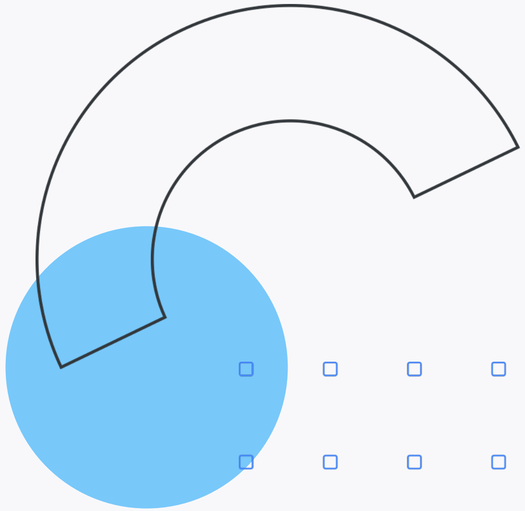
Databend

Databend 新 Hashtable 的设计和实现

主讲人：崔子龙

2022.11

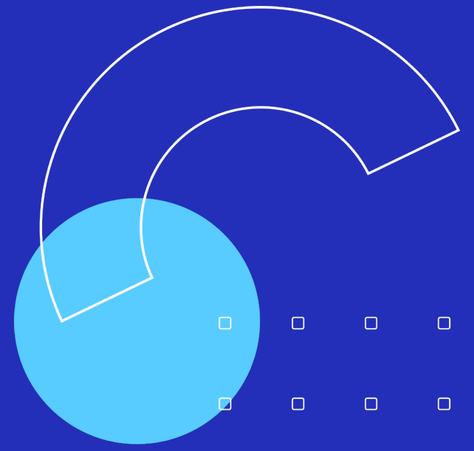


Data infra 研究社

自我介绍

Databend

- 崔子龙
  - 开源之夏2022“为 Databend 增加 String Adaptive Hash Table”项目参与者



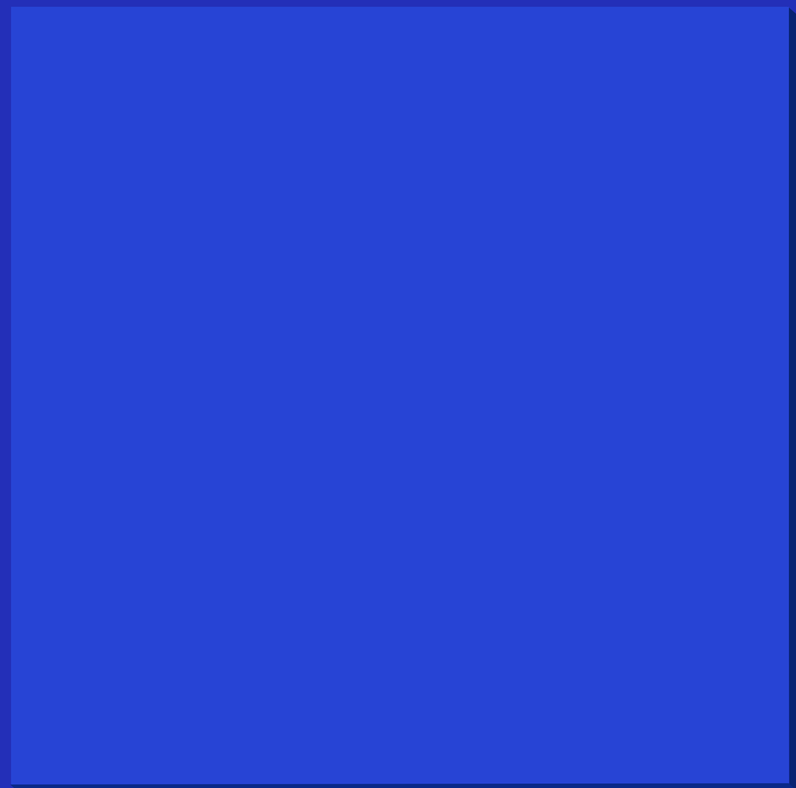
Databend

目录

CONTENTS

Databend

设计目标



Data infra 研究社

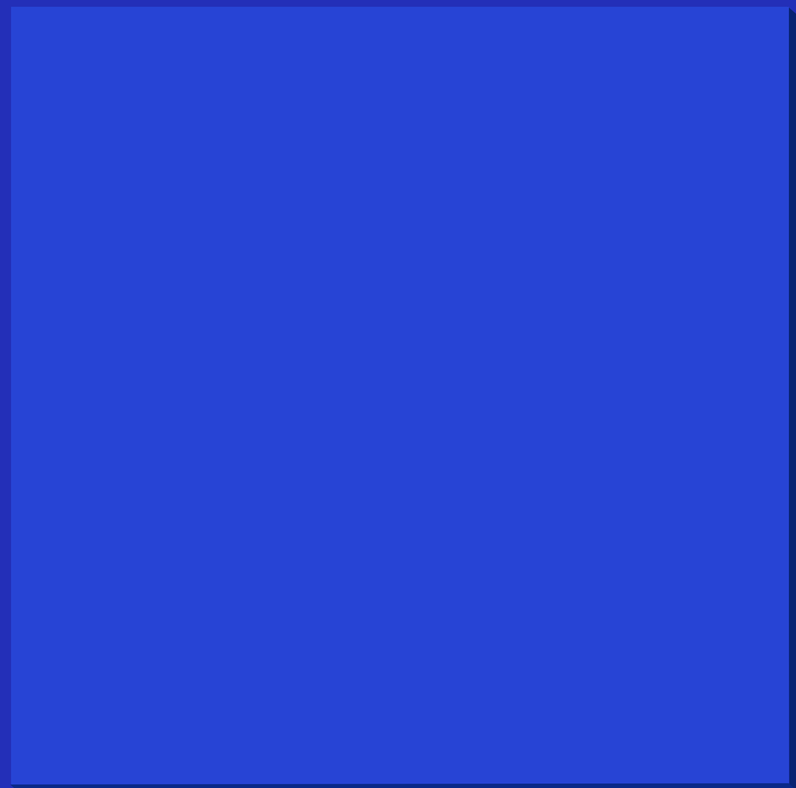
