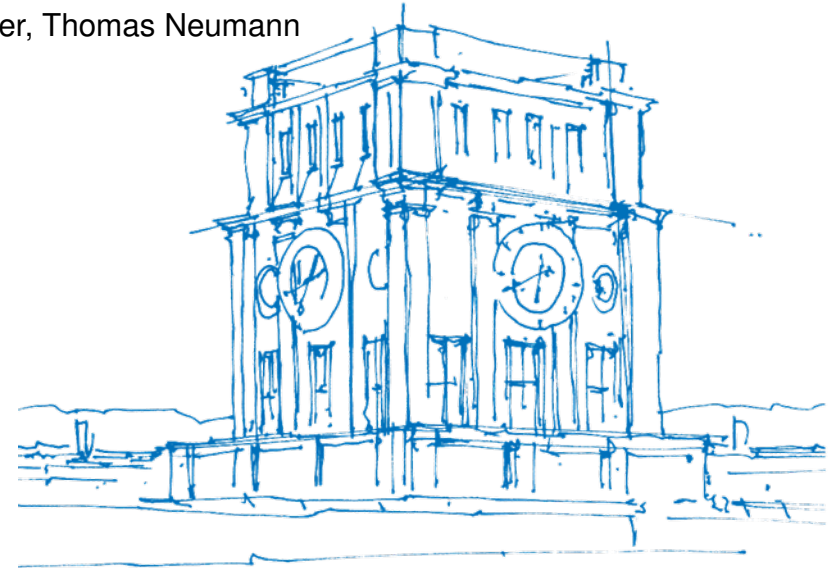


# TardisDB: Extending SQL to Support Versioning

Maximilian E. Schüle, Josef Schmeißer, Thomas Blum, Alfons Kemper, Thomas Neumann  
June 18 - 27, 2021, Virtual Event, China



*TUM Uhrenturm*

# Wikipedia: Version Control with Meta Tables



WIKIPEDIA  
The Free Encyclopedia

- Main page
- Contents
- Featured content
- Current events
- Random article
- Donate to Wikipedia
- Wikipedia store

## Interaction

- Help
- About Wikipedia
- Community portal
- Recent changes
- Contact page

## Tools

- What links here
- Related changes
- Atom
- Upload file

Not logged in [Talk](#) [Contributions](#) [Create account](#) [Log in](#)

Main Page [Talk](#)

[Read](#) [View source](#)

[View history](#)

## Main Page: Revision history

[? Help](#)

[View logs for this page \(view filter log\)](#)

### Filter revisions

[\[show\]](#)

External tools: [Find addition/removal \(Alternate\)](#) • [Find edits by user](#) • [Page statistics](#) • [Pageviews](#) • [Fix dead links](#)

For any version listed below, click on its date to view it. For more help, see [Help:Page history](#) and [Help:Edit summary](#). (cur) = difference from current version, (prev) = difference from preceding version, **m** = minor edit, **→** = section edit, **←** = automatic edit summary

(newest | [oldest](#)) View (newer 50 | [older 50](#)) ([20](#) | [50](#) | [100](#) | [250](#) | [500](#))

- (cur | prev)  [16:37, 24 March 2019](#) [Ad Orientem](#) (talk | contribs) **m** .. (6,712 bytes) **(+6,622)** .. *(Reverted edits by [Necrothesp](#) (talk) to last version by [K6ka](#)) (Tag: Rollback)*
- (cur | prev)  [16:35, 24 March 2019](#) [Necrothesp](#) (talk | contribs) .. (90 bytes) **(-6,622)** .. *(edit summary removed) (Tag: Replaced)*
- (cur | prev)  [20:20, 24 November 2018](#) [K6ka](#) (talk | contribs) **m** .. (6,712 bytes) **(+6,271)** .. *(Reverted edits by [Killiondude](#) (talk) to last version by [Vanamonde93](#)) (Tag: Rollback)*
- (cur | prev)  [20:19, 24 November 2018](#) [Killiondude](#) (talk | contribs) .. (441 bytes) **(-6,271)** .. *(←Replaced content with '<!-- BANNER ACROSS TOP OF PAGE --> H <div id="mp-topbanner" style="clear:both; position:relative; box-sizing:border-box; width:100%; margin:1....') (Tag: Replaced)*
- (cur | prev)  [20:19, 24 November 2018](#) [Vanamonde93](#) (talk | contribs) .. (6,712 bytes) **(+6,712)** .. *(Undid revision 870437181 by [Killiondude](#) (talk))*

# Wikipedia: Version Control with Meta Tables

WIKIPEDIA  
The Free Encyclopedia

- Main page
- Contents
- Featured content
- Current events
- Random article
- Donate to Wikipedia
- Wikipedia store

Interaction

- Help
- About Wikipedia
- Community portal
- Recent changes
- Contact page

Tools

- What links here
- Related changes
- Atom
- Upload file

Main Page [Talk](#)

## Main Page: Revision history

[View logs for this page \(view filter log\)](#)

### Filter revisions

External tools: [Find addition/removal \(Alternate\)](#) • [Find edits by user](#) • [Page statistics](#)

For any version listed below, click on its date to view it. For more help, see [Help:Revision](#) (prev) = difference from preceding version, **m** = minor edit, **→** = section edit (newest | **oldest**) View (newer 50 | **older 50**) (20 | 50 | 100 | 250 | 500)

- (cur | prev)  **16:37, 24 March 2019** [Ad Orientem \(talk | contribs\)](#) **m** [K6ka](#) (Tag: Rollback)
- (cur | prev)  ~~16:35, 24 March 2019~~ [Necrothesp \(talk | contribs\)](#) **m** [K6ka](#) (Tag: Rollback)
- (cur | prev)  **20:20, 24 November 2018** [K6ka \(talk | contribs\)](#) **m** [Vanamonde93](#) (Tag: Rollback)
- (cur | prev)  **20:19, 24 November 2018** [Killiondude \(talk | contribs\)](#) **m** [Vanamonde93](#) (Tag: Rollback)
- (cur | prev)  **20:19, 24 November 2018** [Vanamonde93 \(talk | contribs\)](#) **m** [Vanamonde93](#) (Tag: Rollback)

```
CREATE TABLE page (
  page_id INT PRIMARY KEY,
  page_title TEXT,
  page_latest INT REFERENCES pagecontent (old_id)
);
CREATE TABLE revision (
  rev_id INT PRIMARY KEY,
  rev_page INT REFERENCES page (page_id),
  rev_text_id INT REFERENCES pagecontent (old_id),
  rev_parent_id INT,
  rev_timestamp TIMESTAMP
);
CREATE TABLE pagecontent (
  old_id INT PRIMARY KEY,
  old_text TEXT
);
```

# Challenges and Approaches

- version control including multiple tables and respecting referential integrity
- implicit version control **inside** of database systems
- SQL integration needed



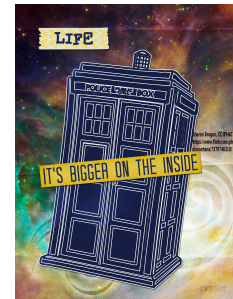
## SQL Integration

Branch Creation/Version Retrieval



## TardisDB

Integration into a Main-Memory DB



## TardisBenchmark

Benchmark Using Text Compression

# SQL Integration



# SQL Integration

- support versions over multiple tables with reference constraints

```
CREATE TABLE users (id INT PRIMARY KEY, name TEXT);  
CREATE TABLE things (id INT PRIMARY KEY, name TEXT, user INT REFERENCES users(id));
```

- downward compatible to SQL-92 (default: master branch)

```
INSERT INTO users VALUES (1, 'Alice'); INSERT INTO things VALUES (21, 'printer', 1);
```

- added statement for branch creation (preserves a database state)

```
CREATE BRANCH mybranch FROM master;
```

- version annotation next to the table in the `from`-clause (default: master branch)

```
INSERT INTO users VERSION mybranch VALUES (2, 'Bob');  
UPDATE things VERSION mybranch SET user=2 WHERE id=21;  
SELECT * FROM users VERSION mybranch;
```

- querying different versions within a branch is possible

```
SELECT a.id, COALESCE(a.name, b.name)  
FROM users VERSION master as a FULL OUTER JOIN users VERSION mybranch as b ON a.id=b.id
```

