

# 关系型数据源-Hadoop-hive数据连接

## 1 概述

Hive是基于Hadoop的一个数据仓库工具，可以将结构化的数据文件映射为一张数据库表，并提供简单的sql查询功能，可以将sql语句转换为MapReduce任务进行运行。Hive是建立在 Hadoop 上的数据仓库基础构架。它提供了一系列的工具，可以用来进行数据提取转化加载（ETL），这是一种可以存储、查询和分析存储在 Hadoop 中的大规模数据的机制。

本文介绍如何在smartbi V9.5版本中连接Hadoop-hive数据库。

## 2 数据连接

### 2.1 配置信息

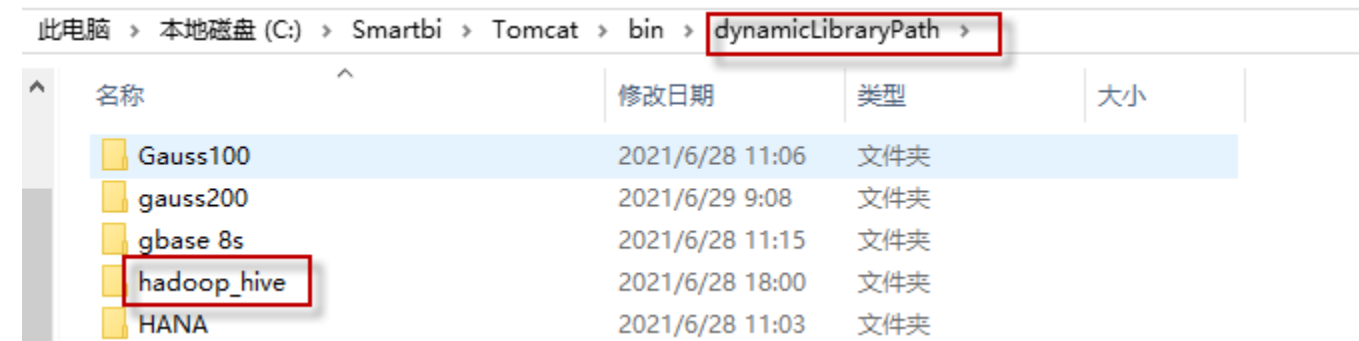
驱动程序类	连接字符串	驱动程序存放目录	支持数据库版本
org.apache.hive.jdbc.HiveDriver	jdbc:hive2://<servername>:10000/default	自定义	hadoop 2.7.2,hive 2.0.0

1) 连接字符串主要信息说明：

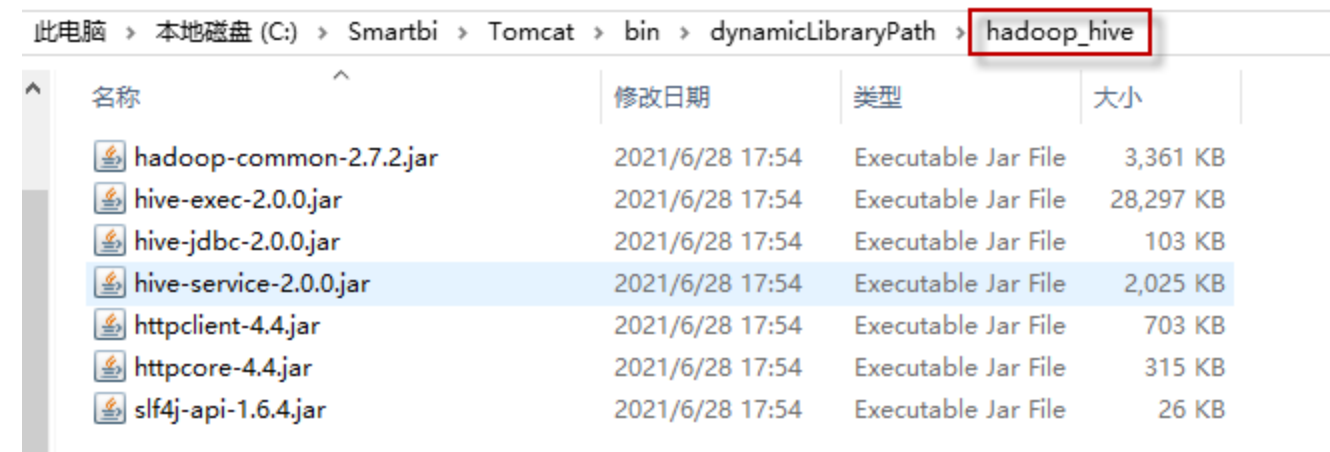
<servername>：数据库的地址，一般可以向数据库管理员获取；  
10000：数据库端口。

