emr快速部署文档-1.2.0

EMR-1.2.0安装

一,环境准备

1.主机规划

IP	主机名	配置	角色
192.168.90.226	master1.emr.lcoal	8c 16g 50+200G	Starry-server Starry-agent NN RM HS
192.168.90.229	master2.emr.lcoal	8c 16g 50+200G	Starry-agent NN RM DN NM
192.168.90.223	node1.emr.lcoal	8c 16g 50+200G	Starry-agent DN NM

2.部署包准备

Shell | 复制代码

```
#解压后目录如下
/data/EMR/
├── EMR
├── EMR-UTILS-1.1.0.22
├── soft
├── starry
```

二,安装步骤

1.环境初始化

1.1 上传安装包到指定目录

Shell | 复制代码

```
#上传emr-1.1.0.tgz/yum-1.1.0.tgz两个安装包到/data/tools并解压
tar zxf /data/tools/EMR-1.2.0.tgz -C /data
```

1.2 配置主机免密登陆

```shell
cat > /data/EMR/soft/ip.list << EOF
192.168.90.226
192.168.90.229
192.168.90.223
EOF

#创建ssh密钥串
ssh-keygen -f ~/.ssh/id_rsa -P "" -t rsa
#开通到所有节点的ssh免密通道（按需输入yes和root密码）
for i in `cat /data/EMR/soft/ip.list`;do ssh-copy-id $i;done
```

1.3 增加本地解析，并同步到所有节点

```shell
cat > /data/EMR/soft/hosts << EOF
#shuqi_emr
192.168.90.226 master1.emr.local master1 yum.emr.local
192.168.90.229 master2.emr.local master2
192.168.90.223 node1.emr.local   node1
EOF

#同步hosts文件到所有节点
for i in `cat /data/EMR/soft/ip.list`;do rsync -av /data/EMR/soft/hosts $i:/etc/;echo --------$i----------;done
for i in `cat /data/EMR/soft/ip.list`;do ssh $i "cat /etc/hosts";echo ----------$i----------;done
```

1.3 配置主机名

```shell
ssh master1 "hostnamectl --static set-hostname master1.emr.local"
ssh master2 "hostnamectl --static set-hostname master2.emr.local"
ssh node1   "hostnamectl --static set-hostname node1.emr.local"
```

1.5 执行系统检测/修复

Shell | 复制代码

```shell
#脚本自动进行检测，按需输入。如系统检测有异常项，执行修复，全部通过即可
cd /data/EMR/soft
#检测
sh opstools.sh checklist
#修复
sh opstools.sh repair
```

1.7 chrony时间同步(可选)

```shell
#所有节点安装chrony服务
for i in `cat /data/EMR/soft/ip.list`;do ssh $i "yum -y install chrony";echo --------$i----------;done

#mastser1作为服务端配置如下
vim /etc/chrony.conf
server ntp.aliyun.com iburst   #有网络可以写阿里云的ntp 订正时间。没有就写自己 手动对正
driftfile /var/lib/chrony/drift
makestep 1.0 3
rtcsync
allow 192.168.90.0/24
local stratum 10
logdir /var/log/chrony

#master2和node1作为客户端配置如下
server master1.emr.local iburst
driftfile /var/lib/chrony/drift
makestep 1.0 3
rtcsync
logdir /var/log/chrony


#启动所有节点chrony服务并enable
for i in `cat /data/EMR/soft/ip.list`;do ssh $i "systemctl start chronyd && systemctl enable chronyd";echo --------$i----------;done

#查看是否同步
for i in `cat /data/EMR/soft/ip.list`;do ssh $i "chronyc sources -v";echo --------$i----------;done
^* 203.107.6.88
带星即可
```

2.starry依赖安装

2.1. 配置EMR安装源

```shell
#配置emr目录为yum源 （master1执行）
sh /data/EMR/soft/install_nginx.sh
#页面访问验证
http://master1.emr.local/


#更改repo模版并分发给所有主机
for i in `cat /data/EMR/soft/ip.list`;do rsync -av /data/EMR/soft/emr.repo $i:/etc/yum.repos.d/;echo --------$i----------;done

#验证yum可用性。是否有starry包
for i in `cat /data/EMR/soft/ip.list`;do ssh $i "yum list | grep starry";echo --------$i----------;done
```

2.2 各节点安装JDK

```shell
#安装jdk（master1执行）
cd /data/EMR/soft/
for i in `cat /data/EMR/soft/ip.list`;do echo $i ;rsync -av /data/EMR/soft/{jdk,jdk1.8.0_241} $i:/opt/third/ ;echo --------$i----------;done
for i in `cat /data/EMR/soft/ip.list`;do echo $i ;rsync -av /data/EMR/soft/default $i:/usr/java/;echo --------$i----------;done
for i in `cat /data/EMR/soft/ip.list`;do echo $i ;rsync -av /data/EMR/soft/mysql-connector-java.jar $i:/usr/share/java/;echo --------$i----------;done
for i in `cat /data/EMR/soft/ip.list`;do ssh $i 'echo "export JAVA_HOME=/opt/third/jdk" >> /etc/profile';echo --------$i----------;done
for i in `cat /data/EMR/soft/ip.list`;do ssh $i 'echo "export PATH=\$PATH:\$JAVA_HOME/bin" >> /etc/profile';echo --------$i----------;done
for i in `cat /data/EMR/soft/ip.list`;do ssh $i 'source /etc/profile &&java -version';echo --------$i----------;done
```

