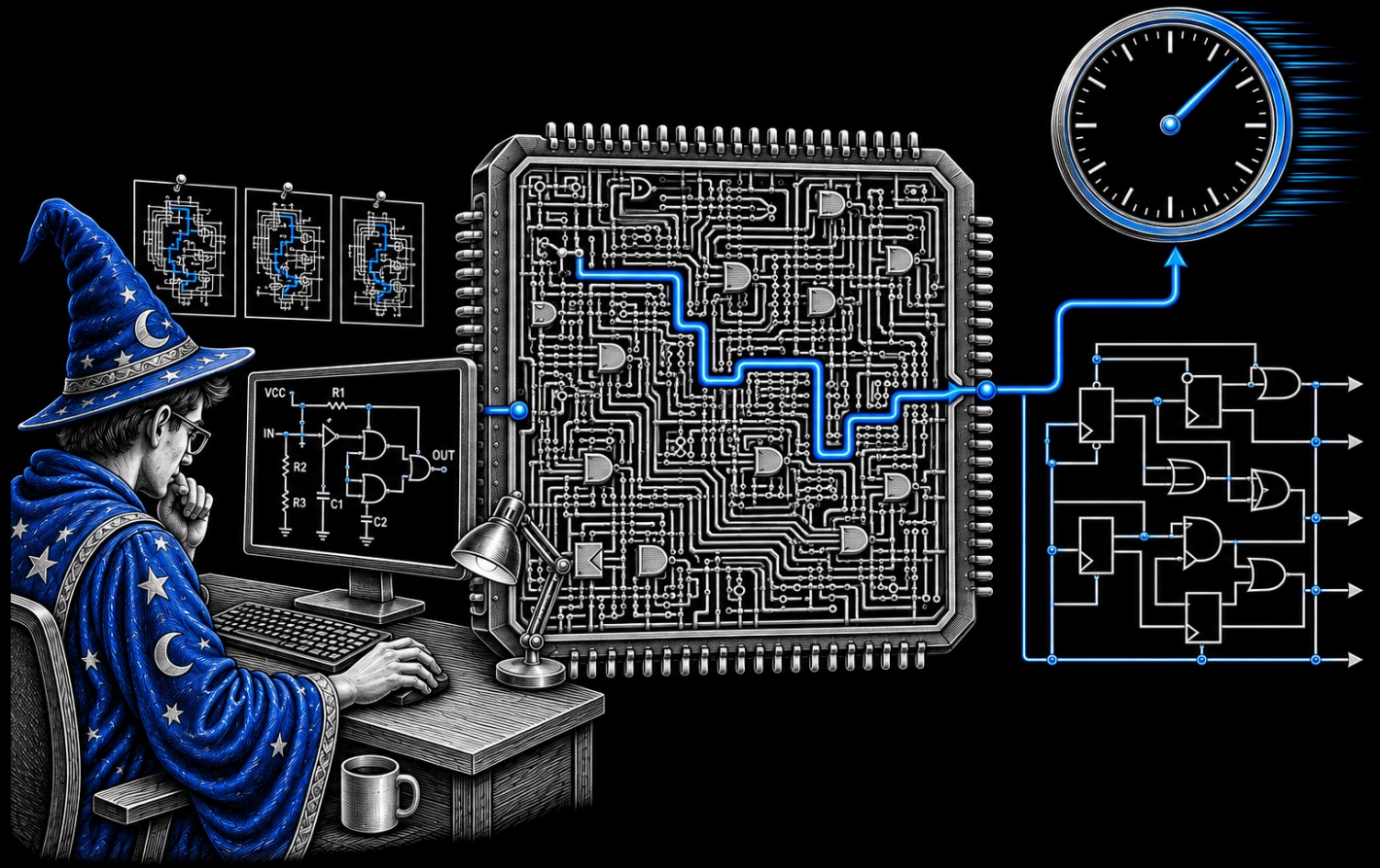
Other industries are now trying to build automatic correctness checks so that AI can be trusted with real work. Chip design has had them for decades.

AMD
together we advance_

Timing Closure: A Wizard’s Domain

Faster clocks can be achieved through expert tuning and repeated builds.

Hard to predict best placement for arbitrary netlists

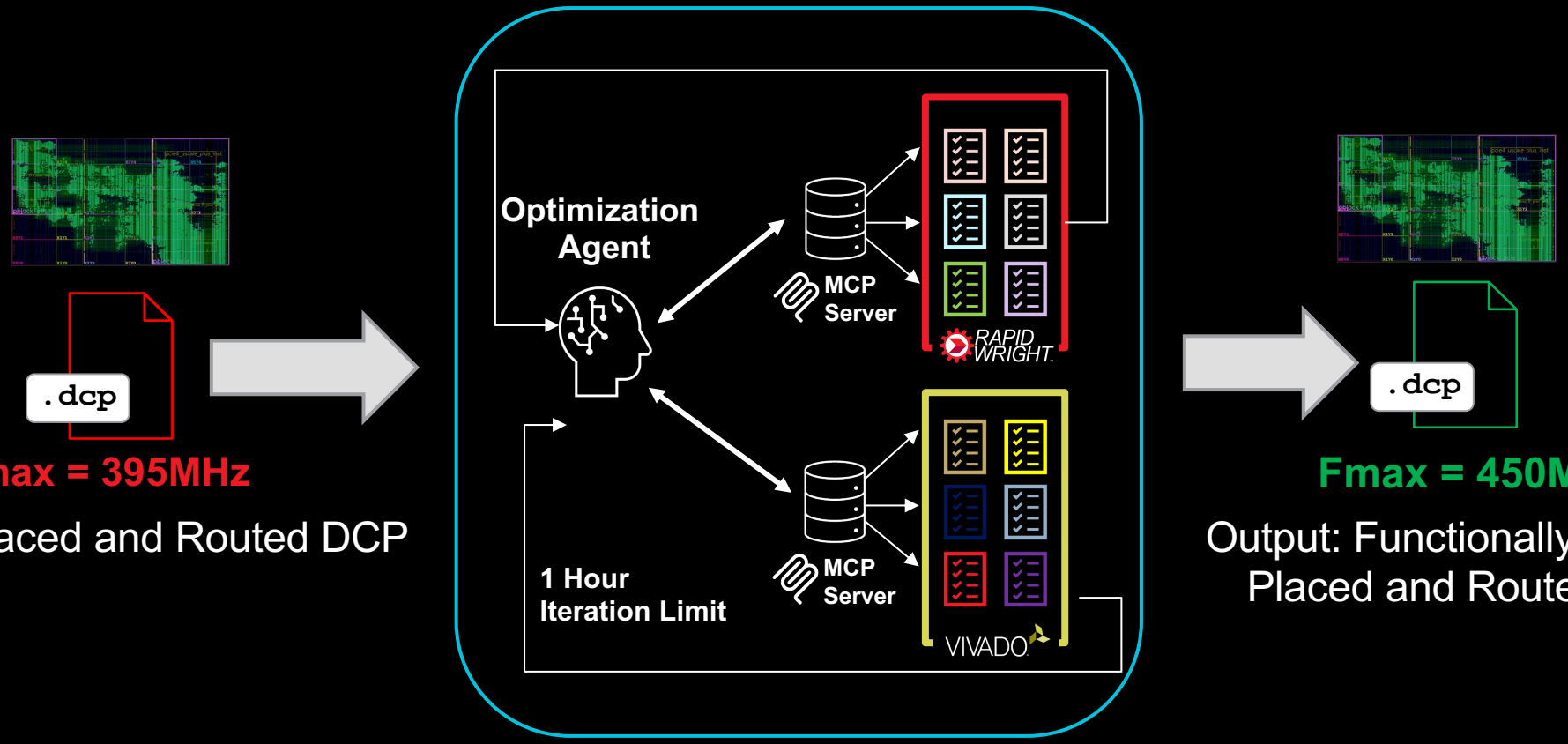


Can agentic loops tame this flow?

