**Druid学习笔记**

**段卫东**

# Druid segment的数据结构

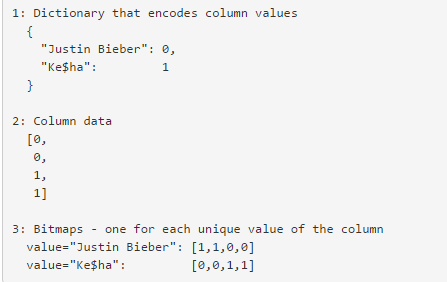
Druid数据存储主要分为三列，timestamp，dimension，metric.

dimension列主要分为三个数据结构。

1.字典。把dimension的value映射成integer。这样可以明显缩小存储。

2.存储值的list。一个List表示每行存储的值的索引。

3.一个bitmap。每个value一个bitmap，还标识该value都存在哪些row中，如果存在1，不存在0.



2 主要用来做groupBy和topN

3主要用来filter。

FILTER机制：