VectorCDC: Accelerating Data Deduplication with Vector Instructions

Sreeharsha Udayashankar, Abdelrahman Baba and Samer Al-Kiswany

USENIX FAST' 25

Introduction

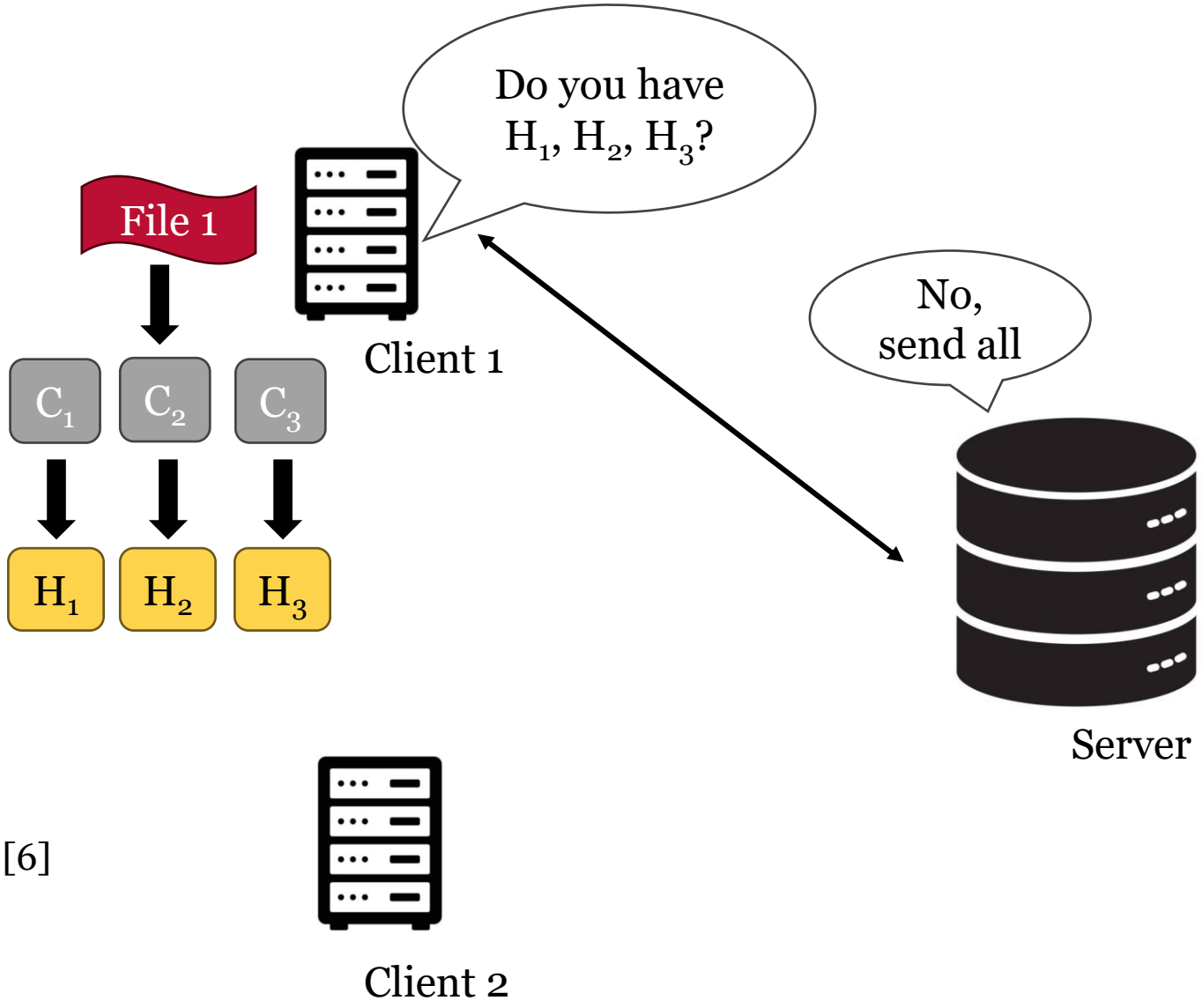
- Data explosion
  - Global data production expected to exceed 180 ZB by 2025 [1]
  - Cloud storage providers

- Mechanisms
  - Distributed file systems [2]
  - Storage Architectures [3]
  - Data Deduplication [4]

[1] Arne Holst. *Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2025*. Statista, 2021.
[2] Sanjay Ghemawat et al. *The Google File System*. SIGOPS Oper. Syst, 2003.
[3] Peter M Chen et al. *RAID: High-performance, reliable secondary storage*. ACM Computing Surveys (CSUR), 1994.
[4] Nagapramod Mandagere et al.. *Demystifying data deduplication*. ACM/IFIP/USENIX Middleware'08 Conference, 2008

Introduction

- Data Deduplication [5]
  - Identify and eliminate duplicate data
- Deduplication Overview
  - File Chunking and Hashing
  - Fingerprint Comparison
  - Data Storage
- Content-Defined Chunking (CDC) [6]
  - Hash-based and Hashless



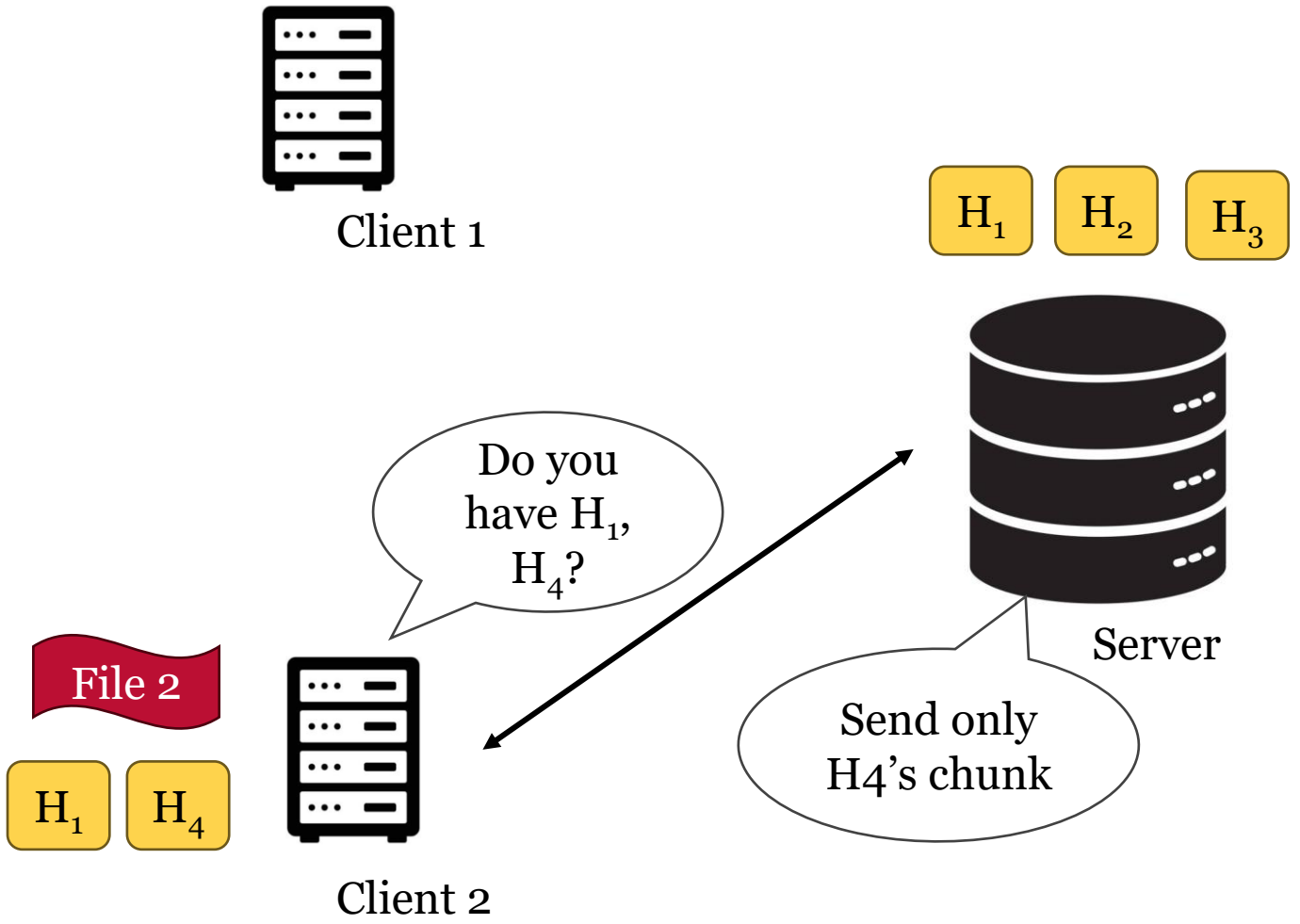
[5] Dutch T Meyer et al. *A study of practical deduplication*. ACM Transactions on Storage (ToS), 2012.

