

エッジ用AIチップ図鑑

中森 章

本稿では、エッジでのAIチップをかなり広範囲に捉えています。本来なら、性能レンジに応じて、エッジ、スマート・カメラ/スマート・スピーカ、センサ、積和演算が高速なマイコンなどの分野に分けて解説すべきかもしれませんが、

しかし、応用分野はユーザが決めるものという立場

から、それらを区別していません。厳密な分類を好む人には奇異に思えることがあるかもしれません。

また、アーキテクチャや実際の性能については不明な場合も多いため、類推が多い点もご容赦ください。

なかもり・あきら

1 Xavier (NVIDIA)

● NVIDIAの高性能エッジAI用GPU「Xavier」

NVIDIAはVoltaなどのサーバ側AI向けGPUを提供しています。

そのNVIDIAがエッジ側、というか、車載での応用を狙ったSoCであるXavier(エクゼビアと発音するらしい)を2016年の9月に発表しています。NVIDIAの発表によれば、Xavierは8個の64ビットArmv8-Aコア(Carmel)と512コアのVoltaアーキテクチャのGPU(CUDA)を1チップに統合しています。当時(2018年)では最高性能のGPUであるVoltaとホストCPUの組み合わせで、期待をもたせてくれました。

● アーキテクチャの概要

Xavierのブロック図を図1⁽¹⁾に示します。512個の

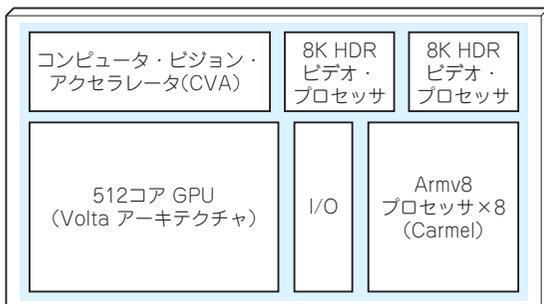


図1⁽¹⁾ NVIDIAのAIエッジ端末向けGPU「Xavier」

CUDAコアですので、同社のワークステーション向けAIアクセラレータGV100の1/10のコア数です。これは、8個のSM(Streaming Multiprocessor)に相当します。Voltaのスケージングの単位であるGPC(Graphics Processing Cluster)は1つで14個のSMを含みますから、8個というのはちょっと不自然です(SMやGPCの詳細は次章のVoltaの項目参照)。実際はどのようなのでしょうか。NVIDIAが公表しているXavierのチップ写真(らしきもの)を見る限り、SMの個数は14個にも見えますが、12個版Turing(T102)もありますし、8個版も可能でしょう。

Xavierは20TOPS(Tera Operations Per Second)の性能ながら20Wの消費電力だそうです。このXavierが1つあれば、現在の自動運転車用AIコンピュータのDRIVE PX2を1/4の電力で置き換えることができるそうです⁽²⁾。

▶ 自動運転用なのか

Xavierは自動運転の要として期待されていました。実際には「Jetson AGX Xavier」「Jetson Xavier NX」という汎用のAI評価ボードとして発売されました。例えば、NVIDIAの「Jetson AGX Xavier」の紹介⁽³⁾では、「宅配および物流ロボット、工場システム、産業用大型UAVなどの自律動作マシンに最適です」となっており、自動運転用としてはトーンが下がった印象もあります。自動車向けの役割は次機種Orin(後述)に譲ったのかもしれませんが。