

EMR 1.2集成LDAP_Kerberos

Kerberos安装

LDAP安装

EMR启动kerberos

EMR启动ranger

1.添加服务配置

2.登录管理页面

4.Ranger配置LDAP

hdfs配置修改，自定义core-site.xml中增加以下配置项

ranger配置common config

ranger配置user config

ranger配置group config

ranger plugins 配置

ranger高级配置

ranger配置保存

v1.0版本，计划简略安装步骤，基于脚本安装完成emr后

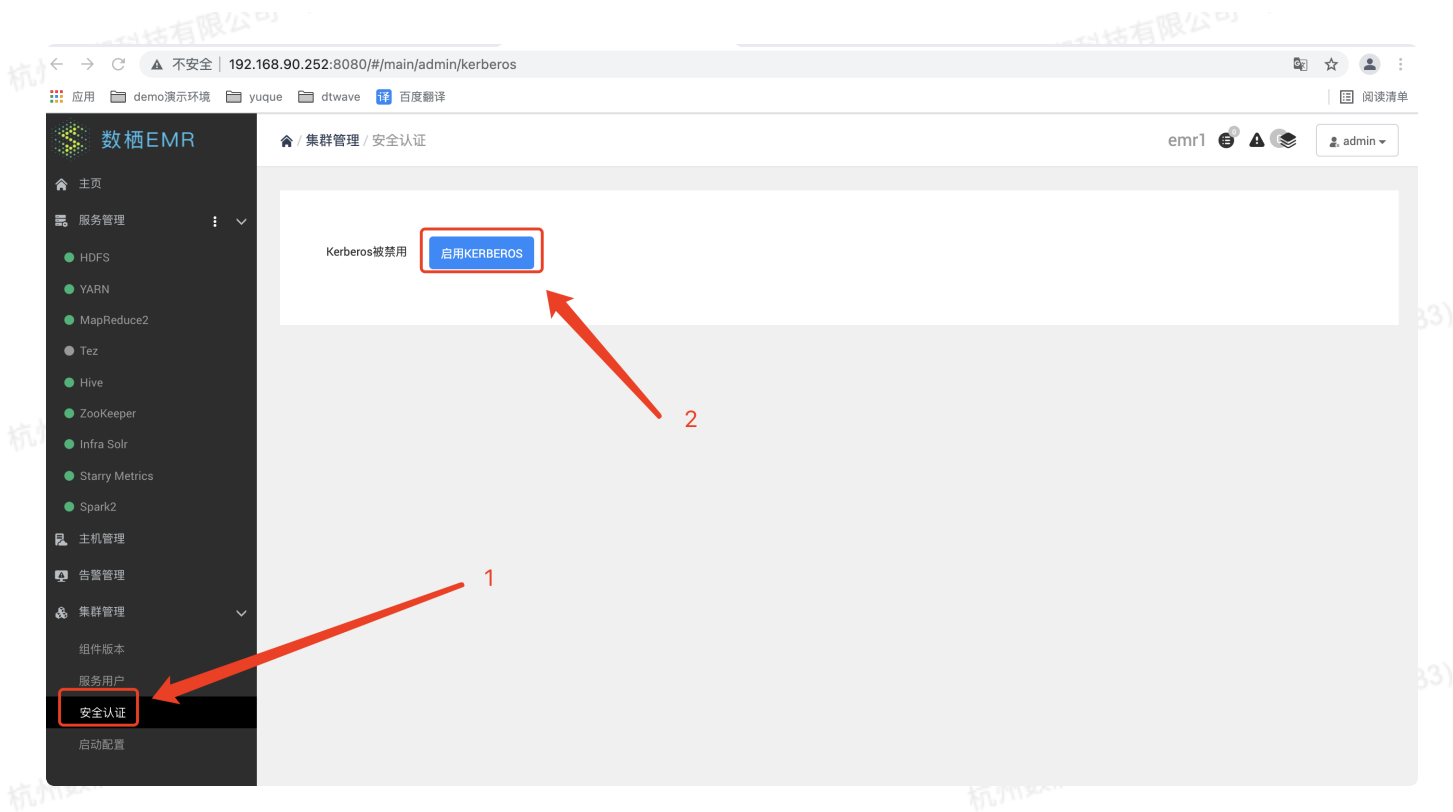
Kerberos安装

```
1 cd /data/EMR/soft
2 sh install_kdc.sh
3
4 #汇总输出如下:
5 ▾ #[install_kdc] - 安装信息汇总:
6 ▾ #[install_kdc] - kdc server: master1.emr.local
7 ▾ #[install_kdc] - kdc domain: HADOOP.COM
8 ▾ #[install_kdc] - kdc 数据库密码: Shuqi@1298
9 ▾ #[install_kdc] - kdc 管理员用户: admin/admin@HADOOP.COM
10 ▾ #[install_kdc] - kdc 管理员密码: Shuqi@1298
11 ▾ #[install_kdc] - kdc 客户端节点: 192.168.90.226 192.168.90.229 192.168.90.223
```

LDAP安装

```
1 cd /data/EMR/soft
2 sh install_ldap.sh
```

EMR启动kerberos



启用Kerberos程序

1 开始使用

2 配置Kerberos

3 安装和测试Kerberos Client

4 配置认证

5 确认配置

6 停止服务

7 安全集群

8 启动和测试服务

开始使用

欢迎来到Starry安全程序，使用此程序可以确保kerberos在集群中保持安全模式，让我们开始吧。

注意：此过程需要重新启动服务并停止集群，同样，根据选择的选项，可能需要安全管理员的支持，请合理规划。

使用什么类型的KDC?

