第一次实验报告

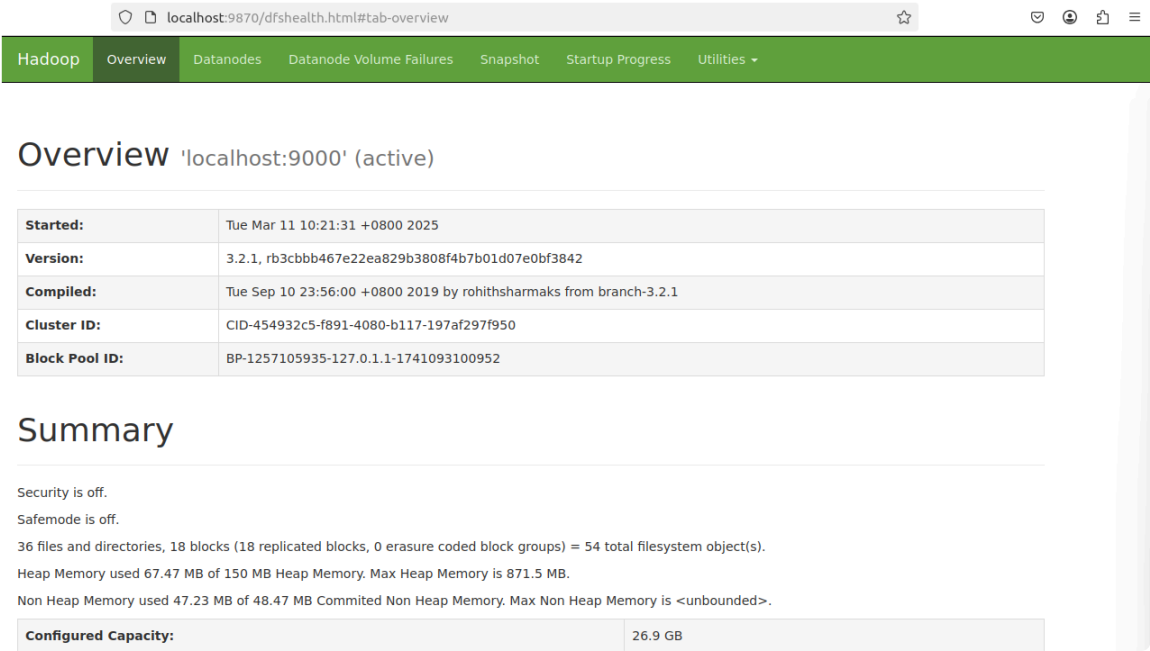| 姓名	学号	邮箱	院系
姜帅	221220115	js13156223725@163.com	计算机学院

任务一：Hadoop伪分布式环境安装与配置，HDFS基本操作

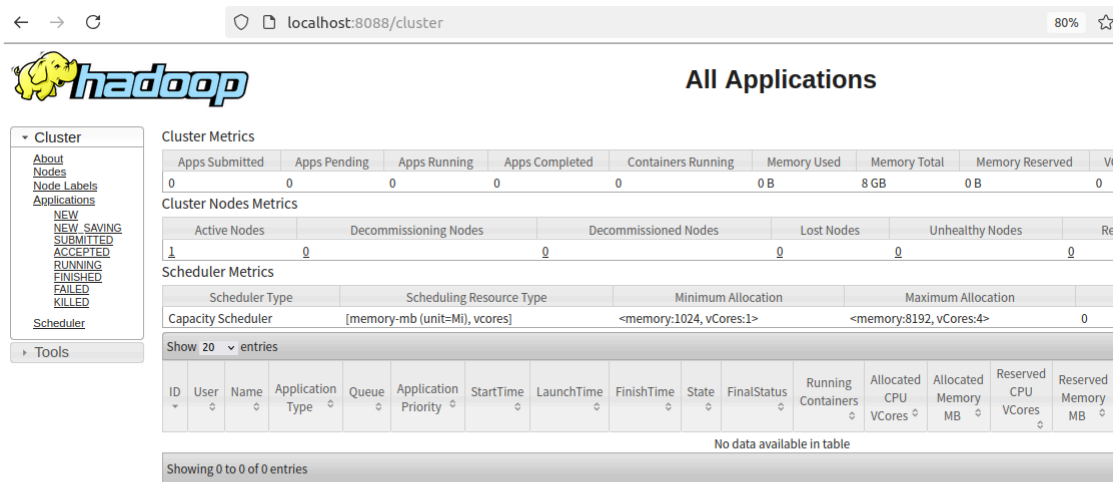参考讲义04，安装对应版本的 `java` 和 `hadoop`，并配置好 `hadoop` 的环境

