

Build Your Own OctopusDB: Blinktopus Edition

Ali Hashaam, Ali Memon, Guzel Mussilova, Pavlo Shevchenko

Scientific Project: Databases for Multi-Dimensional Data, Genomics and Modern Hardware

July 10, 2017

Table of Contents

Motivation/Problem Statement

Background

Conceptual Idea and Implementation

Evaluation Setup and Results

Related Work

Conclusion and Future Work

Demonstration

Motivation

1. Companies need to pick only specialized DBMSs, each tailored to their specific use-case.

Motivation

1. Companies need to pick only specialized DBMSs, each tailored to their specific use-case.
⇒ Need for *one size fits all system* (e.g. HTAP)

Motivation

1. Companies need to pick only specialized DBMSs, each tailored to their specific use-case.
⇒ Need for *one size fits all system* (e.g. HTAP)
2. Support OLAP queries for analysis over real-time data (i.e., freshness).

Motivation

1. Companies need to pick only specialized DBMSs, each tailored to their specific use-case.
⇒ Need for *one size fits all system* (e.g. HTAP)
2. Support OLAP queries for analysis over real-time data (i.e., freshness).
⇒ Explore the techniques related to more interactive queries (e.g. *Approximate Query Processing*)

Background

1. OctopusDB

Background

1. OctopusDB

- uses logs as a primary storage;

Background

1. OctopusDB

- uses logs as a primary storage;
- mimicks several types of systems (OLAP, OLTP, etc.) by representing them as *Storage Views*.

Background

1. OctopusDB

- uses logs as a primary storage;
- mimicks several types of systems (OLAP, OLTP, etc.) by representing them as *Storage Views*.

2. BlinkDB

- successfully integrates AQP techniques into its architecture.

Conceptual Idea and Implementation

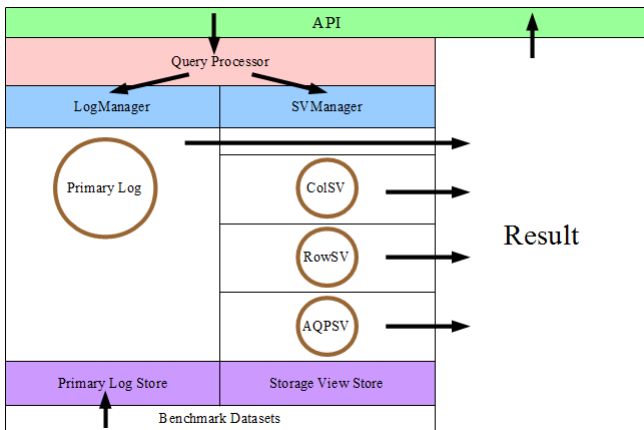


Figure 1: OctopusDB Architecture.

Conceptual Idea and Implementation

Which synopses to pick?

¹<https://datasketches.github.io>

Conceptual Idea and Implementation

Which synopses to pick?

- Equi-depth histograms
 - suitable for range queries;
 - simple to implement and interpret.

¹<https://datasketches.github.io>

Conceptual Idea and Implementation

Which synopses to pick?

- Equi-depth histograms
 - suitable for range queries;
 - simple to implement and interpret.
- Sketches
 - DISTINCT COUNT queries;
 - *HyperLogLog*;
 - *DataSketches* library by *Yahoo!* ¹

¹<https://datasketches.github.io>

Evaluation Setup

Evaluation Setup

Machine

- CentOS Linux 7.1.1503
- Java SDK 8u131-b11-linux-x64
- 2 Intel(r) Xeon (TM) E5-2630 v3s CPU @ 3.2GHz processors (8 cores each) and 1024 GiB memory

Evaluation Setup

Machine

- CentOS Linux 7.1.1503
- Java SDK 8u131-b11-linux-x64
- 2 Intel(r) Xeon (TM) E5-2630 v3s CPU @ 3.2GHz processors (8 cores each) and 1024 GiB memory

Benchmark Datasets

- TPC-H datasets (Orders and Lineitems)

Evaluation Setup

Machine

- CentOS Linux 7.1.1503
- Java SDK 8u131-b11-linux-x64
- 2 Intel(r) Xeon (TM) E5-2630 v3s CPU @ 3.2GHz processors (8 cores each) and 1024 GiB memory