☒ 已存在MIT KDC

☐ 已存在的Active Directory

☐ 已存在的IPA

☐ 手动管理Kerberos principal和keytab

已存在MIT KDC:

需要满足以下条件

☒ Starry服务器和集群主机具有对KDC和KDC管理主机的网络访问权限

☒ 已经拥有KDC的管理员账号

☒ 已在Starry Server主机和集群中的所有主机上设置Java加密扩展(JCE)包

下一步

启用Kerberos程序

配置Kerberos

KERBEROS

KDC

KDC type 已存在MIT KDC

KDC hosts master1.emr.local

Realm name HADOOP.COM

Domains

测试KDC连接 连接成功

Kadmin

Kadmin host master1.emr.local

Admin principal admin/admin@HADOOP.COM

Admin password 密码: admin

保存管理员认证

高级 kerberos-env

高级 krb5-conf

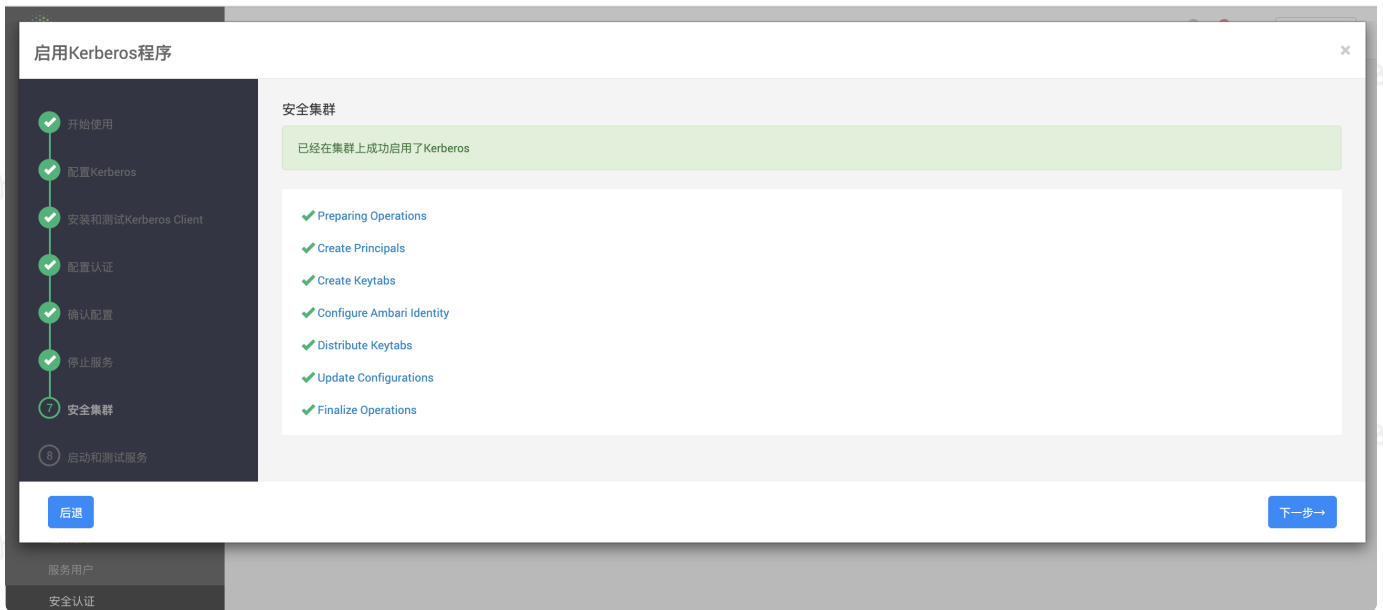
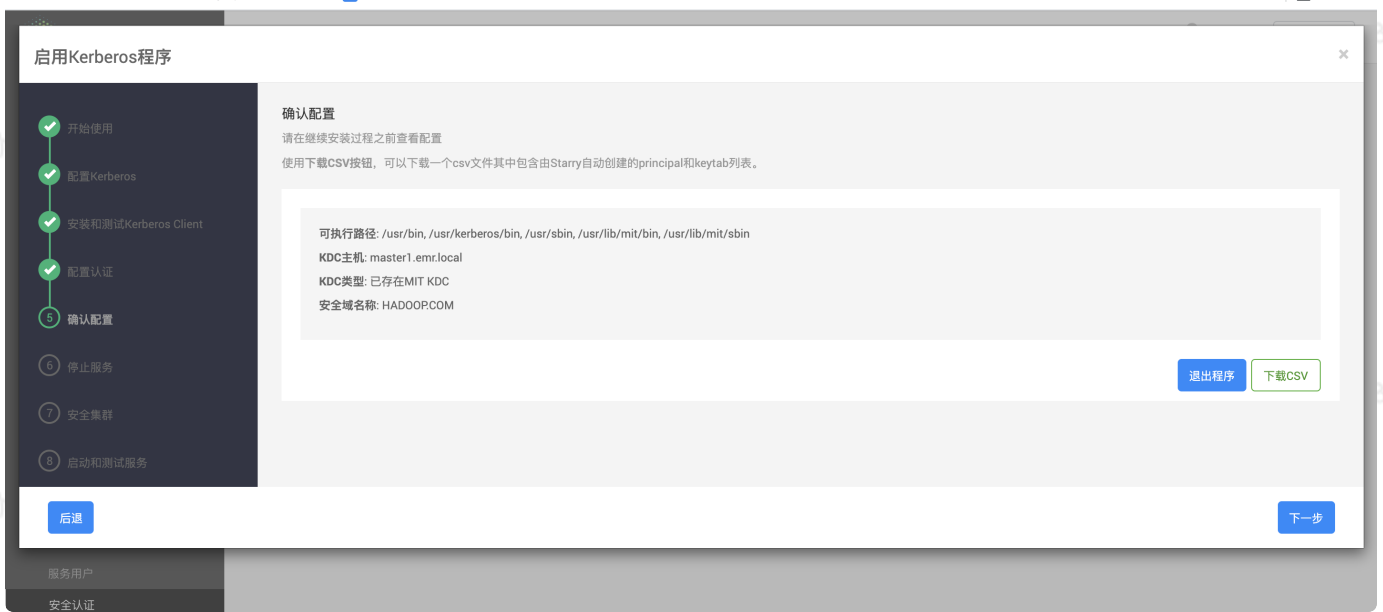
启用Kerberos程序

