·汽车芯片技术专题·

汽车AI芯片发展趋势研究

刘洋　王强　鲍瑞　梅文超　张凯

（中国第一汽车股份有限公司研发总院，长春 130011）

【欢迎引用】刘洋，王强，鲍瑞，等．汽车AI芯片发展趋势研究[J]．汽车文摘，2026(9)：13-21.

【Cite this paper】LIU Y, WANG Q, BAO R, et al. Research on the Development Trend of Automotive AI Chips[J]. Automotive Digest (Chinese), 2026(9): 13-21.

【摘要】为了系统掌握汽车AI芯片产业发展现状、破解技术发展痛点、把握行业未来演进方向，助力国内汽车AI芯片产业自主化、高质量发展，聚焦AI芯片在智能网联汽车中的智能驾驶感知、决策、控制及智能座舱交互等核心应用场景，系统梳理智能网联汽车对AI芯片高算力、低功耗、高安全、高可靠等核心需求。同时分析国内外高通、NVIDIA、特斯拉、地平线、芯擎等主流汽车AI芯片企业的技术布局与产品迭代，结合当前行业发展现状，总结提炼了汽车AI芯片在内存带宽、封装工艺等方面存在的核心技术瓶颈与解决方法。在此基础上，梳理出汽车AI芯片在域融合、异构架构集成、算力增长、存算一体技术落地、先进封装等维度的未来发展趋势，以期为车载AI芯片技术研发、产品迭代、产业布局及行业相关研究工作提供参考与借鉴。

关键词：汽车AI芯片；域融合；存算一体；芯片算力；先进封装

中图分类号：U463.61　　**文献标志码：**A　　**DOI:** 10.19822/j.cnki.1671-6329.20240293

Research on the Development Trend of Automotive AI Chips

Liu Yang, Wang Qiang, Bao Rui, Mei Wenchao, Zhang Kai

（Global R&D Center, China FAW Corporation Limited, Changchun 130013）

【Abstract】To clarify the current state of the automotive AI chip industry, address technological pain points, and identify the future development trends of the industry, while supporting the domestic automotive AI chip industry's independent and high-quality development, this focuses on the core application scenarios of AI chips in intelligent connected vehicles, such as smart driving perception, decision-making, control, and smart cockpit interaction. It systematically reviews the core requirements of intelligent connected vehicles for AI chips, including high computing power, low power consumption, high security, and high reliability. At the same time, it analyzes the technology layouts and product iterations of mainstream domestic and international automotive AI chip companies like Qualcomm, NVIDIA, Tesla, Horizon, and SiEngine Technology, and, based on the current industry situation, summarizes the main technical bottlenecks and solutions for automotive AI chips in areas like memory bandwidth and packaging processes. On this basis, it assesses the future development trends of automotive AI chips in terms of domain fusion, heterogeneous architecture integration, computing power growth, the implementation of computing-storage integration technology, and advanced packaging. The goal is to provide a reference and inspiration for the technological research and development, product iteration, industrial layout and relevant research of vehicle-mounted AI chips.

Key words: Automotive AI chips, Domain fusion, Computing-In-Memory (CIM), Chip computing power, Advanced packaging

0　引言

AI芯片是智能汽车核心零部件[1]，其应用场景包括但不限于智能驾驶的感知处理、决策规划以及座舱的智能交互等[2]。汽车架构正经历着从分布式向域控制及中央集中式的深刻转型[3]，域融合已经成为重要的发展趋势。传统AI芯片可能仅针对单一特定的应用场景，如智驾域或座舱域进行设计。然而，随

着技术进步和需求多样化，智能驾驶与智能座舱的信息交互和协同工作愈发紧密。因此，AI芯片设计也开始向域融合的方向发展，旨在实现更加高效、智能的车辆协同控制与信息娱乐功能。同时，为了保障智能网联汽车的安全行驶和用户的隐私权益，AI芯片还应有更强的安全性，能够抵御各种潜在的安全威胁[4]。此外，高度的可靠性也是汽车AI芯片不可或缺的重要特性，以确保芯片在各种极端环境下稳定运行。

然而，AI芯片应用于汽车行业也面临着一些技术瓶颈。数据传输瓶颈是主要问题之一，随着晶体管数量增加，数据在芯片内部和芯片间移动的成本迅速增加，传统的芯片架构难以满足现代需求。为了解决这些问题，高效的内存架构（如存算一体）应运而生，但同时也增加了设计的复杂性。此外，随着摩尔定律的尺寸微缩趋势放缓，先进封装将在不缩小制程节点的条件下，通过改进封装方式继续提升芯片性能。针对以上问题，学术界也展开了深入研究。如文献[5]阐述了不同存储介质下的存算一体实现方案，文献[6-7]说明了存算一体技术架构在大语言模型和图像解构计算等前沿领域的适应性和可行性。此外，文献[8-9]分别阐述了垂直互连关键技术和铜柱凸点互连技术，对先进封装的互联技术进行了探索与研究。

为了掌握AI芯片在汽车领域的应用情况，本文对汽车AI芯片的研究进展进行综述，梳理多种汽车应用场景对AI芯片的需求，通过对多个代表性芯片企业或车企在芯片方面的发展路线进行分析，并结合AI芯片面临的技术挑战，对汽车行业AI芯片发展趋势进行展望。

