



# AI: Trailblazing the Path of Semiconductor Innovation

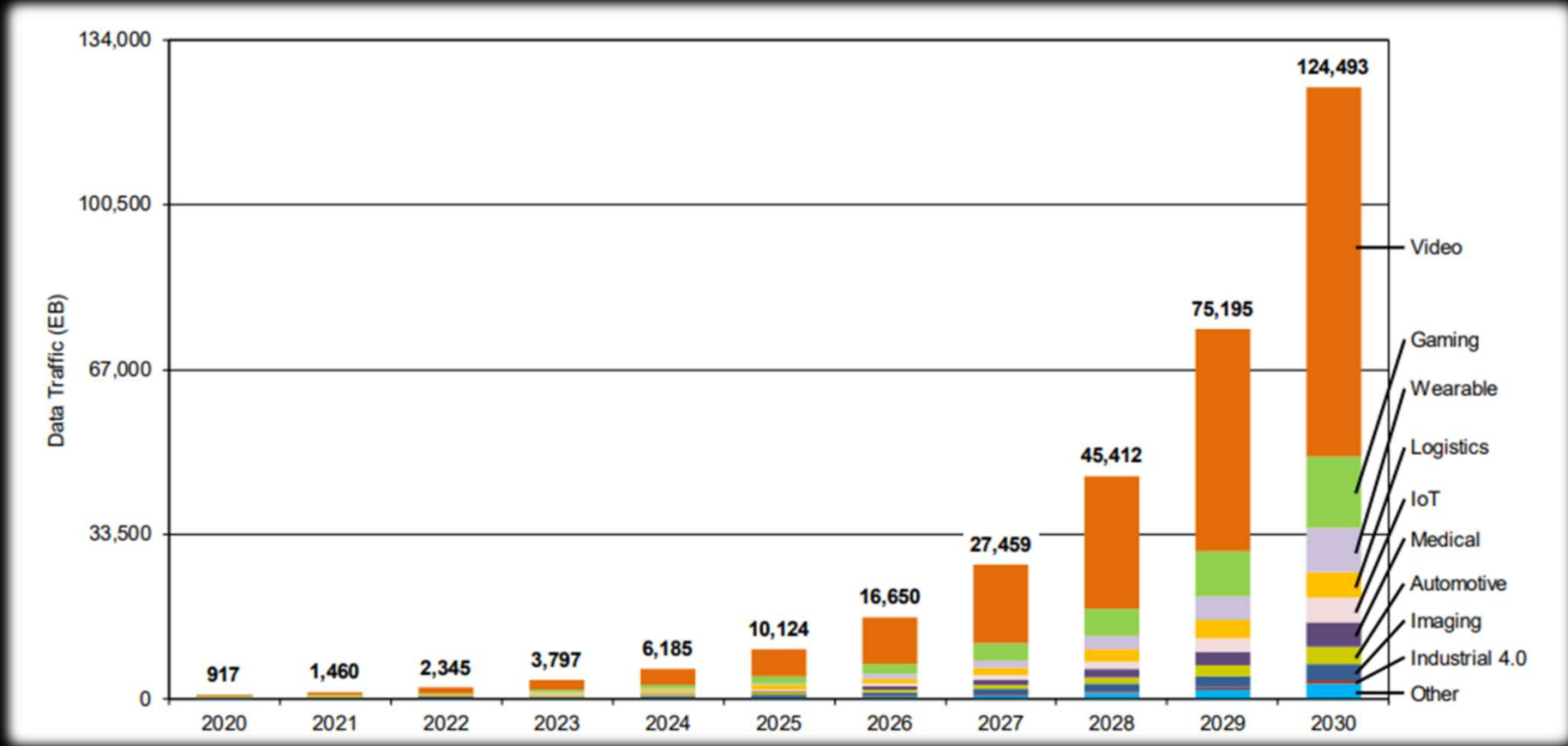
**Vishal Khandelwal**

Chief Architect, Design Technology Group

October 2024



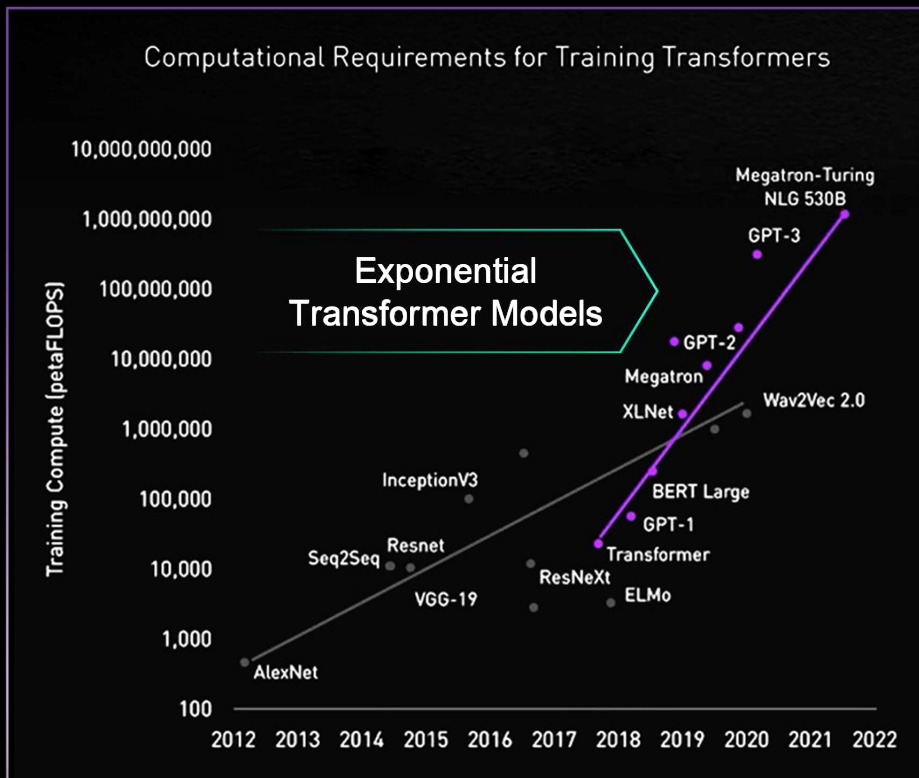
# World is Becoming More Data Intensive



Courtesy: IBS, Impact of Generative AI on Semiconductor Industry, Jan 2024



# Newer Models Further Pushing Limits of Compute



All Models Excluding Transformers:  
**8X / 2 years**

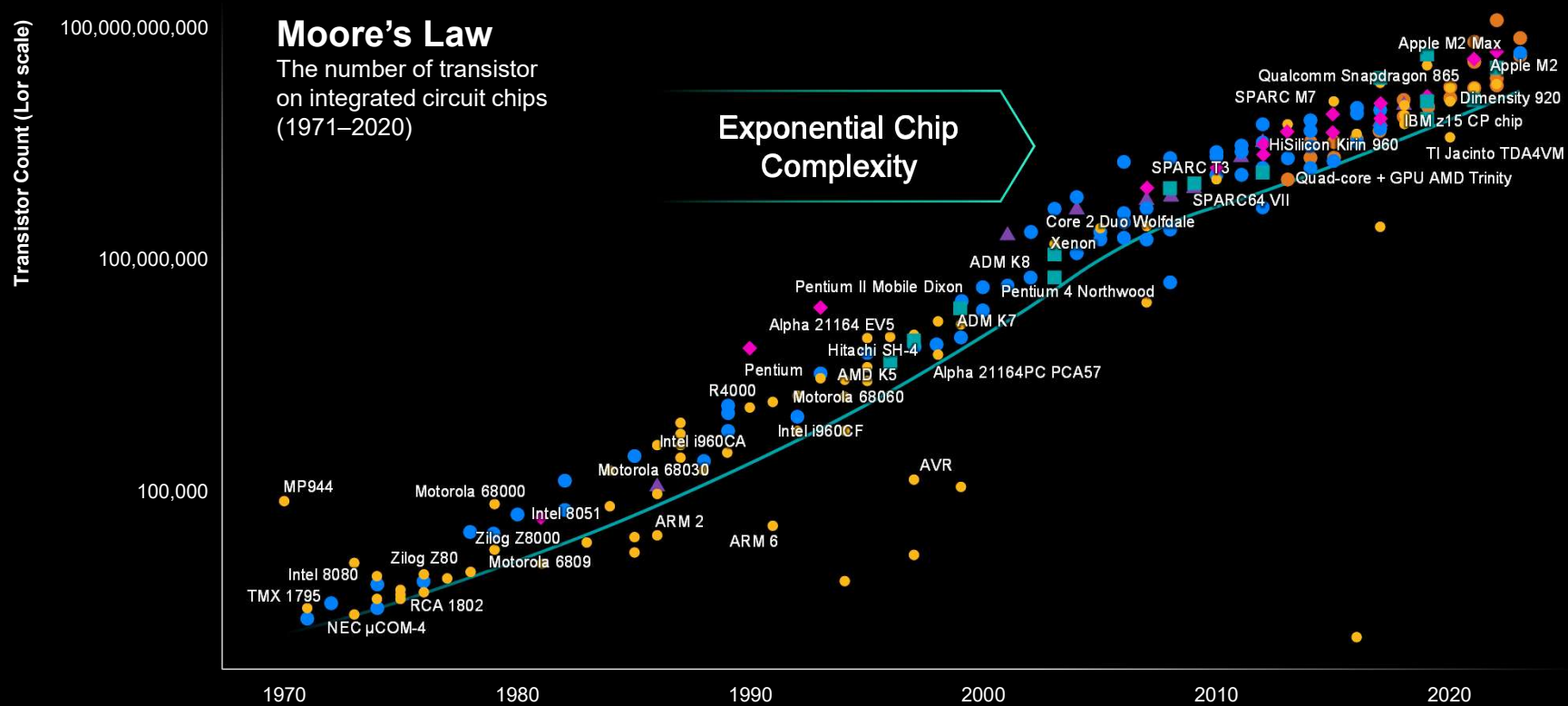
