



北京邮电大学

Beijing University of Posts and Telecommunications

# 华为云鲲鹏大数据基础实验体系 1

## 基于华为云构建大数据实验环境

编写负责人：

鄂海红、张忠宝

参编人员：

朱一凡、宋美娜、程祥、

罗子霄、王浩田、魏文定、刘钟允、汤子辰 等

## 1.1. 大数据实践 1：基于华为云构建大数据实验环境

### 1.1.1. 实验描述

结合云原生大数据平台的主流演进趋势,选择采用基于云服务器模式和云大数据平台模式,并从真实应用场景中分离出实验样例,帮助同学们对理论学习的知识点进行实践。

基于“教育部智能基座项目”资助,华为云为选课同学们可以提供代金券资助同学们使用华为云环境实践学习大数据实验。

在本实验报告中,会通过华为云实操大数据环境部署、数据加载、数据处理及可视化,并提供详细的实验步骤详解,让每位学生可以完整完成大数据实践练习。

本册是大数据实践 1: 基于华为云构建大数据实验环境,实验目标是完成购买云 ECS 服务器,配置完成构建大数据实验环境。

### 1.1.2. 实验目的

- 了解华为云基础操作;
- 购买华为云 ECS;
- 掌握华为云服务器操作;

### 1.1.3. 实验步骤

#### 1. 购买ECS

打开华为云地址: <https://www.huaweicloud.com/>, 点击“登录”, 输入用户名、密码,



如下图:

点击“控制台”，如下图：



点击“服务列表”，如下图：



选择“弹性云服务器 ECS”，如下图：



选择“买弹性云服务器 ECS”，如下图：



选择“按需计费”，“可用区2”，CPU 架构“鲲鹏计算”，选择“鲲鹏通用计算增强型”，

2vCPUs|4GB，如下图：

规格名称	vCPUs   内存	CPU	基准 / 最大带宽	内网收发包	规格参考价
kct.small.1	1vCPUs   1GB	Huawei Kunpeng 920 2.6GHz	0.5/2 Gbit/s	200,000	¥0.12/小时
<b>kct.large.2</b>	<b>2vCPUs   4GB</b>	Huawei Kunpeng 920 2.6GHz	0.8/3 Gbit/s	300,000	¥0.30/小时
kct.large.4	2vCPUs   8GB	Huawei Kunpeng 920 2.6GHz	0.8/3 Gbit/s	300,000	¥0.41/小时
kct.xlarge.2	4vCPUs   8GB	Huawei Kunpeng 920 2.6GHz	1.5/5 Gbit/s	500,000	¥0.60/小时
kct.xlarge.4	4vCPUs   16GB	Huawei Kunpeng 920 2.6GHz	1.5/5 Gbit/s	500,000	¥0.81/小时

配置操作系统和磁盘。选择“公共镜像”，CentOS7.6，系统盘建议配置40GB，购买数量4台，点击“网络配置”，如下图：

规格名称	vCPUs   内存	CPU	基准 / 最大带宽	内网收发包	规格参考价
kct.small.1	1vCPUs   1GB	Huawei Kunpeng 920 2.6GHz	0.5/2 Gbit/s	200,000	¥0.12/小时
<b>kct.large.2</b>	<b>2vCPUs   4GB</b>	Huawei Kunpeng 920 2.6GHz	0.8/3 Gbit/s	300,000	¥0.30/小时
kct.large.4	2vCPUs   8GB	Huawei Kunpeng 920 2.6GHz	0.8/3 Gbit/s	300,000	¥0.41/小时
kct.xlarge.2	4vCPUs   8GB	Huawei Kunpeng 920 2.6GHz	1.5/5 Gbit/s	500,000	¥0.60/小时
kct.xlarge.4	4vCPUs   16GB	Huawei Kunpeng 920 2.6GHz	1.5/5 Gbit/s	500,000	¥0.81/小时

配置网络，网络选择“vpc-default”，安全组选择“Sys-default”，“现在购买”，选择“全动态

BGP”，“按流量计费”，“5M”，点击“高级配置”，如下图：

规格名称	vCPUs   内存	CPU	基准 / 最大带宽	内网收发包	规格参考价
kct.small.1	1vCPUs   1GB	Huawei Kunpeng 920 2.6GHz	0.5/2 Gbit/s	200,000	¥0.12/小时
<b>kct.large.2</b>	<b>2vCPUs   4GB</b>	Huawei Kunpeng 920 2.6GHz	0.8/3 Gbit/s	300,000	¥0.30/小时
kct.large.4	2vCPUs   8GB	Huawei Kunpeng 920 2.6GHz	0.8/3 Gbit/s	300,000	¥0.41/小时
kct.xlarge.2	4vCPUs   8GB	Huawei Kunpeng 920 2.6GHz	1.5/5 Gbit/s	500,000	¥0.60/小时
kct.xlarge.4	4vCPUs   16GB	Huawei Kunpeng 920 2.6GHz	1.5/5 Gbit/s	500,000	¥0.81/小时

