



Faculty of Information Technology

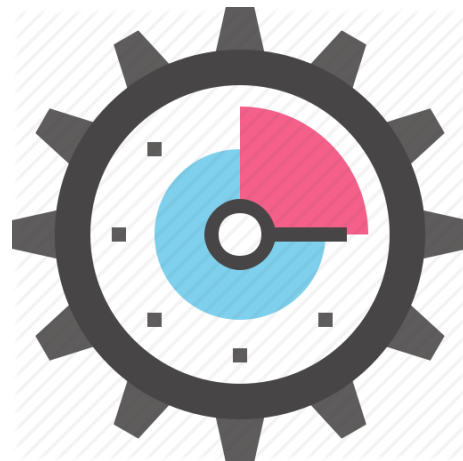
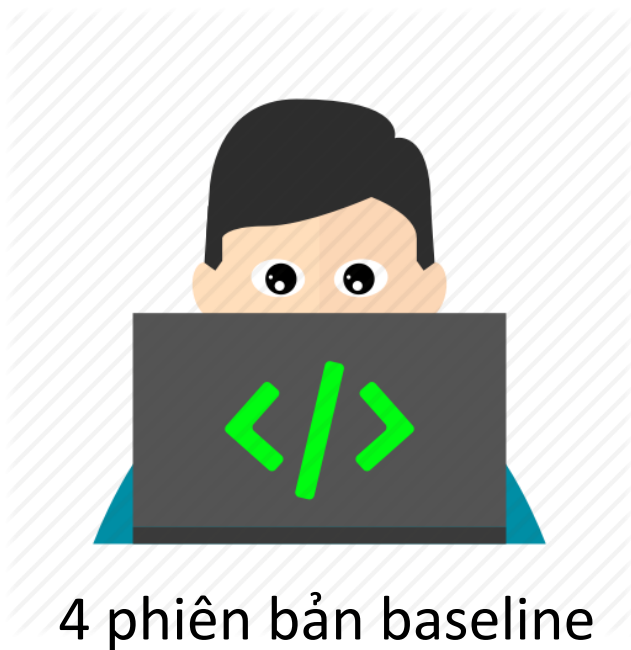
Môn học : **Lập trình song song trên GPU**
Lớp CQ2016/2 – Học kỳ I/2019-2020
GV: **Trần Trung Kiên**

TỐI ƯU HÓA RADIX SORT TRÊN GPU

<Báo cáo đồ án cuối kỳ ngày 10-01-2020>

Nhóm 04: Đặng Phương Nam – Lê Minh Nghĩa

NỘI DUNG CHÍNH



nBits = 4, hist block size = 256 và scan block size = 256

Version	Description	Run Time (ms)
baseline01	Cài đặt tuần tự thuật toán Radix Sort tuần tự.	2283.303
baseline02	Cài đặt song song 2 bước histogram và scan của thuật toán Radix Sort tuần tự.	1546.865
baseline03	Cài đặt song song thuật toán Radix Sort với k = 1 bit.	584.170
baseline04	Cài đặt tuần tự ý tưởng thuật toán Radix Sort song song trong bài báo của NVIDIA.	2779.344
thrust	Thuật toán sort song song trong thư viện Thrust.	80.575

4 PHIÊN BẢN BASE LINE



```
Input size: 16777217
Num bits per digit: 4
Hist block size: 256, scan block size: 256

Radix sort by host
Time: 2283.303 ms
```

Baseline01

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 256, scan block size: 256

Radix sort by host
Time: 2280.748 ms

Radix sort by device
Time: 1546.865 ms
CORRECT :)
```

Baseline02

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 256, scan block size: 256

Radix sort by host
Time: 2293.046 ms

Radix sort by device
Time: 584.170 ms
CORRECT :)
```