[6] Athicha Muthitacharoen et al. *A low-bandwidth network file system*. SOSP, 2001.

[7] Yucheng Zhang et al. *AE: An asymmetric extremum content defined chunking algorithm for fast and bandwidth-efficient data deduplication*. INFOCOM, 2015.

Introduction

- Data Deduplication [5]
  - Identify and eliminate duplicate data
- Deduplication Overview
  - File Chunking and Hashing
  - Fingerprint Comparison
  - Data Storage
- Content-Defined Chunking (CDC) [6]
  - Hash-based and Hashless



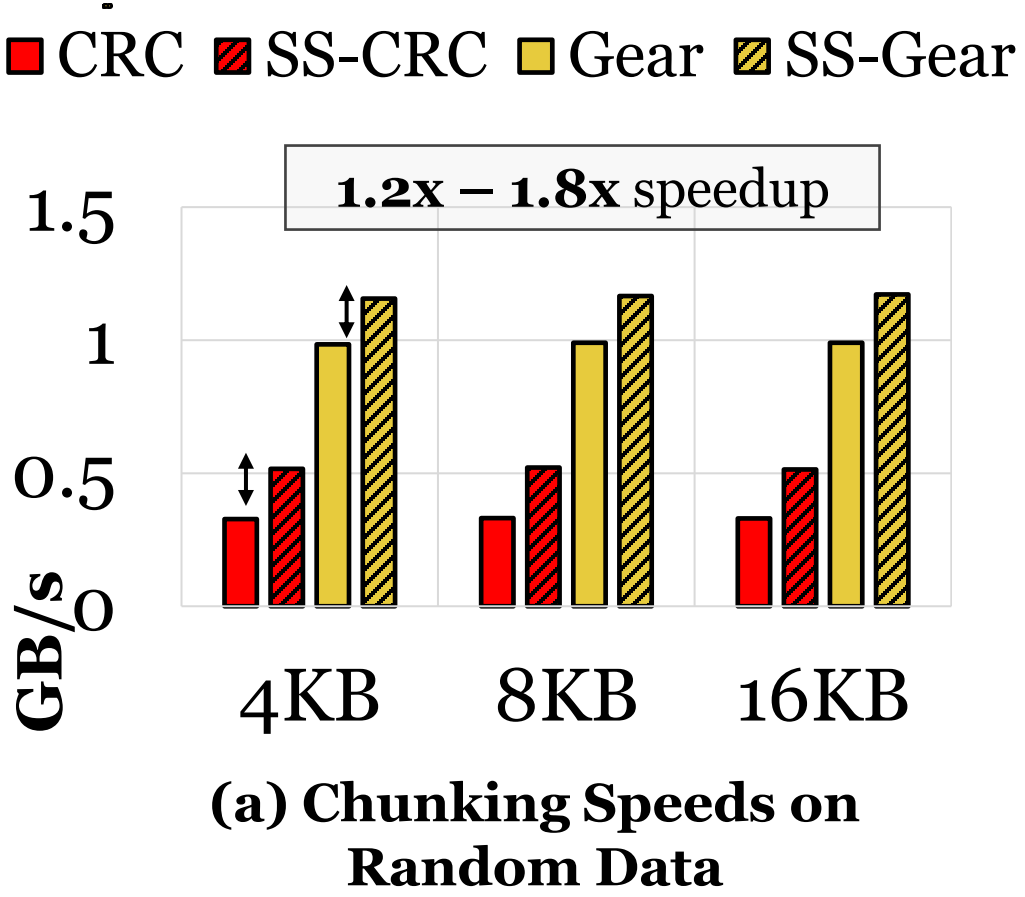
[5] Dutch T Meyer et al. *A study of practical deduplication*. ACM Transactions on Storage (ToS), 2012.

[6] Athicha Muthitacharoen et al. *A low-bandwidth network file system*. SOSP, 2001.

[7] Yucheng Zhang et al. *AE: An asymmetric extremum content defined chunking algorithm for fast and bandwidth-efficient data deduplication*. INFOCOM, 2015.

Motivation – Vector Accelerated CDC

- CDC is computationally intensive
  - **Idea:** Accelerate with SIMD (AVX / SSE) CPU instructions
- Existing approaches
  - SS-CDC [8]
  - Low speedups despite using AVX-512 instructions

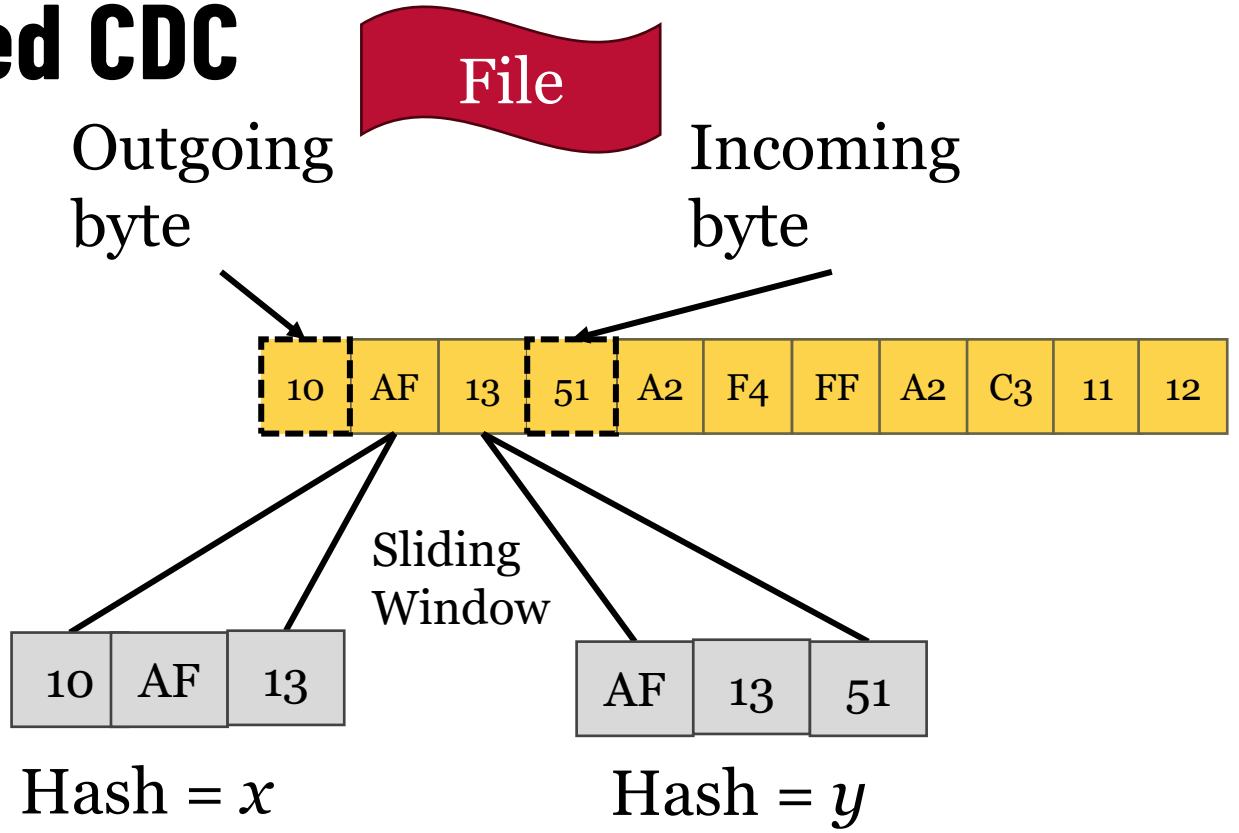


(a) Chunking Speeds on Random Data

[8] Fan Ni et al. *SS-CDC: A two-stage parallel content-defined chunking for deduplicating backup storage*. SYSTOR, 2019.