Transformer AI Models:  
**275X / 2 years**

Context-Aware Transformer  
Models Come at a Price

Source: <https://blogs.nvidia.com/blog/2022/03/25/what-is-a-transformer-model/>

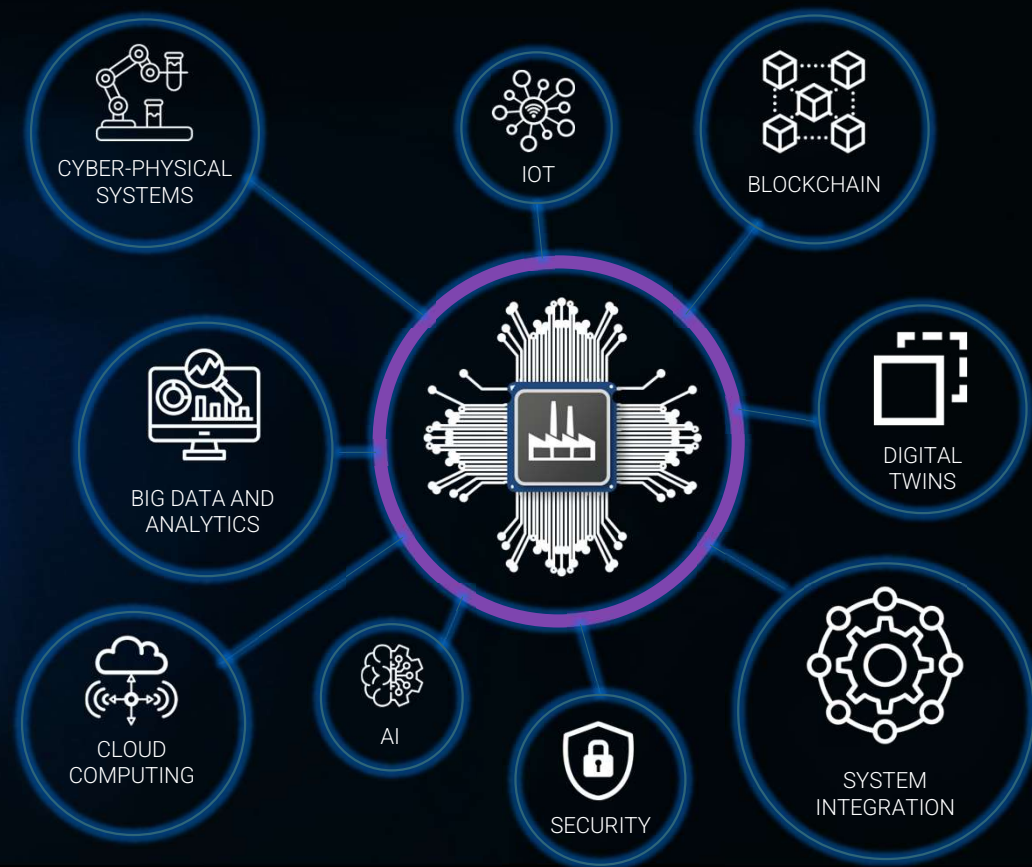


# In Turn, Pushing the Limits of Chip Design



Source: Wikipedia – Transistor Count: [https://interludeone.com/posts/2021-04-21-chips/chips\\_files/figure-html/unnamed-chunk-4-1.png](https://interludeone.com/posts/2021-04-21-chips/chips_files/figure-html/unnamed-chunk-4-1.png)