2) 添加自定义驱动方法：

a、首先进入到<Smartbi\_PATH>\Tomcat\bin\dynamicLibraryPath文件夹下，手动添加hadoop\_hive文件夹。



b、然后进入hadoop\_hive文件夹下，将驱动放到该文件夹下（驱动包可以从数据库官网获取或向数据库管理员获取）。



c、最后在数据连接时需要点击自定义选项，选择hadoop\_hive。

驱动程序存放目录

产品内置

自定义

hadoop\_hive

驱动程序类\*

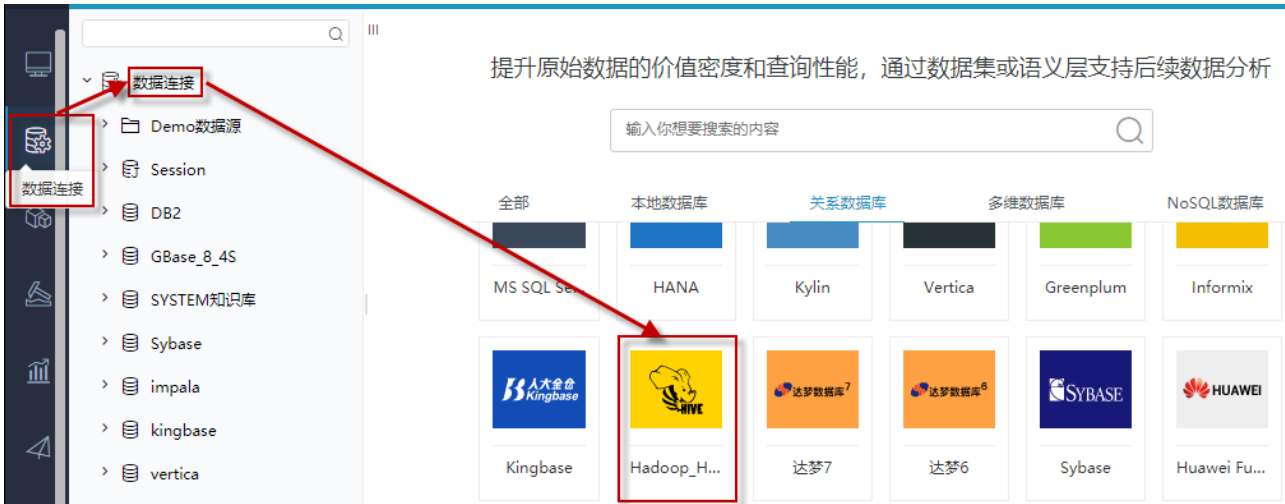
org.apache.hive.jdbc.HiveDriver

连接字符串\*

jdbc:hive2://<servername>:10000/default

2.2 连接步骤

1) 登录Smartbi企业报表分析平台，选择 [数据连接>关系数据库](#) ，点击 [Hadoop\\_Hive](#) 图标进入数据库连接界面。如下图所示：



2) 根据 [2.1章节](#) 提供的配置信息，输入数据库对应的信息。如下图所示：

新建关系数据源

名称\*

hadoop\_hive

别名

驱动程序类型\*

Hadoop\_Hive

驱动程序存放目录

产品内置

自定义

hadoop\_hive

驱动程序类\*

org.apache.hive.jdbc.HiveDriver

连接字符串\*

jdbc:hive2://10.10.204.130:10000/default

Thrift URI

验证类型

静态

动态

用户名

root

密码

\*\*\*\*\*

测试连接(T)

保存(S)

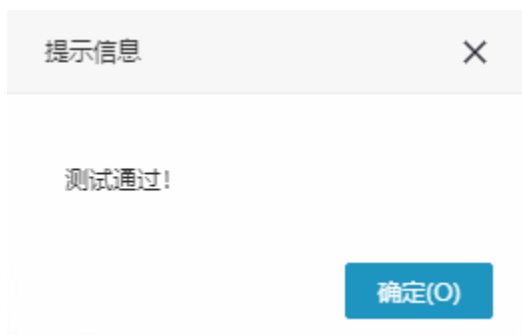
关闭(C)

说明：

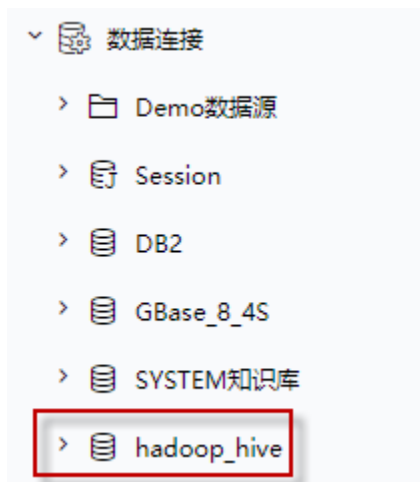
- a、名称是数据连接名称，用户可以自定义，不可为空。
- b、用户名和密码是连接字符串中配置的数据库的连接用户名和密码，可为空，一般可以联系数据库管理员获取。

2.3 测试连接

1) 信息正确输入后，点击 [测试连接](#) ，若出现如下图的 [测试通过](#) 弹出框，则表示可以成功连接上数据库。如下图所示：



2) 测试连接成功后，点击右下角的 **保存** 按钮，选择数据源的保存位置，保存成功后，该数据连接即添加成功。如下图所示：



3) 数据库连接成功后，可以参考[快速新建透视分析](#)、[快速新建即席查询](#)、[快速新建电子表格](#)、等[快速入门](#)使用方式查看数据。

 smartbi 支持连接 Hadoop-hive集群。