

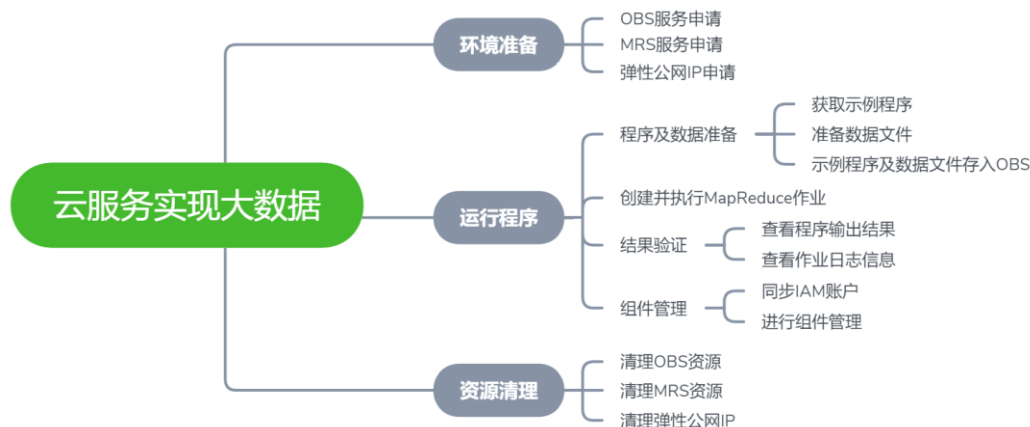
# 南京信息工程大学 实验（实习）报告

实验（实习）名称 华为实验二：云服务实现大数据 （实习）日期 5-8 得分           
指导教师 孙乐 专业 软件工程（中外合作办学） 年级 2020 级 班次 2  
姓名 颜晓雨 学号 202083020070

## 实验目的：

- 理解云服务及其实践。
- 了解云服务的一般使用流程。
- 掌握基于云服务搭建大数据环境。
- 掌握基于 MRS 的作业管理及组件管理。
- 掌握云资源的清理。

## 实验内容：



## 实验步骤：

### 1. 实验环境准备

表2-1 实验设备配套关系

名称	版本	规格	备注
OBS	无	多AZ   标准存储   私有	
MRS	1.9.2	分析集群	
EIP	无	全动态BGP   按流量计费	

2.结果

桶名称	特色功能	存储类别	区域	数据冗余...	存储用量...	桶策略	对象数量...	创建时间
obs-7bcd		标准存储	华北-北京四	多AZ存储	0 byte	私有桶	0	2023/04/22 ...

弹性公网IP	监控	状态	安全	类型	带宽	带宽详情	已绑定实例	操作
120.46.202.16 97c7cd4b-dbafe...		未绑定		全动态BGP	bandwidth-7...	按流量计费 50 Mbit/s	未绑定实例...	<a href="#">绑定</a> <a href="#">解绑</a> <a href="#">更多</a>

对象名称	所属桶	大小	状态	操作
1.txt	obs-7bcd	28 bytes	成功	<a href="#">清除记录</a>
2.txt	obs-7bcd	33 bytes	成功	<a href="#">清除记录</a>
hadoop-mapreduce-...	obs-7bcd	294.86 KB	成功	<a href="#">清除记录</a>

<input type="checkbox"/>	_SUC...	标准存储	0 byte	未加密	--	2023/04/22 20...	<a href="#">下载</a> <a href="#">分享</a> <a href="#">更多</a>
<input type="checkbox"/>	part-r...	标准存储	55 bytes	未加密	--	2023/04/22 20...	<a href="#">下载</a> <a href="#">分享</a> <a href="#">更多</a>

<input type="checkbox"/>	mr-wordcount bd2a9cbb-0c2f-4e53-ba9d...	hid_f1h...	MapReduce	运行中	未定	default	2023/04/22 20:52:35 ...	<a href="#">查看日志</a> <a href="#">查看详情</a> <a href="#">更多</a>
--------------------------	--	------------	-----------	-----	----	---------	-------------------------	--

MapReduce服务-控制台

https://console.huaweicloud.com/mrs/?agencyId=0c7a4cb9cf1f45cd8cd9394522f377b8&region=cn-north-4&lo...