1 AI芯片概述

AI芯片是人工智能系统中最核心的组成部分，其提供了强大的计算能力和数据处理能力，通过训练和推理实现人工智能算法的运算和处理，可以完成图像语音识别、机器翻译以及自动驾驶等功能[10-11]。

AI芯片主要包括3类：一是经过软硬件优化可以高效支持AI应用的通用芯片，如视觉处理器（Graphics Processing Unit，GPU），现场可编程逻辑门阵列（Field Programmable Gate Array，FPGA）；二是专门为特定的AI产品或者服务而设计的专用集成电路（Application Specific Integrated Circuit，ASIC）芯片；三是受生物脑启发设计的神经形态计算芯片[12-13]，具体分类及特点如表1所示。

表1 AI芯片的分类与特点

芯片类型	定制化程度	优点	缺点	适用领域	代表芯片
GPU	通用型	通用性强，浮点运算能力和并行计算能力强，速度快	功耗高，管理能力弱，生产成本高	图像处理，高性能计算	NVIDIA Tesla V100
FPGA	半定制化	灵活性高，功耗低，可编程性强，实时性好	生产成本高	算法更新频繁的专用领域	Xilinx V7-690T
ASIC	全定制化	功耗低，计算性能高，量产成本低	开发周期长，扩展性差，技术风险大	市场需求量大的专用领域	Google Edge TPU
类脑芯片	模拟人脑	功耗低，可扩展性好，认知能力强，通信效率高	算力低，主流算法的适配性差，开发技术难度大	类脑应用领域	IBM Truenorth

2 AI芯片在汽车行业的应用

2.1 AI芯片应用于智能驾驶

根据国际及国内的汽车驾驶自动化分级标准[14-15]，自动驾驶可分为L0级至L5级6个等级，如表2所示。

在智能驾驶方面，AI芯片主要用于处理智能驾驶系统中一系列复杂且关键的任务，主要包括：（1）环境感知：智能驾驶汽车需要通过摄像头、激光雷达、4D毫米波雷达等多种传感器来获取道路和环境的信息。环境感知算法需要处理和分析海量的图像与点云数据，以识别车辆、行人以及障碍物，这要求AI芯片具备强大的图像处理和数据分析能力；（2）传感器数据融合：传感器融合算法则进一步整合来自不同传感器的信息，提高感知的准确性和鲁棒性，这同样需要芯片支持高效的跨模态数据处理；（3）高精度定位：高精度定位算法依赖于卫星导航、地面特征匹配、惯性测量单元（Inertial Measurement Unit，IMU）等技术，确保车辆在任何环境下均能准确感知自身的位置与姿态，这对AI芯片的浮点运算能力和并行处理效率提出了极高要求；（4）路径规划与决策：路径规划与决策算法则基于实时感知信息，预测未来的交通状况，同时考虑安全及效率，计算出最佳行驶路径，该过程同样需要强大的计算能力来支撑复杂

表2　自动驾驶等级分类

等级	自动化水平	具体内容
L0	无自动化驾驶	由人类驾驶者全权操作汽车，在行驶过程中可以得到警告和保护系统的辅助
L1	驾驶支援	通过驾驶环境对转向盘和加减速中的操作提供驾驶支援，其他的驾驶动作都由人类驾驶员进行操作
L2	部分自动化驾驶	系统只是在特定场景下给驾驶员提供协助，车辆行驶决策权在驾驶员
L3	有条件的自动驾驶	可以在特定条件下实现自动驾驶，如高速公路或专用道路上
L4	高度自动化驾驶	可以在特定区域或环境内实现完全的自动驾驶，无需人工接管
L5	完全自动化	整个行车过程全部由车辆完成，无需任何人工干预

的模拟与优化。

因此，智能驾驶的AI芯片不仅要具备超高的算力，以应对上述算法带来的密集计算挑战，还应具有低功耗、高能效比的优点，以适应汽车在特定应用场景下的能源限制和长期运行需求。此外，AI芯片还应具备良好的热管理能力，同时在长时间、高负荷运行中保持高可靠性和安全性，以确保智能驾驶系统稳定运行并保障驾乘人员的出行安全。

2.2　AI芯片应用于智能座舱

随着汽车座舱进入智能时代，一芯多屏、多屏互联等功能需求使座舱芯片成为智能座舱的关键核心。

2.2.1　重点技术

2.2.1.1　一芯多屏

在传统座舱解决方案中，中控屏和仪表板等系统相互独立，一般由单一芯片驱动单个系统。随着座舱智能化发展，座舱域控制器进一步集成仪表板、抬头显示器（Head-Up Display，HUD）、流媒体后视镜等系统，实现单一系统级芯片（System on Chip，SoC）支持多系统多屏显示[16]。相较于传统的单芯单屏方案，一芯多屏方案在成本与通信方面更具优势。一芯多屏方案中，多屏交互信息在芯片内部完成传输，改变了多个操作系统之间的通信传输方式，节省了通信时间。此外，由于数字液晶仪表板涉及车辆安全，需要满足ASIL-B标准，通常采用QNX或Linux系统。娱乐系统需要运行丰富的娱乐功能，通常采用Android系统。单芯单屏方案中，多种芯片可以分别支撑不同系统，不存在跨系统问题，而单芯多屏方案由于跨系统运行可能会产生计算效率降低的问题。