2.3 mysql安装

依赖系统yum源

```shell
#安装mysql（master1执行）
sh /data/EMR/soft/install_mysql.sh
```

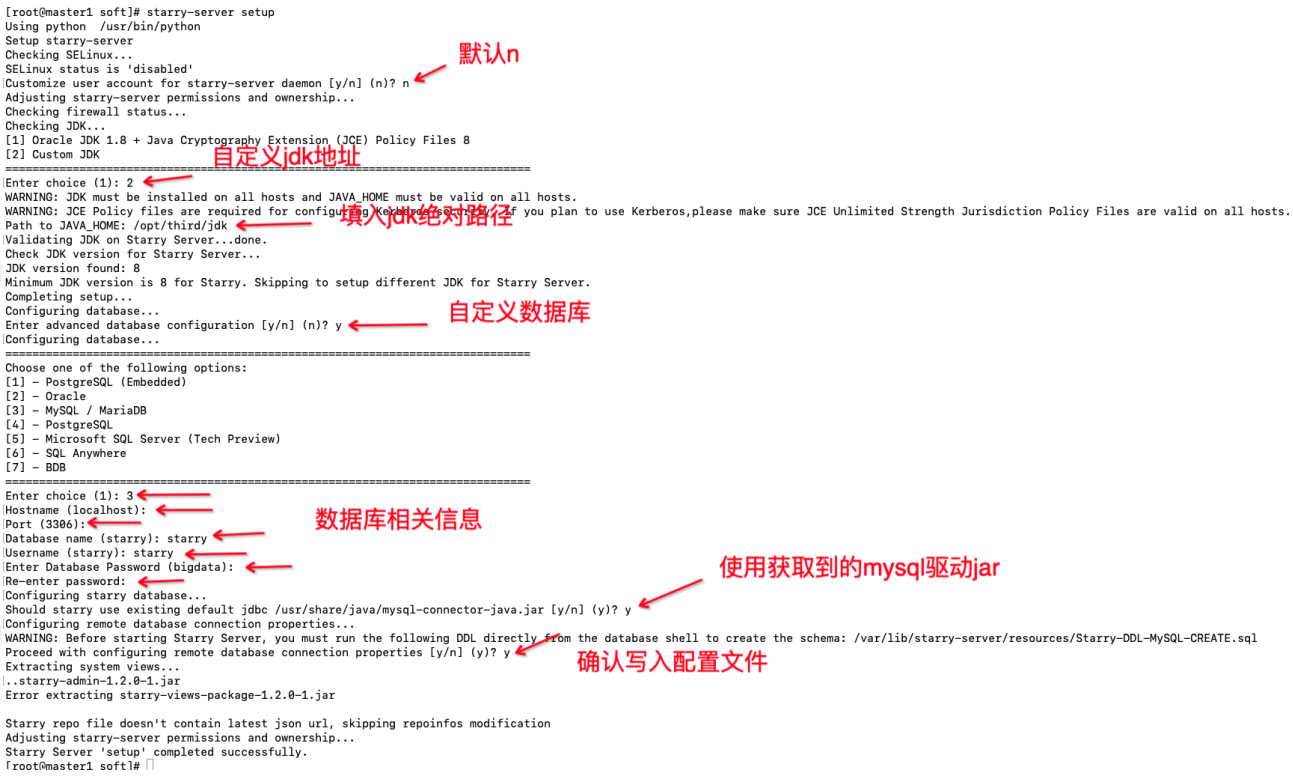
3.staryy安装启动

3.1starry-server安装配置

Shell | 复制代码

```shell
#starry-server安装
yum install -y starry-server

#配置license
vim /etc/profile
export SECURITY_LICENSE=xxxxxxxxxx

source /etc/profile

#starry-server配置
[root@master1 soft]# starry-server setup
Using python  /usr/bin/python
Setup starry-server
Checking SELinux...
SELinux status is 'disabled'
Customize user account for starry-server daemon [y/n] (n)? n
Adjusting starry-server permissions and ownership...
Checking firewall status...
Checking JDK...
[1] Oracle JDK 1.8 + Java Cryptography Extension (JCE) Policy Files 8
[2] Custom JDK
================================================================================
Enter choice (1): 2
WARNING: JDK must be installed on all hosts and JAVA_HOME must be valid on all hosts.
WARNING: JCE Policy files are required for configuring Kerberos security. If you plan to use Kerberos,please make sure JCE Unlimited Strength Jurisdiction Policy Files are valid on all hosts.
Path to JAVA_HOME: /opt/third/jdk
Validating JDK on Starry Server...done.
Check JDK version for Starry Server...
JDK version found: 8
Minimum JDK version is 8 for Starry. Skipping to setup different JDK for Starry Server.
Completing setup...
Configuring database...
Enter advanced database configuration [y/n] (n)? y
Configuring database...
================================================================================
Choose one of the following options:
[1] - PostgreSQL (Embedded)
[2] - Oracle
[3] - MySQL / MariaDB
```

```
[4] - PostgreSQL
[5] - Microsoft SQL Server (Tech Preview)
[6] - SQL Anywhere
[7] - BDB
================================================================================
Enter choice (1): 3
Hostname (localhost):
Port (3306):
Database name (starry):
Username (starry):
Enter Database Password (Shuq1-1298):
Re-enter password:
Configuring starry database...
Should starry use existing default jdbc /usr/share/java/mysql-connector-java.jar [y/n] (y)? y
Configuring remote database connection properties...
WARNING: Before starting Starry Server, you must run the following DDL directly from the database shell to create the schema: /var/lib/starry-server/resources/Starry-DDL-MySQL-CREATE.sql
Proceed with configuring remote database connection properties [y/n] (y)? y
Extracting system views...
starry-admin-1.1.0.0.0.jar

Starry repo file doesn't contain latest json url, skipping repoinfos modification
Adjusting starry-server permissions and ownership...
Starry Server 'setup' completed successfully.



#初始化 starry 数据库连接配置