安装和测试Kerberos Client

Kerberos Client正在安装和测试之后请等待

安装Kerberos Client 测试Kerberos Client 68%

后退 下一步



EMR启动ranger

1.添加服务配置

添加服务程序

选择服务

指定主机

分配Slave和Client

自定义服务

检查

安装、启动和测试

总结

选择文件系统

选择要在集群上安装的文件系统

服务	版本	描述
☒ HDFS	3.1.1.3.1	Apache Hadoop Distributed File System

选择服务

选择在集群上要安装服务

服务	版本	描述
☒ YARN + MapReduce2	3.1.0	Apache Hadoop NextGen MapReduce (YARN)
☒ Tez	0.9.0.3.1	Tez is the next generation Hadoop Query Processing framework written on top of YARN.
☒ Hive	3.0.0.3.1	Data warehouse system for ad-hoc queries & analysis of large datasets and table & storage management service
☐ HBase	2.0.0.3.1	Non-relational distributed database and centralized service for configuration management & synchronization
☒ ZooKeeper	3.4.9.3.1	Centralized service which provides highly reliable distributed coordination
☒ Infra Solr	0.1.0	Core shared service used by Starry managed components.
☒ Starry Metrics	0.1.0	A system for metrics collection that provides storage and retrieval capability for metrics collected from the cluster
☐ Kafka	1.0.0.3.1	A high-throughput distributed messaging system
☒ Ranger	1.2.0.3.1	Comprehensive security for Hadoop
☐ Ranger KMS	1.2.0	Key Management Server
☒ Spark2	2.3.0	Apache Spark 2.3 is a fast and general engine for large-scale data processing.
☐ Clickhouse	21.3.4.25	open source distributed column-oriented DBMS.
☐ Flink	1.10.0	Apache Flink is a framework and distributed processing engine for stateful computations over unbounded and bounded data streams.
☐ Flink-1.11.2	1.11.2	Apache Flink is a framework and distributed processing engine for stateful computations over unbounded and bounded data streams.

添加服务程序

选择服务

指定主机

分配Slave和Client

自定义服务

检查

安装、启动和测试

总结

分配Slave和Client

将Slave和Client分配给要运行它们的主机。被分配Master的主机将显示。

主机	全选 清空	全选 清空	全选 清空	全选 清空	全选 清空	全选 清空
master1.emr.local	☐ DataNode	☐ NFSGateway	☐ NodeManager	☐ Livy for Spark3 Server	☐ Spark2 Thrift Server	☒ Ranger Tagsync
master2.emr.local	☒ DataNode	☐ NFSGateway	☒ NodeManager	☐ Livy for Spark3 Server	☐ Spark2 Thrift Server	☐ Ranger Tagsync
node1.emr.local	☒ DataNode	☐ NFSGateway	☒ NodeManager	☐ Livy for Spark3 Server	☐ Spark2 Thrift Server	☐ Ranger Tagsync

每页: 25

1-3 的 3

后测

下一步

jdbc:mysql://192.168.90.112:3306/ranger?characterEncoding=utf8&useSSL=false

添加服务程序

master1.emr.local:8080/#/main/service/add/step4

RANGER ADMIN RANGER USER INFO RANGER PLUGIN RANGER AUDIT RANGER TAGSYNC 高级配置 1

Ranger Admin

DB FLAVOR
MYSQL

Ranger DB host
master1.emr.local

Driver class name for a JDBC Ranger database
com.mysql.jdbc.Driver

Ranger DB name
ranger

Ranger DB password
app@1298

Ranger DB username
app

JDBC connect string for a Ranger database
jdbc:mysql://master1.emr.local:3306/ranger?characterEncoding=utf8&useSSL=false