2.2.1.2　虚拟化与硬隔离

在智能座舱的SoC芯片中，一般采用虚拟化（Hypervisor）或者芯片内部硬隔离2种技术来解决一芯多屏的跨系统运行问题。

虚拟化是指在操作系统和硬件之间插入Hypervisor层，根据虚拟机对计算单元、内存等的调度，实现对硬件资源的动态分配。该方式比较灵活，对硬件资源的利用效率较高。然而，每个虚拟机实例均需要运行客户端操作系统的完整副本以及其中包含的大量应用程序。从实际运行的角度来说，由此产生的沉重负载将会影响其工作效率及性能表现。

硬隔离是指在座舱SoC芯片设计时提前完成硬件资源划分。该方法不影响系统性能，功能安全隔离更可靠。然而，因不同系统不能混用，采用该方法会造成资源浪费且无法实现系统功能的扩展。

2.2.1.3　多模态交互

在座舱域的范围内，一般采用驾驶员监控系统（Drive Monitor System，DMS）摄像头实现驾驶员状态的感知，采用乘员监督系统（Occupancy Monitoring System，OMS）摄像头实现舱内乘客的状态感知，采用麦克风阵列实现多区域音源检测与语音识别。目前已经有DMS、OMS、麦克风阵列融合方案，实现了多模态识别与交互功能需求。此方案下，智能座舱域控制器（Cockpit Domain Controller，CDC）的AI算力需支持多模态数据的识别与人机互动。

2.2.2　AI芯片特征需求

足够的芯片算力。座舱芯片不仅需要处理传统的车辆管理任务，还需支持越来越多的娱乐和信息服务，如高清视频流、实时数据处理和复杂的用户界面交互。因此，座舱芯片需具有高算力、高性能图形图像渲染GPU、高性能音质和音效处理能力，支持多模态AI推理。

良好的功耗和散热管理。由于汽车系统需要长时间运行，对芯片的功耗和散热管理提出了极高要求。

生成式AI的集成。通过集成AI处理器，智能座舱能够提供更智能的语音交互、面部识别以及情境感知等功能。这些AI功能能够在本地快速处理，减少对

云端的依赖，从而降低延迟并提高响应速度。此外，还要求AI芯片具有更好的连接性。随着通信技术的快速发展，智能座舱芯片需要具备更强的通信能力和更快的数据处理速度。

多屏显示与高分辨率支持。目前，车内屏幕数量日益增加，除了传统的仪表板和中控屏，还包括副驾娱乐屏和后排娱乐屏等，这些屏幕通常要求高分辨率和高刷新率，对芯片的图像处理能力提出了更高要求。

智能座舱芯片的高端市场以高通为代表，其芯片在性能、处理速度和图像渲染能力上均占优势，这使高端智能座舱能够支持更多屏显、更流畅的运行和更丰富的画面。而中低端市场则以瑞萨电子、恩智浦、德州仪器及部分中国企业（如杰发科技和芯驰科技）为代表，其产品虽然在性能上不及高通等消费级芯片，但在车规级安全和成本效益上具有一定优势。

3 汽车AI芯片发展现状

各大代表性汽车芯片企业的发展路径呈现出多元化的特点。例如，高通在智能座舱领域拥有较大的市场占有率，同时其积极布局智能驾驶和域融合方向技术。NVIDIA在智能驾驶领域展现出强大的市场竞争力，其芯片算力一直处于行业领先水平，基于其高算力AI芯片继续布局智驾方向，未来可能会联动云端整体布局车云一体化方案。特斯拉在自动驾驶领域取得了显著进展，其自研的FSD芯片为自动驾驶提供了强有力的支持。国内车载芯片企业也正加速发展并持续扩大市场份额，地平线是智能驾驶领域的标杆企业，芯擎是智能座舱芯片领域的代表性企业。蔚来、小鹏等国内新能源车企也通过自主研发和合作投资等方式，不断提升自身的芯片研发能力。为了分析AI芯片在汽车领域的发展趋势，选择高通、NVIDA、特斯拉、地平线和芯擎作为不同技术路径的代表性企业，对其在汽车AI芯片方面的发展进行阐述。

3.1 高通

自2014年起，高通先后推出了602A、骁龙820A、SA8155P和SA8295P和SA8397P五代座舱芯片。依托其在安卓生态系统中的领先地位，高通在高端汽车座舱芯片市场迅速占据主导地位。随着智能座舱业务逐渐成熟，高通将业务优势延伸至智能驾驶和域融合领域，全面布局座舱芯片、智驾芯片、域融合芯片。

2021年，高通推出第一代智驾芯片Ride（即SA8540P）。2022年，高通推出第二代Ride系列智驾芯片，包含SA8620P、SA8650P 2款SoC，其中SA8650旗舰版AI算力为100×10^{12}次操作/s。