- branch deletion statement to release memory

```
DELETE BRANCH mybranch;
```

# TardisDB



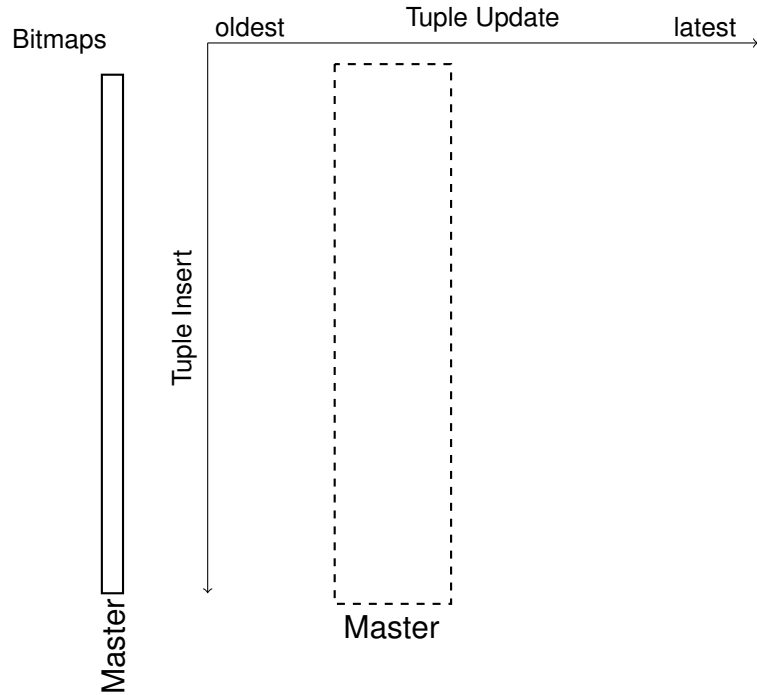
# TardisDB: Concept

- **Time and Relative Dimensions in Databases**
  - two dimensions: inserts and updates
- use bitmaps for insertions and deletions
  - bitmaps for every branch indicate included tuples
- reuses multi-version concurrency control for updates
  - every tuple is marked by the creator branch for table scans
  - prioritised master branch: no further operation needed
  - updates happen in place, previous versions are chained in buffers



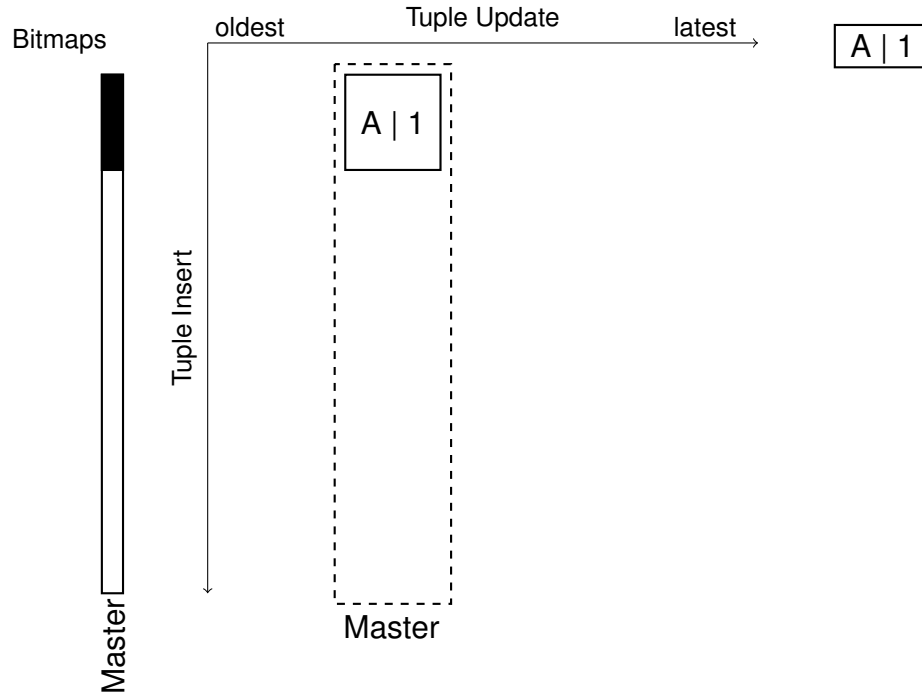


# TardisDB: Concept



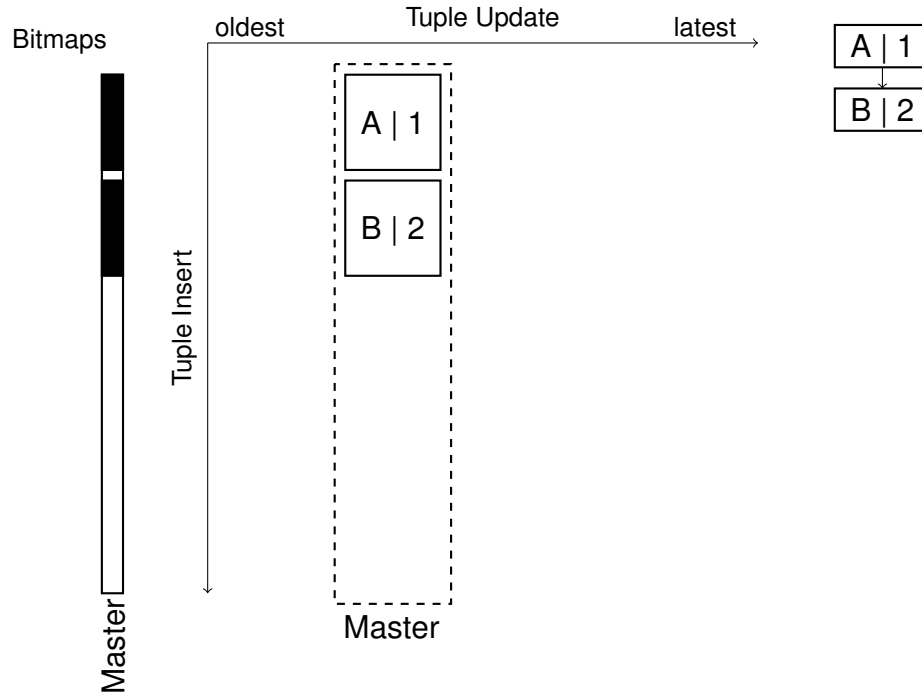
Time	Master	Branch 1	Branch 2	Branch 3
------	--------	----------	----------	----------

# TardisDB: Concept



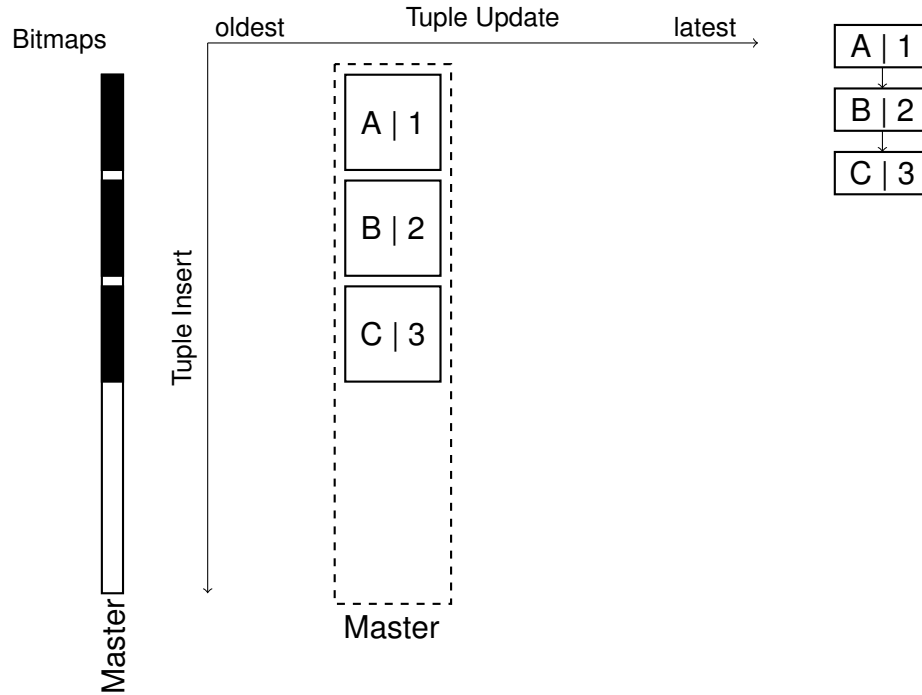
Time	Master	Branch 1	Branch 2	Branch 3
1	insert A			

# TardisDB: Concept



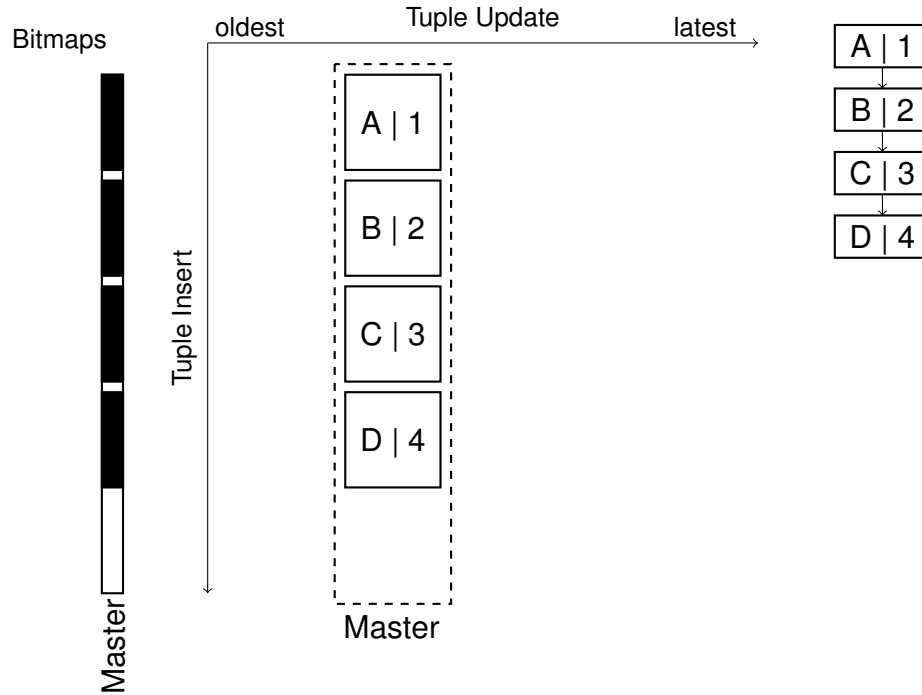
Time	Master	Branch 1	Branch 2	Branch 3
1	insert A			
2	insert B			

