Druid虚拟表设计

背景：

druid存储数据是按照预先定义的各个维度，按照一定的时间粒度进行聚合。因此，时间粒度越粗，维度总基数越少，聚合效果越明显，聚合后的数据量越少。业务上需要的维度总数据比较多，多达20个维度，而并不总是要同时用到这20个维度，常用的查询一次仅用到几个维度。所有的查询查这一张表，会让查询响应时间比较久。

**方案**：

根据大表衍生出很多小表，每个小表针对某一类查询。同时每类表都会按照聚合的时间粒度进行聚合汇总。

**具体设计要求：**

1.配置文件摄入配置。（配置文件的内容可以是一个json对象，或者是一个包含以下json对象的json数据）

|  |  |  |
| --- | --- | --- |
| 字段名称 | 类型 | 说明 |
| vds | String | 虚拟的表名（sql查询的依据） |
| bds | String | 基本的表（衍生该ds数据） |
| ds | String | 表名称 |
| dimensions | List | 元素是string类型的，表示该ds聚合的维度集合 |
| metrics | List | 元素是string类型的，表示该ds中metric的集合 |
| granularity | string | 该ds的时间粒度。取值：day,hour,month |

进阶：1.metrics中的元素可以支持string类型以外的json对象，进行更细致化的metric定制需求。