深入Fast-Cubing及TopN

2016.4

史少锋

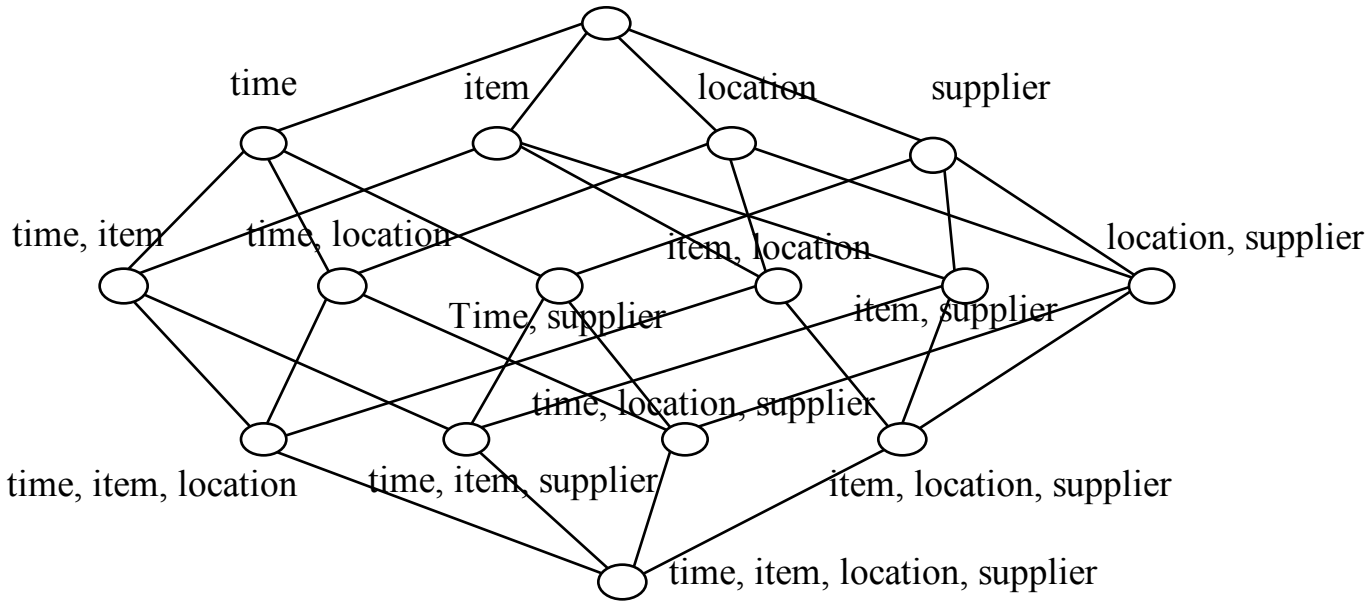
shaofeng@kyligence.io

Apache Kylin v1.5 New Features

- Plug-in Architecture
- New MR Cube Engine with fast cubing (1.5x faster)
- New HBase Storage with parallel scan (2x faster)
- Near real-time analysis (experimental)
- User defined aggregations
- Approximate Top-N Measure
- Excel / PowerBI / Zeppelin integration

