

山东大学 计算机科学与技术 学院

大数据分析实践 课程实验报告

队名: Five in Big Data	队长: 何青青	学号: 202300130262
队员 1:	刘逸宁	学号: 202300130160
队员 2:	邱琚	学号: 202300130063
队员 3:	徐瑞良	学号: 202200120100
队员 4:	常智博	学号: 202200130005
实验题目: 电子表格实践		
实验学时: 2		实验日期: 2025/10/17-10/31
实验目标: 按照下面的教程运行 datasread 代码: https://github.com/datasread/datasread-web		
实验环境: Linux		
实验步骤与内容: 1、安装必要工具 尝试安装 Docker 时, 发现安装失败, 所以需要安装必要工具, 更新包管理器并安装 curl。		
<pre>cs18@games101vm:~\$ curl -fsSL https://get.docker.com -o get-docker.sh Command 'curl' not found, but can be installed with: sudo snap install curl # version 8.13.0, or sudo apt install curl See 'snap info curl' for additional versions. cs18@games101vm:~\$ sudo apt update [sudo] password for cs18: Hit:1 http://security.ubuntu.com/ubuntu bionic-security InRelease Hit:2 http://us.archive.ubuntu.com/ubuntu bionic InRelease Hit:3 http://us.archive.ubuntu.com/ubuntu bionic-updates InRelease Hit:4 http://us.archive.ubuntu.com/ubuntu bionic-backports InRelease Reading package lists... Done Building dependency tree Reading state information... Done 218 packages can be upgraded. Run 'apt list --upgradable' to see them. cs18@games101vm:~\$ sudo apt install curl -y Reading package lists... Done Building dependency tree Reading state information... Done The following NEW packages will be installed: curl 0 upgraded, 1 newly installed, 0 to remove and 218 not upgraded. Need to get 159 kB of archives. After this operation, 398 kB of additional disk space will be used. Get:1 http://us.archive.ubuntu.com/ubuntu bionic-updates/main amd64 curl amd64 7.58.0-2ubuntu3.24 [159 kB] Fetched 159 kB in 5s (33.7 kB/s) Selecting previously unselected package curl. (Reading database ... 179837 files and directories currently installed.) Preparing to unpack .../curl_7.58.0-2ubuntu3.24_amd64.deb ... Unpacking curl (7.58.0-2ubuntu3.24) ... Setting up curl (7.58.0-2ubuntu3.24) ... Processing triggers for man-db (2.8.3-2ubuntu0.1) ... cs18@games101vm:~\$ curl --version curl 7.58.0 (x86_64-pc-linux-gnu) libcurl/7.58.0 OpenSSL/1.1.1 zlib/1.2.11 libidn2/2.0.4 libpsl/0.19.1 (+libidn2/2.0.4) nghttp2/1.30.0 librtmp/2.3 Release-Date: 2018-01-24 Protocols: dict file ftp ftps gopher http https imap imaps ldap ldaps pop3 pop3s rtsp smb smbs smtp smtps telnet tftp Features: AsynchDNS IDN IPv6 Largefile GSS-API Kerberos SPNEGO NTLM NTLM_WB SSL libz TLS-SRP HTTP2 UnixSockets HTTP S-proxy PSL</pre>		

