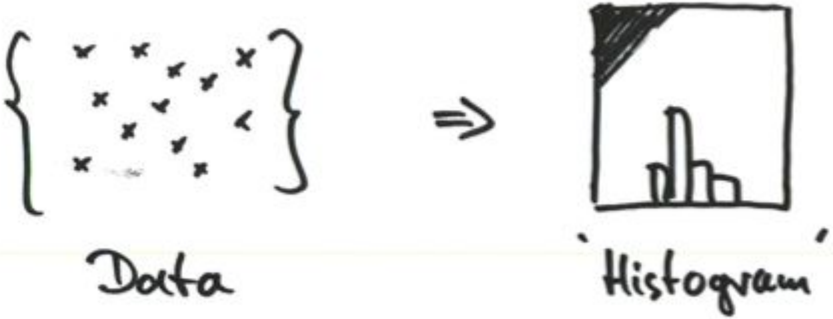
The State of the Histogram

Heinrich Hartmann

SLO Conf - May 2021

Heinrich Hartmann - SRE @ zalando (🧦)

Why Histograms?



- small
- accurate percentiles
- mergeable

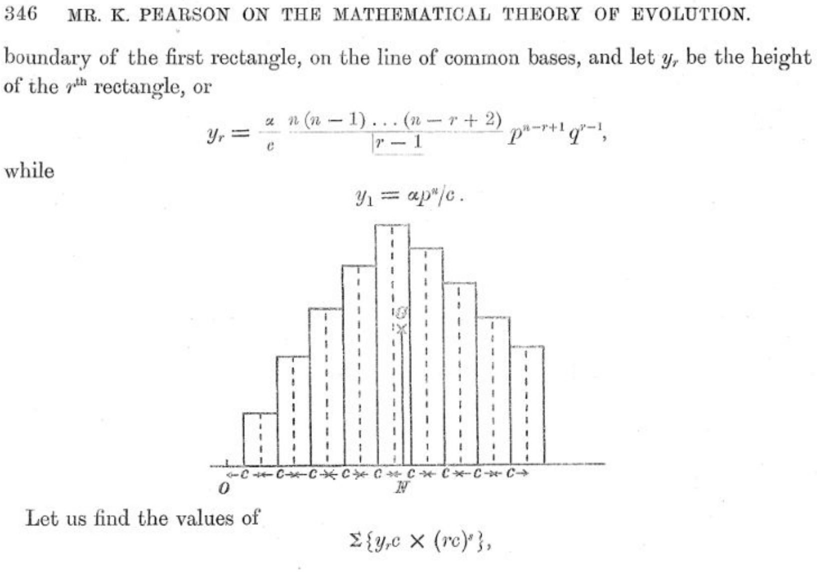
Application: Latency measurements & SLOs.

⇒ Latency SLOs done right.

- @FOSDEM 2019
- Circonus/blog.

History of

Histograms



346 MR. K. PEARSON ON THE MATHEMATICAL THEORY OF EVOLUTION.

boundary of the first rectangle, on the line of common bases, and let y_r be the height of the r^{th} rectangle, or

$$y_r = \frac{\alpha}{c} \frac{n(n-1)\ldots(n-r+2)}{\underline{|r-1}} p^{n-r+1} q^{r-1},$$