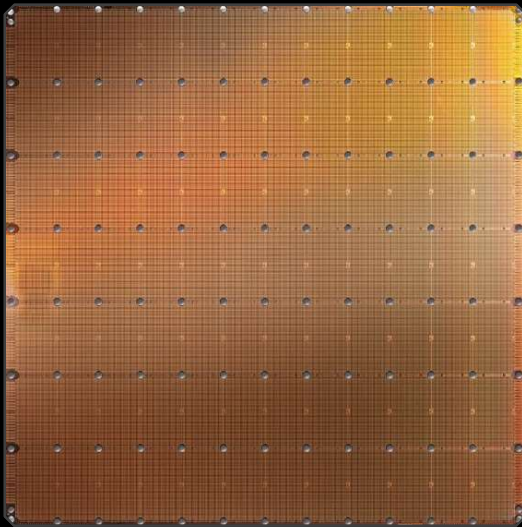
# The Fourth Industrial Revolution – Powered by Silicon



# **Hardware** is at the heart of the **AI revolution**

# Examples of AI “Super-chips”

**Cerebras WSE - 2**



2.6 trillion transistors  
TSMC 7nm  
850,000 AI-optimized cores

**Graphcore GC200 IPU**



59.4Bn transistors  
TSMC 7nm @ 823mm<sup>2</sup>  
1472 independent processor cores

## Data center chips for deep learning training and inference

- Trillions of transistors
- Hundreds of thousands of processing elements

## Edge IP (primarily) for deep learning inference

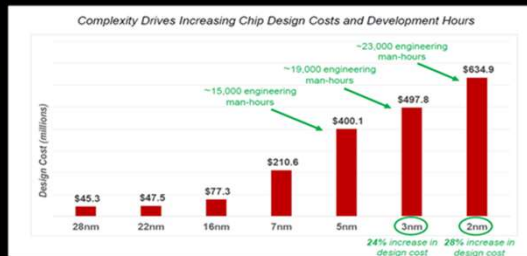
- Mixed scalar/vector/spatial compute
- Ultra energy efficient:  
Several TOP/s/W



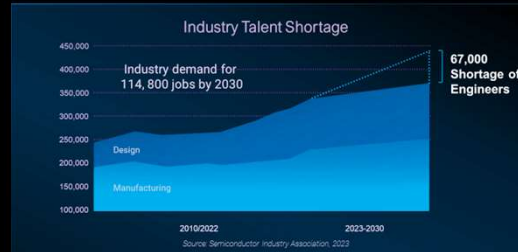
# Why AI, Why Now?

Chip Design Complexity, Cost and Labor Shortage Drive the need for Productivity

## Growing Demand & Systemic Complexity



## 30% Talent Shortage By 2030



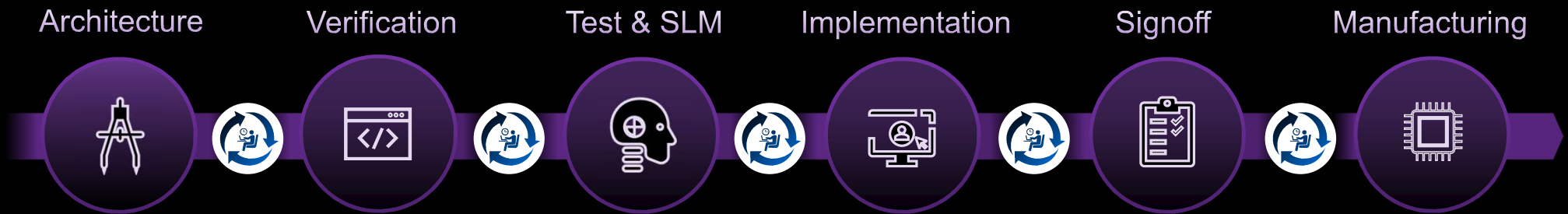
## Catalyst for AI Design



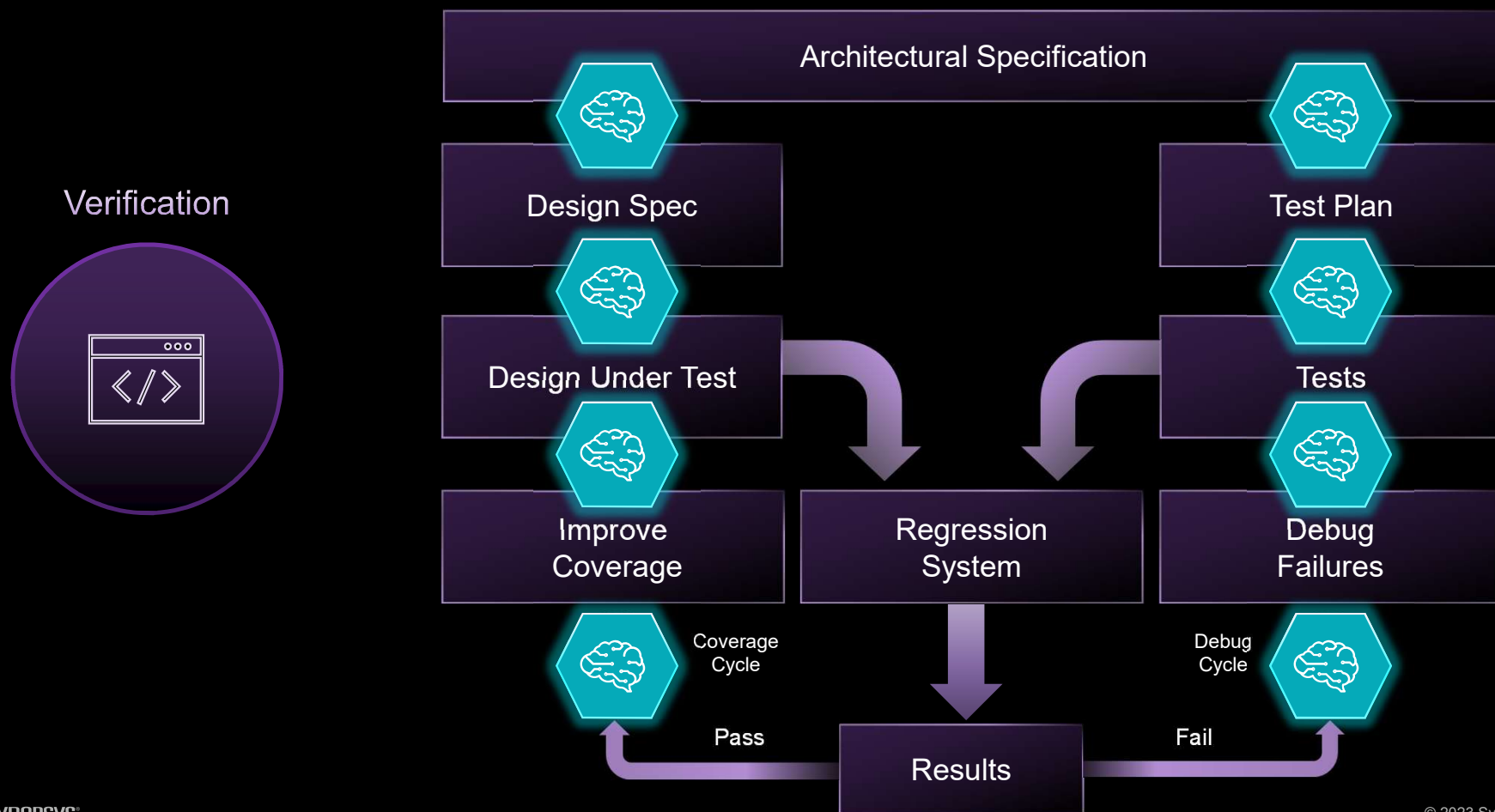
Chip Design Complexity, Cost and Labor Shortage Drive the Need for Productivity