Fast Cubing Algorithm

Cube – Balance Between Space and Time



0-D(apex) cuboid

1-D cuboids

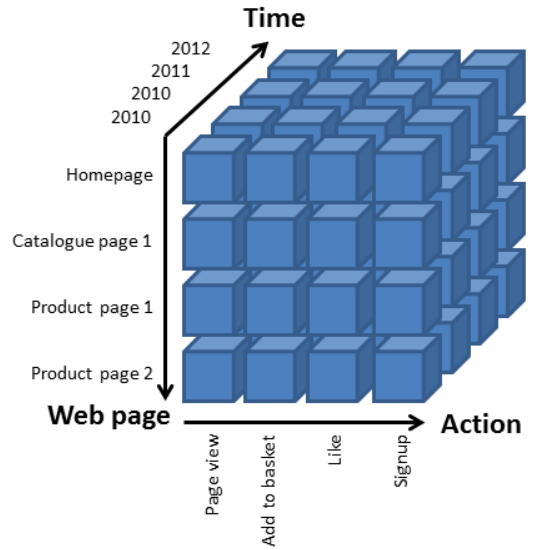
2-D cuboids

3-D cuboids

4-D(base) cuboid

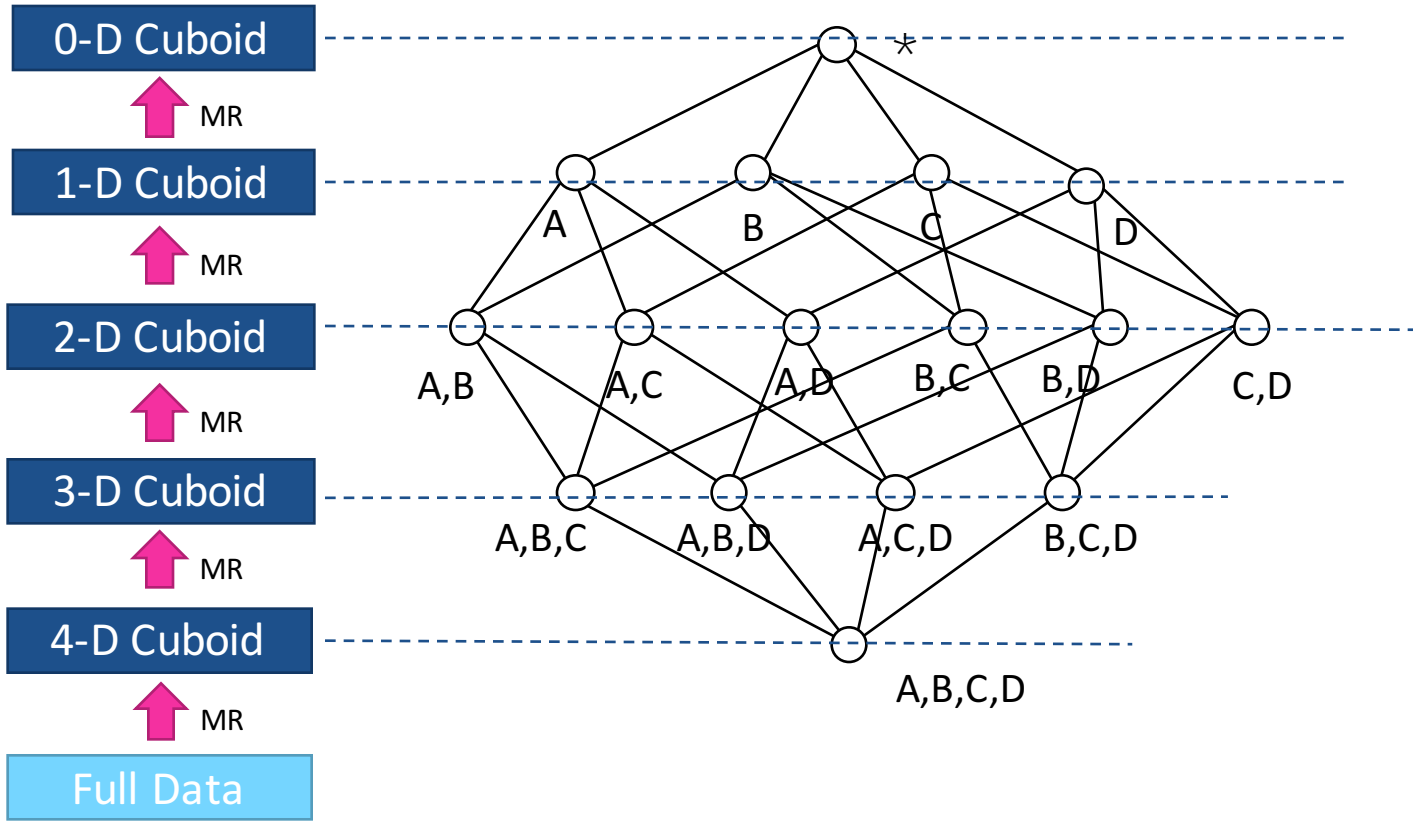
OLAP Cube

- Cuboid = one combination of dimensions
- Cube = all combination of dimensions (all cuboids)