while

$$y_1 = \alpha p^n / c\,.$$

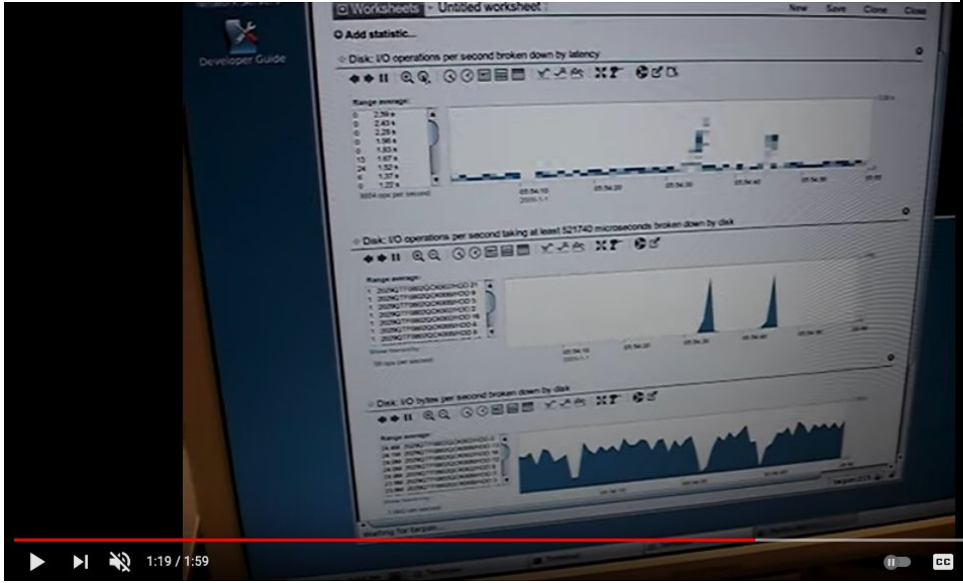
c
c
c
c
c
c
c
c
c
O
N
G

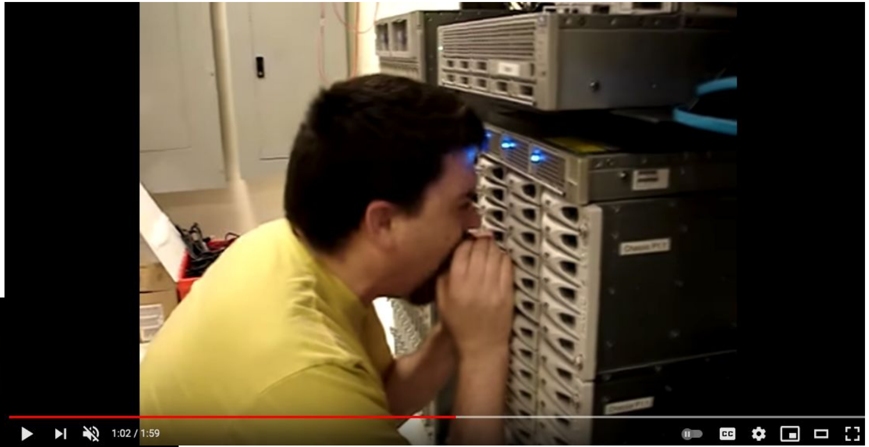
Let us find the values of

$$\Sigma\{y_r c \times (rc)^s\},$$

Pearson, K. (1895). "Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material". *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. **186**: 343–414.

