

高速缓存库

在Smartbi中高速缓存库主要是为了加速数据查询和分析。


- [支持范围](#)
- [配置步骤](#)
- [各数据库的连接详情](#)

支持范围

高速缓存库支持下面几种数据库：


- GaussDB（DWS）
- Presto+Hive
- SelectDB
- SmartbiMpp，广州思迈特软件有限公司研发的高速缓存库
- SmartbiMppMD，广州思迈特软件有限公司研发的高速缓存库
- Vertica
- 星环
- StarRocks

采购高速缓存MPP模块后，可以根据实际情况选择以上几种数据库中的一种作为高速缓存库。

-  1、高速缓存库支持将表抽取至ClickHouse多副本集群，抽取方式与之前抽取至分布式集群相同，选定分区字段即可。

配置步骤

高速缓存库在Smartbi中的配置方法如下：

- （1）启动服务器，在浏览器输入Smartbi地址，进行登录；
- （2）输入用户名密码，登录平台；
- （3）在“系统导航栏”选择 **数据连接**，在“资源目录区”的高速缓存库更多操作，选择 **打开**，或双击 **高速缓存库**；



(4) 进入“高速缓存库”界面

高速缓存库

名称*

SmartbiCache

别名

高速缓存库

驱动程序类型*

SmartbiMpp

选择驱动程序

☒ 产品内置
 ☐ 自定义

驱动程序类*

smartbi.jdbc.SmartbiMppCH

连接字符串*

jdbc:smartbimppch://10.10.110.33:31824/default?socket_timeout=1000000

登录验证方式*

用户名密码

验证类型

☒ 静态
 ☐ 动态

用户名

default

密码

高级 >

测试连接(T)

保存(S)

关闭(C)

(5) 据实际使用的数据库，修改相应连接属性，连接属性详情请参见 [各数据库的连接详情](#)；

(6) 点击保存，保存配置。

各数据库的连接详情

数据库类型	配置说明
GaussDB (DWS)	<div> <div> <div>高速缓存库</div> <div> <div>名称*</div> <div>SmartbiCache</div> </div> <div> <div>别名</div> <div>高速缓存库</div> </div> <div> <div>驱动程序类型*</div> <div>GaussDB(DWS)</div> </div> <div> <div>选择驱动程序</div> <div> <input checked="" type="radio"/> 产品内置 <input type="radio"/> 自定义 </div> </div> <div> <div>驱动程序类*</div> <div>com.huawei.gauss200.jdbc.Driver</div> </div> <div> <div>连接字符串*</div> <div>jdbc:gaussdb://<servername>:8000/<database></div> </div> <div> <div>登录验证方式*</div> <div> <div>用户名密码</div> <div>验证类型</div> <div> <input checked="" type="radio"/> 静态 <input type="radio"/> 动态 </div> </div> </div> <div> <div>用户名</div> <div>smartbi</div> </div> <div> <div>密码</div> <div>*****</div> </div> <div>高级 ></div> <div> <div>测试连接(T)</div> <div>保存(S)</div> <div>关闭(C)</div> </div> </div> <div> <div>1、切换到GaussDB(DWS)</div> <div>2、配置GaussDB(DWS)的连接信息，根据实际情况填写IP、端口、数据库名称</div> <div>3、GaussDB(DWS)的用户名</div> <div>4、GaussDB(DWS)的密码</div> </div> </div>
Presto+Hive	

名称*	SmartbiCache
别名	高速缓存库
驱动程序类型*	Presto+Hive
驱动程序存放目录	<input type="radio"/> 产品内置 (产品未内置驱动) <input checked="" type="radio"/> 自定义 HADOOP_HIVE
链接方式*	用户名密码
Hive 连接字符串: *	jdbc:hive2://smartbi-hadoop:10000/default 1、Hive的连接地址、ip、端口、数据库
Hive 用户名: *	root 2、Hive的用户名
Hive 密码: *	**** 3、Hive的密码