2、安装 Docker 和 docker-compose

在尝试使用官方脚本安装 Docker 时，发现安装失败，警告显示我的 Ubuntu 系统版本（bionic）已经停止支持，但脚本仍在继续安装，不过安装过程中出现了问题，docker-model-plugin 包找不到。

```
cs18@games101vm:~$ sudo sh get-docker.sh
[sudo] password for cs18:
# Executing docker install script, commit: 86415efcfe5f8d966625843da41a0f798238cce5

DEPRECATION WARNING
  This Linux distribution (ubuntu bionic) reached end-of-life and is no longer supported by this script.
  No updates or security fixes will be released for this distribution, and users are recommended to upgrade to a currently maintained version of ubuntu.

Press Ctrl+C now to abort this script, or wait for the installation to continue.
+ sh -c apt-get -qq update >/dev/null
+ sh -c DEBIAN_FRONTEND=noninteractive apt-get -y -qq install ca-certificates curl >/dev/null
+ sh -c install -m 0755 -d /etc/apt/keyrings
+ sh -c curl -fsSL "https://download.docker.com/linux/ubuntu/gpg" -o /etc/apt/keyrings/docker.asc
+ sh -c chmod a+r /etc/apt/keyrings/docker.asc
+ sh -c echo "deb [arch=amd64 signed-by=/etc/apt/keyrings/docker.asc] https://download.docker.com/linux/ubuntu bionic stable" > /etc/apt/sources.list.d/docker.list
+ sh -c apt-get -qq update >/dev/null
+ sh -c DEBIAN_FRONTEND=noninteractive apt-get -y -qq install docker-ce docker-ce-cli containerd.io docker-compose-plugin docker-ce-rootless-extras docker-buildx-plugin docker-model-plugin >/dev/null
E: Unable to locate package docker-model-plugin
cs18@games101vm:~$ docker --version

Command 'docker' not found, but can be installed with:

sudo snap install docker      # version 27.5.1, or
sudo apt install docker.io

See 'snap info docker' for additional versions.
```

尝试使用多种方法解决问题后，选择使用系统自带的 docker.io 包进行安装

```
cs18@games101vm:~$ sudo apt install docker.io -y
Reading package lists... Done
Building dependency tree
Reading state information... Done
```

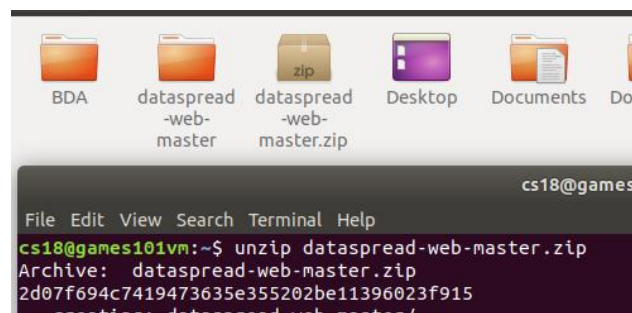
```
cs18@games101vm:~$ sudo apt install docker-compose -y
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
```

查看版本信息，可以看到安装成功：

```
cs18@games101vm:~$ docker --version
Docker version 20.10.21, build 20.10.21-0ubuntu1~18.04.3
cs18@games101vm:~$ docker-compose --version
docker-compose version 1.17.1, build unknown
```

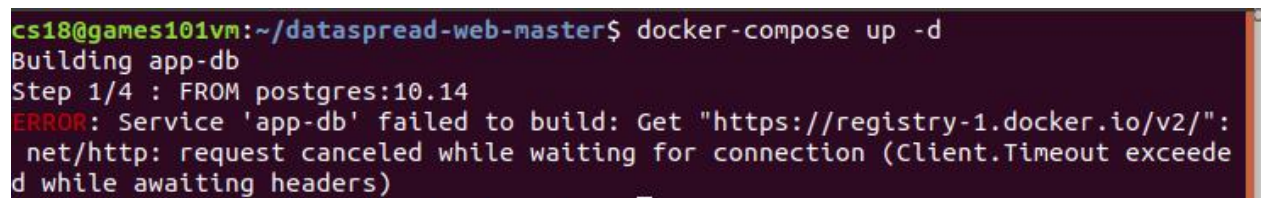

3、部署 DataSpread

在 <https://github.com/dataspread/dataspread-web> 中下载 zip 包并且放入虚拟机中后，使用 unzip 命令成功解压，如下图：



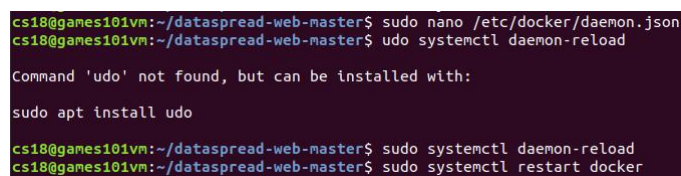
4、构建并启动 DataSpread 服务

使用命令 `docker-compose up -d` 启动服务时，始终显示无法从 Docker Hub 下载 PostgreSQL 镜像。

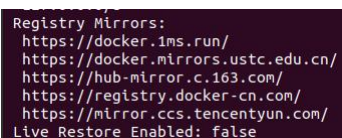


尝试检查网络连接后，无法解决问题，经过多次尝试，最终找到一个可使用的镜像源 <https://docker.lms.run/> 用于解决 Docker 镜像拉取的问题