1895

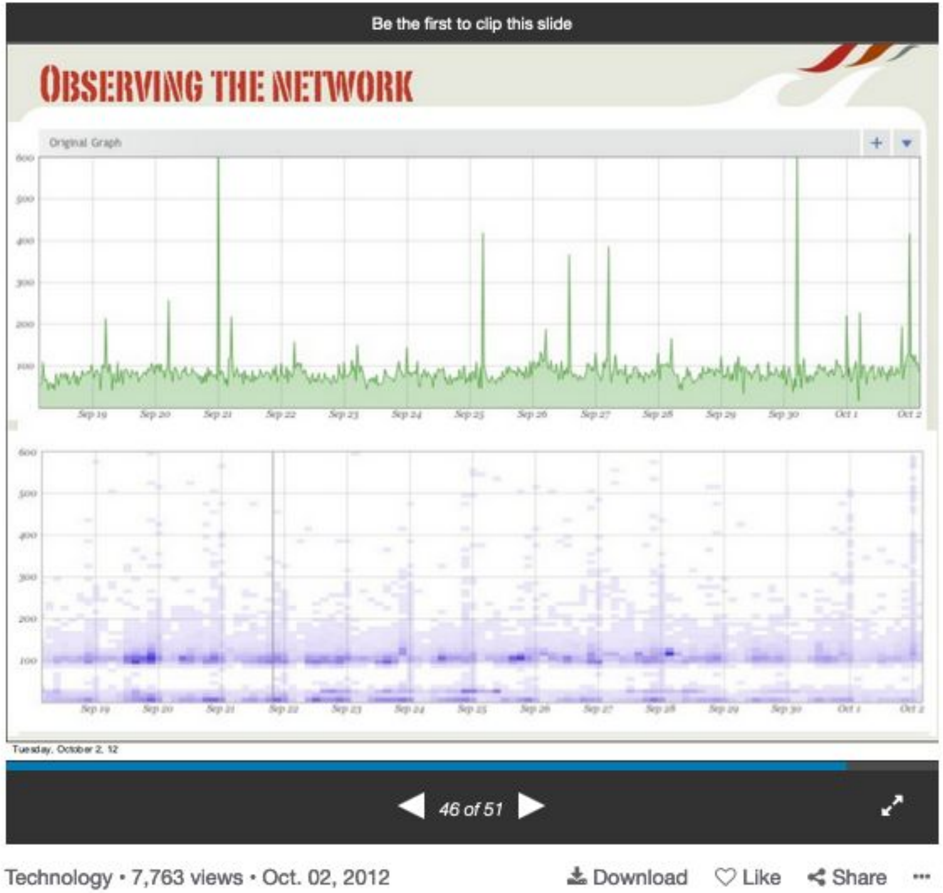

Developer Guide
Worksheets · Untitled worksheet
New Save Clone Close
Add statistic...
Disk: I/O operations per second broken down by latency
Range average:
Disk: I/O operations per second taking at least 521740 microseconds broken down by disk
Range average:
Disk: I/O bytes per second broken down by disk
Range average:
1:19 / 1:59



