

---

## AMPERE COMPUTING 1H 2024 ROADMAP UPDATE

---

STEVE McDOWELL, CHIEF ANALYST  
MAY 29, 2024

### CONTEXT

---

Ampere Computing released its [annual roadmap update](#), outlining significant progress and strategic initiatives aimed at driving innovation in sustainable, power-efficient computing for cloud and AI applications.

The update is a reaction to the growing power and energy demands of AI, emphasizes Ampere's collaboration with Qualcomm Technologies, and introduces its next-generation AmpereOne platform.

Let's take a deeper look at what was announced.

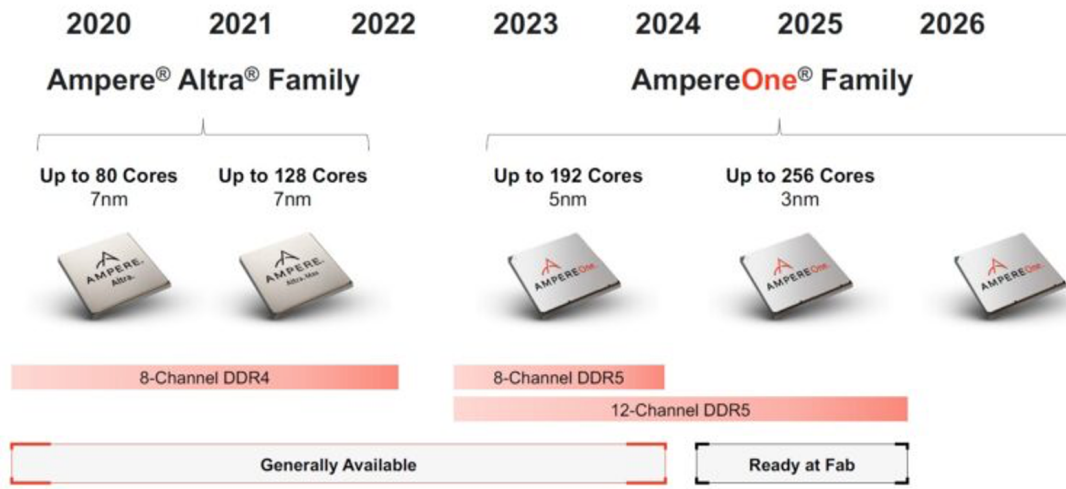
### AMPEREONE PROCESSOR

---

Ampere's latest innovation, the AmpereOne processor, is a significant update to the company's roadmap for sustainable and power-efficient computing. It is designed to meet the increasing demands of cloud and AI workloads with unprecedented performance and efficiency.

## Powerful Roadmap with Rapid Innovation

Continued Ampere Innovation Leads to Breakaway Performance Using Unique Architecture



### Key Features and Specifications:

- 256-Core CPU:** The new AmpereOne processor is a 256-core CPU, designed to handle intensive AI and cloud-native applications efficiently.
- 12-DRAM Channels:** The new processor provides 12 memory channels, enhancing memory bandwidth and supporting high-performance computing needs. This platform continues Ampere's existing 192-core, 12-channel memory platform, which is still expected later this year.
- N3 Process Node:** The AmpereOne processor utilizes the N3 process node, which leverages advanced semiconductor manufacturing technology to achieve higher performance and energy efficiency.
- AI Integration:** The processor supports a range of workloads, from traditional cloud-native applications to AI inferencing.
- Performance Metrics:** The new AmpereOne processor outpaces AMD Genoa by 50% and Bergamo by 15% in performance per watt.
- OEM and ODM Platforms:** Ampere announced that new OEM and ODM platforms featuring the AmpereOne processor will be shipping within a few months.

## QUALCOMM RELATIONSHIP

Ampere announced a new strategic collaboration with Qualcomm to enhance AI inferencing capabilities. The partnership leverages the strengths of both companies

to create a joint solution that combines Ampere's efficient, high-performance CPUs with Qualcomm's advanced Qualcomm Cloud AI 100 inference solutions.

1. **High-Performance AI Inferencing:** The primary objective of the collaboration is to develop a solution for AI inferencing, mainly focused on handling large language models (LLMs). By integrating Qualcomm's AI 100 inference technology with Ampere's CPUs, the joint solution provides robust performance for the industry's largest generative AI models.
2. **Power Efficiency:** A core tenet of this partnership is power efficiency. Qualcomm's Cloud AI 100 is known for its low power consumption, and when combined with Ampere's energy-efficient CPUs, the solution delivers high-performance AI inferencing while minimizing power usage. This focus on efficiency is particularly critical as AI workloads become increasingly power-intensive.
3. **Scalability:** The joint solution is designed to scale efficiently, addressing the needs of data centers that manage extensive and complex AI workloads.

---

## ANALYSIS

---

Ampere's technical capabilities must be tempered by the on-going shifts in the Arm-based server market. While Ampere has enjoyed its position as the sole Arm-architecture processor vendor to nearly every public cloud provider, excluding AWS, that exclusivity is ending. Microsoft recently announced its internally developed Cobalt 100 Arm-based processor, while Google introduced its Arm-based Axion processor.

These moves will ultimately leave Oracle as Ampere's only significant cloud customer. At the same time, there's been a considerable demand for Arm-based servers in the on-prem data center.

Ampere Computing's latest updates give the company a robust roadmap centered on sustainable, power-efficient computing tailored for the cloud and AI. With strategic collaborations, product launches, and a commitment to addressing the energy challenges of AI, Ampere is poised to make significant strides in the semiconductor industry. The company's focus on performance per watt and flexible, scalable solutions are precisely what it needs to be doing to grow its business. However, any potential forecasts should be tempered with the realities of the Arm-based server market.



© Copyright 2024 NAND Research.

NAND Research is a registered trademark of NAND Research LLC, All Rights Reserved.

This document may not be reproduced, distributed, or modified, in physical or electronic form, without the express written consent of NAND Research. Questions about licensing or use of this document should be directed to [info@nand-research.com](mailto:info@nand-research.com).

The information contained within this document was believed by NAND Research to be reliable and is provided for informational purposes only. The content may contain technical inaccuracies, omissions, or typographical errors. This document reflects the opinions of NAND Research, which is subject to change. NAND Research does not warranty or otherwise guarantee the accuracy of the information contained within.

NAND Research is a technology-focused industry analyst firm providing research, customer content, market and competitive intelligence, and custom deliverables to technology vendors, investors, and end-customer IT organizations.

Contact NAND Research via email at [info@nand-research.com](mailto:info@nand-research.com) or visit our website at [nand-research.com](http://nand-research.com).