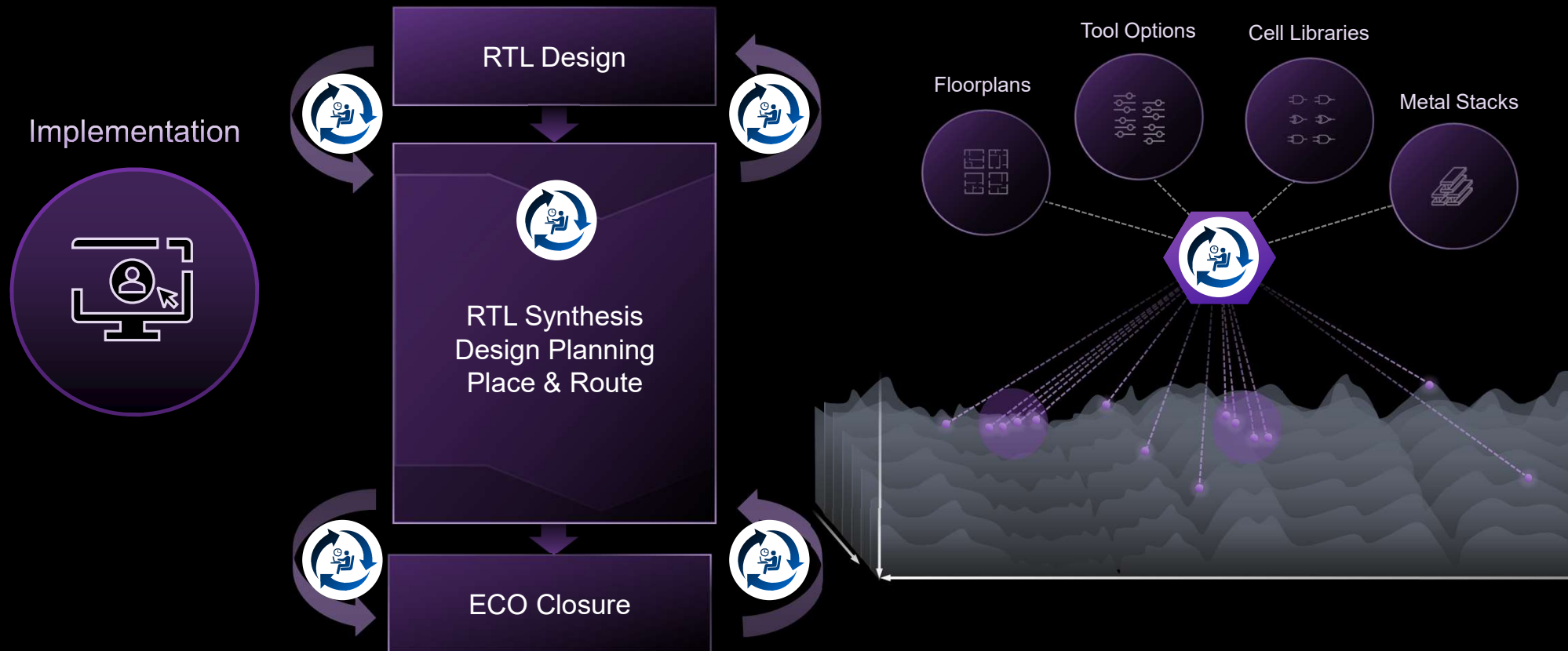
# EDA Workflow Offers Opportunities for AI