AMD
together we advance_

Case Study: AI-Driven Optimization for Timing Signoff

"Iterate with Vivado to fine-tune the design placement in order to increase frequency."

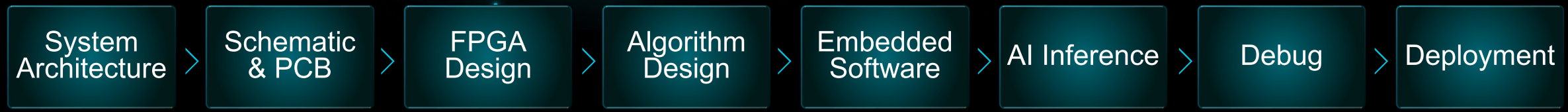


Agent Reasoning + Tool Feedback Enables Design Specific Optimizations

AMD
together we advance_

AMD Ross

AGENTIC AI ASSISTANT FOR EMBEDDED DEVELOPMENT



Inside AMD we are looking for solutions in both *FPGA and ASIC* design flows

AMD
together we advance_

How AMD Ross™ Turns Intent Into Action

Natural-language interaction & connected tools

ASK **Client-agnostic interface**
VS Code · Claude Code · Cursor · Devin · OpenCode CLI · any MCP-compatible client

INTERPRET **AMD Ross™ agentic workflows**
Agent skills encode proven AMD methodologies

CONNECT **MCPs + Knowledge Base**
Docs · PDM · Vivado™ · ChipScope™ · Vitis™ · ROCm™ · Ryzen™ AI SW

INFORM or ACT **Action layer**
Guidance · Reporting · Execution · Autonomous flows

Today: Knowledge Base · Vivado · Vitis HLS
Q4: Power Design Manager · ChipScope · Vitis AI

AMD VIVADO™

- Vivado Rev-Control
- RTL-Lint
- Timing Methodology
- Synth Report
- NoC Debug
- ILA Debug
- VIO-debug
- Synth Rep.

AMD VITIS™ HLS

14 skills

- HLS Synth
- HLS Architect
- HLS Flow
- Pragma Scope
- Perf Pragma
- HLS Optimize
- MATLAB to C++
- HLS Line Buffer
- HLS Impl Rep.
- HLS Flattenable
- HLS Dataflow
- HLS Cosim
- HLS Burst
- HLS Stencil
- Custom Skill
- + More every month

20+
AMD-expert skills available today

800+
early-access customers

Monthly
additional tools & flow capabilities

Not an exhaustive list of AMD-authored skills

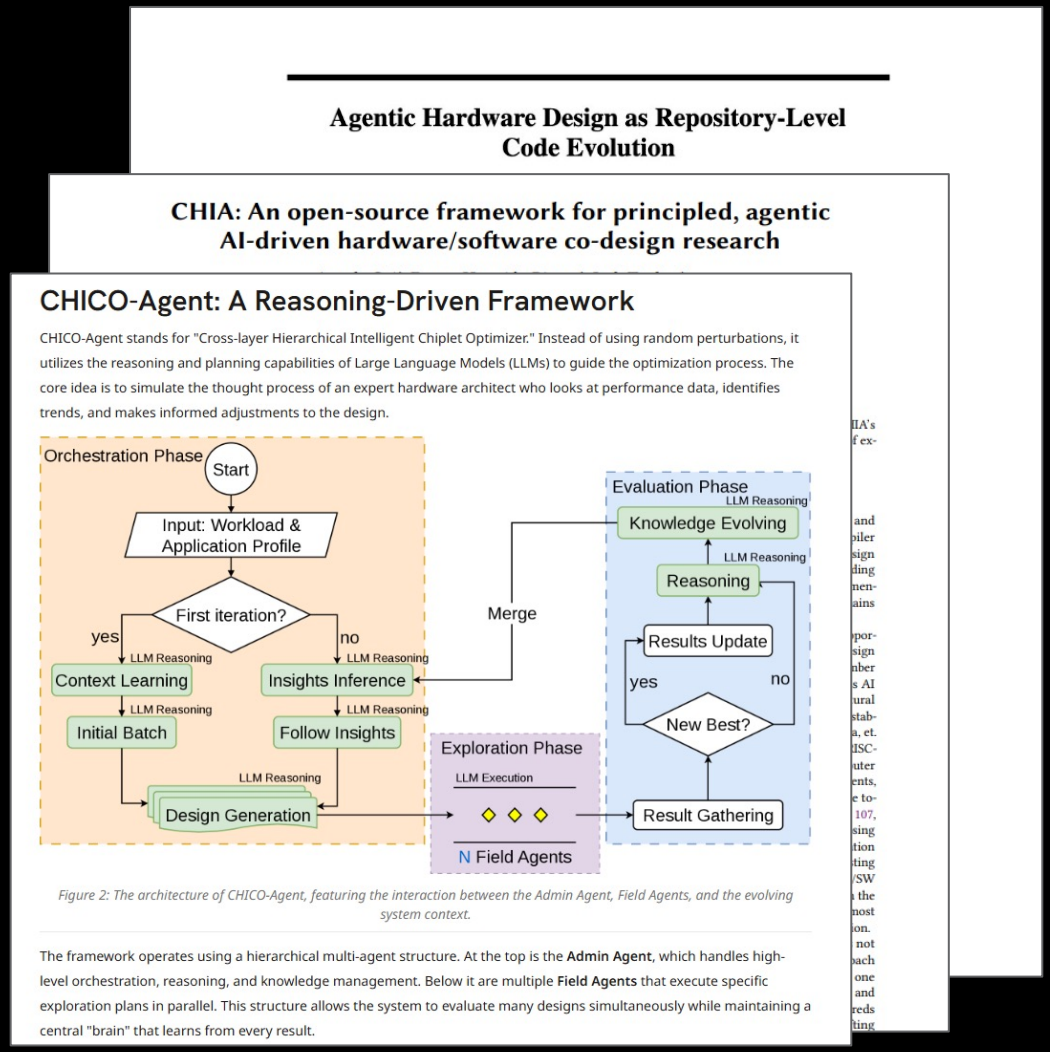
AMD
together we advance_

A Future with End-to-End Optimization

What prevents us from simply expanding this and letting agents iterate on the whole design?

3 Agentic Bottlenecks

- EDA Runtime
- Verification
- Organization



Agentic Hardware Design as Repository-Level Code Evolution

CHIA: An open-source framework for principled, agentic AI-driven hardware/software co-design research

CHICO-Agent: A Reasoning-Driven Framework

CHICO-Agent stands for "Cross-layer Hierarchical Intelligent Chiplet Optimizer." Instead of using random perturbations, it utilizes the reasoning and planning capabilities of Large Language Models (LLMs) to guide the optimization process. The core idea is to simulate the thought process of an expert hardware architect who looks at performance data, identifies trends, and makes informed adjustments to the design.

Figure 2: The architecture of CHICO-Agent, featuring the interaction between the Admin Agent, Field Agents, and the evolving system context.

