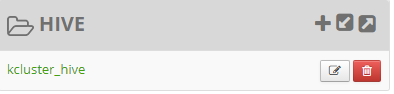
# 概述：

Ranger mastk主要是用来对敏感信息进行保护，不让显示敏感信息。

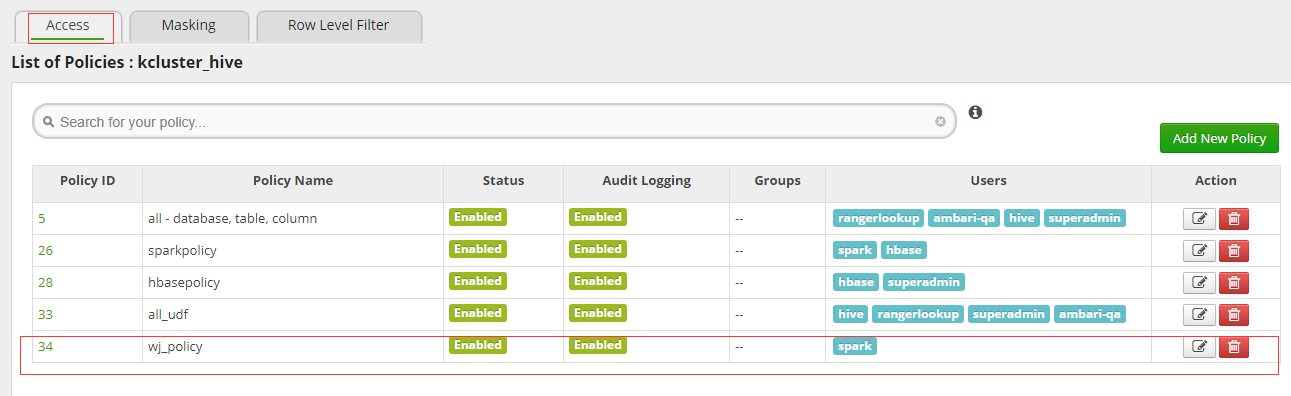
# 操作 mask特性

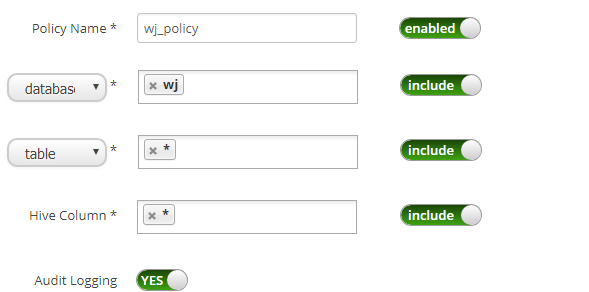
使用spark用户，让spark用户对hive的wj库，t1表的 name字段产生遮罩，没法展现：

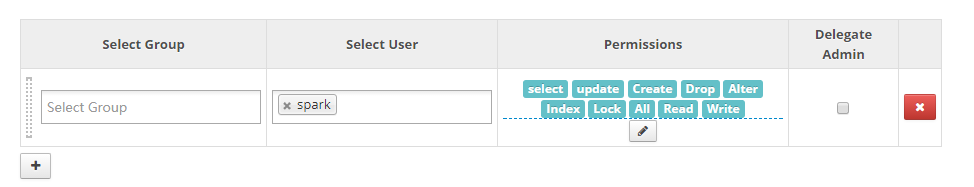
## （1）通过Ambari页面进入Ranger UI管理页面



## 让Spark用户对wj库具有所有权限吧，应该读写权限就行，这里贪省事了





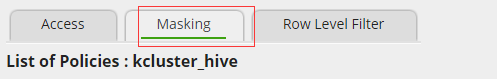


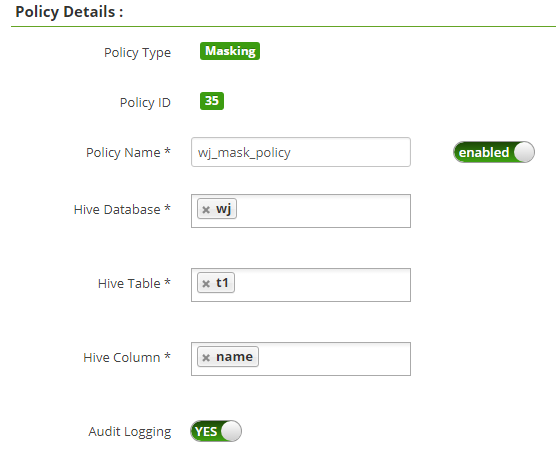
这里缓存spark票据，访问jdbc:hive2://10.110.18.48:10000/default;principal=hive/idap-agent-48.idap.com@IDAP.COM是可以看到wj库下面t1,t2,t3的