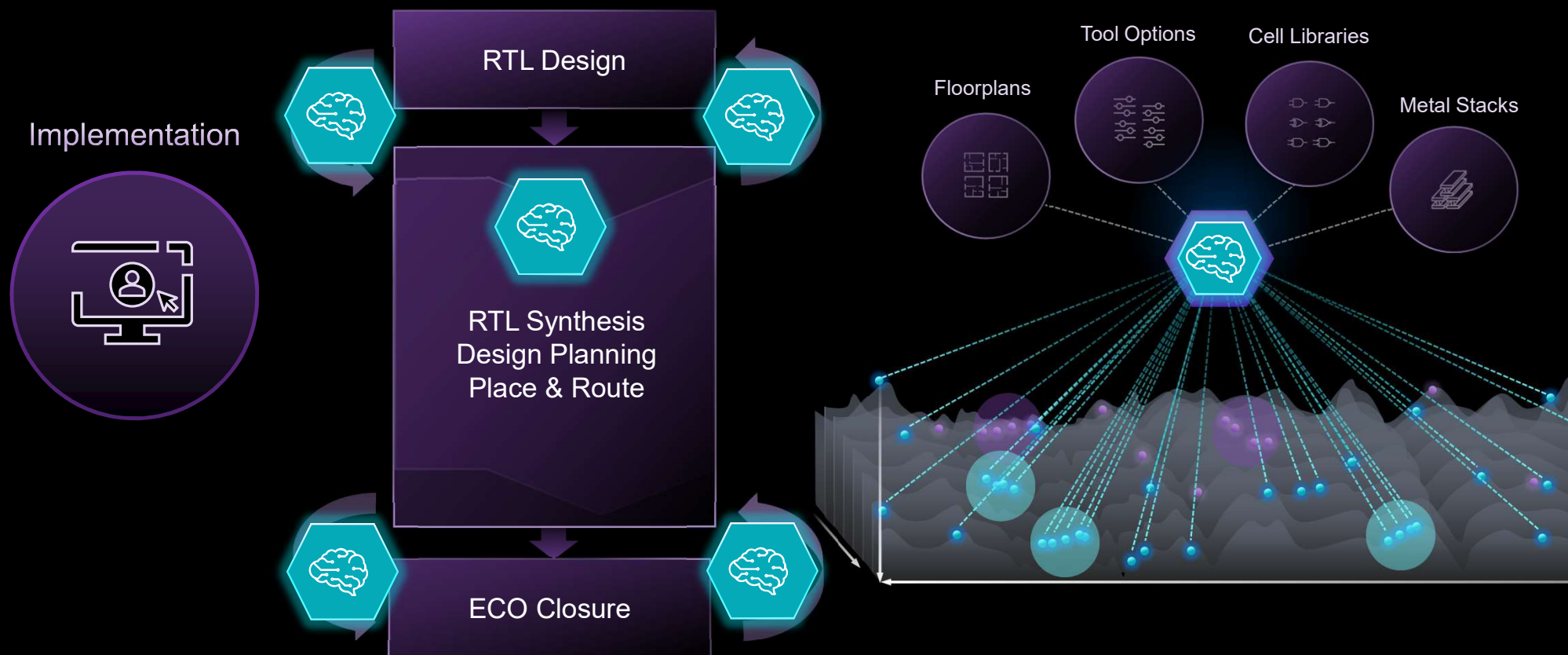
# EDA Workflow Offers Opportunities for AI



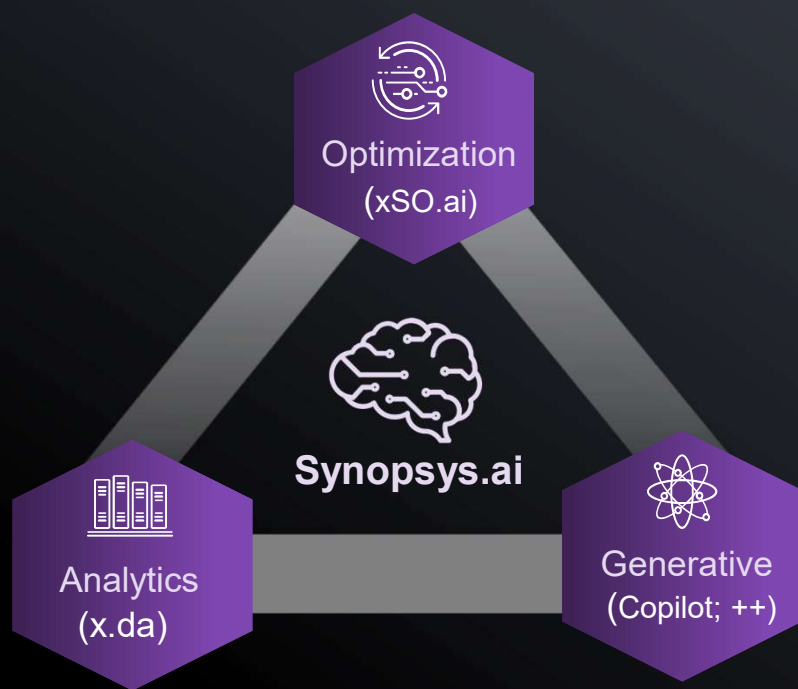
# EDA Workflow Offers Opportunities for AI



# EDA Workflow Offers Opportunities for AI



# Pervasive AI: Synopsys Pioneered AI Driven Chip Design



## AI-Driven Optimization (2020+)

Better, Faster Results

*DSO.ai, VSO.ai, ASO.ai, TSO.ai, 3DSO.ai*

## AI-Driven Analytics (2022+)

Fast Analysis of Data for actionable Insights

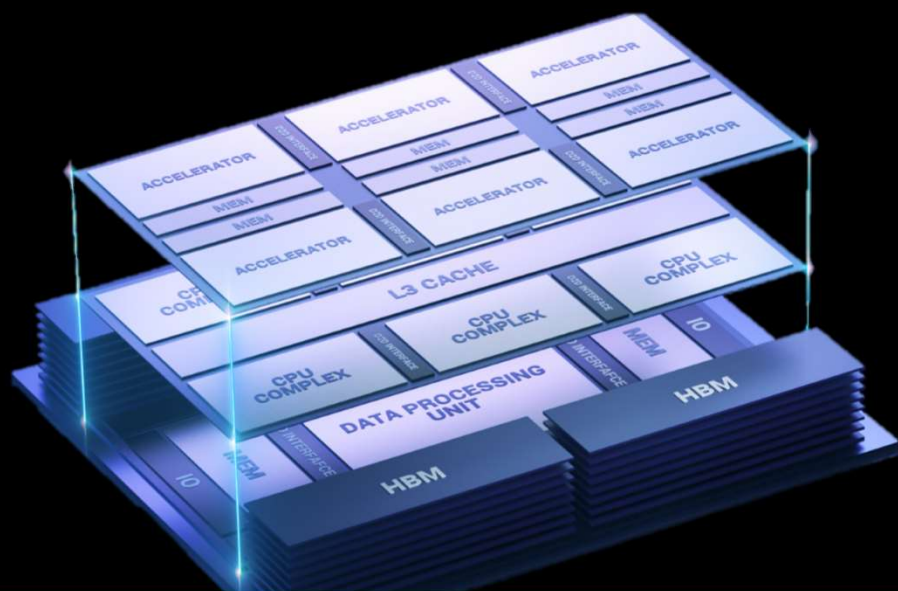
*Design.da, Silicon.da, Fab.da*

## LLM Based Generative AI (2023+)

Powered by LLMs for Collaborative and Generative Capabilities

# Trillions of Devices: Evolving to System(s) OF Chips

Highly heterogeneous, multi-die systems



## Motivation for Multi-Die Systems



Accelerated scaling of system functionality at a cost-effective price



Reduced risk & time-to-market by re-using proven designs/die



Lower system power while increasing throughput



Rapid creation of new product variants for flexible portfolio management

## New Dimensions, Expanding Challenges...



# Chip Design Workflows Have Large Search Spaces

Architectural  
Design

Structural  
Design

Logic  
Design

Layout  
Design

System Architecture

Design Capture

Verification

Implementation

Signoff

Test & Silicon Lifecycle Mgmt.