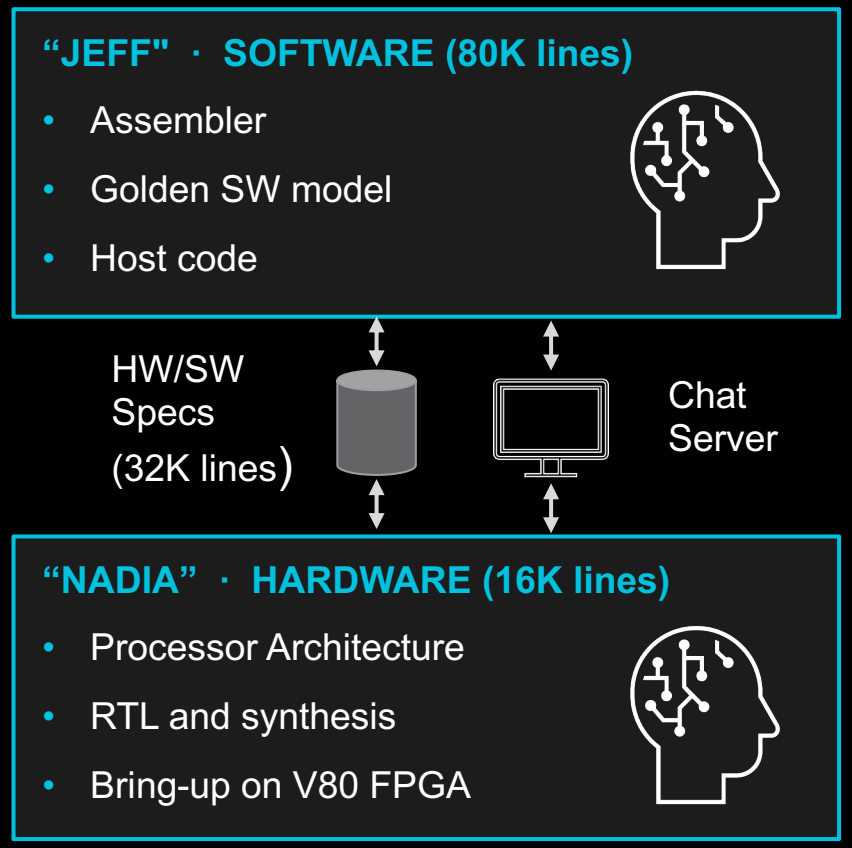
The framework operates using a hierarchical multi-agent structure. At the top is the **Admin Agent**, which handles high-level orchestration, reasoning, and knowledge management. Below it are multiple **Field Agents** that execute specific exploration plans in parallel. This structure allows the system to evaluate many designs simultaneously while maintaining a central "brain" that learns from every result.

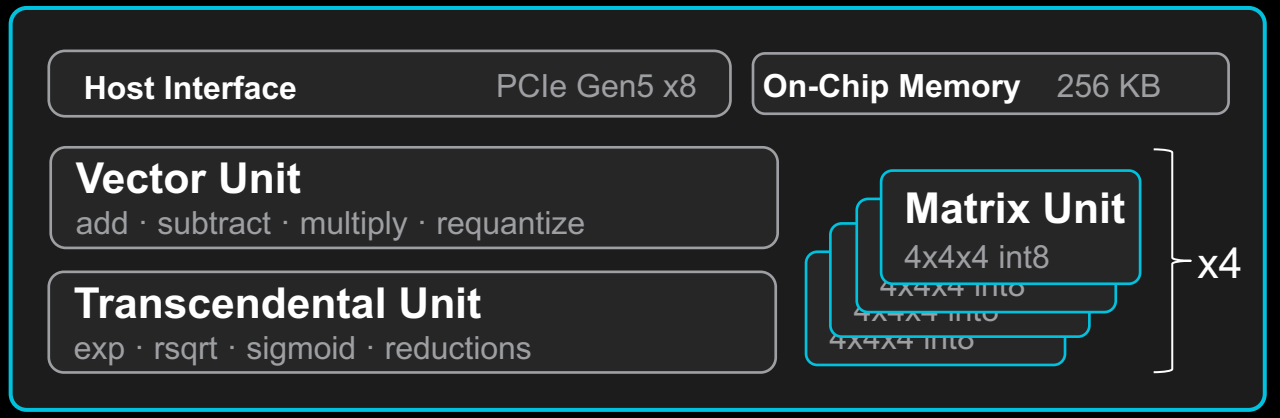
Recent works building larger agentic loops

AMD
together we advance_

Case Study: Two Agents Working Autonomously

"Build me a widget that automates my leaving home routine."

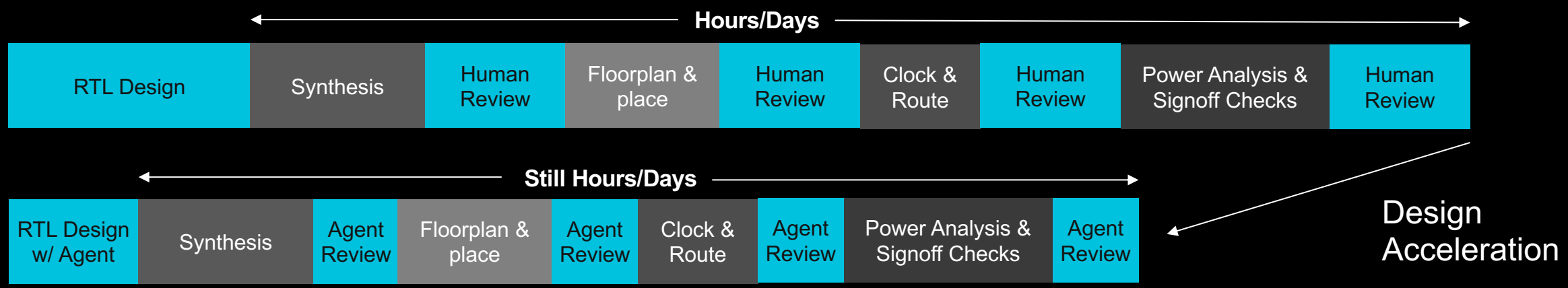




- 100 MHz Custom processor (13 opcodes)
- TinyJev Decision model (600M weights)
- 53 Design Iterations over 8 agent-days
- 21 minutes per decision !?!?!

AMD
together we advance_

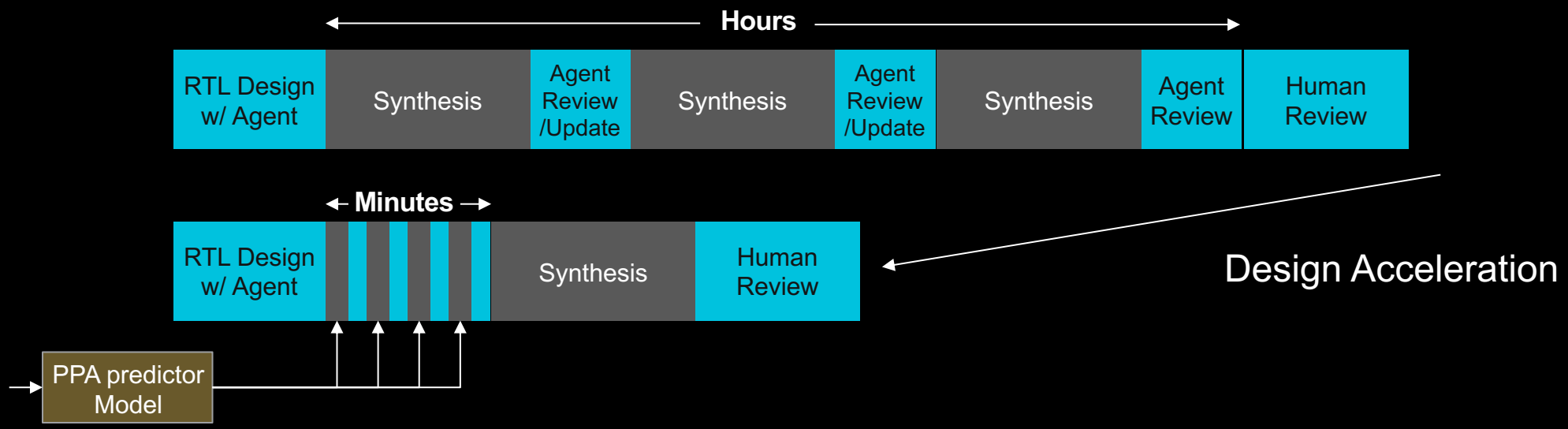
The EDA Runtime Bottleneck



Agents can reduce human latency — but throughput is still limited by traditional EDA!