Motivation – Vector Accelerated CDC

- Fundamental inefficiencies
  - Rolling hash algorithms [8]
- Dependency between adjacent bytes
  - *Solution:* Process different regions of the file with SIMD
  - Expensive scatter/gather instructions



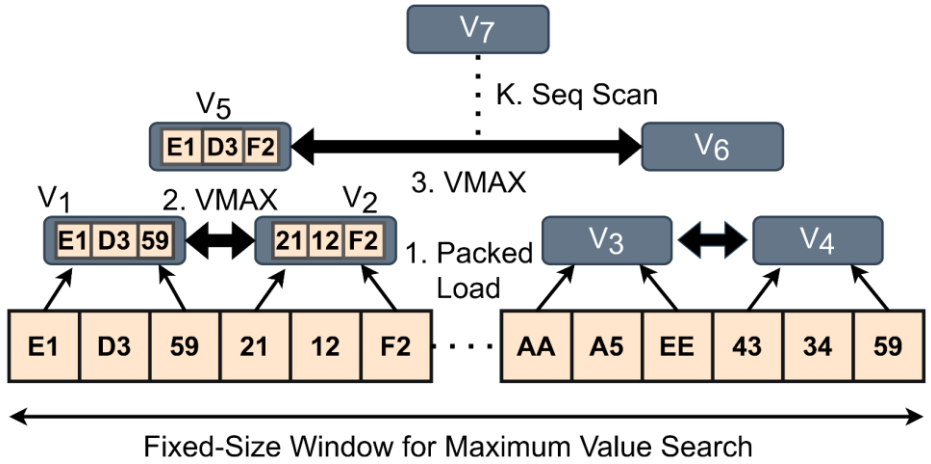
$$y = x + f(51) - g(10)$$

$f(x)$ and $g(x)$ are functions

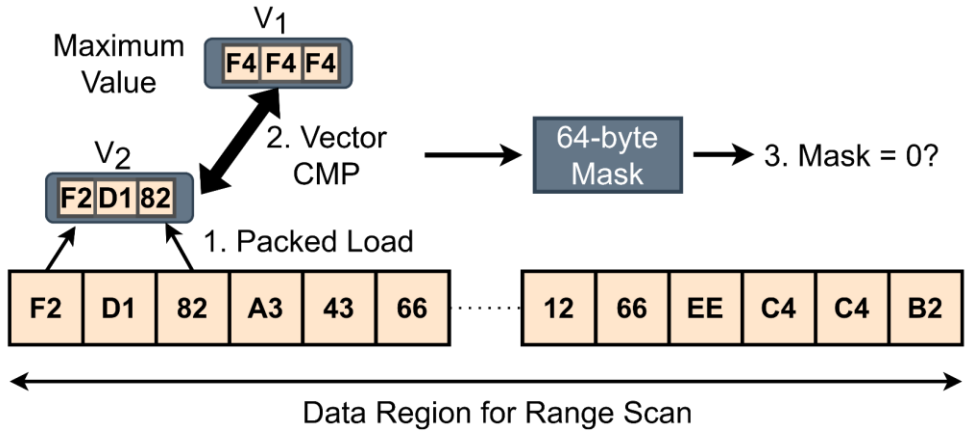
[8] Fan Ni et al. *SS-CDC: A two-stage parallel content-defined chunking for deduplicating backup storage*. SYSTOR, 2019.

VectorCDC

- New vector acceleration method for hashless CDC
  - Use AVX-friendly *tree-based search* and *packed scanning*
    - Compatible with a wide range of existing hashless CDC
  - **21x higher throughput** over SS-CDC
    - No impact on deduplication space savings



a) Tree-based Search



b) Packed Scanning

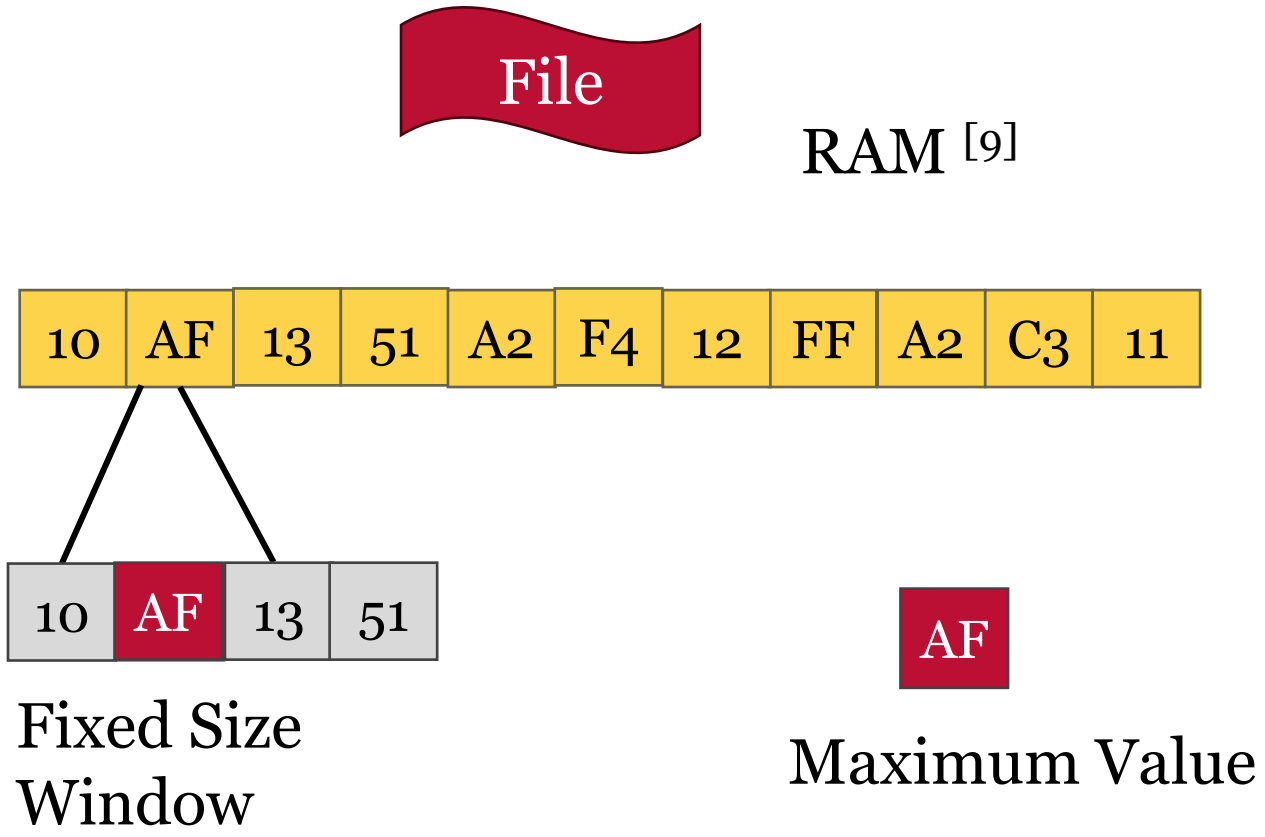
Outline

- Introduction
- Background
  - Hashless CDC
  - Vector Instructions
- Design
- Evaluation
- Conclusion