Brendan Gregg @ Sun -- Shouting at the DataCenter

2009

CIRCONUS



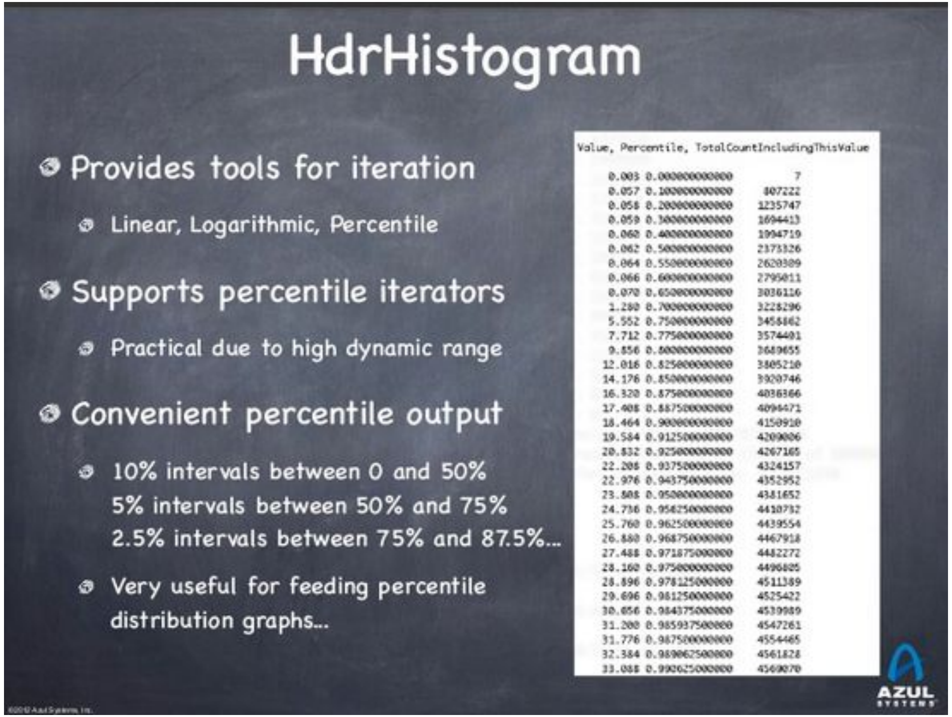
OBSERVING THE NETWORK

Tuesday, October 2, 12

Technology • 7,763 views • Oct. 02, 2012

Theo Schlossnagle @ Circonus -- Monitoring & Observability

2012



HdrHistogram

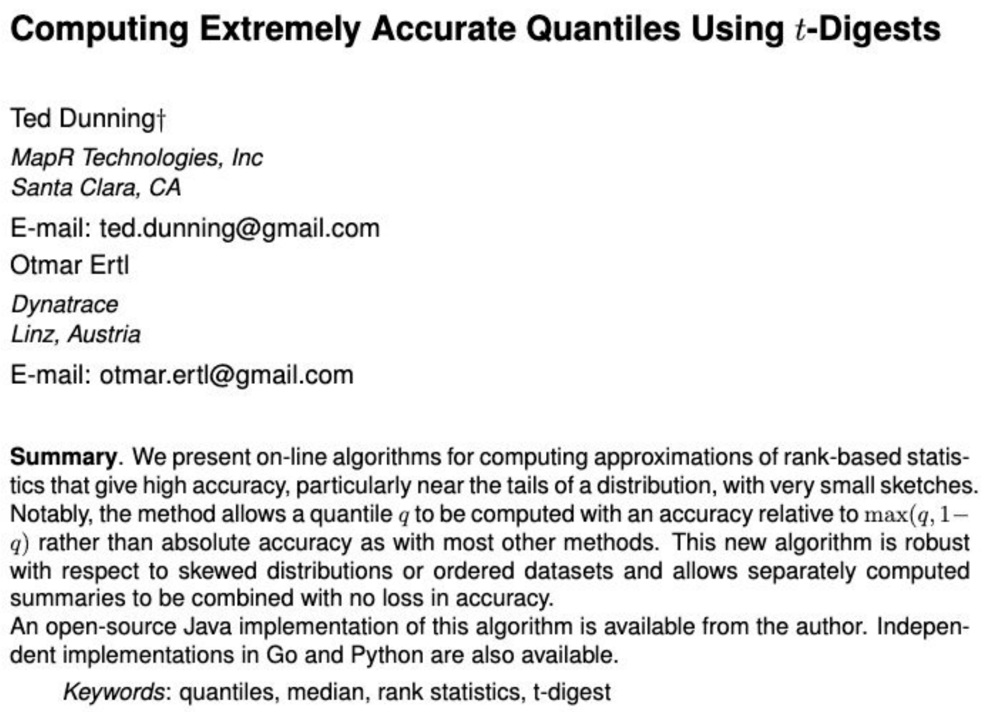
- Provides tools for iteration
  - Linear, Logarithmic, Percentile
- Supports percentile iterators
  - Practical due to high dynamic range
- Convenient percentile output
  - 10% intervals between 0 and 50%
    5% intervals between 50% and 75%
    2.5% intervals between 75% and 87.5%...
  - Very useful for feeding percentile distribution graphs...

```
Value, Percentile, TotalCountIncludingThisValue

    0.003 0.000000000000          7
    0.057 0.100000000000     807222
    0.058 0.200000000000    1235747
    0.059 0.300000000000    1694413
    0.060 0.400000000000    1994719
    0.062 0.500000000000    2373326
    0.064 0.550000000000    2620309
    0.066 0.600000000000    2795011
    0.070 0.650000000000    3036116
    1.280 0.700000000000    3228296
    5.552 0.750000000000    3458862
    7.712 0.775000000000    3574401
    9.856 0.800000000000    3689655
   12.016 0.825000000000    3805210
   14.176 0.850000000000    3920746
   16.320 0.875000000000    4036366
   17.408 0.887500000000    4094471
   18.464 0.900000000000    4150910
   19.584 0.912500000000    4209006
   20.832 0.925000000000    4267165
   22.208 0.937500000000    4324157
   22.976 0.943750000000    4352952
   23.808 0.950000000000    4381652
   24.736 0.956250000000    4410732
   25.760 0.962500000000    4439554
   26.880 0.968750000000    4467918
   27.488 0.971875000000    4482272
   28.160 0.975000000000    4496805
   28.896 0.978125000000    4511389
   29.696 0.981250000000    4525422
   30.656 0.984375000000    4539989
   31.200 0.985937500000    4547261
   31.776 0.987500000000    4554465
   32.384 0.989062500000    4561828
   33.088 0.990625000000    4569070
```

