华为云 DDM 实践作业

通过本次实践作业，学会如何使用华为云 DDM 进行数据表拆分。

选择拆分算法

拆分算法即将逻辑表中数据拆分到多个数据库分片上的算法，DDM 支持 hash 和 range 两种拆分算法。

- **hash**

将数据均匀分布在各个分片。

适用于 SQL 查询条件使用“=”、“IN”之类运算符相对较多的场景。

- **range**

将数据表内的记录按照算法原数据的取值范围进行分片存储。

适用于 SQL 查询条件使用“>”、“<”、“BETWEEN ... AND ...”之类运算符相对较多的场景。

拆分算法的使用，需要结合业务查询场景进行评估，以选择最佳拆分算法，提升 DDM 效率。

参照下表：

拆分算法	hash		range	
拆分键	表字段	表字段+日期函数	表字段	表字段+日期函数
详细说明	根据指定的表字段将数据平均拆分到各个分片上。	根据指定的表字段+日期函数将数据平均拆分到各个分片上。 表字段必须是日期类型（date、datetime、timestamp）	将数据表内的记录按照算法元数据定义的规则将数据拆分到指定的分片上。	根据指定的表字段+日期函数将数据按照算法元数据的规则将数据拆分到各个分片上。 表字段必须是日期类型（date、datetime、timestamp）
适用场景	适用于需要将数据均匀分布的场景，例如：银行类客户业务应用，业务逻辑主体是客户，可使用客户对应的表	需要按时间（年、月、日、周及其组合）对数据进行拆分的场景，例如：游戏类的应用，可使用玩家对应的表字段（例	适合范围类操作较多的场景，例如：电商类应用，如果业务场景是围绕商家做活动进行，业务逻辑主体是活动日期，可使用	例如日志分析场景，日志系统中可能包含各类复杂的信息，这时您可以选择时间字段作为拆分键，然后对拆分键使用日期函数拆分。

	字段（例如客户号）作为拆分键，详情参见如下示例。	如玩家注册时间）作为拆分键，按日、月、年等函数分片，方便统计和查询某日、月玩家的操作数据，帮助游戏厂家做大数据分析。	活动日期对应的表字段（例如活动名称、日期范围）作为拆分键，方便统计某周期内销量等情况。	为了方便日志清理和转储，采用 range 拆分算法，对时间字段用日期函数转换成年，表示按年存储到各个分片上，详情参见如下示例。

选择拆分键

拆分键是在水平拆分逻辑表的过程中，用于生成路由结果的表字段，指定表字段后，可以进一步选择日期函数，也可以手动输入“日期函数(字段名)”，数据表字段必须是日期类型（date、datetime、timestamp），日期函数适用于需要按时间（年、月、日、周及其组合）对数据进行拆分的场景。

DDM 根据拆分键与拆分算法计算路由结果，对分片表的数据进行自动水平拆分，分发到数据分片中。

选取拆分算法与拆分键一般遵循以下规则：

- 尽可能使数据均匀分布到各个分片上。
- 该拆分键是最频繁或者最重要的查询条件。
- 优选主键作为拆分键，因为主键作为查询条件时，查询速度最快。

明确业务场景下拆分方法

分片表的数据量一般都达到千万级，因此选择合适的拆分算法和拆分键非常重要。如果能找到业务主体，并且确定绝大部分的数据库操作都是围绕这个主体的数据进行的，那么可以选择这个主体所对应的表字段作为拆分键，进行水平拆分。

业务逻辑主体与实际应用场景相关，下列场景都有明确的业务逻辑主体。

1. 银行类客户业务应用，业务逻辑主体是客户，可使用客户对应的表字段（例如客户号）作为拆分键。部分业务系统的业务场景为围绕银行卡/账号的，可以选取卡/账号作为拆分键。

2. 电商类应用，如果业务场景是围绕商品进行操作，业务逻辑主体是商品，可使用商品对应的表字段（例如商品编码）作为拆分键。

3. 游戏类的应用，主要围绕玩家数据进行操作，业务逻辑主体是玩家，可使用玩家对应的表字段（例如玩家 id）作为拆分键。

以银行类客户业务为例，建表 SQL 语句如下：

```
CREATE TABLE PERSONALACCOUNT (
ACCOUNT VARCHAR(20) NOT NULL PRIMARY KEY,
NAME VARCHAR(60) NOT NULL,
TYPE VARCHAR(10) NOT NULL,
AVAILABLEBALANCE DECIMAL(18,2) NOT NULL,
STATUS CHAR(1) NOT NULL,
CARDNO VARCHAR(24) NOT NULL,
CUSTOMID VARCHAR(15) NOT NULL
) ENGINE=INNODB DEFAULT CHARSET=UTF8;
```

无明确业务场景下拆分方法

如果业务场景中找不到合适的主体，也可以选择那些数据分布较为均匀的属性所对应的表字段作为拆分键。

例如日志分析场景，日志系统中可能包含各类复杂的信息，这时您可以选择时间字段作为拆分键。

选择时间字段作为拆分键时，支持对拆分键使用日期函数拆分。为了方便清理和转储，采用 range 拆分算法，对时间字段用日期函数转换成年，表示按年存储到各个分片上。

建表 SQL 语句如下：

```
CREATE TABLE LOG (
LOGTIME DATETIME NOT NULL,
LOGSOURCESYSTEM VARCHAR(100),
```

```
LOGDETAIL VARCHAR(10000)
);
```

了解更多信息，请访问 [分布式数据库中间件 DDM 主页]