HDFS的Web接口

作者：陆建华，刘东旭，李华盛，陈博洋，李建敦

版本： 0.0.1-SNAPSHOT

时间：2016年12月8日

# 开发环境

|  |  |
| --- | --- |
| OS | Fedora release 25 (Twenty Five) |
| 架构 | AMD64 |
| Kernel | kernel-4.8.11-300.fc25.x86\_64 |
| JRE | 1.8.0\_112-release-408-b2 amd64 |
| IntelliJ IDEA | 2016.3 |
| Apache Maven | 3.3.9 |

# 消息队列环境

|  |  |
| --- | --- |
| OS | CentOS Linux release 7.2.1511 (Core) |
| 架构 | AMD64 |
| Kernel | 3.10.0-327.36.2.el7.x86\_64 |
| Docker | 1.10.3 |
| Image | webcenter/activemq:latest(9286304d8105) |

# 使用者

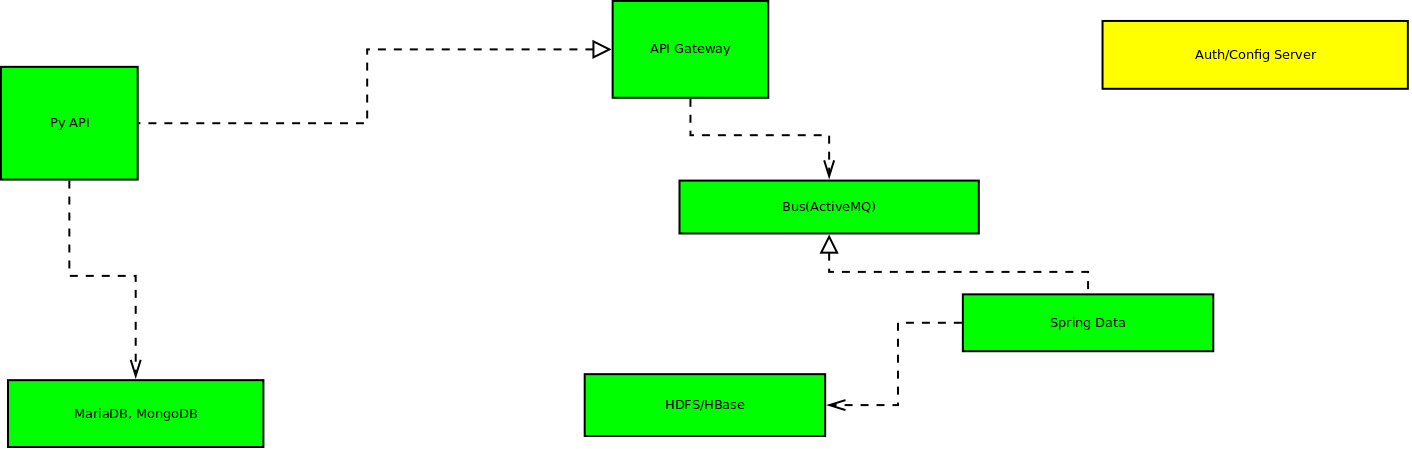
这个服务目前支持向HDFS增加数据，只需要以POST方式发送如下JSON文件即可：

|  |
| --- |
| {  "name": "save", // 操作的名称（目前只至此保存）  "hdfs\_path": "/", // HDFS的目标目录  "options": {  "data.csv": { // 追加写入的目标文件  "1": "row\_data\_in\_line\_1", // 写入的行  "2": "row\_data\_in\_line\_2",  "3": "row\_data\_in\_line\_3"  }  }  } |