# TardisDB: Concept



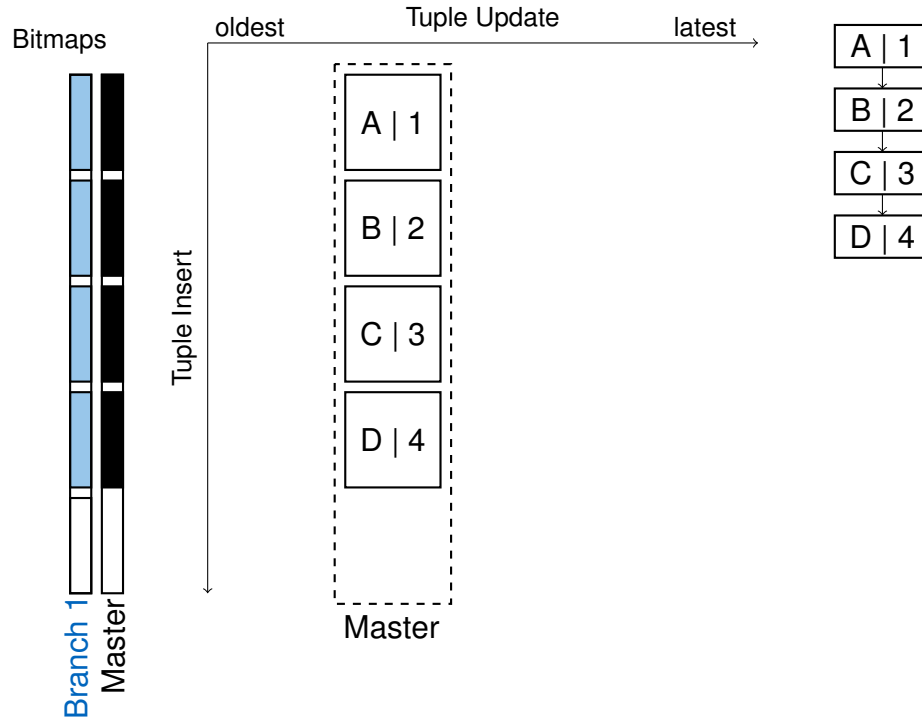
Time	Master	Branch 1	Branch 2	Branch 3
1	insert A			
2	insert B			
3	insert C			

# TardisDB: Concept



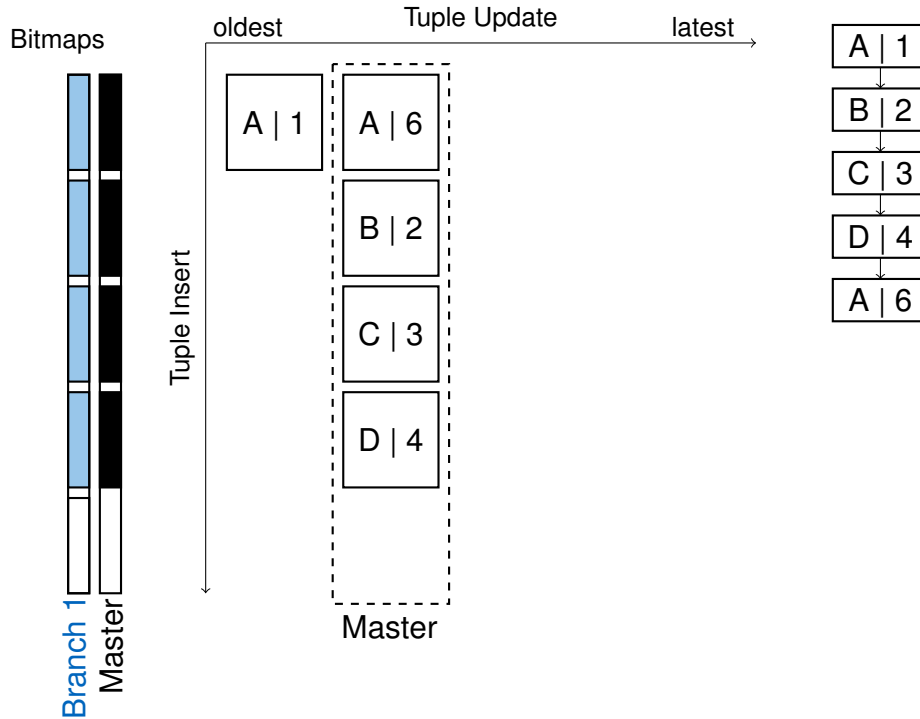
Time	Master	Branch 1	Branch 2	Branch 3
1	insert A			
2	insert B			
3	insert C			
4	insert D			

# TardisDB: Concept



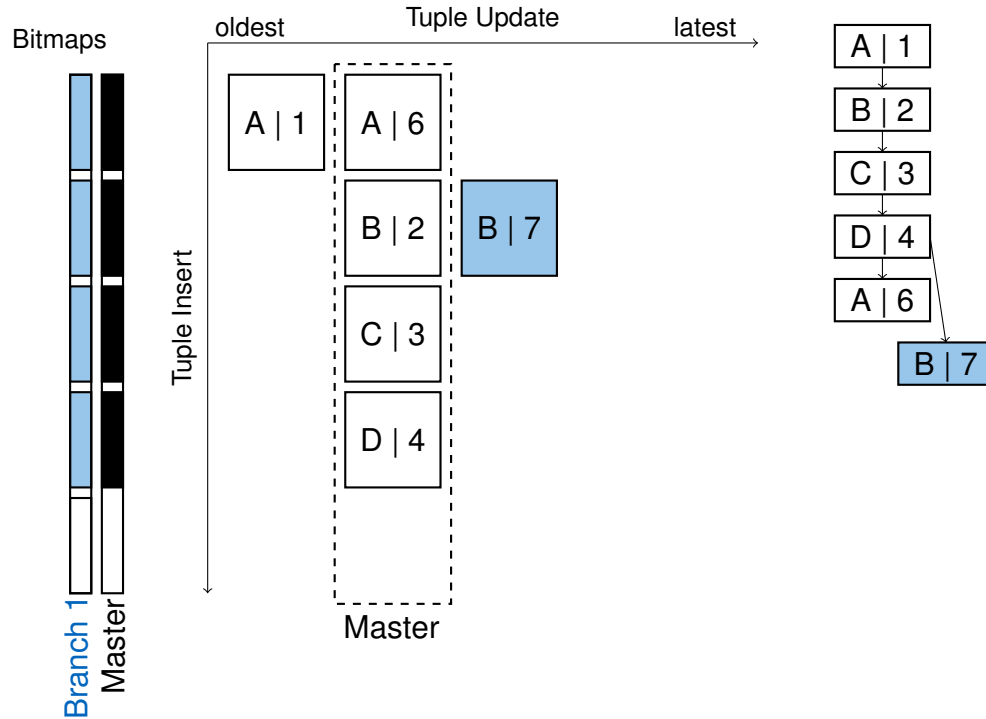
Time	Master	Branch 1	Branch 2	Branch 3
1	insert A			
2	insert B			
3	insert C			
4	insert D			
5	branch 1			

# TardisDB: Concept



Time	Master	Branch 1	Branch 2	Branch 3
1	insert A			
2	insert B			
3	insert C			
4	insert D			
5	branch 1			
6	update A			

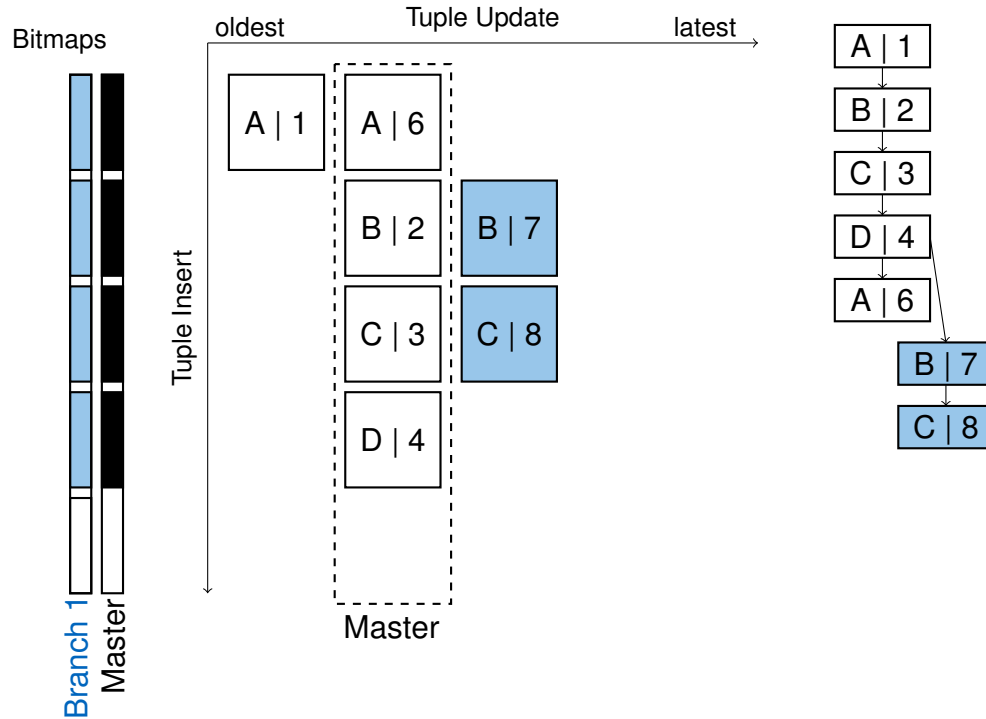
# TardisDB: Concept



Time	Master	Branch 1	Branch 2	Branch 3
1	insert A			
2	insert B			
3	insert C			
4	insert D			
5	branch 1			
6	update A			
7		update B		

