

Metadata Considered Harmful ... to Deduplication

Xing Lin, Fred Douglass, Jim Li, Xudong Li,
Robert Ricci, Stephen Smaldone, and Grant Wallace

University of Utah, EMC Corporation, Nankai University



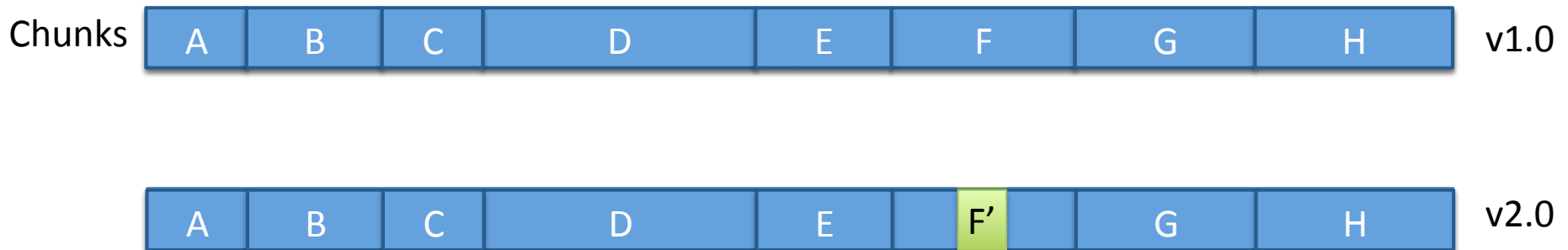
Deduplication

Idea: identify duplicate data blocks and store a single copy



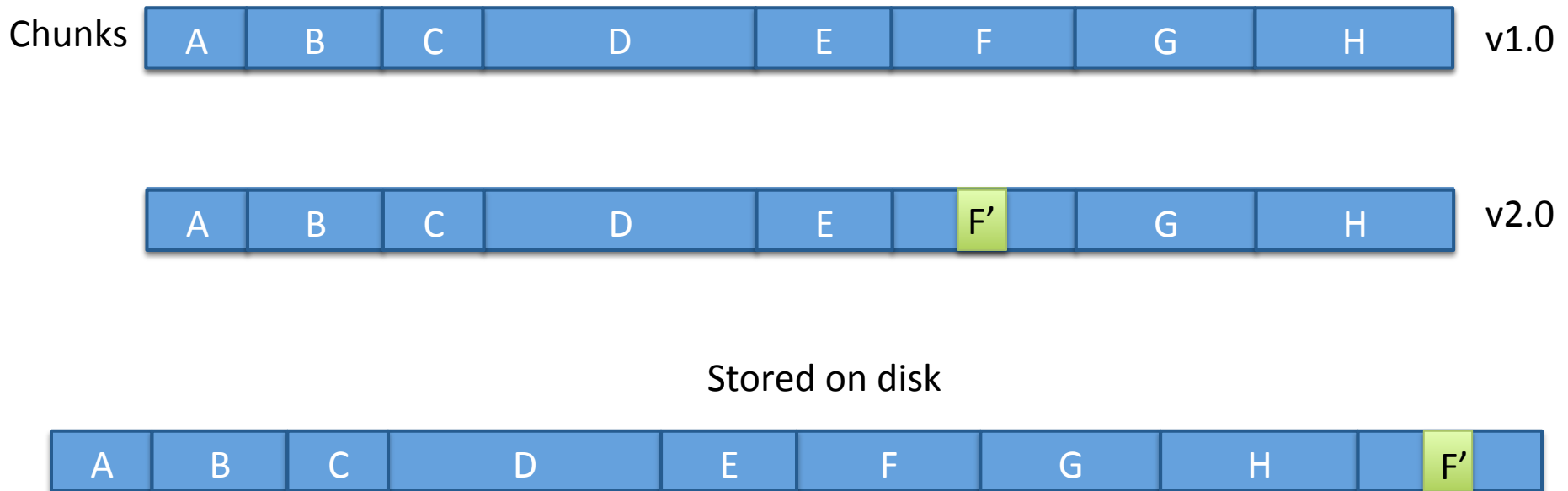
Deduplication

Idea: identify duplicate data blocks and store a single copy



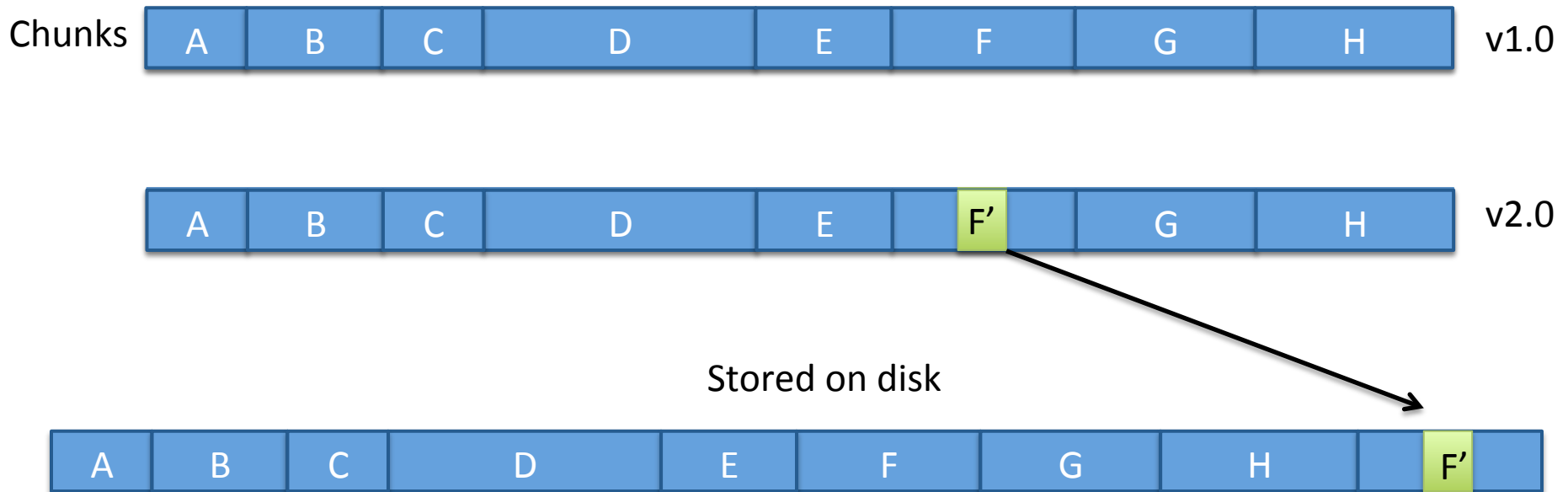
Deduplication

Idea: identify duplicate data blocks and store a single copy



Deduplication

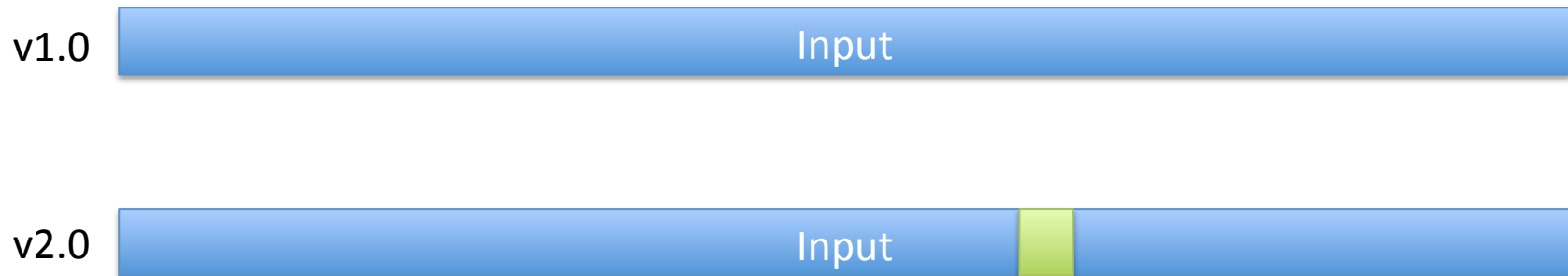
Idea: identify duplicate data blocks and store a single copy



Eliminate nearly all of v2.0 on disk

Expected Deduplication

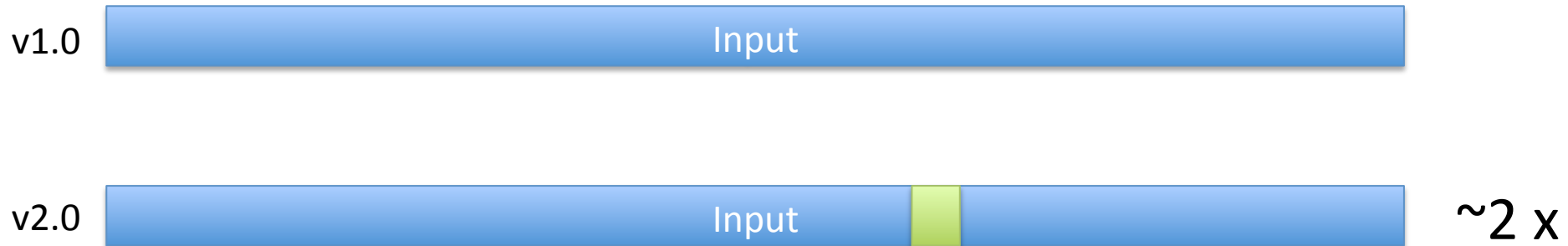
What if we have many versions?



Expected Deduplication

What if we have many versions?

