Semiconductors – AIG Spring 2026

Molly Liu, Kris Deng, Derek Sun, Edward Liang, Matthew Choe
March 21st, 2026

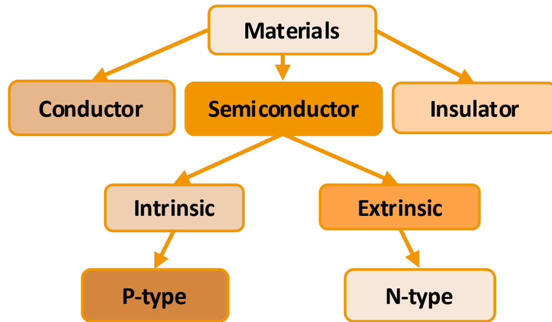
VP: Joseph Yu
Lead Mentor: Oliver Zhao

What is a Semiconductor?

A semiconductor is a material whose electrical conductivity can be precisely controlled, making it the foundation of all modern electronics

What makes a semi conductor

Semiconductors are made of silicon, a material that can be engineered to turn electricity on and off on demand. This precise control is what powers every smartphone, AI system on the planet.

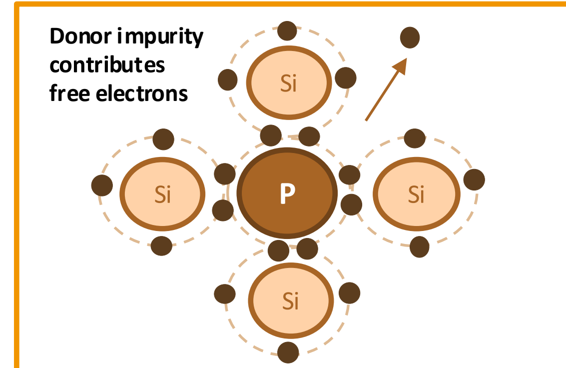


Silicon is the only one that matters for semiconductor

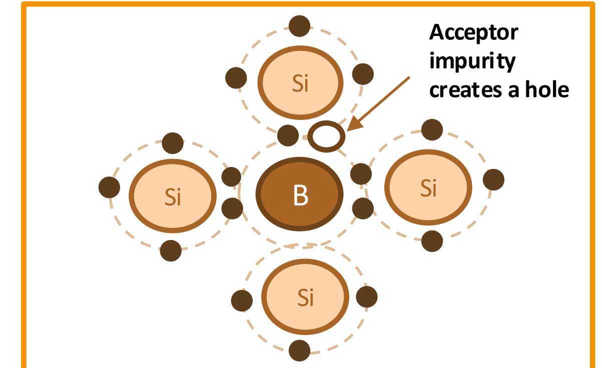
- **Cheap & Abundant:** Earth's Most common element
- **Thermally Stable:** Withstands ~1400 C conditions
- **Programmable Conductivity:** Only material whose conductivity can be precisely engineered by design

Metal *E.g. Copper*	Si *Silicon*	Glass *E.g. Rubber*
Always On	Switchable	Always off

How Doping Works

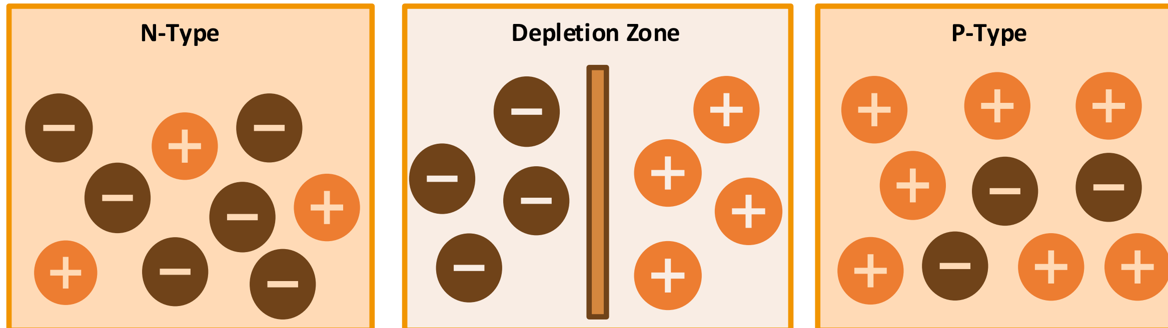


Phosphorus (5e⁻) → 1 free electron → negative carrier (~10^{15} e⁻/cm³)



Boron (3e⁻) → missing bond = hole → positive carrier (~10^{15} holes/cm³)

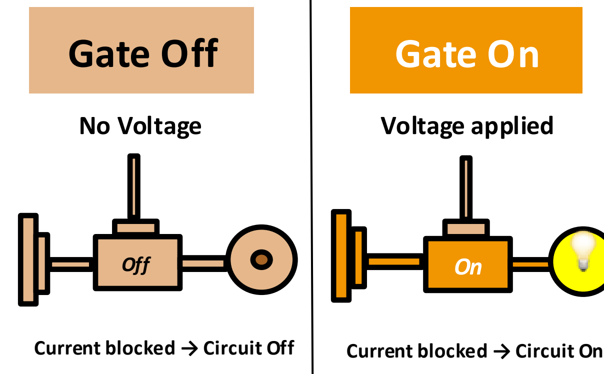
What is a PN Junction?