Silicon Manufacturing



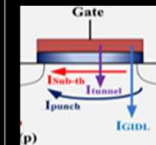
## Performance

- Frequency, WNS, TNS,...
- CTS latency,...



## Area

- Die area, std cell area,...
- Area by cell VT,...



## Power Leakage

- Leakage
- Leakage by cell type, by VT,...

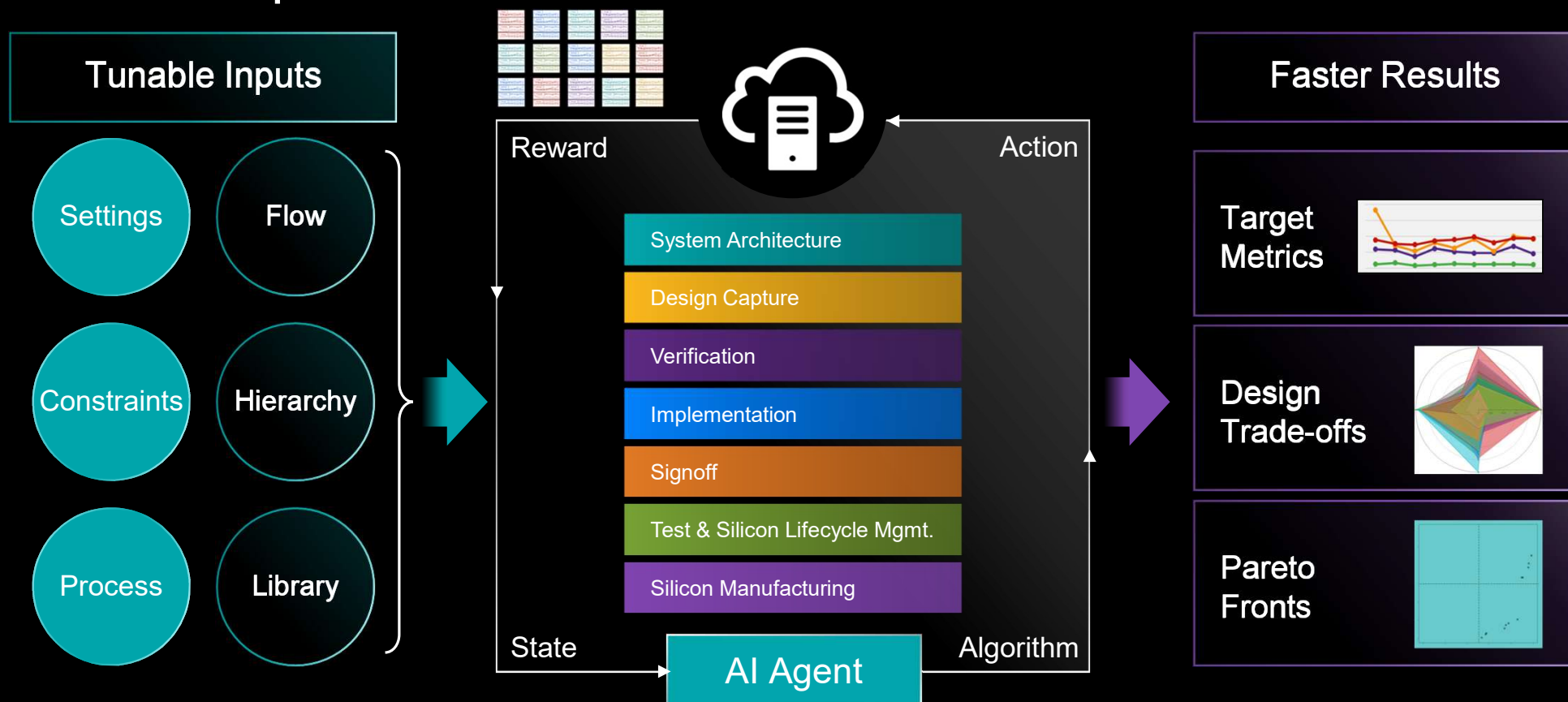
Source : SV SNUG '23 Proceedings

Example :  $10^{25}$  for latest Arm CPU

## Can AI Help?



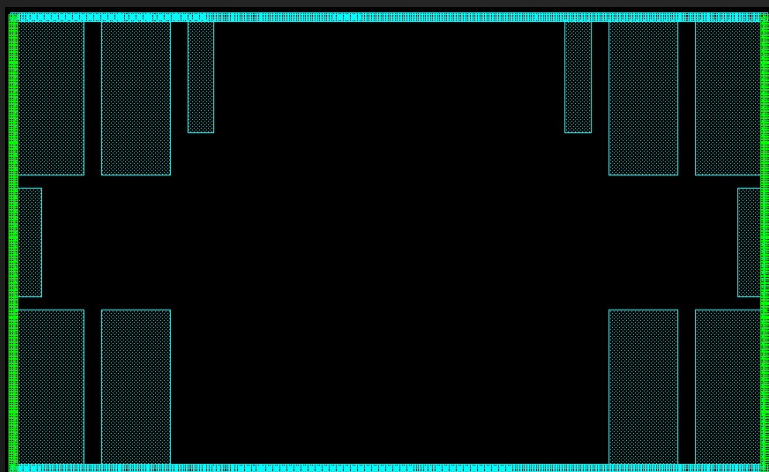
# Digital Implementation: Applying AI to Navigate Chip Design Solution Space