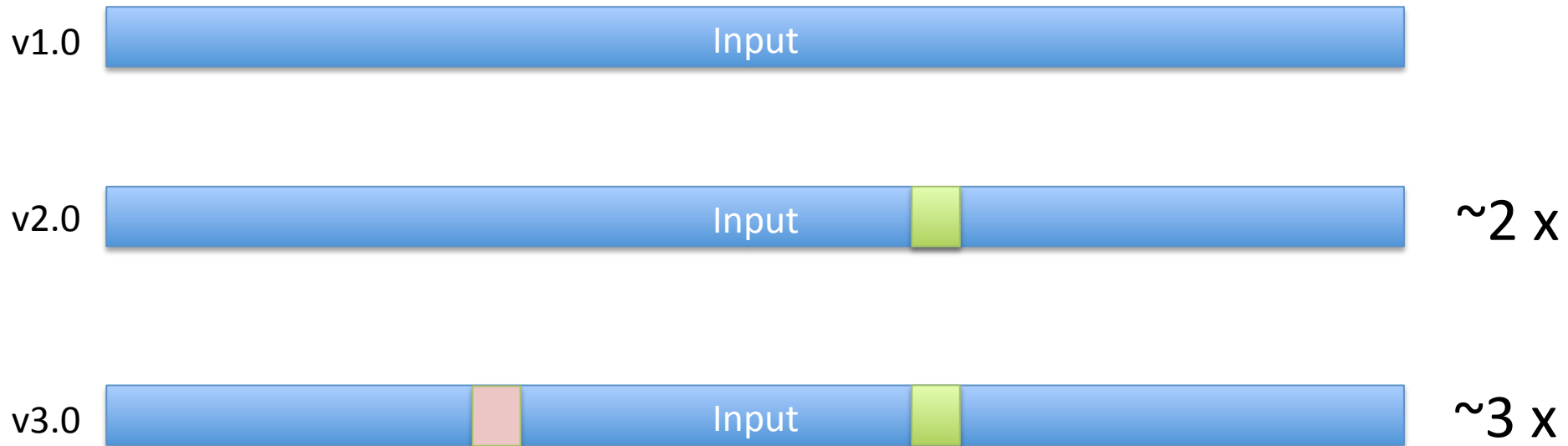
Dedup ratio = original_size/post_dedup_size



Expected Deduplication

What if we have many versions?

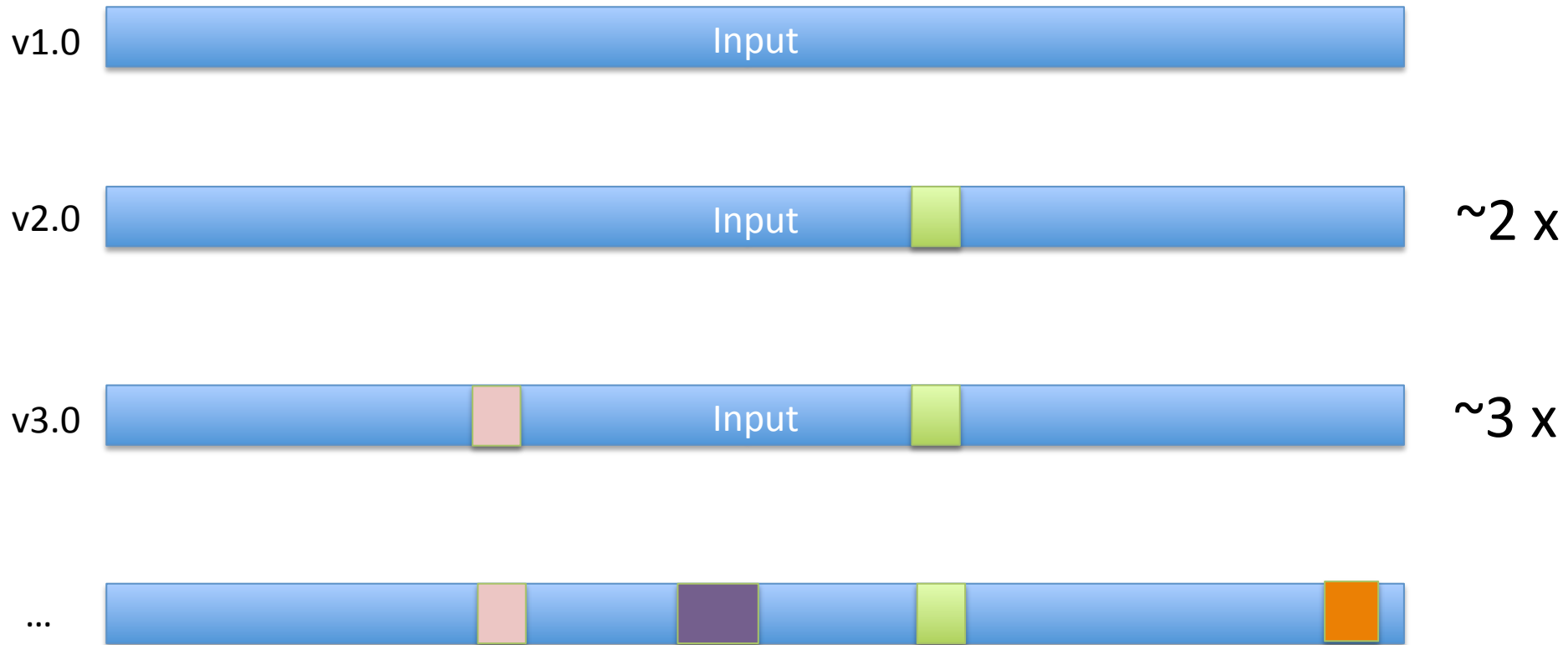
Dedup ratio = original_size/post_dedup_size



Expected Deduplication

What if we have many versions?

Dedup ratio = original_size/post_dedup_size



Great (Deduplication) Expectations

- ***In Reality***
 - **308** versions of Linux source code: **2 x**
 - Other examples of awful deduplication

Great (Deduplication) Expectations

- ***In Reality***

- **308** versions of Linux source code: **2 x**
- Other examples of awful deduplication

- ***Contributions:***

- Identify and categorize barriers to deduplication
- Solutions
 - EMC Data Domain (industrial experience)
 - GNU tar (academic research)

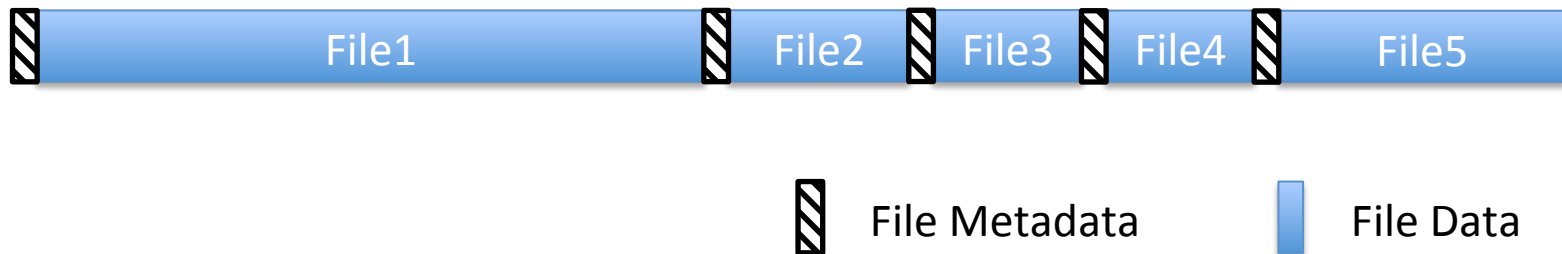
Metadata Impacts Deduplication

Metadata Impacts Deduplication

- Case 1: metadata changes
 - the input is an aggregate of many small user files, interleaved with file metadata
 - Metadata: file path, timestamps, etc.

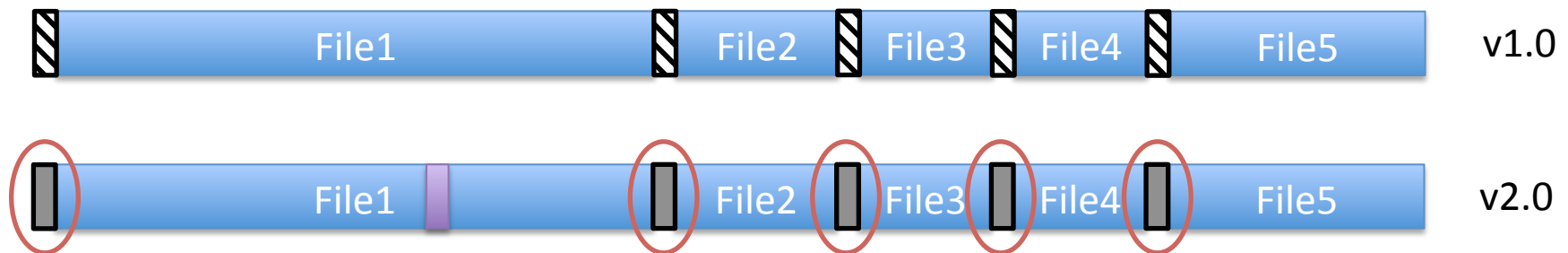
Metadata Impacts Deduplication

- Case 1: metadata changes
 - the input is an aggregate of many small user files, interleaved with file metadata
 - Metadata: file path, timestamps, etc.



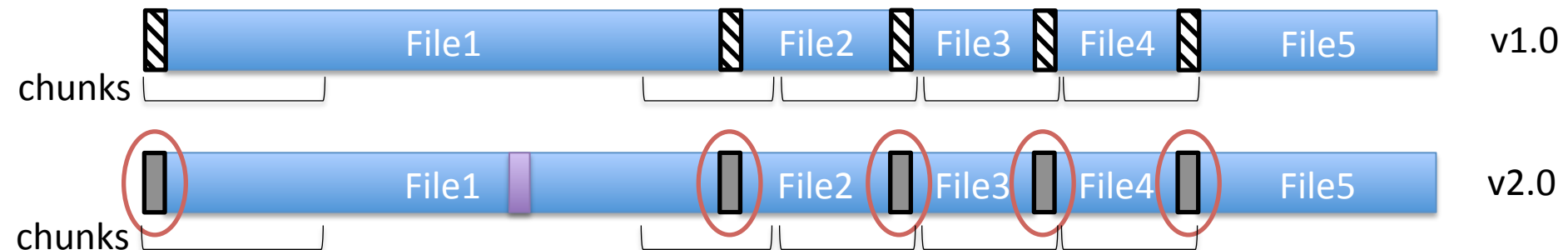
Metadata Impacts Deduplication

- Case 1: metadata changes
 - the input is an aggregate of many small user files, interleaved with file metadata
 - Metadata: file path, timestamps, etc.



