**多维引擎缓存的使用说明**

作者：霍琦

目录

[1. 多维缓存的意义 2](#_Toc425874850)

[1.1. 多维分析 2](#_Toc425874851)

[1.2. 意义 2](#_Toc425874852)

[2. 参数说明 3](#_Toc425874853)

[2.1. 刷新时间 3](#_Toc425874854)

[2.2. 替换策略 3](#_Toc425874855)

[2.3. 内存占比阈值 4](#_Toc425874856)

[2.4. 保存参数 5](#_Toc425874857)

[2.5. 强制卸载缓存 5](#_Toc425874858)

[3. 缓存实体分类 6](#_Toc425874859)

[3.1. 多维定义缓存 6](#_Toc425874860)

[3.2. 维度成员缓存 7](#_Toc425874861)

[3.3. SQL缓存 7](#_Toc425874862)

[3.4. 聚合对象缓存 7](#_Toc425874863)

# 多维缓存的意义

## 多维分析

从多个角度对事实发生的数据做分析，称为多维分析。其中涉及到的模型包括立方体、维度、指标等，从报表查询开始的执行流程基本是这样的，实施人员需要对业务报表有清晰的理解，也要掌握生产库中的业务数据结构，对生产数据进行ETL，生成可用于分析的数据到数据仓库中，建立多维分析模型，将模型置于报表中的相应位置，做多维分析。

## 意义

根据多维分析的定义，可以看到一次多维分析中涉及到的实体包括，立方体维度等分析模型，指标聚合值等数据，由于要和数据库打交道，还要包括数据库查询模型等。对于业务比较复杂的分析，一次分析涉及到的实体数量会很多，执行时间也会随着业务的复杂程度而变化。而在多维分析的应用场景中，数据仓库的数据以及分析模型的变化频率很低，很多情况基本都是以星期、月这样的时间单位在变化。此时缓存的意义就体现出来，对于变化不是很频繁的数据、模型，分析引擎会将其缓存在服务器上，以便于多次查询分析中使用，这样做可以明显提升分析效率，提高用户对产品的体验性，并且可以更充分地利用服务器资源。

# 参数说明

首先来介绍一下缓存管理中使用到的几个参数。



## 刷新时间

刷新时间指缓存池的刷新时间，即缓存的失效时间。这个参数以秒为单位。对于基本不会变的缓存实体，可以把时间设置的相对长一些。

## 替换策略



缓存中的数据不是全部都有意义的，当缓存实体占满缓存池，并且又有新的实体需要缓存时，应该对缓存池中的一些对象实施卸载策略，也称为缓存替换策略。替换策略分为两种，最近最少使用和先进先出。

最近最少使用：当需要缓存池需要卸载对象时，选择那些最近使用频率最低的缓存实体，将其移出缓存池。

先进先出：缓存池中的实体按照排队方式进入到缓存池，当缓存池需要卸载对象时，选择最先进入缓存池，排在队伍头的缓存实体，将其移出缓存池。

## 内存占比阈值



内存占比阈值，指当此应用程序（不只是分析引擎）使用的内存占分配给JAVA虚拟机的最大内存的比例达到一定程度后，会启动缓存卸载程序，直到占比满足设置的参数。比如，当内存占比阈值设置为60%时，如果使用内存占JAVA虚拟机的最大内存的比例达到60%时，会卸载缓存池中的某些缓存实体。

此处需要说明的是，由于虚拟机是动态地分配内存及垃圾回收，所以设置的比例参数不一定能准确地反映当前内存使用情况，只是一个大概的比例设置。

## 保存参数



修改缓存参数后，点击【保存】，使缓存设置更新。

## 强制卸载缓存

因为某些原因，希望缓存立刻失效，可以点击【强制卸载】，立刻卸除缓存池中的所有缓存实体。

# 缓存实体分类

如下图所示，多维分析涉及到的缓存实体包括：多维定义缓存，维度成员缓存、SQL缓存、聚合对象缓存，接下来会一一介绍。

 各个缓存实体之间是有父子关系的，即最上层的实体被卸载后，属于其的子实体也会被卸载出缓存池，具体的父子关系会在每个实体中说明。



## 多维定义缓存

多维定义指对当前分析涉及到的多维模型的统称，该实体属于所有缓存实体中的父类，如果多维定义缓存被卸除，属于其的维度成员、SQL对象、聚合对象等缓存实体也会被卸除。

多维定义缓存可以不启用，设置位置在分析建模工具中所属的立方体模型，如下：



在分析建模工具中修改多维模型，或者在缓存管理中强制卸载，会卸除多维定义缓存。

## 维度成员缓存

一次查询分析涉及到的立方体可能包含多个维度，维度中又包含多个成员，在报表展现、计算聚合值等过程中都会加载维度成员，维度成员缓存就是避免多次加载。

维度成员缓存一定会启用，多维定义缓存被卸除后，或者在缓存管理中强制卸载，会卸除维度成员缓存。

## SQL缓存

事实表或维表有可能是某一张具体的物理表，也有可能是语义模型，而语义模型最终在取数时表现的是一段SQL语句，对于解析出的元定义SQL模型，分析引擎会将它缓存起来，以便于重用。

因为SQL对应的是事实表或维表的元定义，所以在元定义中可以设置SQL缓存不启用。



如果SQL缓存被启用，多维定义缓存被卸除后，或者在缓存管理中强制卸载，会卸除SQL缓存。

## 聚合对象缓存

聚合对象指的是某个指标对于某些维度成员的聚合值，比如，{[门店].[北京],[时间].[2015年],[Measures].[销售量]}，北京的门店在2015年的销售量就是一个聚合值，也是聚合对象。聚合值有可能在几次查询分析中被多次使用，所以需要缓存。

聚合对象缓存可以不启用，设置位置在分析建模工具中所属的立方体模型，如下。



如果聚合对象缓存被启用，多维定义缓存被卸除后，或者在缓存管理中强制卸载，会卸除聚合对象缓存。