Team Pink Final Presentation

20200445 박은하
20200927 장유진
20200220 오상윤
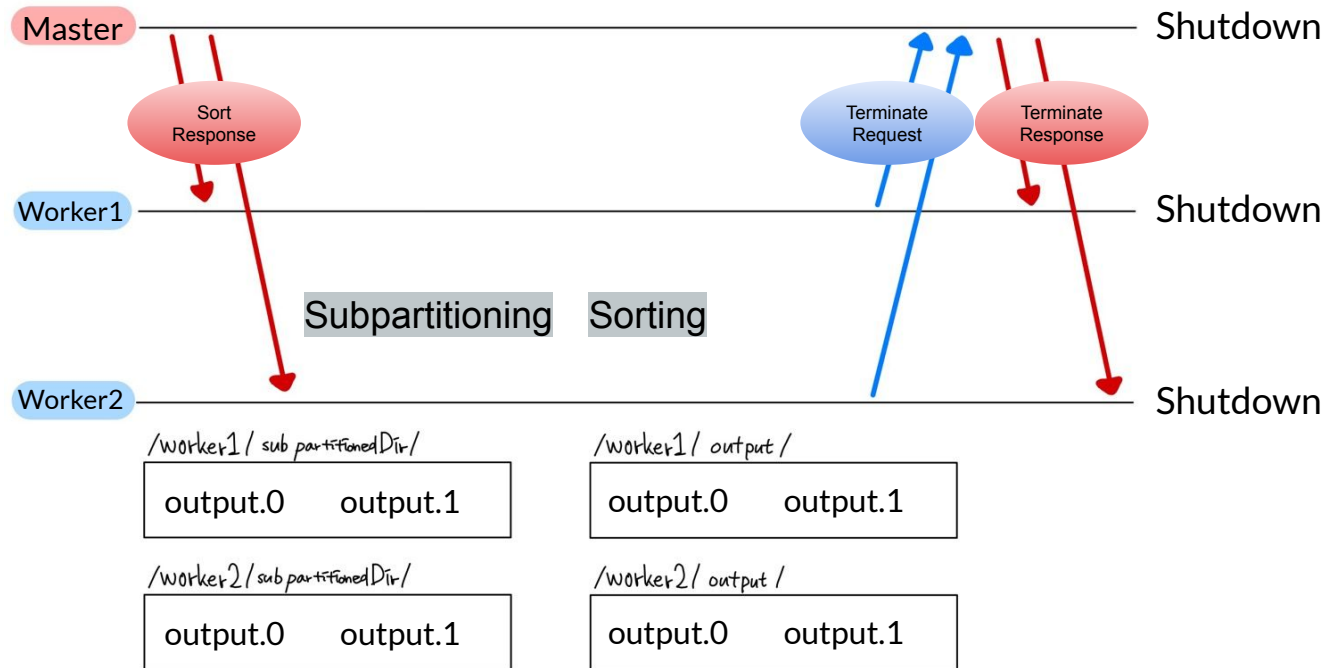
Design