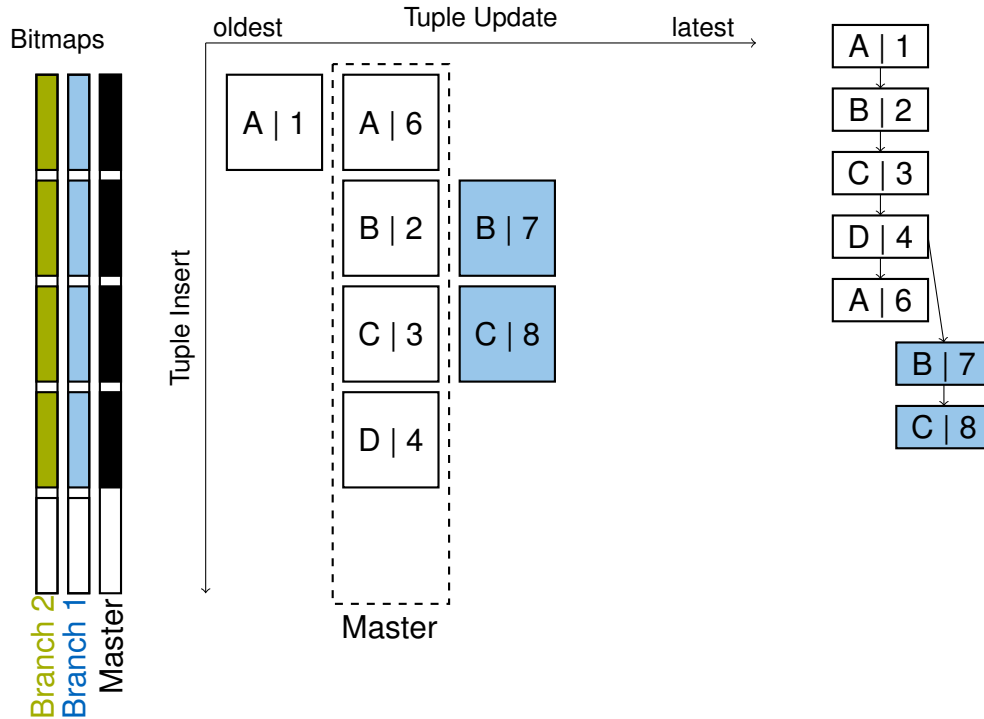


# TardisDB: Concept



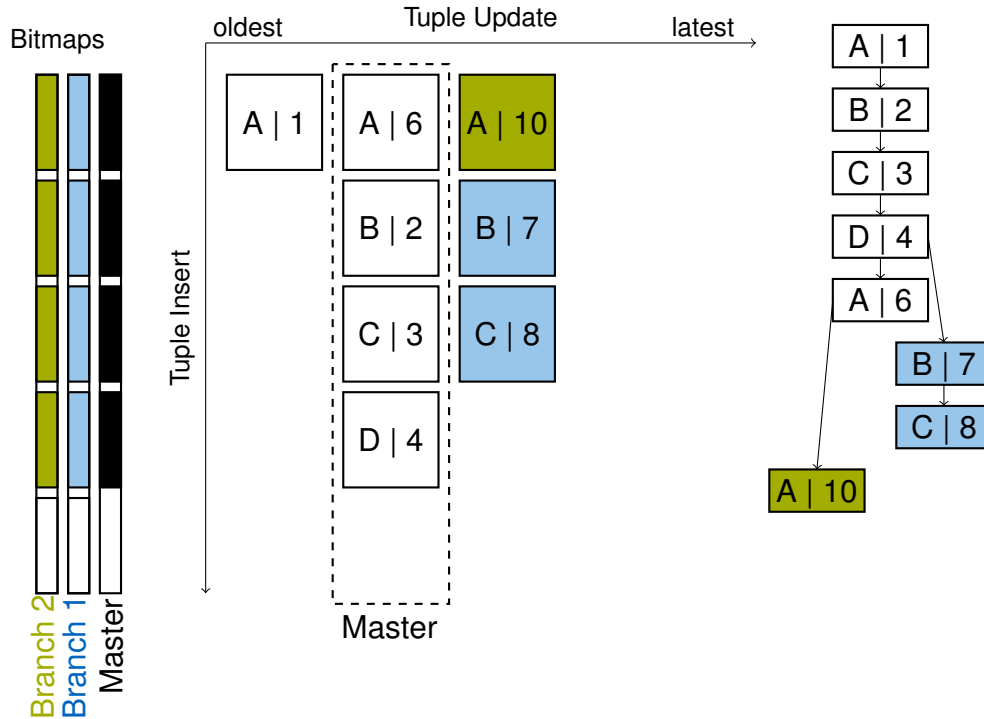
Time	Master	Branch 1	Branch 2	Branch 3
1	insert A			
2	insert B			
3	insert C			
4	insert D			
5	branch 1			
6	update A			
7		update B		
8		update C		

# TardisDB: Concept



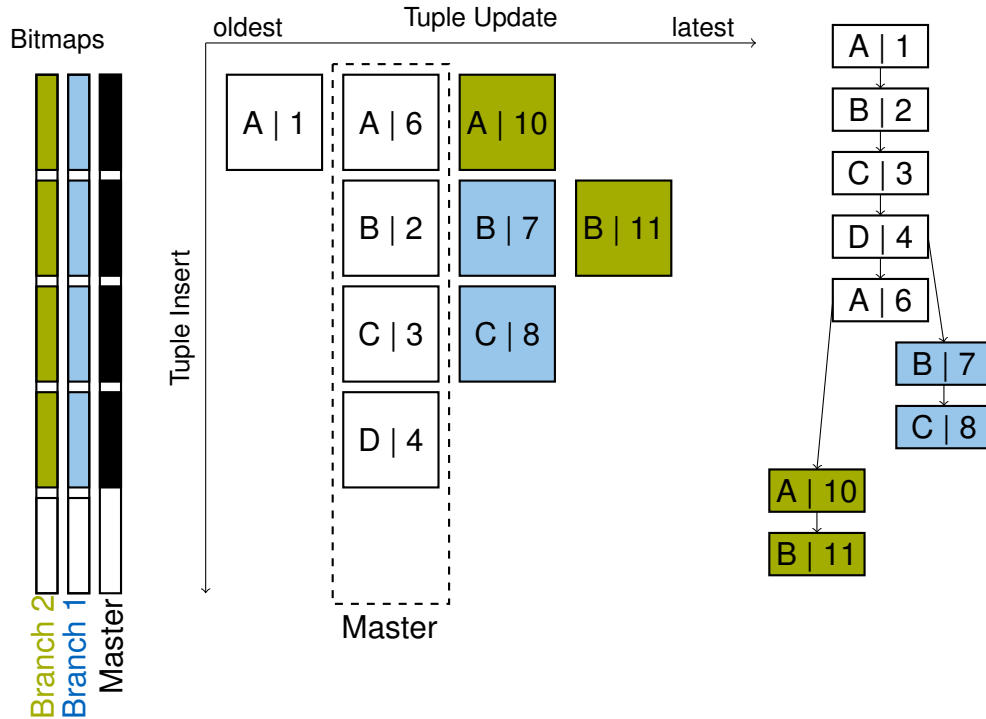
Time	Master	Branch 1	Branch 2	Branch 3
1	insert A			
2	insert B			
3	insert C			
4	insert D			
5	branch 1			
6	update A			
7		update B		
8		update C		
9	branch 2			