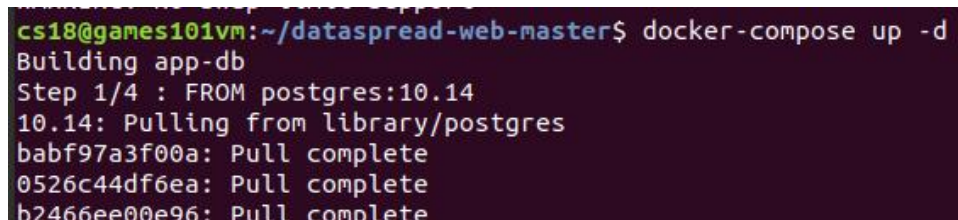
修改 Docker 配置文件后重启 Docker 服务



配置完成后，执行命令 `docker info` 验证是否生效，可以看到已经成功添加镜像仓库



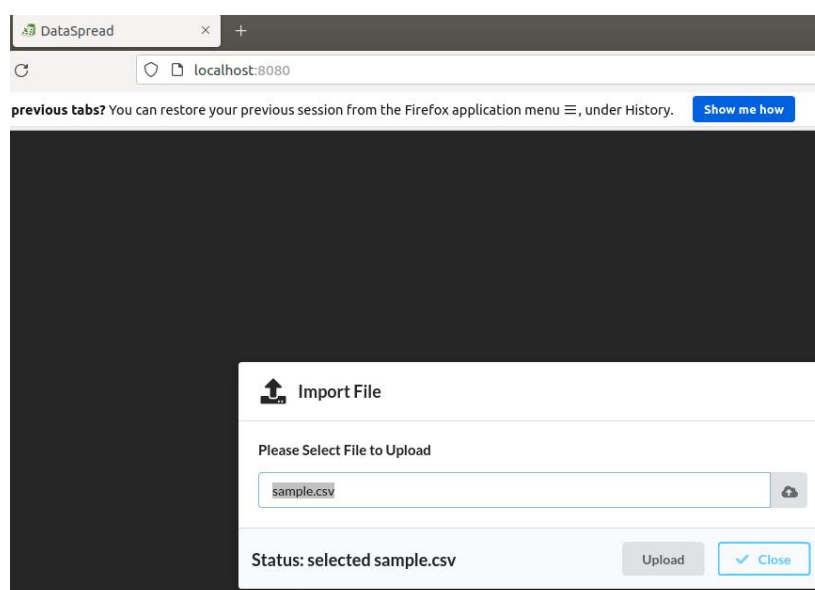
重新使用 Docker Compose 启动服务：



服务启动成功



访问网站 localhost:8080



导入 sample.csv，查看结果图：

	A	B
1	a	1.0
2	b	2.0
3	c	3.0
4		
5		
6		
7		

结论分析与体会：

1. 本次实验在 Ubuntu 环境下完成 DataSpread 部署与 sample.csv 数据导入，因系统版本限制,改用原生 docker.io 与 docker-compose 实现工具兼容,通过添加国内镜像源解决 Docker 镜像拉取超时问题，最终成功访问 localhost:8080 并完成数据导入，达成实验目标。
2. 实验中解决了工具缺失、Docker 安装失败、镜像拉取超时等问题，通过补装工具、选用系统原生软件、补充镜像源等方式逐步排查，积累了旧系统配置与 Docker 网络问题处理的经验。
3. 实验中体会到 Docker 部署效率优势，也明确其对环境的敏感性，需通过命令排查状态；团队协作核对参数、记录操作，减少了失误，并且各成员发挥所长，一起解决实验中遇到的问题，将理论知识应用于实践，深化了对容器化工具与大数据工具部署的理解。