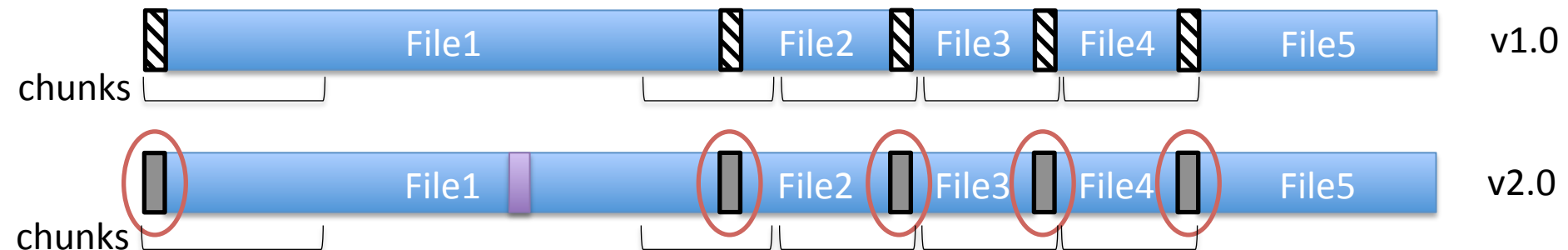
Metadata Impacts Deduplication

- Case 1: metadata changes
 - the input is an aggregate of many small user files, interleaved with file metadata
 - Metadata: file path, timestamps, etc.



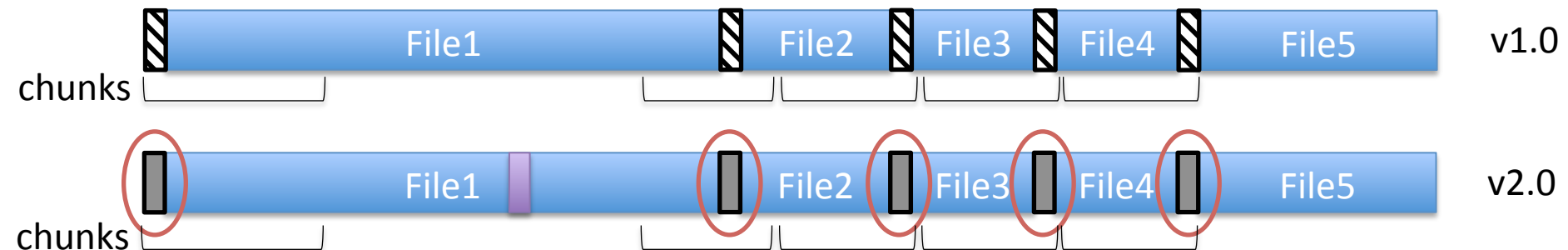
Metadata Impacts Deduplication

- Case 1: metadata changes
 - the input is an aggregate of many small user files, interleaved with file metadata
 - Metadata: file path, timestamps, etc.
 - GNU tar, EMC NetWorker, Oracle RMAN backups



Metadata Impacts Deduplication

- Case 1: metadata changes
 - the input is an aggregate of many small user files, interleaved with file metadata
 - Metadata: file path, timestamps, etc.
 - GNU tar, EMC NetWorker, Oracle RMAN backups
 - Videos suffer from a similar problem (*next talk*)



Metadata Impacts Deduplication

- Case 2: metadata location changes
 - The input is encoded in (fixed-size) blocks and metadata is inserted for each block

Metadata Impacts Deduplication

- Case 2: metadata location changes
 - The input is encoded in (fixed-size) blocks and metadata is inserted for each block
 - Data insertion/deletion lead to metadata shifts
 - Tape format

Metadata Impacts Deduplication

- Case 2: metadata location changes
 - The input is encoded in (fixed-size) blocks and metadata is inserted for each block
 - Data insertion/deletion lead to metadata shifts
 - Tape format

v1.0

Metadata Impacts Deduplication

- Case 2: metadata location changes
 - The input is encoded in (fixed-size) blocks and metadata is inserted for each block
 - Data insertion/deletion lead to metadata shifts
 - Tape format



v1.0

Metadata Impacts Deduplication

- Case 2: metadata location changes
 - The input is encoded in (fixed-size) blocks and metadata is inserted for each block
 - Data insertion/deletion lead to metadata shifts
 - Tape format



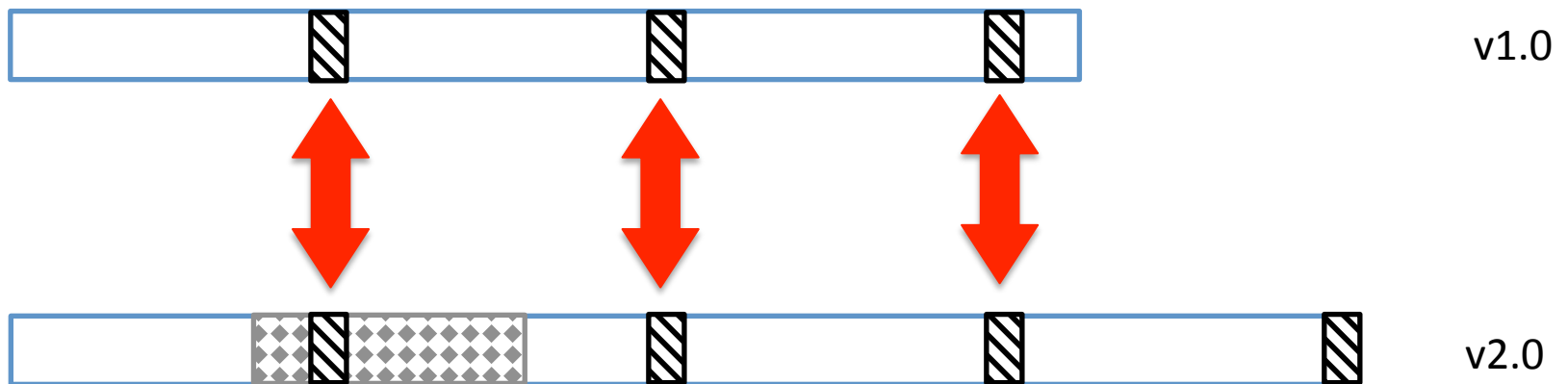
v1.0



v2.0

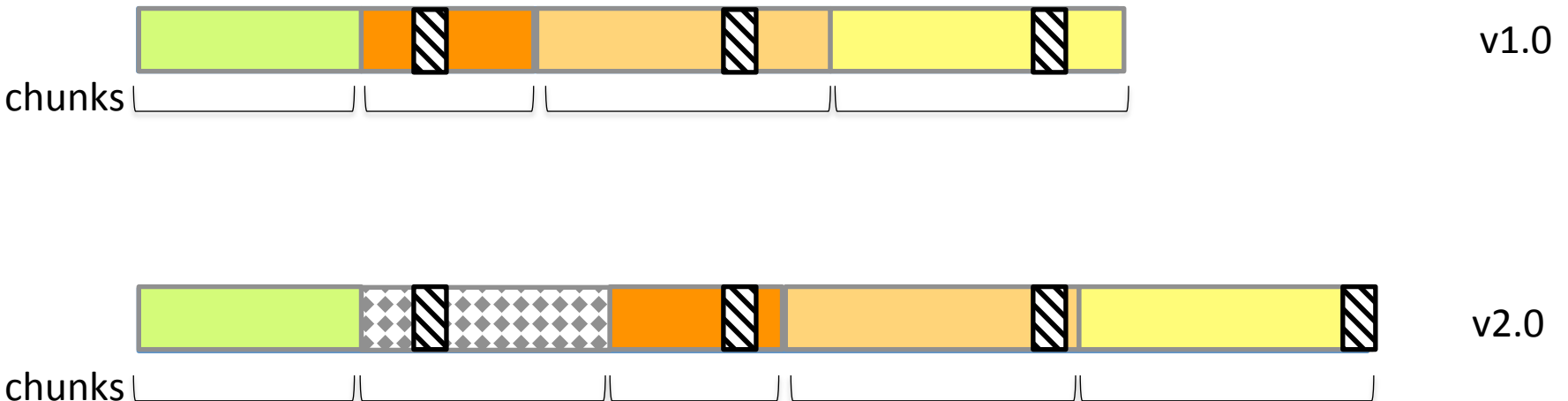
Metadata Impacts Deduplication

- Case 2: metadata location changes
 - The input is encoded in (fixed-size) blocks and metadata is inserted for each block
 - Data insertion/deletion lead to metadata shifts
 - Tape format



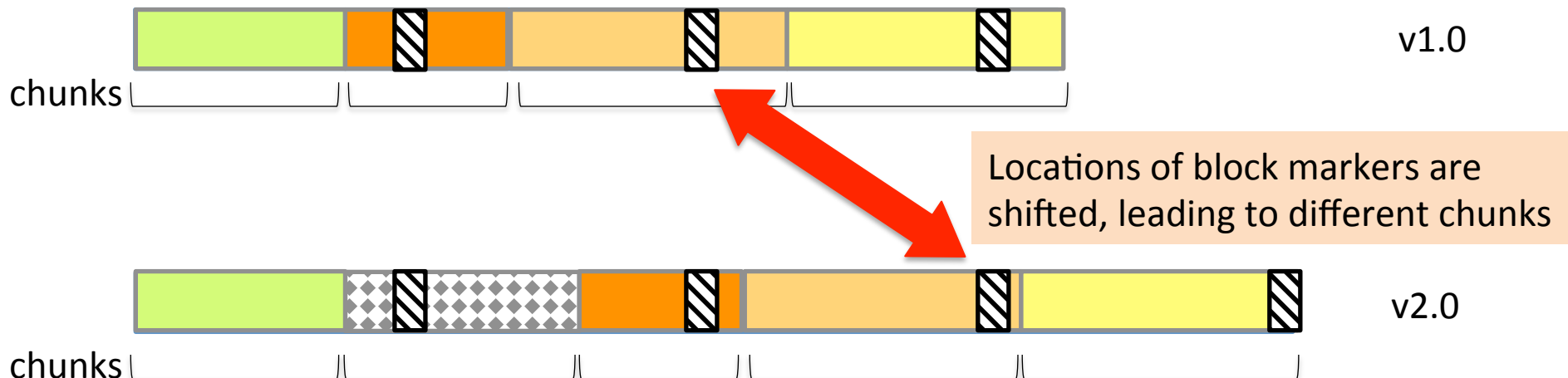
Metadata Impacts Deduplication

- Case 2: metadata location changes
 - The input is encoded in (fixed-size) blocks and metadata is inserted for each block
 - Data insertion/deletion lead to metadata shifts
 - Tape format



Metadata Impacts Deduplication

- Case 2: metadata location changes
 - The input is encoded in (fixed-size) blocks and metadata is inserted for each block
 - Data insertion/deletion lead to metadata shifts
 - Tape format



Solution: Separate Metadata from Data

Solution: Separate Metadata from Data

- Three approaches:
 - ***Recommended***: design deduplication-friendly formats
 - Case study: EMC NetWorker

Solution: Separate Metadata from Data

- Three approaches:
 - ***Recommended***: design deduplication-friendly formats
 - Case study: EMC NetWorker
 - ***Transparent***: application-level post-processing
 - Case study: GNU tar