# TardisDB: Concept



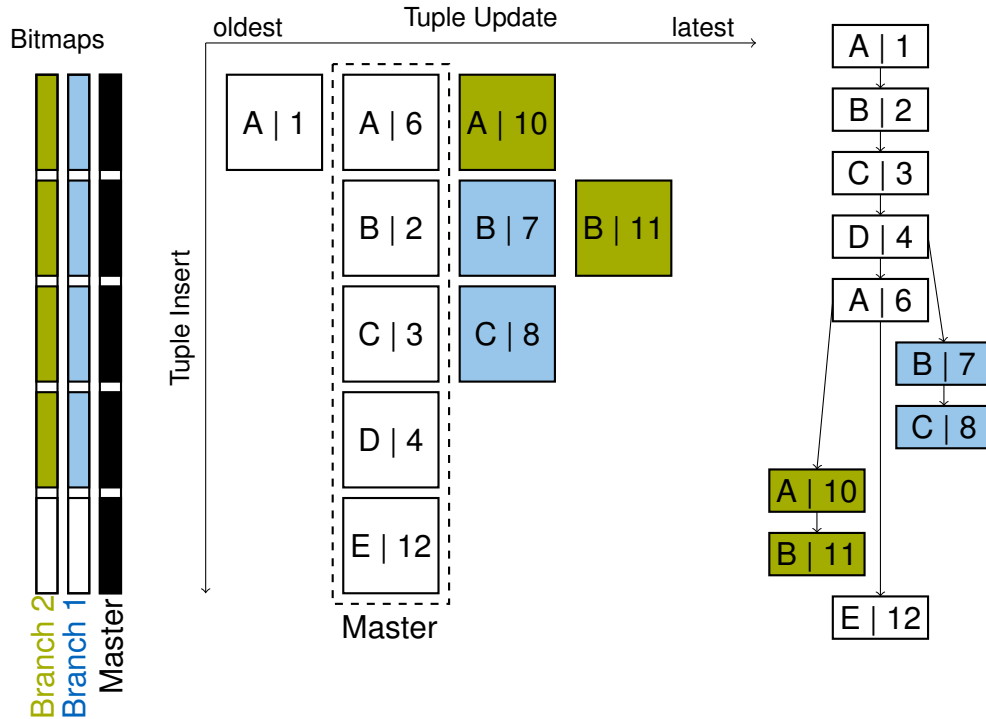
Time	Master	Branch 1	Branch 2	Branch 3
1	insert A			
2	insert B			
3	insert C			
4	insert D			
5	branch 1			
6	update A			
7		update B		
8		update C		
9	branch 2			
10			update A	

# TardisDB: Concept



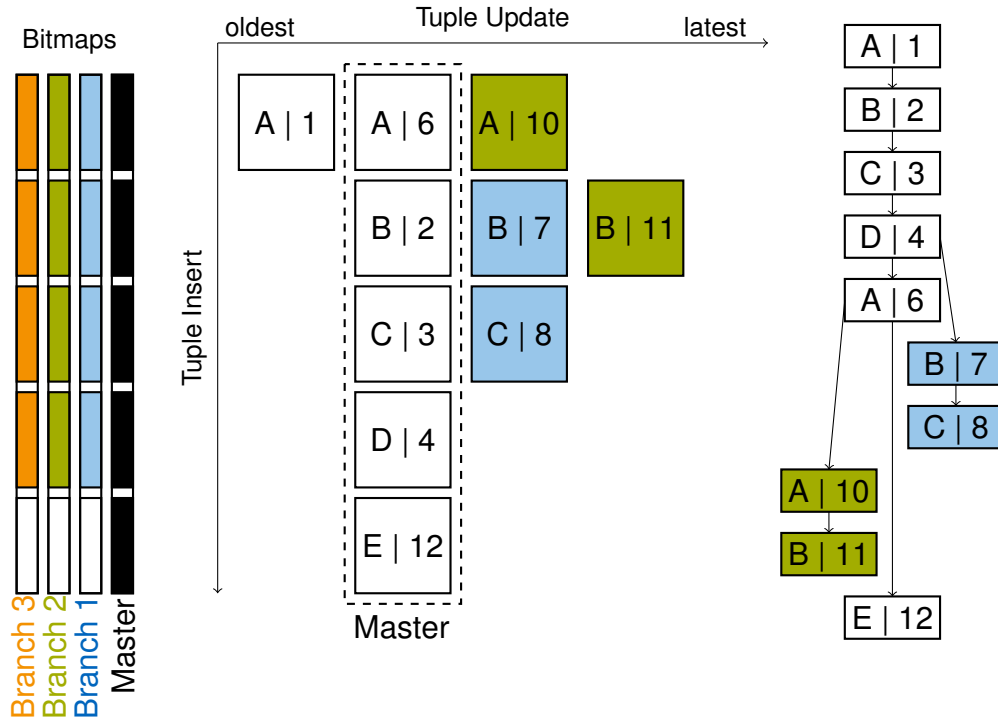
Time	Master	Branch 1	Branch 2	Branch 3
1	insert A			
2	insert B			
3	insert C			
4	insert D			
5	branch 1			
6	update A			
7		update B		
8		update C		
9	branch 2			
10			update A	
11			update B	

# TardisDB: Concept



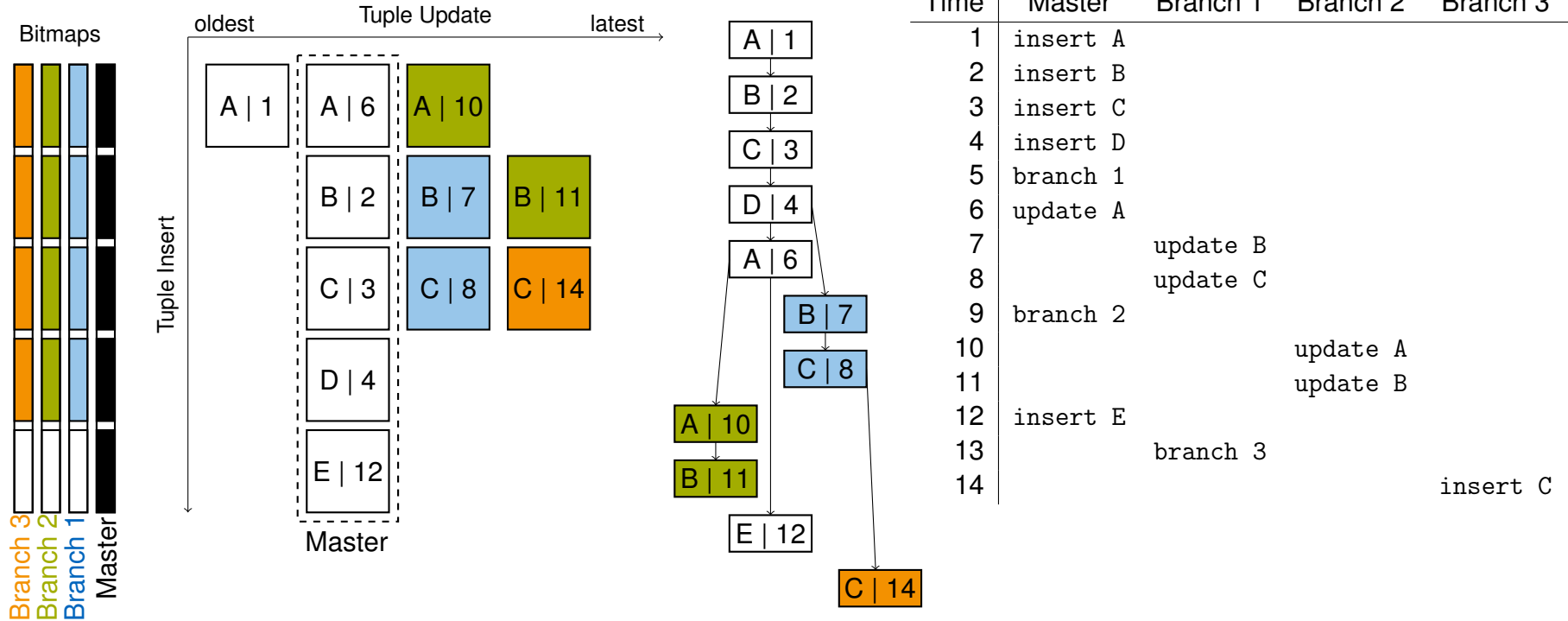
Time	Master	Branch 1	Branch 2	Branch 3
1	insert A			
2	insert B			
3	insert C			
4	insert D			
5	branch 1			
6	update A			
7		update B		
8		update C		
9	branch 2			
10		update A		
11		update B		
12	insert E			