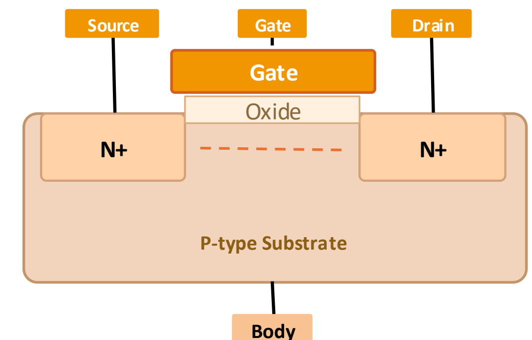
When N-type meets P-type silicon, electrons fill the holes at the boundary, creating a depletion zone, a natural barrier with a built-in electric field. This barrier is the foundation of every transistor.

How a Transistor Works

A transistor is a switch – controlled by a voltage



MOSFET Transistor



Source: Britannica, Semiconductor Fundamentals; Intel Corporation, How Transistors Work; Texas Instruments, Semiconductor Device Fundamentals; Kasap, S.O., Principles of Electronic Materials and Devices; TSMC Technology Overview; Sedra & Smith

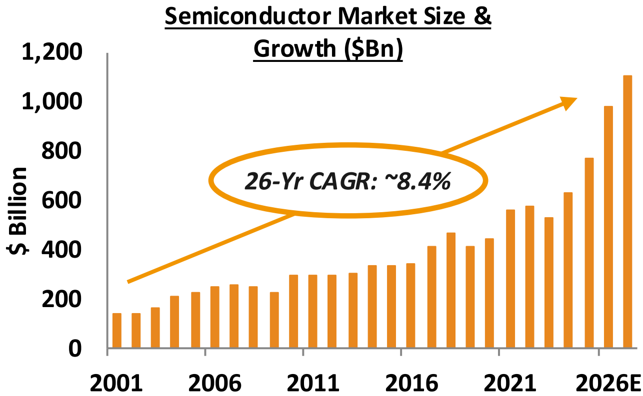
AIG ASCEND INVESTMENT GROUP

Market Outlook & Moore's Law

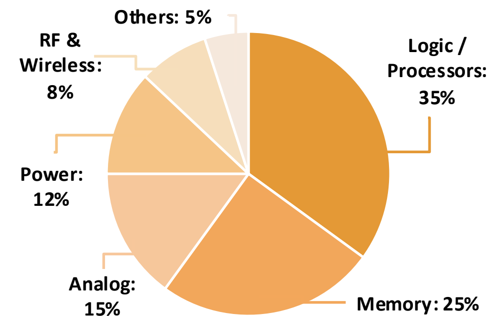
50 Years of Shrinking Transistors Built a $1T Industry. What Happens When They Can't Shrink Anymore?"

Market Outlook

Semiconductor Market Size & Growth ($Bn)



Semiconductor Markets by End Market (2024)



Logic	Memory	Analog	Power	RF & Wireless	Others
Processes data and runs all computing and AI workloads. Fastest-growing segment. (Ex. GPU, CPUs, AI Chips)	Stores and retrieves data at high speed. Highly cyclical with boom & bust cycle. *(Ex. DRAM, NAND flash)*	Converts real-world signals into digital data. Critical for EVs and IoT. *(Ex. Sensors, mixed-signal ICs)*	Regulates power flow in devices and EVs. Growing with electrification. *(Ex. Power Management ICs)*	Transmits radio signals enabling wireless connectivity. Driven by 5G expansion. *(Ex. 5G modems, RF chips)*	Specialty chips for niche markets like medical devices, industrial, and aerospace. *(Ex. Optoelectronics, MEMS)*

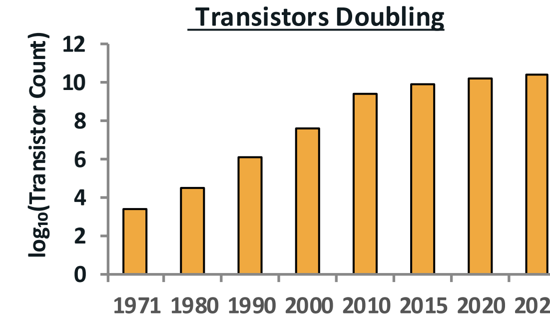
▲ Tailwinds	▼ Headwinds
AI infrastructure spending boom + 5G global rollout	Boom and bust inventory cycles, high cost
EV adoption driving chip demand	Fab cost increasing to around $20-$30 billion
CHIPS Act government investment	Customer concentration & geopolitical risk

Moore's Law and Scaling

Gordon Moore's Observation (1965)

Transistor density doubles every ~2 years while costs fall—guiding the industry for over 50 years.

Transistors Doubling



↘ Cost per Transistor — What Did It Actually Cost?

1971	1990
~$1.00	~$0.0001
1 transistor = $1	*10,000x cheaper*

2005	2023
~$0.000001	~$0.000000000001
1,000,000x cheaper	*100 billion x cheaper*

100 billion x cheaper in 52 years powered by Moore's Law

Key Milestones

1971	1993	2012	2023
Intel 4004	Intel Pentium	Ivy Bridge	Apple M3 3nm
2,300 transistors	3.1 million	1.4 billion	25 billion

Physical limits

- Transistors approaching atom-scale (~2nm)
- Quantum tunneling causes leakage at <5nm
- Heat density rising exponentially
- Cost per transistor no longer falling

Beyond Moore's law

- 3D chip stacking (HBM, TSMC SoIC)
- Chiplets & advanced packaging
- AI-specific accelerators (GPUs, TPUs)
- Photonics & quantum chips (long-term)

