

Google annonce sa première puce ARM

Google a récemment dévoilé sa première puce ARM, baptisée Axion, lors de l'événement Google Cloud Next '24 à Las Vegas. Cette puce sur mesure est destinée à soutenir les travaux d'intelligence artificielle de Google dans ses centres de données, tout en offrant une alternative compétitive aux solutions d'Intel et Nvidia.

Axion est conçu pour prendre en charge les charges de travail d'IA de Google avant d'être étendu aux clients professionnels de Google Cloud plus tard dans l'année. Déjà en service dans des services tels que les publicités YouTube et le moteur Google Earth, Axion promet des performances jusqu'à 50% supérieures aux processeurs x86, avec une efficacité énergétique jusqu'à 60% meilleure que les instances actuelles x86. De plus, Axion vise à faciliter le transfert des charges de travail existantes vers l'architecture ARM.

Pour se démarquer, Axion s'appuie sur Titanium, un système de microcontrôleurs supportant les opérations de sécurité et de réseau, ainsi que sur Hyperdisk, offrant des fonctionnalités avancées pour les applications critiques. Cette annonce fait suite à celle de Microsoft, qui a récemment dévoilé ses propres puces personnalisées pour son infrastructure cloud, mettant en avant une tendance vers des solutions de silicium sur mesure dans le secteur.

Google ne prévoit pas de vendre directement ces puces, mais de les mettre à disposition via ses services cloud, offrant ainsi aux entreprises la possibilité de les louer et de les utiliser pour leurs propres charges de travail cloud et d'IA.