点击“申请扩大配额”，简单填写内容（例如：配额不足，无法购买对应服务，需要申请更多配额），约1小时后会分配配额。



配置密码，将服务器名称设置为“姓名首字母\_学号”，例如：lxd\_2021140808，自行设置root 登录密码，云备份选择“暂不购买”，点击“确认配置”，点中“我已经阅读并同意”，点击“立即购买”，如下图：



注：本次需购买 4 个 ECS，每个 ECS 规格相同，其中 1 个主节点、3 个从节点。

点击“我已经阅读并同意”，可以点击“返回云服务器列表”，创建过程需要等待几



分钟。

## 2. 购买OBS

进入控制台，选择“对象存储服务”，如下图：



选择“创建桶”，如下图：



将桶名称设置为“姓名首字母\_学号”，例如：lxd\_2021140808，选择“标准存储”，点击“立即创建”，如下图：



区域: 华北-北京四

不同区域的资源之间内网互不相通, 请选择靠近您业务的区域, 可以降低网络时延, 提高访问速度。桶创建成功后不支持变更区域, 请谨慎选择。

数据冗余存储策略: **多AZ存储** 单AZ存储

多AZ存储能提高您的数据可用性, 同时会采用相对较高的计费标准。 [价格详情](#)

多AZ存储属性一旦启用, 后续无法修改。

桶名称: obs-

命名规则:

- 需全局唯一, 不能与已有的任何桶名称重复。
- 删除桶或并行文件系统后, 需要等待30分钟才能创建同名桶或并行文件系统。
- 长度范围为3到63个字符, 支持小写字母、数字、中划线 (-)、英文句号 (.)。
- 禁止两个英文句号 (.) 或英文句号 (.) 和中划线 (-) 相邻, 禁止以英文句号 (.) 和中划线 (-) 开头或结尾。
- 禁止使用IP地址。
- 如果名称中包含英文句号 (.), 访问桶或对象时可能会进行安全证书校验。

存储类别: **标准存储** 低频访问存储 归档存储

适用于有大量热点文件或小文件, 且需要频繁访问 (平均一个月多次) 并快速获取数据的业务场景。

上传对象时, 对象默认与桶的存储类别相同, 也可以根据适用场景修改。 [了解更多](#)

桶策略: **私有** 公共读 公共读写

桶的拥有者拥有完全控制权, 其他用户在未经授权的情况下均无访问权限。

默认加密: **关闭** 开启

**推荐** 密钥管理全免费, 核心数据更安全。

归档数据直读: **关闭** 开启

通过归档数据直读, 您可以直接下载存储类别为归档存储的对象, 而无需提前恢复。归档数据直读会收取相应的费用。 [价格详情](#)

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。

按用量收费  
创建免费, 使用阶段按用量收费。 [了解计费详情](#)

**立即创建**

选择“创建并行文件系统”，将文件系统名称设置为“姓名首字母\_学号”，例如：

lxd\_2021140808，点击“自定义立即创建”，如下图：

对象存储服务 **对象存储服务** 开启声明 **创建并行文件系统**

对象存储

**并行文件系统**

我的套餐

数据快速服务

在控制台上您还可以创建99个并行文件系统。

请输入文件系统名称

文件系统名称	区域	存储	文件	创建时间	操作
暂无数据					

进入创建的 OBS 桶，如下图：



复制该参数，保存到本地文档，如下图：



获取 AK/SK，点击“我的凭证”，选择“访问秘钥”，如下图：



点击新增访问秘钥，根据提示进行操作。



华为云智能基座在线实验答疑

QQ群1：471519925（2021大数据技术基础）、QQ群2：125545864（2022-大数据技术基础）、QQ群3：295477031（2023-大数据技术基础）