Solution: Separate Metadata from Data

- Three approaches:
 - ***Recommended***: design deduplication-friendly formats
 - Case study: EMC NetWorker
 - ***Transparent***: application-level post-processing
 - Case study: GNU tar
 - ***Last resort***: format-aware deduplication
 - Case studies: 1) virtual tape libraries (VTL)
2) Oracle RMAN backups

Solution: Separate Metadata from Data

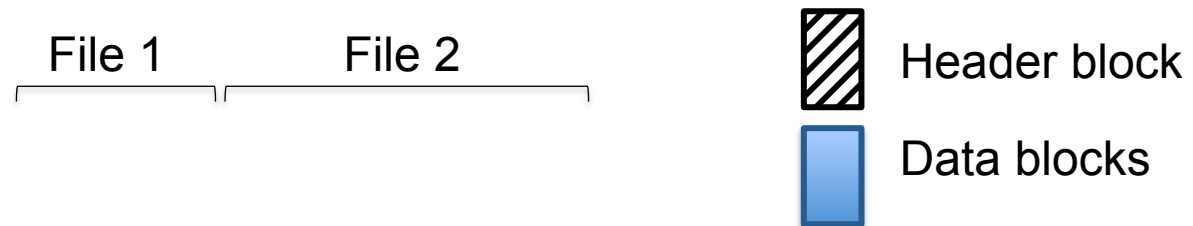
- Three approaches:
 - ***Recommended***: design deduplication-friendly formats
 - Case study: EMC NetWorker
 - ***Transparent***: application-level post-processing
 - **Case study: GNU tar**
 - ***Last resort***: format-aware deduplication
 - Case studies: **1) virtual tape libraries (VTL)**
2) Oracle RMAN backups

Application-level Post-processing

- tar (tape archive)
 - Collects files into one archive file
 - File system archiving, source code distribution, ...
- GNU tar format
 - A sequence of entries, one per file, each containing a file header and data blocks
 - Header block: file path, size, modification time

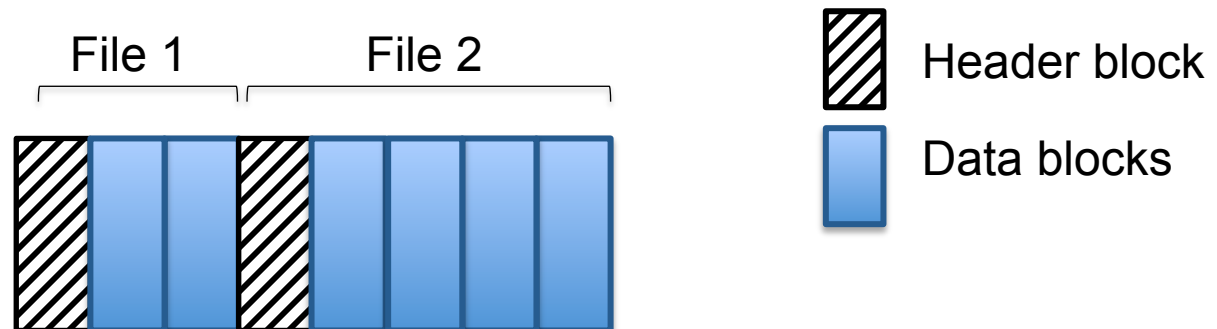
Application-level Post-processing

- tar (tape archive)
 - Collects files into one archive file
 - File system archiving, source code distribution, ...
- GNU tar format
 - A sequence of entries, one per file, each containing a file header and data blocks
 - Header block: file path, size, modification time



Application-level Post-processing

- tar (*t*ape *a*rchive)
 - Collects files into one archive file
 - File system archiving, source code distribution, ...
- GNU tar format
 - A sequence of entries, one per file, each containing a file header and data blocks
 - Header block: file path, size, modification time



Metadata Changes with GNU tar

SHA1s	linux-2.6.39.3/README	a735c31cef6d19d56de6824131527fdce04ead47
	linux-2.6.39.4/README	a735c31cef6d19d56de6824131527fdce04ead47

Metadata Changes with GNU tar

SHA1s

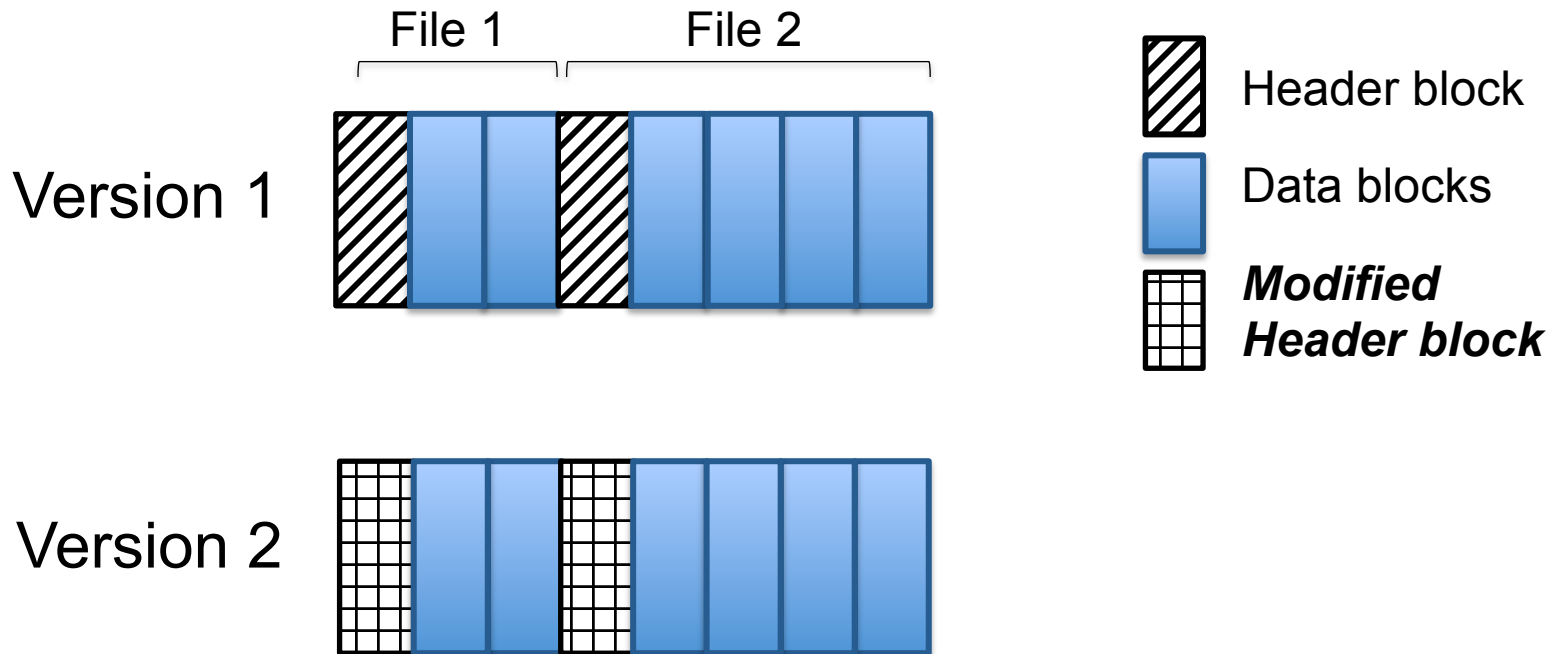
linux-2.6.39.3/README	a735c31cef6d19d56de6824131527fdce04ead47
linux-2.6.39.4/README	a735c31cef6d19d56de6824131527fdce04ead47
linux-2.6.39.3/README	-rw-rw-r-- 1 root root 17525 Jul 9 2011
linux-2.6.39.4/README	-rw-rw-r-- 1 root root 17525 Aug 3 2011

Metadata Changes with GNU tar

SHA1s

```
linux-2.6.39.3/README a735c31cef6d19d56de6824131527fdce04ead47  
linux-2.6.39.4/README a735c31cef6d19d56de6824131527fdce04ead47
```

```
linux-2.6.39.3/README -rw-rw-r-- 1 root root 17525 Jul 9 2011  
linux-2.6.39.4/README -rw-rw-r-- 1 root root 17525 Aug 3 2011
```