2023年，高通推出了中央计算平台芯片Ride Flex（即SA8775），可以实现座舱智驾一体化功能。2024年，高通推出面向汽车的Snapdragon Cockpit Elite（骁龙座舱至尊版）和Snapdragon Ride Elite（骁龙Ride至尊版）平台。骁龙Ride至尊版平台是高通为汽车智能驾驶专属打造的灵活架构平台，可帮助整车企业在同一SoC无缝运行数字座舱和智能驾驶，实现舱驾融合，有效降低了硬件成本。2026年，高通推出车规级中央计算芯片SA8797P，支持舱驾一体融合计算，单颗算力可达600×10^{12}次操作/s，适用于高算力AI推理、端侧大模型运行及环境感知与决策执行。

3.2 NVIDIA

2015年，NVIDIA发布了Drive智能驾驶开发计算平台。2019年首次公布了Orin芯片并在2022年正式推出，目前已应用于多种车型，如蔚来ET7、小米SU7、小鹏P7、上汽飞凡F7和智己L7等。2022年，NVIDIA发布了新一代集中式车载计算平台Thor，其可在单个安全、可靠的系统上运行高级驾驶员辅助应用和车载信息娱乐应用，最高算力可达$2\ 000\times10^{12}$次操作/s，能够进行多域计算，具备跨域融合能力。理想、小鹏、极氪、比亚迪等多家国内头部新能源车企应用NVIDIA Thor平台。

3.3 特斯拉

特斯拉专注于通过高精度的摄像头捕捉周围环境的原始视频数据。特斯拉认为借助先进的图像处理技术和深度学习算法，摄像头足以提供丰富且准确的信息支持自动驾驶系统的决策和规划。

为了实现该技术路线，特斯拉投入了大量资源来收集和分析行驶数据。每一辆特斯拉车辆在行驶过程中均会持续收集道路信息、交通状况、天气条件以及车辆自身的动态数据并实时上传至特斯拉的数据中心。通集中处理和分析采集的数据信息，可以不断优化其自动驾驶算法和模型，提高系统的准确性和可靠性。

特斯拉建立了一个完善的数据闭环体系。自动驾驶系统不仅可以根据当前的数据进行决策和规划，还会在行驶过程中不断收集新的数据，并将数据用于进一步的算法优化和模型训练。这种持

续迭代的过程使得特斯拉的自动驾驶技术能够不断适应多种复杂场景和路况，提高了自动驾驶的智能化水平。

特斯拉智驾硬件平台完成了从初代HW1.0到HW4.0的多代演进，最新一代AI5芯片（原HW5.0）已进入量产阶段。HW4.0方案配备2枚FSD芯片，总算力可达720×10^{12}次操作/s；新一代AI5芯片由台积电与三星协同代工生产，单芯片运算算力区间为2 000×10^{12}~2 500×10^{12}次操作/s，整体算力相较HW4.0单芯片提升约5倍，智驾数据处理与决策能力实现跨越式升级。

3.4 地平线

地平线依托自主研发的BPU架构，持续迭代征程系列智能驾驶芯片，完成了从入门级辅助驾驶到大算力高阶自动驾驶的完整技术进阶，已经发展为国内智能驾驶AI芯片的领军企业。其新一代征程6系列智能驾驶AI芯片，算力覆盖10×10^{12}~560×10^{12}次操作/s，可以灵活满足从高级辅助驾驶到全场景智能辅助驾驶的差异化量产需求。此外，地平线也积极布局域融合芯片，已发布地平线星空芯片，可以用一颗芯片支持传统智驾与座舱两套系统，实现了高效的一体化部署。

3.5 芯擎

芯擎是国内少数同时布局智能座舱与智能驾驶高算力车规芯片的本土企业，于2021年推出龍鹰一号（SE1000）芯片，这是国内首款实现大规模量产的7 nm车规级智能座舱芯片，凭借领先的制程工艺、稳定的车规可靠性与强大的多媒体处理能力，在吉利、长安、中国一汽、大众等国内外车企的部分车型中应用，累计装车量突破百万辆，稳居国产中端乘用车座舱芯片市场首位。芯擎后续迭代推出龍鹰二号（SE2000）5 nm高阶座舱芯片，面向舱驾融合场景优化架构设计，支持更高等级AI交互与多域协同计算。在智能驾驶领域，芯擎推出星辰一号（AD1000）7 nm车规级辅助驾驶芯片，形成了“龍鹰座舱系列+星辰驾驶系列”双线并行的完整产品矩阵。

3.6 小结

通过分析高通、NVIDIA、特斯拉、地平线和芯擎等代表性企业在汽车AI芯片领域的发展进程，结果表明：各大企业均追求更高算力，并且积极推动技术架构向域融合方向发展。

4 AI芯片的技术挑战与解决方法

4.1 内存瓶颈

内存墙是指内存的容量或传输带宽有限，从而严重限制中央处理器（Central Processing Unit，CPU）性能发挥的现象[17]。随着处理器技术的飞速发展，处理器的计算速度显著提升。然而，内存技术的发展却相对滞后，导致内存带宽和访问速度无法与处理器的计算速度相匹配，如图1所示[18]。图1展示了不同代际互连，内存带宽和峰值算力每秒浮点运算次数（Floating-point Operations Per Second，FLOPS）的发展。图1表明，带宽增长非常缓慢。纵坐标是归一化后的系数，如使用R10000系统对硬件峰值FLOPS进行归一化。

