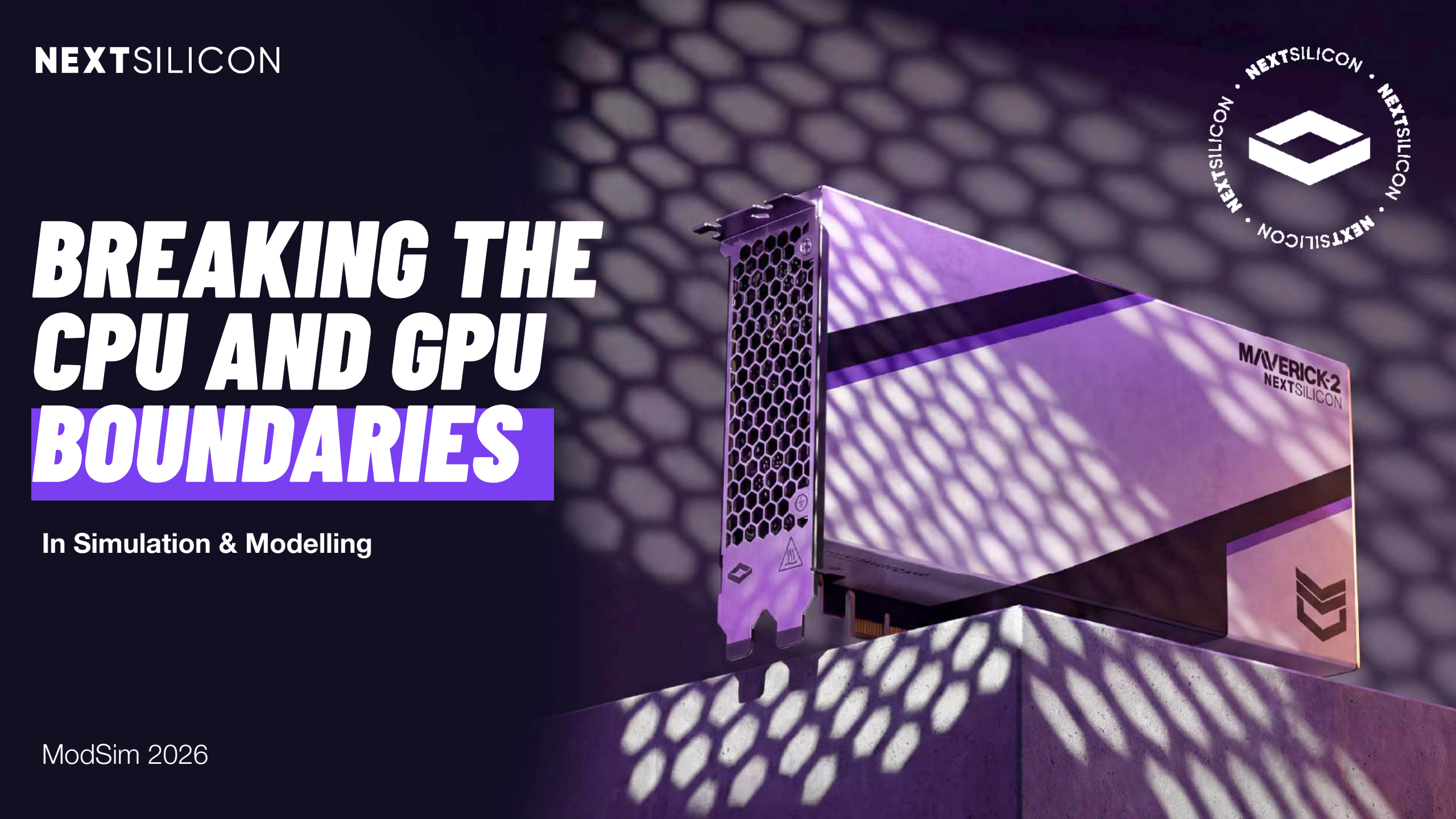
NEXTSILICON

BREAKING THE CPU AND GPU BOUNDARIES

In Simulation & Modelling



ModSim 2026

PRESENTER

HI! I'M ILAN

VP architecture
& Co-Founder

A bit about me

I **co-founded NextSilicon in 2018** and am its **VP Architecture**.