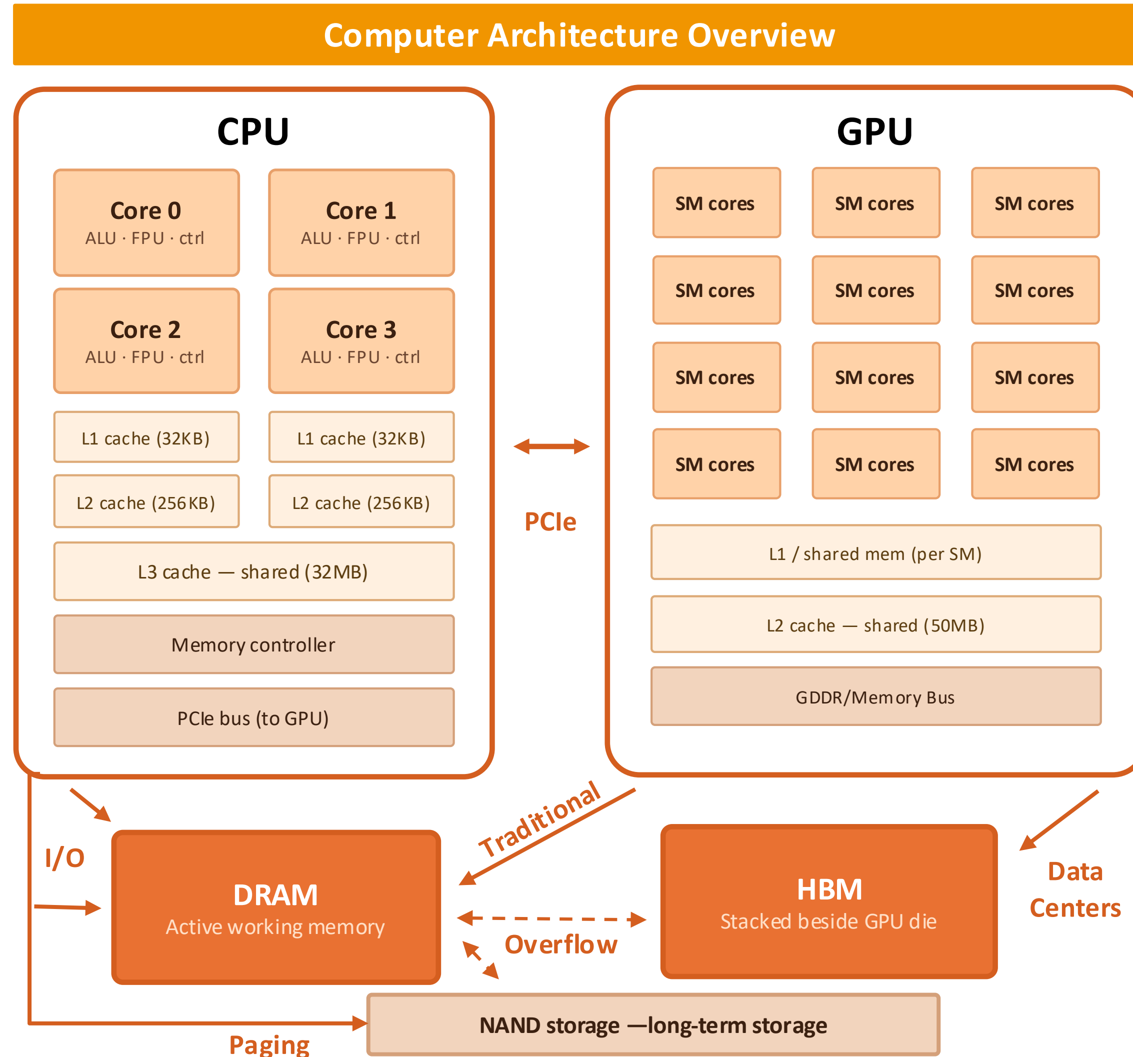
Source: WSTS Autumn 2025 Forecast; SIA 2024 Annual Report; Deloitte 2026 Semiconductor Outlook; Moore, G.E., Electronics Magazine 1965; Our World in Data; Intel, 50 Years of Moore's Law; McKinsey & Company 2024

ASCEND INVESTMENT GROUP

Semi-conductors: Generalized Chips

Semiconductors in Computer Architercture and the Interactions between the Internal Computer Architecture