Setup Database and Database User
No

TEST CONNECTION 连接成功

添加服务程序

master1.emr.local:8080/#/main/service/add/step4

RANGER ADMIN RANGER USER INFO RANGER PLUGIN RANGER AUDIT RANGER TAGSYNC 高级配置 1

Admin Settings

Ranger Admin host
master2.emr.local

Ranger Admin username for Ambari
amb_ranger_admin

Ranger Admin user's password for Ambari
ranger@1298

Ranger Usersync user's password
ranger@1298

Ranger Tagsync user's password
ranger@1298

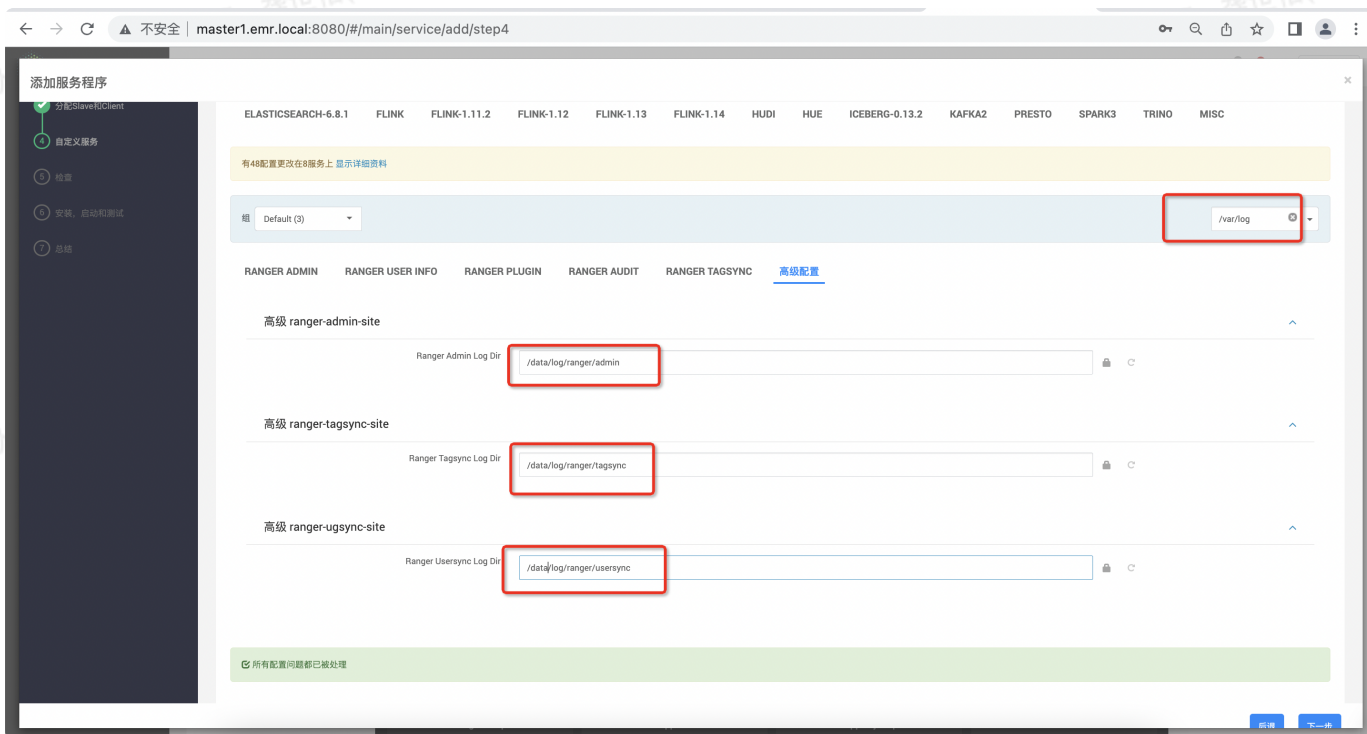
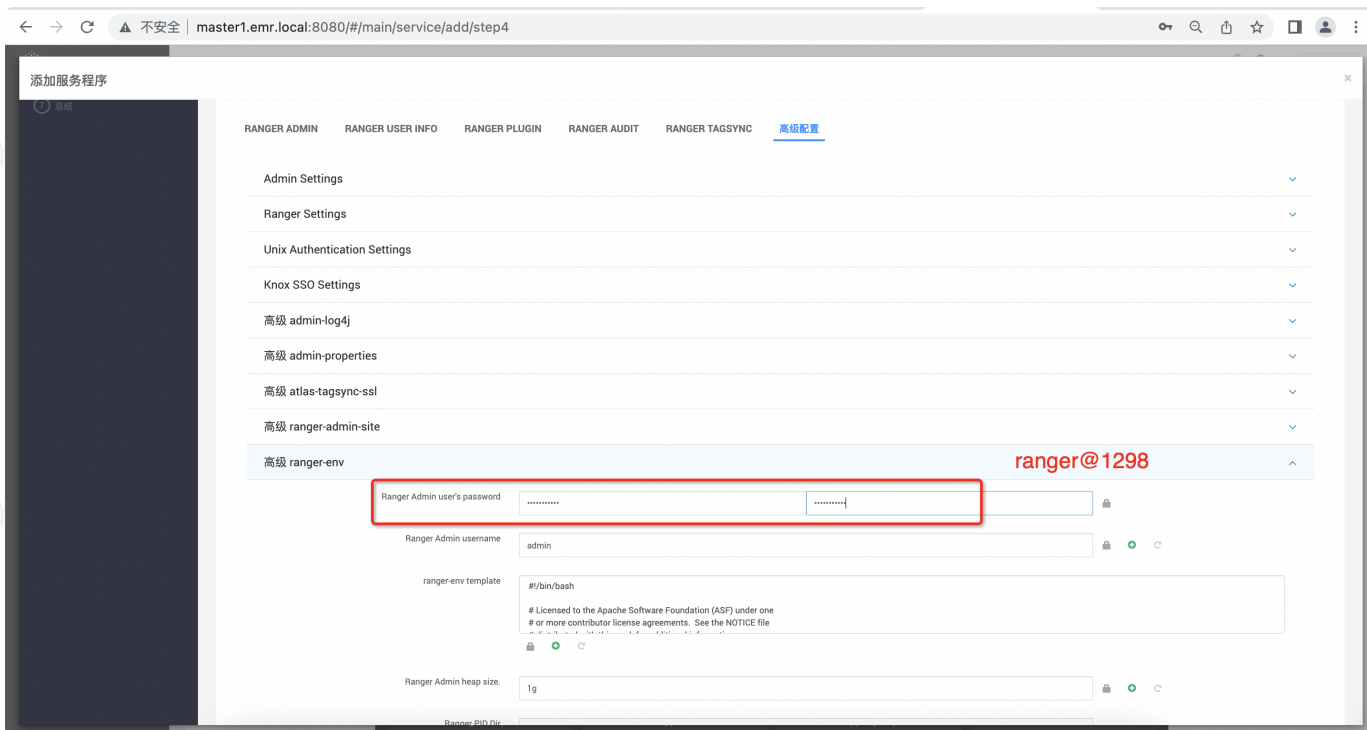
Ranger KMS keyadmin user's password
ranger@1298

