

# CIC 灼识



## 全球 AI 推理芯片行业报告

© 2026 CIC 灼识版权所有。本文件包含高度机密信息，仅供我方客户专属使用。  
未经 CIC 灼识书面许可，严禁以任何形式传阅、引用、复制或转载本文任何内容。

## 摘要

训练与推理是 AI 芯片的两大核心计算任务。训练需在模型投入实际应用前，处理海量数据并优化参数以完成模型构建。因此，在 AI SoC 行业发展初期，训练是行业的核心。然而，随着 AI 模型（尤其是大语言模型）在性能与实用性上不断提升，市场需求持续拓展，行业如今更聚焦实际落地应用，AI 推理芯片的重要性也与日俱增。

## 目录

### 1. 行业概览

#### 1.1 行业定义

#### 1.2 行业结构

#### 1.3 行业规模与增长

### 2. 核心增长驱动因素与发展趋势

#### 2.1 核心驱动因素

#### 2.2 未来展望

## 1. 行业概览

### 1.1 行业定义

AI 芯片行业正从以训练为核心向以推理为核心的范式发生根本性转变，AI 推理芯片的需求正急剧增长。

AI 推理芯片专为高能效、低时延设计，可实现近乎即时的处理结果。

### 1.2 行业结构

AI 推理芯片可部署于云端、边缘及端侧场景，不同场景需根据差异化需求进行定制化芯片设计：

**云端 AI 推理芯片**通常用于数据中心，负责处理大规模、高密度、高并发的集中式推理任务。因此，这类芯片优先考量高算力、广泛适用性、灵活性及可扩展性。

**边缘 AI 推理芯片**部署于靠近数据源的边缘服务器、网关或基站，执行实时本地推理，需要在高性能与能效之间做好平衡，以保障低时延、数据安全及运行稳定性。

**端侧 AI 推理芯片**直接应用于终端用户设备，例如智能手机、智能汽车及智能家居等消费电子设备。

### 1.3 行业规模与增长

2024 年，全球 AI 推理芯片市场规模达到 6,067 亿元人民币，细分市场规如下：

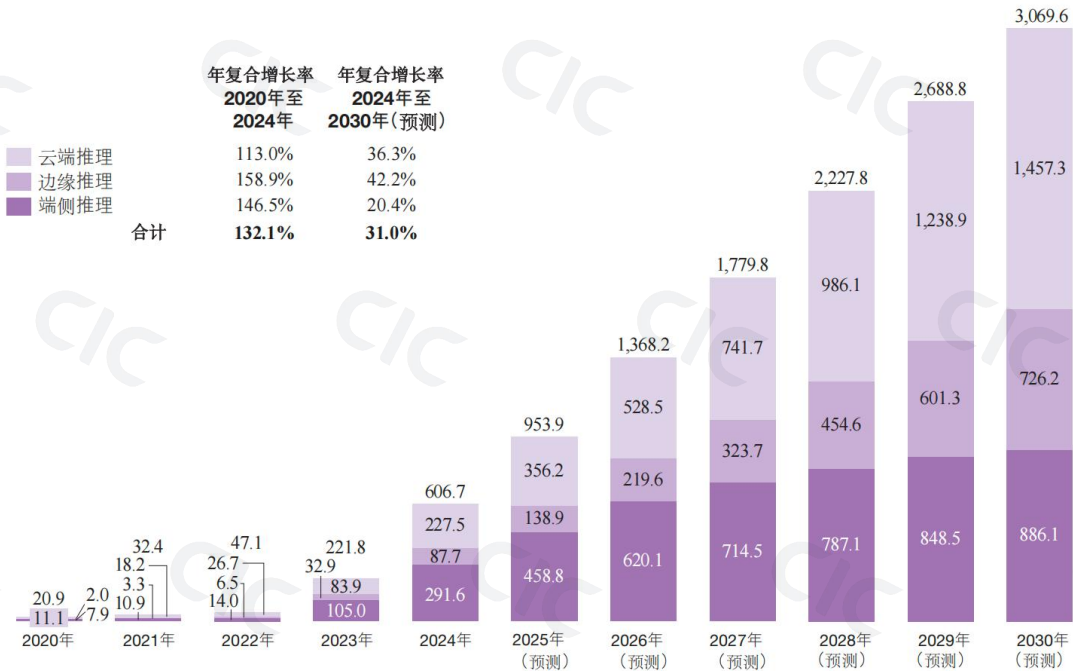
**端侧推理:** 2,916 亿元人民币

**云端推理:** 2,275 亿元人民币

**边缘推理:** 877 亿元人民币

全球云端、边缘及端AI推理芯片市场规模，2020年至2030年（预测）

(人民币十亿元)



资料来源: CIC 灼识、市场参与者访谈、行业出版物、政府统计资料、上市公司公开文件、

## 新闻

2024 年，中国 AI 推理芯片行业规模达到 1,608 亿元人民币，细分行业规模如

下:

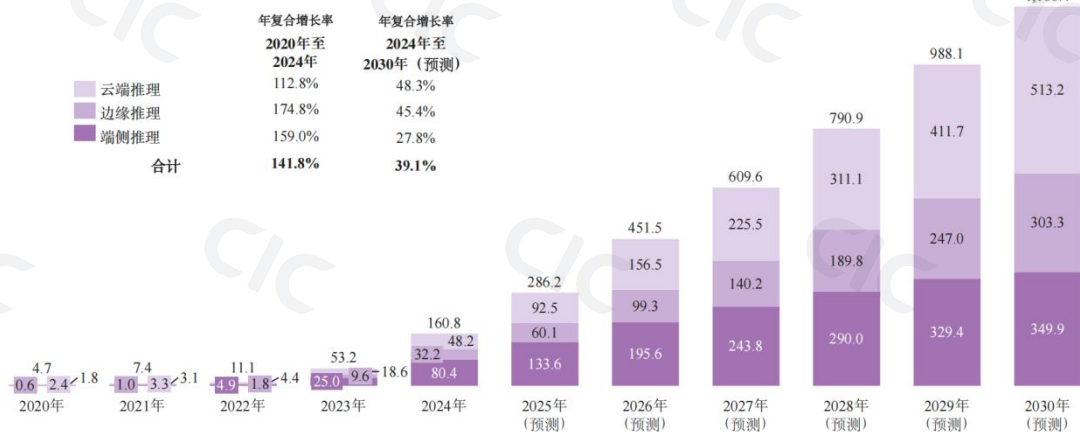
**端侧推理:** 804 亿元人民币

**云端推理:** 482 亿元人民币

**边缘推理: 322 亿元人民币**

中国云端、边缘及端侧AI推理芯片行业规模，2020年至2030年（预测）

(人民币十亿元)



资料来源: CIC 灼识、市场参与者访谈、行业出版物、政府统计资料、上市公司公开文件、

新闻

## 2. 核心增长驱动因素与发展趋势

### 2.1 核心驱动因素

**AI 智能设备渗透率提升：**全球 AI 智能设备渗透率从 2020 年的不足 1% 增长至 2024 年的 9.4%，预计 2030 年将超过 44%，这将推动边缘及端侧 AI 推理计算需求大幅增长，让边缘和端侧 AI 推理芯片成为此次智能化转型的关键支撑。

**数据量激增与低时延需求：**智能驾驶、机器人、工业控制等实时应用要求毫秒级处理，而传统云端架构受时延与带宽限制，往往难以满足。通过将 AI 推理芯片直接部署在边缘侧，系统可本地处理数据，实现即时、可靠、协同的响应。因此，边缘推理已成为核心基础设施，推动愈发智能化的设备快速发展。

**数据合规推动本地化处理：**全球愈发严格的数据监管法规已使数据成为高价值战略资产。为在运营效率与严格合规之间取得平衡，企业越来越多地采用本地化处理，而非云端传输。因此，边缘及端侧 AI 推理芯片成为关键基础设施，为满足现代监管要求提供所需的安全性与数据闭环处理能力。

### 2.2 未来展望

预计到 2030 年，全球 AI 推理芯片行业规模将达到 30,696 亿元人民币，2024 至 2030 年预测期内年复合增长率为 31.0% (2024-2030E)。

预计到 2030 年，中国 AI 推理芯片行业规模将达到 11,664 亿元人民币，2024 至 2030 年预测期内年复合增长率为 39.1% (2024-2030E)。

## 关于 CIC 灼识

CIC 灼识咨询（简称“CIC 灼识”）是一家专业咨询机构，围绕投融资全生命周期，提供定制化一站式全流程服务。公司在全球各大市场主导打造多个行业首创的标杆 IPO 项目，业绩稳居世界前列。同时在全类专业细分赛道中，拥有无可匹敌的资源触达能力与深度全覆盖研究实力。

CIC 灼识助力企业优化具备规模化潜力的商业模式，塑造极具说服力的资本市场价值叙事，畅通对接全球资本市场的路径。同时作为投资机构信赖的尽职调查合作伙伴，输出精细化行业研判视角，并直通各领域权威专家资源，助力客户精准锁定高价值机遇、有效规避核心重大风险。

CIC 灼识团队深耕金融服务、人工智能、大数据、互联网、高新技术、医疗健康、教育、文娱、消费品、交通运输物流、能源电力、环境与建筑科技、化工、工业制造、农业等多元领域，实时掌握深度一线市场动态，能够为客户独家输出贴合细分行业、可落地执行的专业洞察结论。

## CIC 灼识报告 & 行业概览

CIC 灼识搭建了一套严谨的多元化研究框架，整合一手调研与二手资料，为所有分析研判筑牢根基。一手调研主要深度对接行业权威专家与一线从业者，重点深耕供应链金融领域。二手研究则汇总梳理各大权威机构的公开数据，数据来源包括：中华人民共和国国家统计局、国家金融监督管理总局（SAFR，原中国

银行业监督管理委员会)、中国证券监督管理委员会, 以及上市公司公开披露文件。

我们运用自研专属数据分析体系对收集到的信息进行加工处理, 并通过多渠道研究数据交叉比对验证研究结论, 确保分析过程严谨、结果真实可靠。

本报告中展示的所有统计数据均可核验追溯, 全部基于报告出具当日可获取的有效信息整理而成。

本篇内容摘编提炼自 CIC 灼识深度行业研究报告精华, 聚焦各细分赛道的供需走势、核心增长驱动因素、研发创新趋势与行业未来发展前景等核心内容, 同时融合专家访谈、市场实地调研、行业数据解析等多维度专业研判成果。

## 免责声明

本报告由 CIC 灼识依据截至出具当日可获取的信息编制。本报告仅作参考之用, 内容不具备最终定论效力, 亦不得被解读为确定性结论。

本报告所载全部内容, 均不构成且不得视作投资建议、投资推荐, 亦非开展任何投资活动的要约、招揽或劝导。

凡因使用或依赖本报告所载信息, 直接或间接引发任何损失、损害及各类索赔诉求的, CIC 灼识特此明确免除一切相关责任。



CIC 灼识 | 全球 AI 推理芯片行业报告

## 联系我们

如需了解本报告更多详情,或咨询 CIC 灼识的各项专业服务,欢迎访问 [CIC 灼识官方网站](#),亦可发送邮件至: [marketing@cninsights.com](mailto:marketing@cninsights.com)。