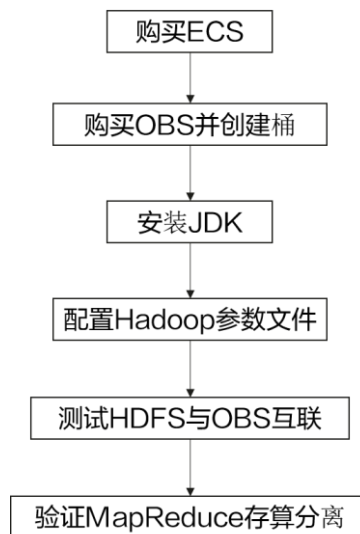


# 南京信息工程大学 实验（实习）报告

实验（实习）名称 华为实验一：云主机实现大数据 （实习）日期 4-18 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：

学习大数据领域中常用的 NoSQL 数据库 Hbase 的使用本实验通过购买华为云 ECS 和在华为云上选取对象存储服务 OBS 服务，为后续实验提供搭建环境基础。本实验的基本步骤包含：购买并配置 ECS；购买 OBS 并获取 AK、SK 信息；搭建 Hadoop 集群；搭建 MapReduce 集群并验证存算分离。

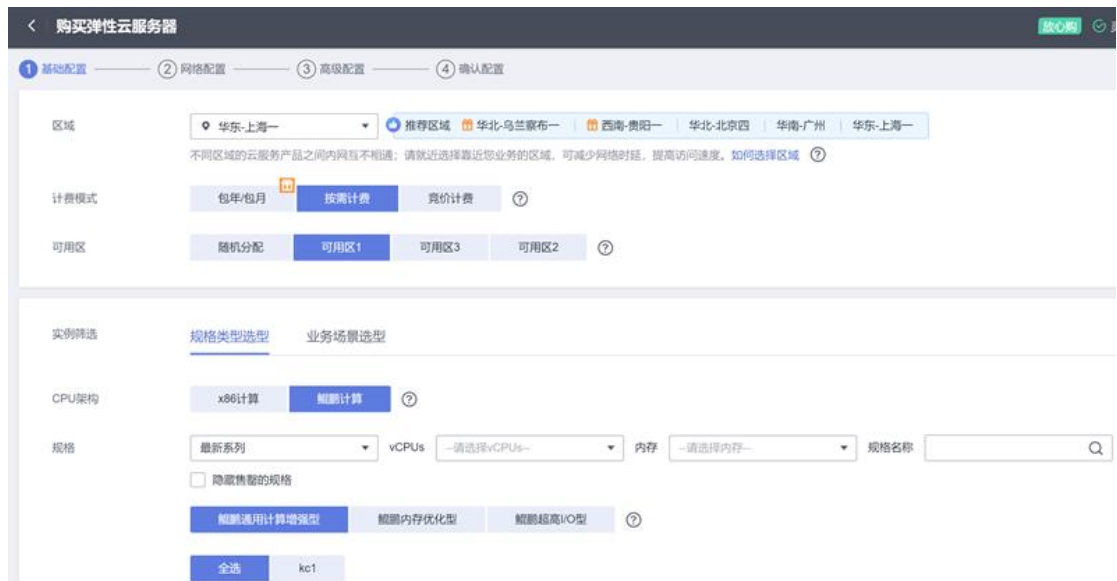
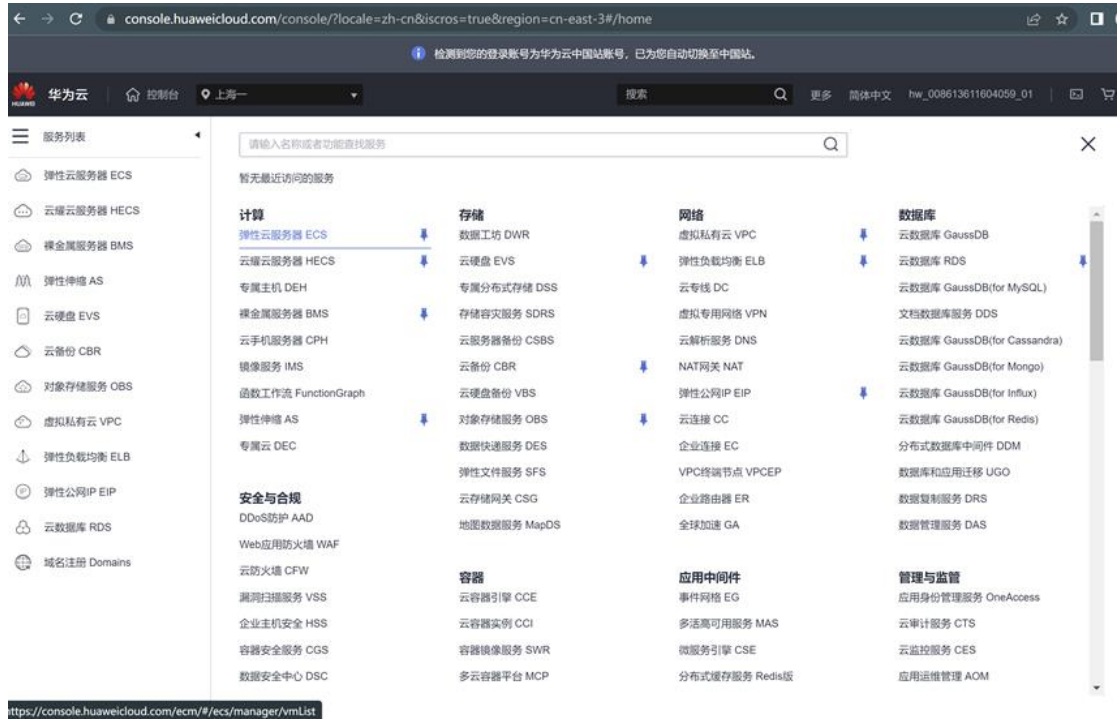