Baseline03

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 256, scan block size: 256

Radix sort by thrust
Time: 71.611 ms

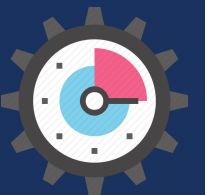
Radix sort by host (#paper: sequential radix sort)
Time: 2790.938 ms
CORRECT :)
```

Baseline04

nBits = 4, hist block size = 256 và scan block size = 256

Version	Description	Run Time # Step 1 (ms/loop)	Run Time # Step 2 (ms/loop)	Run Time # Step 3 (ms/loop)	Run Time Total (ms)
baseline04-v1	Cài đặt tuần tự ý tưởng thuật toán Radix Sort song song trong bài báo của NVIDIA.	135.680	5.601	201.208	2772.958
baseline04-v2	Cài đặt song song bước tính local histogram và exclusive scan trên mảng một chiều gồm các cột nối với nhau.	2.802	0.883	224.110	1866.060
baseline04-v3	Cài đặt song song bước scatter dữ liệu xuống mảng output, với in-block-sort dùng work-inefficient scan .	2.671	0.868	49.173	470.487
baseline04-v4	Cài đặt song song bước scatter dữ liệu xuống mảng output, với in-block-sort dùng work-inefficient scan (mỗi thread xử lý 2 phần tử).	2.678	0.872	45.780	442.680

QUÁ TRÌNH TỐI ƯU HÓA CHO PHIÊN BẢN BASELINE04

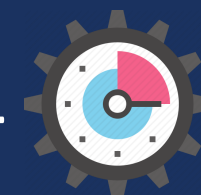


nBits = 4, hist block size = 256 và scan block size = 256

Version	Description	Run Time # Step 1 (ms/loop)	Run Time # Step 2 (ms/loop)	Run Time # Step 3 (ms/loop)	Run Time Total (ms)
baseline04-v5	Cài đặt song song bước scatter dữ liệu xuống mảng output, với in-block-sort dùng work-efficient scan (mỗi thread xử lý 2 phần tử).	2.798	0.867	48.805	466.951
baseline04-v6	Cài đặt song song bước scatter dữ liệu xuống mảng output, với in-block-sort dùng warp shuffles scan .	2.671	0.867	22.801	258.824
baseline04-v7	Cài đặt song song bước scatter dữ liệu xuống mảng output, với in-block-sort dùng warp shuffles scan . (mỗi thread xử lý 2 phần tử).	2.673	0.868	20.783	243.245
thrust	Thuật toán sort song song trong thư viện Thrust.				72.695

Phiên bản tốt nhất mà nhóm cài đặt được là phiên bản **baseline04-v7**.

QUÁ TRÌNH TỐI ƯU HÓA CHO PHIÊN BẢN BASELINE04



```

Radix sort by thrust
Time: 71.374 ms

Radix sort by host (#paper: sequential radix sort)
#1 (iteration):
+ Step 1. Local histogram. Time: 136.402 ms
+ Step 2. Exclusive scan. Time: 6.192 ms
+ Step 3. Scatter. Time: 214.075 ms
#2 (iteration):
+ Step 1. Local histogram. Time: 135.721 ms
+ Step 2. Exclusive scan. Time: 5.689 ms
+ Step 3. Scatter. Time: 199.286 ms
#3 (iteration):
+ Step 1. Local histogram. Time: 135.735 ms
+ Step 2. Exclusive scan. Time: 5.271 ms
+ Step 3. Scatter. Time: 199.408 ms
#4 (iteration):
+ Step 1. Local histogram. Time: 135.387 ms
+ Step 2. Exclusive scan. Time: 5.681 ms
+ Step 3. Scatter. Time: 199.420 ms
#5 (iteration):
+ Step 1. Local histogram. Time: 135.835 ms
+ Step 2. Exclusive scan. Time: 5.342 ms
+ Step 3. Scatter. Time: 199.402 ms
#6 (iteration):
+ Step 1. Local histogram. Time: 135.352 ms
+ Step 2. Exclusive scan. Time: 5.681 ms
+ Step 3. Scatter. Time: 199.481 ms
#7 (iteration):
+ Step 1. Local histogram. Time: 135.472 ms
+ Step 2. Exclusive scan. Time: 5.290 ms
+ Step 3. Scatter. Time: 199.460 ms
#8 (iteration):
+ Step 1. Local histogram. Time: 135.537 ms
+ Step 2. Exclusive scan. Time: 5.666 ms
+ Step 3. Scatter. Time: 199.133 ms
Time: 2772.958 ms
CORRECT :)

```

Baseline04-v1