设计目标

Databend

- 仅需要插入、删除和迭代
- 键包括整数和字符串两类
- String Adaptive Hash Table：根据字符串键的长度优化哈希表
- 新的 API

Databend

实现细节

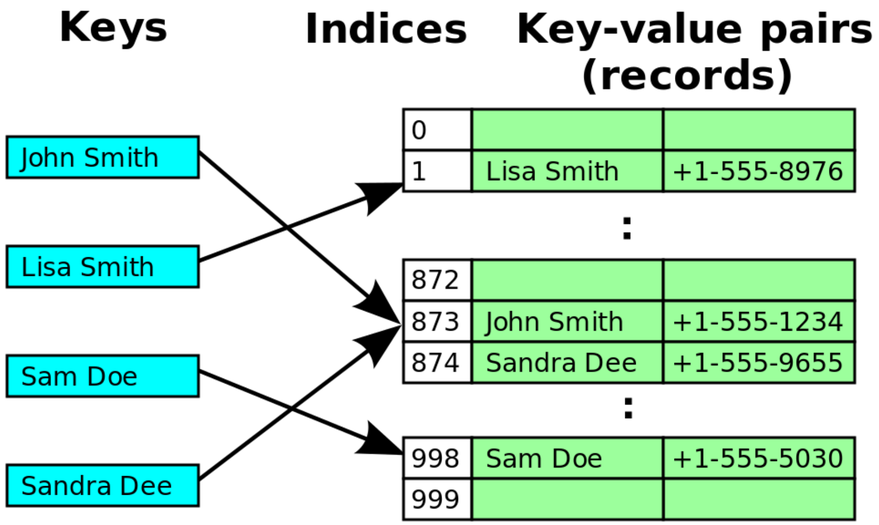


Data infra 研究社

碰撞处理

Databend

- 线性探测
- 二倍扩容



https://upload.wikimedia.org/wikipedia/commons/9/90/HASHTB12.svg

String Adaptive Hash Table

Databend

- https: //doi.org/10.3390/app10061915