# Digital Implementation: High Performance CPU Core

## AI-driven Optimization Results

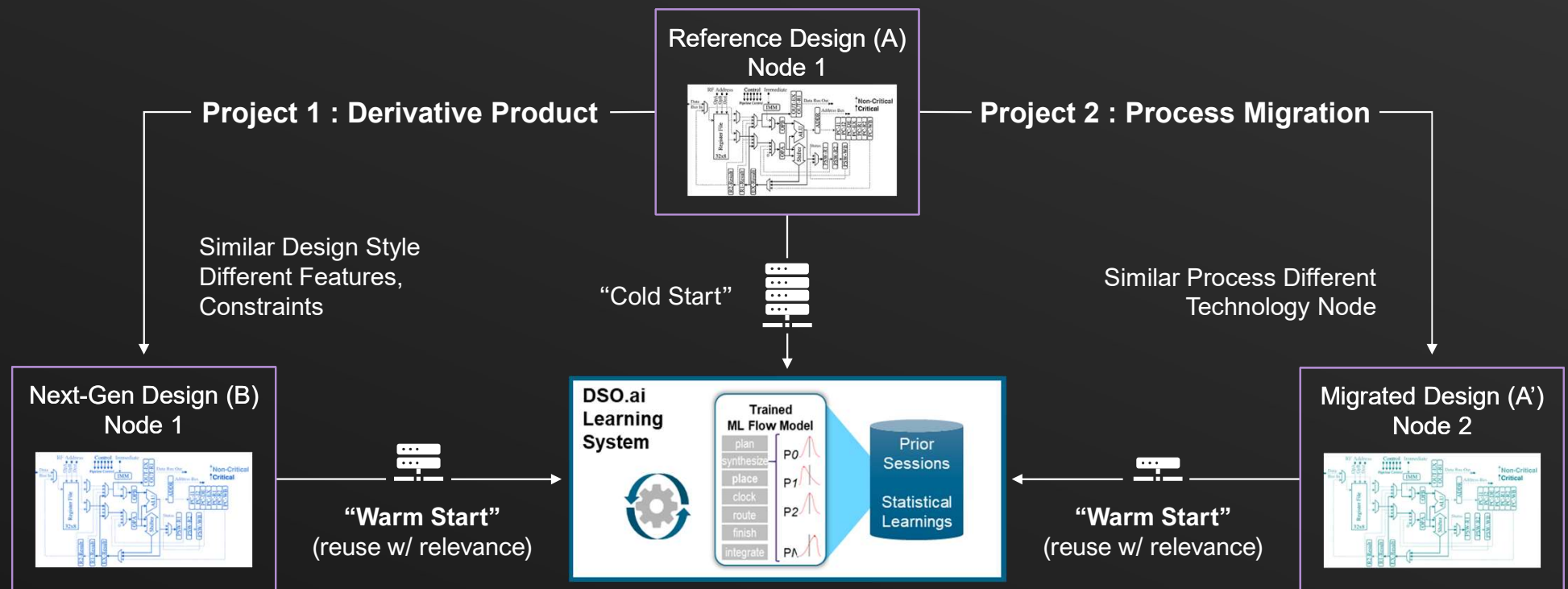
- RISC-V based “Big Core” targeted for data center applications
- Size: 426um x 255um (single core only)
- Technology process: 5nm



	Perf.	Power	Area
<b>Target</b> User Expectation	1.95Ghz	30mW	MET
<b>Baseline</b> OOTB RISC-V Reference Flow	1.75Ghz	29.8mW	MET
 <b>2 days, 90 runs, 0 human!</b>			
<b>DSO.ai</b> AI-Driven RISC-V Reference Flow	<b>1.95Ghz</b>	<b>27.9mW</b>	<b>MET</b>

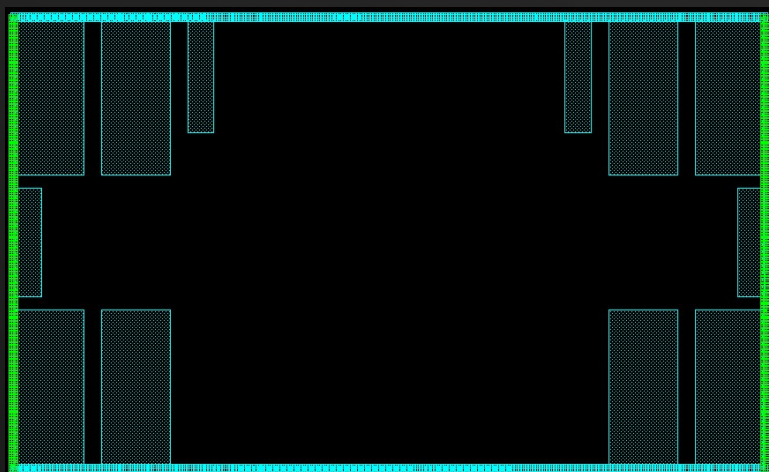
# Digital Implementation: Cross-Design Learning for Project Reuse

## Boost Design Team Productivity, Improve Compute Efficiency



# Digital Implementation: Node Migration using AI

- RISC-V based “Big Core” targeted for data center applications
- Size: 404um x 242um (10% shrink)
- Technology: 5nm → 4nm



	Perf.	Power	Area
<b>Target</b> User Expectation	2.1Ghz	30mW	-10%
<b>Baseline</b> OOTB RISC-V Reference Flow	1.85Ghz	28.4mW	-10%
<b>DSO.ai</b> AI-Driven RISC-V Reference Flow	2.15Ghz	29.4mW	-10%



1 days, 15 runs, 0 human!

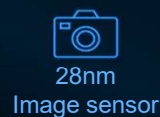
# What's next in Digital Implementation?

## EDA R&D perspective

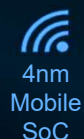
- Design Learning Driven Auto-Convergence
  - Graduate from point AI/ML applications
  - EDA tools become a fully learning-driven system to drive end-end convergence
- Auto-convergence to target different user personas
  - Virtual designers
  - “Green” designers
  - “Expert” designers

# Market Leaders Realizing Significant Gains from Synopsys.ai

DSO.ai



12%  
Area  
Shrink



25%  
Lower  
Power

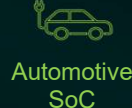


20%  
Faster  
TAT

VSO.ai



2X  
Tests  
Reduction



15%  
Coverage  
Boost

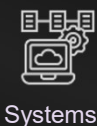


2X  
Faster  
TAT

TSO.ai



45%  
Pattern Count  
Reduction



18%  
Pattern Count  
Reduction



25%  
Pattern Count  
Reduction

ASO.ai



10%  
Area Optimization



2X  
Analog IP  
Node Migration



>5X  
TAT reduction  
Analog Circuit Opt



# AI-Driven Data Analytics Improving Productivity and QoR

Unified Manufacturing Analytics

## Design.da



Hyperscaler

**60%**

Better  
TNS

**20%**

Lower  
Leakage

## Fab.da



Automotive

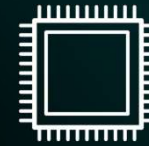
**80%**

Lower  
False  
Alarm

**12K Hrs.**

Equipment  
Time  
Save / Yr.