Computer Architecture Overview



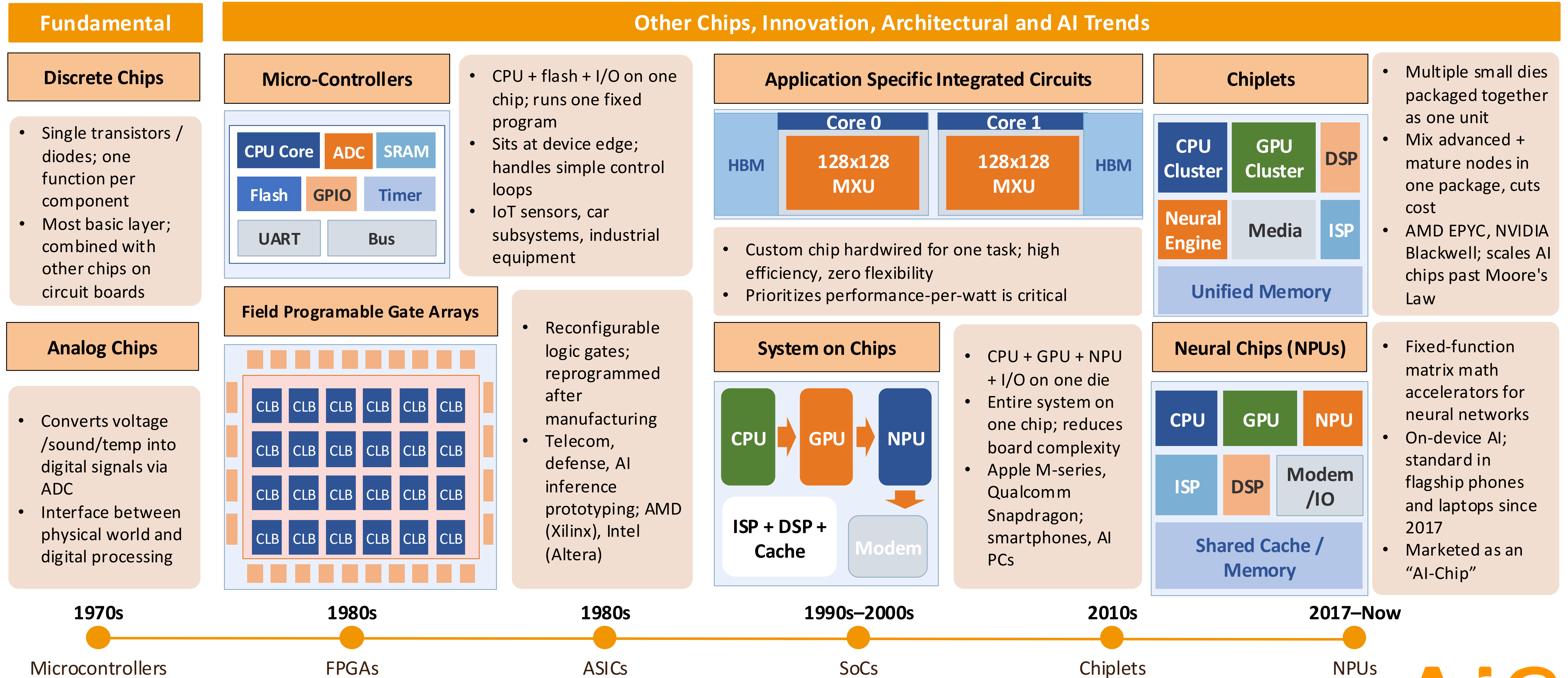
Different Types of Chips

	What is it?	How it works	Applications to AI
Central Processing Unit (CPU)	*Has a small number of powerful, complex cores for general purpose processes*	*Sequential, single-thread performance handles one complex task at a time per core, very fast*	*NEW AI-orientated innovation: NPU cores and chiplets*
Graphics Processing Unit (GPU)	*Streaming Multiprocessors (SMs): Thousands of Small Shader Cores*	***Parallelism**: ability to run thousands of matrix operations simultaneously*	***Matrix multiplication**: multiply huge grids of numbers together, millions of times*
DRAM (Dynamic Random Access Memory)	*Capacitor-based, volatile (loses data when power is off), arranged in rows and columns*	*Data is stored in capacitors **Dynamic**: Capacitors lose charge, so memory is refreshed 1000s/second.*	*AI servers require hundreds of gigabytes of DRAM to feed GPU workloads.*
HBM (High Bandwidth Memory)	*DRAM dies stacked in 3D connected via a silicon interposer, implanted with GPU*	*Higher Bandwidth Lower Latency Lower Capacity*	*Standard DRAM can't feed data fast enough to keep thousands of GPU cores busy during AI training.*
NAND Storage	*Non-Volatile Storage: Traps electrons in floating-gate transistors to store data permanently*	*Stores data as electrical charge in floating-gate cells; programmed and erased in large blocks, trading high-density, persistent storage*	*Massive amounts of AI training datasets rely on NAND/SSDs*

Source: Congressional Research Service (R47508), Semiconductor Industry Association 2025: State of the Industry Report, Investopedia, Stanford University, Santa Clara University, Google Cloud, StoredBits,

Semi-conductors: Chip Innovation

The Evolution from Analog Discretes and Microcontrollers Through ASICs, SoCs, and Chiplets to NPUs and Chiplets