UNIVERSITY OF WATERLOO

Background - Hashless CDC

- Hashless CDC
  - AE [7]
  - RAM [9]
  - MAXP [10]



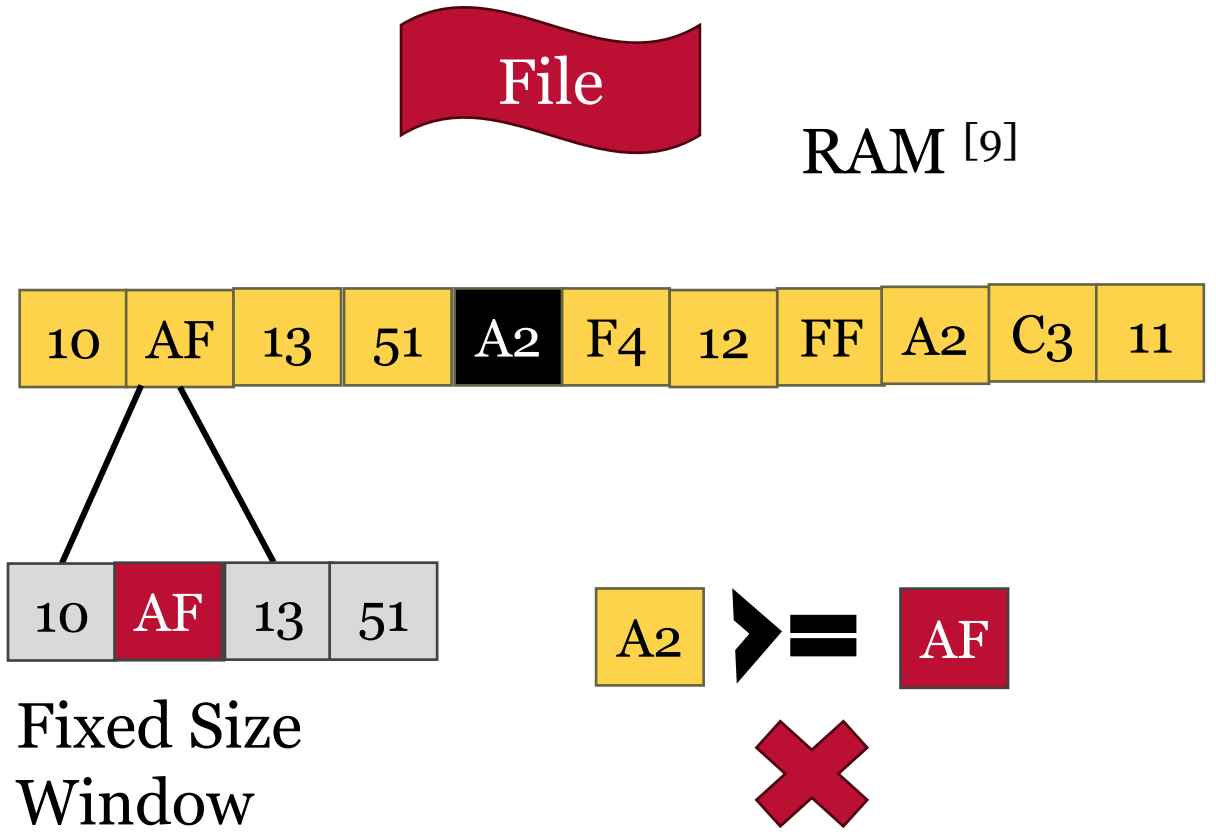
[7] Yucheng Zhang et al. *AE: An asymmetric extremum content defined chunking algorithm for fast and bandwidth-efficient data deduplication*. INFOCOM, 2015.

[9] Ryan Widodo et al. *A new content-defined chunking algorithm for data deduplication in cloud storage*. Future Generation Computer Systems, 2017

[10] Nikolaj Bjørner et al. *Content-dependent chunking for differential compression, the local maximum approach*. Journal of Computer and System Sciences, 2010

Background - Hashless CDC

- Hashless CDC
  - AE [7]
  - RAM [9]
  - MAXP [10]



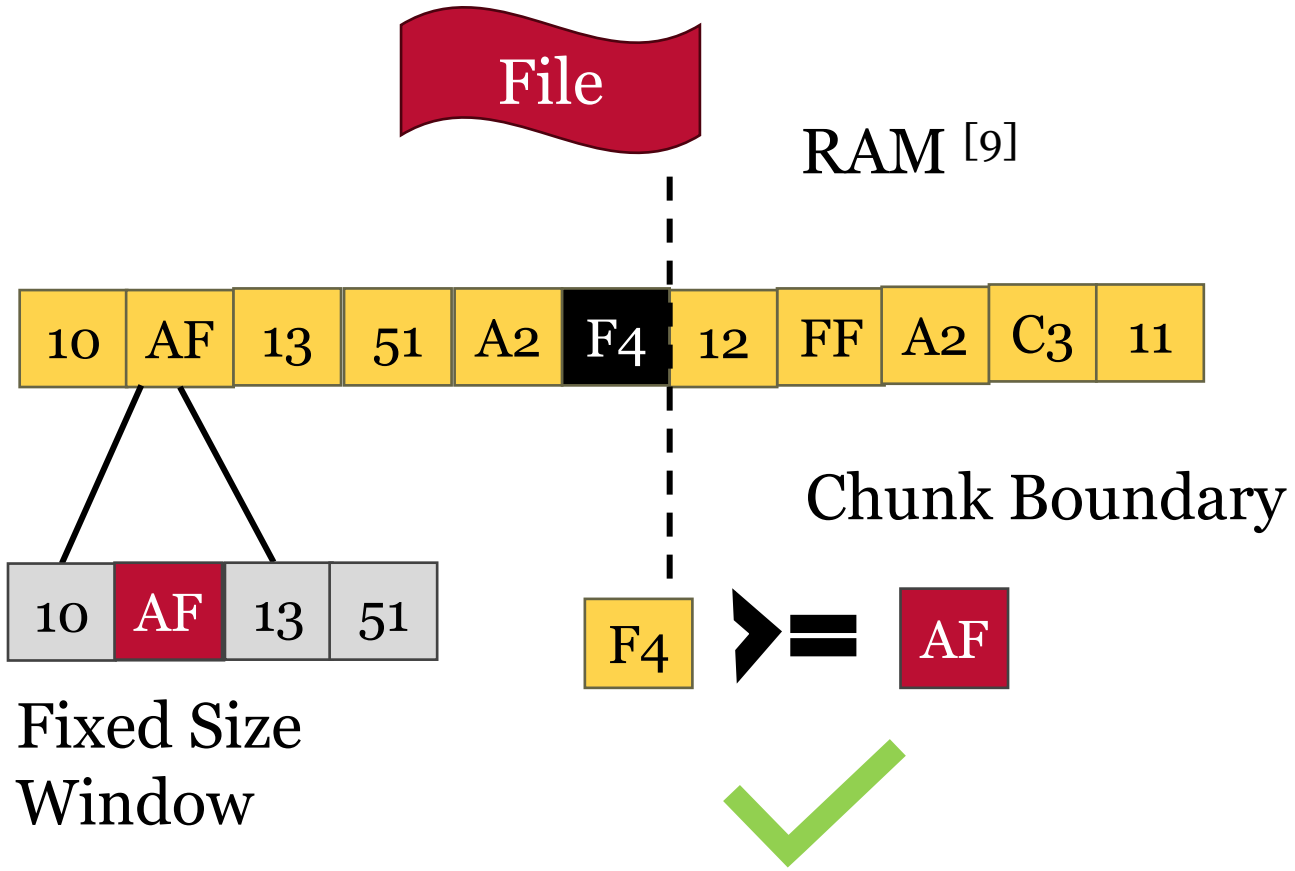
[7] Yucheng Zhang et al. *AE: An asymmetric extremum content defined chunking algorithm for fast and bandwidth-efficient data deduplication*. INFOCOM, 2015.