华为云 控制台 北京四

返回集群

作业类型  
MapReduce

作业名称  
mr-wordcount

container-localizer-syslog stderr **stdout** syslog

193 Reduce input groups=0

194 Reduce shuffle bytes=137

195 Reduce input records=10

196 Reduce output records=8

197 Spilled Records=0

198 Shuffled Maps =2

199 Failed Shuffles=0

200 Merged Map outputs=2

201 GC time elapsed (ms)=173

202 CPU time spent (ms)=3850

203 Physical memory (bytes) snapshot=1254912000

204 Virtual memory (bytes) snapshot=12294066176

205 Total committed heap usage (bytes)=994204160

206 Shuffle Errors

207 BAD\_ID=0

208 CONNECTION=0

209 IO\_ERROR=0

210 WRONG\_LENGTH=0

211 WRONG\_MAP=0

212 WRONG\_REDUCE=0

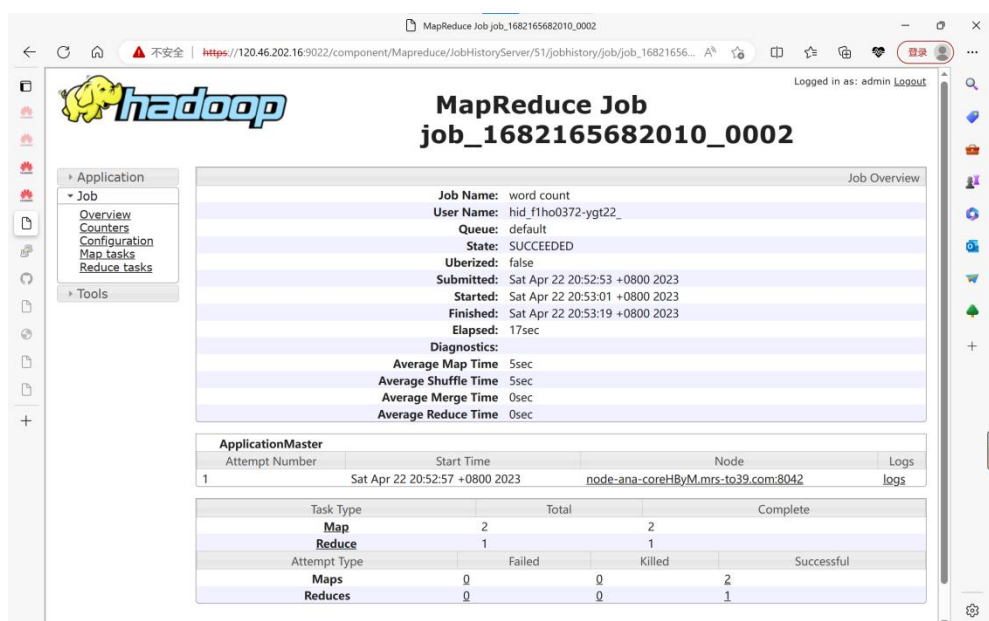
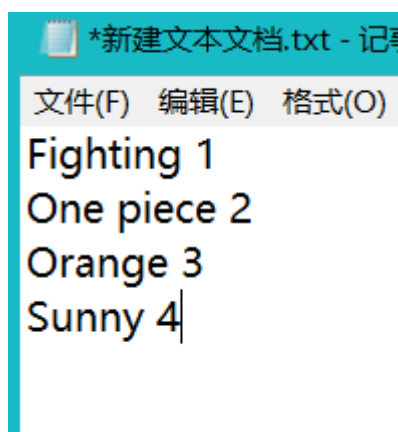
213 File Input Format Counters

214 Bytes Read=61

215 File Output Format Counters

216 Bytes Written=55

217



### 3. 实验总结

本次实验基于已经搭建好的大数据环境，运行了一个 MapReduce 程序，读取 OBS 中的数据进行计算，结果存储在 OBS 中。实验选用了 Hadoop 自带的 WordCount 示例程序，数据集文件及程序输出结果文件存放到 OBS 中，使用 MRS 构建大数据环境、作业运行示例程序及组件管理。

实验流程如下：首先，将待处理数据集存储到 OBS 中，然后将示例程序也存储到 OBS 中。接着，创建 MapReduce 类型作业执行示例程序，等待作业执行完成后，下载程序输出文件并查看结果。最后，使用组件管理查看作业执行日志。

本次实验的目的是通过实际操作，了解如何在 MRS 上运行 MapReduce 程序，掌握如何使用 OBS 存储数据集和程序输出结果文件，以及如何使用组件管理查看作业执行日志。通过本次实验，我深刻认识到了大数据技术的重要性，也更加熟悉了 MRS 平台的使用方法。

实验时间：2 机时