

# Computer Architecture:

## HW1: 1D Data Parallelism

Computer Architecture 2019 (Prof. Chih-Wei Liu)

CUDA Tutorial HW1

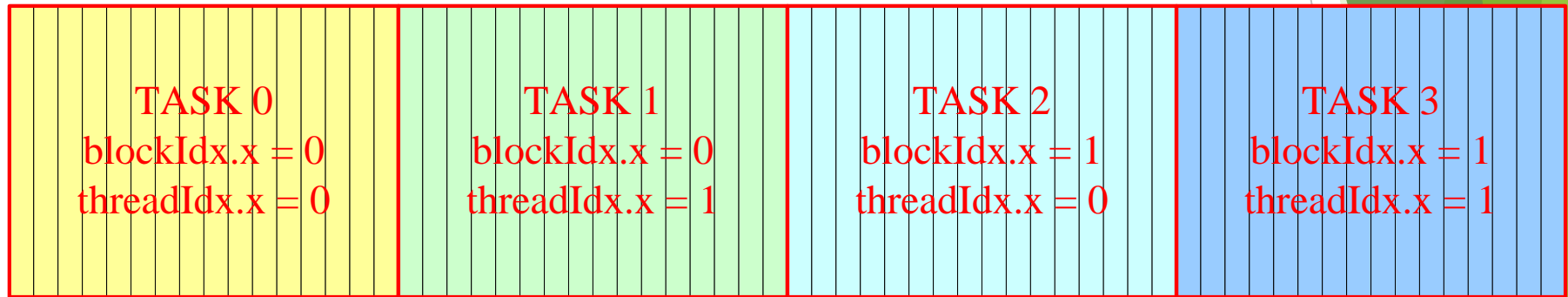
# 回顧CUDA的平行運算

- ▶ 平行程式可被指定的block與thread數量切割成若干部分
  - ▶ `Cuda_Kernel <<<nBlk,nTid>>>(..)`
  - ▶ 工作(Tasks)總數 =  $nBlk * nTid$
- ▶ 每個工作(Task)擁有不同的block id與thread id
  - ▶ 被處理的資料也依據這些id做適度切割
- ▶ 這些工作(Task)是動態排程地執行
  - ▶ 在同一個block下的工作(Tasks)是會被安排在同一個Streaming multiprocessor (SM)被執行

# 一維資料切割-方法一

When processing 64 integer data:

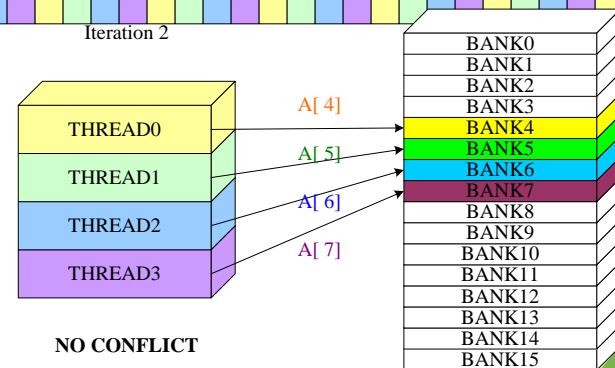
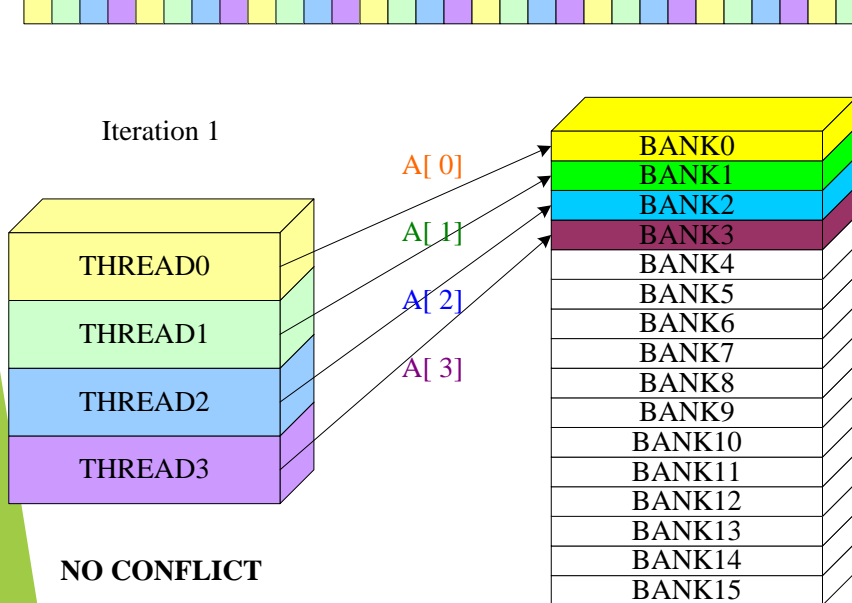
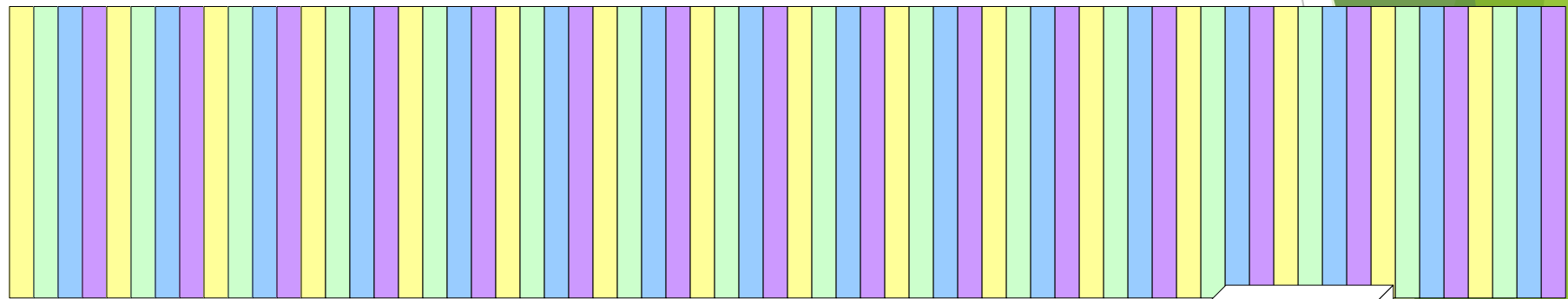
```
cuda_kernel<<<2, 2>>>(...)
```