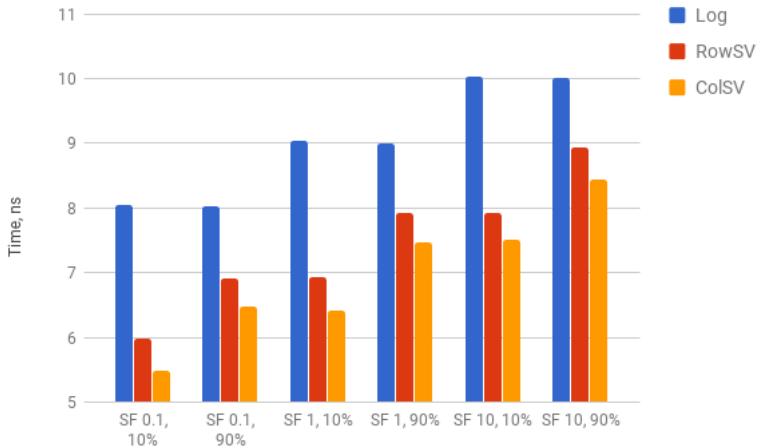
Benchmark Datasets

- TPC-H datasets (Orders and Lineitems)

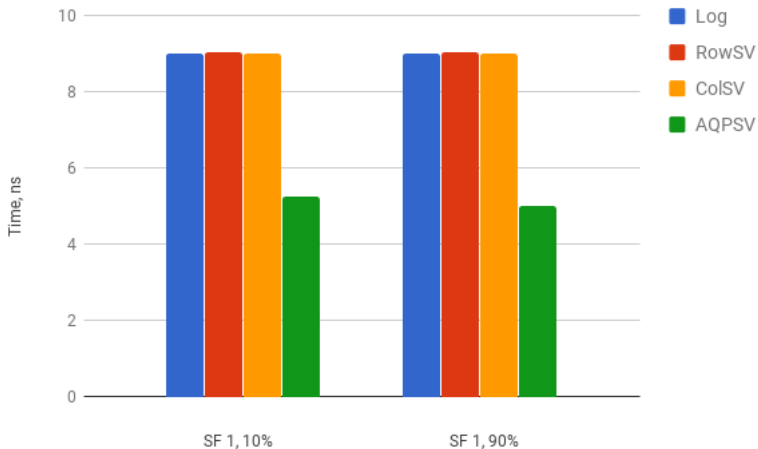
Experiments

1. Average response time for a range query on the Orders table with various scaling factors and predicate selectivity.
2. Average response time for a count-range query on the Orders table. Comparison with an equi-depth histogram.
3. Average response time for a count distinct query on the Orders table. Comparison with a HLL sketch.

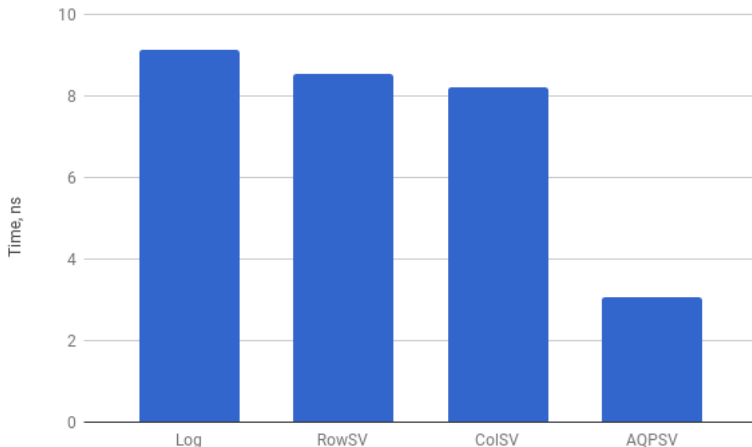
Results. Experiment 1



Results. Experiment 2



Results. Experiment 3



Challenges



Challenges

[illegible]