## 实验内容：



### 实验步骤:

1.准备实验环境



CPU架构

x86计算

鲲鹏计算

规格

最新系列

vCPUs

--请选择vCPUs--

内存

--请选择内存--

规格名称

☐ 隐藏售罄的规格

鲲鹏通用计算增强型

鲲鹏内存优化型

鲲鹏超高IO型

全选

kc1

展开帮助

实例类型	规格名称	vCPUs	内存	CPU	基准 / 最大带宽	内网收发包	IPv6	规格参考价
<input type="radio"/> 鲲鹏通用计...	kc1.small.1	1vCPUs	1GiB	Huawei Kun...	0.5 / 2 Gbit/s	20万PPS	否	¥0.12/小时
<input checked="" type="radio"/> 鲲鹏通用计...	kc1.large.2	2vCPUs	4GiB	Huawei Kun...	0.8 / 3 Gbit/s	30万PPS	否	¥0.30/小时
<input type="radio"/> 鲲鹏通用计...	kc1.large.4	2vCPUs	8GiB	Huawei Kun...	0.8 / 3 Gbit/s	30万PPS	否	¥0.41/小时
<input type="radio"/> 鲲鹏通用计...	kc1.xlarge.2	4vCPUs	8GiB	Huawei Kun...	1.5 / 5 Gbit/s	50万PPS	否	¥0.60/小时
<input type="radio"/> 鲲鹏通用计...	kc1.xlarge.4	4vCPUs	16GiB	Huawei Kun...	1.5 / 5 Gbit/s	50万PPS	否	¥0.81/小时
<input type="radio"/> 鲲鹏通用计...	kc1.2xlarge.2	8vCPUs	16GiB	Huawei Kun...	3 / 7 Gbit/s	80万PPS	否	¥1.20/小时
<input type="radio"/> 鲲鹏通用计...	kc1.2xlarge.4	8vCPUs	32GiB	Huawei Kun...	3 / 7 Gbit/s	80万PPS	否	¥1.63/小时

购买弹性云服务器

放心购

灵活调整

1 基础配置

2 网络配置

3 高级配置

4 确认配置

网络

vpc-default(192.168.0.0/16)