Gil Tene @ Azul -- How not to measure latency

2014



Computing Extremely Accurate Quantiles Using t-Digests

Ted Dunning†

MapR Technologies, Inc
Santa Clara, CA

E-mail: ted.dunning@gmail.com

Otmar Ertl

Dynatrace
Linz, Austria

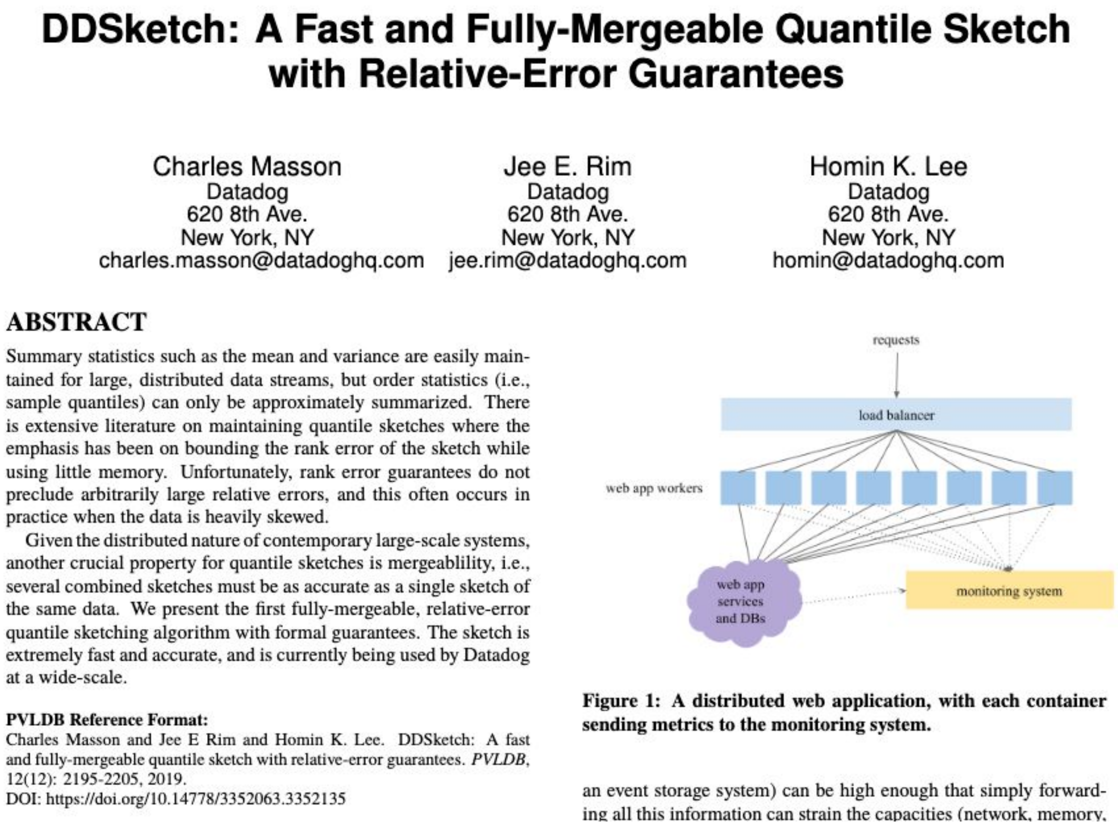
E-mail: otmar.ertl@gmail.com

Summary. We present on-line algorithms for computing approximations of rank-based statistics that give high accuracy, particularly near the tails of a distribution, with very small sketches. Notably, the method allows a quantile q to be computed with an accuracy relative to $\max(q, 1-q)$ rather than absolute accuracy as with most other methods. This new algorithm is robust with respect to skewed distributions or ordered datasets and allows separately computed summaries to be combined with no loss in accuracy.

An open-source Java implementation of this algorithm is available from the author. Independent implementations in Go and Python are also available.

Keywords: quantiles, median, rank statistics, t-digest

2014



DDSketch: A Fast and Fully-Mergeable Quantile Sketch with Relative-Error Guarantees

Charles Masson
Datadog
620 8th Ave.
New York, NY
charles.masson@datadoghq.com

Jee E. Rim
Datadog
620 8th Ave.
New York, NY
jee.rim@datadoghq.com

Homin K. Lee
Datadog
620 8th Ave.
New York, NY
homin@datadoghq.com

ABSTRACT

Summary statistics such as the mean and variance are easily maintained for large, distributed data streams, but order statistics (i.e., sample quantiles) can only be approximately summarized. There is extensive literature on maintaining quantile sketches where the emphasis has been on bounding the rank error of the sketch while using little memory. Unfortunately, rank error guarantees do not preclude arbitrarily large relative errors, and this often occurs in practice when the data is heavily skewed.