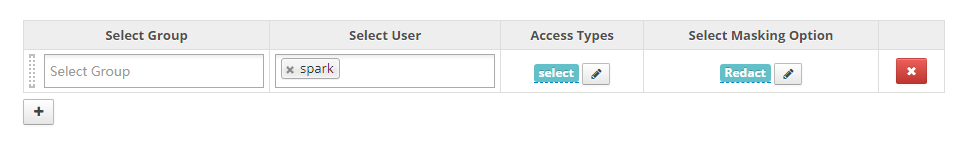




## 然后配置遮罩层

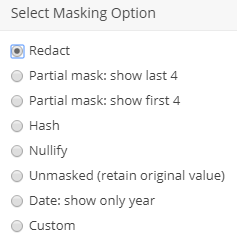






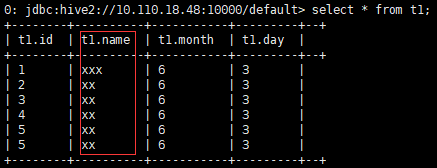
注意：这里有很多遮罩操作，这里我会一一测试：

### 测试Redact



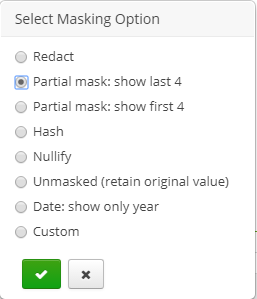
选择Redact，进入10.110.18.48测试环境

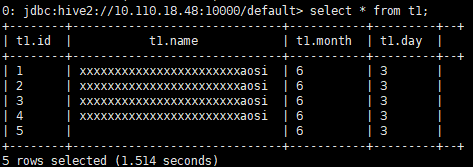
由于我们是对wj库的t1表的name字段遮罩的，我们就查看t1表：



可以看到Redact操作对name字段打码了，我们无法知道name信息。Redact是编辑的意思。

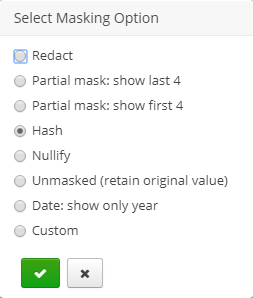
### 测试Partial mask：show last 4

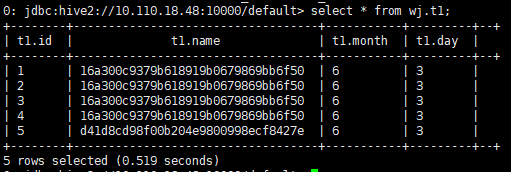




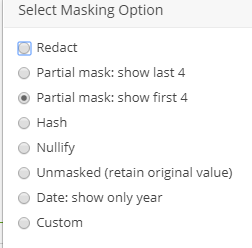
可以看到，它只显示后面四个字节，前面字节打码了。

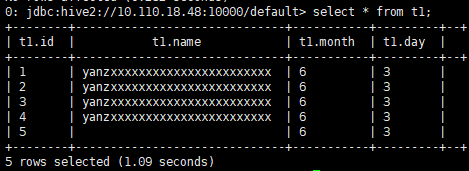
### 测试hash



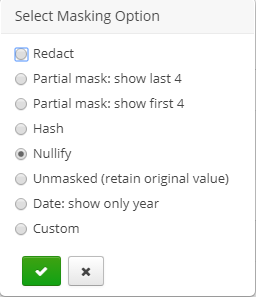


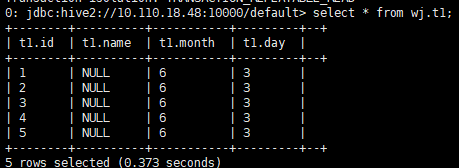
### 测试Partial mask:show first 4



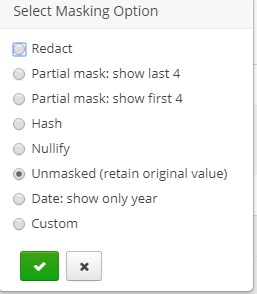