starry-server setup --jdbc-db=mysql --jdbc-driver=/usr/share/java/mysql-connector-java.jar
#初始化starry库
mysql -ustarry -pShuq1-1298  starry < /var/lib/starry-server/resources/Starry-DDL-MySQL-CREATE.sql;
#启动starry-server
starry-server start
echo "source /etc/profile && starry-server restart" >> /etc/rc.d/rc.local
```



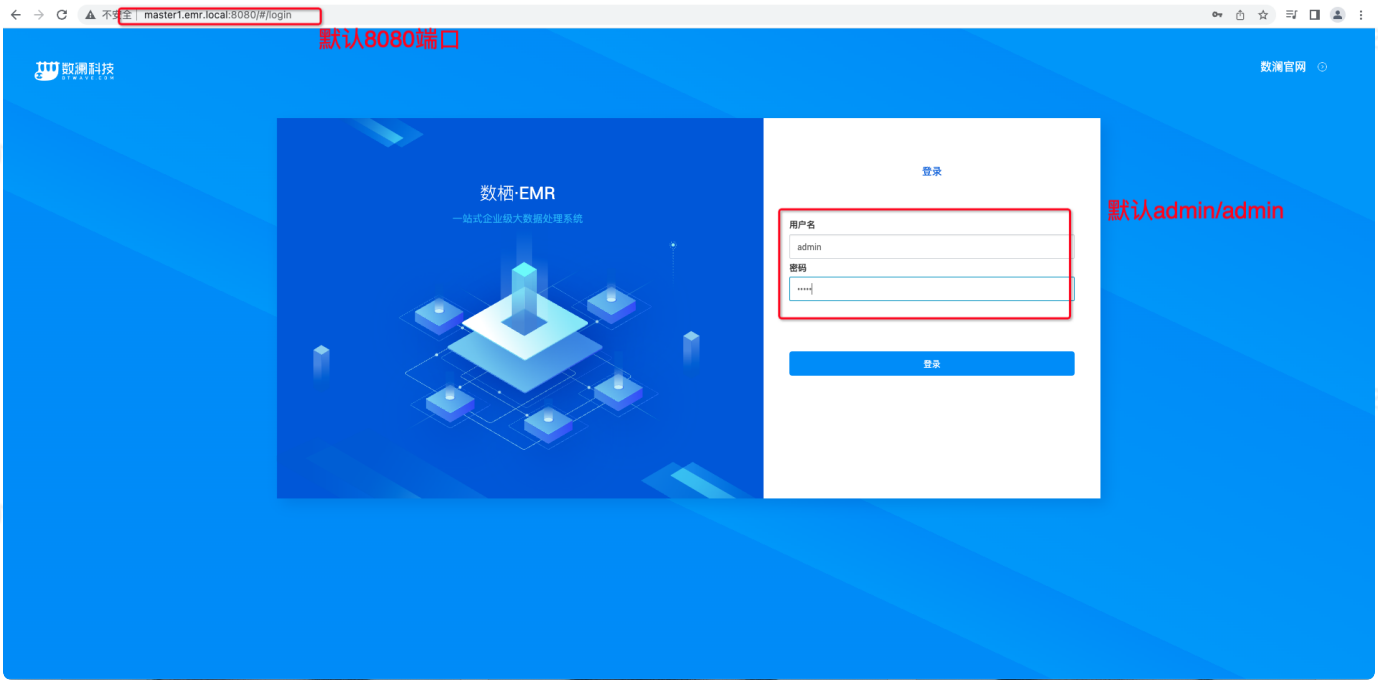

```
[root@master1 soft]# starry-server setup
Using python  /usr/bin/python
Setup starry-server
Checking SELinux...
SELinux status is 'disabled'
Customize user account for starry-server daemon [y/n] (n)? n
Adjusting starry-server permissions and ownership...
Checking firewall status...
Checking JDK...
[1] Oracle JDK 1.8 + Java Cryptography Extension (JCE) Policy Files 8
[2] Custom JDK
================================================================================
Enter choice (1): 2
WARNING: JDK must be installed on all hosts and JAVA_HOME must be valid on all hosts.
WARNING: JCE Policy files are required for configuring Kerberos security. If you plan to use Kerberos,please make sure JCE Unlimited Strength Jurisdiction Policy Files are valid on all hosts.
Path to JAVA_HOME: /opt/third/jdk
Validating JDK on Starry Server...done.
Check JDK version for Starry Server...
JDK version found: 8
Minimum JDK version is 8 for Starry. Skipping to setup different JDK for Starry Server.
Completing setup...
Configuring database...
Enter advanced database configuration [y/n] (n)? y
Configuring database...
================================================================================
Choose one of the following options:
[1] - PostgreSQL (Embedded)
[2] - Oracle
[3] - MySQL / MariaDB
[4] - PostgreSQL
[5] - Microsoft SQL Server (Tech Preview)
[6] - SQL Anywhere
[7] - BDB
================================================================================
Enter choice (1): 3
Hostname (localhost):
Port (3306):
Database name (starry): starry
Username (starry): starry
Enter Database Password (bigdata):
Re-enter password:
Configuring starry database...
Should starry use existing default jdbc /usr/share/java/mysql-connector-java.jar [y/n] (y)? y
Configuring remote database connection properties...
WARNING: Before starting Starry Server, you must run the following DDL directly from the database shell to create the schema: /var/lib/starry-server/resources/Starry-DDL-MySQL-CREATE.sql
Proceed with configuring remote database connection properties [y/n] (y)? y
Extracting system views...
..starry-admin-1.2.0-1.jar
Error extracting starry-views-package-1.2.0-1.jar

Starry repo file doesn't contain latest json url, skipping repoinfos modification
Adjusting starry-server permissions and ownership...
Starry Server 'setup' completed successfully.
[root@master1 soft]#
```