Design

Design

Master ── Shutdown

Sort
Response

Terminate
Request

Terminate
Response

Worker1 ── Shutdown

Subpartitioning   Sorting

Worker2 ── Shutdown

/worker1/ subpartitionedDir/

output.0	output.1

/worker1/ output /

output.0	output.1

/worker2/ subpartitionedDir/

output.0	output.1

/worker2/ output /

output.0	output.1
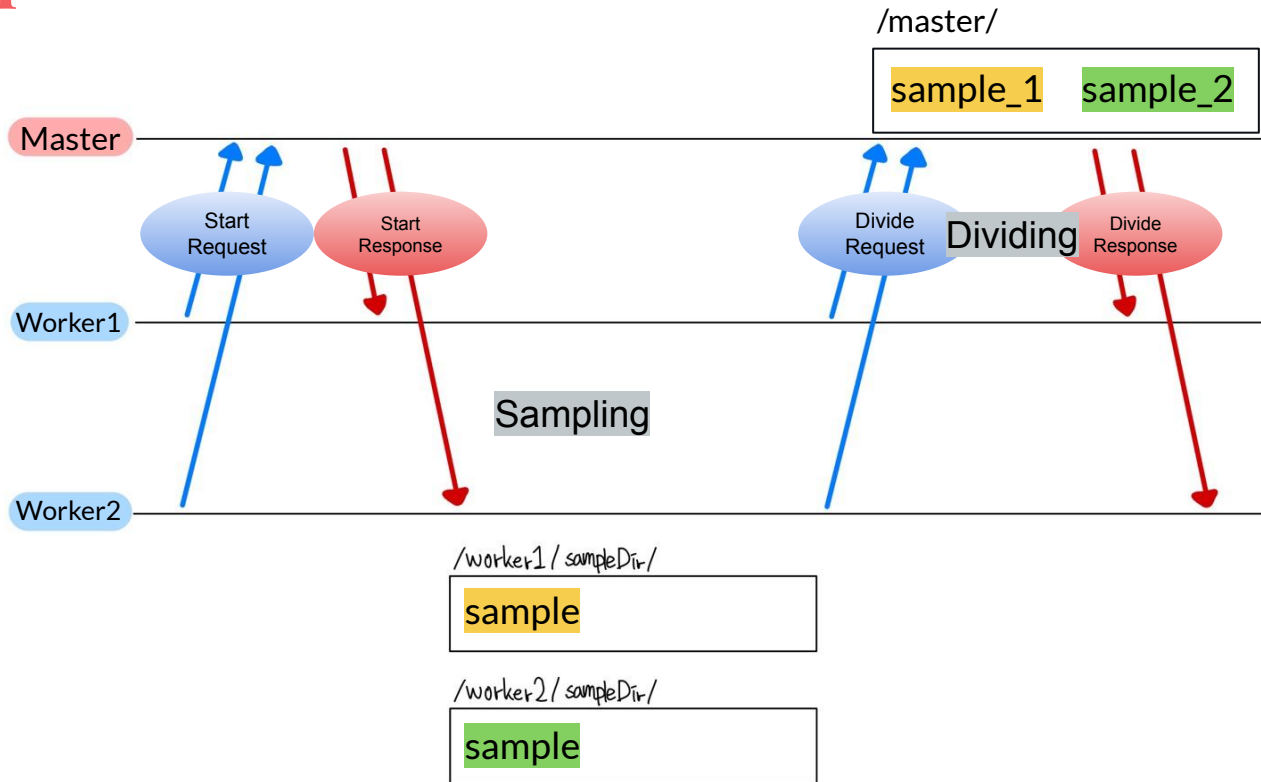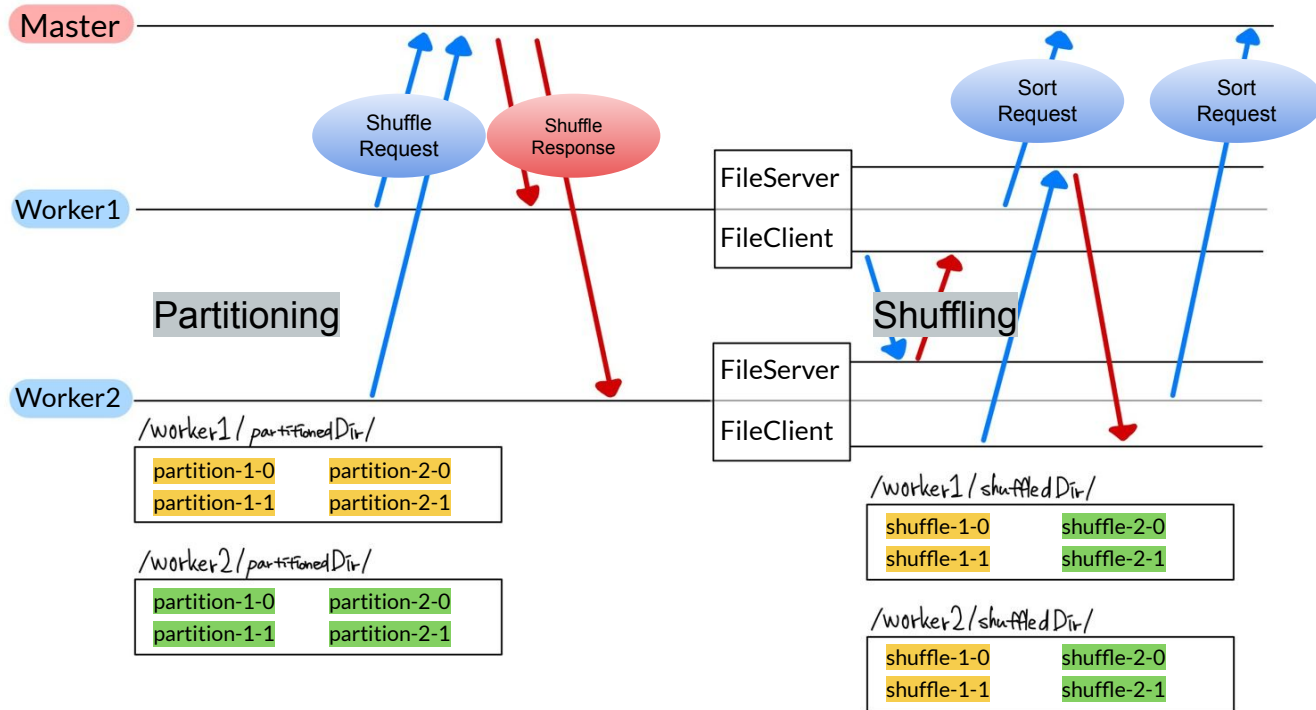
Experiment for correctness