- Base vs. aggregate cells; ancestor vs. descendant cells; parent vs. child cells
  1. (9/15, milk, Urbana, Dairy_land) - **<time, item, location, supplier>**
  2. (9/15, milk, Urbana, *) - **<time, item, location>**
  3. (*, milk, Urbana, *) - **<item, location>**
  4. (*, milk, Chicago, *) - **<item, location>**
  5. (*, milk, *, *) - **<item>**

Layered Cubing

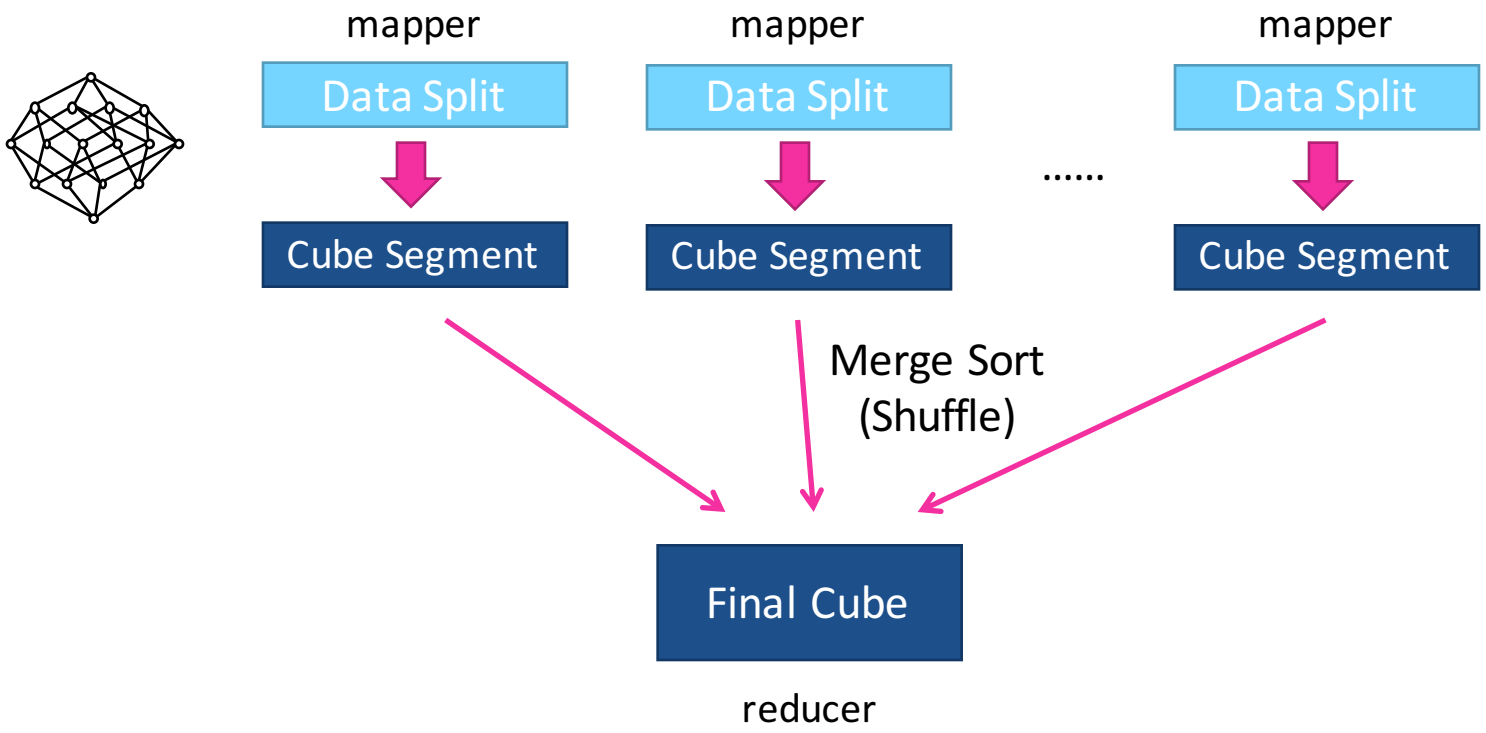


Layered Cubing Summary

- Pros
  - Depends on MR to sort and aggregate -> Simple
  - Low mapper CPU/Memory consumption -> Stable

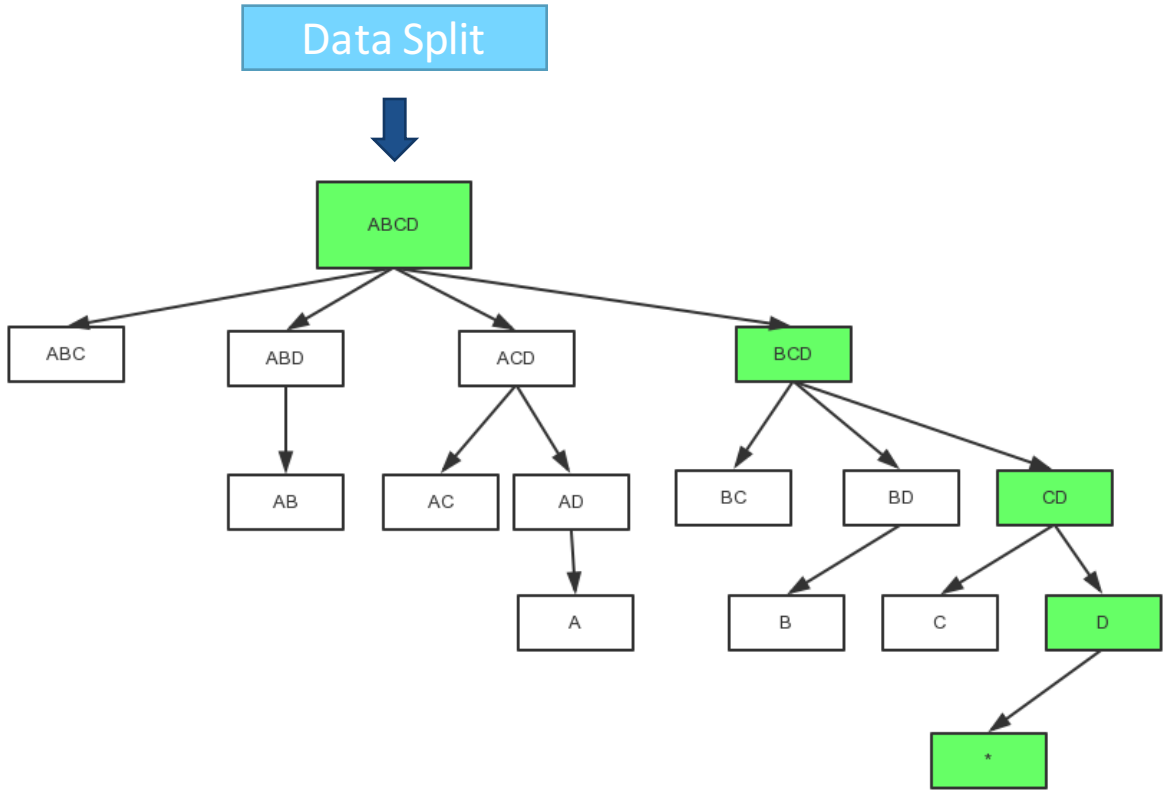
- Cons
  - High network IO
  - Job submission overhead
  - Total duration is long

