Adaptive Online Aggregation with Randomness Detection

Weijie Zhao[1], Zhiyi Huang[1], Florin Rusu[1], Kai Zeng[2], Rahul Potharaju[3]

[1]University of California Merced, [2]Alibaba, [3]Microsoft

wzhao23,zhuang29,frusu@ucmerced.edu, kai.zeng@alibaba-inc.com, rapoth@microsoft.com

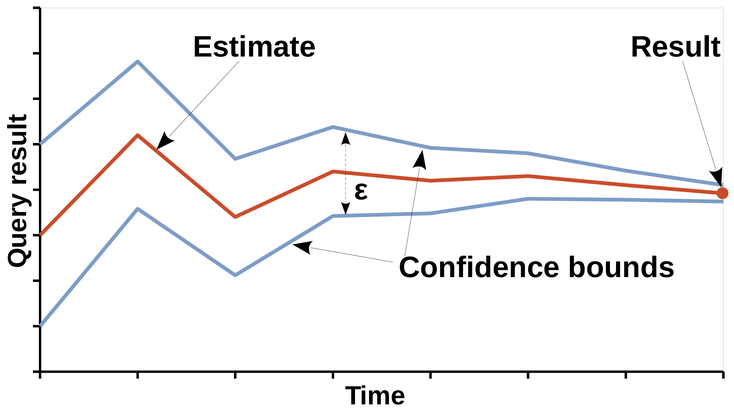
Approximate Query Processing

Advantages	Disadvantages
Faster execution	Inaccurate
Non-blocking	Heavy preprocessing

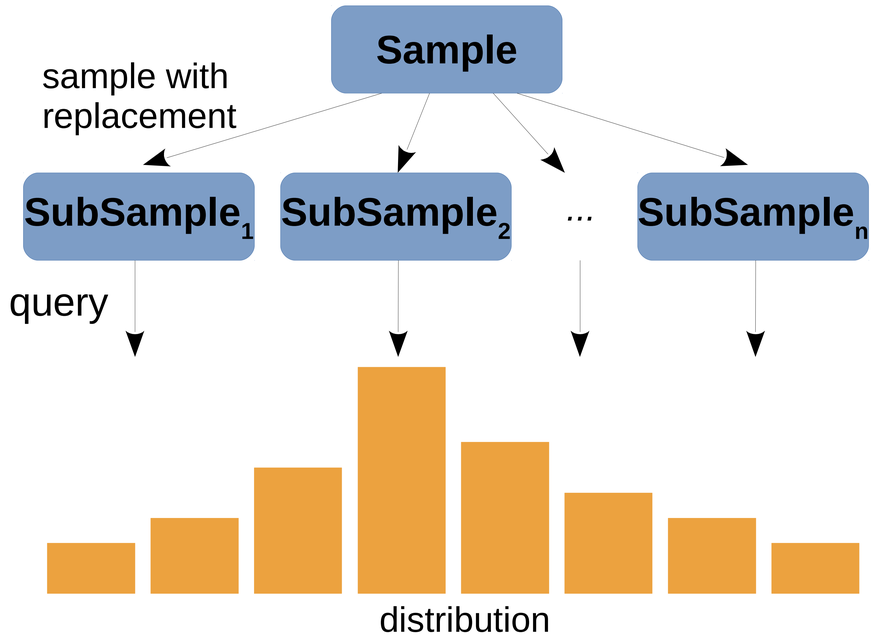
Many application can tolerate some errors