从内存带宽方面来看，尽管带宽持续提升，但其与处理器计算速度的提升幅度相比，仍存在较大差距。以动态随机存取存储器（Dynamic Random-Access Memory，DRAM）为例，其带宽的增长速度远跟不上处理器对数据读取和写入的需求。在访问速度上，内存的物理特性决定了数据从内存存储单元传输到处理器缓存这一过程存在不可忽视的延迟。当前主

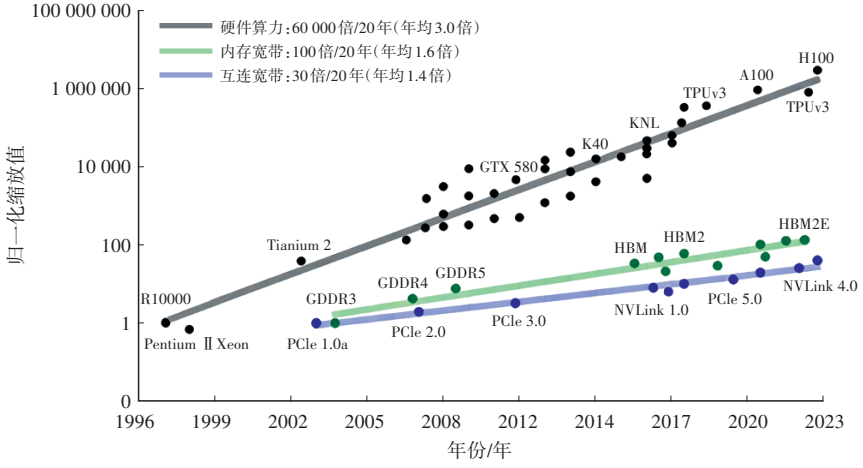


图1 存储计算性能“剪刀差”[18]

流的DRAM内存访问延迟通常在60~90 ns甚至更高，阻碍了处理器的运算速度。

处理器与内存之间的性能差异，使得在计算机系统处理任务时，处理器通常需要耗费大量时间等待数据从内存中加载至高速缓存中。以复杂的数据分析任务为例，处理器需频繁地从内存中读取大量的数据进行运算，由于内存访问速度远低于处理器运算速度，处理器在大部分时间内处于空转状态，导致计算机系统的资源利用率低下，降低了系统的整体性能表现。在实时性要求极高的应用场景中，如自动驾驶汽车的控制系统，高访存时延可能导致车辆对路况变化的响应不及时，从而引发严重的安全隐患。

存算一体（Computing-In-Memory，CIM）技术是解决上述问题的关键途径。其核心思想是将计算单元与存储单元集成，通过在存储器内部直接执行运算操作，从而避免频繁的数据传输，提高计算效率[19-20]。这种架构可以直接在存储器中执行计算操作，而不需要将数据从存储器中加载到计算单元进行处理。相比于传统的冯·诺伊曼体系架构，存算一体支持更高性能、更低功耗地实现计算，并且可以提高芯片的集成度，是未来大数据计算芯片架构的主流方向[21]。

静态随机存取存储器（Static Random-Access Memory，SRAM）具有高速读写的特性，常作为存算一体架构的首选存储介质[22-24]。此外，各类存算一体芯片架构的发展促使更多新式存储方式得以应用。其中，相变储存器（Phase Change Memory，PCM）、磁阻随机存储器（Magnetoresistive RAM，MRAM）、电阻式存储器（Resistive RAM，RRAM）以及铁电存储器（Ferroelectric RAM，FRAM）[25]的出现为破解内存墙问题提供了新思路。

4.2 CMOS工艺和器件瓶颈

比例缩小原则是指在芯片面积不变的条件下，通过等比例缩小金属氧化物半导体场效应晶体管器件的栅极宽度、沟道宽度等器件尺寸和电学参数，同时保证晶体管性能和尺寸，从而实现晶体管密度的提升，并且工作电压等比例减小，进而能够使得器件在保持功耗不变的情况下获得更高的速度，即能够在原有技术的基础上小幅度改进器件性能，推动摩尔定律的发展。直到21世纪初，等比缩减一直是芯片发展的主流方向。

然而，随着摩尔定律的持续发展，互补金属氧化物半导体（Complementary Metal Oxide Semiconductor，CMOS）晶体管的尺寸已接近物理极限，基于一维静电势的等比例缩小理论将不再适用。这导致了如量子效应、隧穿效应和短沟道效应等问题[26]，严重影响晶体管性能。此外，随着CMOS晶体管尺寸的缩小，虽然单个晶体管的功耗有所降低，但整体系统的功耗却可能因晶体管数量的增加而上升。特别是在高集成度和高频率工况下，功耗问题尤为突出。CMOS工艺和器件瓶颈的解决方法包括先进封装和Chiplet。

