**对把语义模型当作大横表使用的场景做性能评估**

作者：霍琦

目录

[1. 背景 2](#_Toc431213078)

[2. 说明点 2](#_Toc431213079)

# 背景

在有些分析场景下，用户希望把事实表和维表的数据放在一起，存在于一张物理表中，这种表被称为大横表。在我们的分析平台上，大横表可以包括语义模型、数据仓库元数据，也可以是物理表。这样做的好处是简化数据仓库设计，简化数据模型，对用户屏蔽在数据上的多维概念，方便ETL操作。

但是这样做会在多维分析中，造成性能瓶颈。举个例子，如果某个语义模型作为大横表被立方体使用，立方体中的维度和指标都会使用这个语义模型，在解析维度成员、计算聚合值时都会执行这个语义模型的取数。如果这个语义模型的执行时间很长或数据量很大，那么很明显，较长的执行时间会随着使用维度的增加而增加，并最终造成性能问题。

由于使用的语义模型无法确定数据量，并且语义模型可能会包含参数，执行的结果集会随着参数的变化而变化，所以暂时不考虑缓存语义模型数据。

目前的方案是，在建立分析模型时，对模型做校验，当立方体和维度共用一个语义模型时，如果语义模型的执行时间大于配置中指定的最大时间，或结果集数量大于配置中指定的最大数量，会对用户提示。同时，在性能监控开启时，对这种模型的执行会记录下日志，方便调查性能问题。

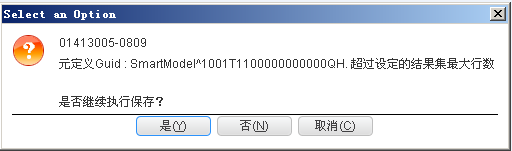
# 说明点

#### 分析模型评估

在设计分析模型时，如下这种情况会造成指标和维度共用一个大横表，即同一个语义模型。



在这种情况下，如果语义模型的执行时间大于配置文件中的最大时间，或者结果集数量大于配置文件中的最大数量，会提示。

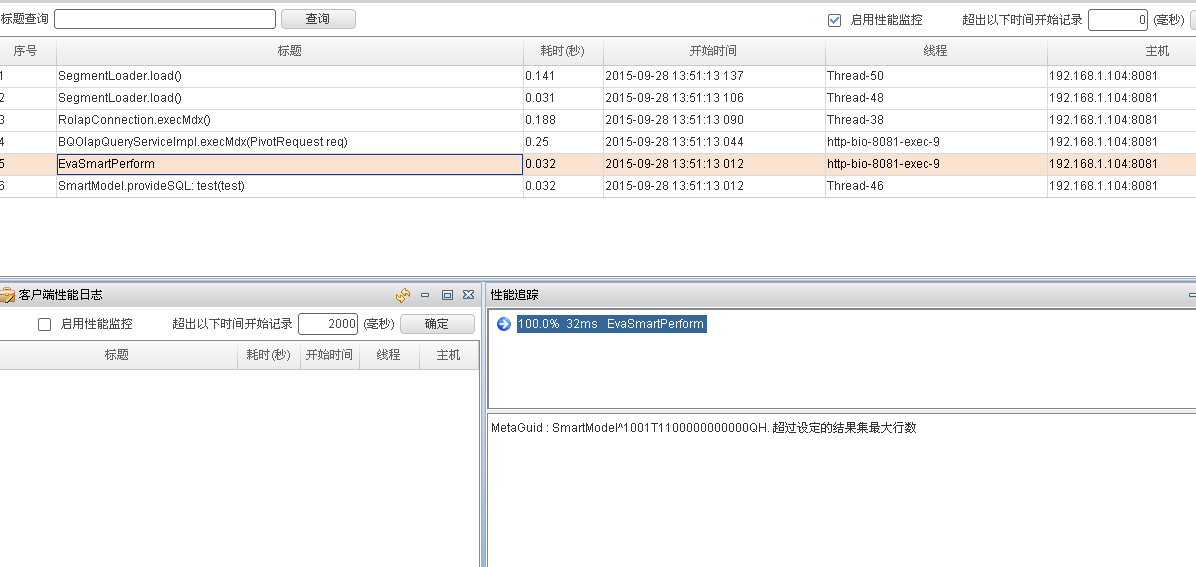


可以看出，提示了具体的哪个语义模型，会造成性能的问题点描述。

这种情况，可以选择否，重新设计，也可以直接保存。

#### 性能监控评估

当透视表中引用了上面定义的这种分析模型，在运行时如果发生性能问题，可以打开性能监控查看。



EvaSmartPerform这行就是做的语义模型性能评估，在性能追踪部分同样可以看到问题描述。需要说明的是，当不打开性能监控时，是不会做性能评估的，所以无需担心会造成性能问题。

#### 配置文件

上面我们多次提到配置文件中的最大时间和结果集的最大数量，这个配置文件的路径是：/uapHome/resources/ae/olap/olapenv.properties。

配置项是：

最大执行时间：com.ufida.bi.olapsrv.maxSmartExecTime=3，单位是秒，默认3秒。

结果集最大数量：com.ufida.bi.olapsrv.maxSmartResultSize=10000，默认10000行。