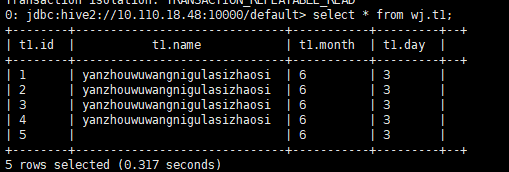
### 测试Nulify





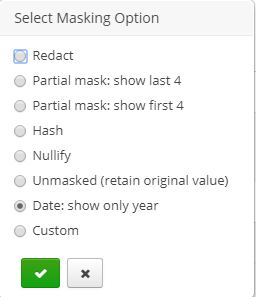
测试Unmasked(retain original value)

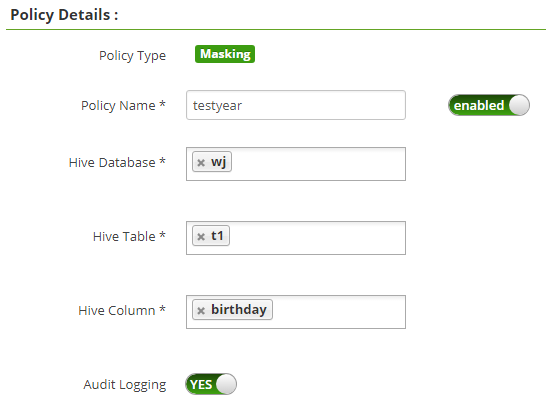




发现一个意外，虽然是不打码了，但是如果name字段包含中文，这里就显示空了。

### 测试date:show only year

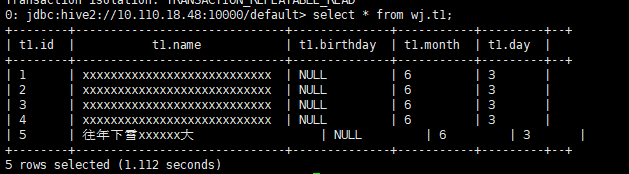




alter table t1 add birthday date;

update t1 set birthday=cast ('1989-11-14' as date) where id=1;

然后切换回spark票据

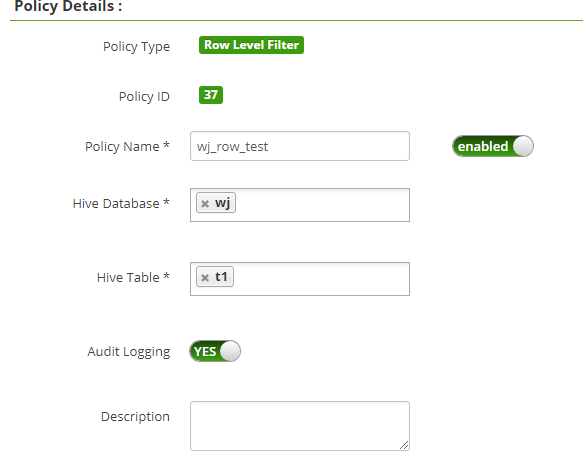


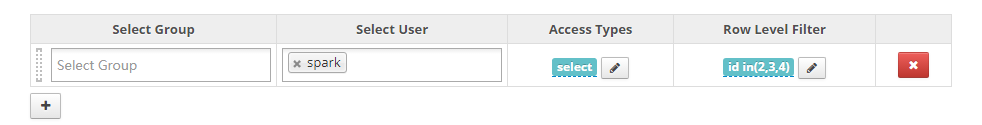
这里特效没有显示出来，birthday应该显示年才对，但所有的都是空

在这个测试里面，name字段采用了Mask.custom特性，name打码了，但还是不支持中文。

# 操作Row level Filter

## (1)添加一个行级别过滤

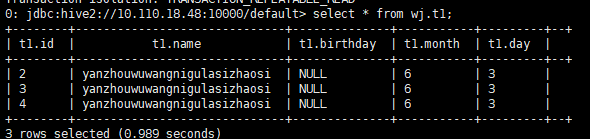




## **(2)切换spark用户票据**

kinit -kt spark.headless.keytab [spark-kcluster@IDAP.COM](mailto:spark-kcluster@IDAP.COM)