4.2.1 先进封装

当前芯片工艺节点中，进一步缩小晶体管尺寸难度极大且成本较高。采用先进封装技术可以实现对芯片进行封装级重构，有效提高系统集成密度。

（1）2.5D封装：2.5D封装是指通过在芯片之间插入中介层来进行高密度互连的一种异构芯片封装。在2.5D封装中，裸晶堆栈并排放置在具有硅通孔的中介层顶部，通过中介层实现芯片之间的连接。其物理结构上是3D的，通过结构上的堆叠，能够节省封装空间。

（2）3D封装：3D封装又称为叠层芯片封装技术，是指在不改变封装体尺寸的前提下，在同一个封装体内于垂直方向叠放多个芯片的封装技术[20,27]。3D集成电路（Integrated Circuit，IC）封装不使用TSV，支持更小的尺寸，可容纳多个异构裸片且各功能模块可采用不同的制程节点，可以有效降低成本并提高产品上市速度[28]。

4.2.2 芯粒技术

芯粒（Chiplet）是一种功能电路块。Chiplet技术将功能丰富且面积较大的芯片裸片拆分为多个芯粒，并将这些具有特定功能的芯粒通过先进封装的形式进行组合，最终形成一个系统芯片[29-30]。Chiplet技术将大芯片分成面积更小的芯片，有助于改善良品率，从而降低制造成本，同时提高芯片设计效率及灵活性。

5 汽车AI芯片发展趋势

5.1 域融合

在传统的分布式架构中，智能座舱和智能驾驶由2个独立的电子控制单元（Electronic Control Unit，ECU）控制，分别负责车载娱乐功能和自动驾驶功能，2个ECU通过车载通信网络相连。而将CDC与智能驾驶域控制器（Active-Directory Domain Controller，ADC）融合至一个中央计算平台是电子电气架构演进

的趋势，这种架构也被称为舱驾一体。

舱驾一体的实现过程可以分为3个阶段：(1)域控硬件融合，即One Box/Two Board的方案。座舱域和智驾域的板子和接口相互独立，仅在硬件层面将2块PCB板集成于较大的域控中，板间通过PCIe或者以太网进行通信。从架构角度来看，One Box/Two Board和Two Box的方案区别不明显，仅节省了部分线束和ECU壳体成本，核心的座舱域和智驾域的功能需求和开发仍相互独立。(2)在一块印制电路板(Printed Cirwit Board，PCB)上同时集成座舱和智驾芯片，实现One box/One board。该设计能进一步减小域控的尺寸，同时提高座舱和SoC之间的通信效率，通过算力共享提升性能，但是对于硬件设计的能力要求较高。(3)使用一个SoC同时实现座舱和智驾的功能，即One Chip方案，这也是舱驾一体的最终形态。硬件层面，与One Board方案相比，One Chip方案节省了一个SoC，可以最大程度节约成本。在软件层面，座舱和智驾运行在统一的软件架构下，提升了开发效率同时可以带来更多创新功能。

虽然单SoC芯片的舱驾一体架构方案才能实现真正的座舱和智驾融合，但是座舱和智驾在功能需求、功能安全要求、信息安全要求以及对不同类型算力需求的侧重点等方面的要求存在差异，这就对单SoC芯片实现舱驾一体提出了极大挑战。目前较为可行的方案是将部分域的功能集合于高性能计算单元，再逐渐聚合更多功能域。如舱泊一体方案将泊车功能融合于座舱，座舱域控制器接收泊车信号，节省了泊车控制器成本。而行泊一体方案中行车和泊车共用于一个域控制器。从适用领域来看，舱泊方案更适合于基础泊车功能融合，而高阶泊车需要行车感知系统的参与和功能安全等级要求，更适合采用行泊一体方案。

5.2 算力增长

随着汽车各应用场景对计算需求的不断增长，AI芯片正朝着更大算力的方向演进。芯片设计需要不断优化，以提高运算速度和能效比，从而支撑起日益庞大的数据处理量。

如图2所示，AI芯片算力实现了快速增长，业界已普遍将峰值计算能力视为评估AI芯片性能的关键指标，由此引发了一场围绕算力的激烈竞争。各大汽车企业也陆续投身芯片自研行列，或是与芯片厂直接合作，共同研发设计芯片，提升对芯片的自主掌握能力。



图2 汽车AI芯片算力增长趋势

AI芯片从单一域应用向跨域应用转变，意味着AI芯片需要处理更多的数据和信息，因此对作为运算核心的AI芯片算力提出了更高的要求。面对由软件定义汽车所带来的更为复杂多变的新设计要求，汽车AI芯片亟需具备远超以往的更高的计算性能，以确保能够高效处理海量的数据运算和复杂的实时任务。在计算性能需求的驱动下，一种新的计算架构——多核异构计算应运而生。区别于传统同构计算(仅包含单一类型计算单元，如全部由中央处理器组成)，异构计算整合了多种各具优势的计算资源，利用各个计算单元协同完成计算任务。因此，异构计算在计算效率与可扩展性方面具备显著优势[31-32]。