通过手机号码，接收短信验证码，点击“确定”，如下图：



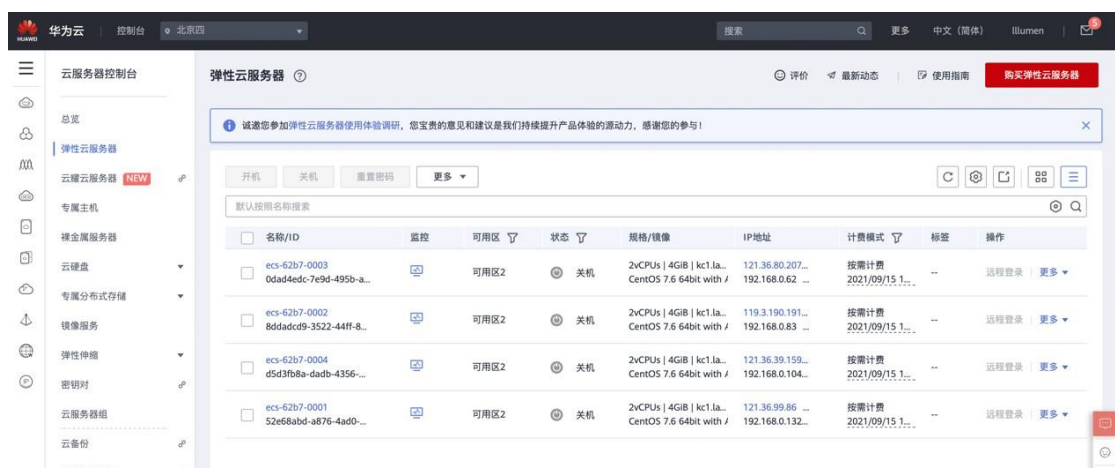
操作完成后，得到文件，打开即可得到 AK/SK。



A	B	C
User Name	Access Key Id	Secret Access Key
hw...	8MIA... K	Ttsx... JJ709eRQX1

### 3. 启动与关闭服务器

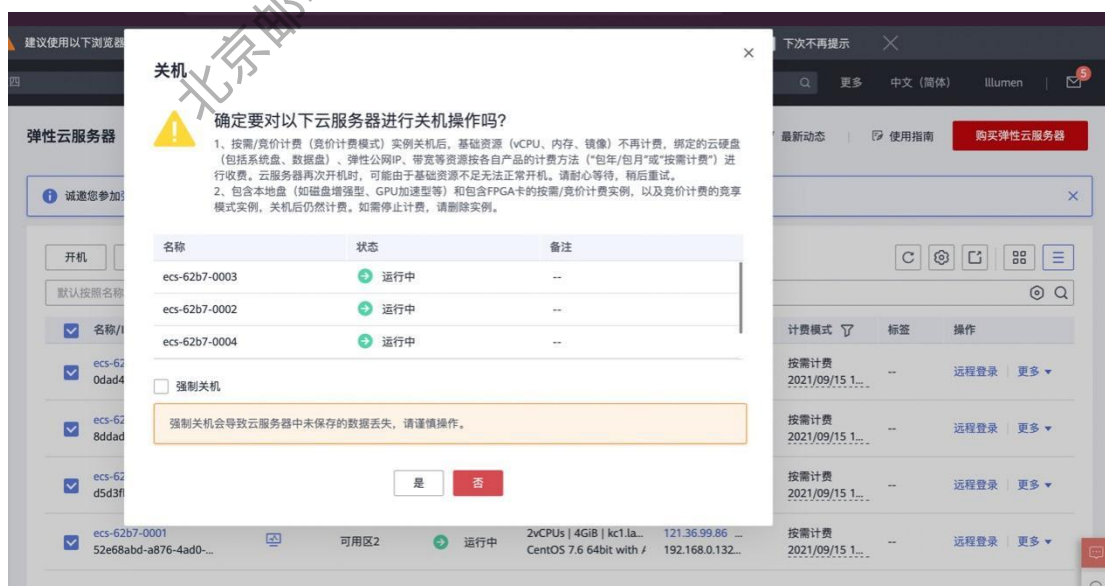
从服务列表中打开弹性云服务器。



选择四个申请到的服务器，点击开机按钮，启动服务器。下图所示，服务器为运行状态。

<input type="checkbox"/>	名称/ID	监控	可用区	状态	规格/镜像	IP地址	计费模式	标签	操作
<input type="checkbox"/>	ecs-62b7-0003 0dad4dc-7e9d-495b-a...		可用区2	运行中	2vCPUs   4GiB   kc1.la... CentOS 7.6 64bit with /	121.36.80.207... 192.168.0.62 ...	按需计费 2021/09/15 1...	--	远程登录 更多
<input type="checkbox"/>	ecs-62b7-0002 8ddadcd9-3522-44ff-8...		可用区2	运行中	2vCPUs   4GiB   kc1.la... CentOS 7.6 64bit with /	119.3.190.191... 192.168.0.83 ...	按需计费 2021/09/15 1...	--	远程登录 更多
<input type="checkbox"/>	ecs-62b7-0004 d5d3fb8a-dadb-4356-...		可用区2	运行中	2vCPUs   4GiB   kc1.la... CentOS 7.6 64bit with /	121.36.39.159... 192.168.0.104...	按需计费 2021/09/15 1...	--	远程登录 更多
<input type="checkbox"/>	ecs-62b7-0001 52e68abd-a876-4ad0-...		可用区2	运行中	2vCPUs   4GiB   kc1.la... CentOS 7.6 64bit with /	121.36.99.86 ... 192.168.0.132...	按需计费 2021/09/15 1...	--	远程登录 更多

关闭服务器，选择四个服务器后，点击关机。在不使用服务器时，建议将服务器关机，以节省费用。



### 1.1.4. 实验结果与评分标准

实验结束后应获得以下资源：

1. 四台以自己姓名学号命名的服务器；
2. 以姓名学号命名的分布式对象存储 OBS 桶；
3. 保存在本地的 endpoint（之后实验需使用）；
4. 保存在本地的 AK/SK 文件（之后实验需使用）。

北京邮电大学《大数据技术基础》课程教研室