Fast (in-mem) Cubing



Inside Mapper

- 深度优先遍历
- 多路并发



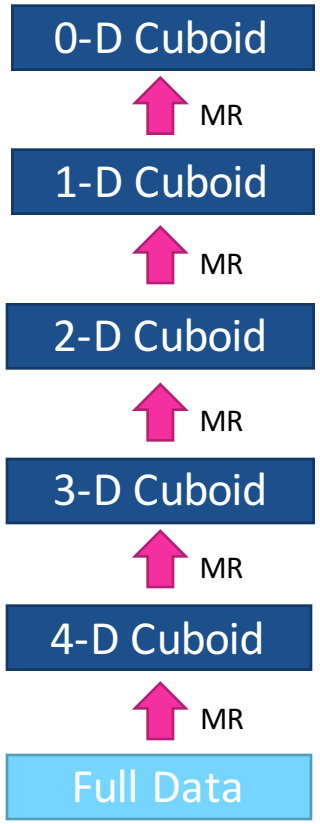
Fast Cubing Summary

- Pros
  - Lesser network pressure
  - Independent cubing algorithm that can be reused by Streaming, Spark.
  - 30%-50% faster than layered cubing
- Cons
  - Code complexity
  - High mapper CPU/Mem consumption
  - May fail at reduce phase in extreme case

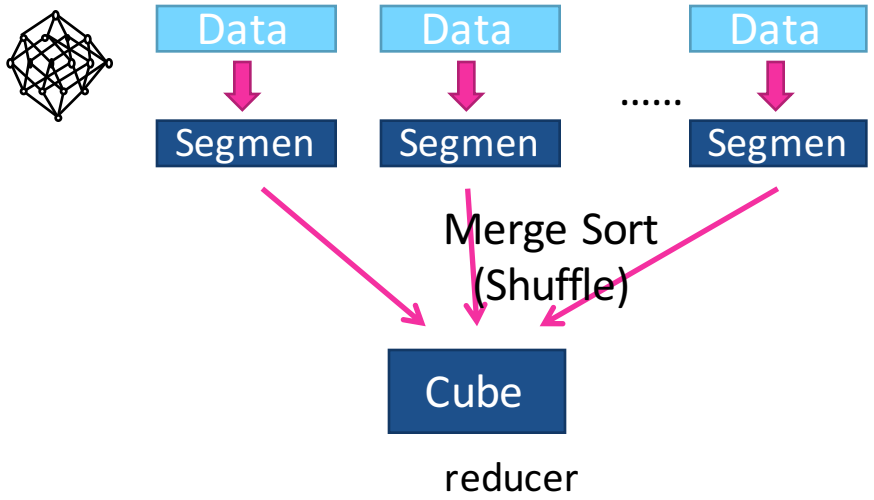
Make fast cubing faster

- Give more Mem to Mappers
  - By default Kylin requests 3Gb mem for mapper; Configure more if possible with "**kylin.job.cubing.inmem.mrjob_conf_override"**
- Balance between parallel computing and map side aggregation
  - By default Kylin request 1 mapper per 32 Mb source data;
  - If cube expansion is low, feed more data to a mapper can gain better overall performance; vice versa.
  - Customize "dfs.block.size" in kylin_hive_conf.xml

Which is the Best?