C

subnet-default(192.168.0.0/24)

C

自动分配IP地址

可用私有IP数量250个

如需创建新的虚拟私有云，您可前往控制台创建。

扩展网卡

增加一块网卡

您还可以增加 1 块网卡

安全组

default(45bb2434-c67a-4d1d-859f-4cac3c90dece)

C

新建安全组

安全组类似防火墙功能，是一个逻辑上的分组，用于设置网络访问控制。  
请确保所选安全组已放通22端口（Linux SSH登录），3389端口（Windows远程登录）和ICMP协议（Ping）。[配置安全组规则](#)

隐藏安全组规则

入方向规则

出方向规则

安全组名称	优先级	策略	协议端口	类型	源地址	描述
default	100	允许	全部	IPv6	default	--
	100	允许	全部	IPv4	default	--

华为云

控制台

搜索

更多

简体中文

hw\_008613611604059\_01

购买弹性云服务器

放心购

灵活调整

1 基础配置

2 网络配置

3 高级配置

4 确认配置

使用云服务需要先进行实名认证，手机扫码认证仅需1分钟，立即认证

配置

基础配置

计费模式

按量计费

区域

北京西

可用区

可用区1

规格

鲲鹏通用计算增强型 | kc1.large.2 ...

镜像

CentOS 7.6 64bit with ARM

主机安全

免费开启主机安全基础防护

系统盘

通用型SSD, 40GiB

网络配置

虚拟私有云

弹性公网IP

全动态BGP | 计费方式: 按流量计费 带宽: 5 Mbit/s

全动态BGP | 计费方式: 按流量计费 ...

default

主网卡

subnet-default(192.168.0.0/24)

高级配置

云服务器名称

ecs-d986

登录凭证

密码

云服务器组

--

生成Open API最佳实践脚本

配置费用 ¥1.3552/小时 + 弹性公网IP流量费用 ¥0.80/GB

上一步

立即购买



## 任务提交成功!

您的弹性云服务器已经开始创建。

[返回购买](#)

[返回云服务器列表](#)

华为云

控制台

搜索

更多

简体中文

hw\_008613611604059\_01

对象存储服务

总览 **NEW**

桶列表

并行文件系统

总用量 **NEW**

资源包管理 **NEW**

我的套餐

数据快速服务

云存储网关

媒体转码服务

CDN

对象存储迁移服务

桶列表 **桶列表** 开源声明

评价

流程引导

使用指南

任务中心

**创建桶**

购买资源包

诚邀您参加对象存储服务使用体验调研。您宝贵的意见和建议是我们持续提升产品体验的源动力，感谢您的参与!

1

**创建存储桶**

桶是OBS中存储对象的数据单元。用于存储您的数据。您可以通过控制台、API、SDK等方式进行操作。

**创建桶**

2

**配置存储桶**

通过配置桶信息，让存储桶符合您的使用要求。包括：访问权限控制、生命周期管理、防盗链、静态网站托管、日志等。

>>了解更多配置

3

**管理对象**

对象是OBS中数据的基本单位。OBS提供上传、下载、列举、搜索、分享、断点续传等基础功能，满足您各个场景的对象管理要求。

>>了解更多

4

**数据监控**

您可以通过OBS控制台或者云监控服务（Cloud Eye）监控桶的流量统计和请求次数等指标。方便您及时了解目前资源的使用状况，并合理规划使用计划。

>>更多监控项

5

**购买资源包**

OBS为您提供多种规格的资源包，让您享受更优惠的计费。

购买资源包

您的帐号及帐号下的所有IAM用户可创建的桶和并行文件系统的上限为100个。建议您使用桶内文件夹管理项目区分，合理进行桶规划和使用。[了解更多](#)

您还可以创建100个桶。

通过指定属性的关键字搜索

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	操作
-----	------	------	----	---------	---------	-----	---------	------	----

