Nvidia Tegra X2
Architecture and Design

Jeff Barker
Barry Wu

Rochester Institute of Technology

Spring 2017
Agenda

▶ Overview
▶ CPU Complex (4x ARM A57 + 2x Denver 2)
▶ GPU Complex (Pascal GPU)
▶ Applications
  ▶ Nvidia Jetson TX2
  ▶ Nvidia Drive PX2
▶ Closing Remarks and Questions
Nvidia Tegra X2 Overview

- Announced in April 2017 as credit card sized system-on-module
- Improved on previous generation Tegra X1 module with:
  - Two additional Denver 2 CPU Cores
  - 8GB 128-bit LPDDR4 RAM (compared to 4GB 64-bit)
  - 32GB eMMC storage (compared to 16GB)
  - Pascal GPU architecture (compared to Maxwell)
  - Improved video encoding and decoding
ARM CPU Cores

- Quad Core ARM A57 64-bit Processor Cluster
- Multi-thread optimized ARMv8 architecture
- Superscalar, variable-length, out-of-order pipeline
- Dynamic branch prediction with Branch Target Buffer (BTB) and Global History Buffer RAMs, a return stack, and an indirect predictor
- 48KB 3-way set-associative L1 Instruction cache per core
- 32KB 2-way set-associative L1 Data cache per core
- 2MB 16-way set-associative shared L2 cache
Denver 2 CPU Cores

- Dual Core Denver 2 64-bit Processor Cluster
- Single-thread optimized ARMv8 architecture
- 7-wide Superscalar architecture
- Dynamic branch prediction with Branch Target Buffer (BTB) and Global History Buffer RAMs, a return stack, and an indirect predictor
- 128KB 4-way set-associative L1 Instruction cache per core
- 64KB 4-way set-associative L1 Data cache per core
- 2MB 16-way set-associative shared L2 cache
Pascal GPU Micro-architecture
Pascal GPU

- 16-nm FinFET manufacturing process
- 2 Streaming Multiprocessors (SM)
- 4 Streaming Multiprocessor Blocks (SMP) per SM
- 32 CUDA cores per SMP (256 CUDA cores total)
- Simplified data path
- Improved scheduling and overlapping load/store instruction
- Improved addressing and page fault handling
Jetson TX2 System-On-Module Block Diagram
Jetson Carrier Board
Jetson Carrier Board (cont.)

- Interfaces
  - HDMI 2.0, 2x DSI display ports
  - CSI2 D-PHY 1.2 camera port
  - CAN, UART, SPI, I2C, I2S, GPIO
  - USB 3.0, USB 2.0
  - Gigabit Ethernet, 802.11ac WLAN, Bluetooth

- Applications
  - Computer Vision
  - Deep Learning/Neural Networks
  - Autonomous Drones
  - Robotics
Nvidia Drive PX2
Nvidia Drive PX2 (cont.)

- Fully autonomous driving
- Auto-pilot Functionality
  - Perception
  - Reasoning
  - Mapping
  - Driving
- Co-pilot Functionality
  - Face detection
  - Head tracking
  - Gaze tracking
  - Lip reading
- Multimedia Control
Closing Remarks

- Extremely capable embedded platform processor
  - Two high power single thread optimized cores
  - Four low power multiple thread optimized cores
  - Nvidia Pascal GPU
- High performance for low power draw
  - Claimed 1.5 TFLOPS at 15 Watts
  - Twice the performance per watt compared to the Tegra X1
- Wide range of applications, especially with Jetson
- Higher cost for increased performance compared to other embedded platforms
Questions?