



Optimizing Presto for Uber-scale

Vivek Bharathan

Uber's mission is to
ignite opportunity by
setting the world in
motion.

10K

Cities

15B

Cumulative Trips

6.9B

Trips/2019

Data informs every decision at Uber



**Community
Operations**



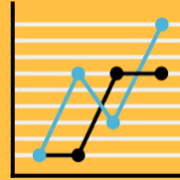
**Marketplace
Pricing**



Eats



Compliance



Growth Marketing



Data Science



Presto @ Uber-scale

12K

Monthly Active Users

400K

Queries/day

35PB

HDFS data processed/day

2.6K

Nodes

12

Clusters

Agenda

- **Partial Aggregation Pushdown**
- **Approximate Queries**
- **JVM tuning & Memory Management
(Lessons Learned)**



Problem Statement

Motivation

Data consistency checks & Health checks account for 28% of Presto queries at Uber

A large number of these perform simple aggregations over table partitions

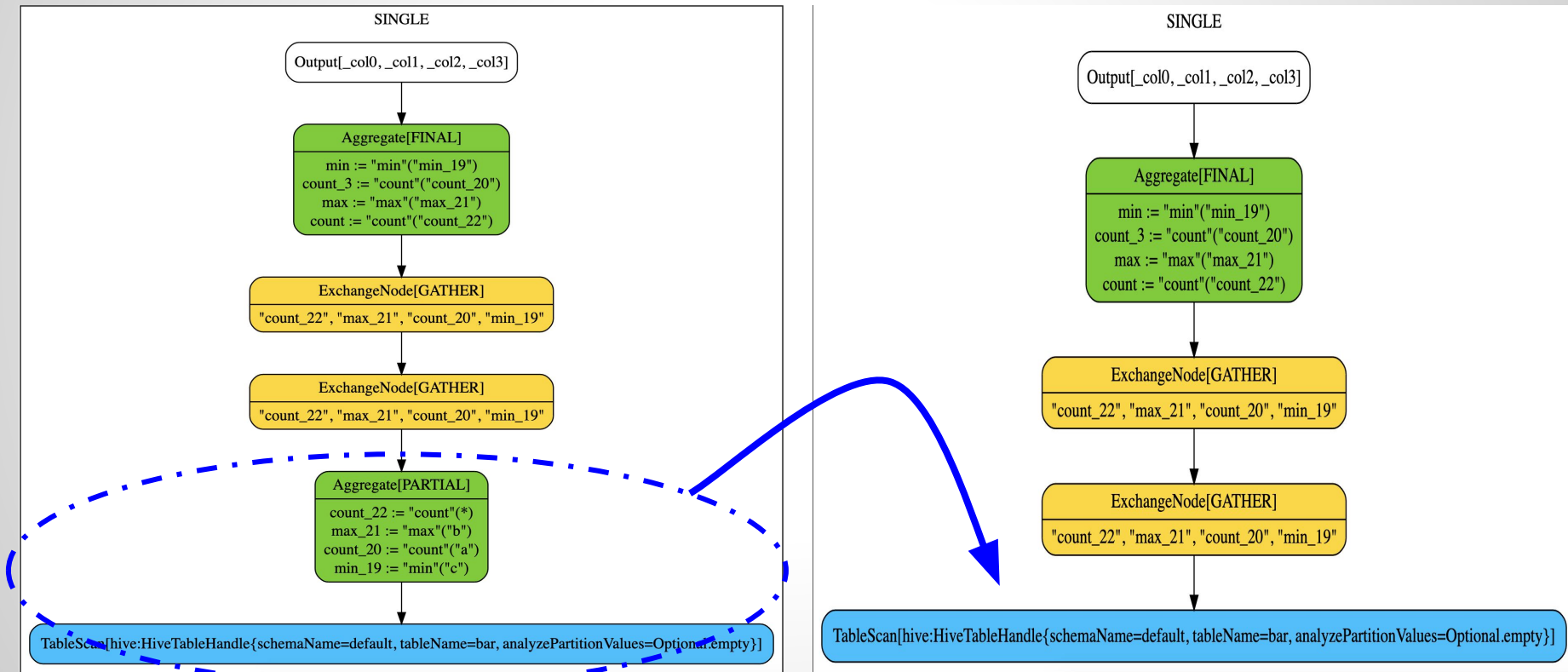
Proposed Solution

Typical Hadoop file formats serialize some metadata (statistics) as part of the file's header/footer

