Instituto Tecnológico de Costa Rica

Área Académica de Ingeniería en Computadores

Curso: CE-4302 Arquitectura de Computadores II

Taller 4

Realizado por:

David Eduardo Gómez Vargas – 2015028430

Profesor:

M.Sc. Ing. Jeferson González Gómez

Fecha: Cartago, Octubre 24, 2018

Evaluación Taller 4. CUDA

Micro-investigación

1. ¿Qué es CUDA?

NVIDIA creó una nueva familia de procesadores gráficos para computación de propósito general, contiene ALUs que se diseñaron para cumplir con los requisitos IEEE para aritmética de punto flotante de precisión simple y se diseñaron para usar un conjunto de instrucciones adaptado para computación general en lugar de específicamente para gráficos. Todas estas características de la arquitectura CUDA se agregaron para crear una GPU que sobresaliera en el cómputo, además de tener un buen desempeño en las tareas de gráficos tradicionales (Sanders & Kandrot, 2010).

2. ¿Qué es un kernel en CUDA y cómo se define?

Una función que se ejecuta en un GPU normalmente se llama un kernel (Sanders & Kandrot, 2010). Se define con el identificador __global__ para indicarle al compilador que es una función que va a ejecutarse en un GPU y no en el procesador normal.

3. ¿Cómo se maneja el trabajo a procesar en CUDA? ¿Cómo se asignan los hilos y bloques?

El trabajo de ejecución se realiza por medio de bloques que son copias de la función kernel, además de que cada bloque puede ejecutar una cantidad definida de hilos. Se definen estas dos variables cuando se llama al kernel <<bloques, hilos>>.

Se llama a la colección de bloques que lanzamos en la GPU un grid que puede ser una colección de bloques de una o dos dimensiones. Cada copia del kernel puede determinar qué bloque se está ejecutando con la variable integrada blockIdx. Del

mismo modo, puede determinar el tamaño del grid mediante el uso de la variable integrada gridDim (Sanders & Kandrot, 2010).

4. Investigue sobre la plataforma Jetson TX2 ¿cómo está compuesta la arquitectura de la plataforma a nivel de hardware?

El dispositivo NVIDIA Jetson TX2 es un sistema integrado en el módulo (SoM) con NVIDIA Denver2 + de cuatro núcleos ARM Cortex-A57, LPDDR4 de 8 GB de 128 bits y 256 núcleos integrados de GPU Pascal. Jetson TX2 es útil para implementar visión por computadora y deep learning, ejecuta Linux y proporciona más de 1TFLOPS de rendimiento de cómputo FP16 en menos de 7.5 vatios de potencia.

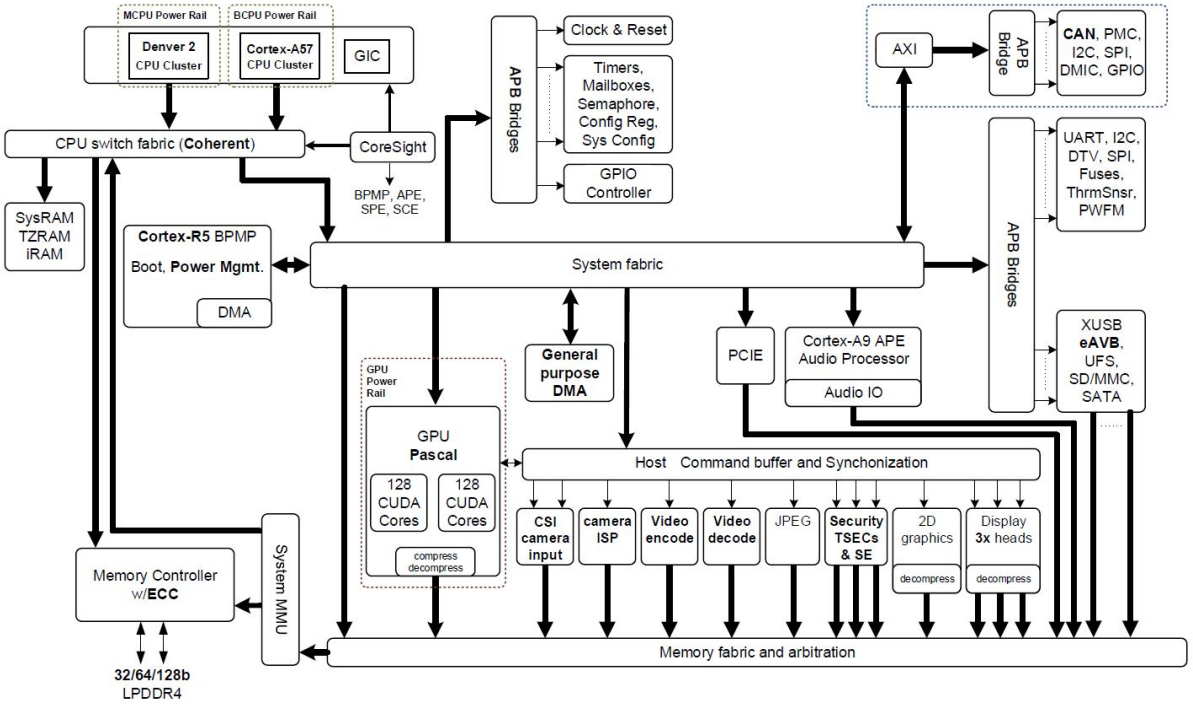

Figura 1. NVIDIA Jetson TX2 Tegra “Parker” SoC. Tomado de https://devblogs.nvidia.com/jetson-tx2-delivers-twice-intelligence-edge/

Ejemplo CUDA

1. Analice el código vecadd.cu y el Makefile correspondiente. A partir del análisis del código, extraiga cuáles son los pasos generales para la generación de aplicaciones utilizando CUDA.

- Crear la función kernel.
- Definir la cantidad de bloques e hilos.
- Definir la memoria para los parámetros de CUDA.
- Llamar la función kernel.

2. Analice el código fuente del kernel vecadd.cu. A partir del análisis del código, determine ¿Qué operación se realiza con los vectores de entrada? ¿Cómo se identifica cada elemento a ser procesado en paralelo y de qué forma se realiza el procesamiento paralelo?

Se realiza la suma por índice de ambos vectores. La identificación de cada elemento se realiza con la fórmula blockIdx.x * blockDim.x + threadIdx.x, que realiza la selección del índice por medio de la cantidad de bloques e hilos que se envía como parámetro.

3. Realice la ejecución de la aplicación vecadd. ¿Qué hace finalmente la aplicación?

Se realiza la suma por índice de ambos vectores en paralelo y serial.

4. Cambie la cantidad de hilos por bloque y el tamaño del vector. Compare el desempeño antes al menos 5 casos diferentes.

Tamaño del Vector	Tiempo Paralelo	Tiempo Serial	Hilos
1000000	0.000424	0.088763	250
10000000	0.000415	0.790896	500
10000000	0.000367	0.783707	1000
20000000	0.000546	1.483126	1000
20000000	0.000502	1.493025	500
20000000	0.00051	1.506799	250

Ejercicios prácticos

El código de los ejercicios prácticos se encuentra en el repositorio:

https://github.com/Daedgomez/CE4302-CUDA

Referencias

Sanders, J., & Kandrot, E. (2010). *CUDA by example: an introduction to general-purpose GPU programming*. Addison-Wesley Professional.

Información sobre Jetson tx2 tomada de:

https://devblogs.nvidia.com/jetson-tx2-delivers-twice-intelligence-edge/