通过采用相同的制造工艺，将AI芯片集成于同一块芯片上，并结合片上网络(Network on Chip，NoC)技

术，构建SoC级别的多核异构架构，进而实现多域融合。在此异构架构下开展的异构计算，可以通过合理的任务分配与协同工作，充分发挥各计算单元的独特优势，极大地提高计算效率。目前，异构计算正处于快速发展阶段，已在汽车、人工智能及科学计算等多个前沿领域实现广泛应用，必将是未来AI芯片提升运算速度与运算效能的重要发展方向。

5.3 存算一体

AI芯片的设计也在逐步突破传统冯·诺伊曼体系结构，向存算一体的新型架构发展。存算一体架构可以显著提高数据的访问速度和计算效率，为AI应用提供了更加强劲的性能支持。

5.4 先进封装

在封装技术方面，AI芯片也在从传统封装向更先进的2.5D或3D封装或Chiplet技术发展。先进封装技术不仅提高了芯片的集成度和互连密度，还使得不同功能模块的灵活组合成为可能，降低了设计成本，提高了生产效率和良品率。

5.5 小结

展望未来，AI芯片应用于汽车行业的发展趋势将呈现出以下3个特点：一是随着智能网联汽车的不断发展和普及，AI芯片的需求将持续增长；二是AI芯片行业将不断涌现出新的技术和产品，包括更高效的算法、更先进的芯片制造技术、更强大的计算能力等；三是行业将面临更多的竞争和合作机会，各大整车企业正通过与芯片企业合作或自研芯片等方式，积极寻求AI芯片自主开发能力，以满足不断增长的市场需求。同时，也需要不断攻克技术瓶颈，推动AI芯片在汽车行业的广泛应用和深入发展。

6 结束语

本文对AI芯片在汽车行业的主要应用场景智能驾驶和智能座舱进行了阐述和分析，梳理出汽车AI芯片的技术需求。以高通，NVIDIA和特斯拉等企业为代表，分析了汽车AI芯片的发展趋势，即技术架构向域融合发展，并且不断追求更高算力。针对汽车AI芯片面临的技术挑战，总结了存算一体、先进封装和Chiplet技术等解决方法，这些也是AI芯片的重要发展方向。然而，这些先进技术尚未成熟，仍存在着一系列问题亟需解决，如存算一体面临着原有编译程序难以直接适配、可靠性与纠错机制缺失、数据一致性与同步等方面的难题；先进封装的3D高密度堆叠带来散热集中、信号串扰等问题；应用多Chiplet可能导致故障叠加，需新增故障隔离、冗余备份、动态降级等机制，设计复杂度大幅提升。

虽然仍面临着许多技术挑战，但是汽车AI芯片发展前景广阔、潜力巨大，工艺更优、性能更强的AI芯片必定会助力智能网联汽车的深入发展。

参 考 文 献

[1] 赵瑞, 夏显召, 吴海文, 等. 智能驾驶人工智能芯片的计算性能测试研究[J]. 标准科学, 2024(10): 65-70.

[2] AI对汽车芯片设计的影响[J]. 中国集成电路, 2020, 29(Z1): 84-87+93.

[3] 林澍, 陈宇锋, 黄丽军. 电动汽车域控制器发展分析与未来展望[J]. 汽车实用技术, 2023, 48(11): 207-212.

[4] 张微, 邵帅. 汽车用芯片技术现状和发展趋势[J]. 金属功能材料, 2023, 30(5): 107-112.

[5] 李嘉宁, 姚鹏, 揭路, 等. 存算一体技术研究现状[J]. 电子学报, 2024, 52(4): 1103-1117.

[6] 何斯琪, 穆琛, 陈迟晓. 基于存算一体集成芯片的大语言模型专用硬件架构[J]. 中兴通讯技术, 2024, 30(2): 37-42.

[7] 邵帅, 王榆升, 李家玺, 等. 面向图像解构计算的存算一体混合存储技术研究[J]. 科技视界, 2024, 14(20): 35-37.

[8] 韩传悦. 面向先进封装的垂直互连关键技术探索与应用[D]. 成都: 电子科技大学, 2024.

[9] 杨彦章, 陈志华, 熊伟, 等. 先进封装高速电镀铜柱的研究[J]. 印制电路信息, 2024, 32(S2): 231-238.

[10] 赵玥, 肖梦燕, 罗军, 等. 人工智能芯片及测评体系分析[J]. 电子与封装, 2023, 23(5): 31-37.

[11] 朱海涛. AI芯片的应用与发展趋势[J]. 中国安全防范技术与应用, 2019(5): 44-49.

[12] 商惠敏. 人工智能芯片产业技术发展研究[J]. 全球科技经济瞭望, 2021, 36(12): 24-30.

[13] 尹首一, 郭珩, 魏少军. 人工智能芯片发展的现状及趋势[J]. 科技导报, 2018, 36(17): 45-51.

[14] SAE International. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles: SAE J3016: 2021[S/OL]. (2021-04-30) [2026-08-13]. https://www.sae.org/standards/j3016_202104-taxonomy-definitions-terms-related-driving-automation-systems-road-motor-vehicles.

[15] 国家市场监督管理总局, 国家标准化管理委员会. 汽车驾驶自动化分级:GB/T 40429—2021[S]. 北京: 中国标准出版社, 2021.