```

Radix sort by thrust
Time: 71.495 ms

Radix sort by device
#1 (iteration):
+ Step 1. Local histogram. Time: 2.743 ms
+ Step 2. Exclusive scan. Time: 0.894 ms
+ Step 3. Scatter. Time: 231.174 ms
#2 (iteration):
+ Step 1. Local histogram. Time: 2.780 ms
+ Step 2. Exclusive scan. Time: 0.880 ms
+ Step 3. Scatter. Time: 222.343 ms
#3 (iteration):
+ Step 1. Local histogram. Time: 2.777 ms
+ Step 2. Exclusive scan. Time: 0.885 ms
+ Step 3. Scatter. Time: 221.448 ms
#4 (iteration):
+ Step 1. Local histogram. Time: 2.783 ms
+ Step 2. Exclusive scan. Time: 0.883 ms
+ Step 3. Scatter. Time: 222.229 ms
#5 (iteration):
+ Step 1. Local histogram. Time: 2.785 ms
+ Step 2. Exclusive scan. Time: 0.882 ms
+ Step 3. Scatter. Time: 222.343 ms
#6 (iteration):
+ Step 1. Local histogram. Time: 2.779 ms
+ Step 2. Exclusive scan. Time: 0.882 ms
+ Step 3. Scatter. Time: 220.099 ms
#7 (iteration):
+ Step 1. Local histogram. Time: 2.780 ms
+ Step 2. Exclusive scan. Time: 0.884 ms
+ Step 3. Scatter. Time: 221.839 ms
#8 (iteration):
+ Step 1. Local histogram. Time: 2.995 ms
+ Step 2. Exclusive scan. Time: 0.876 ms
+ Step 3. Scatter. Time: 231.405 ms
Time: 1866.060 ms
CORRECT :)

```

Baseline04-v2

```

Radix sort by thrust
Time: 72.081 ms

Radix sort by device
#1 (iteration):
+ Step 1. Local histogram. Time: 2.708 ms
+ Step 2. Exclusive scan. Time: 0.871 ms
+ Step 3. Scatter. Time: 49.260 ms
#2 (iteration):
+ Step 1. Local histogram. Time: 2.635 ms
+ Step 2. Exclusive scan. Time: 0.867 ms
+ Step 3. Scatter. Time: 49.266 ms
#3 (iteration):
+ Step 1. Local histogram. Time: 2.632 ms
+ Step 2. Exclusive scan. Time: 0.865 ms
+ Step 3. Scatter. Time: 49.264 ms
#4 (iteration):
+ Step 1. Local histogram. Time: 2.636 ms
+ Step 2. Exclusive scan. Time: 0.869 ms
+ Step 3. Scatter. Time: 49.265 ms
#5 (iteration):
+ Step 1. Local histogram. Time: 2.636 ms
+ Step 2. Exclusive scan. Time: 0.867 ms
+ Step 3. Scatter. Time: 49.264 ms
#6 (iteration):
+ Step 1. Local histogram. Time: 2.639 ms
+ Step 2. Exclusive scan. Time: 0.878 ms
+ Step 3. Scatter. Time: 49.261 ms
#7 (iteration):
+ Step 1. Local histogram. Time: 2.637 ms
+ Step 2. Exclusive scan. Time: 0.867 ms
+ Step 3. Scatter. Time: 49.268 ms
#8 (iteration):
+ Step 1. Local histogram. Time: 2.848 ms
+ Step 2. Exclusive scan. Time: 0.865 ms
+ Step 3. Scatter. Time: 48.537 ms
Time: 470.487 ms
CORRECT :)

```

Baseline04-v3


```
Radix sort by thrust
Time: 71.516 ms
```

```
Radix sort by device
```

```
#1 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.700 ms
+ Step 2. Exclusive scan. Time: 0.882 ms
+ Step 3. Scatter. Time: 45.861 ms
```

