

3、部署新组件：Hadoop

- 部署Hadoop服务
 - 1、上传安装包
 - 2、执行脚本导入镜像
 - 3、下载配置文件到本地
 - 4、部署Hadoop服务
 - 5、测试验证Hadoop

Hadoop用于存储挖掘实验的节点中间数据，Hadoop非必选组件，可不安装。
在V95版本中，没有存储节点中间数据，所有对任何节点做修改，都要从头开始执行，这会浪费大量时间。
在V96版本中，存储了节点中间数据，所有对于任何节点，只要它上游节点都执行完成，都可以从该节点执行，不需要从头开始执行，提高了效率。

部署Hadoop服务

1、上传安装包

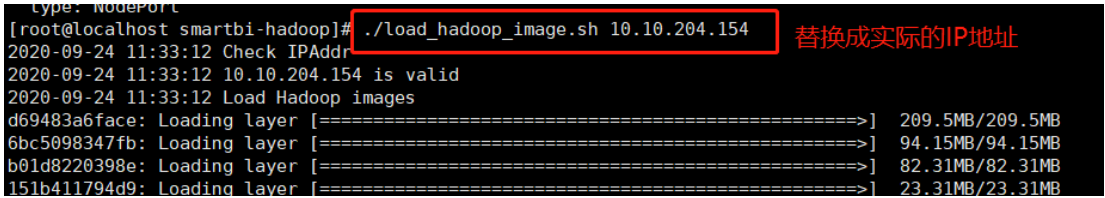
获取并上传安装包到<高性能版本安装目录>/smartbi_data/目录下，并解压。例如：

```
cd /data/Smartbi_All/smartbi_data/      #
tar -zxvf smartbi-hadoop.tar.gz
```

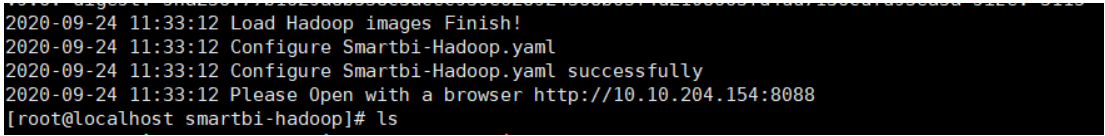
2、执行脚本导入镜像

```
cd /data/Smartbi_All/smartbi_data/smartbi-hadoop
./load_hadoop_image.sh [RancherIP]      #RancherIP
```