OR

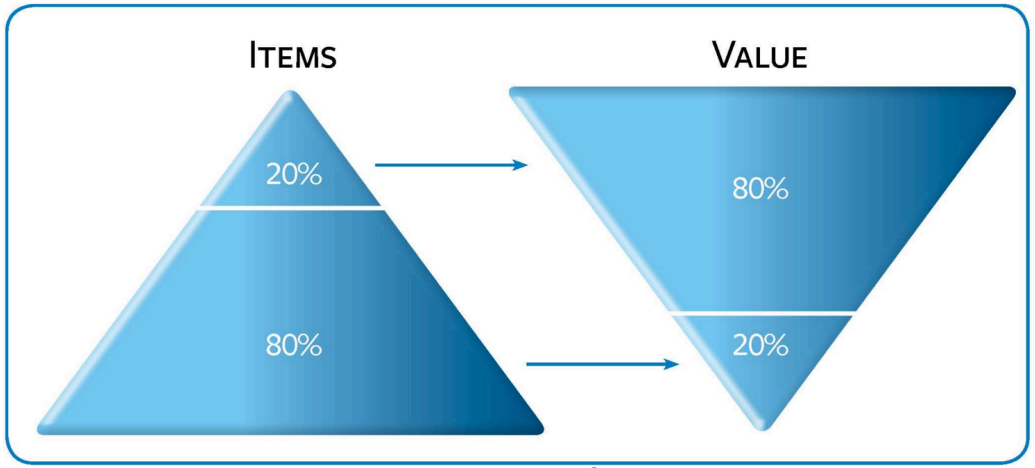


Let Kylin Select!

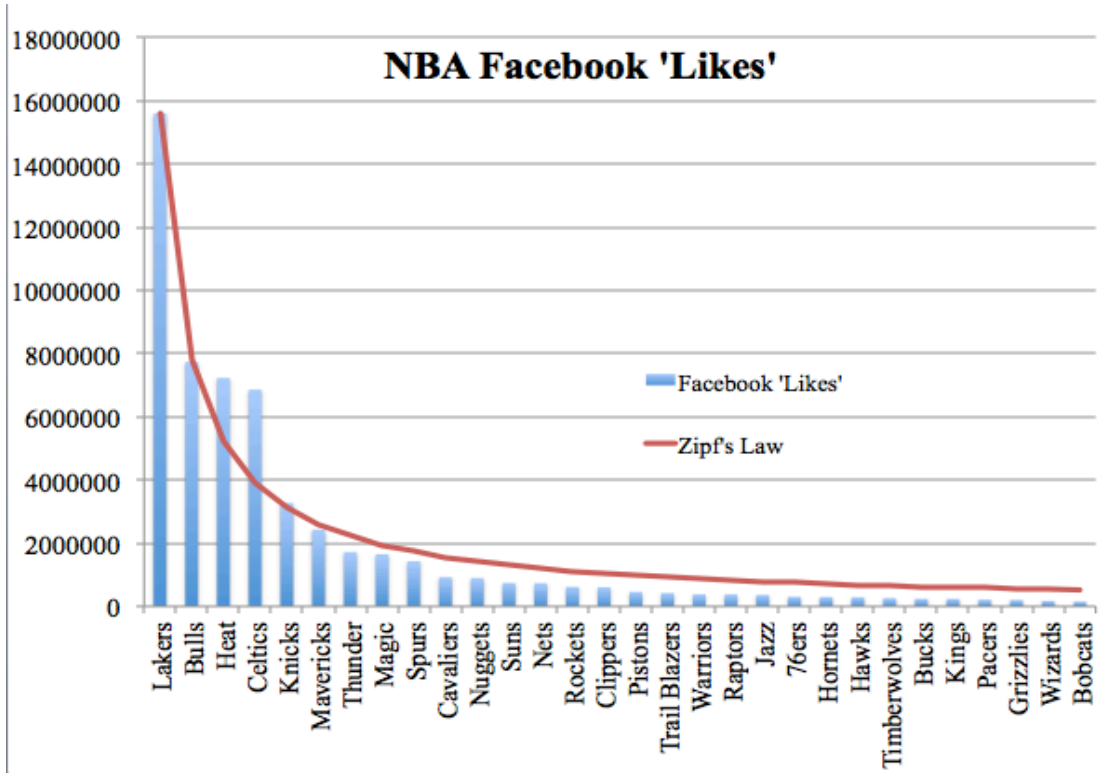
- Smartly detect before cubing
  - Estimate the micro segment size of each mapper
  - Estimate the size after reduce
  - Get the key overlap ratio
- Decide cubing algorithm
  - High key overlap: by layer cubing
  - Low key overlap: fast cubing
  - Threshold is configurable with *“kylin.cube.algorithm.auto.threshold”*