[9] Ryan Widodo et al. *A new content-defined chunking algorithm for data deduplication in cloud storage*. Future Generation Computer Systems, 2017

[10] Nikolaj Bjørner et al. *Content-dependent chunking for differential compression, the local maximum approach*. Journal of Computer and System Sciences, 2010

Background - Hashless CDC

- Hashless CDC
  - AE [7]
  - RAM [9]
  - MAXP [10]



[7] Yucheng Zhang et al. *AE: An asymmetric extremum content defined chunking algorithm for fast and bandwidth-efficient data deduplication*. INFOCOM, 2015.

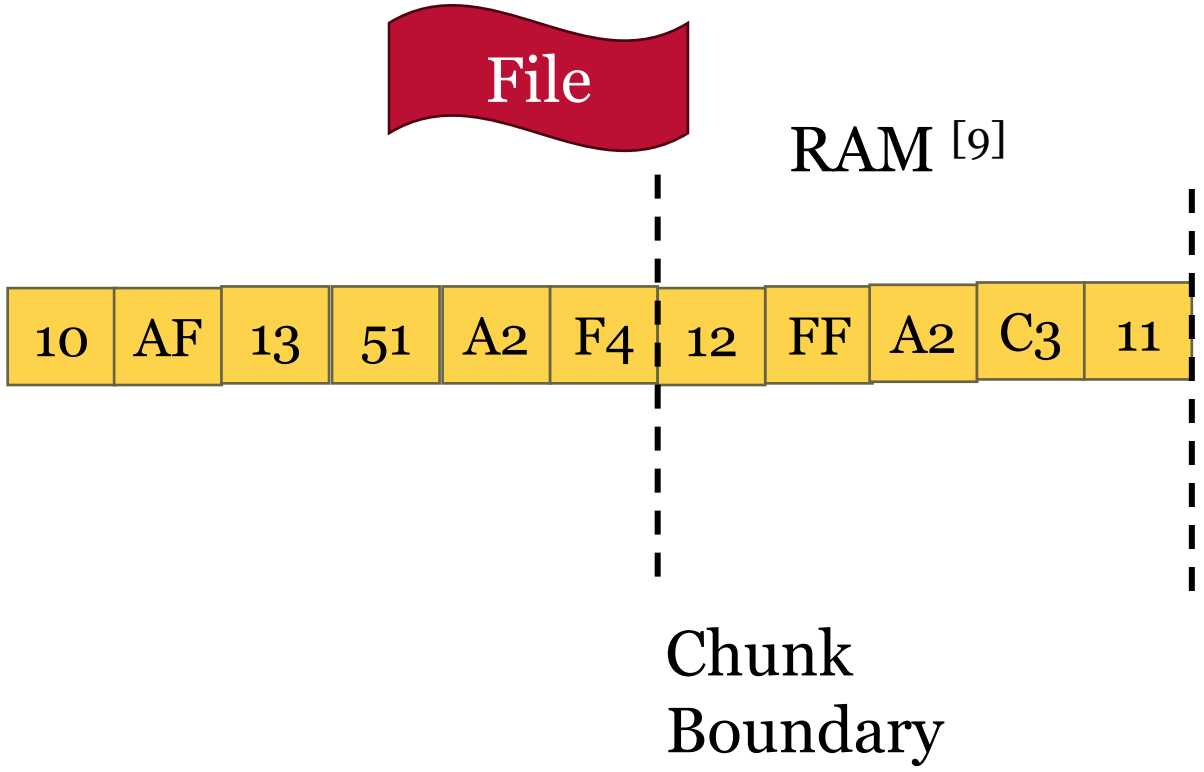
[9] Ryan Widodo et al. *A new content-defined chunking algorithm for data deduplication in cloud storage*. Future Generation Computer Systems, 2017

[10] Nikolaj Bjørner et al. *Content-dependent chunking for differential compression, the local maximum approach*. Journal of Computer and System Sciences, 2010

UNIVERSITY OF WATERLOO

Background - Hashless CDC

- Hashless CDC
  - AE [7]
  - RAM [9]
  - MAXP [10]

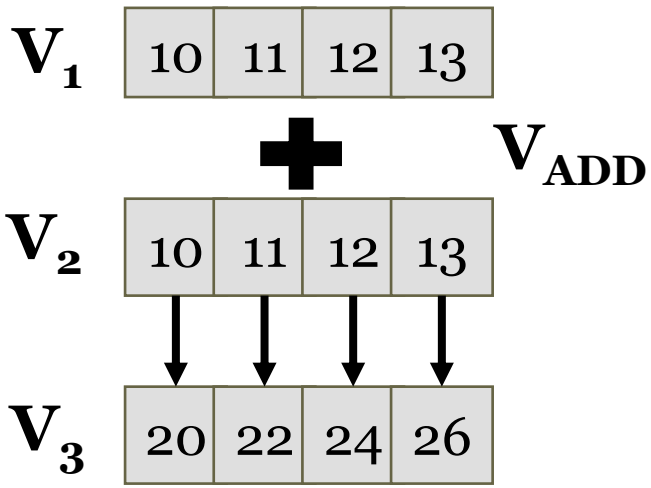


[7] Yucheng Zhang et al. *AE: An asymmetric extremum content defined chunking algorithm for fast and bandwidth-efficient data deduplication*. INFOCOM, 2015.
[9] Ryan Widodo et al. *A new content-defined chunking algorithm for data deduplication in cloud storage*. Future Generation Computer Systems, 2017
[10] Nikolaj Bjørner et al. *Content-dependent chunking for differential compression, the local maximum approach*. Journal of Computer and System Sciences, 2010

UNIVERSITY OF WATERLOO

Background – Vector Instructions

- Special CPU instructions with SIMD capabilities [11]
  - Used in math / multimedia applications
- Vector registers
  - 128 – 512 bits (16 – 64 bytes) wide
    - SSE-128
    - AVX-256
    - AVX-512