# 一維資料切割-方法二

When processing 64 integer data:

```
cuda_kernel<<<1, 4>>>(...)
```



# Example 2

## cuda\_wrapper.cu

```
#include <cuda.h>
#include <cuda_runtime.h>
#include <stdio.h>
#include "IndexSave.h"

__global__ void someKernel(int N, IndexSave *dInd)
{
    int i=0;
    int TotalThread = blockDim.x*gridDim.x;
    int stripe = N / TotalThread;
    int head = (blockIdx.x*blockDim.x + threadIdx.x)*stripe;
    int LoopLim = head+stripe;

    for(i=head ; i<LoopLim ; i++ )
    {
        dInd[i].blockInd_x = blockIdx.x;
        dInd[i].threadInd_x = threadIdx.x;
        dInd[i].head = head;
        dInd[i].stripe = stripe;
    }
}
```

紀錄task執行的情況

方法一  
的一維  
資料切  
割

```
#include <cuda.h>
#include <cuda_runtime.h>
#include <stdio.h>
#include "IndexSave.h"

__global__ void someKernelIntleave(int N, IndexSave *dInd)
{
    int i=0;
    int stripe = blockDim.x*gridDim.x;
    int head = blockIdx.x*blockDim.x + threadIdx.x;

    for(i=head ; i<N ; i+=stripe )
    {
        dInd[i].blockInd_x = blockIdx.x;
        dInd[i].threadInd_x = threadIdx.x;
        dInd[i].head = head;
        dInd[i].stripe = stripe;
    }
}
```

紀錄task執行的情況

方法二  
的一維  
資料切  
割

# HW1:一維亂數序列產生器

## 程式說明&作業規定

### ► 程式說明

- 每個亂數序列，由rand()函式產生原始亂數(random seeds)，再將每個亂數自乘**LOOP**次(在parameter.h定義)得到結果(CPU side程式碼請參考function\_1(...))
- 相關參數請不要更動(像是block size、thread numbers、常數)

### ► 作業規定

- 請補上**空缺的程式碼**(留有註解暗示)
- 請做出**方法一**與**方法二**的資料切割方式
- 請將kernel執行的blockId與threadId使用**example2**的方式，儲存在IndexSave的資料結構中(請參考**example2**程式碼)，助教僅檢查IndexSave中此兩變量的值是否正確。
- 繳交檔案：main.cu，方法一的 device1.cu，方法二的 device2.cu