```
#2 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.658 ms
+ Step 2. Exclusive scan. Time: 0.865 ms
+ Step 3. Scatter. Time: 45.843 ms
```

```
#3 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.653 ms
+ Step 2. Exclusive scan. Time: 0.861 ms
+ Step 3. Scatter. Time: 45.860 ms
```

```
#4 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.635 ms
+ Step 2. Exclusive scan. Time: 0.868 ms
+ Step 3. Scatter. Time: 45.845 ms
```

```
#5 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.636 ms
+ Step 2. Exclusive scan. Time: 0.876 ms
+ Step 3. Scatter. Time: 45.853 ms
```

```
#6 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.648 ms
+ Step 2. Exclusive scan. Time: 0.866 ms
+ Step 3. Scatter. Time: 45.849 ms
```

```
#7 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.639 ms
+ Step 2. Exclusive scan. Time: 0.875 ms
+ Step 3. Scatter. Time: 45.863 ms
```

```
#8 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.853 ms
+ Step 2. Exclusive scan. Time: 0.882 ms
+ Step 3. Scatter. Time: 45.265 ms
```

```
Time: 442.680 ms
```

```
CORRECT :)
```

Baseline04-v4

```
Radix sort by thrust
Time: 71.367 ms
```

```
Radix sort by device
```

```
#1 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.702 ms
+ Step 2. Exclusive scan. Time: 0.872 ms
+ Step 3. Scatter. Time: 48.892 ms
```

```
#2 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.636 ms
+ Step 2. Exclusive scan. Time: 0.866 ms
+ Step 3. Scatter. Time: 48.878 ms
```

```
#3 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.635 ms
+ Step 2. Exclusive scan. Time: 0.867 ms
+ Step 3. Scatter. Time: 48.882 ms
```

```
#4 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.636 ms
+ Step 2. Exclusive scan. Time: 0.865 ms
+ Step 3. Scatter. Time: 48.886 ms
```

```
#5 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.651 ms
+ Step 2. Exclusive scan. Time: 0.878 ms
+ Step 3. Scatter. Time: 48.878 ms
```

```
#6 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.638 ms
+ Step 2. Exclusive scan. Time: 0.865 ms
+ Step 3. Scatter. Time: 48.887 ms
```

```
#7 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.637 ms
+ Step 2. Exclusive scan. Time: 0.864 ms
+ Step 3. Scatter. Time: 48.874 ms
```

```
#8 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.851 ms
+ Step 2. Exclusive scan. Time: 0.861 ms
+ Step 3. Scatter. Time: 48.264 ms
```

```
Time: 466.951 ms
```

```
CORRECT :)
```

Baseline04-v5

```
Radix sort by thrust
Time: 71.573 ms
```

```
Radix sort by device
```

```
#1 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.703 ms
+ Step 2. Exclusive scan. Time: 0.882 ms
+ Step 3. Scatter. Time: 22.888 ms
```

```
#2 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.635 ms
+ Step 2. Exclusive scan. Time: 0.863 ms
+ Step 3. Scatter. Time: 22.875 ms
```

```
#3 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.638 ms
+ Step 2. Exclusive scan. Time: 0.866 ms
+ Step 3. Scatter. Time: 22.865 ms
```

```
#4 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.635 ms
+ Step 2. Exclusive scan. Time: 0.861 ms
+ Step 3. Scatter. Time: 22.891 ms
```

```
#5 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.637 ms
+ Step 2. Exclusive scan. Time: 0.865 ms
+ Step 3. Scatter. Time: 22.863 ms
```

```
#6 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.638 ms
+ Step 2. Exclusive scan. Time: 0.866 ms
+ Step 3. Scatter. Time: 22.873 ms
```

```
#7 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.637 ms
+ Step 2. Exclusive scan. Time: 0.869 ms
+ Step 3. Scatter. Time: 22.870 ms
```

```
#8 (iteration):
```

```
+ Step 1. Local histogram. Time: 2.848 ms
+ Step 2. Exclusive scan. Time: 0.863 ms
+ Step 3. Scatter. Time: 22.282 ms
```