4 workers with individually 2 inputs.
Each file is 32MB (sum of the all inputs is 256MB)
- Master : vm 7
- Workers : vm 3/4/5/6

Experiment for correctness

```
INFO: [terminate] All workers terminated. Master terminate
Total output bytes is 256000000
[success] Total time: 134 s (02:14), completed Dec 10, 2022 4:39:27 PM
```

All Loggers of master : 134 s
(실제는 더 작을 것이다)

```
INFO: shutdown(): Keep watching me shut you down!!!
[success] Total time: 44 s, completed Dec 10, 2022 4:39:15 PM
```

All Loggers of one worker : 44 s

Experiment for correctness – Start


```
[info] running network.Master 4
Dec 10, 2022 4:37:34 PM network.NetworkServer open
INFO: Server started, listening on 50051
2.2.2.107:50051
```

```
Workers IP address: 2.2.2.103 2.2.2.106 2.2.2.104 2.2.2.105
```

Experiment for correctness – Directory

Sample

```
pink@vm03:/tmp/sampleDir$ ls
sample
```

PartitionedDir

```
pink@vm03:/tmp/partitionedDir$ ls -al
total 45288
drwxrwxr-x  2 pink pink      4096 Dec 11 21:08 .
drwxrwxrwt 25 root root      4096 Dec 11 21:08 ..
-rw-rw-r--  1 pink pink  6499800 Dec 11 21:08 partition-1-0
-rw-rw-r--  1 pink pink  6499800 Dec 11 21:08 partition-1-1
-rw-rw-r--  1 pink pink 11430300 Dec 11 21:08 partition-2-0
-rw-rw-r--  1 pink pink 11430300 Dec 11 21:08 partition-2-1
-rw-rw-r--  1 pink pink  5249000 Dec 11 21:08 partition-3-0
-rw-rw-r--  1 pink pink  5249000 Dec 11 21:08 partition-3-1
```