My 30 years' experience include senior staff engineer (software) at **Mellanox Technologies**, distinguished architect at **Integrity Project**, and senior engineer at Nocturnus Systems and Cogniview Systems.

I have authored many patents in computer architecture, and hold a B.Sc. degree in **physics and computer science** from the Hebrew University of Jerusalem.


NEXTSILICON
THE NEXTGEN
COMPUTE PLATFORM
MAVERICK·2

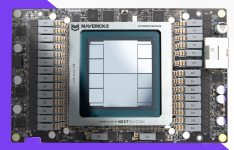
MAVERICK·2

DATAFLOW ACCELERATOR

MAVERICK·2 BY NEXTSILICON · MAVERICK·2 BY NEXTSILICON ·

ACCELERATING VANGUARD

MAVERICK·2



128x Maverick-2 installed on Sandia National Labs' Vanguard system

NEXTSILICON × Sandia National Laboratories

EVERYONE MAKES ONE

Usual Suspects

NVIDIA

intel

AMD

Qualcomm

EVERYONE MAKES ONE

Usual Suspects

NVIDIA
intel
AMD
Qualcomm

Hyperscalers

Google
amazon
Meta
Microsoft

EVERYONE MAKES ONE

Usual Suspects

NVIDIA
intel
AMD
Qualcomm

Hyperscalers

Google
amazon
Meta
Microsoft

AI labs

OpenAI
ANTHROP\C

WHY?

Usual Suspects

NVIDIA
intel
AMD
Qualcomm

Hyperscalers

Google
amazon
Meta
Microsoft

AI labs

OpenAI
ANTHROP\C

Global Tech Giants

HUAWEI
Alibaba

SQUEEZE A LITTLE MORE

↓

OPTIMIZE

↓

SPECIALIZE

Usual Suspects

NVIDIA. intel.

AMD Qualcomm

Hyperscalers

amazon

Meta Microsoft

AI labs

OpenAI

ANTHROP\C

Global Tech Giants

HUAWEI Alibaba

SPECIALIZE
EXTREME EXAMPLES

Google
Frozen V2

Gemini-specific ASIC
bakes the model into the silicon

Taalas
(Acq by AMD)

ROM weights
on last mask

SPECIALIZE
EXTREME EXAMPLES

Google
Frozen V2

Gemini-specific ASIC
bakes the model into the silicon

Taalas
(Acq by AMD)