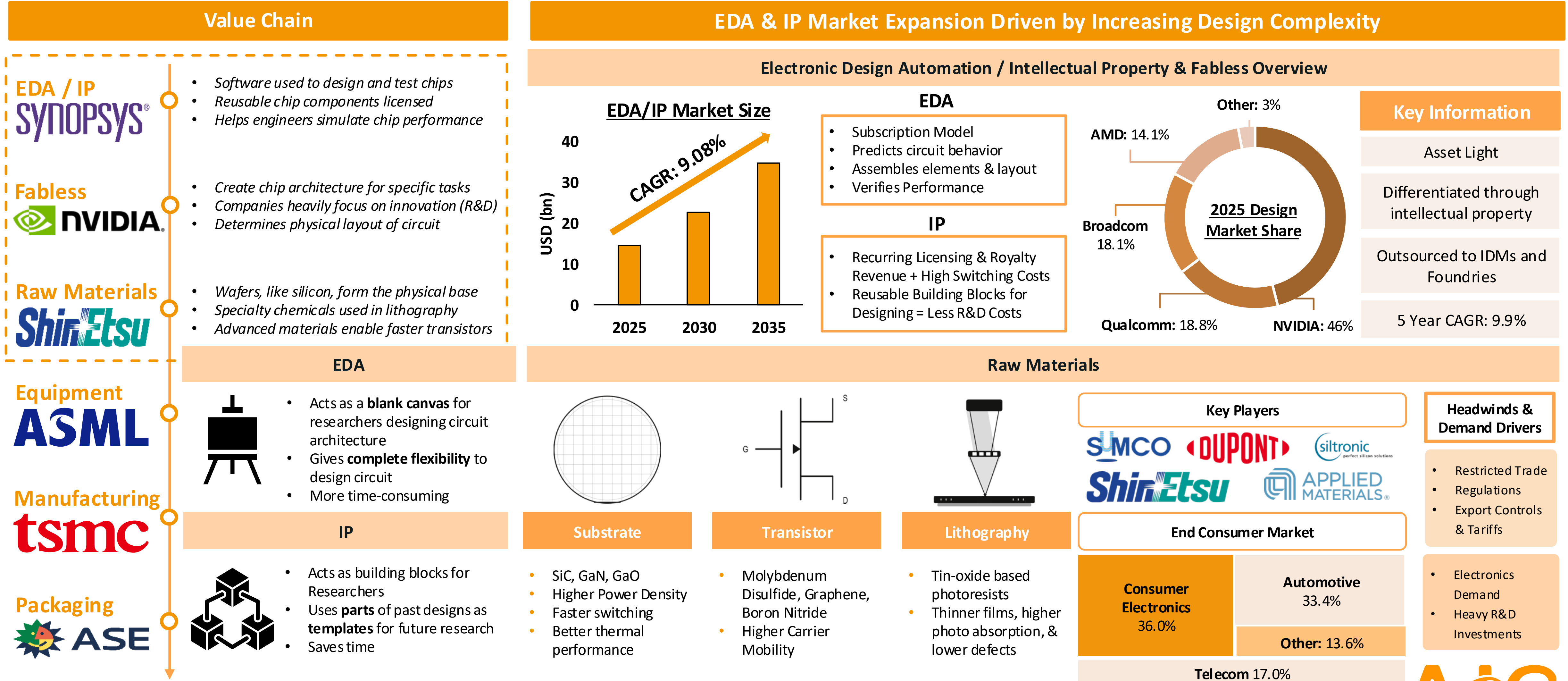


Source: Congressional Research Service (R47508), Semiconductor Industry Association 2025: State of the Industry Report, Investopedia, Stanford University, Santa Clara University, Google Cloud, StoredBits,

AIG ASCEND INVESTMENT GROUP

The Semiconductor Ecosystem: Idea to Integrated Circuit

Semiconductors Have a Very Broad Value Chain with Different Business Models and Different Valuation Methods, Changing Analysis



Sources: Refinitiv LSEG Workplace, Factset, Boston Consulting Group, Semiconductor Industry Association

AIG
ASCEND INVESTMENT GROUP

The Semiconductor Ecosystem: Idea to Integrated Circuit

ASML's Pricing Power Through Its Monopoly of Raw Materials Has Cascading Effects Across the Broader Semiconductor Industry

The Logistics of Chip Manufacturing

EDA / IP – Synopsys

Fabless – NVIDIA

Raw Materials – Shin-Etsu

Equipment – ASML
- *Machines that print and etch patterns*
- *Lithography systems create designs*
- *Facilitate deposition, cleaning, and layers*

Manufacturing – tsmc
- *Factories produce chips in multi-step process*
- *Transistors built layer by layer on wafers*
- *Requires extremely clean environments*

Packaging – ASE
- *Individual chips cut and enclosed in packaging*
- *Electrical connections link chips to environment*
- *Advanced packaging connects multiple chips*

Extreme Ultra Violent Lithography Market

Customer Demographic

Foundries: 36.1%
IDM: 63.9%

IDM: Intel, Samsung, Texas Instruments

Foundries: TSMC & Global Foundries

ASML	**Customers**
Needs customers for prolonged AI Demand Regulation Risk	Need ASML for PP&E Pre-negotiated Prices Government-backed

Market Structure
- Monopoly (ASML) & High Barriers: $5.3Bn
- Few Key Buyers: TSMC, Intel, Samsung

Key Risks
- Customer Concentration & Export Restrictions
- Cyclical semiconductor CapEx Spending

ASML Stock Price

Stock Value (Euro): 0, 400, 800, 1200, 1600
2018, 2020, 2022, 2024, 2026

AI Demand → EUV Demand → ASML Pricing Power

Packaging

Circuit Application
- **Before:** Circuit works, but is dangerous and vulnerable to operate
- **After:** Protective layers increases durability of circuit and increases user safety

1970s – Dual In-line Package First Yield: 10-50%

1990s – Ball-grid Array First Yield: 99%+

2010s – 2.5/3-D Integrated Circuits: 95%+

Manufacturing

Node Range: 3nm – 7nm
Gate wraps 3 sides of fin structure
Key Players: TSMC, Samsung, Intel

GAAFET (Next Gen – 2nm+)
Node Range: 0nm – 2nm
Gate Wraps All 4 Sides (Nanosheet)
Key Players: TSMC (N2, 2025), Samsung SF2

+30-50% Wafer ASP Uplift TSMC N2 vs. N3	**~2-3x EDA / IP Spend Synopsys & Cadence benefit from complexity**	**$25-30B+ CapEx Per Fab Per leading-edge fab (TSMC Arizona)**
TSMC Now: 30,000 wafers/month 3x by 2027	10-15% Speed Increase, 25-30% Power Reduction, 1.15x Chip Density (N3E)	AMD, NVIDIA, Apple