# TardisDB: Concept

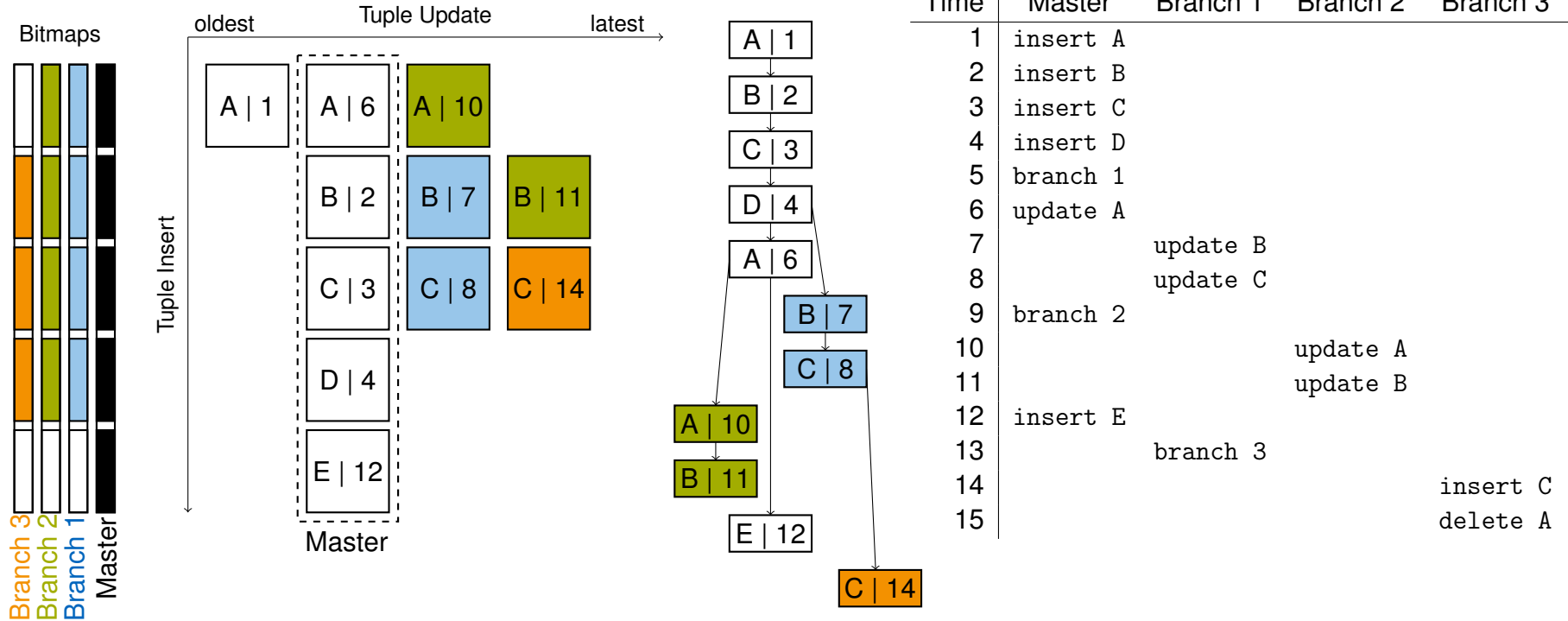


Time	Master	Branch 1	Branch 2	Branch 3
1	insert A			
2	insert B			
3	insert C			
4	insert D			
5	branch 1			
6	update A			
7		update B		
8		update C		
9	branch 2			
10			update A	
11			update B	
12	insert E			
13		branch 3		