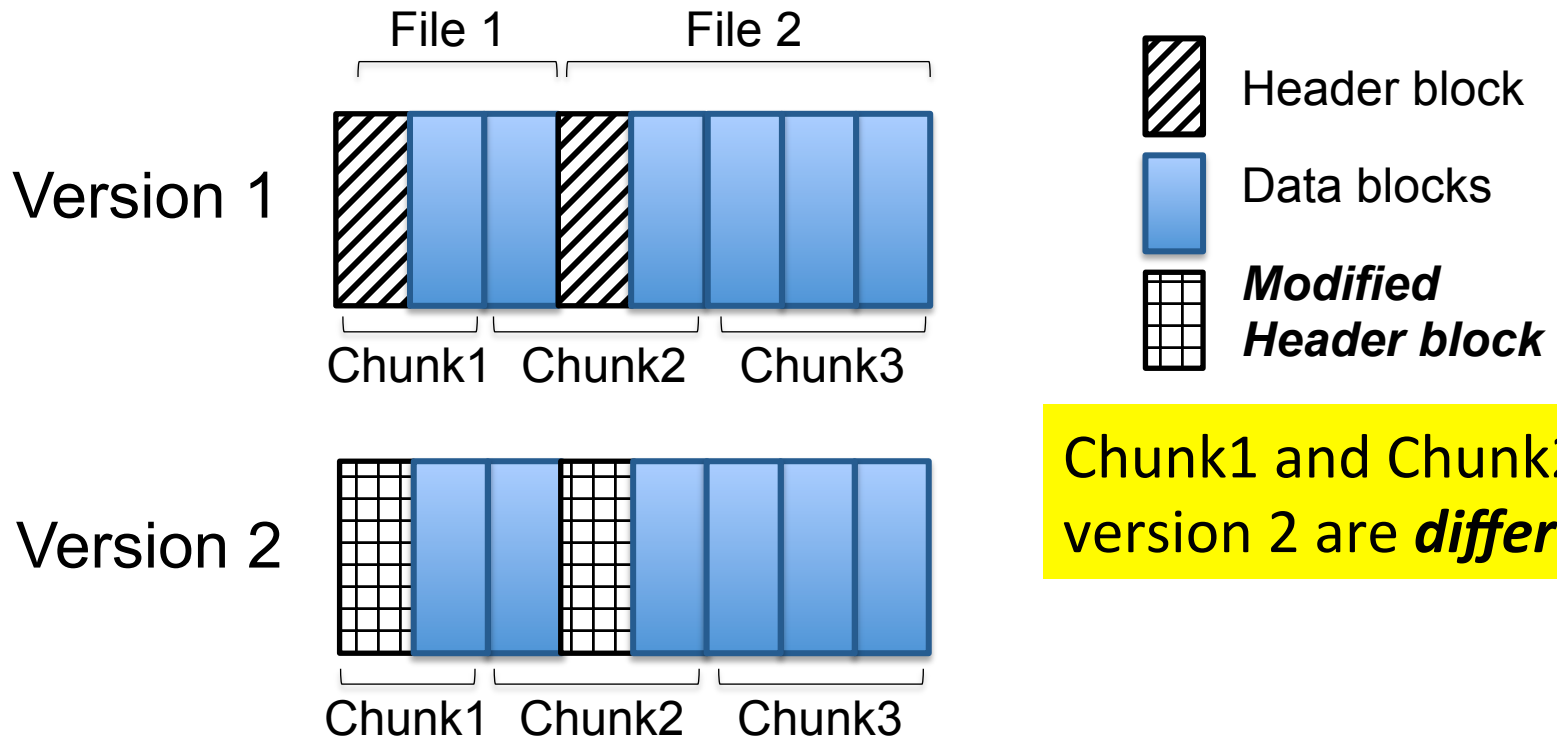


Metadata Changes with GNU tar

SHA1s

```
linux-2.6.39.3/README a735c31cef6d19d56de6824131527fdce04ead47  
linux-2.6.39.4/README a735c31cef6d19d56de6824131527fdce04ead47
```

```
linux-2.6.39.3/README -rw-rw-r-- 1 root root 17525 Jul 9 2011  
linux-2.6.39.4/README -rw-rw-r-- 1 root root 17525 Aug 3 2011
```



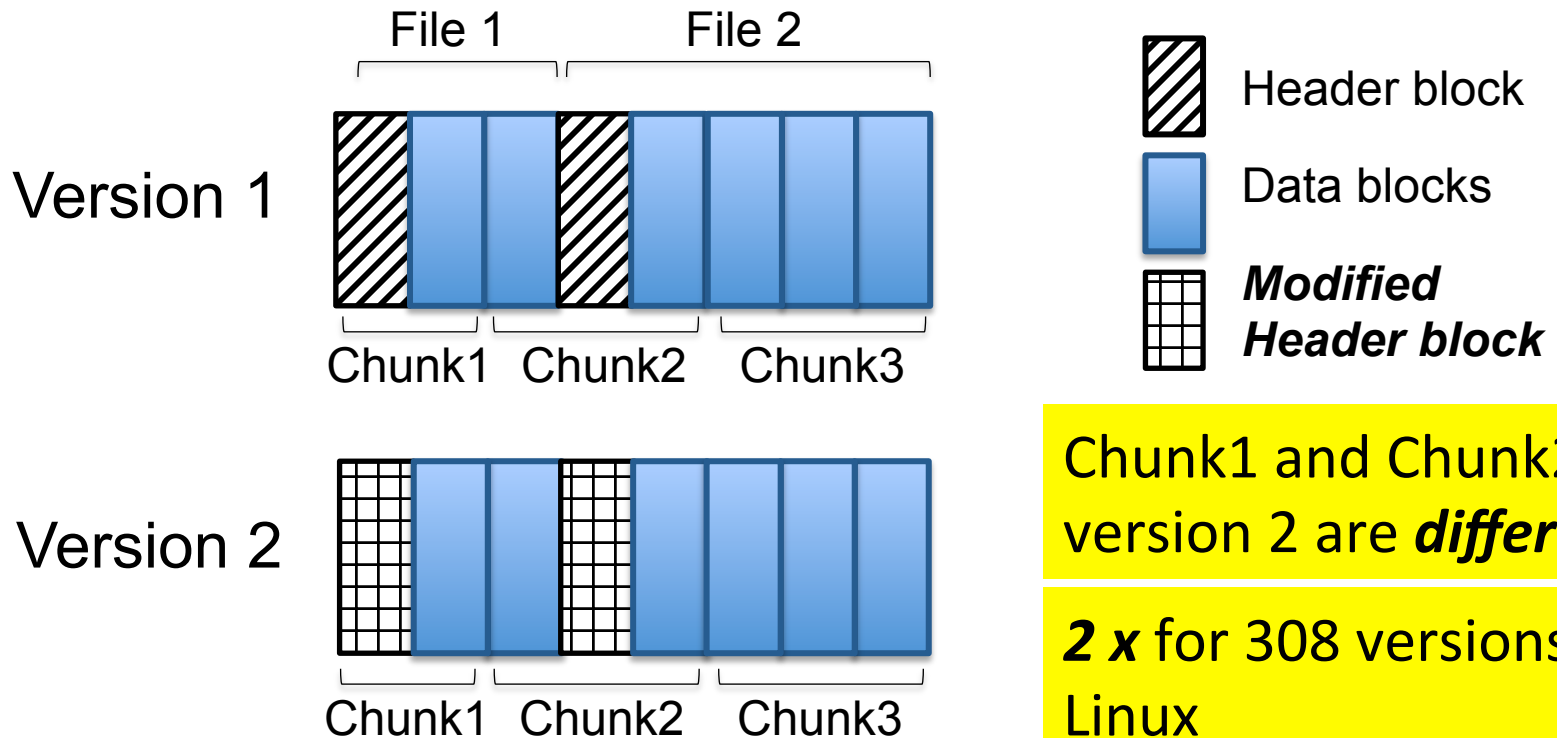
Chunk1 and Chunk2 in version 2 are ***different***.

Metadata Changes with GNU tar

SHA1s

```
linux-2.6.39.3/README a735c31cef6d19d56de6824131527fdce04ead47  
linux-2.6.39.4/README a735c31cef6d19d56de6824131527fdce04ead47
```

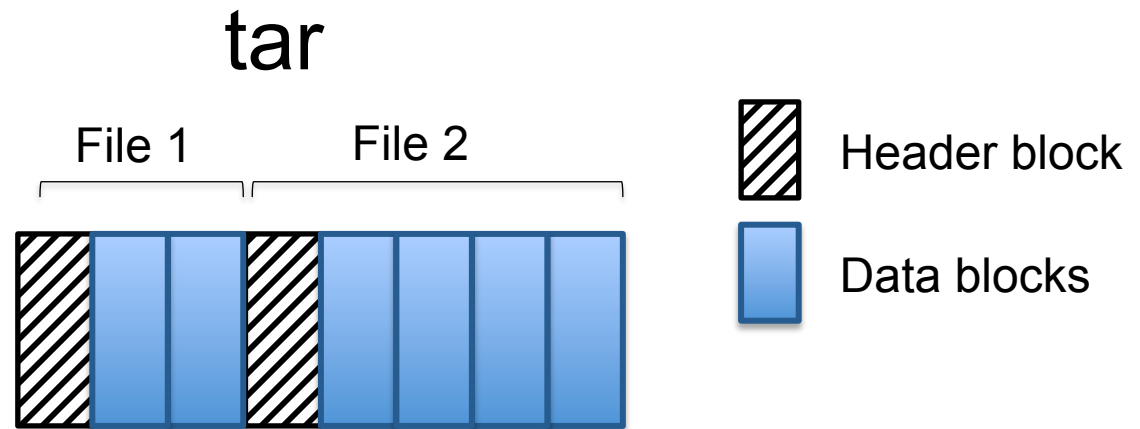
```
linux-2.6.39.3/README -rw-rw-r-- 1 root root 17525 Jul 9 2011  
linux-2.6.39.4/README -rw-rw-r-- 1 root root 17525 Aug 3 2011
```



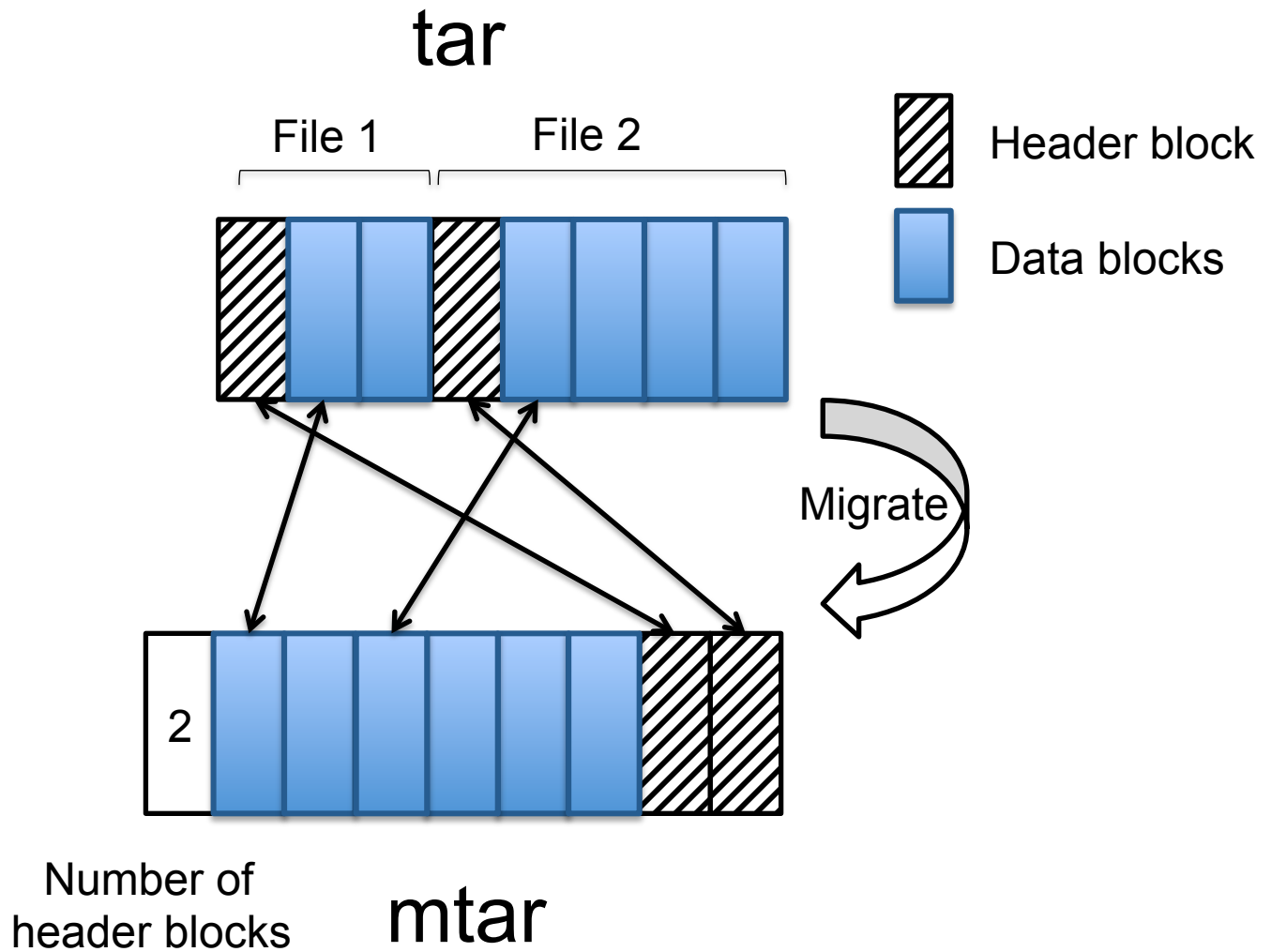
Chunk1 and Chunk2 in version 2 are **different**.