华为云

控制台

搜索

更多

简体中文

hw\_008613611604059\_01

默认存储类别

**标准存储**

适合高性能，高可靠。高可用，频繁访问场景

多AZ存储

单AZ存储

图片处理

**低频访问存储**

适合高可靠，低成本，较少访问场景

多AZ存储

单AZ存储

图片处理

归档存储

适合长期存储，平均一年访问一次

单AZ存储

费用参考

创建桶时选择的存储类别会作为上传对象的默认存储类别。[了解存储类别差异](#)

桶策略

私有

公共读

公共读写

复制桶策略

桶的拥有者拥有完全控制权限。其他用户在未经授权的情况下均无访问权限。

归档数据直读

开启

关闭

关闭归档直读。归档存储类别的数据要先恢复才能访问。归档存储数据恢复和访问会收取相应的费用。[价格详情](#)

服务端加密

不开启加密

SSE-KMS

建议开启加密，核心数据更安全。如果您使用KMS加密模式，超过免费配额会收取相应费用。[价格详情](#)

开启默认加密后，上传到当前桶的对象会默认加密。您也可以在此桶创建完成之后在桶概览页面调整默认加密配置。

标签

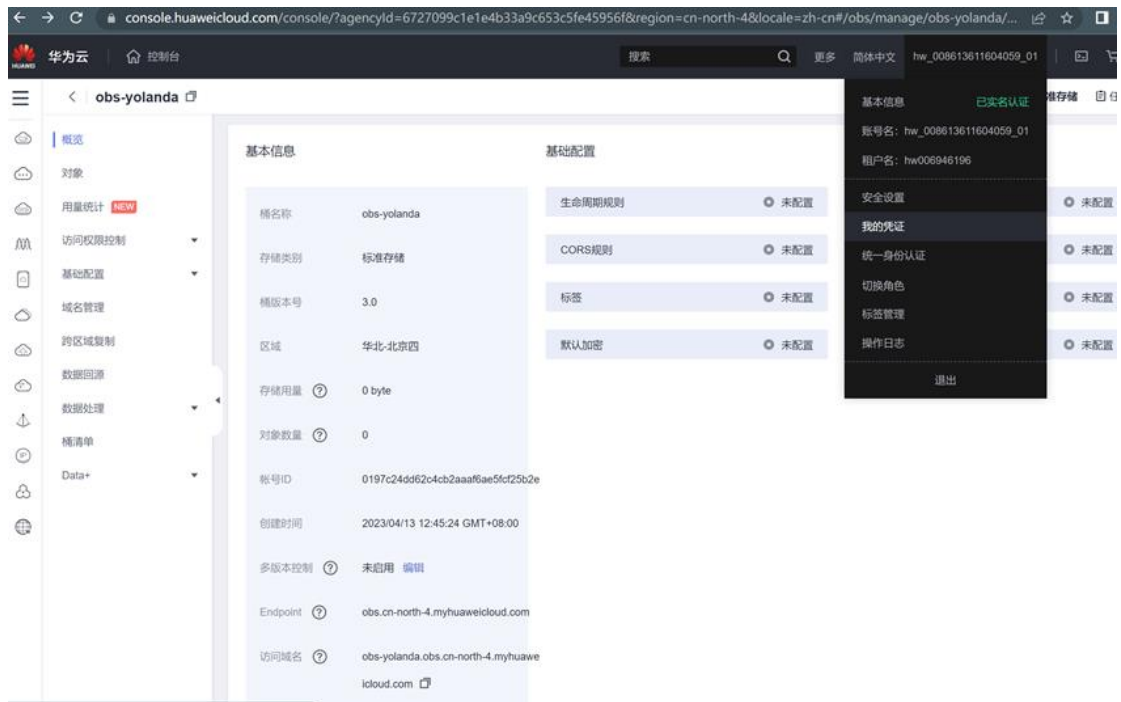
如果您需要使用同一标签标识多种云资源，即所有服务均可在标签输入框下拉选择同一标签，建议在TMS中创建预定义标签。[查看预定义标签](#)

标签键

标签值

您还可以添加10个标签。

存储包超值购



## 2. 搭建 Hadoop 集群

32-bit x86: [putty-0.78.tar.gz](#) [\(signature\)](#)

**Unix source archive**

-tar.gz: [putty-0.78.tar.gz](#) [\(signature\)](#)

---

**Alternative binary files**

The installer packages above will provide versions of all of these (except PuTTYtel and pterm), but you can download standalone binaries one by one if you prefer.