# TardisDB: Concept

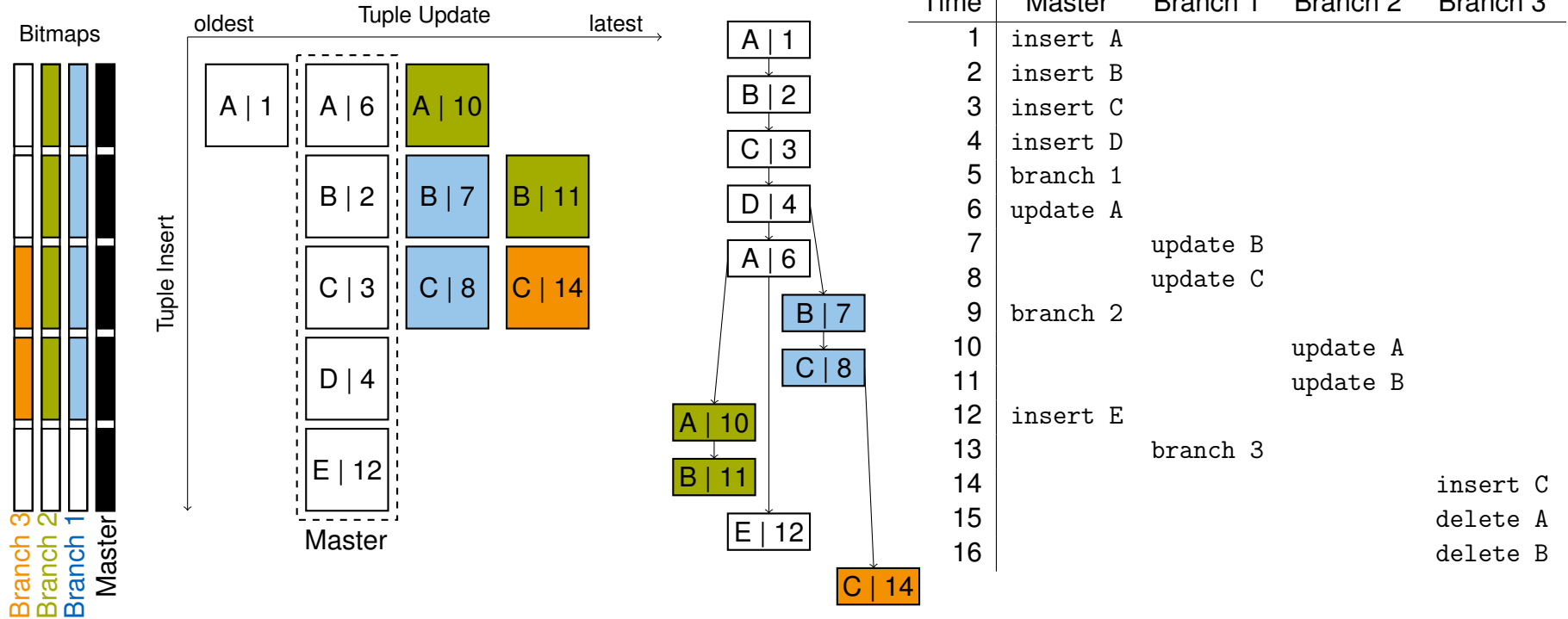


# TardisDB: Concept

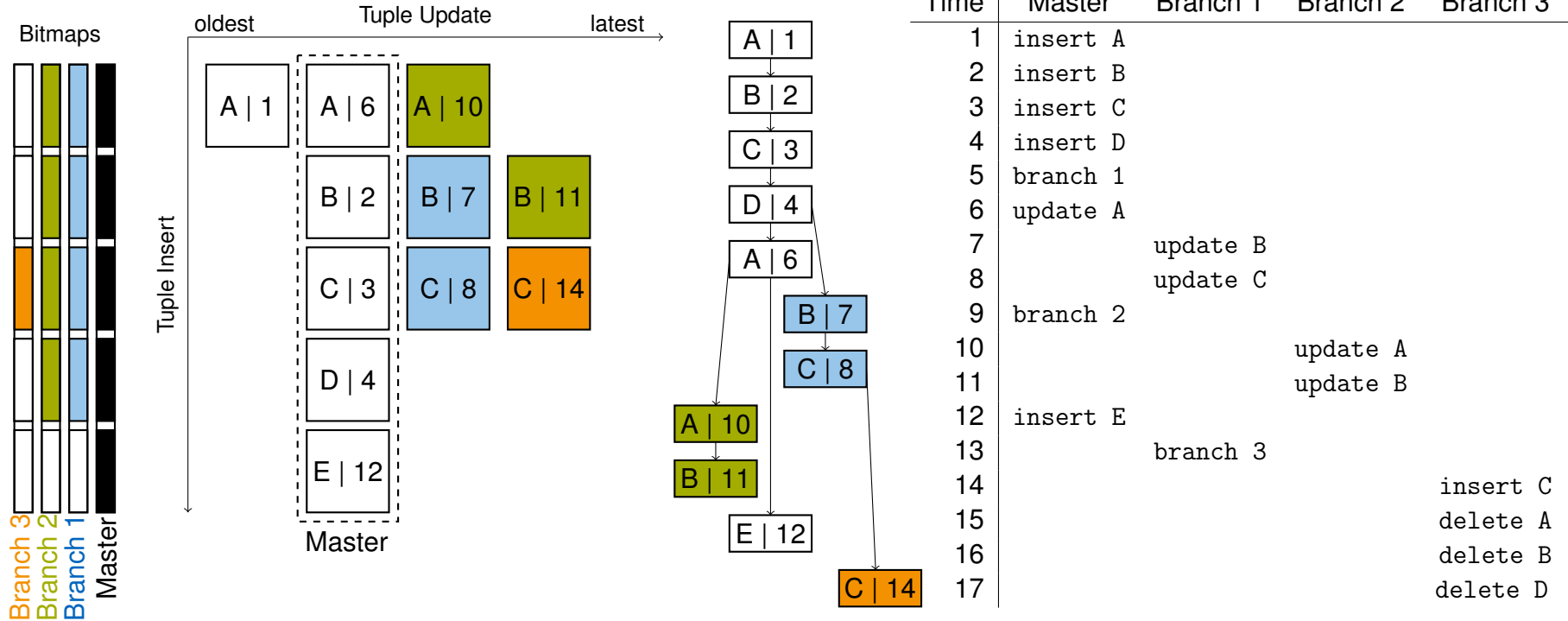




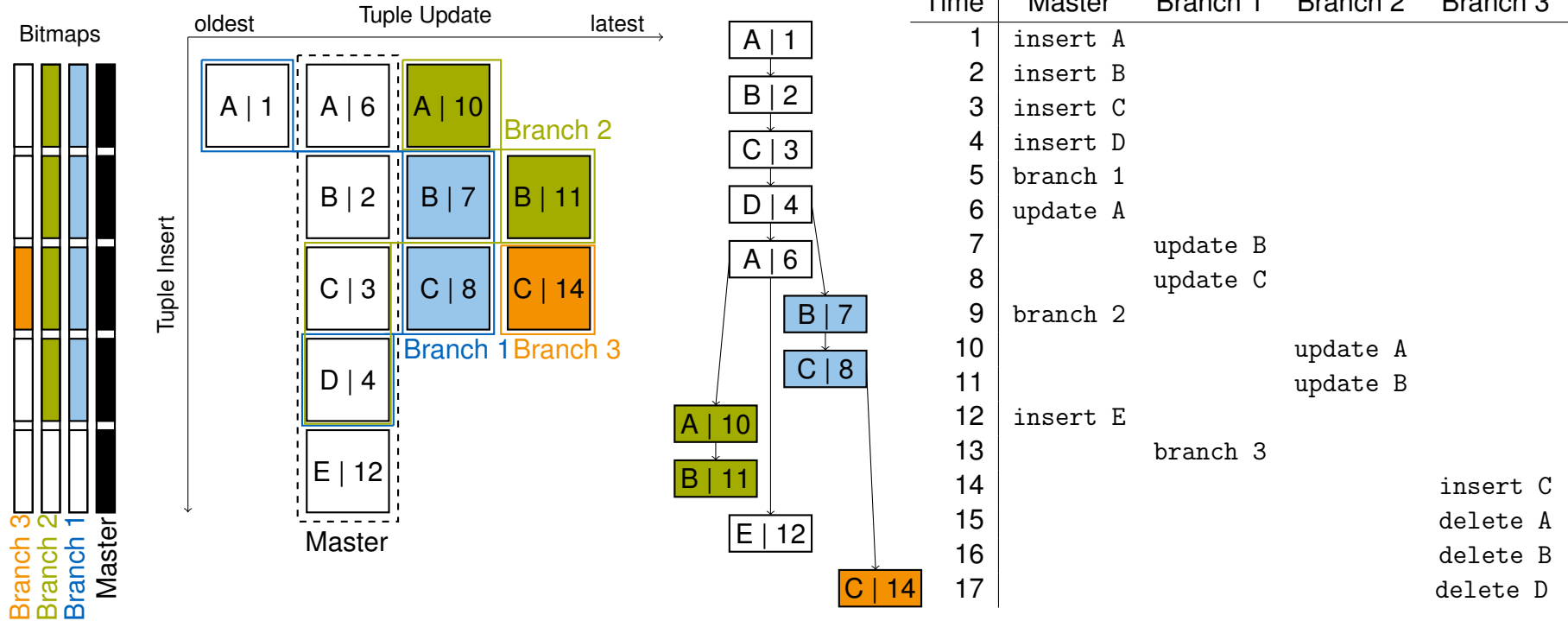
# TardisDB: Concept



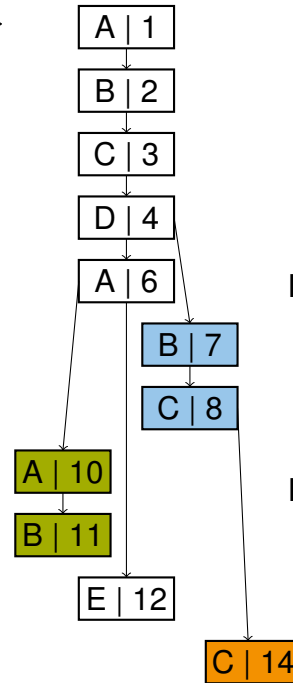
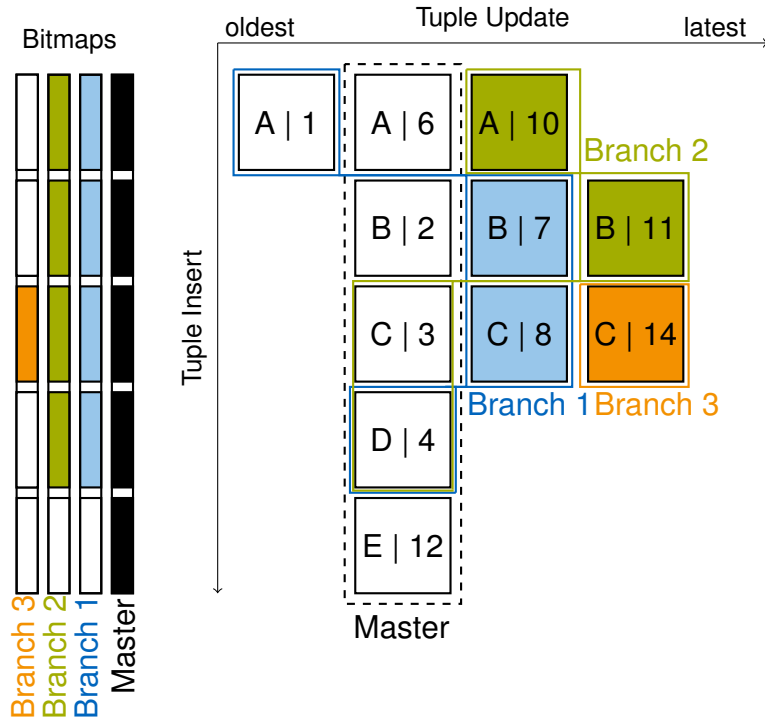
# TardisDB: Concept



# TardisDB: Concept



# TardisDB: Concept



## Produce Tuple $t$ in Branch $b$

timestamp  $ts()$  for every branch and every tuple

1. Check bit in bitmap
2. evaluate predicate  $active(t, b)$ 
  - created by the branch itself ( $created(b, t)$ )
  - or active in an ancestor branch before branching took place
3. take latest entry (with highest timestamp)

## Predicate $active()$

$active(t, b) \Leftrightarrow created(b, t) \vee$

$$\bigvee_{p \in parent(b)} active(t, p) \wedge ts(t) < ts(b)$$

## Examples

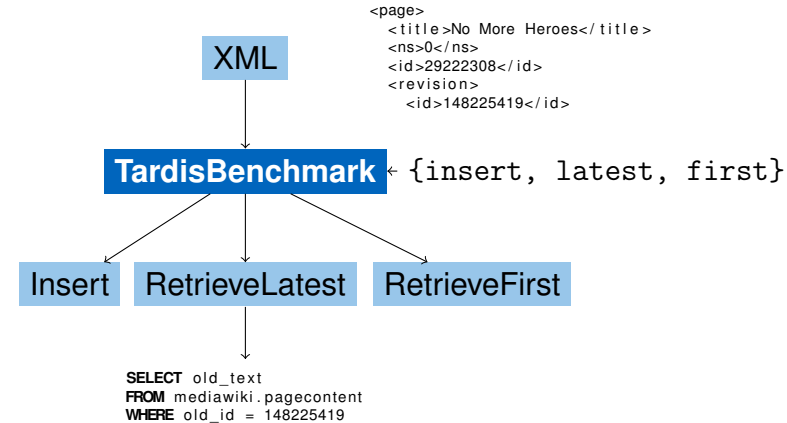
- A in Branch 3: bit not set, not visible
- C in Branch 3: visible, latest tuple is C|14
- A in Branch 1 ( $ts(b) = 5$ ): visible, A|6 not active