Challenges

```

500000000 linesitem were loaded
500000000 linesitem were loaded
500000000 linesitem were loaded
500000000 linesitem were loaded
Start Export count 1
Create CalcW for 10%
Create BaseW for 10%
Start 10%
Log ready
Flow ready
Col ready

!!!! Result: LOG: 1.0700225801610 ; ROW: 0.4727587 ; COL: 5.194757287 !!!

Create CalcW for 90%
Create BaseW for 90%
Start 90%
Log ready
Flow ready
Col ready

!!!! Result: LOG: 1.9852704547610 ; ROW: 0.668200268 ; COL: 2.728436
0000 !!!

campaign: gpus(45,10) - Blinktopus blinktopus, Blinktopus Tests

File: DB View Search Terminal: Help
top: 19:43:59 up 5 days, 11:28, 5 users, load average: 1.01, 3.04, 1.05
tasks: 986 total, 1 running, 986 sleeping, 0 stopped, 0 zombie
Meminfo: 9.2 m, 0.0 m, 0.0 m, 90.0 m, 0.0 m, 0.0 m, 0.0 m, 0.0 m
Mem Mem: 100077228total, 8337371free, 2208841used, 2370872 buff/cache
Mem Swap: 4134380 total, 2866604 free, 3387784 used, 83382808avail Mem

PID PPID PR NI VIRT RES SHR SPC INCR OVRFLW TIME COMMAND
18864 campera 20 0 0 8337 0 1944 10206 5 165 2 20 5788.41 1m
21750 blinker 20 0 127466 7886 1488 5 1.0 0.1 0 80.56 10 0
50 root 20 0 0 0 0 0 0 0 0.3 0.1 0 0.04.82 re ms/7
60 root 20 0 0 0 0 0 0 0 0.3 0.1 0 0.04.82 re ms/7
625493 root 20 0 183600 20884 332 5 0.3 0.1 0 2.29.55 wrdp
125500 campera 20 0 481936 95448 6616 6 0.3 0.1 0 2.33.17 kvmc
2 root 20 0 195660 3940 2272 2 0.0 0.0 0 0.18.38 systemd
3 root 20 0 0 0 0 0 0 0 0.0 0.0 0 0.00.26 kthreadd
5 root 20 0 0 0 0 0 0 0 0.0 0.0 0 0.00.50 kiofltrpdr
6 root 20 0 0 0 0 0 0 0 0.0 0.0 0 0.00.00 kworker/0:0
8 root 20 0 0 0 0 0 0 0 0.0 0.0 0 0.00.00 kworker/0:0
10 root 20 0 0 0 0 0 0 0 0.0 0.0 0 0.00.33 migration/0
11 root 20 0 0 0 0 0 0 0 0.0 0.0 0 0.00.00 rcuob/0
12 root 20 0 0 0 0 0 0 0 0.0 0.0 0 0.00.00 rcuob/1
13 root 20 0 0 0 0 0 0 0 0.0 0.0 0 0.00.00 rcuob/2
14 root 20 0 0 0 0 0 0 0 0.0 0.0 0 0.00.00 rcuob/3
15 root 20 0 0 0 0 0 0 0 0.0 0.0 0 0.00.00 rcuob/4
16 root 20 0 0 0 0 0 0 0 0.0 0.0 0 0.00.00 rcuob/5

```

5788,41 min = 96 hours = 4 days

Related Work

Related Work

1. Apache Samza

- logs as a primary storage;
- replicates logs on multiple nodes.

Related Work

1. Apache Samza

- logs as a primary storage;
- replicates logs on multiple nodes.

2. Rodent Store

- represents data in the various physical layouts;
- provides DBAs a high-level interface to specify the data physical representation by means of storage algebra.

Related Work

1. Apache Samza

- logs as a primary storage;
- replicates logs on multiple nodes.

2. Rodent Store

- represents data in the various physical layouts;
- provides DBAs a high-level interface to specify the data physical representation by means of storage algebra.

3. Snappy Data

- AQP Support;
- uses numerous types of synopses (samples, sketches);
- user defines the level of accuracy and the number of column sets to approximate the results.

Conclusion

Conclusion

- Systems with adaptive layouts can be effectively combined with AQP techniques.

Conclusion

- Systems with adaptive layouts can be effectively combined with AQP techniques.
- OLAP queries can benefit from AQP techniques.

Conclusion

- Systems with adaptive layouts can be effectively combined with AQP techniques.
- OLAP queries can benefit from AQP techniques.
- Non-optimized central log as a primary storage is quite prohibitive.

Future Work

- optimize centralized log (e.g. log replication, garbage collection);

Future Work

- optimize centralized log (e.g. log replication, garbage collection);
- evaluate the efficiency of the concurrency control scheme of OctopusDB;

Future Work

- optimize centralized log (e.g. log replication, garbage collection);
- evaluate the efficiency of the concurrency control scheme of OctopusDB;
- evaluate the memory footprint of histograms and sketches;

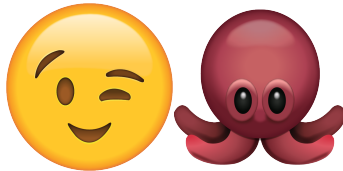
Future Work

- optimize centralized log (e.g. log replication, garbage collection);
- evaluate the efficiency of the concurrency control scheme of OctopusDB;
- evaluate the memory footprint of histograms and sketches;
- extend Blinktopus architecture to support transactional model;

Future Work

- optimize centralized log (e.g. log replication, garbage collection);
- evaluate the efficiency of the concurrency control scheme of OctopusDB;
- evaluate the memory footprint of histograms and sketches;
- extend Blinktopus architecture to support transactional model;
- extend query classes by implementing sample-based data synopses.

Demonstration



Thank you!

Questions? Recommendations? Remarks?