- 根据字符串长度确定不同的存储方式。

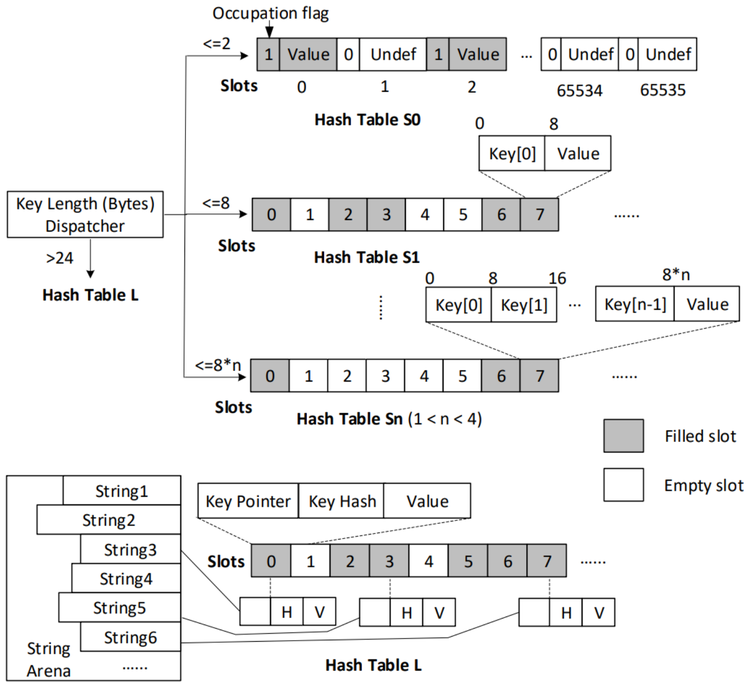


Figure 5. Architecture of SAHA with multiple hash tables to store strings with different length ranges.

哈希函数

Databend

- AHasher
- CRC
- 32 bit or 64 bit?

内存分配

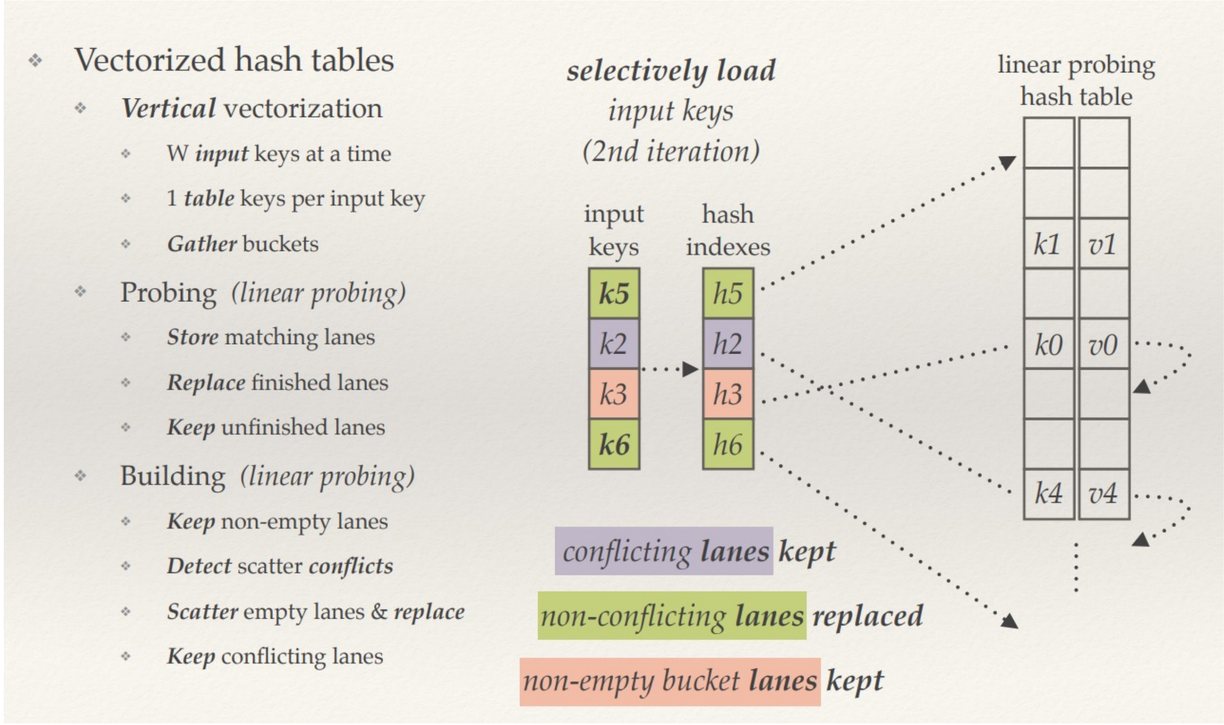
Databend

- mmap：MAP_POPULATE
- mremap：madvise, MADV_POPULATE_WRITE (since Linux Kernel 5.14)