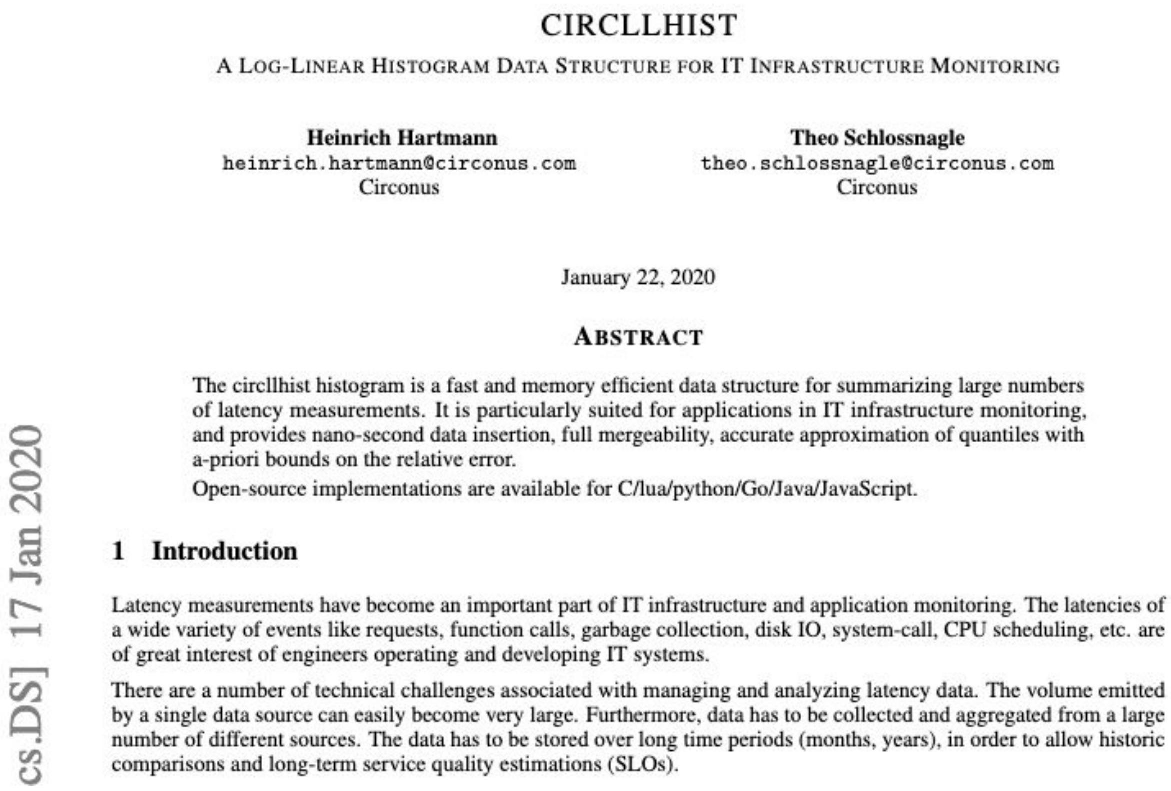
Given the distributed nature of contemporary large-scale systems, another crucial property for quantile sketches is mergeablility, i.e., several combined sketches must be as accurate as a single sketch of the same data. We present the first fully-mergeable, relative-error quantile sketching algorithm with formal guarantees. The sketch is extremely fast and accurate, and is currently being used by Datadog at a wide-scale.

PVLDB Reference Format:
Charles Masson and Jee E Rim and Homin K. Lee. DDSketch: A fast and fully-mergeable quantile sketch with relative-error guarantees. *PVLDB*, 12(12): 2195-2205, 2019.
DOI: https://doi.org/10.14778/3352063.3352135

Figure 1: A distributed web application, with each container sending metrics to the monitoring system.

an event storage system) can be high enough that simply forwarding all this information can strain the capacities (network, memory,

2019



CIRCLLHIST

A Log-Linear Histogram Data Structure for IT Infrastructure Monitoring

Heinrich Hartmann
heinrich.hartmann@circonus.com
Circonus

Theo Schlossnagle
theo.schlossnagle@circonus.com
Circonus

January 22, 2020

Abstract

The circllhist histogram is a fast and memory efficient data structure for summarizing large numbers of latency measurements. It is particularly suited for applications in IT infrastructure monitoring, and provides nano-second data insertion, full mergeability, accurate approximation of quantiles with a-priori bounds on the relative error.

Open-source implementations are available for C/lua/python/Go/Java/JavaScript.

1 Introduction

Latency measurements have become an important part of IT infrastructure and application monitoring. The latencies of a wide variety of events like requests, function calls, garbage collection, disk IO, system-call, CPU scheduling, etc. are of great interest of engineers operating and developing IT systems.

There are a number of technical challenges associated with managing and analyzing latency data. The volume emitted by a single data source can easily become very large. Furthermore, data has to be collected and aggregated from a large number of different sources. The data has to be stored over long time periods (months, years), in order to allow historic comparisons and long-term service quality estimations (SLOs).

2020

OpenHistogram

Docs FAQs License

Vendor-neutral log-linear histograms for the compression, mergeability, and analysis of telemetry data

OpenHistogram is an open source data structure for quickly and efficiently summarizing large numbers of measurements.

Get started

Organizations using OpenHistogram

CIRCONUS SmugMug SPARKPOST

2021

So many Histograms!

HdR

Circllhist

openhistogram

t-Digest

DDsketch

Which should you use?