AMD
together we advance_

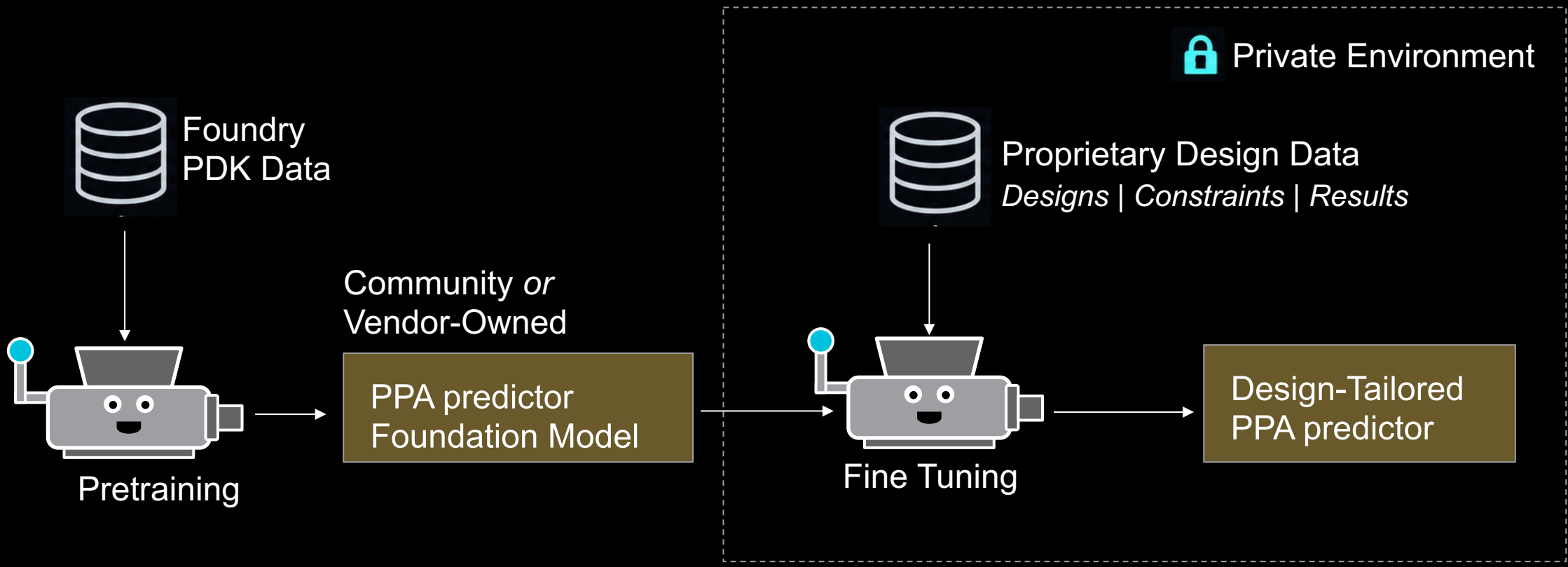
Predictive AI in EDA



Fast Surrogate Models can accelerate design-space exploration

AMD
together we advance_

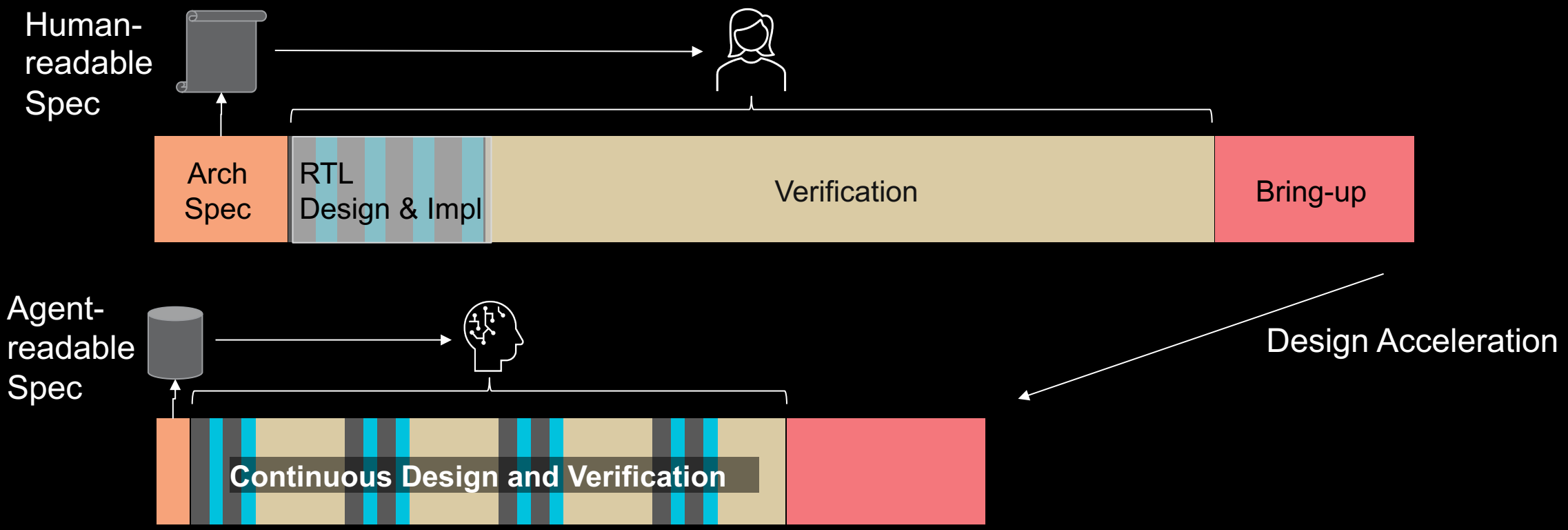
Tailored EDA



Partnership on Foundation Models needed

AMD
together we advance_

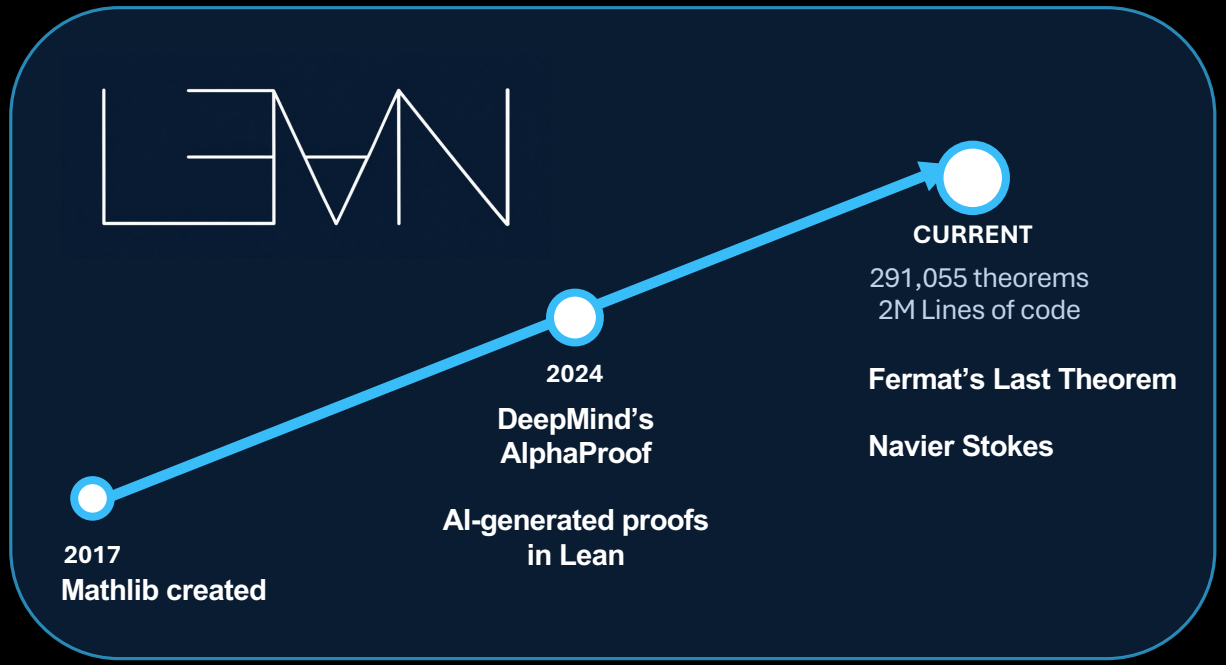
The Verification Bottleneck



Require Agents to provide behavioral certificates along with design artifacts

AMD
together we advance_

New Verification Technologies



Arbitrary mathematical theorems

Sound reasoning and decision methods

Significant existing theory library

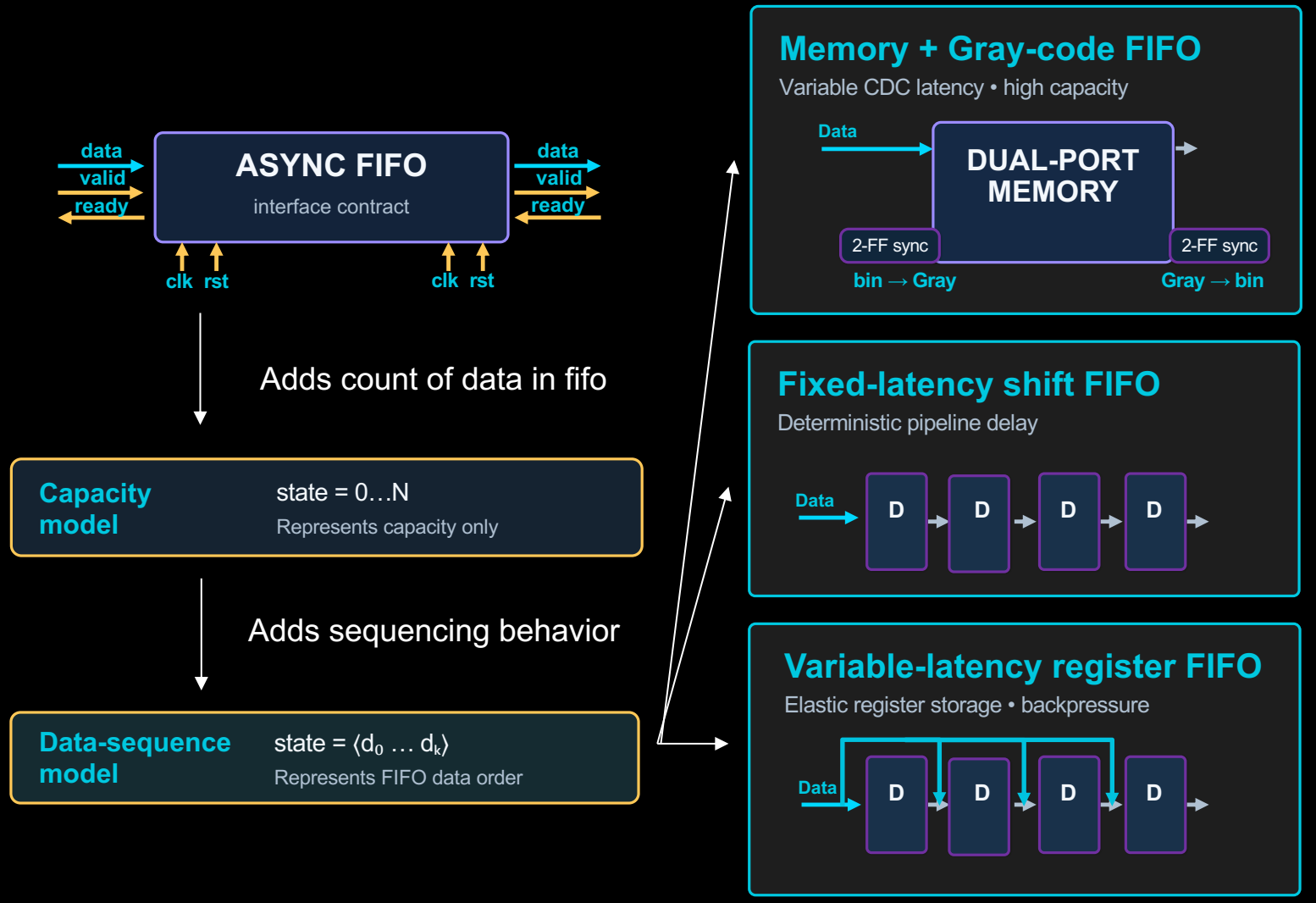
LLMs good at writing theorems

Future Agentic Systems Require New Specification and Verification Paradigms!

AMD
together we advance_

A Simple Lean Example

- Pre-RTL abstract specifications
- Proof by Induction
- Parameterized IP
- Numerical Properties
- Provable Programming Models



AMD
together we advance_

The Organizational Bottleneck

MELVIN CONWAY, 1968
"How Do Committees Invent?"

Communication structure becomes system structure

Team boundaries become interfaces. The shape of the conversation becomes the shape of the product.

FRED P. BROOKS, 1975
The Mythical Man-Month

Delivery capacity does not scale with headcount

Communication paths grow as n(n−1)/2: capacity rises linearly, coordination quadratically.

5 people	10 people	20 people
10 paths	**45 paths**	**190 paths**

> **A parallel machine is limited by its interconnect as much as by its cores.**
>
> **The same holds for the organisation that designs it.**

AMD
together we advance_

Teams Exchange Work Far Slower Than AI Produces It

GENERATION · AN AGENT WORKING

Latency
Seconds to minutes

Throughput
Thousands of tokens per second, in parallel, continuously

Errors
Automatically checked at every step

3–5
orders of magnitude

COORDINATION · BETWEEN TEAMS

Latency
Hours to weeks — a review slot, a sprint, a sync

Throughput
One document, one meeting, one review at a time

Errors
Nothing checks that the meaning survived

We've spent years building chip design processes around managing large teams of people

Our future will be determined by our ability to implement processes in which agents are effective