Ranger Settings

External URL
http://master2.emr.local:6080

Authentication method
☐ LDAP
☐ ACTIVE_DIRECTORY
☒ UNIX
☐ PAM
☐ NONE

HTTP enabled
☒



2.登录管理页面

点击右上侧连接：

<http://192.168.90.252:6080/index.html#!/policymanager/resource>

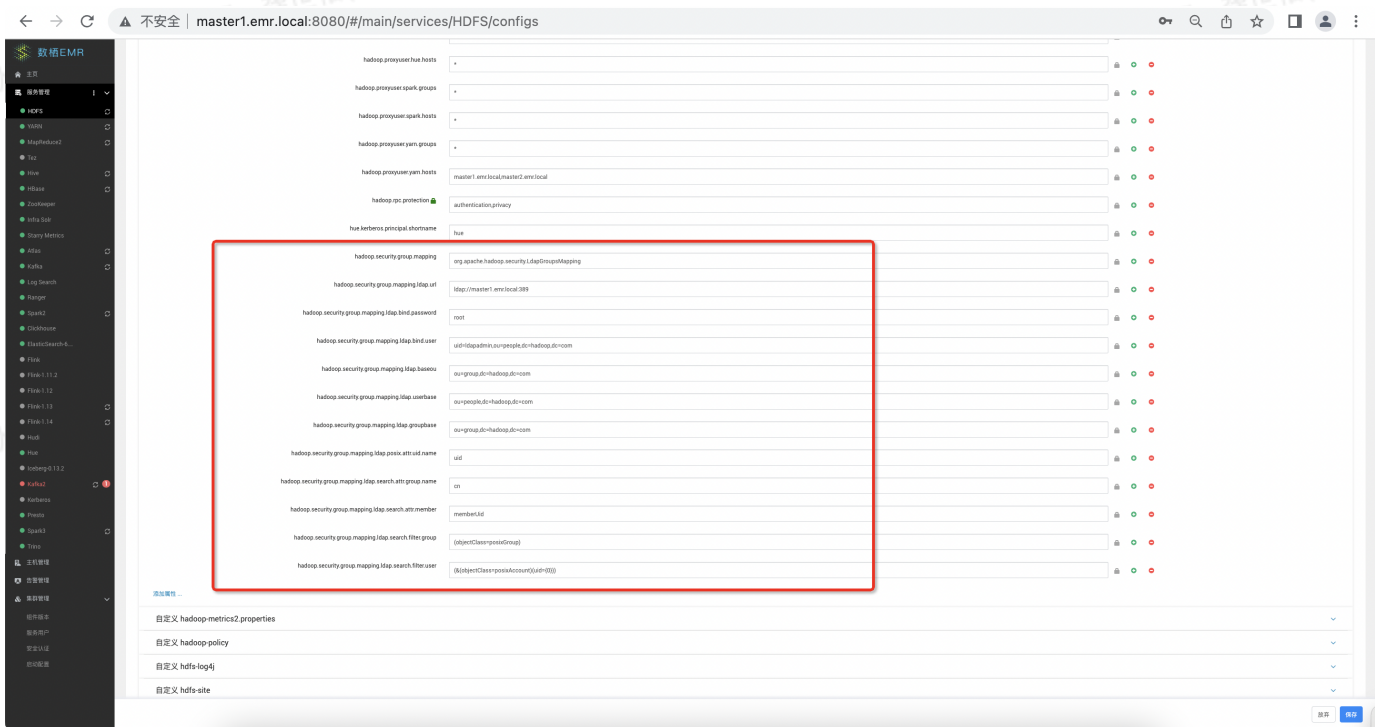
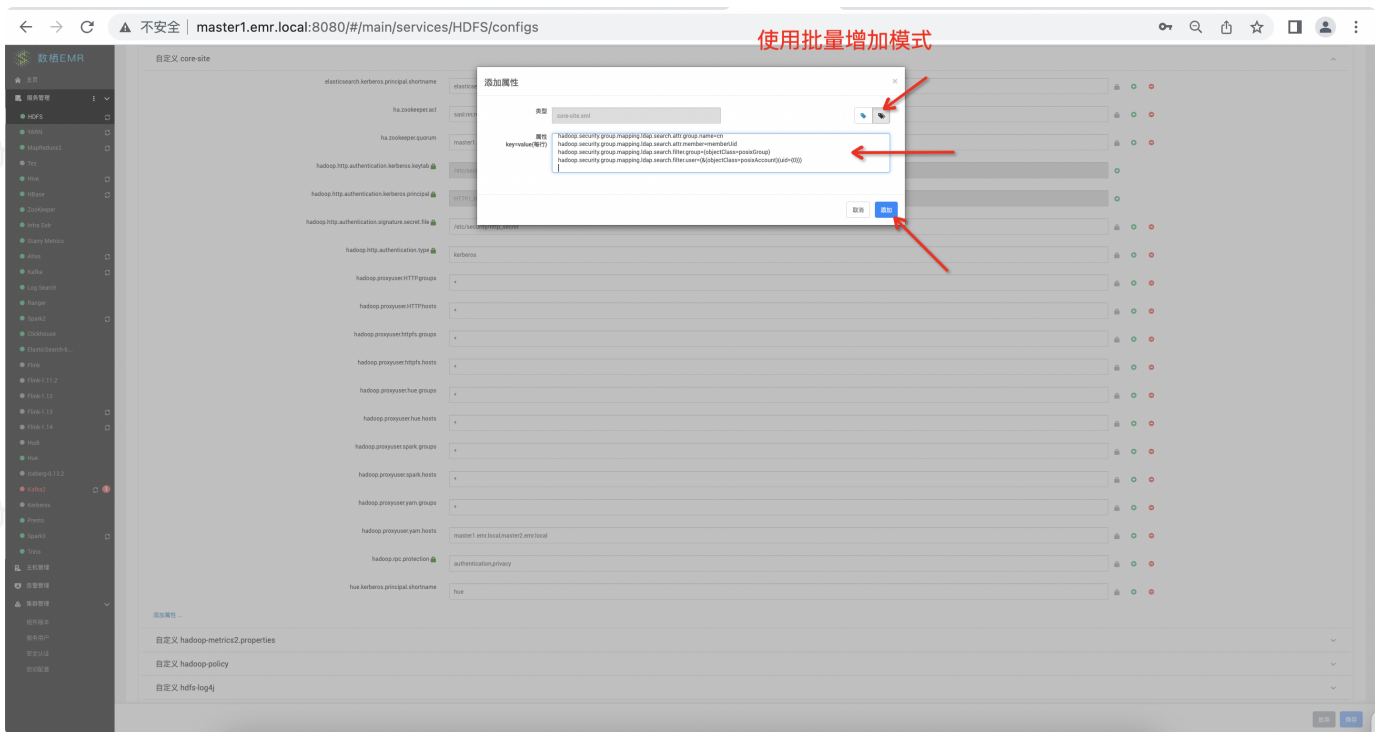
输入用户名admin和密码Shuqi@1298