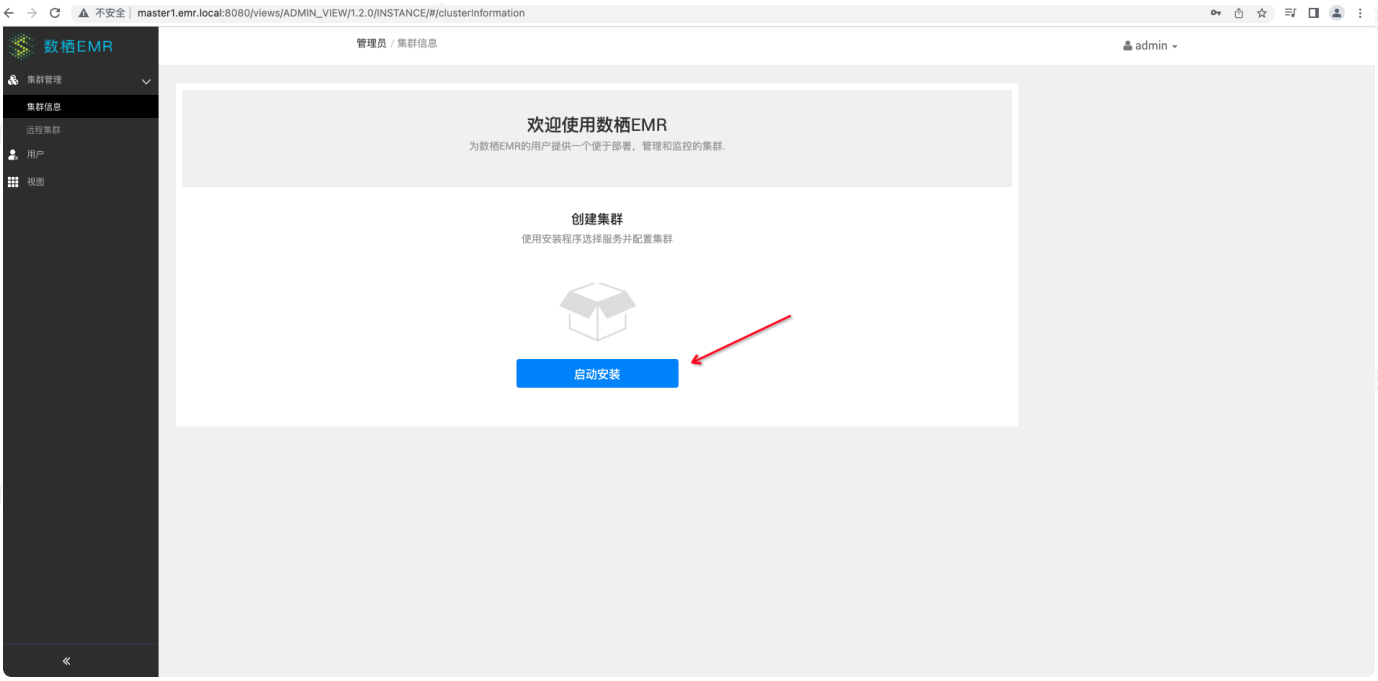
3.3 starry-agent安装配置

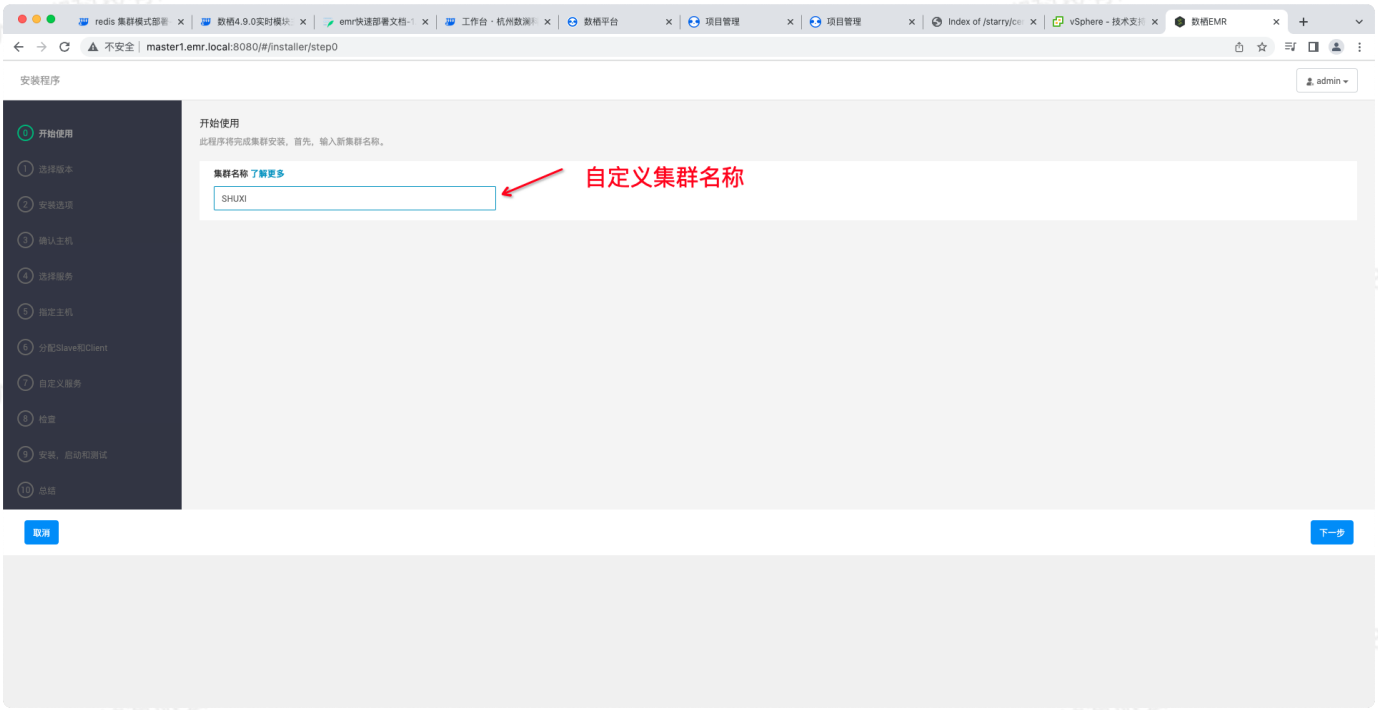
Shell | 复制代码

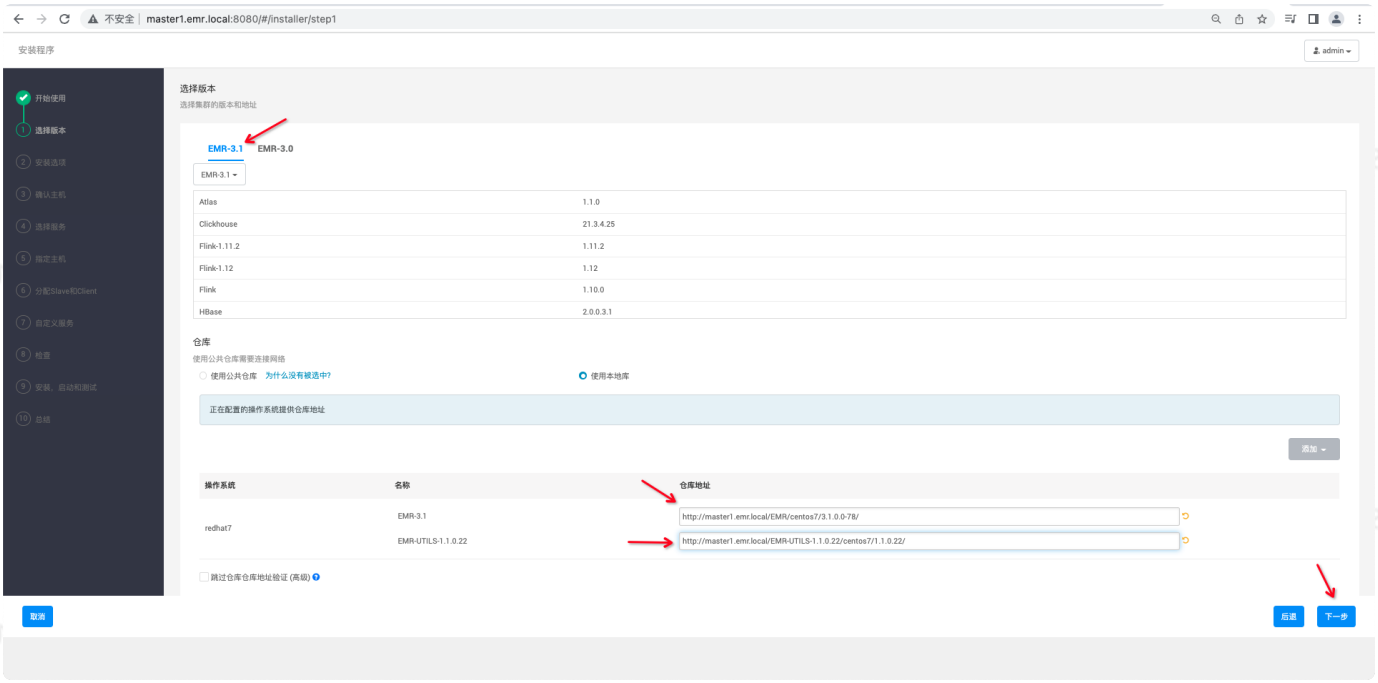
```shell
#starry-agent安装
for i in `cat /data/EMR/soft/ip.list`;do ssh $i 'yum install -y starry-agent';echo --------$i----------;done

#更改客户端配置
for i in `cat /data/EMR/soft/ip.list`;do ssh $i "sed -i '/^hostname=/c hostname=master1' /etc/starry-agent/conf/starry-agent.ini";echo --------$i----------;done
for i in `cat /data/EMR/soft/ip.list`;do ssh $i 'starry-agent start';echo --------$i----------;done
```