## Silicon.da



CPU

**20<sub>sec</sub>**

Peta Byte  
Data  
Rendering

**4.2<sub>B</sub>**

Die  
19K  
wafers

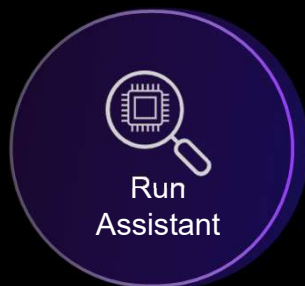


# GenAI-based EDA Apps Driving Wide Scale Productivity



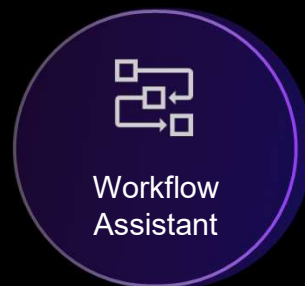
Knowledge  
Assistant

Answer expert questions  
on tools & workflows



Run  
Assistant

In-context Analysis and  
Design Debug



Workflow  
Assistant

Prescriptive Guidance &  
Workflow recommendation



Collateral  
Generation

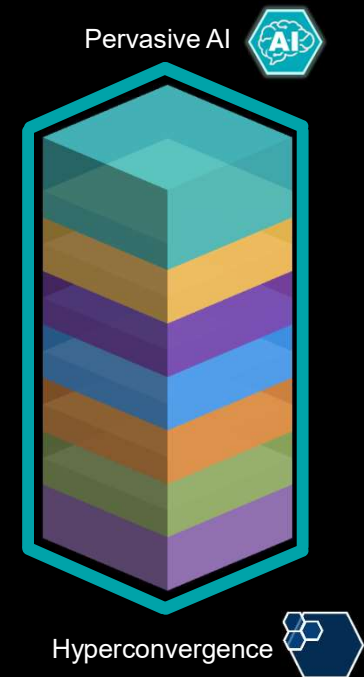
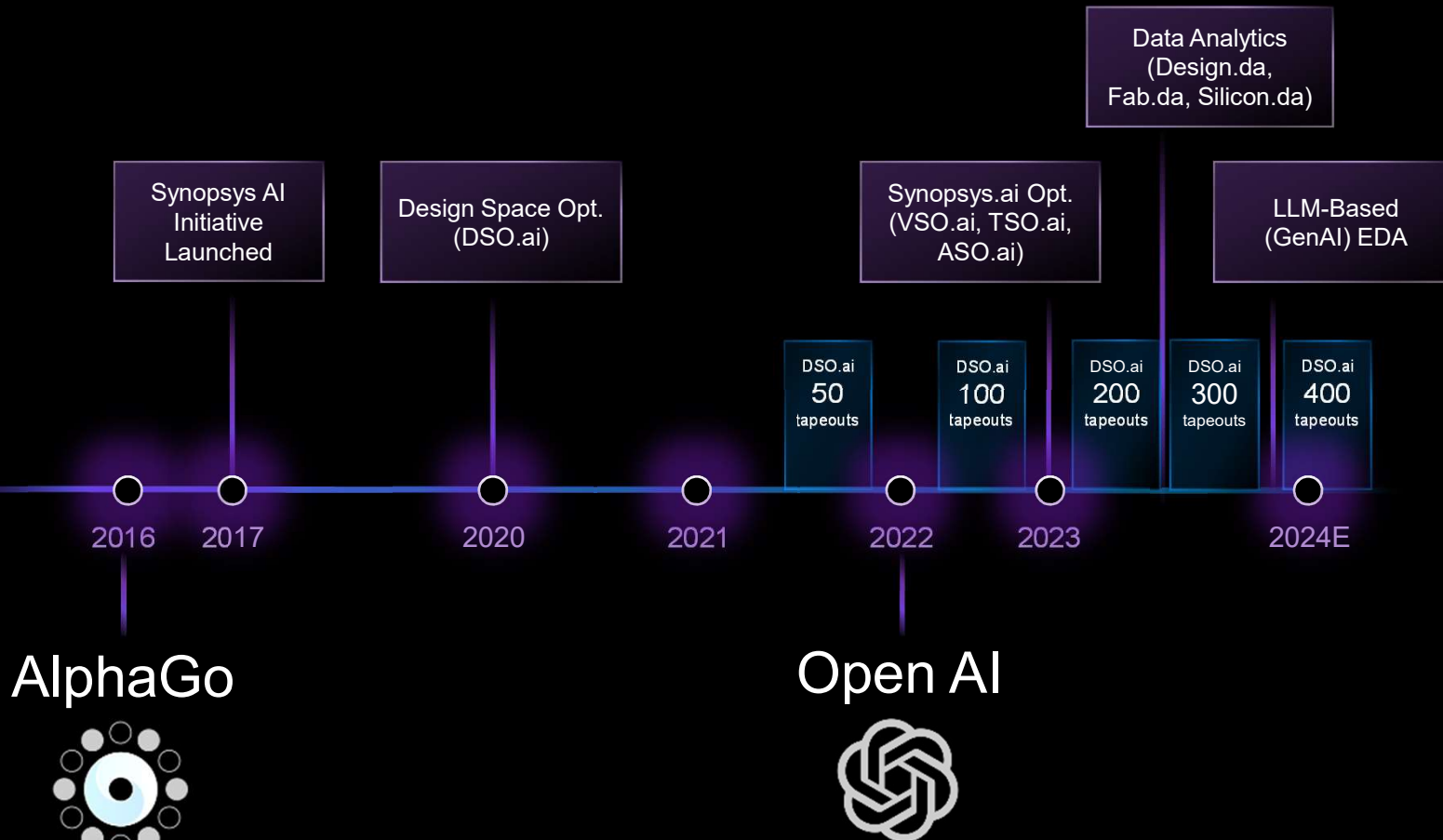
Design / Verification Collateral  
Gen: RTL, Testbench, Assertion

“30% faster ramp-time for junior engineers without having to depend on expert engineers”

“We can focus on the critical tasks while GenAI is taking care of the mundane stuff”

“The responses are at least 2x faster to expert queries than the search process”

# Blistering Pace of AI Innovations





**AI,**  
The Only Way  
Forward

Thank You