4.Ranger配置LDAP

hdfs配置修改，自定义core-site.xml中增加以下配置项

注意：修改hadoop.security.group.mapping.ldap.url配置地址

```
▼ Shell | 复制代码
1  hadoop.security.group.mapping=org.apache.hadoop.security.LdapGroupsMapping
2  hadoop.security.group.mapping.ldap.url=ldap://master1.emr.local:389
3  hadoop.security.group.mapping.ldap.bind.password=Ldapadmin@1298
4  hadoop.security.group.mapping.ldap.bind.user=uid=ldapadmin,ou=people,dc=hadoop,dc=com
5  hadoop.security.group.mapping.ldap.baseou=ou=group,dc=hadoop,dc=com
6  hadoop.security.group.mapping.ldap.userbase=ou=people,dc=hadoop,dc=com
7  hadoop.security.group.mapping.ldap.groupbase=ou=group,dc=hadoop,dc=com
8  hadoop.security.group.mapping.ldap.posix.attr.uid.name=uid
9  hadoop.security.group.mapping.ldap.search.attr.group.name=cn
10 hadoop.security.group.mapping.ldap.search.attr.member=memberUid
11 hadoop.security.group.mapping.ldap.search.filter.group=(objectClass=posixGroup)
12 ▼ hadoop.security.group.mapping.ldap.search.filter.user=(&(objectClass=posixAccount)(uid={0}))
13
```

ranger配置common config

注意：修改ldap server地址

Shell | 复制代码

```
1 ldap://master1.emr.local:389
2 uid=ldapadmin,ou=people,dc=hadoop,dc=com
3 Ldapadmin@1298
```

数栖EMR

主页面

服务管理

- HDFS
- YARN
- MapReduce2
- Tez
- Hive
- HBase
- ZooKeeper
- Infra Solr
- Starry Metrics
- Atlas
- Kafka
- Log Search
- Ranger**
- Spark2
- Clickhouse
- ElasticSearch 6...
- Flink
- Flink-1.11.2
- Flink-1.12
- Flink-1.13
- Flink-1.14
- Hudi
- Hue
- Iceberg-0.13.2

