

The background of the slide is a dark orange color with a white circuit board pattern. The pattern consists of numerous parallel lines and various shapes, including circles and rectangles, representing traces and components on a PCB.

Heterogeneous
Parallel
Programming

Lecture 1.6

Introduction to CUDA

- Kernel-Based SPMD Parallel Programming

Wen-mei Hwu - University of Illinois at Urbana-Champaign



Objective

- To learn the basic concepts involved in a simple CUDA kernel function
 - Declaration
 - Built-in variables
 - Thread index to data index mapping

Example: Vector Addition Kernel

Device Code

```
// Compute vector sum C = A+B  
// Each thread performs one pair-wise addition
```

```
__global__
```

```
void vecAddKernel(float* A, float* B, float*  
    C, int n)
```

```
{
```

```
int i = threadIdx.x+blockDim.x*blockIdx.x;
```

```
    if(i<n) C[i] = A[i] + B[i];
```

```
}
```

Example: Vector Addition Kernel (Host Code)

```
int vecAdd(float* h_A, float* h_B, float* h_C, int n)
{
    // d_A, d_B, d_C allocations and copies omitted
    // Run ceil(n/256.0) blocks of 256 threads each
    vecAddKernel<<<ceil(n/256.0),256>>>(d_A, d_B, d_C, n);
}
```

Host Code

More on Kernel Launch (Host Code)

```
int vecAdd(float* h_A, float* h_B, float* h_C, int n)
{
    dim3 DimGrid((n-1)/256 + 1, 1, 1);
    dim3 DimBlock(256, 1, 1);
    vecAddKernel<<<DimGrid,DimBlock>>>(d_A, d_B, d_C, n);
}
```

Host Code

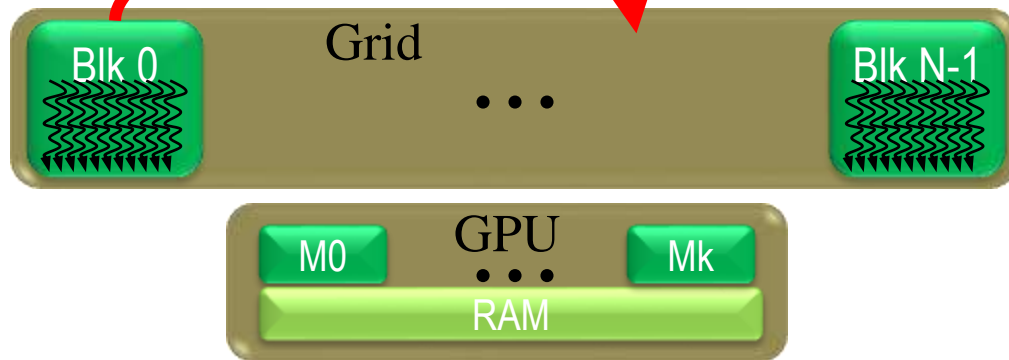
This makes sure that there are enough threads to cover all elements.



Kernel execution in a nutshell

```
__host__  
Void vecAdd(...)  
{  
    dim3 DimGrid(ceil(n/256.0),1,1);  
    dim3 DimBlock(256,1,1);  
    vecAddKernel<<<DimGrid,DimBlock>>>(d_A,d_B  
,d_C,n);  
}
```

```
__global__  
void vecAddKernel(float *A,  
                  float *B, float *C, int n)  
{  
    int i = blockIdx.x * blockDim.x  
          + threadIdx.x;  
    if( i < n ) C[i] = A[i]+B[i];  
}
```



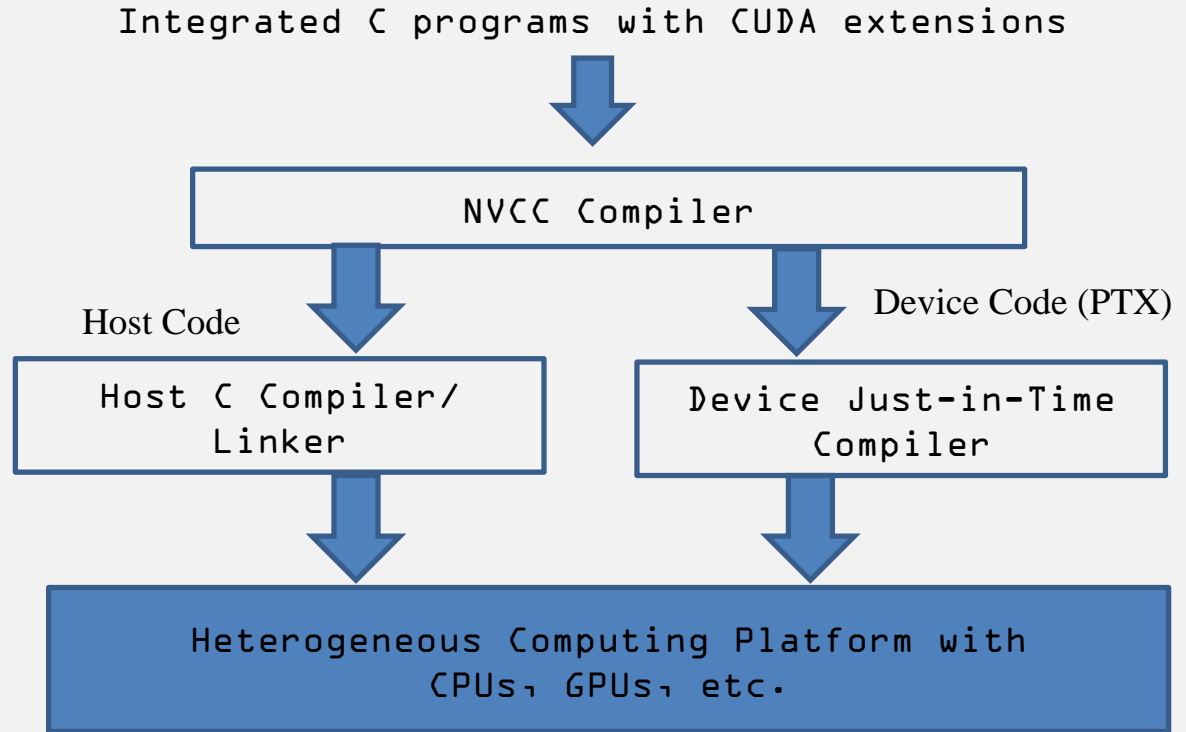
More on CUDA Function Declarations

	Executed on the:	Only callable from the:
<code>__device__ float DeviceFunc()</code>	device	device
<code>__global__ void KernelFunc()</code>	device	host
<code>__host__ float HostFunc()</code>	host	host

- `__global__` defines a kernel function
 - Each “__” consists of two underscore characters
 - A kernel function must return `void`
- `__device__` and `__host__` can be used together
- `__host__` is optional if used alone



Compiling A CUDA Program





To learn more, please read
Chapter 3.

