



Deploying AI for Chip Design & Verification

Matt Maidment

Introduction

Work for a small fabless semiconductor design company

IP Protection is extremely important to the company

Very experienced and pragmatic team

Not first adopters for legally risky technology

Not first adopters for potentially insecure technology

Do not have deep pockets to draw attention from vendors or others

This is a story of pragmatic deployment of generative AI

Timeline

2023

- Generative AI Use Forbidden

2024

- Chatbots using Frontier LLMs emerge
- Generative AI permitted for non-IP tasks and early investigations begin for DV
- Chip Design IP security requires expensive on-prem hosting

2025

- Compelling agentic Verification workflows demonstrated
- Companies trust upload of chip design IP to Frontier LLMs
- Frontier LLMs win on cost and capability
- EDA deployments move to Frontier LLMs
- Coding agents are a must-have for engineers

2026

- LLMs & coding agents improve *dramatically*
- Internal Team building AI-based infrastructure and supercharging individual engineers
- EDA augmenting existing tools, connecting datamodels and developing agentic workflows

2023



2024: Hello! How can I help you today?

H1: IM-based chatbot enabled for non-technical tasks

- Review business correspondence
- Supercharged search engine

H2: IM chatbot permitted for DV investigation

- Good understanding of well-known, open specifications
- Limited ability to write software
- Able to generate test plans and write simple tests

Industry Sentiment: Chip IP + Generative AI → On-Prem hosting

2025: Generative AI Optimism Accelerates

Early 2025 Reveals Exciting EDA Possibilities

- Agentic verification workflows demonstrated to team
- Ambitious EDA AI pricing revealed
- On-prem vs Frontier LLM deployment still in play

Frontier LLMs squash on-prem hosting as an option

- Low token costs, low administration costs, favorable IP security T&C and rapid improvements

Corporate/Legal policy streamlined

- Frontier LLMs with 0 data retention protect IP
- Upload: any information that does not expressly forbid it
- Generate: as long what's generated does not create corporate risk

2025: Team gets the hang of AI

New web-based chatbot okayed for use on production problems

- Download input data from various sources
- Upload artifacts file-by-file
- Cut-and-paste AI outputs into scripts
- Scenarios explored: reason about failures, identify trends in bugs, strategize implementation of complex algorithms, write scripts
- Team experiences rapid improvement of frontier models

Built AI library to predict DV simulation test memory usage and run time

- Using database or previous runs, trained a simple, on-prem model to predict future test run characteristics
- Enables prediction for test runs that were never previously run

2025: Coding agents arrive

September

"Can someone help me install Node.js so I can install and run this program called a coding agent in our datacenter?"

- The coding agent can run any program on our server?
- The coding agent has no affiliation with the company?
- A user has to approve every command?
- A YOLO agent can continuously code with extraordinary skill?
- NO 😊

October

- Embedded coding agent in a container to prevent network access and control file system and command access
- Established responsible use training for policy and use implications
- Initiated pilot to understand what can be done with limitations of secured agents

2025: What is the ROI of LLMs and Coding Agents?

Managers asked “Why should we risk uploading our most valuable IP to external servers (Frontier LLM providers)?

October 2025 to January 2026: Ran coding agent evaluation to assess ROI

- 25 participants coding SystemVerilog testbenches, assertions and coverage, scripts in various languages, web UIs and more.
- Participants reasoned about complex specifications and algorithms to more quickly understand and tackle challenges
- Participants built internal software for debugging and performed data analysis that they would have otherwise never accomplished

February 2026: Coding agent approved

- Following 2 months on-boarded hundreds of engineers & increased daily tokens use >10X

2026: Hands on with EDA

Evaluated first AI EDA tool in early 2026

- Agentic solution driven like a coding agent
- Decided to duplicate experiments with standard coding agent
- Found standard coding agent quite capable and learned quite a bit
 - Frontier LLM reasoning capability improving rapidly but not good enough for big, complex problems
- Users and LLMs spent too much time marshalling data to the LLM
 - Design data is often too large to be fully ingested or effectively analyzed
 - Source code and log files repeatedly searched with awk, sed, grep, rg, bash, Perl, Python
 - Many users worked with the agent to convert spec documents into markdown files
- Result: The tool and the coding agent failed, but...
- Participants impressed by the potential of generative AI

Future: Team Outlook

Evaluation experience clarified path for unlocking coding agents

- Built MCPs for waveform analysis, coverage data analysis, flow execution, bug tracking DB content
- Biggest gap is context: requirements, specifications, wikis, bugs
- Team building ambitious RAG system to organize, ingest and recall documentation

With efficient access to data, teams able to build workflows

- Analyze coverage, generate reports, write new tests
- Write testbench components and tests
- Build long-desired GUIs for debug, code generation and workflows

Using *secure* agents is the right direction

Future: EDA Outlook

EDA still building and learning

- Complex support systems required to enable AI in applications
- Lack built-in agentic security
- When EDA+GenAI combine: Is the value from EDA or the LLM?
- Value proposition is still TBD

Where EDA has value

- Traditional tools have efficient data models and APIs
- Traditional tools are complex. Should benefit from natural language interfaces
- EDA has the talent and experience to tackle big/complex problems ahead of customers
 - How to trust AI from spec to polygons with limited human interaction?