ShuffledDir

```
pink@vm03:/tmp/shuffledDir$ ls
shuffle-1-0-5582065970176649619   shuffle-2-1-2846142947449915089   shuffle-4-0
shuffle-1-1-1648586296099586825   shuffle-3-0-197336383672294280    shuffle-4-1
shuffle-2-0-5146201507111247591   shuffle-3-1-5652756852380941284
```

subpartitionedDir

```
pink@vm03:/tmp/subpartitionedDir$ ls
output.0  output.1
```

Experiment for correctness - sampling

Sample
(Worker)

Sample
(Master)

Sampled Data
(Master)

Experiment for correctness – Dividing

Dividing
(Master)


```
INFO: [Dividing] Get all workers get ranges
Dec 11, 2022 9:08:19 PM network.NetworkServer$ConnectionImpl divide
INFO: [divide] Send ranges to Worker2: (3HUeZy.+ZU,U8Jry?Q}b>)
Dec 11, 2022 9:08:23 PM network.NetworkServer$ConnectionImpl divide
INFO: [divide] Send ranges to Worker3: (U8Jry?Q}b>,di>*+Yu-sQ)
Dec 11, 2022 9:08:24 PM network.NetworkServer$ConnectionImpl divide
INFO: [divide] Send ranges to Worker1: (        ,3HUeZy.+ZU)
Dec 11, 2022 9:08:24 PM network.NetworkServer$ConnectionImpl divide
INFO: [divide] Send ranges to Worker4: (di>*+Yu-sQ,~~~~~~~~~~)
```

Experiment for correctness - partitioning & shuffling

PartitionedDir
(receiver - input)

ShuffledDir
(sender - input)

Experiment for correctness – subpartitioned

Experiment for correctness – Terminate

```
INFO: [terminate] All Workers terminated. Master terminate
Total output bytes is 256000000
```

Master print the sum of all output files sizes from individual workers.
Input was 256MB and the result byte is 256000000.
(This printf is used for testing)

Experiment for correctness – output

vm 03 -
range from : "{Y>:#%`
range to : Vu1s$HZwQ;

vm 06 -
range from : Vu27GO]$g?
range to : k$&X\sS}`T

vm 04 -
range from : k$'a~gf{Z.
range to : pW3|:I8,%

vm 05 -
range from : pW9cf)7^Z@
range to : ~~}+tO+-g}

Experiment for correctness – Each working sorting test

```
[info] SortingTest:
[info] - sort
[info] Run completed in 894 milliseconds.
[info] Total number of tests run: 1
[info] Suites: completed 1, aborted 0
[info] Tests: succeeded 1, failed 0, canceled 0, ignored 0, pending 0
[info] All tests passed.
[success] Total time: 7 s, completed Dec 10, 2022 5:33:16 PM
```

Milestones completed

	plan	progress	final
#1. Understand problem	week 1	First meeting : too lazy -> week 2	week 2
#2. Design module and message	week 3	finish message design at week3 but we didn't consider module design -> week 5	delay by module design week 5

Milestones completed

	plan	progress	final
#3. Implement	week 5	week 6	delay by unexpected error week 7
#4. Test	week 8	week 8	week 8 -> cannot test about many workers

unexpected error:
file path name error(in shuffle), master termination error, worker termination error

Project management

- who is responsible for what?

    : 오상윤 - 구현 총괄, 디자인 총괄, **Client** 네트워크 및 기능 구현

    : 박은하 - 테스트 총괄, **File** 관리 기능 구현

    : 장유진 - 문서화 담당, **Master** 네트워크 및 기능 구현

What we learned from the project

- Do not be **optimistic** -> we have to make more progress at week 1 and 2

- **Assertion** is a powerful checkpoint

  -> we can find the cause of subpartition error immediately

- **Time** is most important resource

- **Communication overhead** was too big -> compared with pintos project

- **Do not bite off more than we can chew** -> too much care about other's code

Thank You

Team Pink