UDH SDK开发使用说明

目录

[UDH SDK开发使用说明 1](#_Toc417557582)

[一、概要 1](#_Toc417557583)

[二、配置说明 1](#_Toc417557584)

[2.1 spring-hadoop-context.xml 1](#_Toc417557585)

[2.2 common.properties 2](#_Toc417557586)

[2.3 hadoop.properties 2](#_Toc417557587)

[三、实例程序说明 3](#_Toc417557588)

[3.1 HDFS访问（HDFSExample.java） 3](#_Toc417557589)

[3.2 HBase访问（HBaseExample.java） 4](#_Toc417557590)

# 一、概要

为方便基于UDH服务进行二次开发，AE产品部提供基于UDH的SDK开发包，这里就如何部署、使用做下说明。

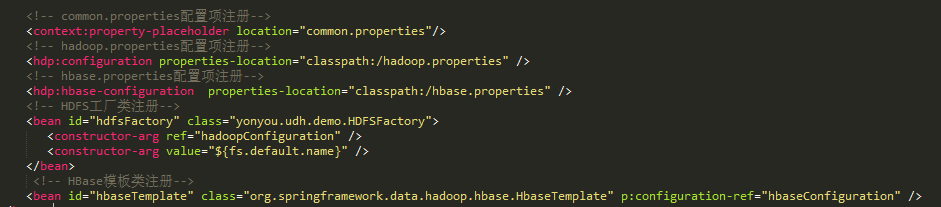
UDH SDK由Spring data for hadoop组件和UDH相关组件程序包组成，提供基于Java的快速开发工具。目前各组件使用版本情况为：spring-data-2.0.4、cdh-5.0.4，SDK版本会随UDH版本升级而升级。

# 二、配置说明

SDK主要有四个配置文件，依次为：spring-hadoop-context.xml、common.properties、hadoop.properties、hbase.properties。

## 2.1 spring-hadoop-context.xml

项目的基础配置文件，其他三个配置文件会在这里注册。



其中hdfsFactory和hbaseTemplate分别是hdfs和hbase的访问入口对象，这里注册为spring bean对象，在应用程序中根据这个bean id可以拿到相应的访问类实例。

## 2.2 common.properties

主要配置一些全局参数，根据实际情况进行配置，比如：



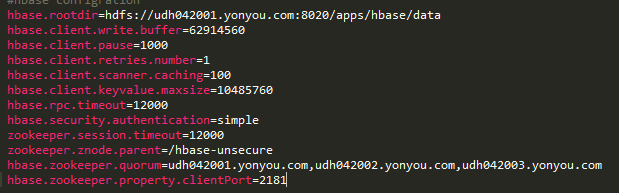
*如果要访问hdfs，该参数为必配项。*

## 2.3 hadoop.properties

主要配置一些hadoop的一些公共参数，根据实际情况进行配置，比如：



hbase.properties配置hbase的相关参数，根据实际情况进行配置，比如：



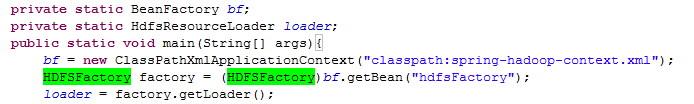
*其中hbase.rootdir，zookeeper.znode.parent，hbase.zookeeper.quorum，hbase.zookeeper.property.clientPort为必配项，其他配置可做适当调整。*

# 三、实例程序说明

## 3.1 HDFS访问（HDFSExample.java）

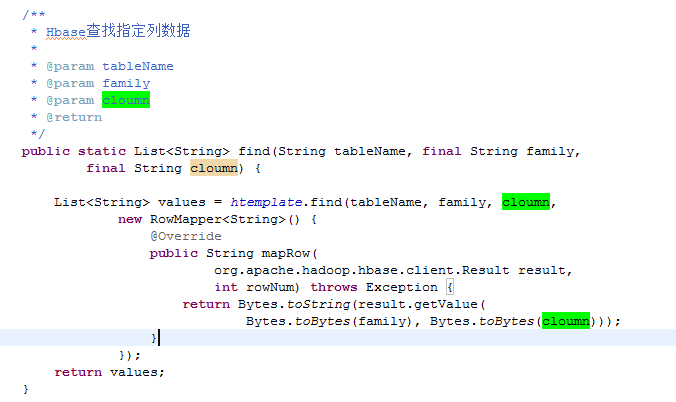


其中loader是hdfs的资源加载器，通过它可以进行hdfs文件及目录的读写操作，它有hdfs的工厂类HDFSFactory构造，具体如下：

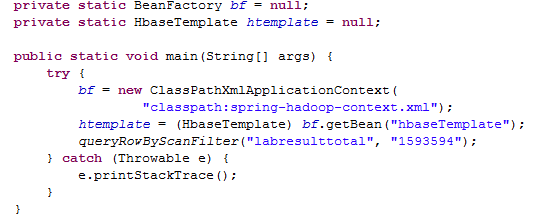


HDFS批量读取文件及目录、写操作在HDFSExample.java中都有说明和实例，这里就不再赘述。

## 3.2 HBase访问（HBaseExample.java）



htemplate是HBase的访问入口，通过它可以进行HBase的读写操作，它的获取方式如下：



在HBaseExample.java中还有提供基于过滤条件查询、全量查询、写数据的示例代码，这里也不再赘述，这里强调一点，HBase的读写操作尽量都通过HbaseTemplate进行，不推荐自己构建HTable，如果使用HTable，必须注意复用及资源及时释放。

如果需要进行DDL操作，则需要构建HBaseAdmin对象，如下：