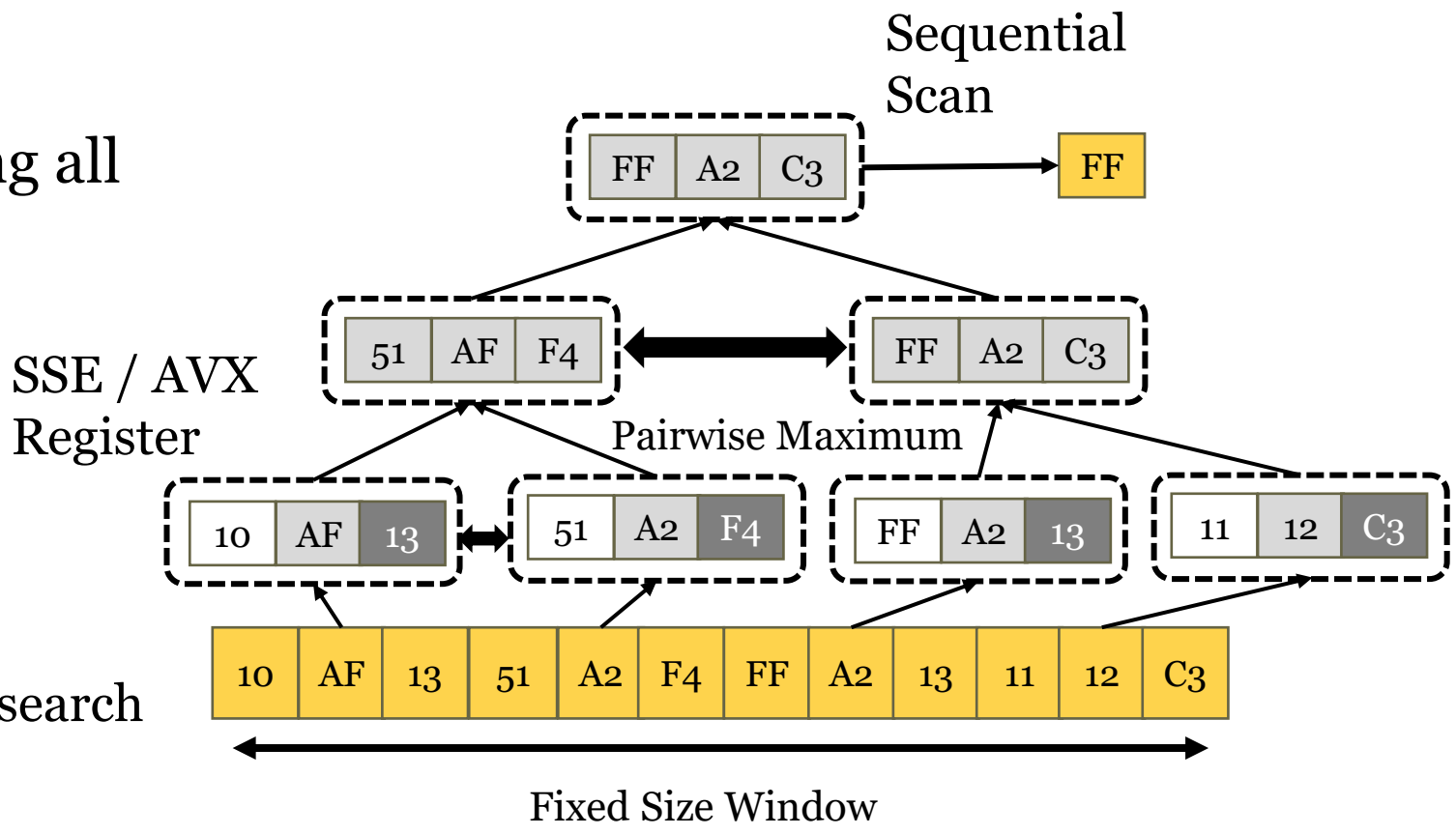
[11] James Smith et al. *Vector instruction set support for conditional operations*. ACM SIGARCH Computer Architecture News, 2000

Outline

- Introduction
- Background
- Design
- Evaluation
- Conclusion

VectorCDC Design

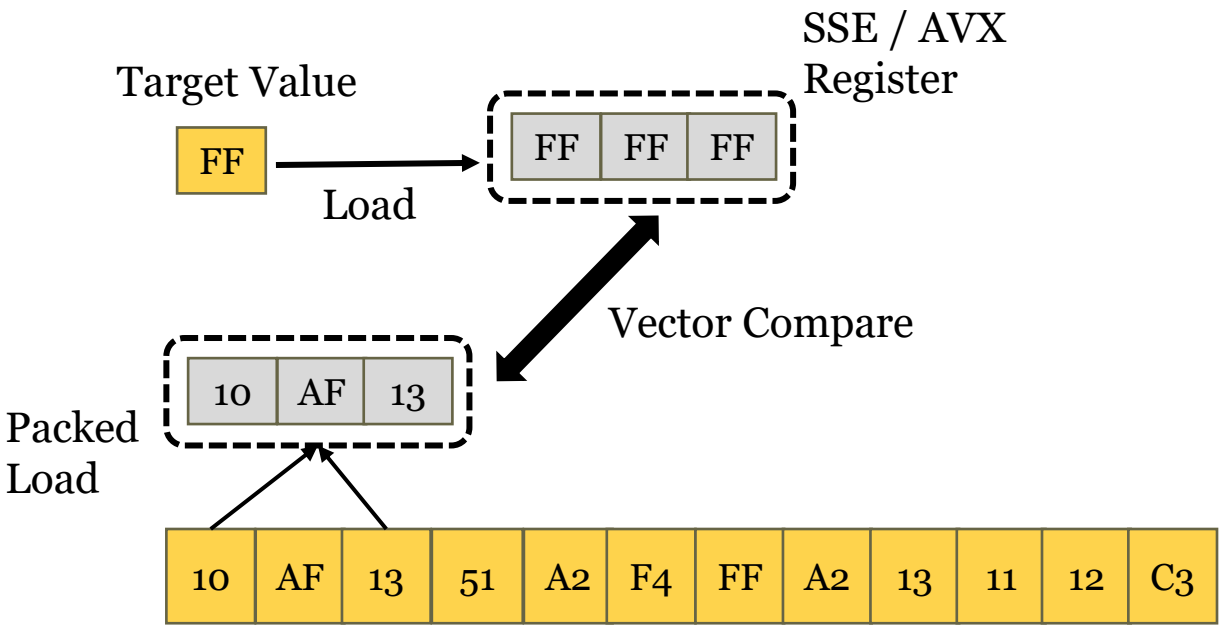
- Identify common phases among all hashless CDC algorithms
  - Extreme Byte Search
  - Range Scan
- Extreme Byte Search
  - Accelerate with novel tree-based search
  - Takes advantage of pipelining



Extreme Byte Search for maximum value

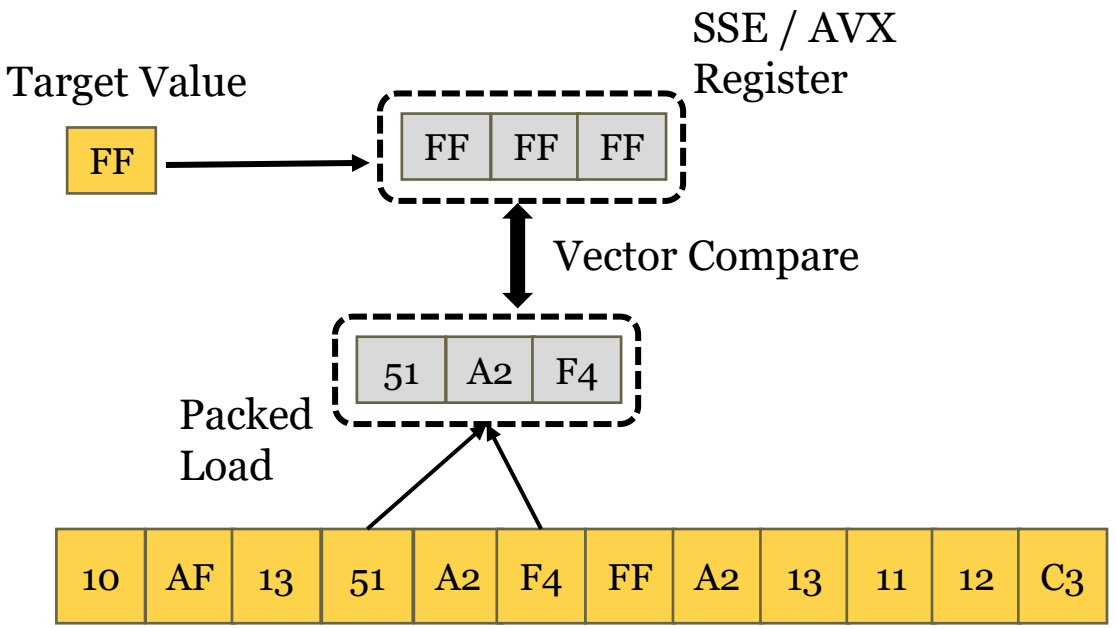
VectorCDC Design

- Identify common phases among all hashless CDC algorithms
  - Extreme Byte Search
  - Range Scan
- Extreme byte search
  - Accelerate with novel tree-based search
  - Takes advantage of pipelining
- Range Scan
  - Packed Scanning