2 x for 308 versions of Linux

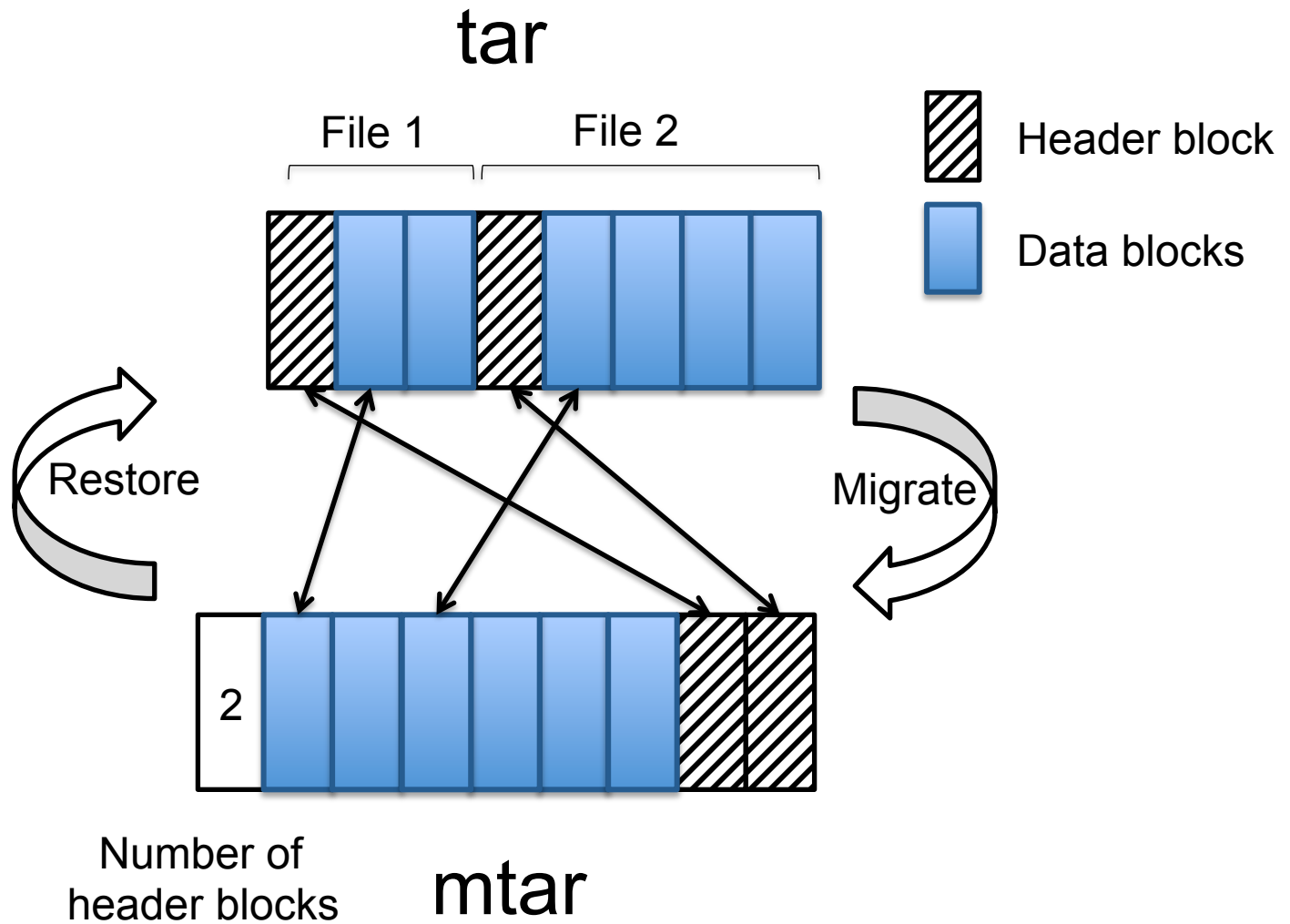
Migratory tar (mtar)



Migratory tar (mtar)



Migratory tar (mtar)