Desireable Properties

(1) small, bounded size

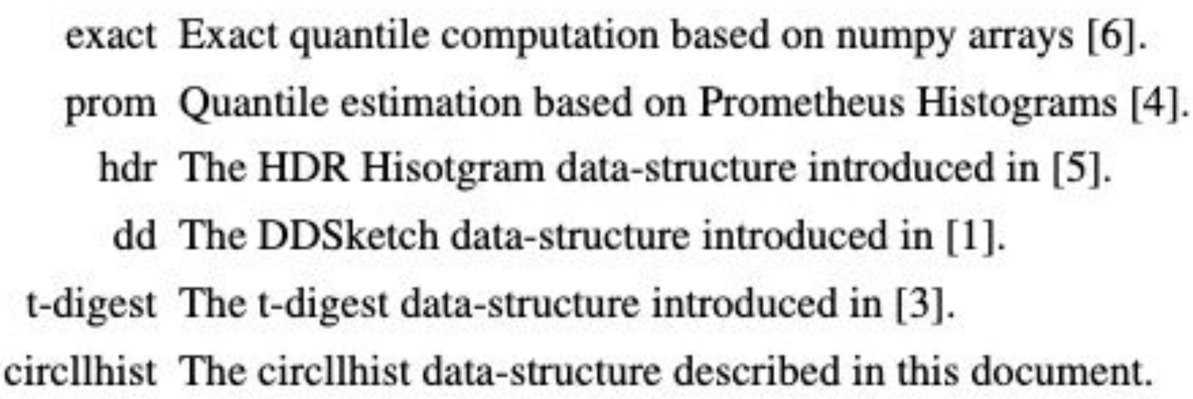
(2) fast insertion

(3) accurate quantiles

(4) mergeability.

(5) zero configuration / compatibility

Evaluation - Contenders



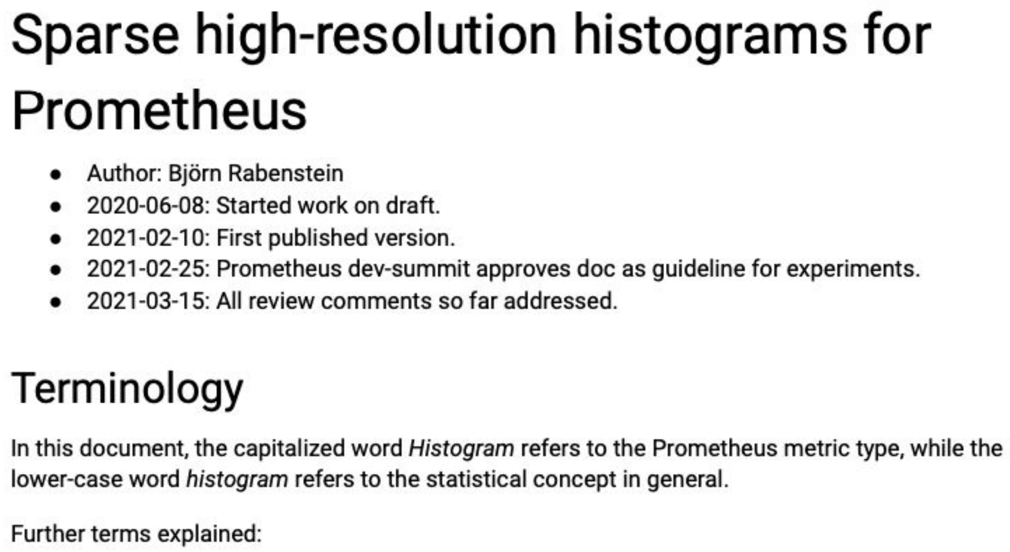
exact	Exact quantile computation based on numpy arrays [6].
prom	Quantile estimation based on Prometheus Histograms [4].
hdr	The HDR Hisotgram data-structure introduced in [5].
dd	The DDSketch data-structure introduced in [1].
t-digest	The t-digest data-structure introduced in [3].
circllhist	The circllhist data-structure described in this document.

Hartmann, Schlossnagle - Circllhist (https://arxiv.org/abs/2001.06561)

Summary

				Log-Histograms		
	Array	Prom. "Histogram"	t-Digest	HdR	DD	Circllhist
size	✗	✓	✓	✓	✓	✓
accuracy	✓	✗	✓	✓	✓	✓
performance	✓	✓	✓	✓	✓	✓
mergeability	✓	✓	?	✓	✓	✓
zero-conf.	✓	✗	O	✗	O	✓

Adoption in Open Source Technology



Sparse high-resolution histograms for Prometheus

- Author: Björn Rabenstein
- 2020-06-08: Started work on draft.
- 2021-02-10: First published version.
- 2021-02-25: Prometheus dev-summit approves doc as guideline for experiments.
- 2021-03-15: All review comments so far addressed.

