



Ara-2：リアルタイム生成AI向けディスクリートNPU

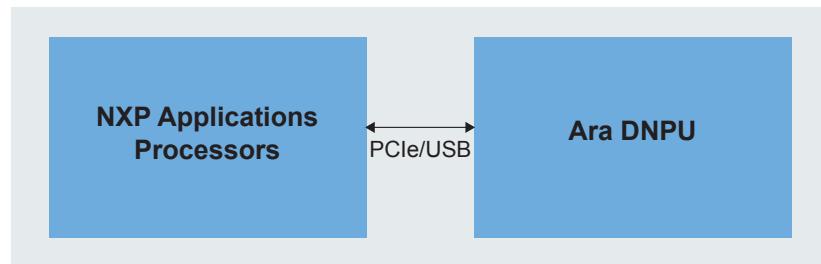
ARA-2-DNPU

Preproduction

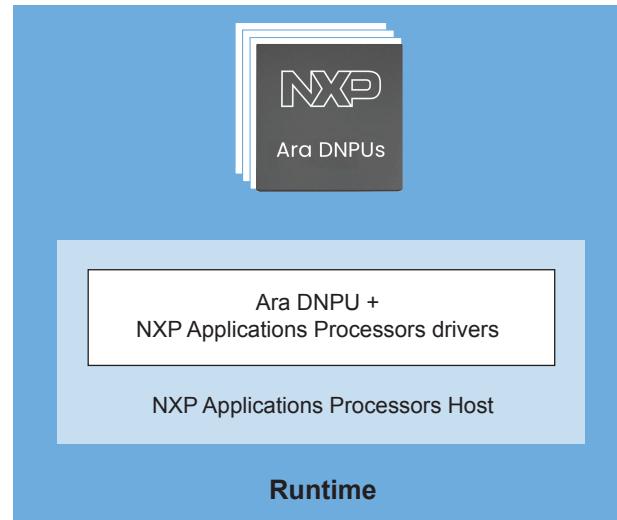
Last Updated: Nov 10, 2025

Ara-2ディスクリート・ニューラル・プロセッシング・ユニット (DNPU) によって、AIを利用するコンピューティング・システムや組込みシステムにおける生成AI、大規模言語モデル (LLM)、視覚言語モデル (VLM) のリアルタイム実行が可能になるため、レイテンシの短縮、運用コストの低減、データ・プライバシーの強化が実現します。その革新的なアーキテクチャと、バランスの取れたコンピューティング、大容量のオンチップ・メモリ、および高いオフチップ帯域幅が組み合わさることで、大規模なモデルを効率的に実行します。

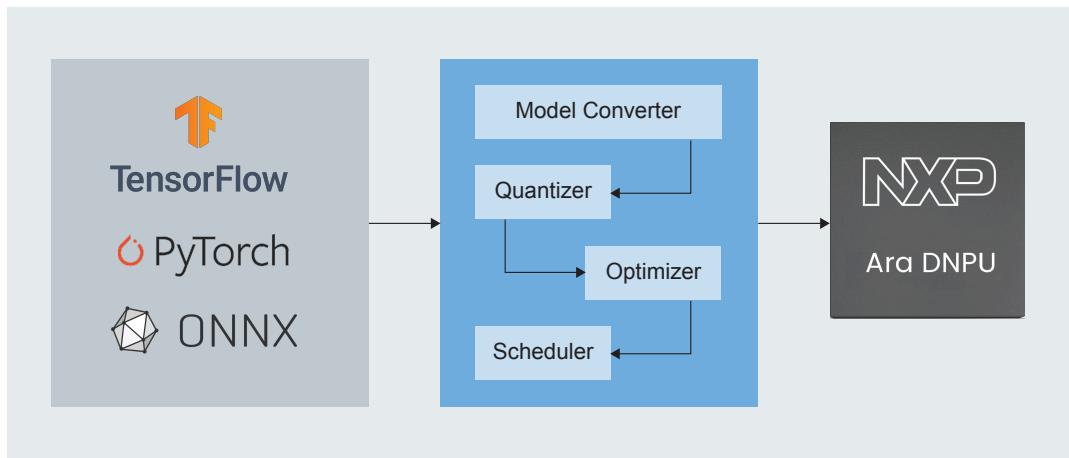
NXPアプリケーション・プロセッサとAra DNPUの接続 Block Diagram



Ara DNPUおよびアプリケーション・プロセッサのホストとドライバの統合 Block Diagram



機械学習の導入フロー Block Diagram



View additional information for [Ara-2 : リアルタイム生成AI向けディスクリートNPU](#).

Note: The information on this document is subject to change without notice.