向量化

Databend

- 起源：Rethinking SIMD Vectorization for In-Memory Databases，
  https://www.cs.columbia.edu/~orestis/sigmod15.pdf
- gather, scatter
- 冲突检测
- 停顿在gather上



http://www.cs.columbia.edu/~orestis/sigmod15slides.pdf

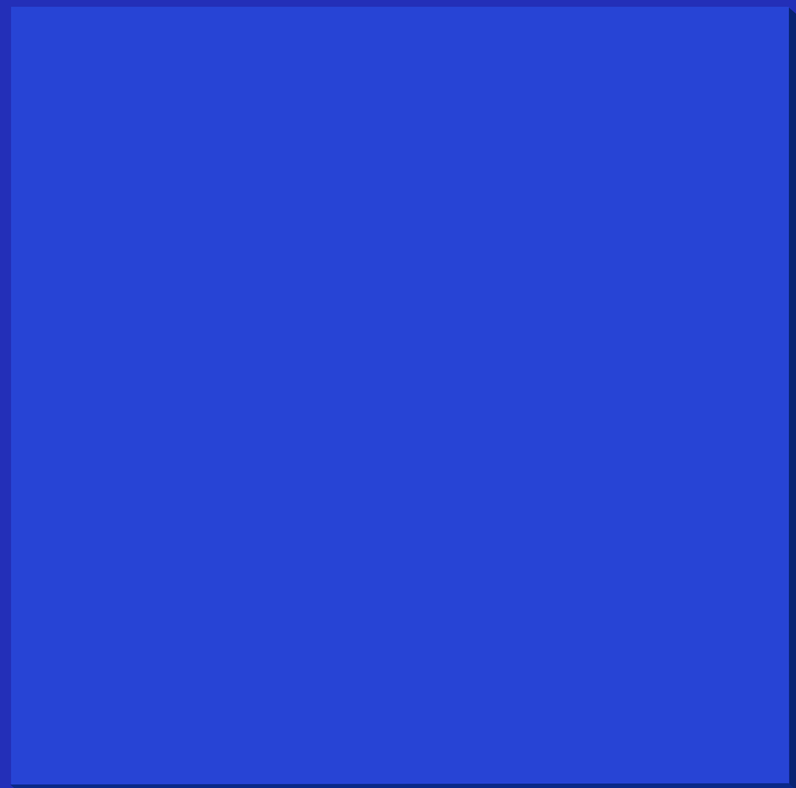
新 API

Databend

- 减少暴露裸指针，用生命周期保证安全性
- GAT

Databend