(Not sure whether you want the 32-bit or the 64-bit version? Read the [FAQ entry](#).)

**putty.exe (the SSH and Telnet client itself)**

64-bit x86: [putty.exe](#) [\(signature\)](#)

64-bit Arm: [putty.exe](#) [\(signature\)](#)

32-bit x86: [putty.exe](#) [\(signature\)](#)

**pscp.exe (an SCP client, i.e. command-line secure file copy)**

64-bit x86: [pscp.exe](#) [\(signature\)](#)

64-bit Arm: [pscp.exe](#) [\(signature\)](#)

32-bit x86: [pscp.exe](#) [\(signature\)](#)


**psftp.exe (an SFTP client, i.e. general file transfer sessions much like FTP)**

64-bit x86: [psftp.exe](#) [\(signature\)](#)

64-bit Arm: [psftp.exe](#) [\(signature\)](#)

32-bit x86: [psftp.exe](#) [\(signature\)](#)

**puttytel.exe (a Telnet-only client)**

 **PuTTY Configuration** ? X

**Category:**

- Session
  - Logging
- Terminal
  - Keyboard
  - Bell
  - Features
- Window
  - Appearance
  - Behaviour
  - Translation
  - Selection
  - Colours
- Connection
  - Data
  - Proxy
  - SSH
  - Serial
  - Telnet
  - Rlogin
  - SUPDUP

**Basic options for your PuTTY session**

Specify the destination you want to connect to

Host Name (or IP address)  Port

Connection type:

☒ SSH ☐ Serial ☐ Other: Telnet

Load, save or delete a stored session

Saved Sessions

Load
Save
Delete

安裝 WinSCP



root@ecs-d986-0001:~

```
login as: root
root@123.249.91.43's password:
Last failed login: Tue Apr 18 09:53:25 CST 2023 from 88.214.25.16 on ssh:notty
There were 189 failed login attempts since the last successful login.
```

Welcome to Huawei Cloud Service

```
[root@ecs-d986-0001 ~]# wget https://bigdata-tools-hw.obs.cn-north-1.myhuaweicloud.com/hadoop-2.8.3.tar.gz
--2023-04-18 10:42:59-- https://bigdata-tools-hw.obs.cn-north-1.myhuaweicloud.com/hadoop-2.8.3.tar.gz
Resolving bigdata-tools-hw.obs.cn-north-1.myhuaweicloud.com (bigdata-tools-hw.obs.cn-north-1.myhuaweicloud.com)... 114.115.192.27, 114.115.192.163, 114.115.192.98
Connecting to bigdata-tools-hw.obs.cn-north-1.myhuaweicloud.com (bigdata-tools-hw.obs.cn-north-1.myhuaweicloud.com)|114.115.192.27|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 244469481 (233M) [application/gzip]
Saving to: 'hadoop-2.8.3.tar.gz'

94% [=====>] 230,555,648 1.24MB/s eta 4s
```