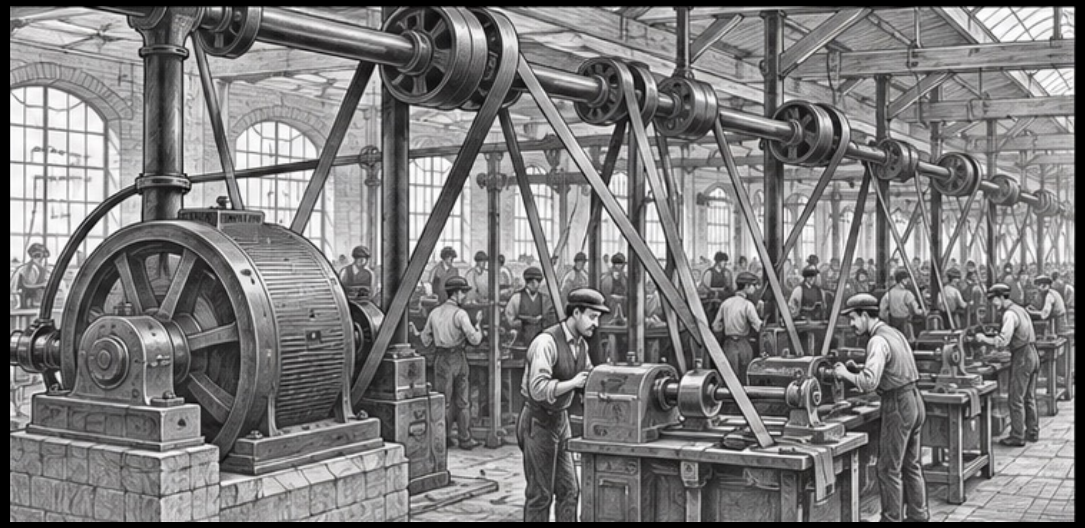
AMD
together we advance_

Adapting to New Technology

1890s

Incorporate in the organization

One large motor - same shafts and layout

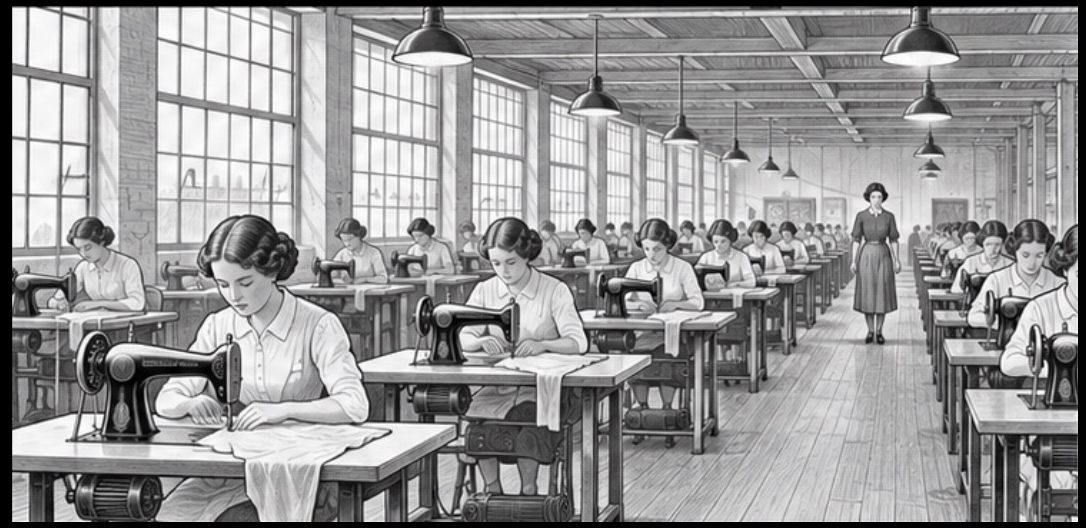


Incremental improvement

1920s

Redesign the organization

Small motors at each machine – new layout

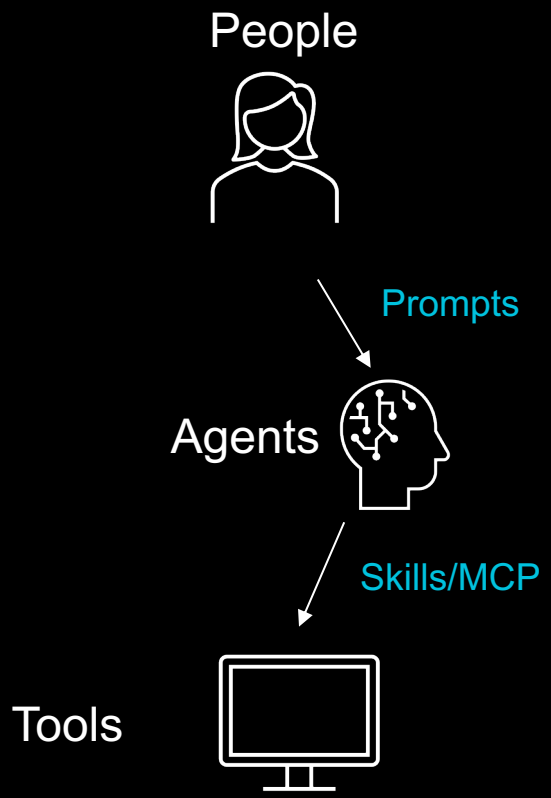


Transformative improvement

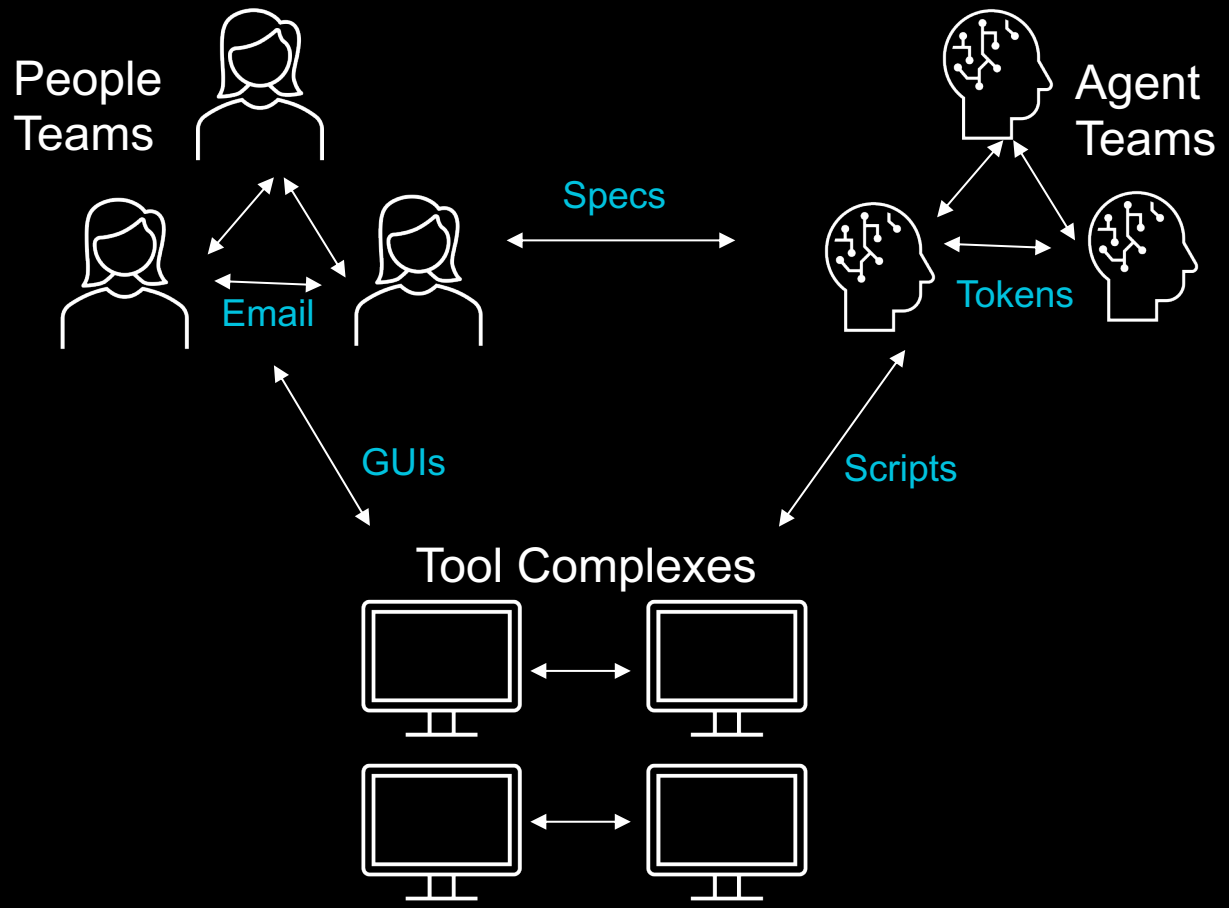
What changes when we redesign work around AI?