Block-level sampling reduces preprocessing but increase errors

Online Aggregation



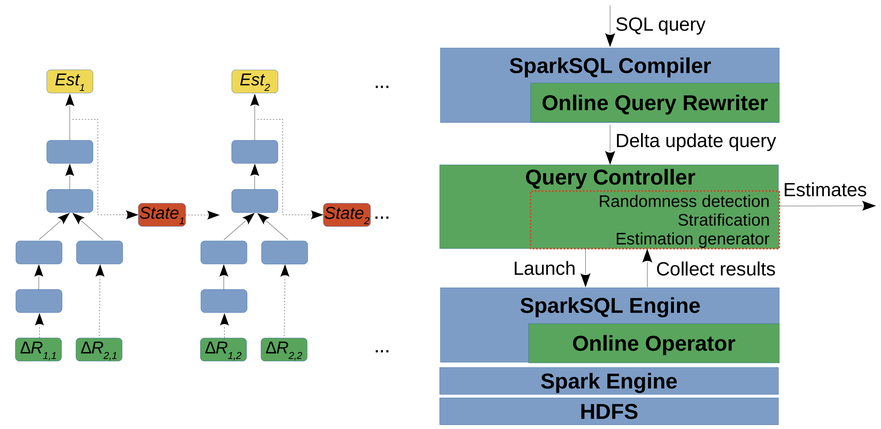
Bootstrap Estimation



Stratification

- Intuitively: stratify on grouping keys!
- Can we also stratify on heavy hit join keys?
- Reduce computations → Spend computing power on more important data

Spark Implementation



Sampling and Estimation

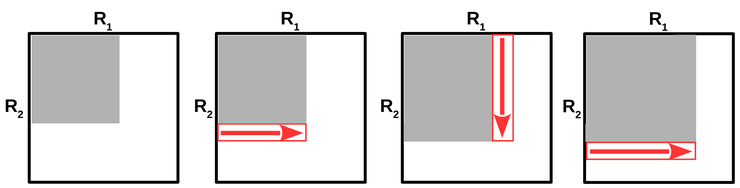
sampling	*preprocessing*	*bound*
tuple-level	shuffling	tight
block-level	none	wide

estimation	*aggregates*	*overhead*
closed-form	limited	exponential to #relations
bootstrap	extended	constant to #resampling

Existing solutions: we cannot get *no preprocessing* and *tight bounds* at the same time

Ideal: *no preprocessing* + *tight bounds* + *extended aggregates*

Ripple Join



Randomness Detection

- Is data shuffling necessary for all queries?
- Is it possible to treat a block-level sample as a tuple-level sample?
- Can we identify when is it possible?

Observation 1: Some attributes are not clustered in the original dataset.

Observation 2: After we have a "*large enough*" block-level sample, it is safe to treat it as a tuple-level sample.

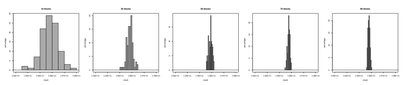
Heuristic:

- Monitor the difference among distributions of each iterations.
- Kullback-Leibler divergence, earth mover's distance.
- Fast to convert ⟷ No statistical guarantee.

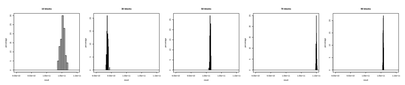
Statistical:

- Are two sample distributions extracted from the same population?
- Chi-squared test, Kolmogorov-Smirnov test.
- Slow to convert ⟷ Statistical guarantee.

Random



Sorted



Experiments

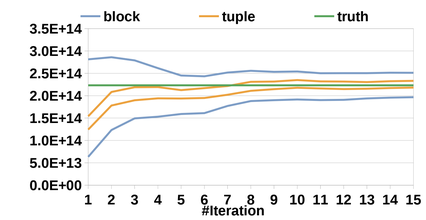
Spark 2.1.0, Hadoop 2.8.2.

9 node cluster, 1 coordinator and 8 working nodes. Each node has 16 cores, 28 GB of memory, and 4 TB of HDD storage.

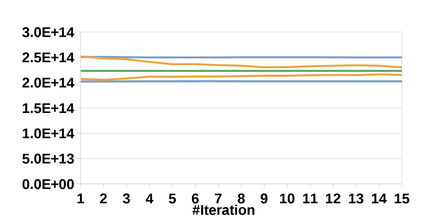
TPC-H: 1 TB

Q3: `SELECT SUM(L_EXTENDEDPRICE * (1 - L_DISCOUNT)) FROM lineitem,orders,customer WHERE L_ORDERKEY = O_ORDERKEY AND O_CUSTKEY = C_CUSTKEY`

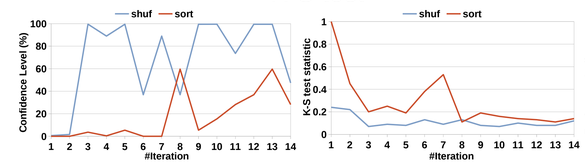
Sorted



Random



K-S test



Earth mover's distance