AnySQL 简介

## 摘要

目前，大部分的实时流系统的业务开发都需要直接调用流系统的API，而这些API一般都只提供基础性的功能，基于这些API进行业务开发还需要写比较多的代码来实现业务逻辑，并且需要业务开发人员对流系统本身也有一定的熟悉程度。为了降低实时流系统的业务开发难度，提高实时流系统的业务开发效率，作者在基于spark-streaming的流系统AnyStream上设计了一种SQL的变种，即AnySQL，利用AnySQL，业务开发人员只需对流系统有基本的了解并熟悉SQL，便可以进行流系统的业务开发，从而极大的方便流上的业务实现。

AnySQL是在Spark-SQL上的一种增强，Spark-SQL上使用的是Hive的解析引擎。因此，在语法结构上，AnySQL非常类似于Hive SQL，除了一些增强的功能。因此，在熟悉AnySQL前，请先认真阅读HiveSQL的语法手册[Hive wiki](https://cwiki.apache.org/confluence/display/Hive/Home;jsessionid=899B53AB7EF6F0323929CA7DF1F9CE39)，以下只对AnySQL在Spark-SQL(即Hive)上的增强做详细数目。

## 语法说明

数据帧(内存表)的定义：

*Identifier* **:=** *SELECT …*

*导出数据到metaQ：*

*导出数据到外部数据库*

*更新外部数据库*

*设置内存表的分区数目*

## 注意事项

1.要极力避免中间表的数据写入基于磁盘的文件系统