- 启动hdfs和yarn，查看web界面：
  - 命令行中输入启动指令


```
js@njujs:~/hadoop_installs/hadoop-3.2.1$ sbin/start-dfs.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [njujs]
js@njujs:~/hadoop_installs/hadoop-3.2.1$ sbin/start-yarn.sh
Starting resourcemanager
Starting nodemanagers
```

  - 查看hdfs的web界面

  - 查看yarn的web界面

- hadoop安装后启动的结果如下

- hdfs基本操作

1. **向hdfs上传文件:** 通过命令 `bin/hdfs dfs-put <path_to_local_file> <path_to_hdfs_file>`,例如,按照讲义04的指引向hdfs上传两个文本文件,输入 `bin/hdfs dfs-put /root/test/*.txt test-in` ,浏览hdfs文件系统,发现文件上传成功:

☐	-rw-r--r--	js	supergroup	26 B	Mar 05 16:09	1	128 MB	file1.txt	🗑
☐	-rw-r--r--	js	supergroup	15 B	Mar 05 16:09	1	128 MB	file2.txt	🗑

2. **删除上传的文件 (附带浏览hfds目录):** 浏览目录指令可以使用 `bin/hdfs dfs -ls <path>`,也可以直接在web图形界面中浏览,同理,对文件的删除操作可以在web中执行,也可以在命令行中输入 `bin/hdfs dfs -rm <path>`,具体操作如下图所示:

任务二：执行Hadoop官方示例程序wordcount

①上传example.txt到HDFS

首先将 `example.txt` 传到本地文件系统，使用hdfs的上传命令：

```
1 bin/hdfs dfs -put ~/test/example.txt test-in
```

如图所示， `example.txt` 成功上传

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
☐	-rw-r--r--	js	supergroup	98 B	Mar 05 16:48	1	128 MB	example.txt	🗑
☐	-rw-r--r--	js	supergroup	26 B	Mar 05 16:09	1	128 MB	file1.txt	🗑
☐	-rw-r--r--	js	supergroup	15 B	Mar 05 16:09	1	128 MB	file2.txt	🗑

Showing 1 to 3 of 3 entries

Previous 1 Next

②使用官方jar包执行wordcount任务

执行命令：

```
1 bin/hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.1.jar
  wordcount test-in/example.txt temp_output
```

③观察MapReduce执行过程

观察控制台上的输出可见，该次执行过程涉及1个map任务和1个reduce任务，此外还可以看到作业从文件系统读取的数据（HDFS 212字节，FILE 111字节），输出（HDFS 73字节，FILE 452809字节），Map、Reduce阶段的记录数，总计耗时（3.7秒），shuffle错误等等。

同时，观察yarn web页面的执行情况如下：

Show 20 ⌄ entries

ID	User	Name	Application Type	Queue	Application Priority	StartTime	LaunchTime	FinishTime	State	FinalStatus	Running Containers	Alloca CPU VCore
application_1741164451699_0001	js	word count	MAPREDUCE	default	0	Wed Mar 5 16:50:11 +0800 2025	Wed Mar 5 16:50:13 +0800 2025	Wed Mar 5 16:50:32 +0800 2025	FINISHED	SUCCEEDED	N/A	N/A

Showing 1 to 1 of 1 entries

④ **执行结果**

```
js@njujs:~/hadoop_installs/hadoop-3.2.1$ bin/hdfs dfs -cat temp_output/part-r-00
000
2025-03-05 16:51:11,551 INFO sasl.SaslDataTransferClient: SASL encryption trust
check: localHostTrusted = false, remoteHostTrusted = false
HBase   2
Hadoop  1
Hive    2
Spark   2
Zookeeper       1
hadoop  3
spark   1
zookeeper       2
js@njujs:~/hadoop_installs/hadoop-3.2.1$
```

任务三：Git安装与基本操作

- 将仓库fork后创建新分支 `new_branch` ，clone到本地。

- 根据任务要求，为 `pkg/ddc/alluxio/ufs_test.go` 下的 `TestUpdateMountTime` 函数添加注释，本地修改如下：

```
910    // TestUpdateMountTime verifies if AlluxioEngine's updateMountTime method correctly updates runtime
911    // It creates a runtime with outdated MountTime, executes the update method, then checks if MountTi
912    //
913    // param: t *testing.T - The testing context used for running the test and reporting failures.
914    //
915    // returns: None (This is a test function and does not return any value.)
916    func TestUpdateMountTime(t *testing.T) {
```