ROM weights
on last mask

Market and industry, not research and innovation

SPECIALIZE
EXTREME EXAMPLES

Google
Frozen V2

Gemini-specific ASIC
bakes the model into the silicon

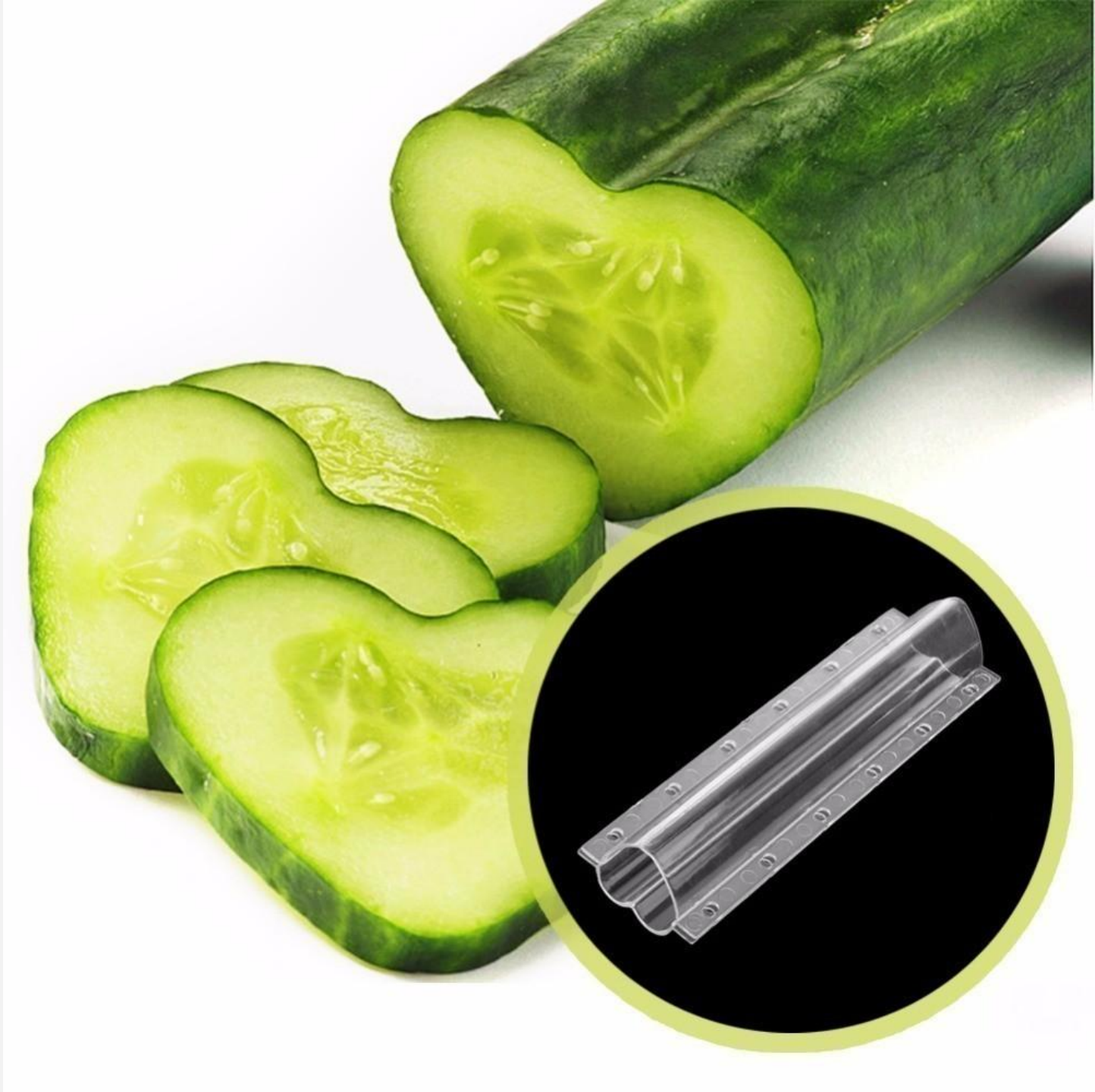
Taalas
(Acq by AMD)

ROM weights
on last mask

Market and industry, not research and innovation

Leads to stagnation, not to progress

BREAK THE MOLD



NOVEL MODELS

The **most common** use cases:

- Casual Generative AI
- Chatbots
- Agentic uses

Depend on **fast-changing novel models**

Public ones, tip of the iceberg:

- **DeepSeek V4** (April 2026)
- **Kimi K3** (July 2026)
- **Minimax** H3 (July 2026)
- **Qwen 3.8** (August 2026) - just last week!

NOVEL MODELS

The **most common** use cases:

- Casual Generative AI
- Chatbots
- Agentic uses

Depend on **fast-changing novel models**

Public ones, tip of the iceberg:

- **DeepSeek V4** (April 2026)
- **Kimi K3** (July 2026)
- **Minimax** H3 (July 2026)
- **Qwen 3.8** (August 2026) - just last week!