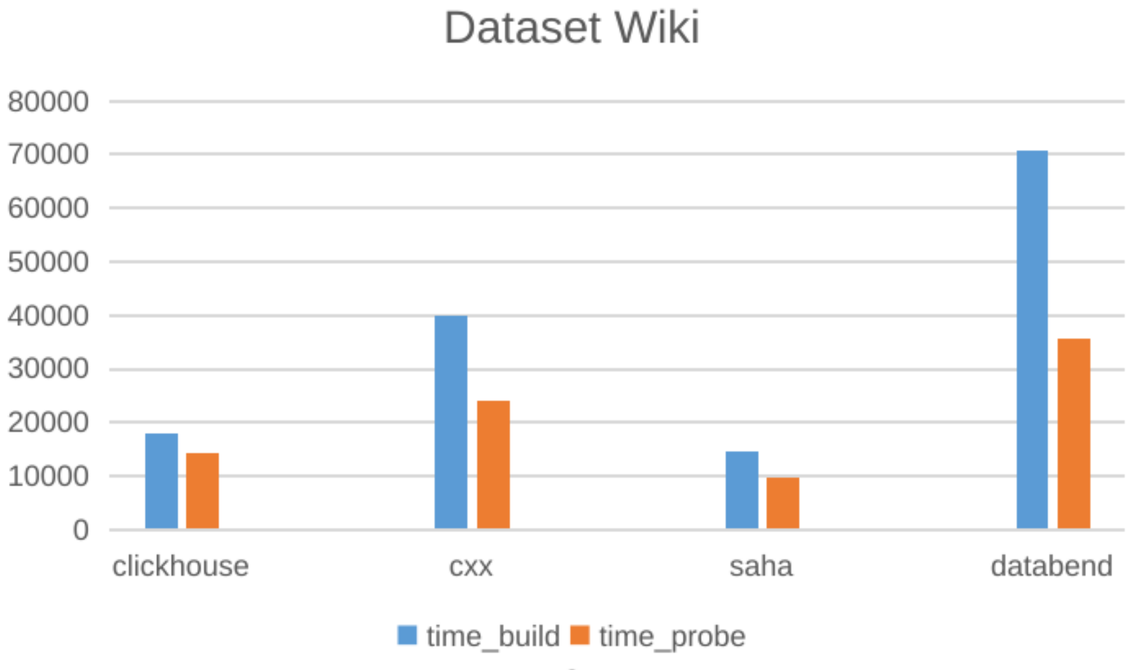
性能分析

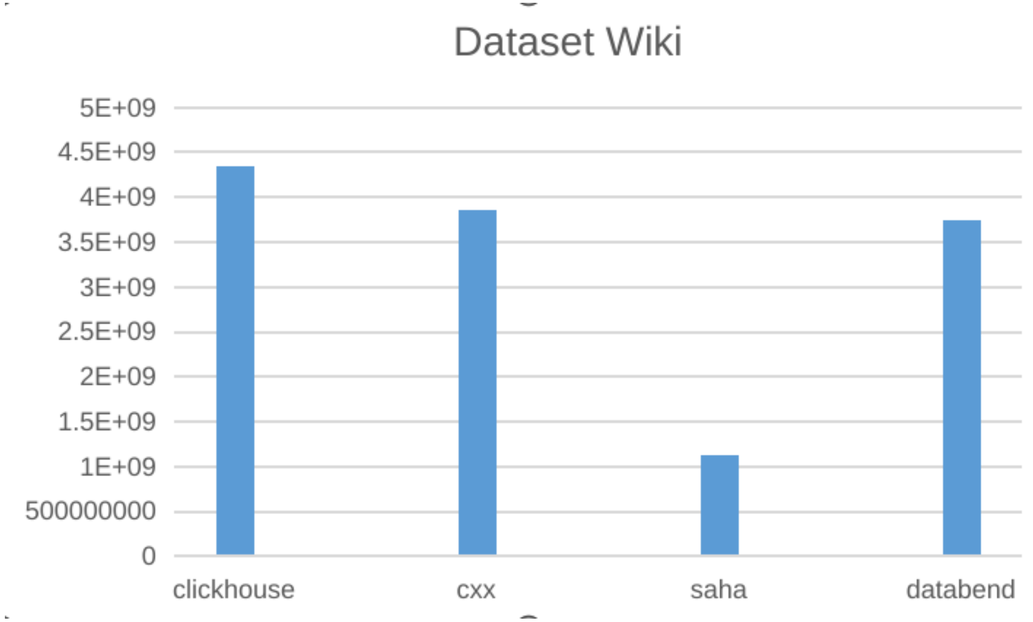


Data infra 研究社

性能分析

Databend

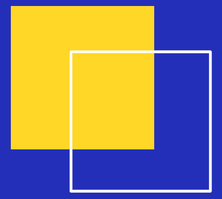




memory

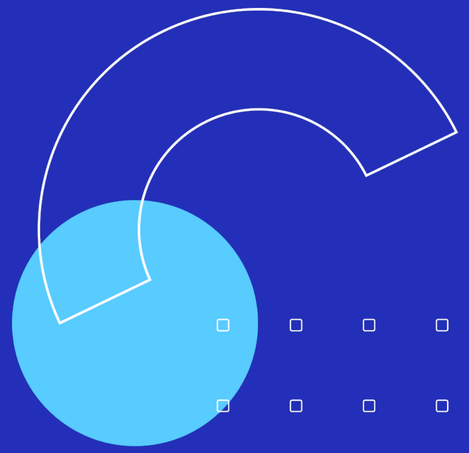
Databend

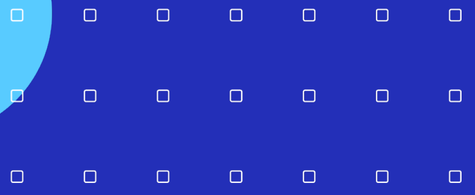
Q & A



Databend

THANKS!





Data infra 研究社