Ranger User Info

Enable User Sync
☒ Yes

Sync Source
LDAP/AD

COMMON CONFIGS USER CONFIGS GROUP CONFIGS 4

LDAP/AD URL
ldap://master1.emr.local:389

Bind User
uid=ldapadmin,ou=people,dc=hadoop,dc=com

Bind User Password
root

Incremental Sync
☒ Yes

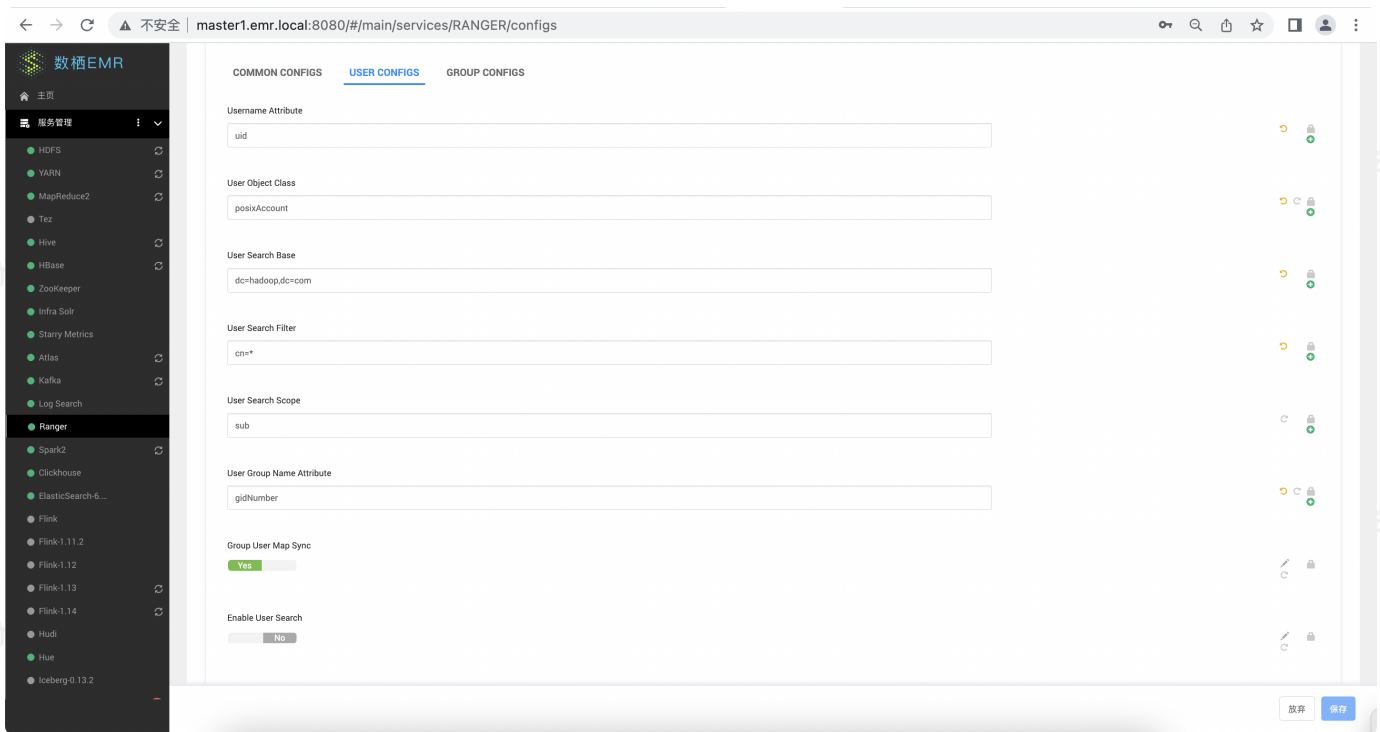
Enable LDAP STARTTLS
☐ No

放弃 保存

ranger配置user config

Shell | 复制代码

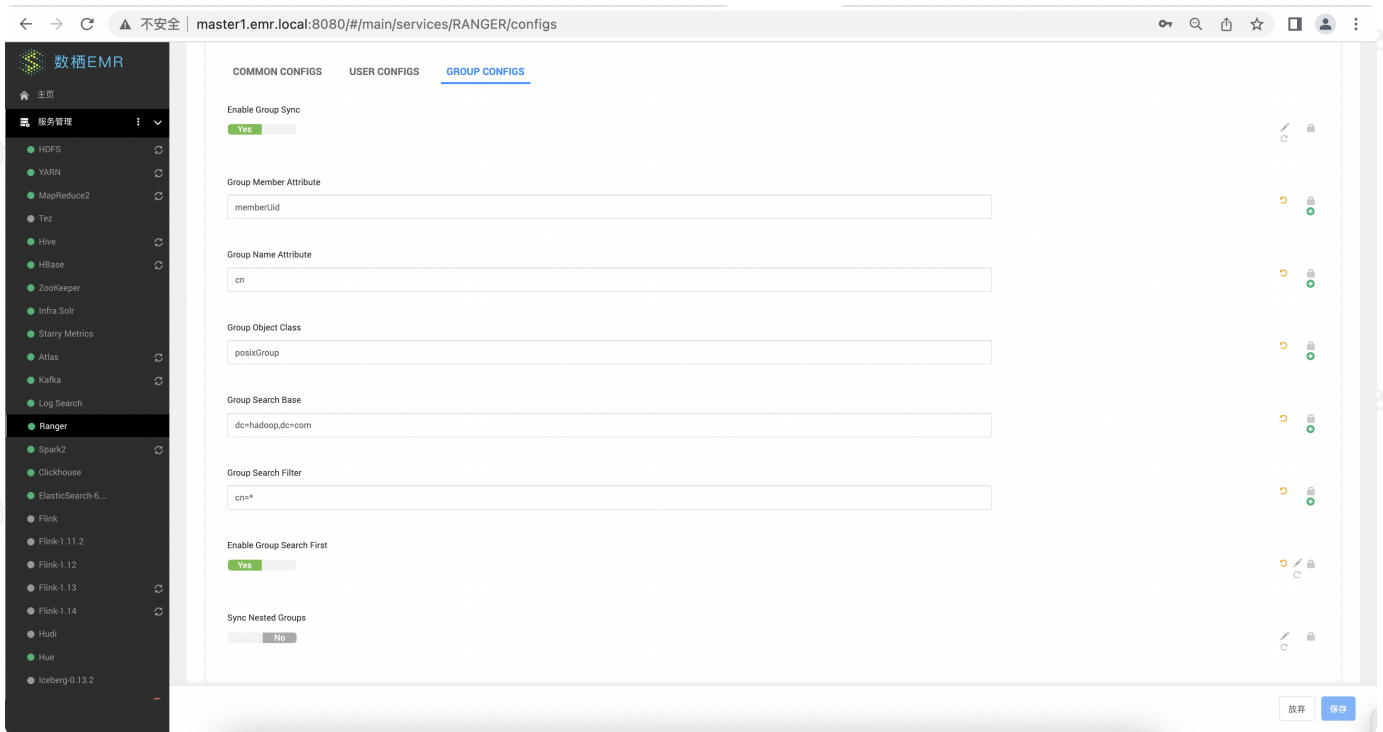
```
1 uid
2 posixAccount
3 dc=hadoop,dc=com
4 cn=*
5 sub
6 gidNumber
```