各项的数据均为String类型。

# 开发者

## 总体架构



其中网关服务（API Gateway）, 总线（Bus），和Spring Data属于这篇文档。

网关服务向客户端提供RESTFull API

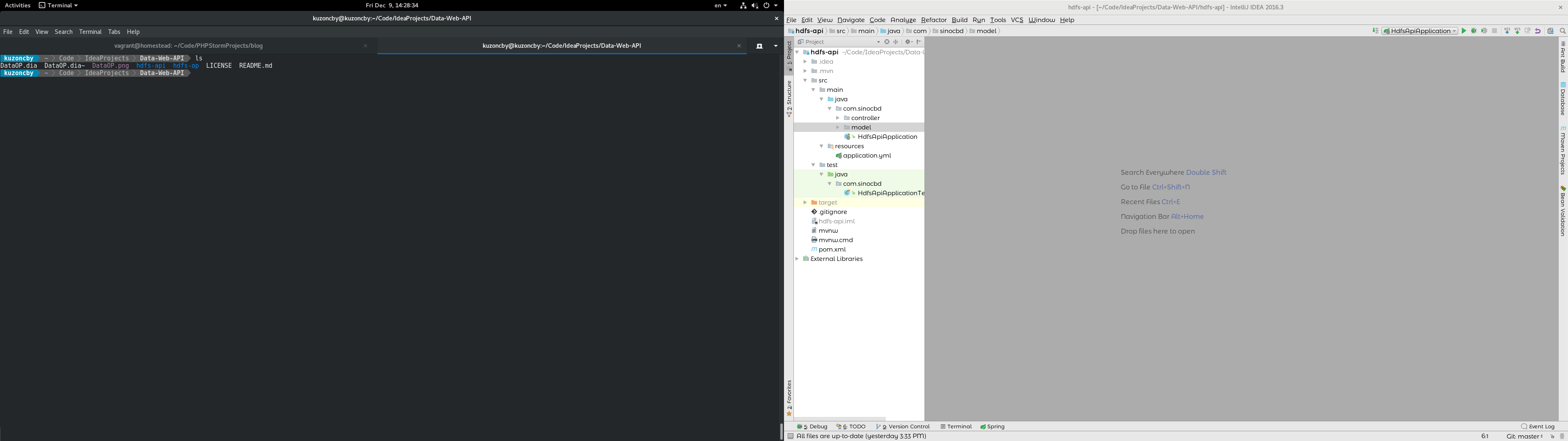
消息总线传递消息

Spring Data操作HDFS

## 起步

请确保安装了第三代的MAVEN

进入项目主目录可以看到两个目录

hdfs-api以及hdfs-op

进入各个项目执行

|  |
| --- |
| $ mvn package |
|  |

构建成功即可开始使用。

## 总线（ActiveMQ消息队列）

这个程序可以起到服务之间的远程通讯，启动消息队列也非常容易。

进入到准备安装消息队列的机器输入：

|  |
| --- |
| docker run --name='activemq' -d \  -e 'ACTIVEMQ\_NAME=amqp-srv1' \  -e 'ACTIVEMQ\_REMOVE\_DEFAULT\_ACCOUNT=true' \  -e 'ACTIVEMQ\_ADMIN\_LOGIN=admin' -e 'ACTIVEMQ\_ADMIN\_PASSWORD=secret' \  -e 'ACTIVEMQ\_WRITE\_LOGIN=producer\_login' -e 'ACTIVEMQ\_WRITE\_PASSWORD=secret' \  -e 'ACTIVEMQ\_READ\_LOGIN=consumer\_login' -e 'ACTIVEMQ\_READ\_PASSWORD=secret' \  -e 'ACTIVEMQ\_JMX\_LOGIN=jmx\_login' -e 'ACTIVEMQ\_JMX\_PASSWORD=secret' \  -e 'ACTIVEMQ\_STATIC\_TOPICS=topic1;topic2;topic3' \  -e 'ACTIVEMQ\_STATIC\_QUEUES=queue1;queue2;queue3' \  -e 'ACTIVEMQ\_MIN\_MEMORY=1024' -e 'ACTIVEMQ\_MAX\_MEMORY=4096' \  -e 'ACTIVEMQ\_ENABLED\_SCHEDULER=true' \  -v /data/activemq:/data/activemq \  -v /var/log/activemq:/var/log/activemq \  -p 8161:8161 \  -p 61616:61616 \  -p 61613:61613 \  webcenter/activemq |
|  |

## HDFS操作（Spring Data）

进入大hdfs-op目录下输入下面的命令启动网关服务

|  |
| --- |
| $ java -jar target/hdfs-op-0.0.1-SNAPSHOT.jar |
|  |

## 网关（API Gateway）

进入大hdfs-api目录下输入下面的命令启动网关服务

|  |
| --- |
| $ java -jar target/hdfs-api-0.0.1-SNAPSHOT.jar |
|  |

## 故障排查及说明

三个服务请按照ActiveMQ,HDFS（Spring Data）, 网关（API Gateway）的顺序启动