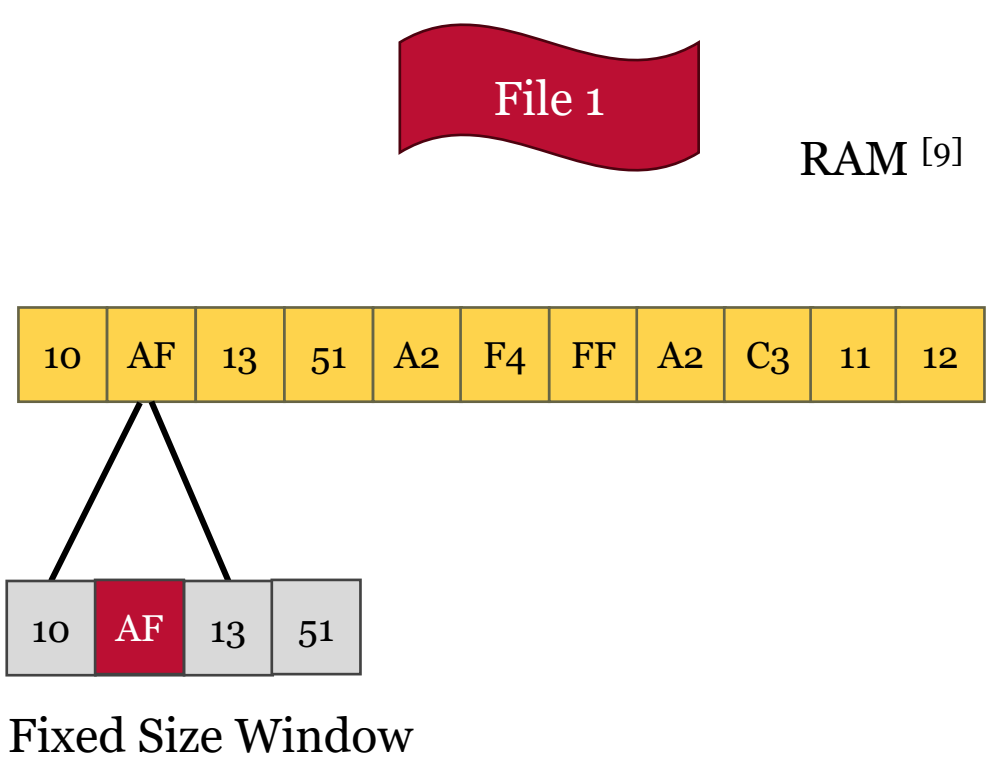
VectorCDC Design

- Identify common phases among all hashless CDC algorithms
  - Extreme Byte Search
  - Range Scan
- Extreme byte search
  - Accelerate with novel tree-based search
  - Takes advantage of pipelining
- Range Scan
  - Packed Scanning

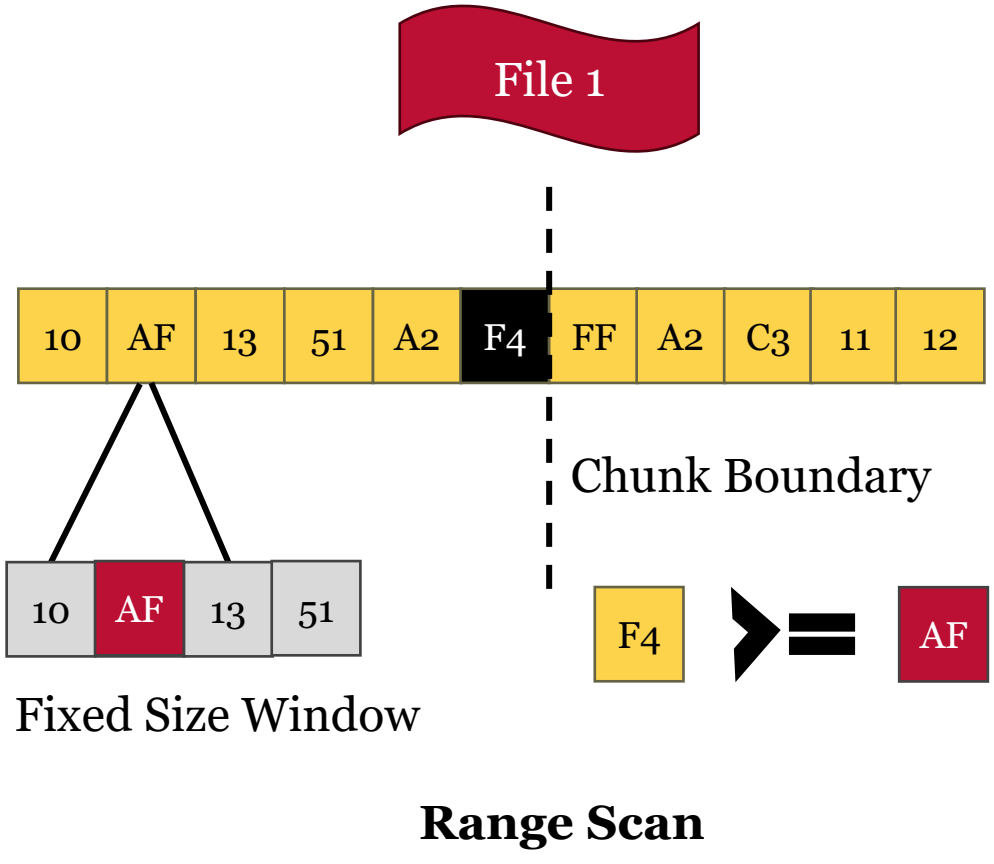


UNIVERSITY OF WATERLOO

Accelerating RAM with VectorCDC



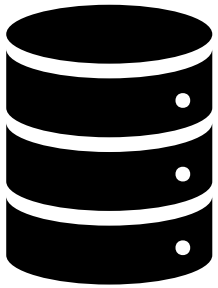
Extreme Byte Search



Range Scan

[9] Ryan Widodo et al. *A new content-defined chunking algorithm for data deduplication in cloud storage*. Future Generation Computer Systems, 2017

Evaluation



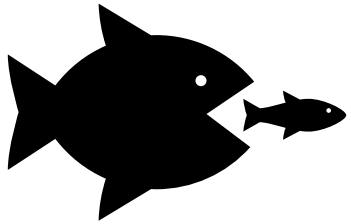
Datasets

Metrics

Space Savings

Speed / Throughput

Backward Compatibility



Alternatives

Hashless

AE [7]

RAM [9]

Hash-based

FastCDC [13]

Rabin's Chunking [6]

SS-CDC [8]

TTTD [12]

[6] Athicha Muthitacharoen et al. *A low-bandwidth network file system*. SOSP, 2001.

[7] Yucheng Zhang et al. *AE: An asymmetric extremum content defined chunking algorithm for fast and bandwidth-efficient data deduplication*. INFOCOM, 2015.

[8] Fan Ni et al. *SS-CDC: A two-stage parallel content-defined chunking for deduplicating backup storage*. SYSTOR, 2019

[9] Ryan NS Widodo et al. *A new content-defined chunking algorithm for data deduplication in cloud storage*. Future Generation Computer Systems, 2017.