root@ecs-d986-0001:~

```
v.obs.cn-north-1.myhuaweicloud.com)|114.115.192.27|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 244469481 (233M) [application/gzip]
Saving to: 'hadoop-2.8.3.tar.gz'
```

```
100%[=====>] 244,469,481 1.22MB/s in 72s
```

```
2023-04-18 10:44:11 (3.22 MB/s) - 'hadoop-2.8.3.tar.gz' saved [244469481/244469481]
```

```
[root@ecs-d986-0001 ~]# wget https://bigdata-tools-hw.obs.cn-north-1.myhuaweicloud.com/OpenJDK8U-jdk_aarch64_linux_hotspot_8u191b12.tar.gz
--2023-04-18 10:49:17-- https://bigdata-tools-hw.obs.cn-north-1.myhuaweicloud.com/OpenJDK8U-jdk_aarch64_linux_hotspot_8u191b12.tar.gz
Resolving bigdata-tools-hw.obs.cn-north-1.myhuaweicloud.com (bigdata-tools-hw.obs.cn-north-1.myhuaweicloud.com)... 114.115.192.98, 114.115.192.163, 114.115.192.27
Connecting to bigdata-tools-hw.obs.cn-north-1.myhuaweicloud.com (bigdata-tools-hw.obs.cn-north-1.myhuaweicloud.com)|114.115.192.98|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 75382012 (72M) [application/gzip]
Saving to: 'OpenJDK8U-jdk_aarch64_linux_hotspot_8u191b12.tar.gz'
```

root@ecs-d986-0003:~

```
login as: root
root@120.46.205.35's password:
Last login: Tue Apr 18 16:29:43 2023 from 112.2.253.45

Welcome to Huawei Cloud Service

[root@ecs-d986-0003 ~]# cat /etc/hosts
::1 localhost localhost.localdomain localhost6 localhost6.localdomain6

127.0.0.1 localhost localhost.localdomain localhost4 localhost4.localdomain4
127.0.0.1 localhost localhost
192.168.0.133 192.168.0.133 ecs-d986-0001
192.168.0.138 192.168.0.138 ecs-d986-0002
192.168.0.21 192.168.0.21 ecs-d986-0003
192.168.0.205 192.168.0.205 ecs-d986-0004
[root@ecs-d986-0003 ~]#
```



```
root4 - root@ecs-d986-0004:/usr/lib/jvm - Xshell 5 (Free for Home/School)
文件(F) 编辑(E) 查看(V) 工具(T) 选项卡(B) 窗口(W) 帮助(H)
ssh://root:*****@123.249.79.175:22
要添加当前会话，点击左侧的箭头按钮。
1 root... 2 root3 3 root2 4 root1
ean.java
./jdk8u191-b12/sample/jmx/jmx-scandir/src/com/sun/jmx/examples/scandir/ScanManagerMXBean.j
ava
./jdk8u191-b12/sample/jmx/jmx-scandir/src/com/sun/jmx/examples/scandir/ResultLogManager.ja
va
./jdk8u191-b12/sample/jmx/jmx-scandir/src/com/sun/jmx/examples/scandir/package.html
./jdk8u191-b12/sample/jmx/jmx-scandir/src/etc/
./jdk8u191-b12/sample/jmx/jmx-scandir/src/etc/testconfig.xml
./jdk8u191-b12/sample/jmx/jmx-scandir/src/etc/password.properties
./jdk8u191-b12/sample/jmx/jmx-scandir/src/etc/access.properties
./jdk8u191-b12/sample/jmx/jmx-scandir/src/etc/management.properties
./jdk8u191-b12/sample/jmx/jmx-scandir/build.xml
./jdk8u191-b12/THIRD_PARTY_README
./jdk8u191-b12/ASSEMBLY_EXCEPTION
./jdk8u191-b12/release
./jdk8u191-b12/lib/
./jdk8u191-b12/lib/jconsole.jar
./jdk8u191-b12/lib/orb.idl
./jdk8u191-b12/lib/tools.jar
./jdk8u191-b12/lib/aarch64/
./jdk8u191-b12/lib/aarch64/libjawt.so
./jdk8u191-b12/lib/aarch64/jli/
./jdk8u191-b12/lib/aarch64/jli/libjli.so
./jdk8u191-b12/lib/ir.idl
./jdk8u191-b12/lib/ct.sym
./jdk8u191-b12/lib/dt.jar
./jdk8u191-b12/lib/jexec
./jdk8u191-b12/include/
./jdk8u191-b12/include/jawt.h
./jdk8u191-b12/include/jni.h
./jdk8u191-b12/include/linux/
./jdk8u191-b12/include/linux/jawt_md.h
./jdk8u191-b12/include/linux/jni_md.h
./jdk8u191-b12/include/jdwpTransport.h
./jdk8u191-b12/include/jvmticmlr.h
./jdk8u191-b12/include/jvmti.h
./jdk8u191-b12/include/classfile_constants.h
./jdk8u191-b12/src.zip
[root@ecs-d986-0004 jvm]# vim /etc/profile
[root@ecs-d986-0004 jvm]# source /etc/profile
[root@ecs-d986-0004 jvm]# java - version
Error: Could not find or load main class - version
[root@ecs-d986-0004 jvm]# source /etc/profile
[root@ecs-d986-0004 jvm]# java -version
openjdk version "1.8.0_232"
OpenJDK Runtime Environment (build 1.8.0_232-b09)
OpenJDK 64-Bit Server VM (build 25.232-b09, mixed mode)
[root@ecs-d986-0004 jvm]#
```