导入数据到HDFS服务器的配置, 要求该服务器上已经安装HDFS环境

服务器地址: *	smartbi-hadoop 4、具有ssh权限的hdfs服务器地址, 如果默认ssh端口不是22, 需要服务器ip: ssh端口
服务器用户(有 HDFS 权限): *	root 5、ssh的用户名
服务器密码: *	***** 6、ssh的密码
SSH 验证类型 (允许空值):	
服务器临时目录: *	/opt 7、抽数时的临时目录, 需要ssh用户具有可读写的权限
HDFS临时目录: *	/tmp 8、抽数时hdfs的临时目录, 需要Hive用户具有此目录的读写权限
Thrift URL: *	thrift://smartbi-hadoop:9083 9、Thrift URL, 在进行跨库查询时需要用到

说明:

如果要提高访问效率, 需要提供HDFS的配置文件: core-site.xml和hdfs-site.xml。
将两个配置文件复制到Presto所在机器后, 修改Presto服务器中的配置文件<PrestoDir>/etc/catalog/smartbi/smartbicache.properties
添加一个键值hive.config.resources=<配置文件的路径>/core-site.xml,<配置文件的路径>/hdfs-site.xml



presto使用注意

- 1、Presto执行多表关联查询时，会把多表拆分成多个执行计划执行，这样会影响查询效率。因此Presto不适合执行的多表关联场景有：跨
- 2、Presto的关联机制：Presto会默认执行广播式的JOIN操作，会将左表拆分到几个工作节点上，然后发送整个右表分别到已拆分好的处理出工作节点的内存限制，进而出错。因此，需要把数据量大的表放在左表才能保证查询正常。
- 3、presto+hive作为高速缓存库时，若hive设置了用户名密码，那么presto的配置中就需使用https证书（证书需根据服务器生成），因此添加https证书的映射参数

跨库联合数据源

名称* SmartbiUnionDB

别名 跨库联合数据源

选择驱动程序 ☒ 产品内置 ☐ 自定义

连接字符串* jdbc:smartbi:uniondb://t-10-10-35-15:8443?SSL=true&SSLKeystorepath=D:/Tomcat/TOMCAT-V95/Presto/ads.keys

登录验证方式* 用户名密码 验证类型 ☒ 静态 ☐ 动态

用户名 root

密码

高级 >

测试连接(T) 保存(S) 关闭(C)

- 4、当高速缓存库驱动程序类型选择了Presto+Hive，系统默认生成的别名是smartbicahe.shive，包括了非法字符“.”，因此在查看高速缓存可。

资源属性

节点ID: SCHEME.SmartbiCache.smartbicahe.shive

名称: smartbicahe.shive

别名: smartbicahe.shive

路径: 数据连接\高速缓存库\smartbicahe.shive

类型名: 模式

描述:

创建人: 系统服务

创建时间: 2019-08-06 18:28:38

图标: 选择

确定(O) 取消(C)

Sele
ctDB

高速缓存库

名称* SmartbiCache

别名 高速缓存库

驱动程序类型* SelectDB 1、切换到SelectDB

选择驱动程序 ☒ 产品内置 ☐ 自定义

驱动程序类* com.mysql.jdbc.Driver 2、配置SelectDB的连接信息，根据实际情况填写IP、端口、数据库名称

连接字符串* jdbc:mysql://<servername>:9030/<database>?useOldAliasMetadataBehavior=true&useUnicode=true&characterEnc

登录验证方式* 用户名密码 验证类型 ☒ 静态 ☐ 动态

用户名 smartbi 3、SelectDB的用户名

密码 ***** 4、SelectDB的密码

高级 >

测试连接(T) 保存(S) 关闭(C)

Smar
tbiM
pp

高速缓存库

名称* SmartbiCache

别名 高速缓存库

驱动程序类型* SmartbiMpp 1、切换到SmartbiMpp

驱动程序存放目录 ☒ 产品内置 ☐ 自定义

驱动程序类* smartbijdbc.SmartbiMppCH

连接字符串* jdbc:smartbimppch://10.10.204.33:8123/default 2、配置SmartbiMpp的连接信息，根据实际情况填写服务器ip、端口、数据库名称

链接方式* 用户名密码 验证类型 ☒ 静态 ☐ 动态

用户名 default 3、SmartbiMpp数据库用户名

密码 ***** 4、SmartbiMpp数据库密码

高级 >

测试连接(T) 保存(S) 关闭(C)