参考下图



执行完成如下图：



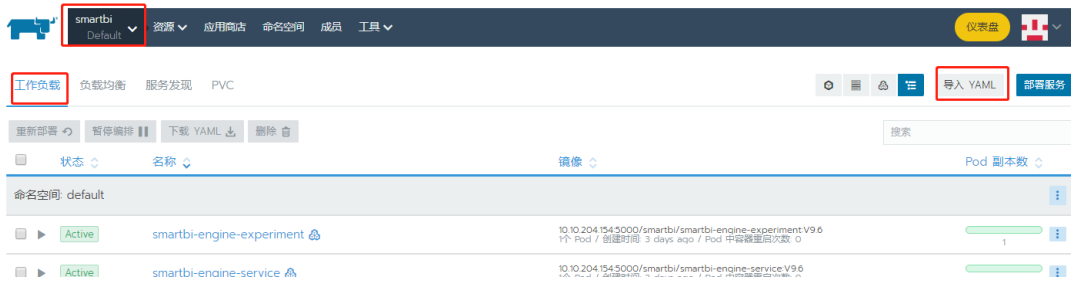
3、下载配置文件到本地

执行完脚本后，使用wincp或者其他工具，将目录中的Smartbi-Hadoop.yaml文件下载到本地。



4、部署Hadoop服务

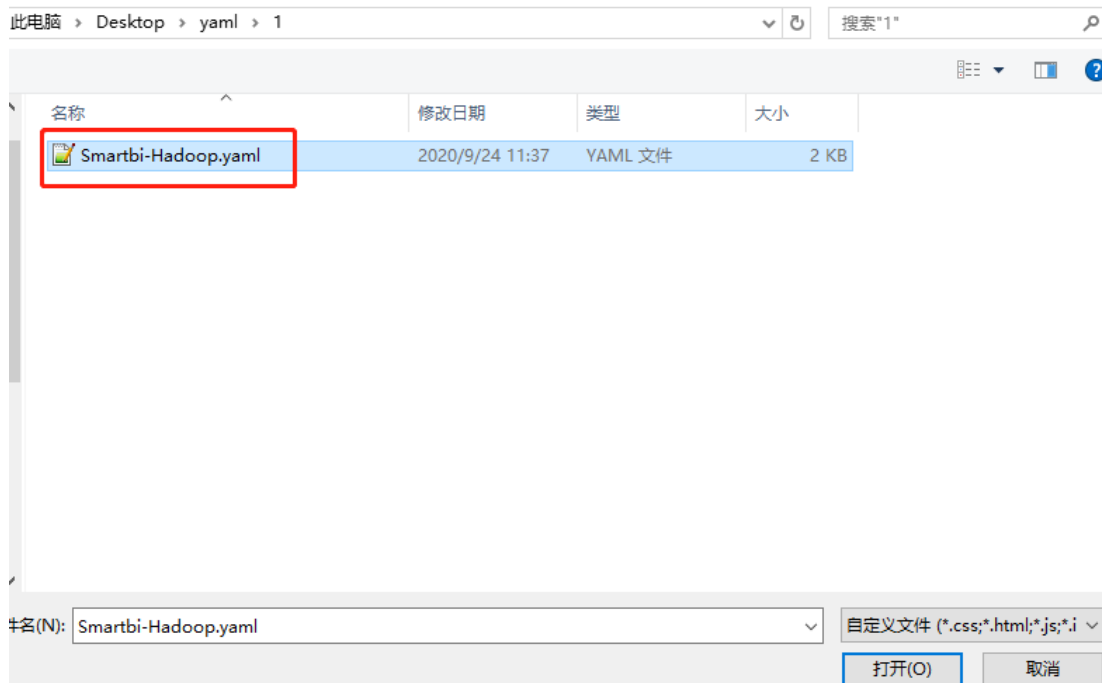
浏览器登陆Rancher控制台，导入Hadoop服务



点击导入YAML，选择从文件读取



选择从服务器中下载的Smartbi-Hadoop.yaml文件



点击导入

导入 YAML

从文件读取

```
1 minikube service
2 apiVersion: v1
3 kind: Service
4 metadata:
5   name: smartbi-hadoop
6   labels:
7     app: smartbi-hadoop
8 spec:
9   ports:
10    - name: hadoop-1
11      port: 9000
12    clusterIP: None
13   selector:
14     app: smartbi-hadoop
15   ---
16
17 apiVersion: apps/v1
18 kind: Deployment
19 metadata:
20   name: smartbi-hadoop
21   namespace: default
22 spec:
23   replicas: 1
24   selector:
```

导入提示

- ① 集群 将资源导入到此集群中
- ② 项目 将资源导入此项目
- ③ 命名空间 将资源导入指定的命名空间

不确定命名空间的资源将被导入到指定的默认命名空间中。
如果您提供了不存在的命名空间，此命名空间将被创建并添加到这个项目中。

默认命名空间 *

default

创建新的命名空间

导入

取消

导入完成后，状态变成绿色Active即表示服务导入成功

工作负载

负载均衡

服务发现

PVC

导入 YAML

部署服务

重新部署

暂停编排

下载 YAML

删除

自

搜索

状态

名称

镜像

Pod 副本数

命名空间: default

Active

smartbi-engine-experiment

10.10.204.154:5000/smartbi/smartbi-engine-experiment V9.6
1个 Pod / 创建时间: 3 days ago / Pod 中容器重启次数: 0

1

Active

smartbi-engine-service

10.10.204.154:5000/smartbi/smartbi-engine-service V9.6
1个 Pod / 创建时间: 3 days ago / Pod 中容器重启次数: 0