AMD
together we advance_

Future Agentic Chip Design



Individuals using Copilots + Agents



Organizations built around Humans and Agents

AMD
together we advance_

What AI Asks of the Hardware

> The demands AI workloads place on silicon,
> … and how improving silicon will improve agents.

AMD
together we advance_

Model Architectures Change Faster Than Silicon

Dense transformers
Attention at scale

Mixture of experts
Leveraging sparsity

Long context & reasoning
Test-time compute

Multimodal
Mixed tensor shapes

Agentic, tool use
Multi-turn sessions

SIX METRICS THAT ALWAYS REMAIN RELEVANT

Latency
time to first token, and every one after

Throughput
humans and agents waiting in real time

Energy
joules per token, chip to data centre

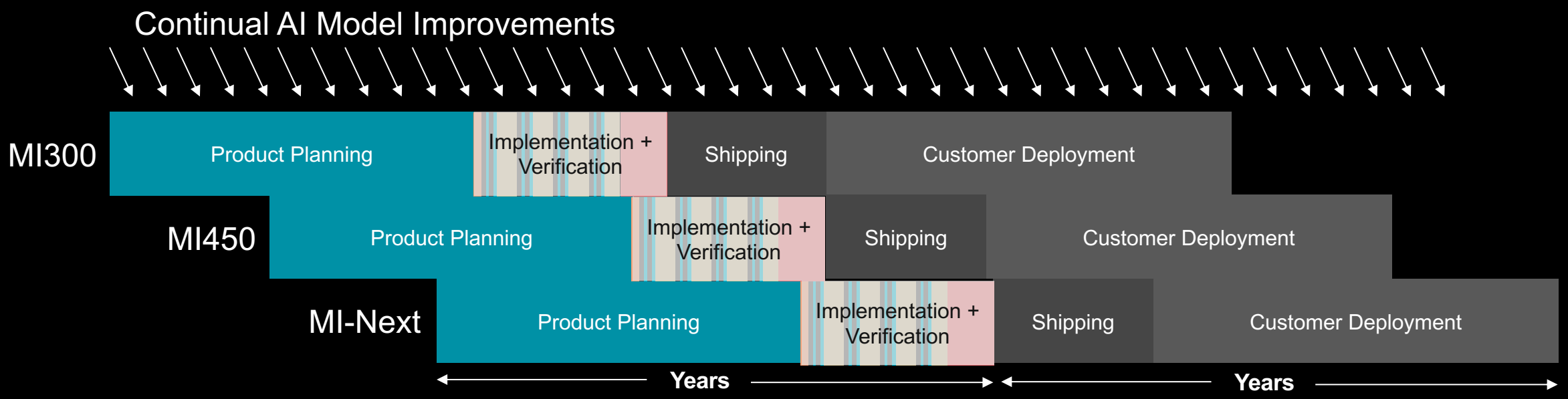
Cost
the unit economics of every deployment

Memory
bandwidth and capacity for weights and KV

Programmability
models change before silicon does

AMD
together we advance_

Chip Design Lifecycle



Optimize new hardware for existing models

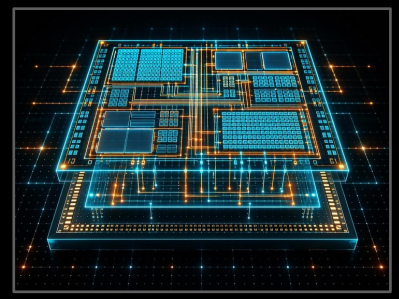
Optimize new models for existing hardware

AMD
together we advance_

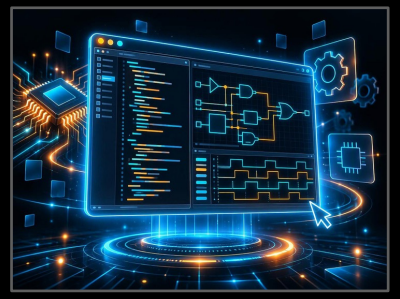
The Next Computers Will Be Co-Designed by AI - For AI



Model Providers



AI Silicon Providers



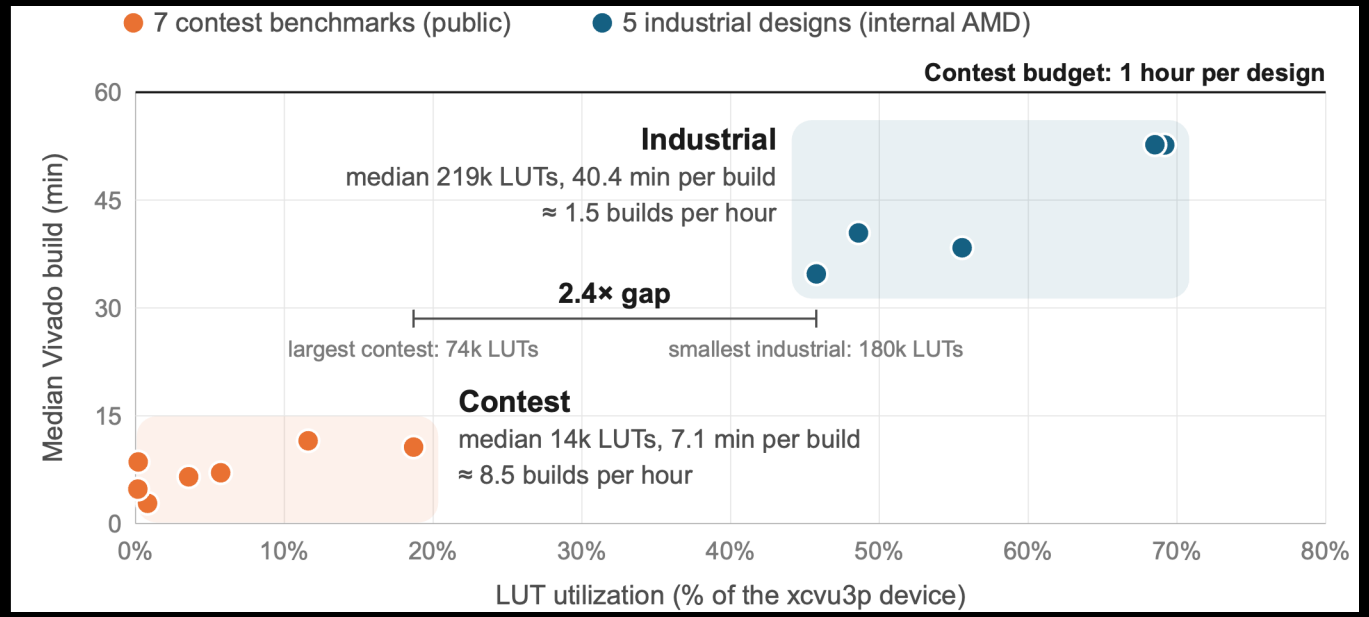
EDA companies

EDA as first-class inference workload. Specific latency, memory, per-cost token requirements.
Improvements will require an ICCAD community effort