[16] 边旭东, 张亦弛, 谢卉瑜. 浅谈智能座舱的“一芯多屏”[J]. 时代汽车, 2021(6): 12-14.

[17] 清华大学, 北京未来芯片技术高精尖创新中心. 人工智能芯片技术白皮书[R]. 北京: 清华大学, 北京未来芯片

技术高精尖创新中心, 2018.
[18] GHOLAMI A, YAO Z, KIM S, et al. AI and Memory Wall[J]. IEEE Micro, 2024, 44 (3): 33-39.
[19] 康旺, 寇竞, 赵巍胜. 存算一体芯片发展现状、趋势与挑战[J]. 中国科学:信息科学, 2024, 54(1): 16-24.
[20] 姚鹏, 宋昌明, 胡杨, 等. 高算力芯片未来技术发展途径[J]. 前瞻科技, 2022, 1(3): 115-129.
[21] 中国移动通信有限公司研究院. 存算一体白皮书[R]. 北京: 中国移动通信有限公司研究院, 2022.
[22] 叶乐, 贾天宇, 陈沛毓, 等. SRAM存算一体芯片研究: 发展与挑战[J]. 中国科学: 信息科学, 2024, 54(1): 25-33.
[23] JHANG C J, XUE C X, HUNG J M, et al. Challenges and Trends of SRAM-Based Computing-In-Memory for AI Edge Devices[J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2021, 68 (5): 1773-1786.
[24] CHANG L, LI C, ZHANG Z, et al. Energy-Efficient Computing-In-Memory Architecture for AI Processor: Device, Circuit, Architecture Perspective[J]. Science China Information Sciences, 2021, 64: 160403.
[25] 李憬. 基于新型存储的超低功耗非易失性处理器设计[D]. 宁波: 宁波大学, 2022.
[26] 陈隆. "后摩尔时代"下的异构计算[J]. 信息化建设, 2024(8): 52-54.
[27] ZHANG Y, LI L, LU Z H, et al. A Survey of Memory Architecture for 3D Chip Multi-Processors[J]. Microprocessors and Microsystems, 2014, 38 (5): 415-430.
[28] 田文超, 谢昊伦, 陈源明, 等. 人工智能芯片先进封装技术[J]. 电子与封装, 2024, 24(1): 21-33.
[29] 张志伟, 田果, 王世权. 先进封装Chiplet技术与AI芯片发展[J]. 中阿科技论坛(中英文), 2023(11): 90-94.
[30] 张衡, 袁杰. 芯粒技术在人工智能芯片设计中的应用探索[J]. 信息与电脑(理论版), 2024, 36(7): 37-39.
[31] 李荣春, 贺周雨, 乔鹏, 等. 面向大规模异构计算平台的MiniGo高效训练方法[J]. 国防科技大学学报, 2024, 46(5): 209-218.
[32] 罗兰. 基于可扩展性和效率的异构计算系统设计与优化[J]. 数字通信世界, 2024(7): 22-24.

(责任编辑　**梵玲**)

《汽车工艺与材料》投稿须知

《汽车工艺与材料》于1986年创刊,是由中国第一汽车集团有限公司主办的国内外公开发行的汽车材料与制造技术类月刊,目前已入选《中文核心期刊要目总览》(第二版)、RCCSE中国准核心学术期刊(B+)、中国核心期刊(遴选)数据库、中国学术期刊综合评价数据库、欧洲学术出版中心数据库(EuroPub)、哥白尼精选期刊数据库(ICI Journals Master List)、EBSCO International数据库、J-Gate数据库。

《汽车工艺与材料》以"为中国报道汽车制造,为汽车引领工艺材料"为办刊宗旨,致力于报道以汽车轻量化技术和智能制造技术为核心的先进制造技术与材料应用技术,重点关注电动汽车蓄电池、电机、电控关键材料技术,燃料电池材料技术,高强度钢、铝镁合金、非金属材料及其成形技术,连接技术,智能装备与绿色制造等,以期通过高质量学术内容的出版和传播助推行业创新技术的交流与发展。

《汽车工艺与材料》杂志关注领先的整车及零部件企业和材料、装备等供应商,及时报道汽车行业最新的产品设计、制造、材料、加工技术、生产装备、检测技术等方面的成功案例。

主要栏目:

AT&M视界、生产现场、材料应用、生产装备、检测技术、数字化园地、行业动态等。

投稿要求:

(1)来稿须具有独创性并与实践相结合,文章字数最好控制在5 000~8 000字。

(2)来稿不能在国内、外公开杂志上发表过,请勿一稿多投。

(3)来稿的试验方法、试验数据、试验结论必须准确、可靠。

(4)来稿须包括以下项目:题名、作者姓名、作者单位、摘要(200字左右)、参考文献等。来稿采用word文档的格式。

(5)来稿文章格式应符合一般科技论文格式,或参考近期本刊所刊登文章格式。

(6)文章必须附有公开发表的、体现本领域最新研究成果的参考文献,且在文中应标注文献引用处。

(7)本刊使用网站投稿,投稿网址:http://qcgyycl.cbpt.cnki.net,咨询电话:0431-82026054。

竭诚欢迎汽车行业及相关各界的专家学者积极向本刊投稿。