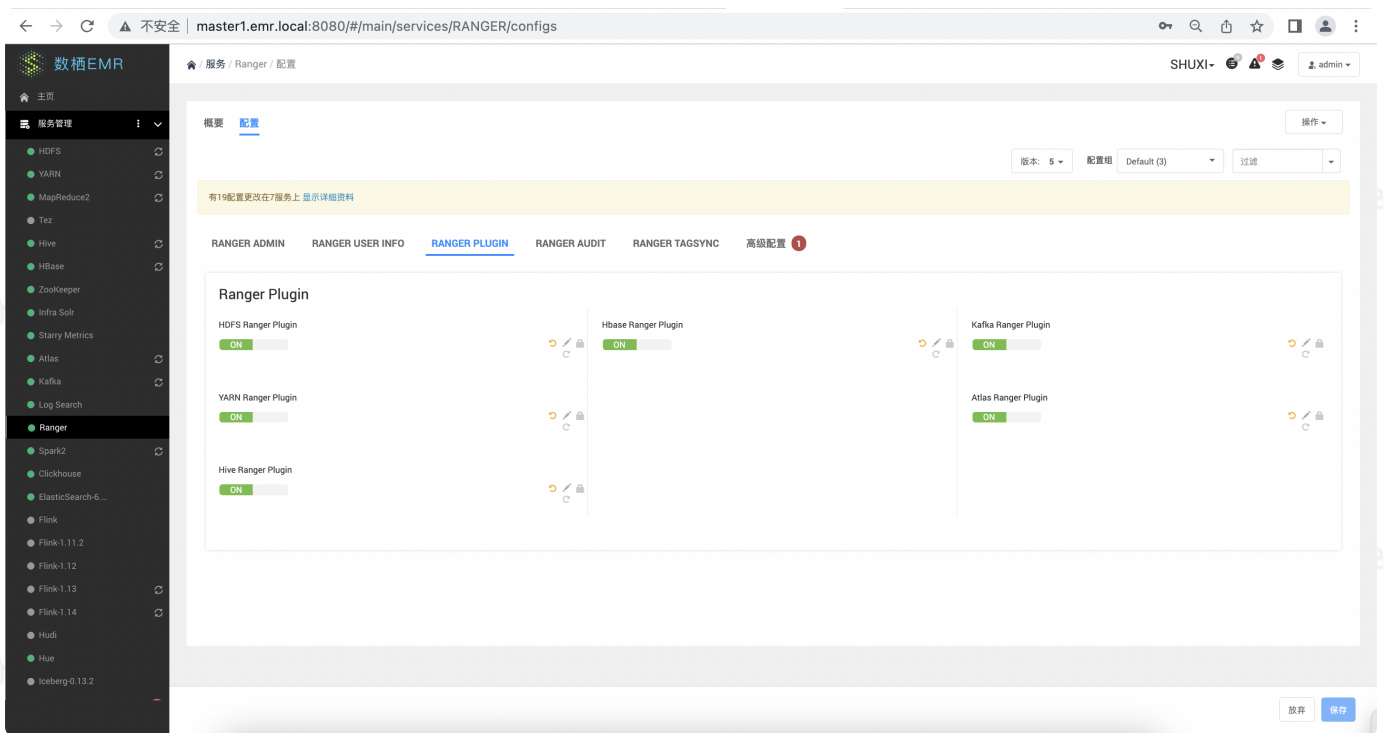
ranger配置group config

▼ Shell 复制代码

```
1 memberUid
2 cn
3 posixGroup
4 dc=hadoop,dc=com
5 cn=*
```



ranger plugins 配置



ranger高级配置

← → ↺ 不安全 | master1.emr.local:8080/#/main/services/RANGER/configs

数栖EMR

服务 / Ranger / 配置

SHUXI

admin

操作

版本: 5 配置组: Default (3) 过滤

有19配置更改在7服务上 显示详细资料

RANGER ADMIN RANGER USER INFO RANGER PLUGIN RANGER AUDIT RANGER TAGSYNC 高级配置

Admin Settings

Ranger Settings

LDAP Settings

LDAP URL

Bind User

Bind User Password

User Search Filter

Group Search Base

Group Search Filter

Ldap Base DN

默认密码: Xin@1298

放弃 保存

ranger配置保存

← → ↺ 不安全 | master1.emr.local:8080/#/main/services/RANGER/configs

数栖EMR

服务 / Ranger / 配置

SHUXI

admin

依赖配置

推荐更改

根据配置更改, Starry推荐以下依赖配置更改
Starry将所有的配置更改更新为推荐值(选中任何配置以保留当前值)

属性	服务	配置组	文件名	原始值	推荐值
☒ ranger.authentication.method	Ranger	Default	ranger-admin-site	UNIX	LDAP
☒ ranger.hdfs.plugin-enabled	HDFS	Default	ranger-hdfs-plugin-properties	No	Yes
☒ REPOSITORY_CONFIG_USERNAME	HDFS	Default	ranger-hdfs-plugin-properties	hadoop	hdfs
☒ dfs.permissions.ContentSummarySubAccess	HDFS	Default	hdfs-site	false	true
☒ dfs.namenode.inode.attributes.provider.class	HDFS	Default	hdfs-site	参数未定义	org.apache.ranger.authorization.hadoop.RangerHdfsAuthorizer
☒ ranger.yarn.plugin-enabled	YARN	Default	ranger-yarn-plugin-properties	No	Yes
☒ yarn.authorization.provider	YARN	Default	yarn-site	参数未定义	org.apache.ranger.authorization.yarn.authoriser.RangerYarnAuthorizer
☒ hive.security.authorization.manager	Hive	Default	hiveserver2-site	org.apache.hadoop.hive.q1.security.authorization.plugin.fallback.FallbackHiveAuthorizerFactory	org.apache.ranger.authorization.hive.authoriser.RangerHiveAuthorizerFactory
☒ hive.server2.enable.doAs	Hive	Default	hive-site	true	false
☒ hive_security.authorization	Hive	Default	hive-env	None	Ranger
☒ ranger.hbase.plugin-enabled	HBase	Default	ranger-hbase-plugin-properties	No	Yes
☒ hbase.coprocessor.master.classes	HBase	Default	hbase-site	org.apache.atlas.hbase.hook.HBaseAtlasCoprocessor,org.apache.hadoop.hbase.security.access.AccessController	org.apache.atlas.hbase.hook.HBaseAtlasCoprocessor,org.apache.ranger.authorization.hbase.RangerAuthorizationCoprocessor

取消 确认

数栖EMR 服务 / Ranger / 配置 SHUXI admin

配置

强烈推荐的配置
请查看以下建议的更改，然后单击属性名称以更改其值。

类型	服务	属性	当前值	描述
警告	Atlas	atlas.graph.storage.hostname	master2.emr.local;node1.emr.local;master1.emr.local	Atlas is configured to use the HBase installed in this cluster. If you would like Atlas to use another HBase instance, please configure this property and HBASE_CONF_DIR variable in atlas-env appropriately.
警告	HDFS	dfs.permissions.enabled	false	dfs.permissions.enabled needs to be set to true if Ranger HDFS Plugin is enabled. If "true", enable permission checking in HDFS. If "false", permission checking is turned off, but all other behavior is unchanged. Switching from one parameter value to the other does not change the mode, owner or group of files or directories. 不建议大于5120MB的值 The maximum allocation for every container request at the RM, in MBs. Memory requests higher than this won't take effect, and will get capped to this value. Consider not using /data partition for storing metrics temporary data. /data partition is already used as hbase.rootdir to store metrics data Temporary directory on the local filesystem. Change this setting to point to a location more permanent than /tmp/ (The /tmp/ directory is often cleared on machine restart). Value is less than the recommended default of 1152 HBase Master Heap Size. In embedded mode, total heap size is sum of master and regionserver heap sizes.
警告	YARN	yarn.scheduler.maximum-allocation-mb	11264	
警告	Starry Metrics	hbase.tmp.dir	/data/lib/starry-metrics-collector/hbase-tmp	
警告	Starry Metrics	hbase_master_heapsize	1024	

取消 继续

此配置保存后修正

数栖EMR 服务 / HDFS / 配置 SHUXI admin

概要 热点图 配置 指标

重启受影响的服务: 11 组件在 3 主机

版本: 31 配置组: Default (3) dfs.permissions.enabled

高级配置

高级 hdfs-site

dfs.permissions.enabled	true
-------------------------	------

放弃 保存