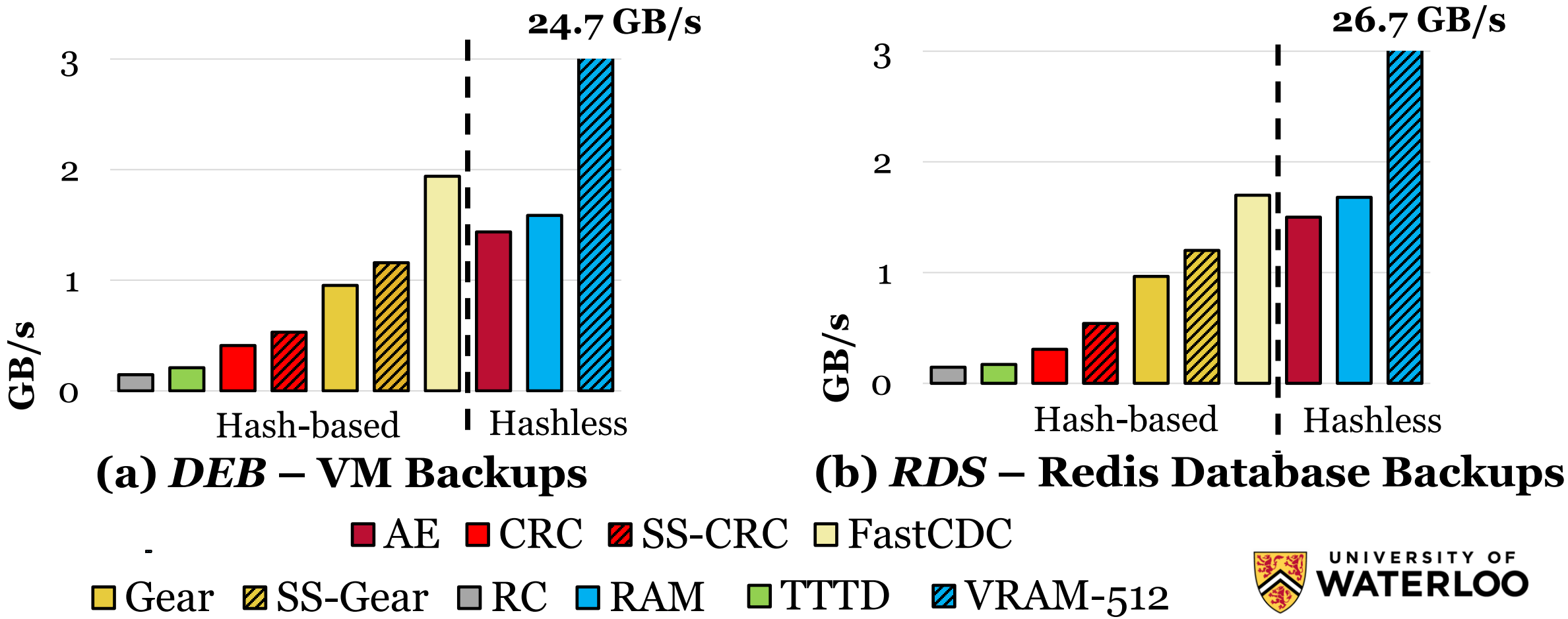
[12] Kave Eshghi et al. *A framework for analyzing and improving content-based chunking algorithms*. Hewlett-Packard Labs Technical Report, 2005

[13] Wen Xia et al. *FastCDC: A fast and efficient content-defined chunking approach for data deduplication*. USENIX ATC, 2016.

Chunking Throughput

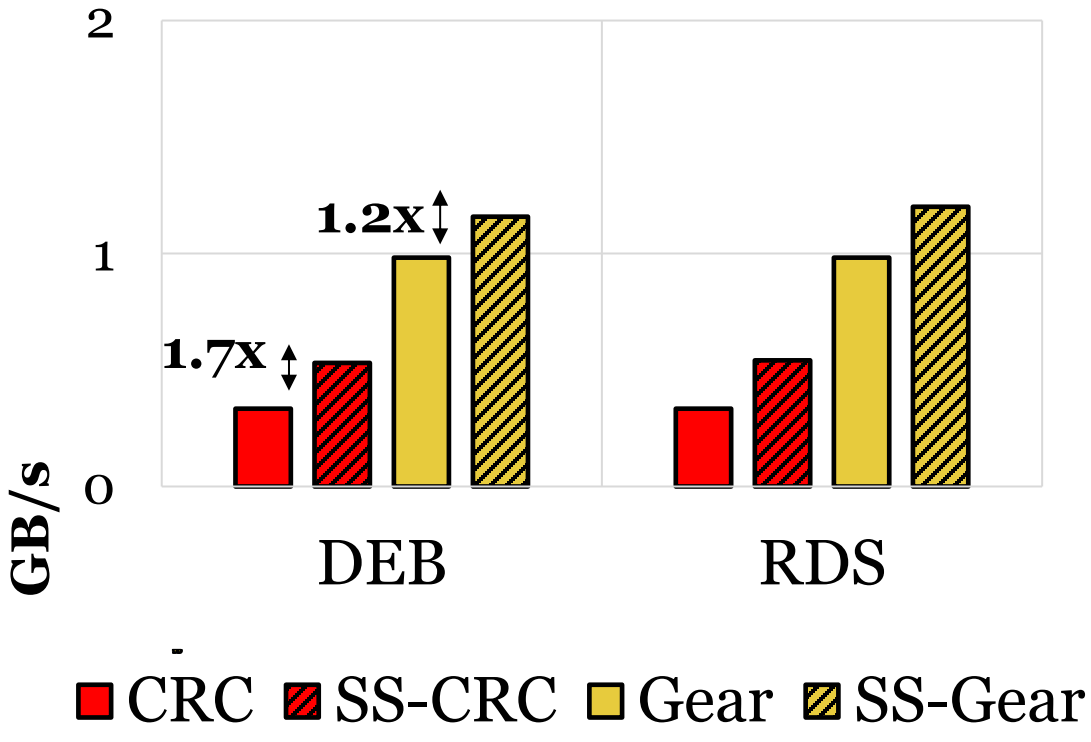
- Configuration: 8 KB chunks

1. *VRAM* is 12-15x faster than alternatives!

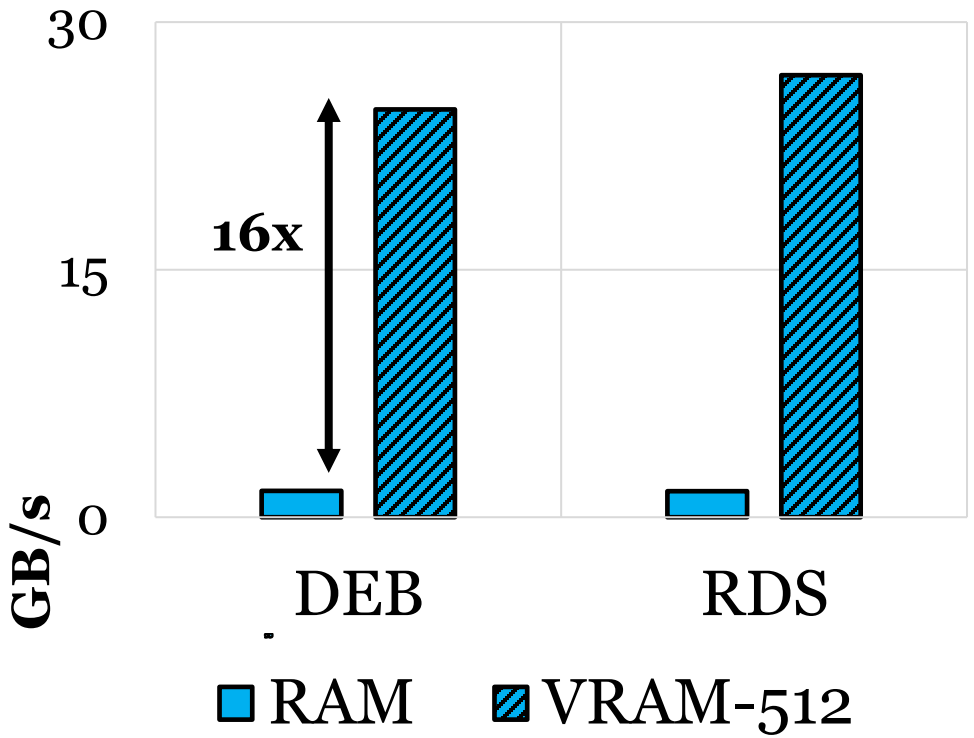


Speedups

2. VectorCDC achieves higher speedups than SS-CDC!



(a) Speedups with SS-CDC



(b) Speedups with VectorCDC

3. *VRAM* is 21x faster than *SS-Gear*!

UNIVERSITY OF WATERLOO

Summary

- Data deduplication is used to improve storage efficiency
  - Content-defined chunking algorithms critical to system performance
- VectorCDC
  - Redesign hashless CDC with SSE/AVX-friendly techniques
  - **21x** higher throughput than state-of-the-art vector accelerated CDC
  - No negative space savings impact
- **Code:** https://github.com/UWASL/dedup-bench

UNIVERSITY OF WATERLOO


UNIVERSITY OF
WATERLOO