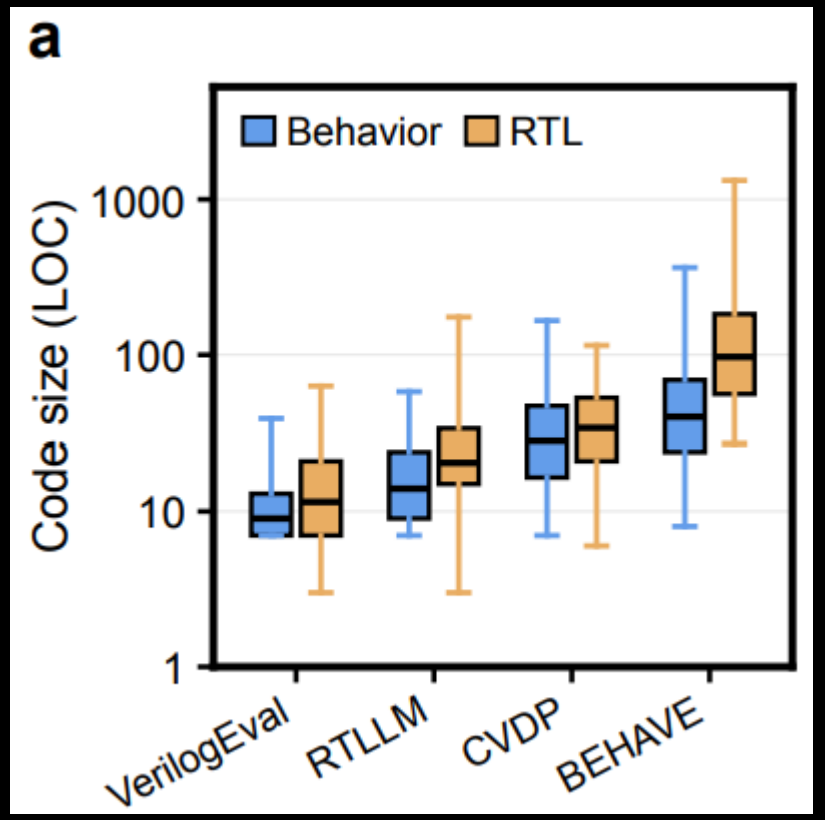
AMD
together we advance_

Benchmarks

- Current benchmarks are still small and quickly saturated.
- Data sets for training are locked inside companies.



AMD design contest



https://arxiv.org/pdf/2609.34785v1

AMD
together we advance_

Chip Design AI Index

A purpose-built index to measure models and agents across the chip design lifecycle

General Artificial Analysis Index benchmarks do not establish performance on chip design tasks.



Proposed task area	What models and agents do	What the index measures
Architecture & high-level design	Interpret specs, explore design trade-offs	Constraint satisfaction, performance estimates
RTL generation & repair	Generate RTL, diagnose and repair bugs	Compile / simulation pass rate, equivalence
Verification & validation	Generate tests / assertions, localize failures	Coverage, bug detection, proof success
Synthesis & optimization	Tune RTL, constraints and synthesis flows	Power, performance and area vs. baseline
Physical design & signoff	Floorplan, place and route, close timing	Timing slack, congestion, DRC / LVS errors
EDA agents & workflows	Operate tools, recover failures, coordinate steps	Task success, runtime, cost, human input

Open Weights LLM for Chip Design that we collaborate across the industry and research community

- industry builds with proprietary data on top.

Source: IBM

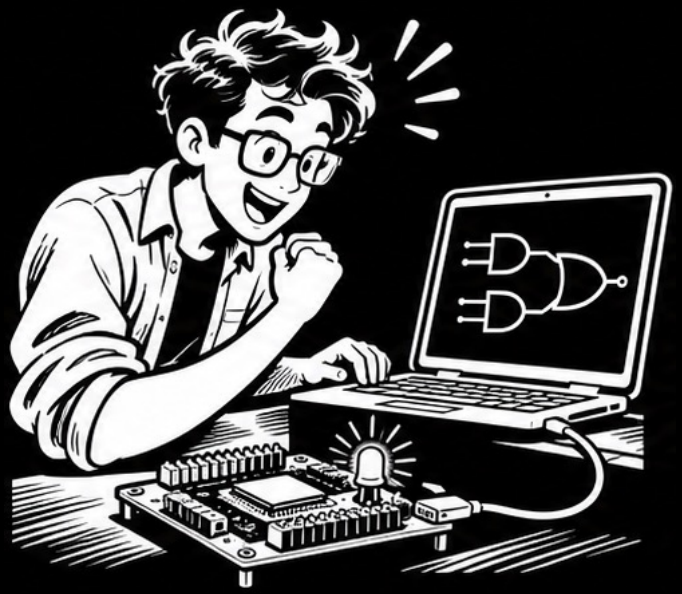
AMD
together we advance_

What the ICCAD Community Should Do

- Give agents goals and a way to check results
- Mine your design database for tribal knowledge
- Invest in new verification technology
- Rethink organizations with agents in mind
- Build open, end-to-end benchmarks together
- Leverage AI hardware to improve EDA

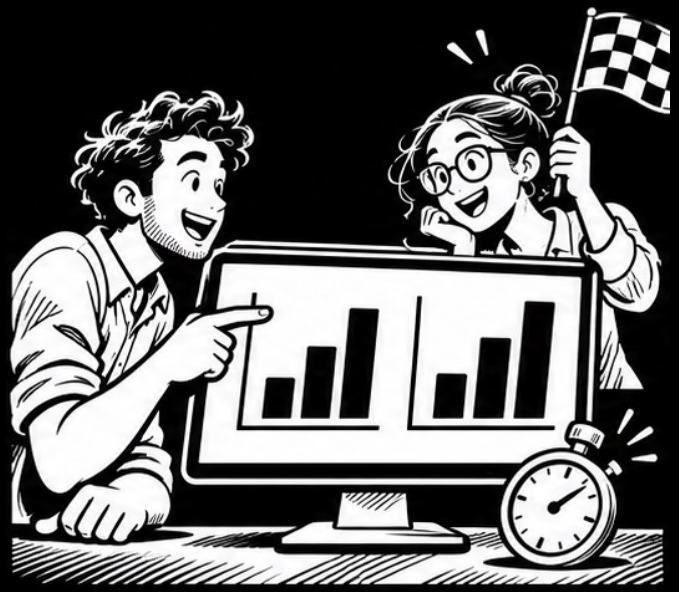
AMD
together we advance_

Paths to Engagement



Build with FPGAs

Get started with hardware design



Benchmark together

Join industry benchmarking efforts



Build the AI EDA community

Collaborate on foundation models for EDA

AMD
together we advance_

[Copyright 2026]

Copyright and Disclaimer

©2026 Advanced Micro Devices, Inc. All rights reserved.

AMD, the AMD Arrow logo, and combinations thereof are trademarks of Advanced Micro Devices, Inc. Other product names used in this publication are for identification purposes only and may be trademarks of their respective companies.

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 releases, for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or 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.

THIS INFORMATION IS PROVIDED 'AS IS." AMD MAKES NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE CONTENTS HEREOF AND ASSUMES NO RESPONSIBILITY FOR ANY INACCURACIES, ERRORS, OR OMISSIONS THAT MAY APPEAR IN THIS INFORMATION. AMD SPECIFICALLY DISCLAIMS ANY IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY, OR FITNESS FOR ANY PARTICULAR PURPOSE. IN NO EVENT WILL AMD BE LIABLE TO ANY PERSON FOR ANY RELIANCE, DIRECT, INDIRECT, SPECIAL, OR OTHER CONSEQUENTIAL DAMAGES ARISING FROM THE USE OF ANY INFORMATION

AMD
together we advance_


AMD