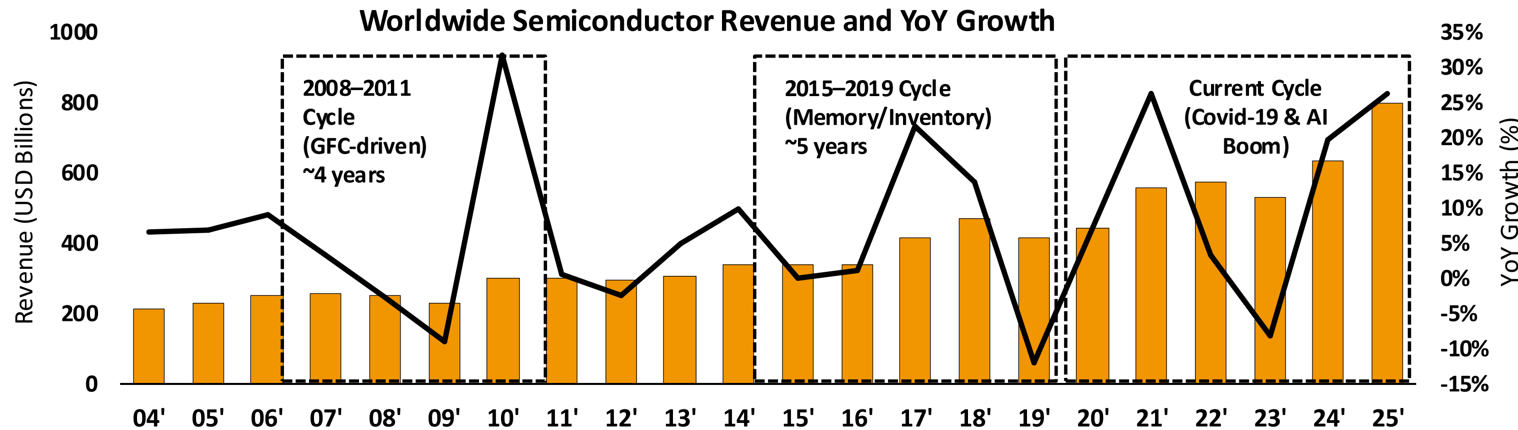
Sources: Synopsys Investor Relations, Jefferies, Bloomberg, NVIDIA Investor Relations, AMD Investor Relations, TSMC Investor Relations, ASE Technology Holding Investor Relations, ASML Holding Investor Relations, Shin-Etsu Chemical Investor Information

ASCEND INVESTMENT GROUP

Semiconductor Cycles: Why Revenue Swings Are Structural

Revenue volatility reflects a recurring mismatch as demand resets quickly, but supply responds with a lag

Worldwide Semiconductor Revenues



Has the current cycle ended?

Shortage 20'-21' — Remote work, online learning, & cloud demand surged. Supply chains were disrupted, creating broad shortage

Correction 22'-24' — Demand cooled and inventories built. Destocking drove the downturn, before AI led a partial recovery

Recovery 25' Beyond — AI remains strong, but the broader market is still mixed. The market is debating a AI-led **supercycle** vs. a **late-cycle rebound**

Why These Cycles Repeat: Fast Demand, Slow Supply

Demand moves faster

- Shortages leads customers **order early**
- Customers tends to **"double order"** to secure supply
- When demand cools -> **cancellation** and hits

Demand Spike → Order Pull-Over → Destocking → Demand Normalize

Adjusts in quarters

Demand & Supply Mismatch

This timing gap creates shortages and oversupply

Supply responds with a lag

- New fabs can require **$15B+** of capex
- Capacity often needs **18–24 months** to ramp
- High fixed costs make output slower to adjust

Capex Decision → Fab Build → Tool Installation → Yield Ramp → Capacity Hits Markey

Adjusts in years

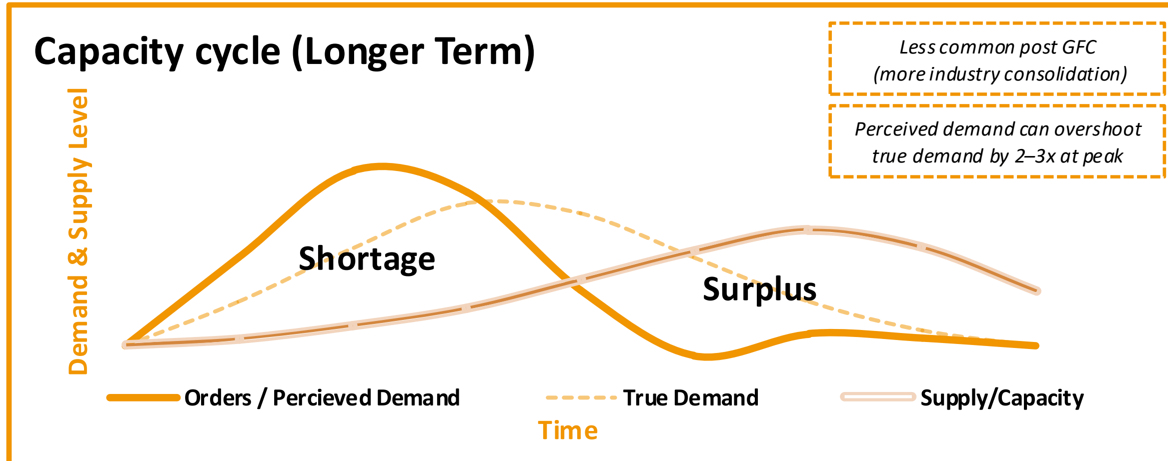
Sources: Nomad Semis; Deloitte; WSTS; EY Report