# TardisBenchmark

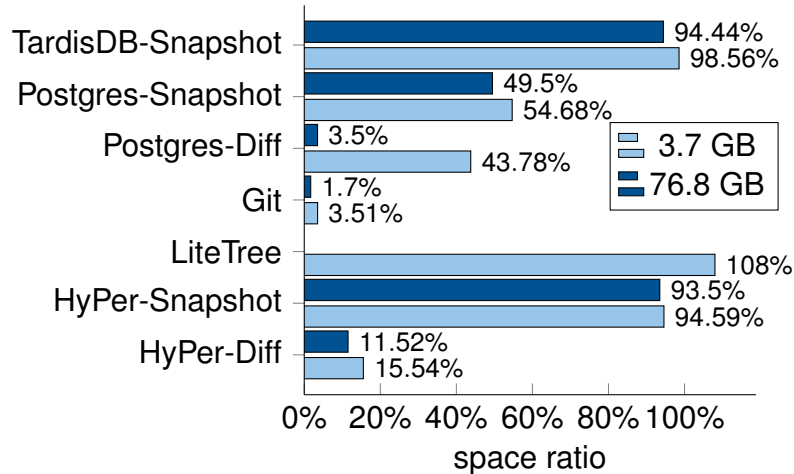


# TardisBenchmark: Concept

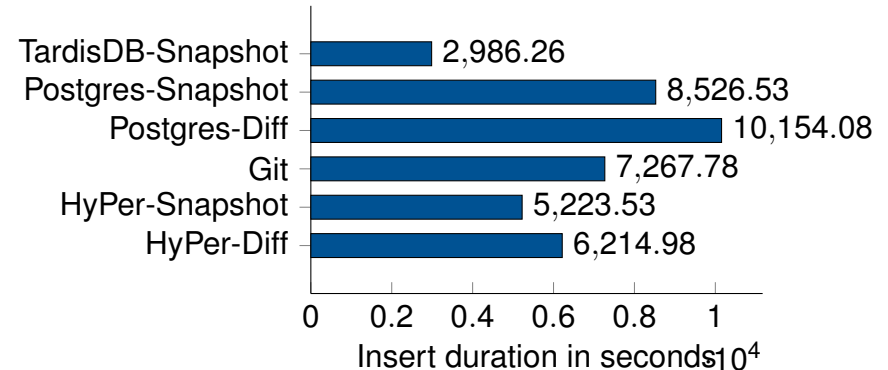
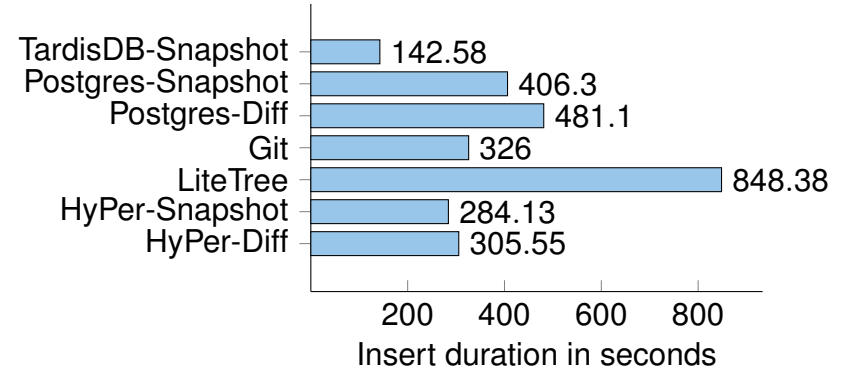
- benchmark based on Wikipedia
  - MediaWiki schema with tables for page, content and revisions
  - XML page dumps as input allows flexible workload (up to 13 TB)
- text compression methods
  - Snapshot: default in MediaWiki, stores every change as a whole
  - Diff: stores changes to the latest version as a deltas of differences
- operations
  - insert: inserts pages, compute deltas
  - retrieveLatest: retrieve latest page version
  - retrieveFirst: apply all deltas to retrieve first version
- Ubuntu 18.04 LTS, Intel Xeon CPU E5-2660 v2 processor, 2.20 GHz (20 cores), 256 GiB DDR4 RAM
- full page edit history from 1 August, 2018: pages 10 up to 2,087 (76.8 GB) and 30,227 up to 30,303 (3.7 GB)



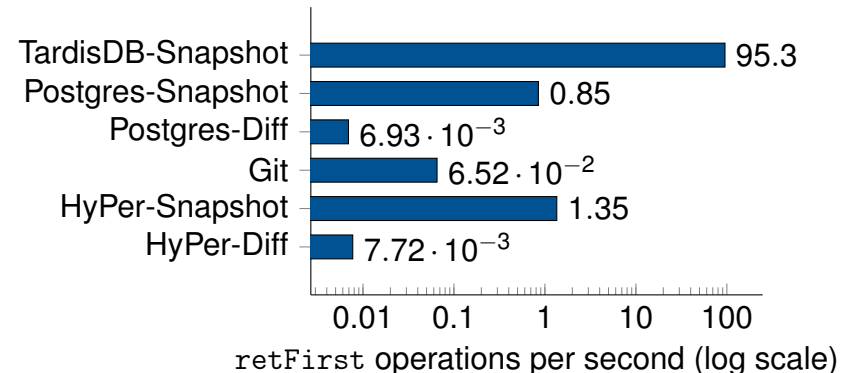
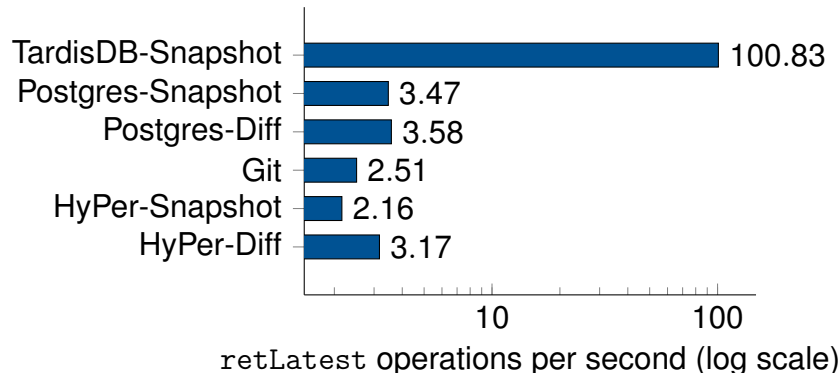
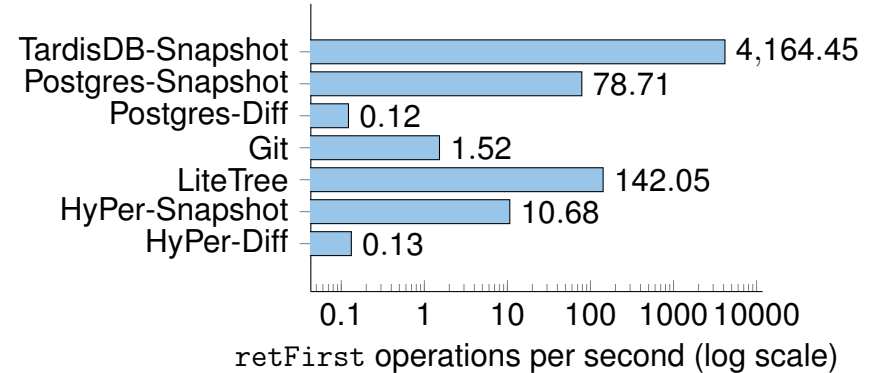
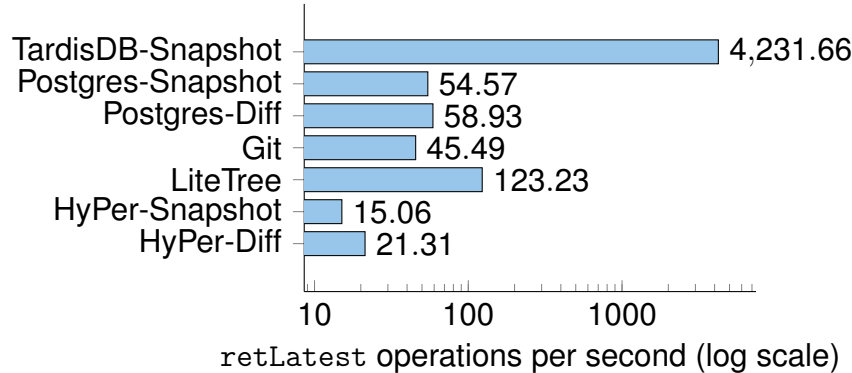
# TardisBenchmark: Insert Performance and Space Requirement



- Diff: slower, therefore less space needed
- Git: space efficient

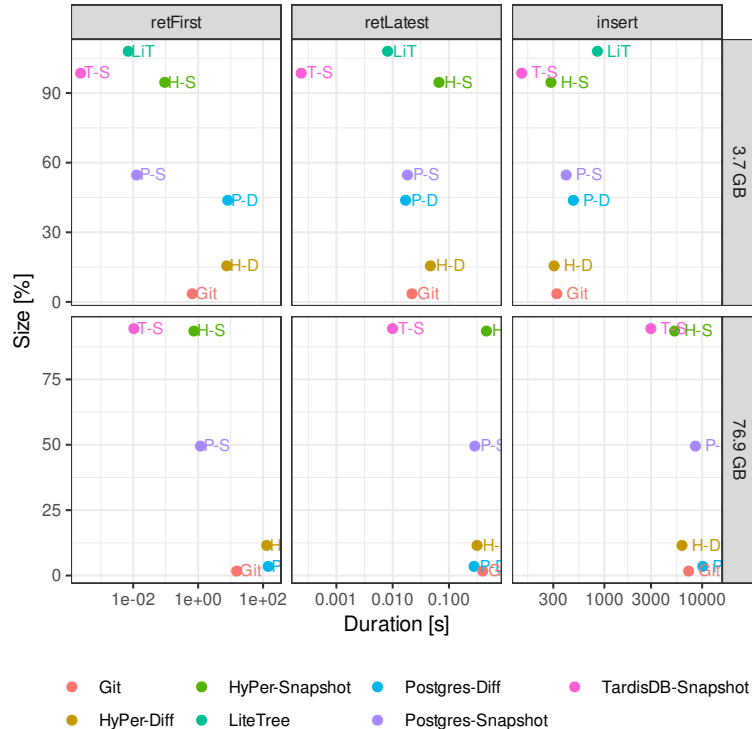


# TardisBenchmark: Retrieve Performance





# Conclusion and Future Work



## Conclusion

- TardisDB showed best performance for retrieving the latest version (as optimised for this workload)
- Git: best compression, good retrieval times
- LiteTree fast, but only allows up to 1024 branches; high space consumption
- PSQL and HyPer: Diff consumed 90 % less space than Snapshot

## Future Work

- combine version chains with difference-based compression to reduce space consumption
- enable multi-version concurrency control in TardisDB