## 查看t1数据

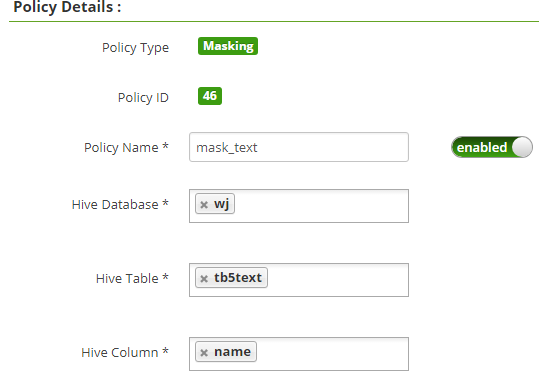


发现只有id=2,3,4的记录才可以查看

# 疑问

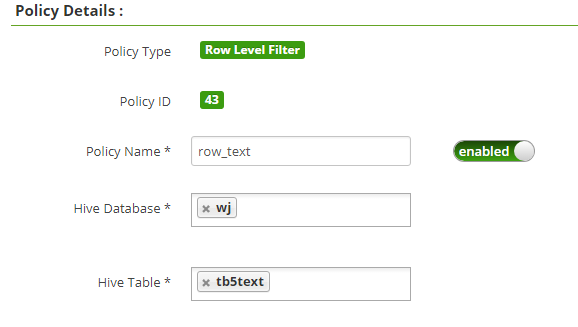
## 1、mask对什么格式的hive表起作用

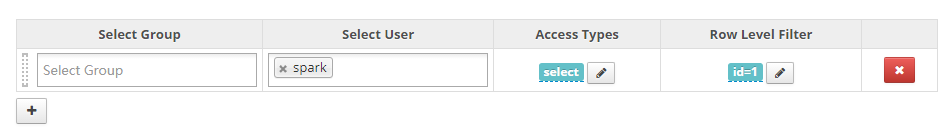
开始验证text格式表格



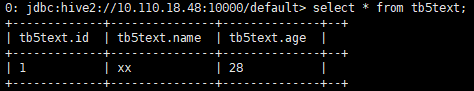


现在对name字段打码了，然后再加行过滤：

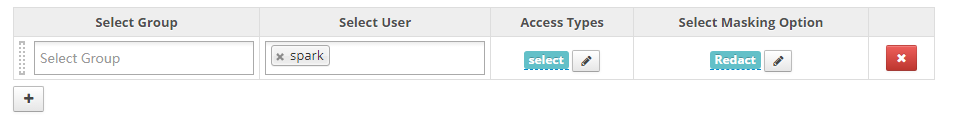
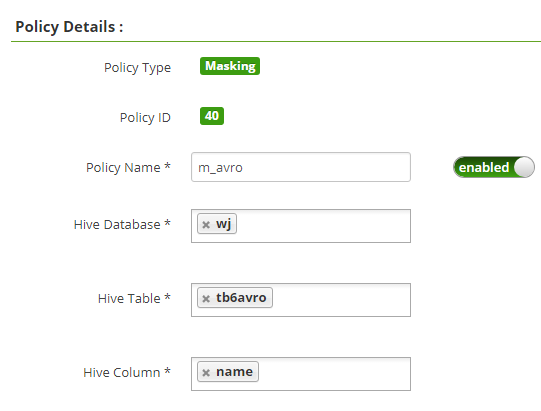




