

VE572 Homework4

Yihao Liu

July 23, 2018

Ex. 1 — Processes and cgroups

1. cgroups (abbreviated from control groups) is a Linux kernel feature that limits, accounts for, and isolates the resource usage (CPU, memory, disk I/O, network, etc.) of a collection of processes.

2. differences:

similarities:

Ex. 2 — MapReduce

1. See the class `Map` in file `MapReduce.java`
2. See the class `Reduce` in file `MapReduce.java`
3. See the function `main` in file `MapReduce.java`
- 4.

data size	5e4	5e5	5e6	5e7
streaming (s)	2.474	3.415	13.464	128.59
java (s)	1.200	2.170	9.184	122.22

Ex. 3

1. add avro in `pom.xml` of maven

```
<!-- https://mvnrepository.com/artifact/org.apache.avro/avro -->
<dependency>
  <groupId>org.apache.avro</groupId>
  <artifactId>avro</artifactId>
  <version>1.8.2</version>
</dependency>
```

2. `grades.avsc`

```
{
  "namespace": "grades.avro",
  "type": "record",
  "name": "Grade",
  "fields": [
    {
      "name": "name",
      "type": "string"
    },
    {
```

```
    "name": "id",  
    "type": "string"  
  },  
  {  
    "name": "score",  
    "type": "int"  
  }  
]  
}
```

2.