仅将文本发送到当前选项卡  
已连接 123.249.79.175:22. SSH2 xterm 90x48 48,27 4 会话 CANU

```
root1 - root@ecs-d986-0001:~ - Xshell 5 (Free for Home/School)
文件(F) 编辑(E) 查看(V) 工具(T) 选项卡(B) 窗口(W) 帮助(H)
ssh://root:*****@123.249.91.43:22
要添加当前会话，点击左侧的箭头按钮。
1 roo...
</property>
<property>
  <name>fs.obs.threads.read.core</name>
  <value>500</value>
</property>
<property>
  <name>fs.obs.threads.read.max</name>
  <value>1000</value>
</property>
<property>
  <name>fs.obs.write.buffer.size</name>
  <value>8192</value>
</property>
<property>
  <name>fs.obs.read.buffer.size</name>
  <value>8192</value>
</property>
<property>
  <name>fs.obs.connection.maximum</name>
  <value>1000</value>
</property>
<property>
  <name>fs.defaultFS</name>
  <value>hdfs://ecs-d986-0001:8020</value>
</property>
<property>
  <name>hadoop.tmp.dir</name>
  <value>/home/modules/hadoop-2.8.3/tmp</value>
</property>
<property>
  <name>fs.obs.access.key</name>
  <value>Q7QIR0WODLI0UPRGRGLD</value>
</property>
<property>
  <name>fs.obs.secret.key</name>
  <value>Cv910LpvvEZZY7DkdNm0zBBn7rY03h8MnXVxAyob</value>
</property>
<property>
  <name>fs.obs.endpoint</name>
  <value>obs-yolanda.obs.cn-north-4.myhuaweicloud.com</value>
</property>
<property>
  <name>fs.obs.buffer.dir</name>
  <value>/home/modules/data/buf</value>
</property>
<property>
  <name>fs.obs.impl</name>
  <value>org.apache.hadoop.fs.obs.OBSFileSystem</value>
-- INSERT --
70,55 52%
```

仅将文本发送到当前选项卡  
已连接 123.249.91.43:22。 SSH2 xterm 90x49 40,55 1 会话 CANU

```
[root@ecs-d986-0001 ~]# hdfs dfs -mkdir /bigdata
23/04/21 14:40:35 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

23/04/21 14:40:35 WARN ipc.Client: Failed to connect to server: ecs-d986-0001/192.168.0.133:8020: try once and fail.
java.net.ConnectException: Connection refused
    at sun.nio.ch.SocketChannelImpl.checkConnect(Native Method)
    at sun.nio.ch.SocketChannelImpl.finishConnect(SocketChannelImpl.java:717)
    at org.apache.hadoop.net.SocketIOWithTimeout.connect(SocketIOWithTimeout.java:206)
    at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:531)
    at org.apache.hadoop.ipc.Client$Connection.setupConnection(Client.java:685)
```

## 测试

The screenshot shows the Huawei Cloud OBS console interface. The left sidebar contains navigation options: 概览 (Overview), 对象 (Object), 用量统计 (Usage Statistics), 访问权限控制 (Access Control), 基础配置 (Basic Configuration), 域名管理 (Domain Management), 跨区域复制 (Cross-Region Replication), 数据回源 (Data Ingestion), 数据处理 (Data Processing), 桶清单 (Bucket Inventory), and Data+.