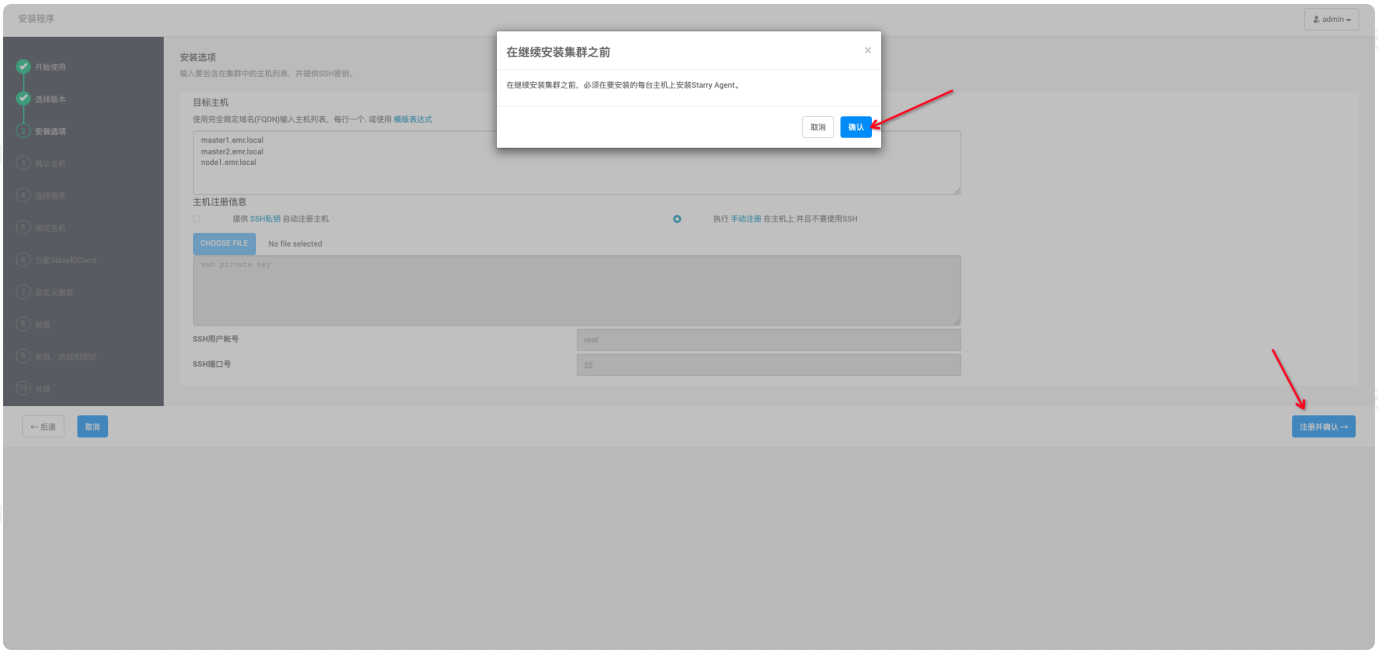
3.4 页面安装Hadoop集群

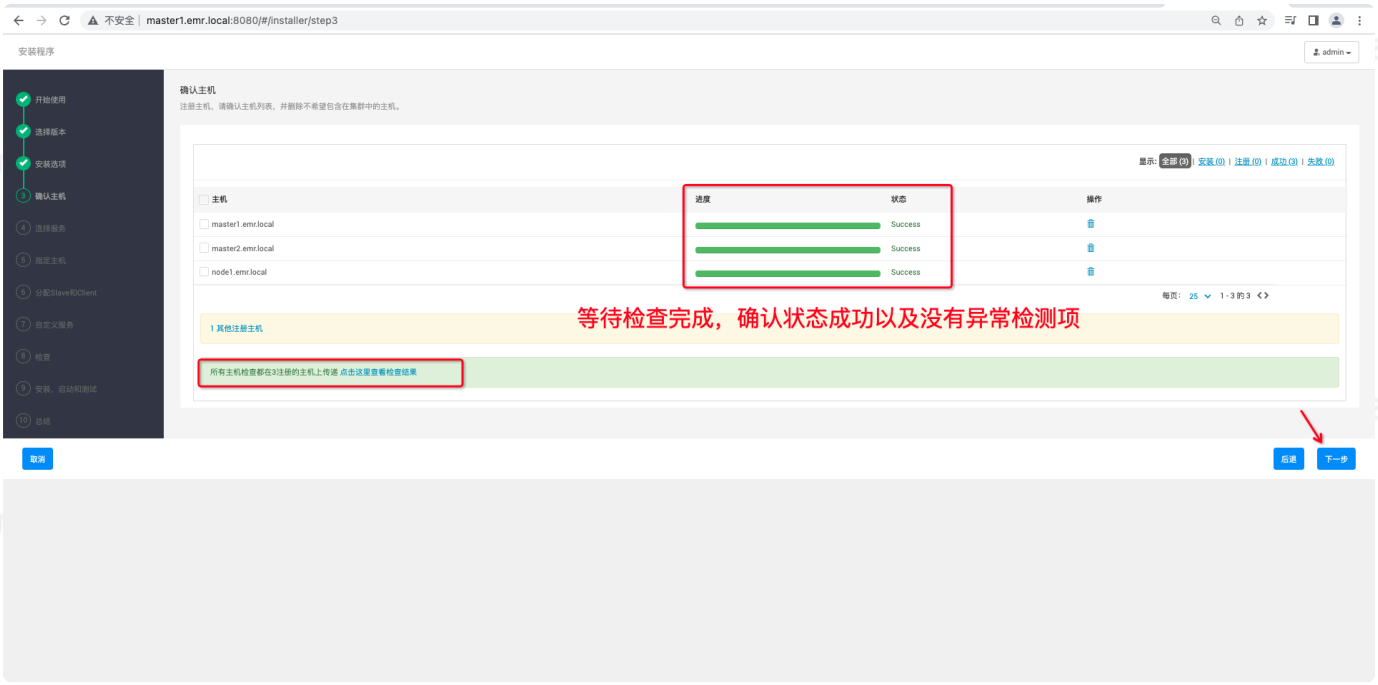




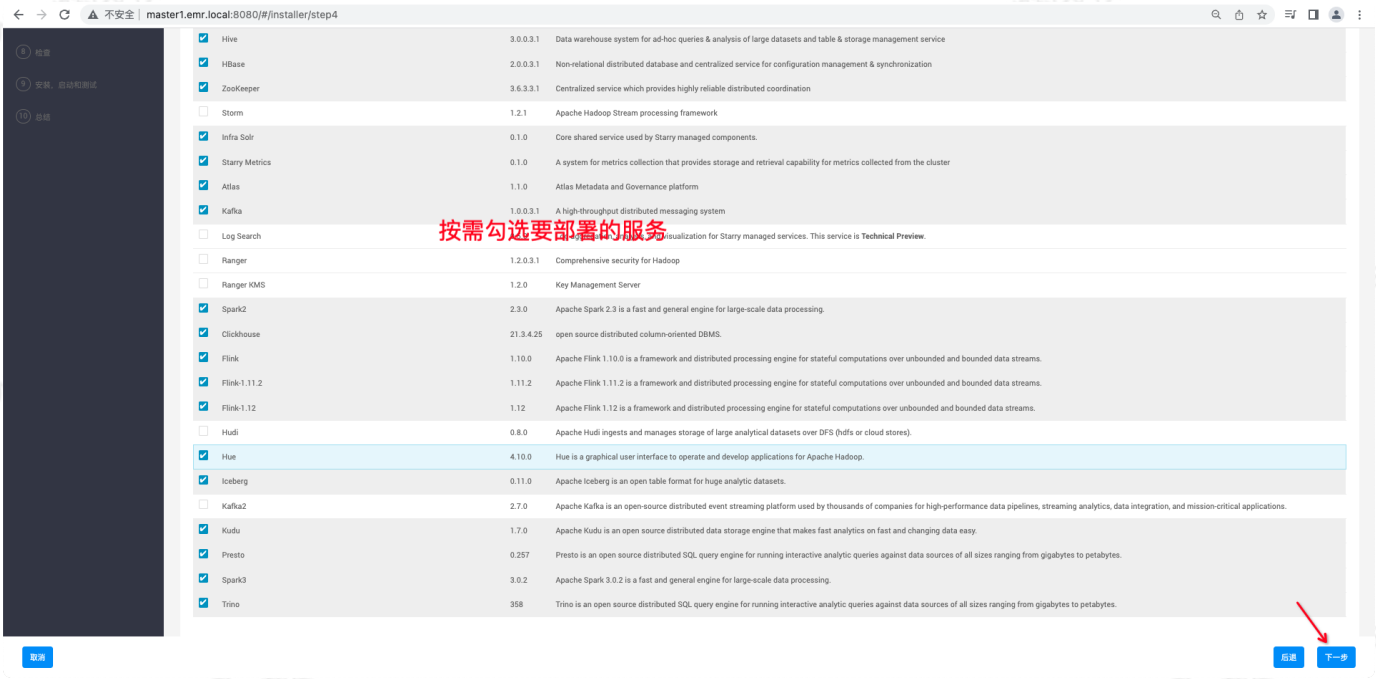


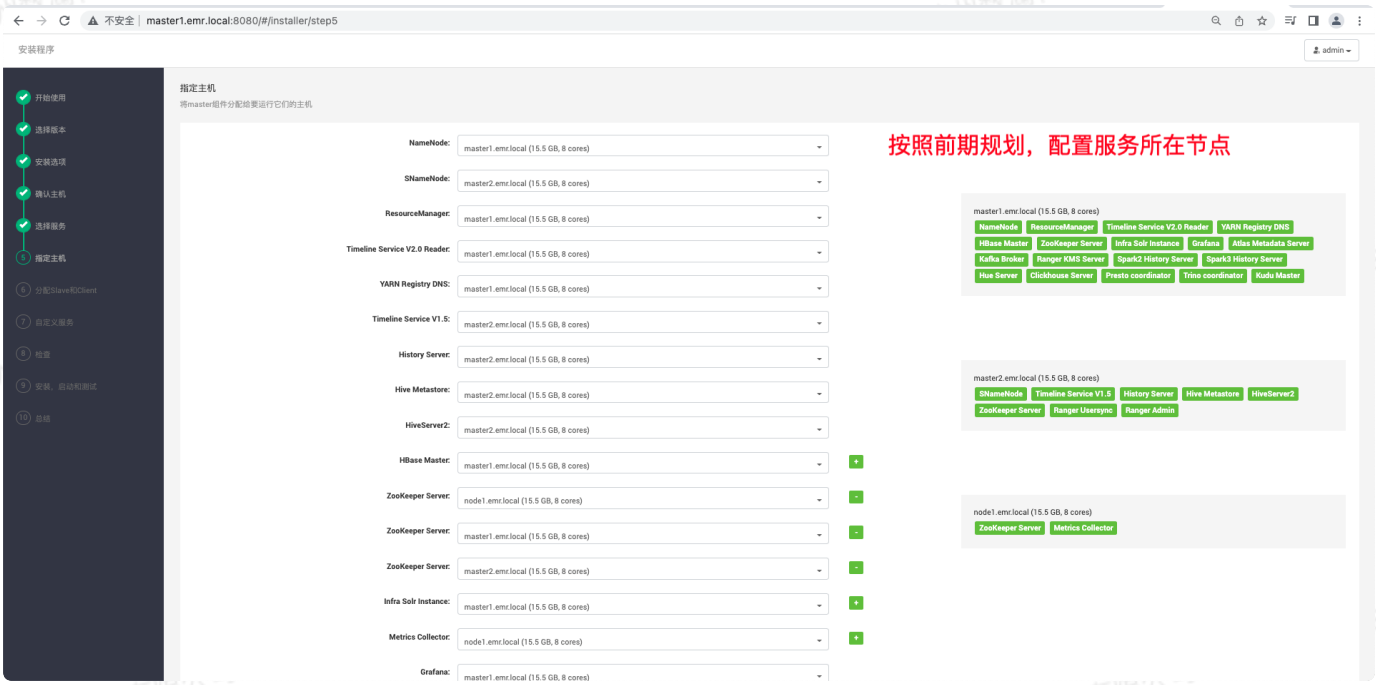