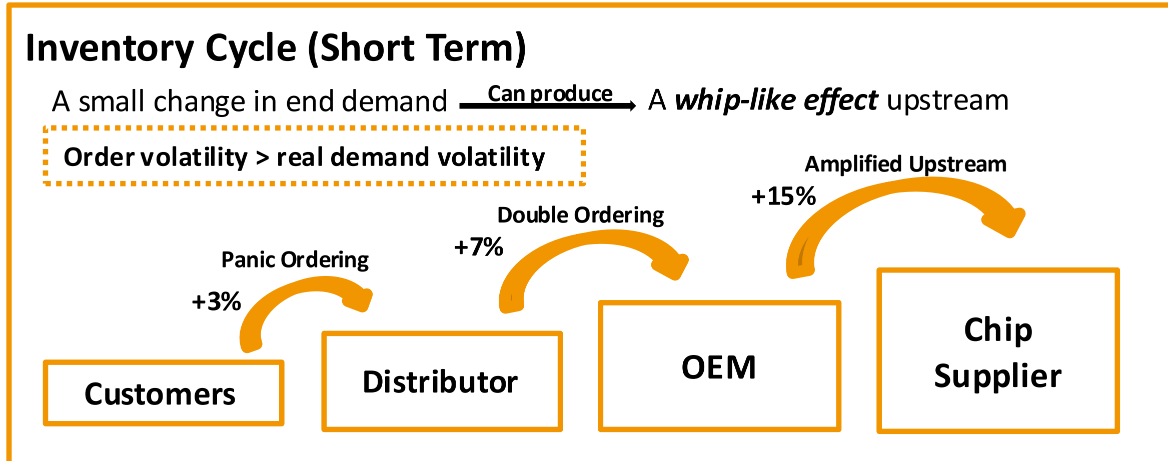
ASCEND INVESTMENT GROUP

Semiconductor Downcycles: Main Structural Drivers and a Covid-19 Case Study

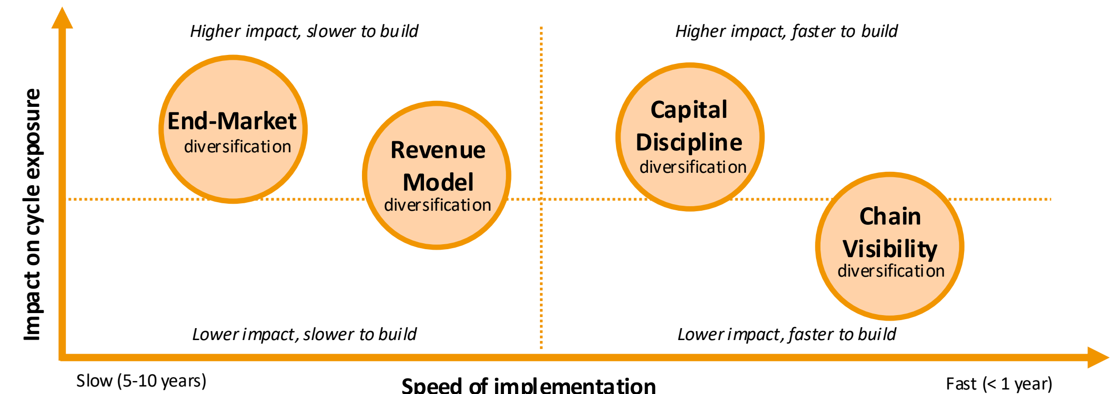
Semiconductor downturns deepen when capacity additions meet weaker end demand and inventory destocking

The 2 Main Drivers of Semis Downcycle





How Companies Reduce Cyclical Risk



Strategy	Why It Works	Company Example
End-Market Diversification	Different markets peak/trough at different times	TI: 10+ years building auto/industrial mix
Capital Discipline & Countercyclical Investing	Downturns = lower costs, less competition for capacity	TSMC/ASML invest in corrections, emerge with share
Revenue Model Diversification	Software can't be double-ordered or destocked	Broadcom: software cushions hardware downturns
Supply Chain Visibility & Multi-Sourcing	Catches inventory buildups 2–3 quarters early	Qualcomm dual-sources TSMC + Samsung by design

Sources: WSTS; EY Report; Semiconductor Engineering

AIG ASCEND INVESTMENT GROUP

Current Trends in the Semiconductor Industry

AI-driven computing is reshaping semiconductor design, highlighting the renewed role of CPUs and the rise of advanced packaging

Trend 1: The Resurgence of CPU

Dominance of GPU

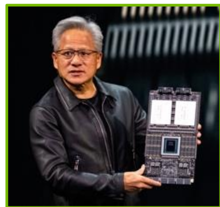
- Accounts roughly 50% of total power consumption in AI data centers
- $10.5bn~$77bn by 2035, 22%CAGR
- Holds 62% of revenue share in global AI accelerator market

CPUs Overshadowed by GPU Growth

- CPU: Sequential processing, limited cores.
- GPU: Parallel processing, abundant cores
- AI computing demands matrix multiplication, parallel processing

CPUs regaining importance Inside AI system

1. CPU limiting GPU performance
2. Need for central coordination, especially data feeding
3. Shift from training (GPU) to inference (CPU)



Standalone launch of Vera Rubin CPU in Q1 2026, validated with the massive multi-year deal with Meta

Vera Rubin NVL72 Rack System

Vera CPU	Rubin GPU	NVlink 6 Switch	BlueField – 4 DPU

Intel re-architecting its strategy

Collaboration with NVIDIA to build a custom Xeon with integrated NV Link, allowing x86 cores for Blackwell and Rubin clusters

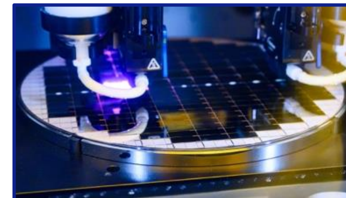
1. Faster communication between x86 processor and large GPU clusters (NVLink 10x faster than traditional PCIe)
2. Intel seeking to keep data-center share