Terminology

In this document, the capitalized word *Histogram* refers to the Prometheus metric type, while the lower-case word *histogram* refers to the statistical concept in general.

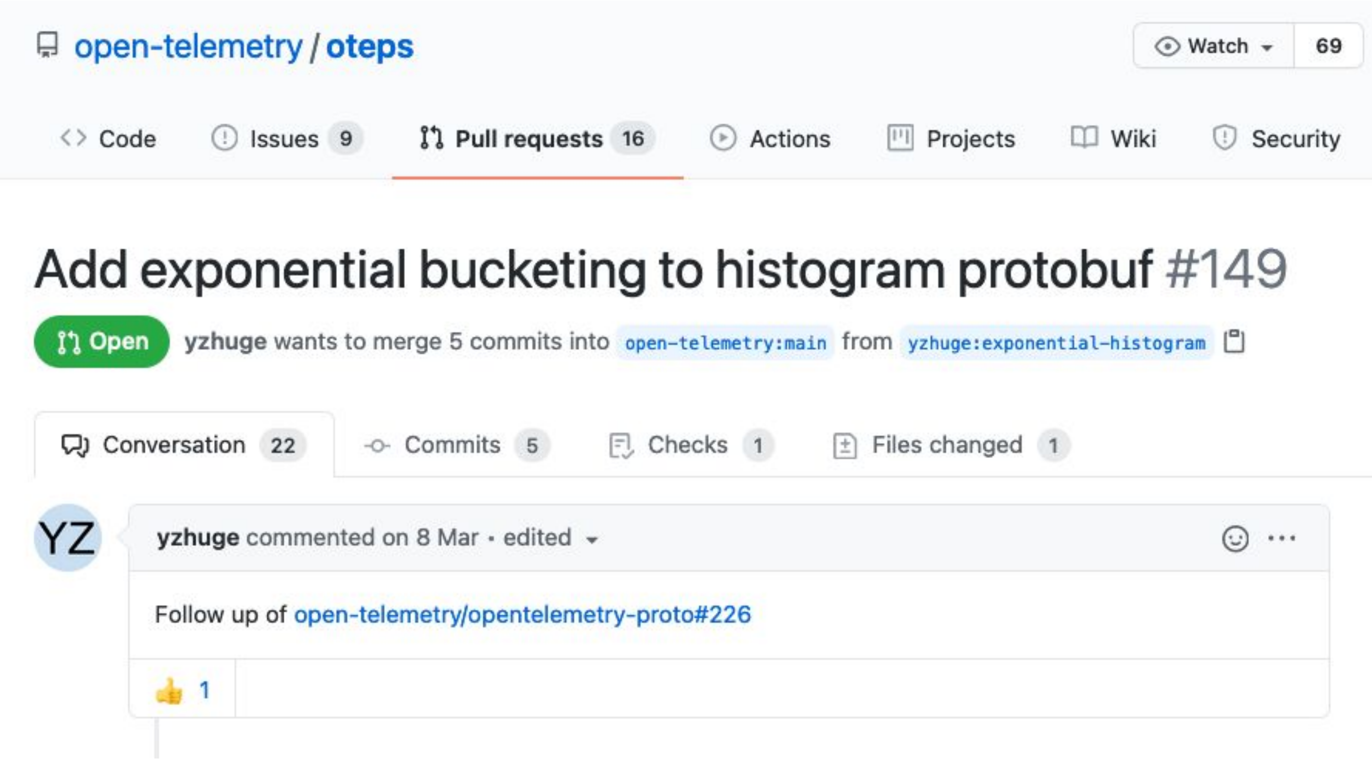
Further terms explained:

Björn Beorn Rabenstein
Principal Software Engineer at Grafana Labs

https://docs.google.com/document/d/1cLNv3aufPZb3fNfaJgdaRBZsInZKKIHo9E6HinJVbpM/edit

2021

OpenTelemetry



Add exponential bucketing to histogram protobuf #149

yzhuge wants to merge 5 commits into `open-telemetry:main` from `yzhuge:exponential-histogram`

yzhuge commented on 8 Mar • edited

Follow up of open-telemetry/opentelemetry-proto#226

Open Telemetry - GitHub

2021