The main content area is titled '对象' (Object) and includes a description: '对象是数据存储在OBS中的基本单位。在OBS中文件和文件夹需要移动到其他位置，推荐下载使用。基于安全合规要求，从浏览器直接访问文件时不...' (Objects are the basic units of data storage in OBS. In OBS, files and folders need to be moved to other locations, and it is recommended to use the download function. Based on security and compliance requirements, when accessing files directly from the browser, it is not recommended...). Below the description are buttons for '上传对象' (Upload Object), '新建文件夹' (Create Folder), and '删除' (Delete).

The '任务中心' (Task Center) section on the right shows a table of upload tasks. The table has columns for '对象名称' (Object Name), '所属桶' (Bucket), '大小' (Size), '状态' (Status), and '操作' (Action). The first row shows an object named '实验报告.doc' (Experiment Report.doc) in the 'obs-yolanda' bucket, with a size of 5.38 MB and a status of '成功' (Success). A '清除记录' (Clear Record) button is available for this task.

对象名称	所属桶	大小	状态	操作
实验报告.doc	obs-yolanda	5.38 MB	成功	清除记录

```

Xshell 5 (Build 0788)
Copyright (c) 2002-2015 NetSarang Computer, Inc. All rights reserved.

Type 'help' to learn how to use Xshell prompt.
[c:\~]$

Connecting to 123.249.91.43:22...
Connection established.
To escape to local shell, press 'Ctrl+Alt+J'.

Last login: Fri Apr 21 15:27:12 2023 from 112.2.2
53.45

Welcome to Huawei Cloud Service

[root@ecs-d986-0001 ~]# hdfs dfs -mkdir /bigdata
23/04/21 15:27:59 WARN util.NativeCodeLoader: Unable to load native-hadoop library for y
r platform... using builtin-java classes where applicable
mkdir: `/bigdata': File exists
[root@ecs-d986-0001 ~]# hdfs dfs -ls /
23/04/21 15:28:01 WARN util.NativeCodeLoader: Unable to load native-hadoop library for y
r platform... using builtin-java classes where applicable
Found 2 items
drwxr-xr-x - root supergroup          0 2023-04-21 15:26 /BigData
drwxr-xr-x - root supergroup          0 2023-04-21 15:24 /bigdata
[root@ecs-d986-0001 ~]# jps
20465 NameNode
21185 Jps
1234 WrapperSimpleApp
20659 SecondaryNameNode
[root@ecs-d986-0001 ~]# █

```

成功

### 3. 实验总结

本次实验通过华为云平台实现了基本的操作，包括服务器环境配置和 Hadoop 集群搭建。随后，我们使用 MapReduce 程序对测试环境和 Hadoop 集群功能进行了测试。最后，我们释放了云服务资源，以熟悉整套系统的搭建过程。

在实验中，我们首先在云平台上创建了一组虚拟机实例，以构建 Hadoop 集群。我们选择了适合需求的虚拟机规格，并为每个实例分配了足够的存储和带宽。一旦创建了虚拟机实例，我们需要为每个实例安装操作系统和必要的软件包。接下来，我们安装和配置了 Hadoop 软件，以便将虚拟机实例连接到 Hadoop 集群。这包括为每个虚拟机实例配置 Hadoop 环境变量，并配置 Hadoop 的核心组件，例如 HDFS。此外，我们还创建了 Hadoop 用户，并为其配置 SSH 密钥对，以便在集群中进行无密码访问。完成 Hadoop 集群的安装和配置后，我们使用 MapReduce 程序对其进行了测试。我们编写了一个简单的 txt 文件，并使用 Hadoop 命令行工具将其提交到集群上进行执行。

最后，当我们完成了 Hadoop 集群的测试和使用后，我们释放了云服务资源，以节省成本。这是通过删除虚拟机实例和清除 Hadoop 集群配置等操作来实现的。

实验时间：2 机时