- 完成任务后提交（commit），并推送（push）到自己仓库的相应分支，如图所示：

```
js@njujs:~/proj1/fluid_njujs$ git checkout new_branch
M       pkg/ddc/alluxio/ufs_test.go
Already on 'new_branch'
Your branch is up to date with 'origin/new_branch'.
js@njujs:~/proj1/fluid_njujs$ git add .
js@njujs:~/proj1/fluid_njujs$ git commit -s -m "Add Notation to TestUpdateMountT
ime in pkg/ddc/alluxio/ufs_test.go"
[new_branch 76088d1e] Add Notation to TestUpdateMountTime in pkg/ddc/alluxio/ufs
_test.go
 1 file changed, 6 insertions(+)
js@njujs:~/proj1/fluid_njujs$ git push origin new_branch
Enumerating objects: 11, done.
Counting objects: 100% (11/11), done.
Delta compression using up to 2 threads
Compressing objects: 100% (6/6), done.
Writing objects: 100% (6/6), 815 bytes | 407.00 KiB/s, done.
Total 6 (delta 5), reused 0 (delta 0)
remote: Resolving deltas: 100% (5/5), completed with 5 local objects.
To github.com:Downtodusk/fluid_njujs.git
   11781753..76088d1e  new_branch -> new_branch
js@njujs:~/proj1/fluid_njujs$
```

- 按照要求，创建一个issue到 `fluid-cloudnative/fluid` ，如下图所示：**issue编号为#4543**

Add comments to TestUpdateMountTime in pkg/ddc/alluxio/ufs_test.go #4543

⊙ Open ⇅ #4544

Downtodusk opened on Mar 6, 2025 ···

Which function do you want to add comments to?

TestUpdateMountTime in pkg/ddc/alluxio/ufs_test.go

What comments do you want to add?

```
// TestUpdateMountTime verifies if AlluxioEngine's updateMountTime method correctly updates runtime's MountTime statu
// It creates a runtime with outdated MountTime, executes the update method, then checks if MountTime gets refreshed timesta
//
// param: t *testing.T - The testing context used for running the test and reporting failures.
//
// returns: None (This is a test function and does not return any value.)
```

☺

- 按照要求，创建一个合并请求 (PR)，并等待合并，如下图所示：**PR编号为#4544**

This PR is to add comments to TestUpdateMountTime in pkg/ddc/alluxio/ufs_test.go #4544

⇅ Open **Downtodusk** wants to merge 1 commit into `fluid-cloudnative:master` from `Downtodusk:new_branch` ⎘

💬 Conversation 3 | ⟜ Commits 1 | ⊟ Checks 3 | ⬆ Files changed 1

Downtodusk commented 2 days ago ···

Ⅰ. Describe what this PR does

This PR is to add comments to TestUpdateMountTime in pkg/ddc/alluxio/ufs_test.go

Ⅱ. Does this pull request fix one issue?

fixes #4543

Ⅲ. Special notes for reviews

☺

⟜ 🐴 Add Notation to TestUpdateMountTime in pkg/ddc/alluxio/ufs_test.go ···   ● 76088d1

○ 合并成功，如下图所示：

This PR is to add comments to TestUpdateMountTime in pkg/ddc/alluxio/ufs_test.go #4544

⑂ Merged **fluid-e2e-bot** merged 1 commit into `fluid-cloudnative:master` from `Downtodusk:new_branch` ⎘ yesterday

💬 Conversation 4 | ⟜ Commits 1 | ⊟ Checks 13 | ⬆ Files changed 1

Downtodusk commented 5 days ago · edited ▾ (Contributor) ···

Ⅰ. Describe what this PR does

This PR is to add comments to TestUpdateMountTime in pkg/ddc/alluxio/ufs_test.go

Ⅱ. Does this pull request fix one issue?

fixes #4543

Ⅲ. Special notes for reviews

☺

⟜ 🐴 Add Notation to TestUpdateMountTime in pkg/ddc/alluxio/ufs_test.go ···   ✓ 76088d1

🏷 🐴 **fluid-e2e-bot** (bot) added the `needs-ok-to-test` label 5 days ago