

# 기말고사 문제2

---

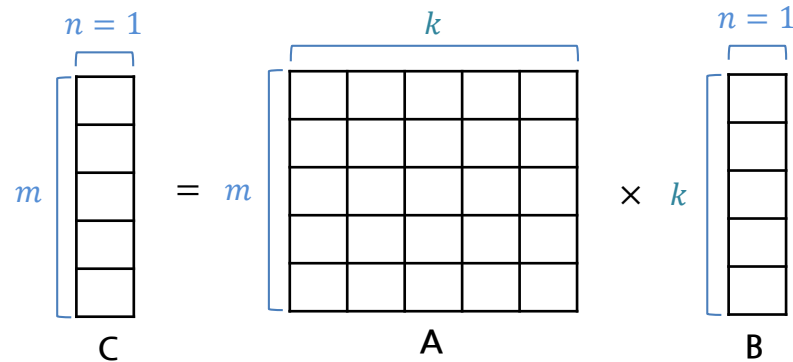
## ■ 문제 내용

- 기말고사 문제2 게시판에서 finalterm\_prog2.zip 파일을 다운로드하여 압축을 푼 뒤 finalterm\_prog2.cu 파일 내에서 “// problem 1”로 표시된 부분을 수정하여 다음 페이지와 같이 shared memory 기반의 행렬과 벡터의 곱셈을 수행하는 CUDA 커널 구현(20점)
  - Device memory 접근 최적화 적용

# 기말고사 문제2

## ■ 행렬과 벡터의 곱셈(matrix and column vector multiplication)

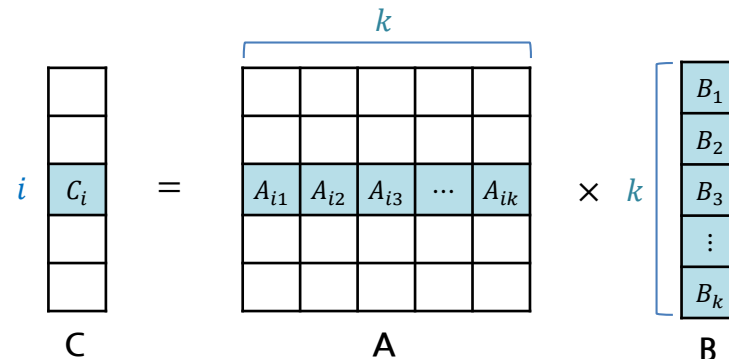
- 행렬  $C(m \times 1) = \text{행렬 } A(m \times k) \times \text{벡터 } B(k \times 1)$



- 벡터  $C$ 의 원소

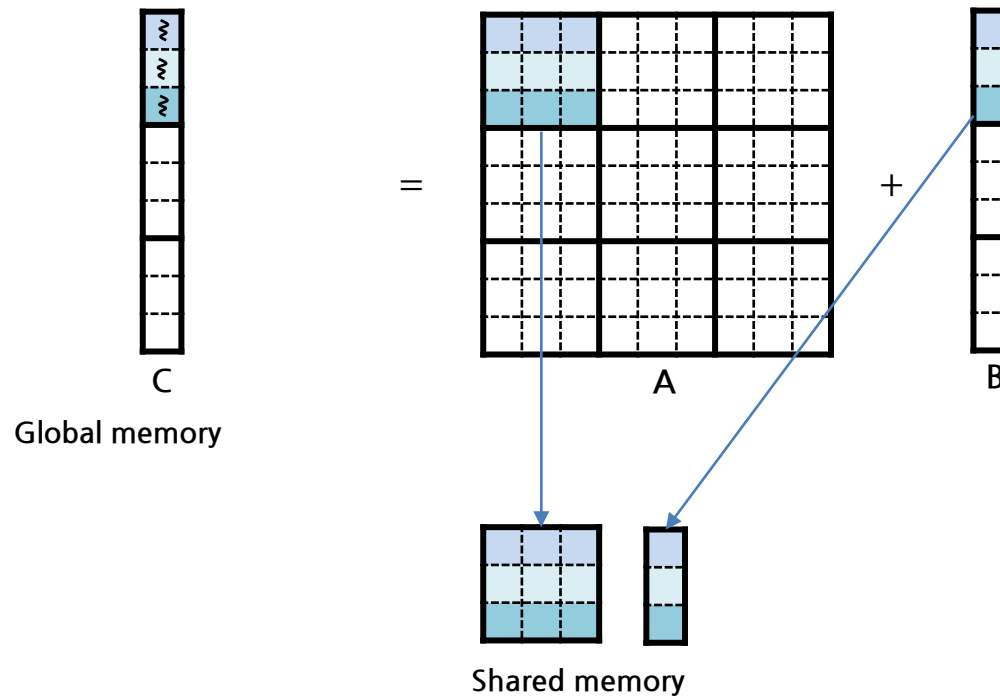
$$C_i = \sum_{p=1}^k A_{ip} B_p$$

$$= A_{i1}B_1 + A_{i2}B_2 + A_{i3}B_3 + \cdots + A_{ik}B_k$$



# 기말고사 문제2

- [예시] Global memory에 있는 행렬 A와 B를 데이터 블록 단위로 분할하여 shared memory에 올려서 행렬 덧셈 연산



# 기말고사 문제2

## ■ 프로그램 실행 결과

```
$ nvcc -o finalterm_prog2 finalterm_prog2.cu DS_timer.cpp
$ qsub job_gpu.sh
$ cat finalterm_prog2.log
*           DS_timer Report           *
* The number of timer = 5, counter = 5
**** Timer report ****
CUDA Total : 2.44700 ms
Computation on device (GPU) : 1.98500 ms
Memory copy: host -> device : 0.45000 ms
Memory copy: device -> host : 0.01200 ms
Computation on host (CPU) : 3.97900 ms
**** Counter report ****
*           End of the report           *
The matrix multiplication on the device (GPU) is the same as the the host (CPU)
```

## ■ 제출 방법

- 작성한 finalterm\_prog2.cu 파일의 이름을 “이름\_학번\_finalterm\_prog2.cu 파일” 형태(예, 홍길동\_202011111\_finalterm\_prog2.cu)로 변경하여 **기말고사 문제2 게시판에 제출**

## ■ 채점

- 부분점수가 적용됨
- 제출하지 않거나 프로그램을 구현하지 않으면 0점
- 프로그램이 정상적으로 동작하면 20점