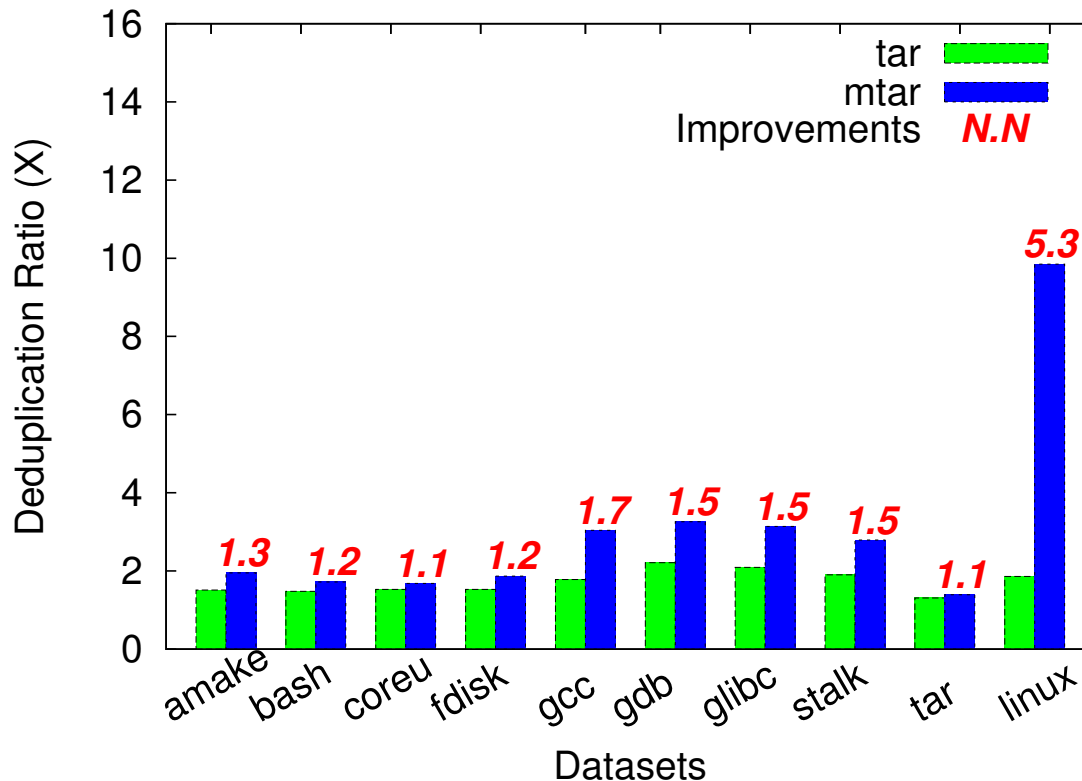


mtar - Evaluation

- 9 GNU software and Linux kernel source code
- Many versions: 13 ~ 308

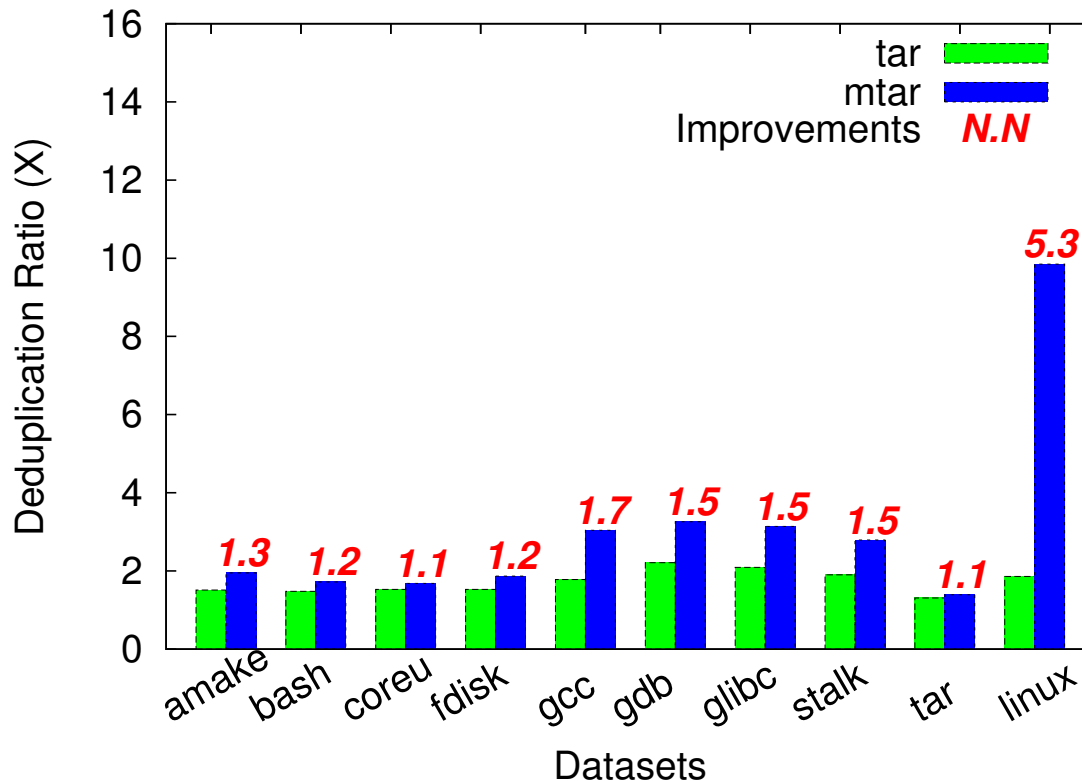
mtar - Evaluation

- 9 GNU software and Linux kernel source code
- Many versions: 13 ~ 308



mtar - Evaluation

- 9 GNU software and Linux kernel source code
- Many versions: 13 ~ 308



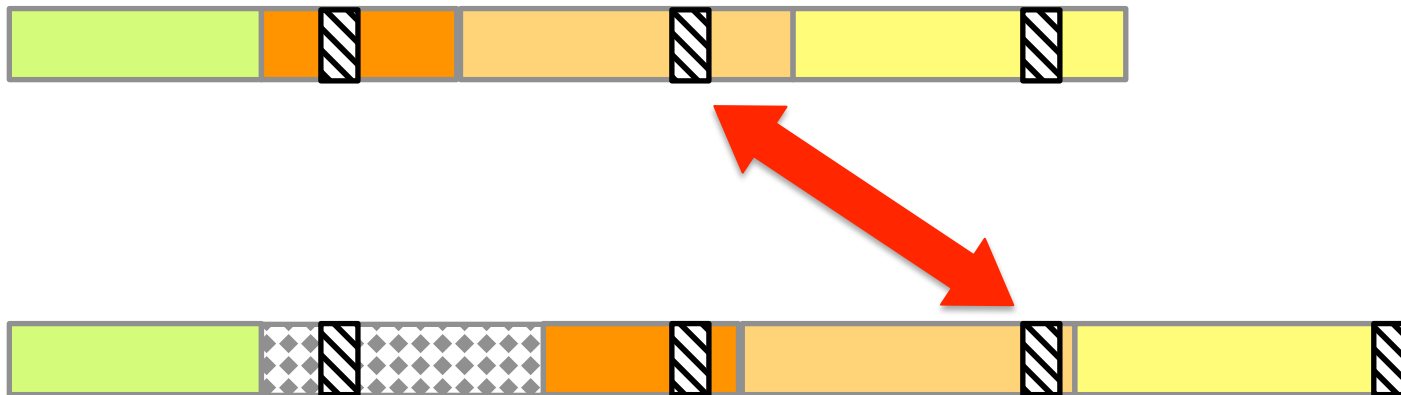
Improvements:
1. across all datasets
2. huge: 1.5-5.3x

Virtual Tape Library (VTL)

- VTL: a disk-based storage system to emulate a tape library
- Tape recorder: write in fixed-size blocks and attach a block marker for each block
 - Block marker: which device and the timestamp a file is written

Virtual Tape Library (VTL)

- VTL: a disk-based storage system to emulate a tape library
- Tape recorder: write in fixed-size blocks and attach a block marker for each block
 - Block marker: which device and the timestamp a file is written
- Data insertion/deletion leads to marker shifts.

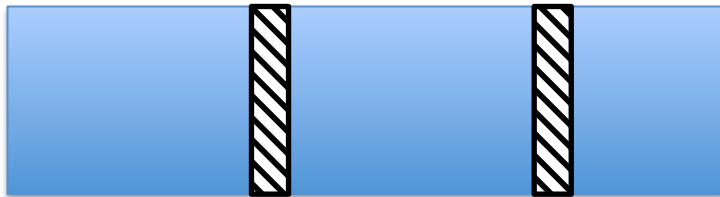


Format-aware Deduplication

- Modify deduplication systems to be aware of file formats; extract metadata and store it separately

Format-aware Deduplication

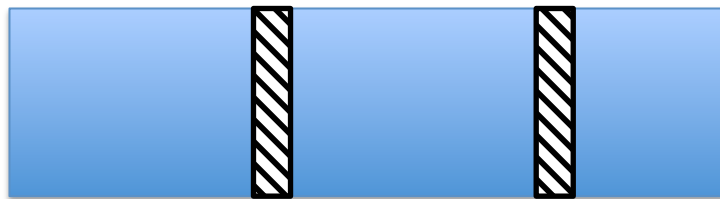
- Modify deduplication systems to be aware of file formats; extract metadata and store it separately



Input received

Format-aware Deduplication

- Modify deduplication systems to be aware of file formats; extract metadata and store it separately



Input received

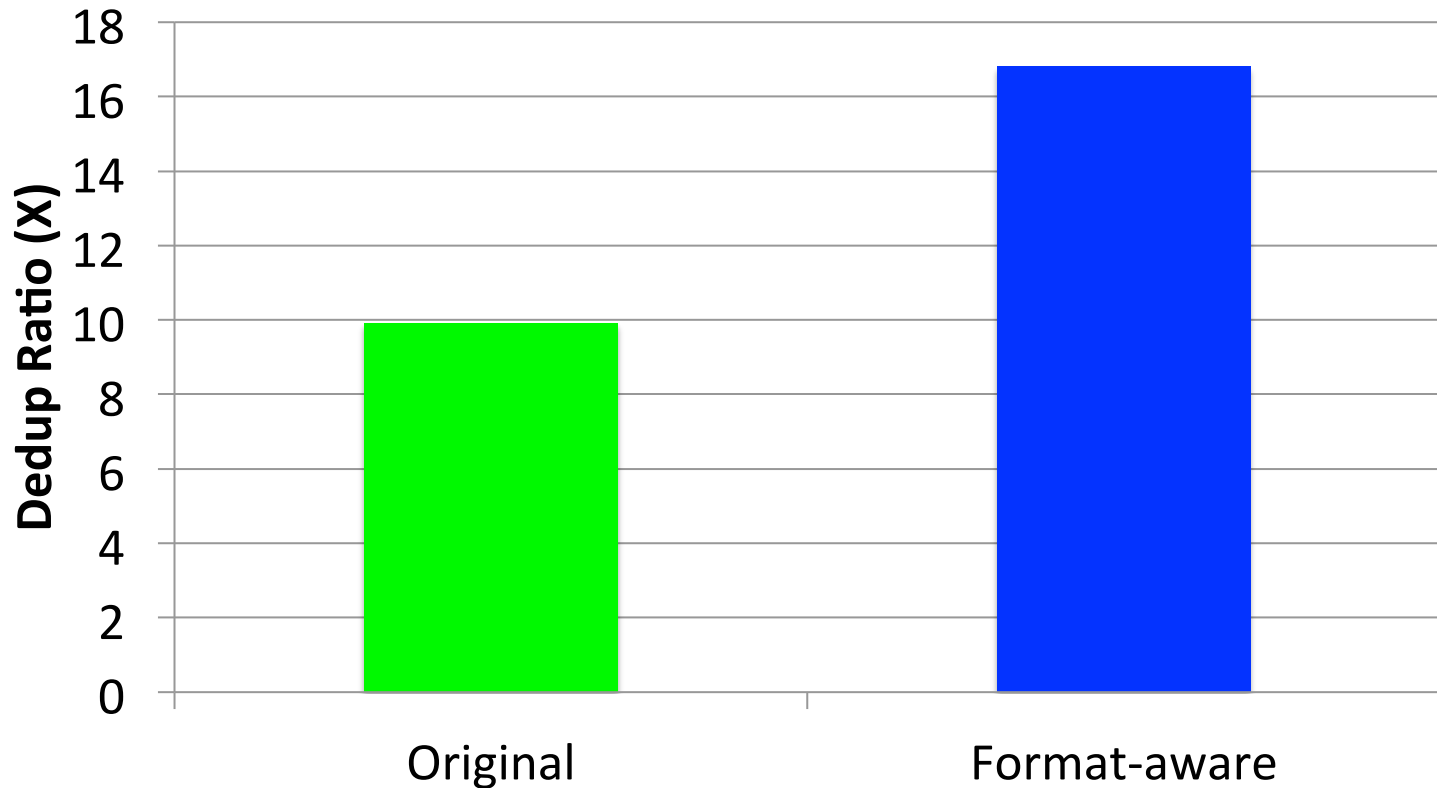


When storing, remove markers



Input stored

VTL Evaluation



75% improvement

More in the Paper

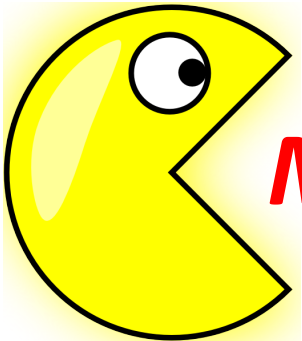
- ***Design deduplication-friendly formats***
 - ***Case study: EMC NetWorker***
- Application-level post-processing
 - Case study: GNU tar
- Format-aware Deduplication
 - Case studies: 1) VTL
 - 2) Oracle RMAN backups**
- mtar available at <https://github.com/xinglin/mtar>

Conclusion

- **Metadata impacts deduplication**
 - Metadata changes more frequently, introducing many *unnecessary* unique chunks
- Solution: separate metadata from data
 - Up to 5× improvements in deduplication