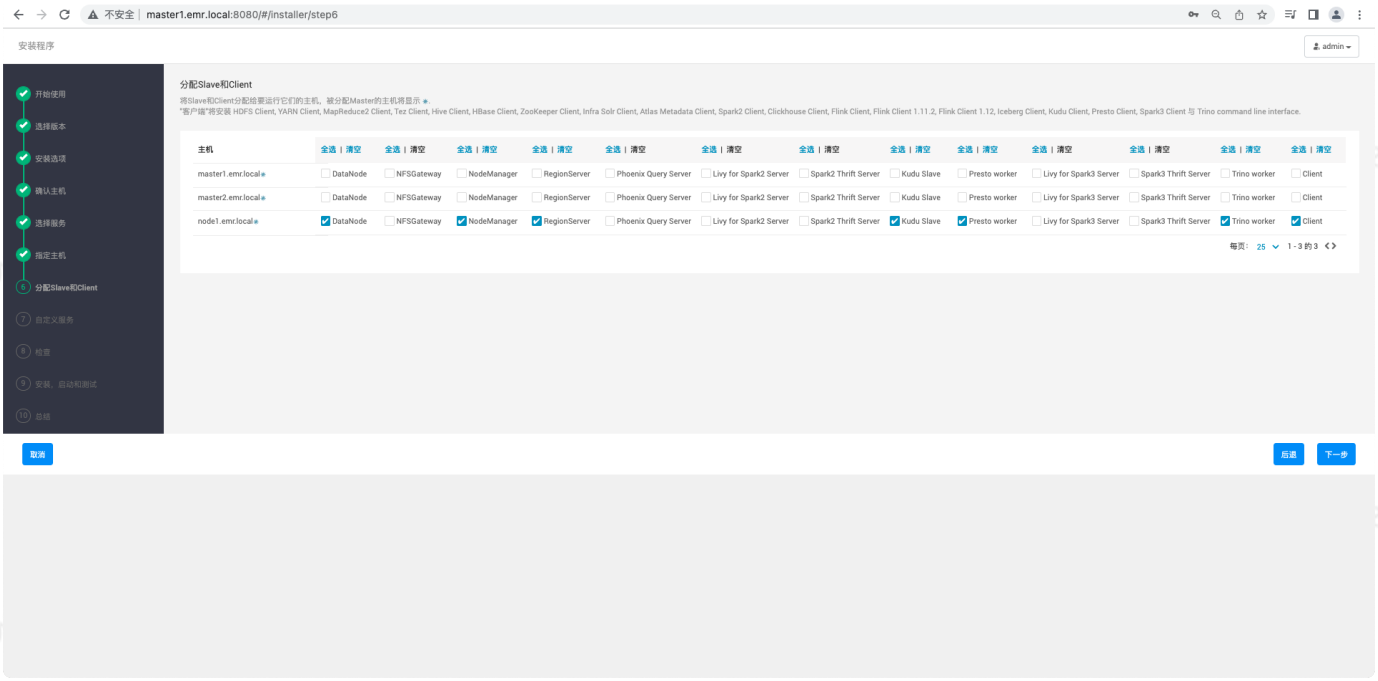


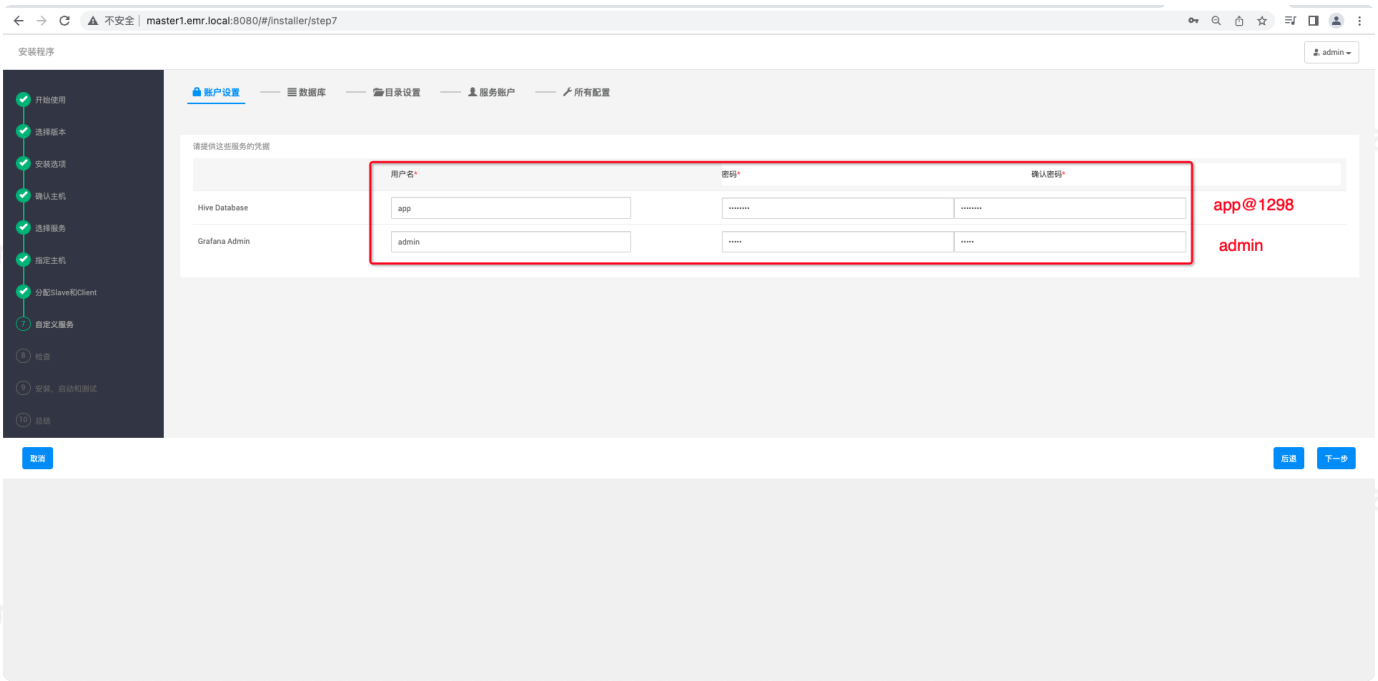


第一次初始化安装集群时，spark3、flink-1.12、clickhouse不能一起部署，该服务依赖部分Hadoop集群的服务正常才能部署成功。可以在部署完成后，再增加服务。