```
Time: 258.824 ms
```

```
CORRECT :)
```

Baseline04-v6


```
Radix sort by thrust
Time: 72.695 ms

Radix sort by device
#1 (iteration):
+ Step 1. Local histogram. Time: 2.703 ms
+ Step 2. Exclusive scan. Time: 0.873 ms
+ Step 3. Scatter. Time: 20.793 ms
#2 (iteration):
+ Step 1. Local histogram. Time: 2.636 ms
+ Step 2. Exclusive scan. Time: 0.866 ms
+ Step 3. Scatter. Time: 20.802 ms
#3 (iteration):
+ Step 1. Local histogram. Time: 2.636 ms
+ Step 2. Exclusive scan. Time: 0.883 ms
+ Step 3. Scatter. Time: 20.793 ms
#4 (iteration):
+ Step 1. Local histogram. Time: 2.650 ms
+ Step 2. Exclusive scan. Time: 0.866 ms
+ Step 3. Scatter. Time: 20.806 ms
#5 (iteration):
+ Step 1. Local histogram. Time: 2.635 ms
+ Step 2. Exclusive scan. Time: 0.870 ms
+ Step 3. Scatter. Time: 20.803 ms
#6 (iteration):
+ Step 1. Local histogram. Time: 2.633 ms
+ Step 2. Exclusive scan. Time: 0.866 ms
+ Step 3. Scatter. Time: 20.806 ms
#7 (iteration):
+ Step 1. Local histogram. Time: 2.638 ms
+ Step 2. Exclusive scan. Time: 0.865 ms
+ Step 3. Scatter. Time: 20.794 ms
#8 (iteration):
+ Step 1. Local histogram. Time: 2.849 ms
+ Step 2. Exclusive scan. Time: 0.859 ms
+ Step 3. Scatter. Time: 20.669 ms
Time: 243.245 ms
CORRECT :)
```

Phiên bản baseline04-v7 với nBits = 4

Hist block size	Scan block size	Run Time (ms)
2	2	6062.347
2	4	4408.337
2	16	2994.594
2	256	2761.198
4	2	4108.598
4	4	3062.705
4	16	2463.188
4	256	2341.852

Hist block size	Scan block size	Run Time (ms)
16	2	2068.141
16	4	1959.383
16	16	1368.400
16	256	1149.069
256	2	381.797
256	4	309.685
256	16	256.146
256	256	242.113

CHẠY PHIÊN BẢN TỐT NHẤT – BLOCK SIZE KHÁC NHAU



Phiên bản baseline04-v7 với nBits = 4 và hist block size = 2

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 2, scan block size: 2

Radix sort by thrust
Time: 71.623 ms

Radix sort by device
Time: 6062.347 ms
CORRECT :)
```

Scan block size = 2

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 2, scan block size: 4

Radix sort by thrust
Time: 71.507 ms

Radix sort by device
Time: 4408.337 ms
CORRECT :)
```

Scan block size = 4

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 2, scan block size: 16

Radix sort by thrust
Time: 71.653 ms

Radix sort by device
Time: 2994.594 ms
CORRECT :)
```

Scan block size = 16

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 2, scan block size: 256

Radix sort by thrust
Time: 71.622 ms

Radix sort by device
Time: 2761.198 ms
CORRECT :)
```

Scan block size = 256

Phiên bản baseline04-v7 với nBits = 4 và hist block size = 4

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 4, scan block size: 2

Radix sort by thrust
Time: 71.465 ms

Radix sort by device
Time: 4108.598 ms
CORRECT :)
```

Scan block size = 2

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 4, scan block size: 4

Radix sort by thrust
Time: 71.877 ms

Radix sort by device
Time: 3062.705 ms
CORRECT :)
```

Scan block size = 4

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 4, scan block size: 16

Radix sort by thrust
Time: 71.365 ms

Radix sort by device
Time: 2463.188 ms
CORRECT :)
```

Scan block size = 16

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 4, scan block size: 256

Radix sort by thrust
Time: 71.588 ms

Radix sort by device
Time: 2341.852 ms
CORRECT :)
```

Scan block size = 256

Phiên bản baseline04-v7 với nBits = 4 và hist block size = 16

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 16, scan block size: 2

Radix sort by thrust
Time: 71.409 ms

Radix sort by device
Time: 2068.141 ms
CORRECT :)
```