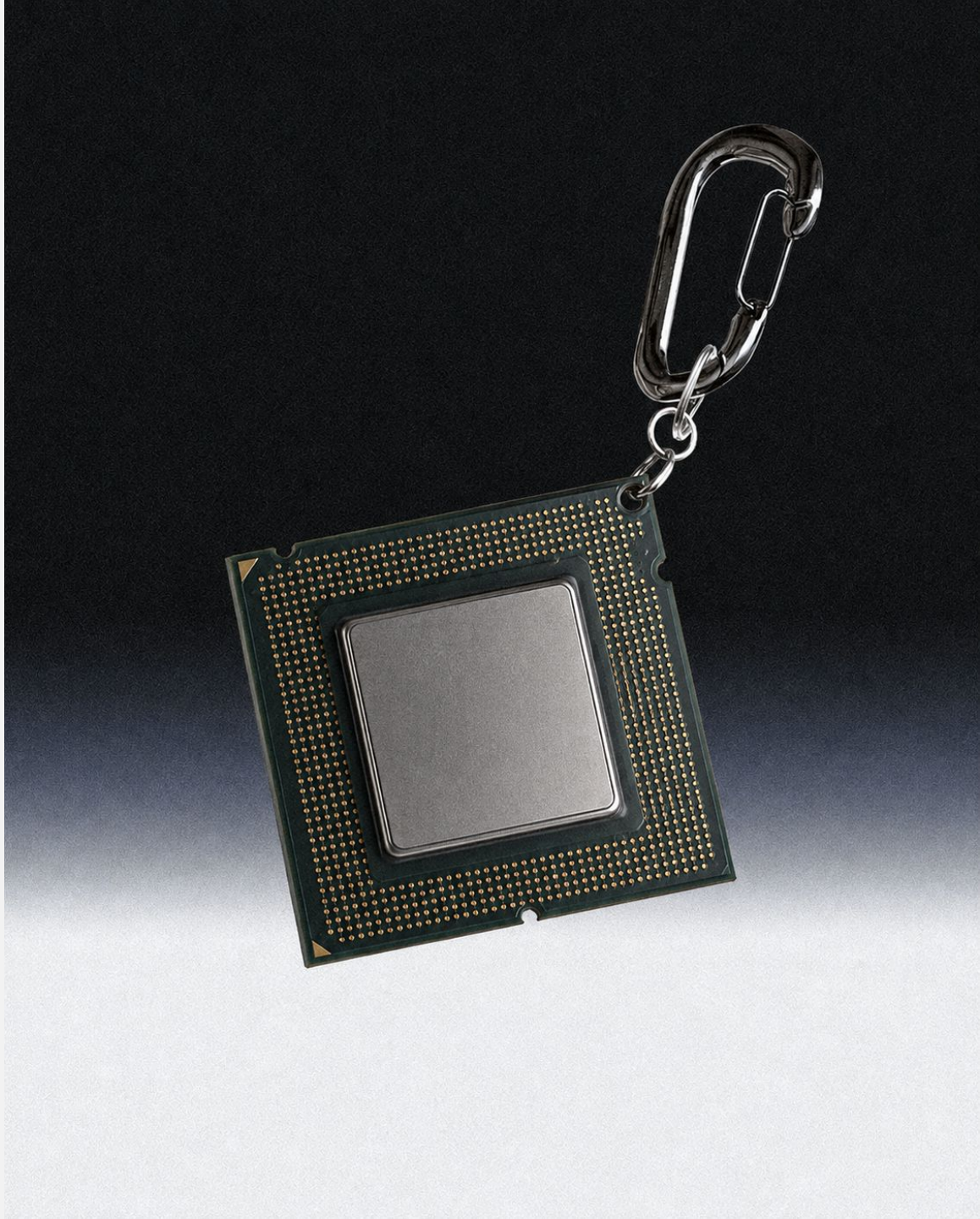
$100M ASIC Makes sense financially?

NOVEL MODELS

MAKE SENSE FINANCIALLY?

Today’s algorithm is tomorrow’s keychain

Re-do **every time** the algorithm changes



GENERAL PURPOSE HARDWARE

Investors circa 2018:

> Why is HPC and general-purposeness interesting at all? Everything is just matrix multiplications

GENERAL PURPOSE HARDWARE

Investors circa 2018:

> Why is HPC and general-purposeness interesting at all? Everything is just matrix multiplications

Recent AI advancements are **far from it**

Activations, Sparse MoE, Hybrid attention, Gated DeltaNet, ‘Adaptive Thinking’, ‘Dreaming V3’

The trend: Future of AI depends on **algorithm variety**

AGENTIC MODSIM

Agentic AI in ModSim

Chemistry question -> script to run GROMACS

Accelerators don’t run GROMACS… or even the script.

Leading to complex systems

A system that can do both!

FUTURE: AGI

AGI Holy Grail

Changing rapidly.

What would it look like 2 years from now?

The future-proof approach allows you to be better prepared for what's next in AGI.

TAKEAWAY

- **Mixed HPC and AI workloads are here:** surrogate models, physics-informed ML, Agentic AI using HPC modelling
- Systems of today have much **to improve**
- **Big opportunities** in general purpose hardware for AI
- Very few contestants - it's **super hard**.

NEXTSILICON

MAVERICK·2

INTELLIGENT COMPUTE ARCHITECTURE®

MAVERICK·2 BY NEXTSILICON · MAVERICK·2 BY NEXTSILICON ·

NEXTSILICON