Approximate Top-N Measure

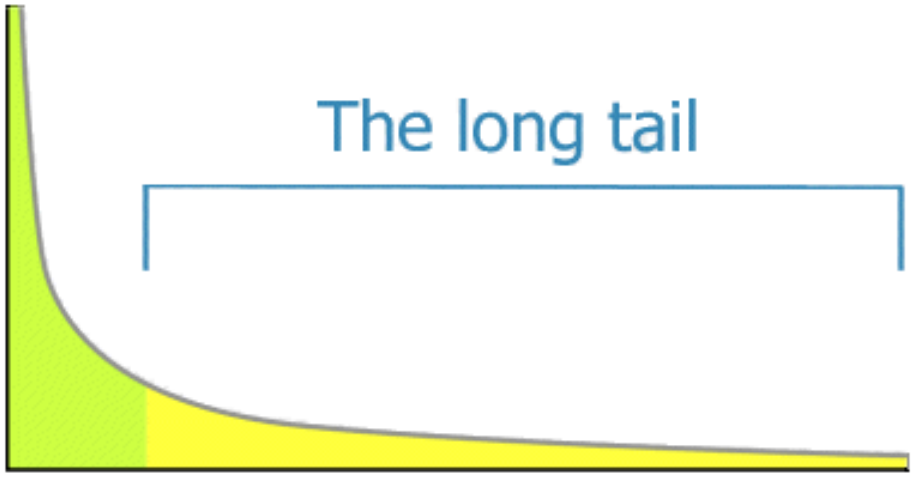
20/80 Rule & Zipf Law



20%的商品，创造80%的价值



Long tail



UHC维度使得Cube膨胀率奇高，分析数据的难度增大

- Long tail -> Ultra High Cardinality Dimension

Sample Top-N query

```
select dt, loc, item, sum(gmv)
from kylin_sales
  where dt='2015-10-01' and loc='BJ'
group by dt, loc, item
  order by sum(gmv) desc
    limit 100
```