Use these statistics to compute the partial aggregates and avoid reading the whole file

Query Plan

select count(*), count(a), max(b), min(c) from default.bar where d=10;



Agenda

- Partial Aggregation Pushdown
- **Approximate Queries**
- JVM tuning & Memory Management (Lessons Learnt)



Problem Statement

Motivation

Classes of queries that do not require precise results

Adhoc development/testing

Exploratory analysis

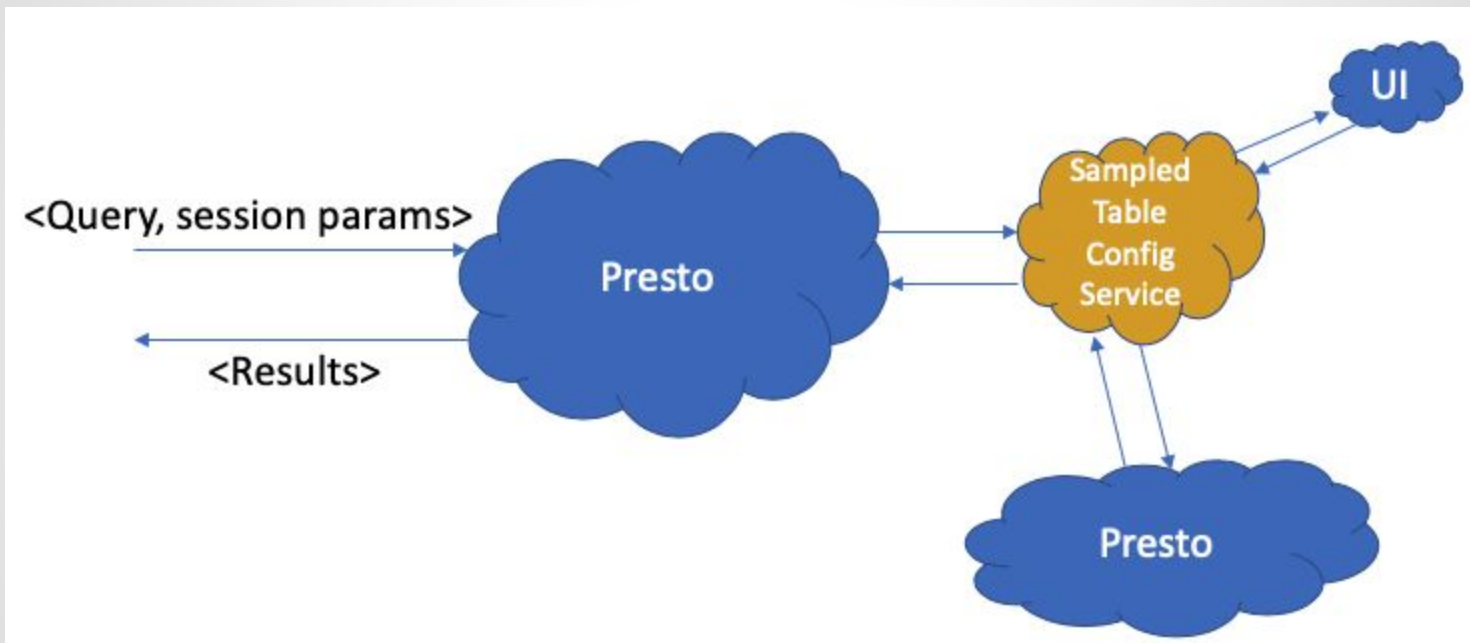
Proposed Solution

Create a sampled version of relevant datasets
(typically events data)

Rewrite the queries to use the sampled tables

Users can opt in

Configuration



Agenda

- Partial Aggregation Pushdown
- Approximate Queries
- **JVM tuning & Memory Management
(Lessons Learnt)**



Problem Statement

Long Garbage Collection Pauses

Old Gen, Stop the World GC

Variable throughput, flaky performance

Operational nightmare

Simple Win

Java 8 -> Java 11

Reduced peak GC pauses from
20 minutes to 8 sec

Increased CPU throughput

Credited to [G1 parallel full GC](#)

Additional monitoring for
consecutive GCs and kill JVM if
not recovered in 1min

Thank you