Conclusion

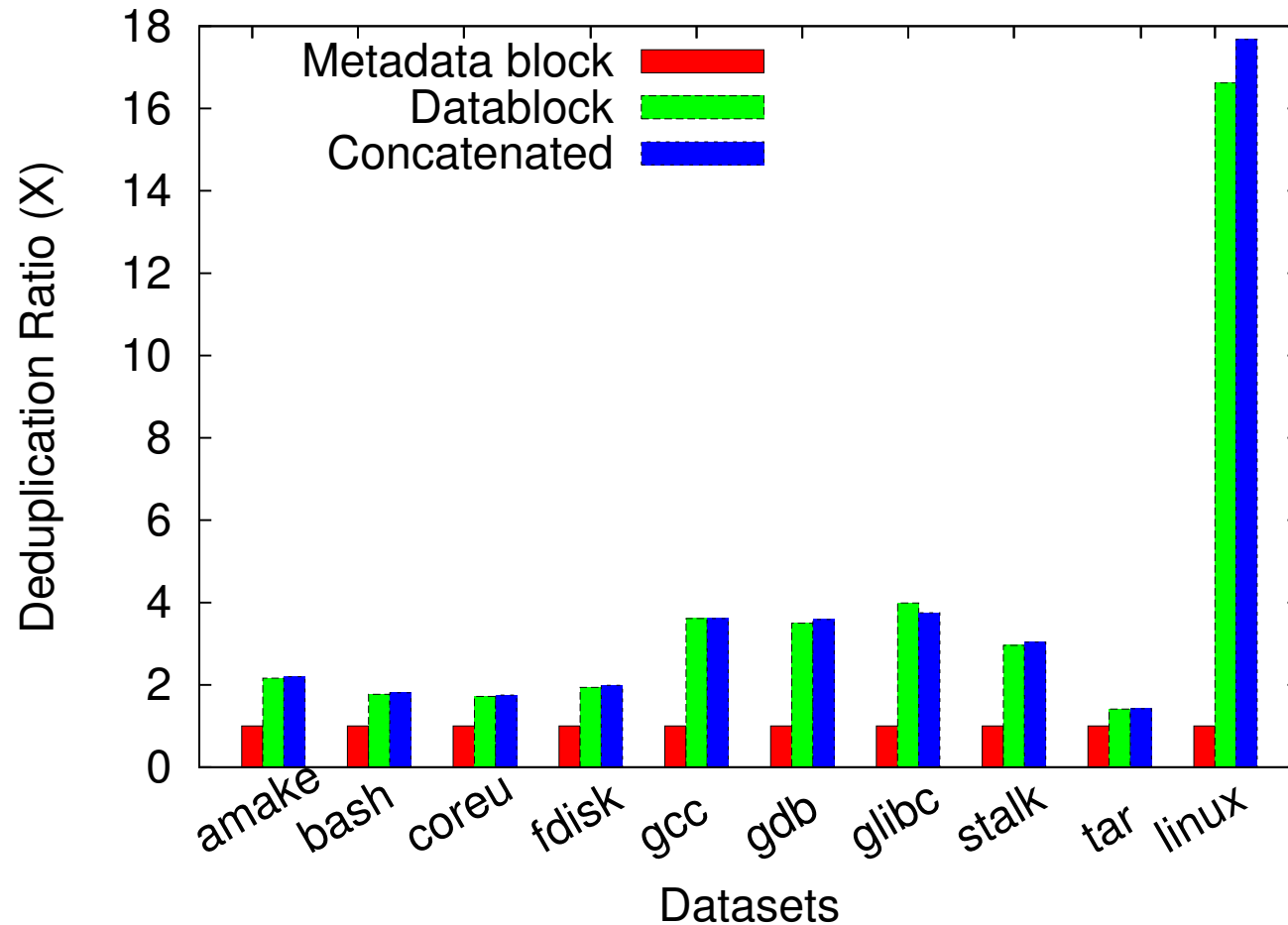
- **Metadata impacts deduplication**
 - Metadata changes more frequently, introducing many *unnecessary* unique chunks
- Solution: separate metadata from data
 - Up to 5× improvements in deduplication



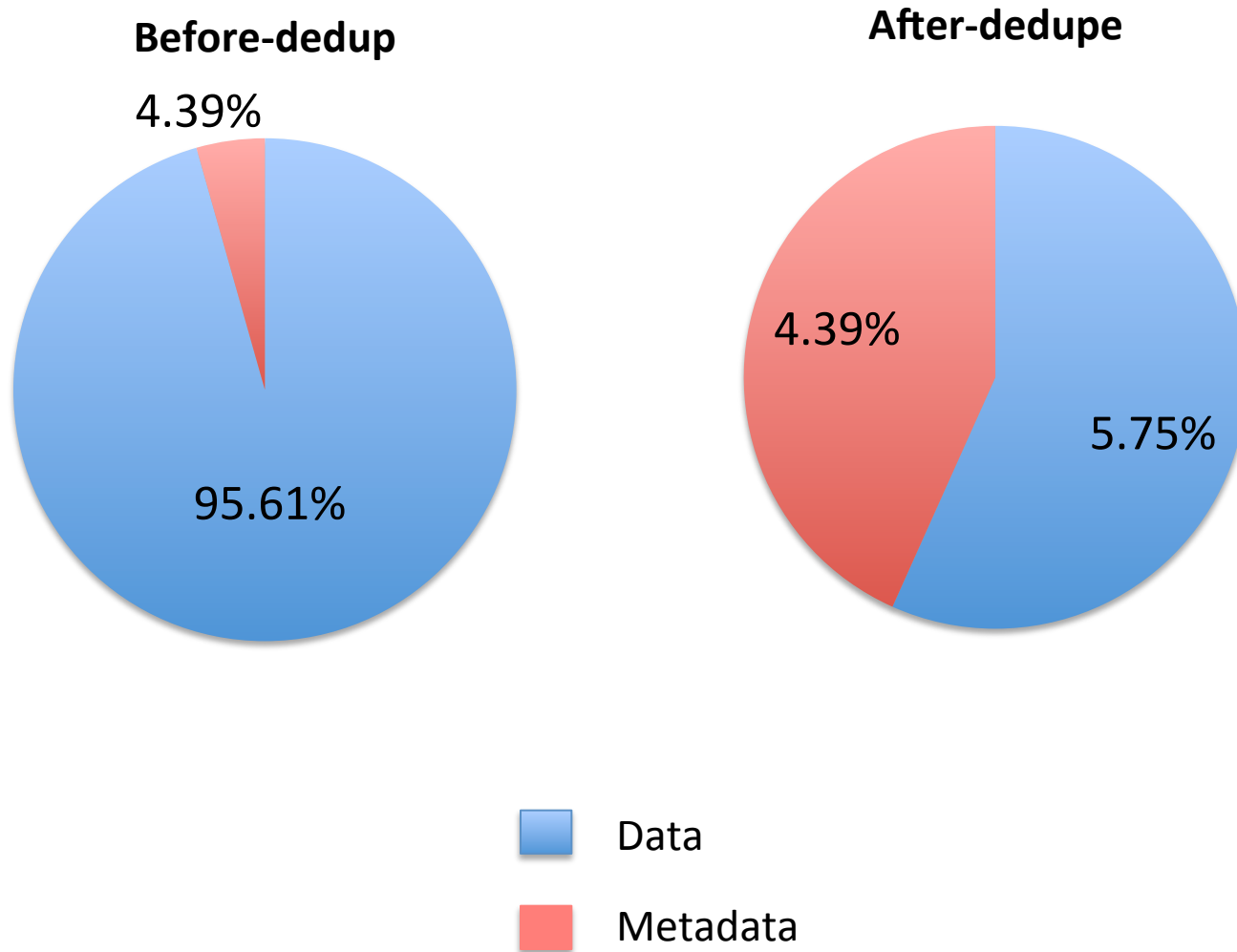
Metadata eats away your dedupe.

Backup slides

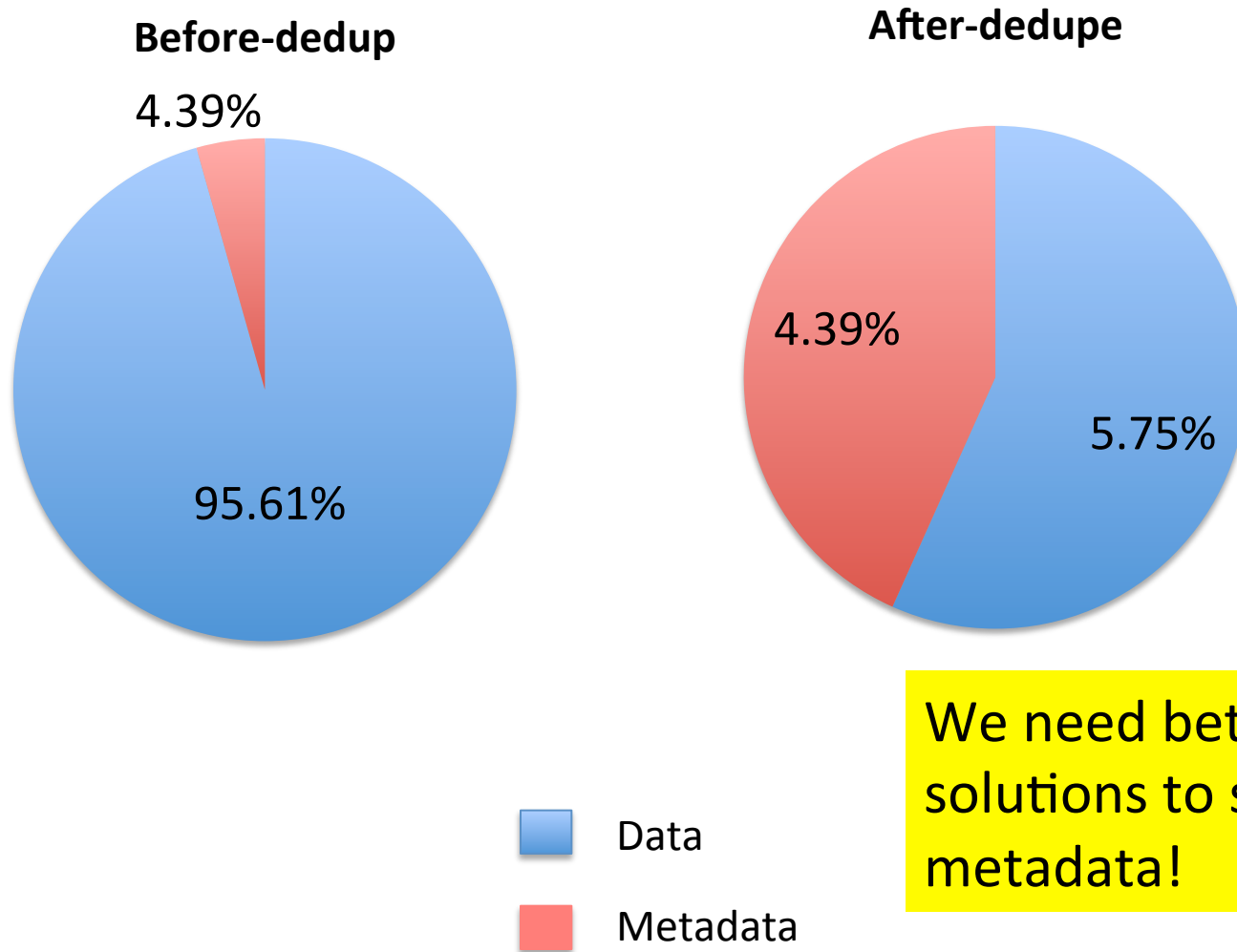
Data vs. Metadata vs. File Content-only



Metadata is STILL There!!!



Metadata is STILL There!!!



We need better solutions to store metadata!

mtar - Evaluation

Software	Versions	Size (MB)
automake	64	304.72
bash	23	276.69
coreutils	37	1284.49
fdisk	13	21.61
gcc	68	20315.45
gdb	32	4004.77
glibc	43	3811.48
smalltalk	33	685.39
tar	21	219.86
linux	308	98444.58

Dataset

Networker Evaluation