1

Active

smartbi-export

10.10.204.154:5000/smartbi/smartbi-export V9.6
32003/tcp

1

Active

smartbi-hadoop

10.10.204.154:5000/smartbi/smartbi-hadoop V9.6
30070/tcp

1

5、测试验证Hadoop

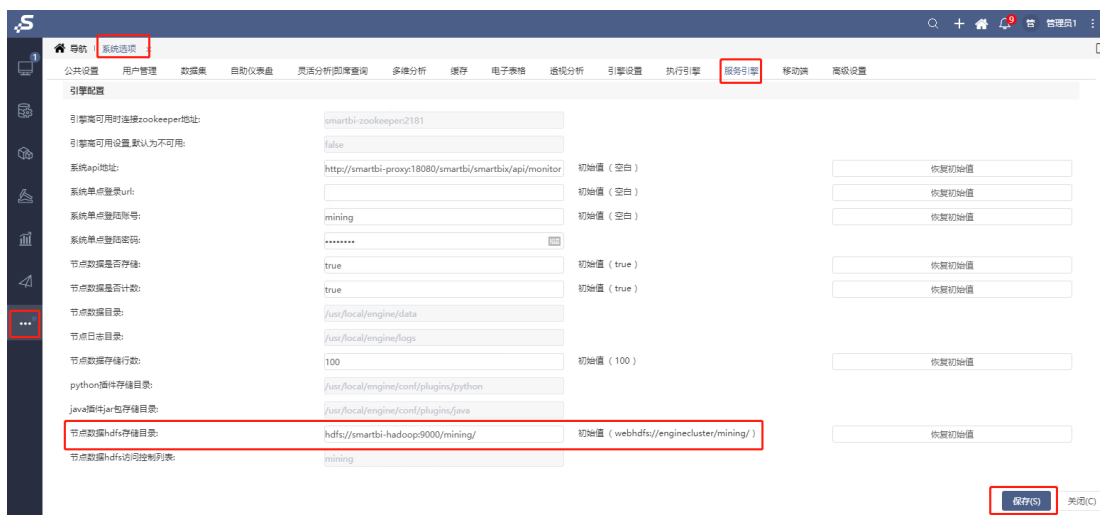
浏览器登陆smartbi，http://IP:31080/smartbi

打开系统运维--系统选项 - 执行引擎 - 引擎配置，节点数据hdfs存储目录设置为hdfs://smartbi-hadoop:9000/mining/，并保存，参考下图

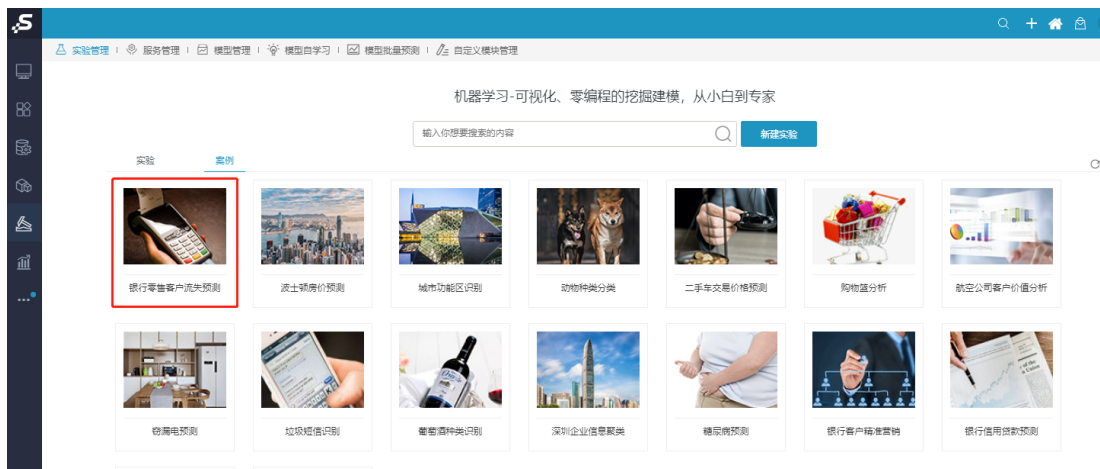
系统API地址:	http://smartbi-proxy:18080/smartbi/smartbi/api/monitor	初始值 (空白)	恢复初始值
系统单点登录url:		初始值 (空白)	恢复初始值
系统单点登录url:	mining	初始值 (空白)	恢复初始值
系统单点登录密码:	*****	初始值 (空白)	恢复初始值
节点数据是否存储:	true	初始值 (true)	恢复初始值
节点数据是否计数:	true	初始值 (true)	恢复初始值
节点数据目录:	/usr/local/smartbi-mining-engine-bin/data		
节点日志目录:	/usr/local/smartbi-mining-engine-bin/logs		
节点数据存储行数:	100	初始值 (100)	恢复初始值
python插件存储目录:	/usr/local/smartbi-mining-engine-bin/conf/plugins/python		
java插件jar存储目录:	/usr/local/smartbi-mining-engine-bin/conf/plugins/java		
节点数据hdfs存储目录:	hdfs://smartbi-hadoop:9000/mining/	初始值 (webhdfs://enginecluster/mining/)	恢复初始值
节点数据hdfs访问控制列表:	mining		

保存(S) 关闭(C)

打开系统运维--系统选项 - 服务引擎，节点数据hdfs存储目录设置为hdfs://smartbi-hadoop:9000/mining/，并保存，参考下图



修改完成后，打开数据挖掘，选择一个案例，运行



运行成功，表示Hadoop部署成功