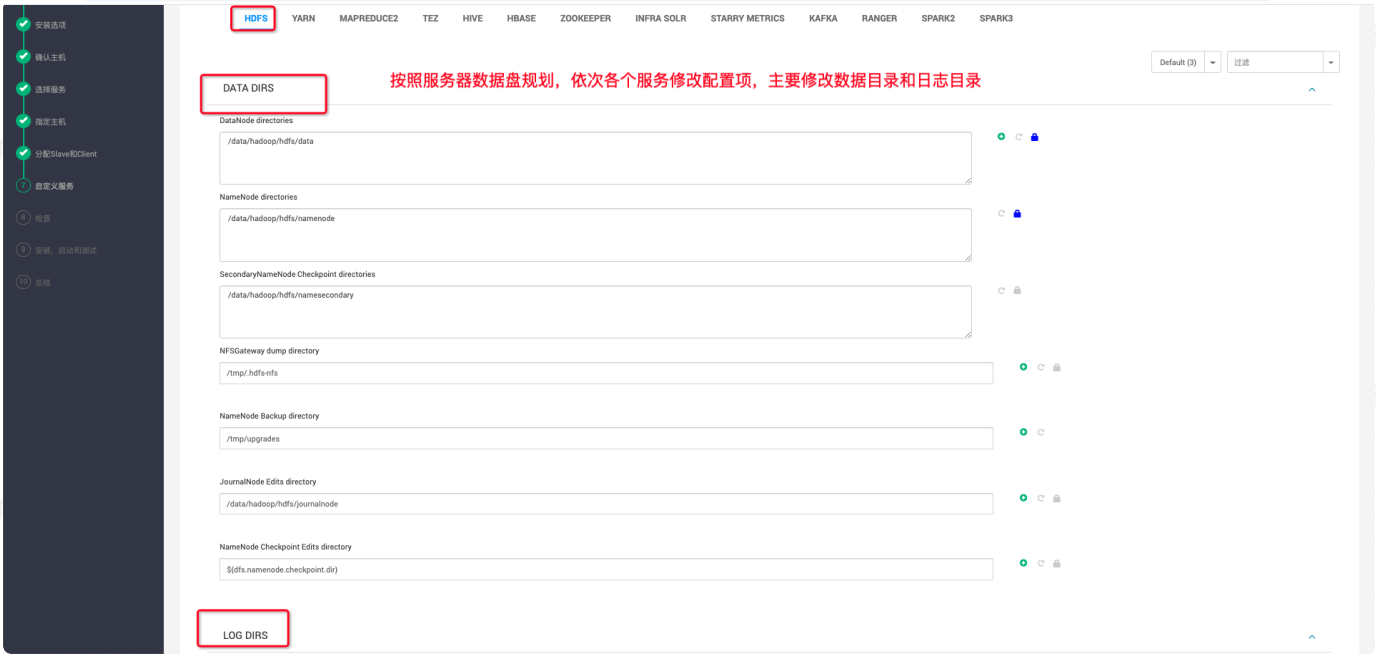


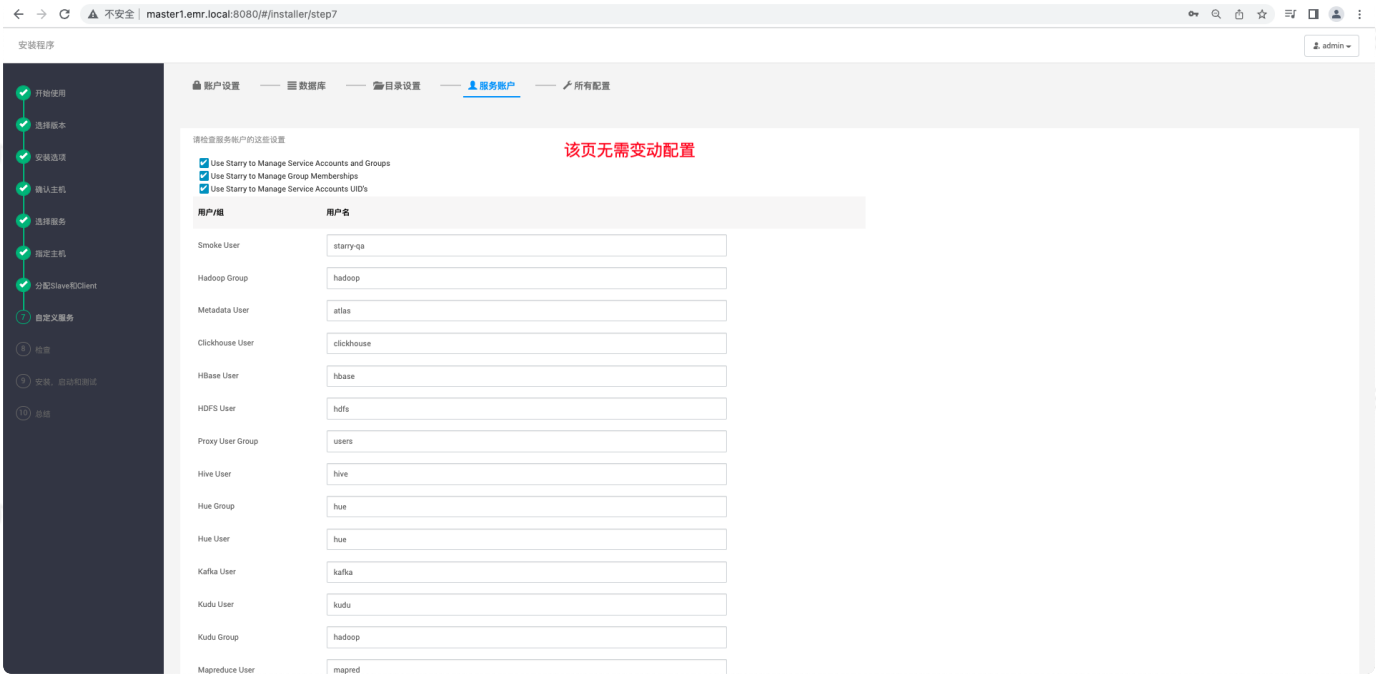


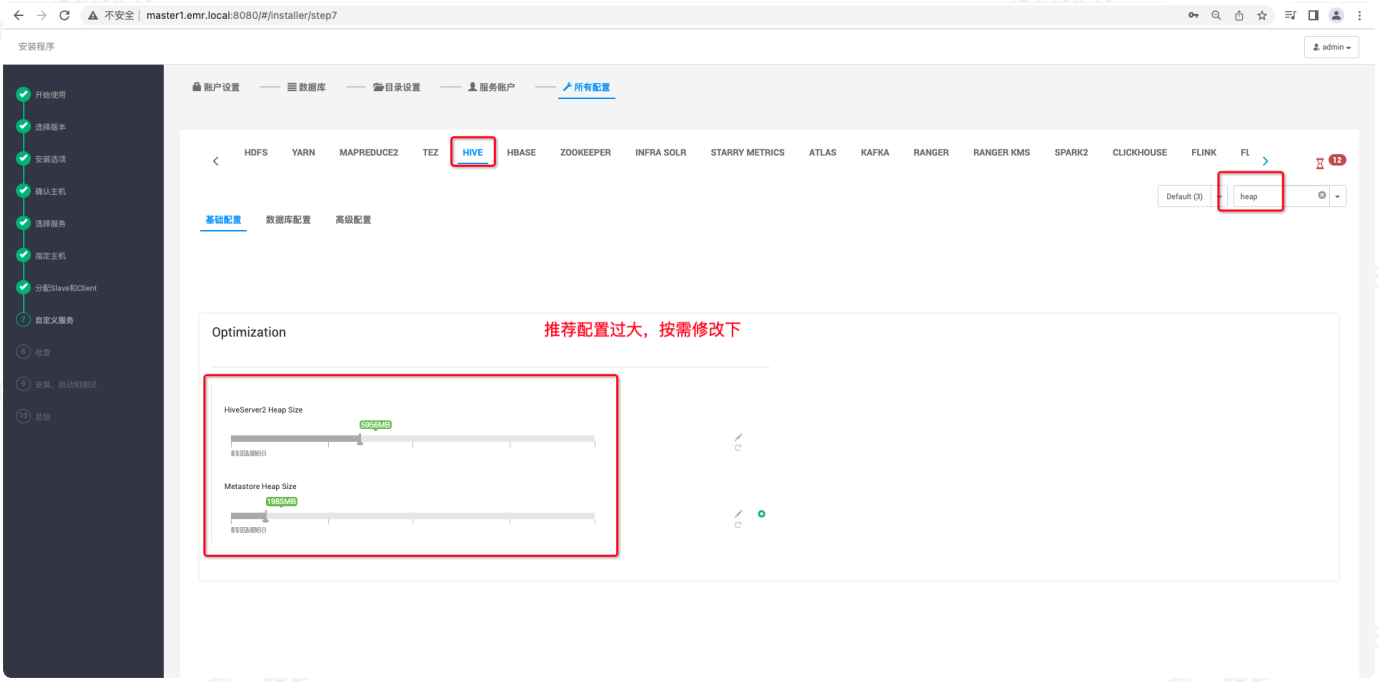


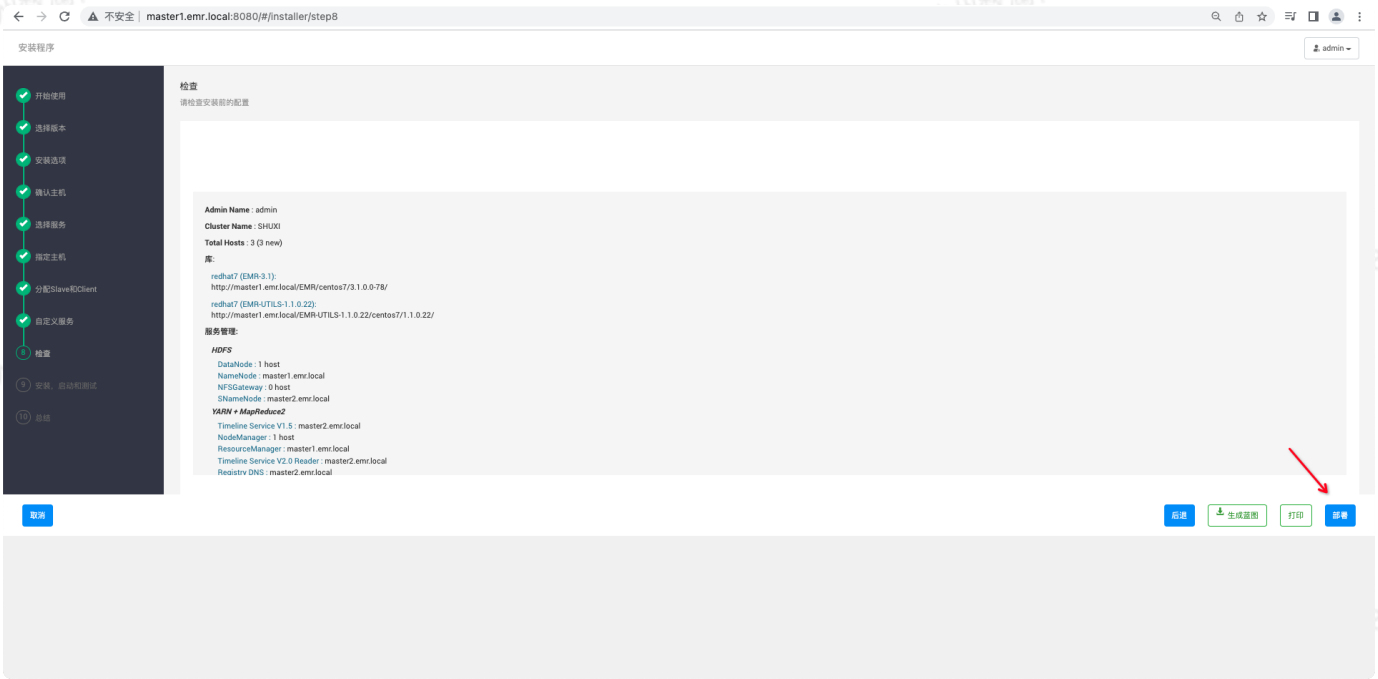


依次修改各个服务的配置，以HDFS为例：









Shell 复制代码

```
注意事项：
1.数据库：hive的元数据库配置增加参数
?characterEncoding=utf8&useSSL=false

2.目录配置：各个服务
调整数据目录和日志目录

3.所有配置：
hive配置中jvm配置不合理，默认配置非常大，注意修改
hive配置中高级配置搜索hive.security.authorization.enabled,去掉该配置。（影响创建函数）
spark配置中搜索spark.ranger.enable，去掉该配置（来回点几次，默认值展示有问题）

```

Shell 复制代码

```
use hive;
alter table COLUMNS_V2 modify column COMMENT varchar(256) character set utf8;
alter table TABLE_PARAMS modify column PARAM_VALUE varchar(4000) character set utf8;
alter table PARTITION_PARAMS  modify column PARAM_VALUE varchar(4000) character set utf8;
alter table PARTITION_KEYS  modify column PKEY_COMMENT varchar(4000) character set utf8;
alter table  INDEX_PARAMS  modify column PARAM_VALUE  varchar(4000) character set utf8;
alter table COLUMNS_V2 modify column COLUMN_NAME varchar(767) character set utf8;
```

Bash 复制代码

```
# EMR安装后HDFS和YARN相关监控指标无法显示问题，可以按如下操作修复。
在HDFS -> Advanced core-site
hadoop.http.authentication.simple.anonymous.allowed=true
在HDFS -> Custom core-site
hadoop.http.authentication.type=simple
hadoop.proxyuser.HTTP.groups=*
set hadoop.proxyuser.knox.groups=*
set hadoop.proxyuser.knox.hosts=*
set hadoop.proxyuser.yarn.hosts=*
```