Trend 2: How Advanced Packaging is Evolving

*Traditional packaging relies on **shrinking transistors on a single chip** to boost performance. Scaling slows (Moore's Law), advanced packaging connects multiple chips to achieve higher performance and flexibility beyond single-chip limits.*

Flip Chip	Wafer Level	Memory 3D TSV	Panel Level
Chips flipped upside down by solder bumps	*Packaging done on wafer before cut into dies*	*Electrical Vertical Connection Through Silicon*	Uses large rectangular panels instead of wafers
Subtypes			Enables higher throughput and mass production
• Chip Scale • Ball Grid Array • Package on Package	• Fan-In: Stay inside chip • Fan-out: Spread outside	• Non-Conductive Film • Liquid Compression Molding	Reduces cost per unit and improves efficiency

ASML has introduced the Twinscan XT: 260, the industry's first lithography scanner built specifically for advanced 3D chip packaging



XT 260 Scanner

Monopoly in EUV lithography machines

XT 260: 400 nm Average: 1600 nm

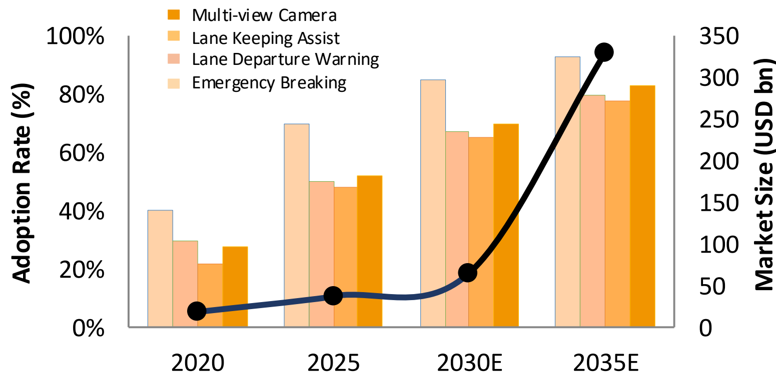
- Dual-stage platform (Exposure & Alignment)
- Up 4x higher productivity
- Throughput of 270 wafers p/h

Sources: NVIDIA, Intel, ASML, McKinsey & Company, Deloitte, Boston Consulting Group, Semiconductor Industry Association (SIA), International Energy Agency (IEA), TSMC, Samsung Electronics

ADAS Evolution: From Feature Growth to System Architecture Transformation

ADAS growth and standardization are driving a shift toward unified workflow and centralized, high-performance vehicle architectures

Trend 3: The Rapid Growth and Standardization of ADAS



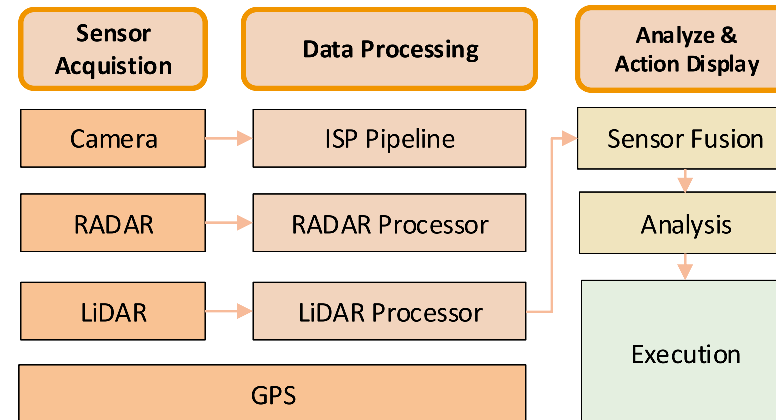
Standardization

- Government safety regulation
- Differentiation Strategy

Rapid Growth

- Advances in sensors and AI chips
- Cost control

The Workflow of ADAS



Sensors

Radar: Radio wave detection
LiDAR: Laser pulses to build 3D map

Processor

ISP: Cleans and enhances raw camera data
Processors: Convert and combine data

CAN and Ethernet in ADAS

Controller Area Network(CAN):
For short, control messages between ECUs
- 1 Mbit/s

- Braking, steering, body control
- Reliable, low cost
- Low bandwidth

The Change for Network

- Addition of cameras, radar, sensor fusion and central compute
- Demands faster data transport
- Bandwidth difference + CPU/GPU context

Ethernet:
High-speed in-vehicle net work for massive data

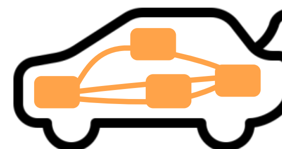
- 100 Mbit/s
- High bandwidth
- Support ADAS backbone

Microcontroller(MCU)

- Control-and-safety chip
- Monitors the system's behavior
- Physically isolated from main SoC
- AURIX TC39xXa, Infineon Tech

Centralized Architecture

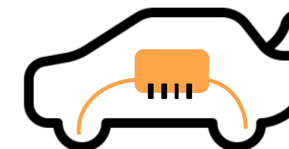
Before Centralization



Multiple electronic control units (ECUs)

- Each has specific roles e.g. braking, steering
- Long wires causing more weight and complexity

After Centralization



Domain Controller (ADAS)

- Sensors like camera, radar
- Demands strong computing
- More overall control. reduces cost

Qualcomm Snapdragon

- **Integrated ADAS SoC** (CPU + GPU + AI)
- Enable **centralized**, high-performance vehicle compute
- **Consolidates** multiple chip functions
- Increases **semiconductor content** per vehicle
- For **centralized automobiles**

Sources: Qualcomm, NVIDIA, Intel, McKinsey & Company, Deloitte, Boston Consulting Group (BCG), International Energy Agency (IEA), Statista, Fortune Business Insights, Britannica

Q&A