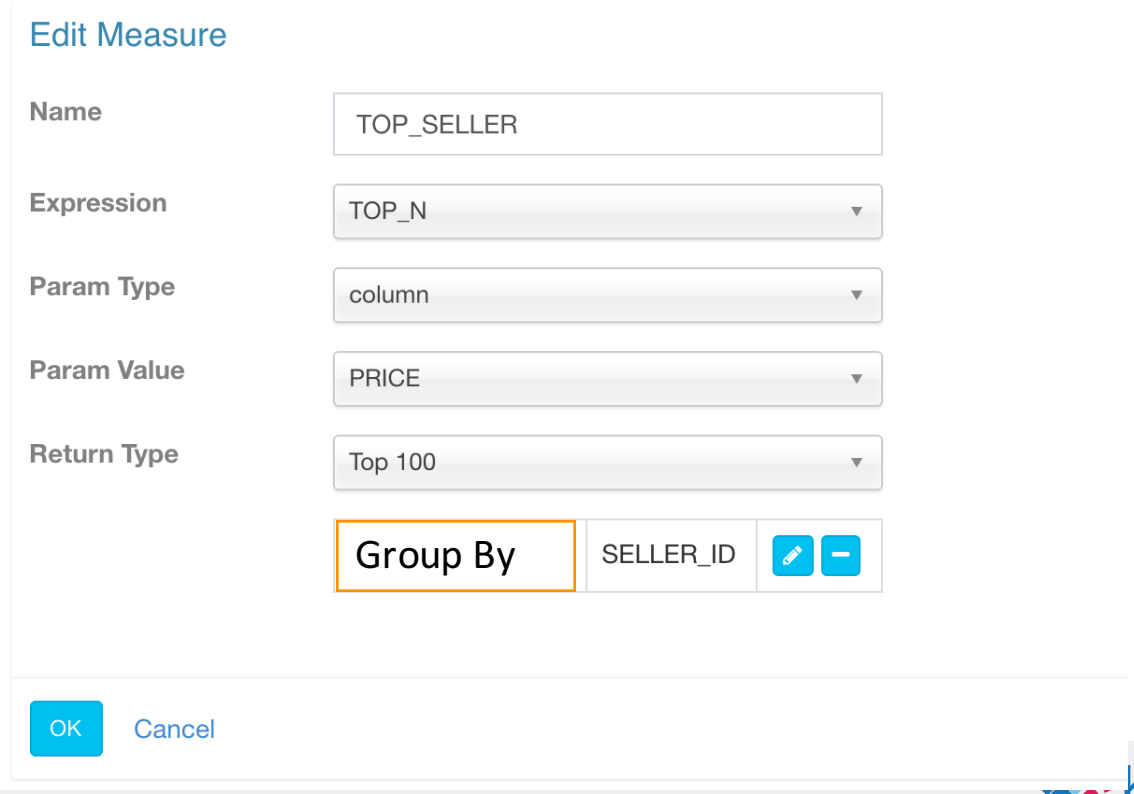
Cube without Top-N

- All “group by” column need be dimension;
- Need fetch all items before sorting and return top records;

DT	LOC	ITEM	Measure Sum(gmv)
2015-10-1	BJ	Item-A	300
2015-10-1	BJ	Item-B	400
2015-10-1	BJ	Item-C	100
2015-10-1	BJ	Item-D	200
2015-10-1	BJ	Item-E	50
2015-10-1	BJ	…	…
2015-10-1	BJ	Item-XXXXXXXXX	200

Approximate Top-N Measure

- Define Top-N as a measure;
- Calculate Top records during cube build;
- Answer Top-N queries directly from pre-calculation.



Approximate Top-N Algorithm

- SpaceSaving Top-N
  - Ahmed Metwally, et al. “Efficient computation of frequent and top-k elements in data streams”. Proceeding ICDT'05 Proceedings of the 10th international conference on Database Theory, 2005.
- A parallel SpaceSaving
  - Massimo Cafaro, et al. “A parallel space saving algorithm for frequent items and the Hurwitz zeta distribution”. Proceeding arXiv: 1401.0702v12 [cs.DS] 19 Setp 2015.

Cube with Top-N

- Move High the cardinality column to Measure, together with metric.
- Records are ordered, tail records are truncated.

DT	LOC	Measure Top-N
2015-10-1	BJ	Item-1,10000
		Item-2,9000
		……
		Item-5000,200
2015-10-1	SH	Item-1,10000
		Item-2,9000
		……
		Item-5000,200

Top-N Merge

DT,LOC	TopN
2015-10-1,BJ	Item A, $500 Item B, $490 Item C, $400 …
2015-10-1,SH	Item A, $800 Item C, $700 Item B, $600 …

……

agg

DT	TopN
2015-10-1	Item A, $1300 Item C, $1100 Item B, $1090 …

Top-N Summary

- Pros
  - Saving Space, cube size can decrease 90%
  - Avoid massive scan, query performance can improve more than 10 times

- Cons
  - Approximate calculation
  - Error accumulates at high level aggregation
  - Error ratio depends on data distribution, space allocated, aggregation level, etc.

About Kylin

- Contact:
  - Shaofeng Shi: shaofeng@kyligence.io
- Kylin Site:
  - http://kylin.io
- 微信公众号：
  - ApacheKylin
- Twitter/微博:
  - @ApacheKylin