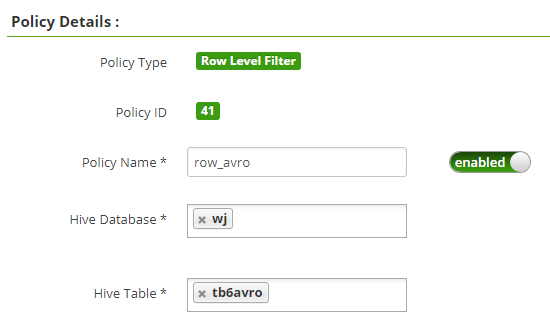
总共tb5text有4条记录，然后我们用text查看：



验证AVRO格式表：

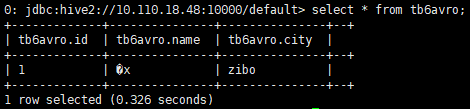


然后看行过滤



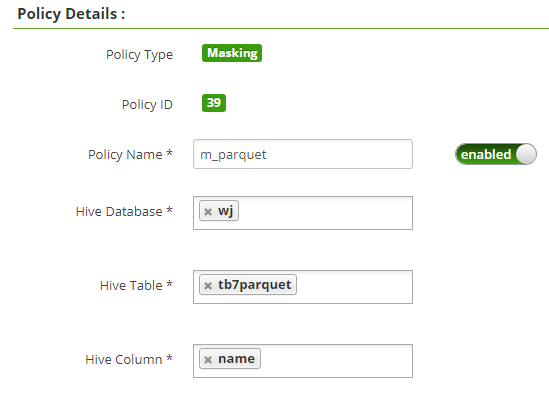


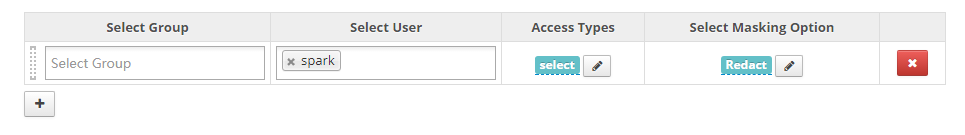
然后现在看看结果：



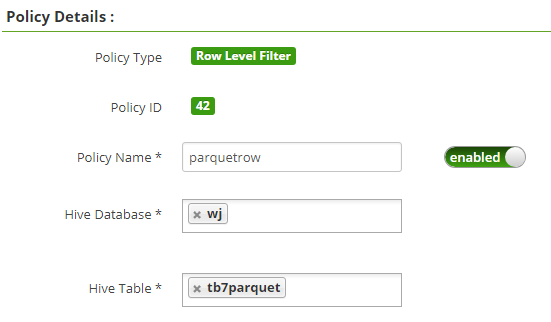
行过滤没有问题，但是遮罩这里name字段有了乱码

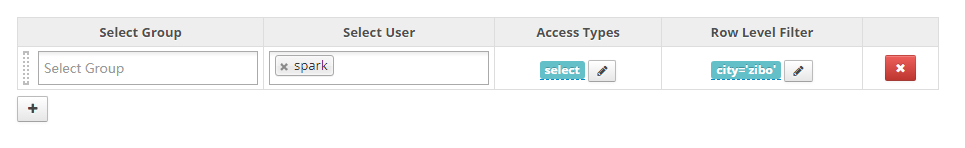
测试parquet格式hive表：



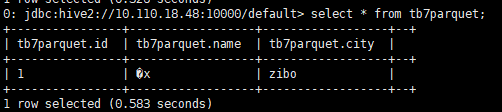


然后看row filter：

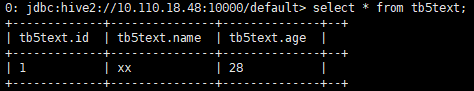


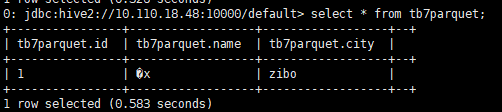


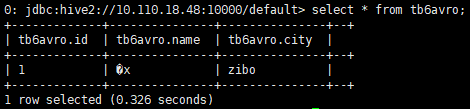
然后看看结果：



## Row lever Filter对什么格式的hive表起作用

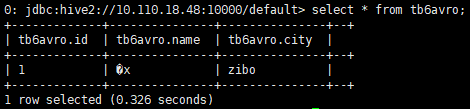


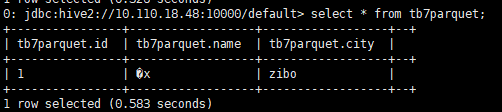




可见row filter 对所有的格式都支持。

## 3、Mask对中文支持不够好？





Name里面包含有汉字，这里显示有乱码。

就算切换到hive票据，这里的name字段也看不到，说明masking 对name 字段有副作用，凡是包含有汉字的，都会破坏数据结构。