在windows环境下使用SmartbiMppMD作为缓存库，在Linux环境下使用 SmartbiMpp作为缓存库。

Smar
tbiM
ppMD

高速缓存库

名称* SmartbiCache

别名 高速缓存库

驱动程序类型* SmartbiMppMD 1、切换到SmartbiMPPMD

驱动程序存放目录 ☒ 产品内置 ☐ 自定义

驱动程序类* smartbijdbc.SmartbiMppMD

连接字符串* jdbc:smartbimppmd://10.10.35.147:50000/smartbimppmd 2、配置SmartbiMPPMD的连接信息，根据实际情况填写服务器IP、端口、数据库名称

链接方式* 用户名密码 验证类型 ☒ 静态 ☐ 动态

用户名 monetdb 3、SmartbiMPPMD数据库用户名

密码 ***** 4、SmartbiMPPMD数据库密码

高级 >

测试连接(T) 保存(S) 关闭(C)

Vertica

高速缓存库

名称*

SmartbiCache

别名

高速缓存库

驱动程序类型*

Vertica

1、选择Vertica

驱动程序存放目录

☒ 产品内置 ☐ 自定义

驱动程序类*

com.vertica.jdbc.Driver

连接字符串*

jdbc:vertica://smartbi-vertica:5433/smartbi

2、按照实际情况填写Vertica的ip、端口和数据库等

链接方式*

用户名密码

验证类型 ☒ 静态 ☐ 动态

用户名

dbadmin

3、Vertica的用户名

密码

4、Vertica的密码

高级 >

测试连接(T)

保存(S)

关闭(C)

星环

名称*

SmartbiCache

别名

高速缓存库

驱动程序类型*

星环

1、选择星环

驱动程序存放目录

☐ 产品内置 (产品来内置驱动) ☒ 自定义 xinghuan

驱动程序类*

org.apache.hive.jdbc.HiveDriver

连接字符串*

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

2、按照实际情况填写星环的ip和端口

链接方式*

用户名密码

验证类型 ☒ 静态 ☐ 动态

用户名

root

3、星环数据库的用户名

密码

4、星环数据库的密码

导入数据到HDFS服务器的配置，要求该服务器上已经安装HDFS环境

5、部署星环客户端的服务器IP地址，如果ssh端口不是22，需要指定端口信息，用冒号隔开，192.168.4.59:22

服务器地址：*

192.168.4.59

服务器用户(有 HDFS 权限)：*

root

6、ssh到部署星环客户端的用户名

服务器密码：*

7、ssh到部署星环客户端的密码

SSH 验证类型 (允许空值)：

服务器临时目录：*

/opt

8、部署星环客户端的服务器上的抽数临时目录，ssh的用户必须要有权限读写

HDFS临时目录：*

/tmp

9、星环hdfs上导出数据临时目录，需要连接星环数据库的用户具有读写权限

高级 >

星环服务器配置详情请参见 [星环客户端配置](#) 内容。

添加JDBC驱动详情请参见 [Smartbi添加星环JDBC驱动](#) 内容。



如果MPP是星环数据库支持分区分桶。

如何开启？需要在系统选项/高级设置中开启：DATAMODEL_EXTRACT_OPEN_PARTITION_SETTING，如果是true，则是开启；false则是关闭。

（具体的使用方法需要参考星环本身的分区分桶）

高速缓存库

名称*

SmartbiCache

别名

高速缓存库

驱动程序类型*

StarRocks

①

选择驱动程序

☒ 产品内置 ☐ 自定义

②

驱动程序类*

com.mysql.jdbc.Driver

连接字符串*

jdbc:mysql://10.10.35.119:9030/ha?useOldAliasMetadataBehavior=true&useUnicode=true&characterEncoding=UTF

登录验证方式*

用户名密码

验证类型 ☒ 静态 ☐ 动态

用户名

root

③

密码

高级 >

测试连接(T)

保存(S)

关闭(C)

**StarRocks作为MPP:**

- 1、适配了社区版2.5。
- 2、导入数据时，应该注意：

- 列的内容不能包含换行符，存在则会删除换行符。
- 列的内容不能包含tab建，因为tab建是作为列分隔符。
- 导入时的数据库表名建议是英文字母、数字，不能包含中文和特殊字符。

- 3、在数据模型 生成日期表，生成到高速缓存库，目前会用时比较久。