Scan block size = 2

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 16, scan block size: 4

Radix sort by thrust
Time: 71.489 ms

Radix sort by device
Time: 1959.383 ms
CORRECT :)
```

Scan block size = 4

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 16, scan block size: 16

Radix sort by thrust
Time: 71.531 ms

Radix sort by device
Time: 1368.400 ms
CORRECT :)
```

Scan block size = 16

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 16, scan block size: 256

Radix sort by thrust
Time: 71.532 ms

Radix sort by device
Time: 1149.069 ms
CORRECT :)
```

Scan block size = 256

Phiên bản baseline04-v7 với nBits = 4 và hist block size = 256

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 256, scan block size: 2

Radix sort by thrust
Time: 71.682 ms

Radix sort by device
Time: 381.797 ms
CORRECT :)
```

Scan block size = 2

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 256, scan block size: 4

Radix sort by thrust
Time: 71.552 ms

Radix sort by device
Time: 309.685 ms
CORRECT :)
```

Scan block size = 4

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 256, scan block size: 16

Radix sort by thrust
Time: 71.465 ms

Radix sort by device
Time: 256.146 ms
CORRECT :)
```

Scan block size = 16

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 256, scan block size: 256

Radix sort by thrust
Time: 71.657 ms

Radix sort by device
Time: 242.113 ms
CORRECT :)
```

Scan block size = 256

Phiên bản baseline04-v7 với hist block size = 256 và scan block size = 256

```
Input size: 16777217
Num bits per digit: 1
Hist block size: 256, scan block size: 256

Radix sort by thrust
Time: 71.458 ms

Radix sort by device
Time: 417.621 ms
CORRECT :)
```

```
Input size: 16777217
Num bits per digit: 4
Hist block size: 256, scan block size: 256

Radix sort by thrust
Time: 71.566 ms

Radix sort by device
Time: 241.928 ms
CORRECT :)
```

nBits	Run Time (ms)
1	417.621
2	289.623
4	241.928
8	414.114

```
Input size: 16777217
Num bits per digit: 2
Hist block size: 256, scan block size: 256

Radix sort by thrust
Time: 71.596 ms

Radix sort by device
Time: 289.623 ms
CORRECT :)
```

```
Input size: 16777217
Num bits per digit: 8
Hist block size: 256, scan block size: 256

Radix sort by thrust
Time: 71.452 ms

Radix sort by device
Time: 313.114 ms
CORRECT :)
```

CHẠY PHIÊN BẢN TỐT NHẤT – NBITS KHÁC NHAU



Phiên bản baseline04-v7
nBits = 4, hist block size = 256 và scan block size = 256

	Run Time (ms)
baseline04-v7	63.893
thrush	51.307

```
*****GPU info*****
Name: GeForce RTX 2080 Ti
Compute capability: 7.5
Num SMs: 68
Max num threads per SM: 1024
Max num warps per SM: 32
GMEM: 11523260416 byte
SMEM per SM: 65536 byte
SMEM per block: 49152 byte
*****

Input size: 16777217

Num bits per digit: 4

Hist block size: 256, scan block size: 256

Radix sort by thrust
Time: 51.307 ms

Radix sort by device
Time: 63.893 ms
CORRECT :)
```

CHẠY PHIÊN BẢN TỐT NHẤT – GPU CÓ 68 SM



- Work-efficient:

Adam O'Donovan, Parallel Prefix Sum on the GPU (Scan), slides 13 – 17.

http://users.umiacs.umd.edu/~ramani/cmsc828e_gpusci/ScanTalk.pdf

- Warp shuffles scan:

Prof. Mike Giles, Oxford University Mathematical Institute, Lecture 4: warp shuffles, and reduction / scan operations, slides 26 – 28.

<https://people.maths.ox.ac.uk/gilesm/cuda/lecs/lec4.pdf>

TÀI LIỆU THAM KHẢO



THANK YOU FOR WATCHING