# A TEMPORAL BENCHMARK IN BENCHBASE

Paul A. Jungwirth

PGConf.dev 2025

14 May 2025

```
1 CREATE EXTENSION btree_gist;
2
3 CREATE TABLE employees (
4   id     int GENERATED BY DEFAULT AS IDENTITY,
5   valid_at daterange,
6   name    text NOT NULL,
7   salary   int NOT NULL,
8   CONSTRAINT employees_pkey
9     PRIMARY KEY (id, valid_at WITHOUT OVERLAPS)
10 );
```

```
1 CREATE EXTENSION btree_gist;
2
3 CREATE TABLE employees (
4   id     int GENERATED BY DEFAULT AS IDENTITY,
5   valid_at daterange,
6   name    text NOT NULL,
7   salary   int NOT NULL,
8   CONSTRAINT employees_pkey
9    PRIMARY KEY (id, valid_at WITHOUT OVERLAPS)
10 );
```

```
1 CREATE EXTENSION btree_gist;
2
3 CREATE TABLE employees (
4   id     int GENERATED BY DEFAULT AS IDENTITY,
5   valid_at daterange,
6   name    text NOT NULL,
7   salary   int NOT NULL,
8   CONSTRAINT employees_pkey
9    PRIMARY KEY (id, valid_at WITHOUT OVERLAPS)
10 );
```

```
1 CREATE EXTENSION btree_gist;
2
3 CREATE TABLE employees (
4   id     int GENERATED BY DEFAULT AS IDENTITY,
5   valid_at daterange,
6   name    text NOT NULL,
7   salary   int NOT NULL,
8   CONSTRAINT employees_pkey
9     PRIMARY KEY (id, valid_at WITHOUT OVERLAPS)
10 );
```

```
EMPLOYEES(SSN, LAST_NAME, FIRST_NAME, ANNUAL_SALARY)

POSITIONS(PCN, JOB_TITLE_CODE1)

INCUMBENTS(SSN, PCN, START_DATE, END_DATE)

JOB_TITLES(JOB_TITLE_CODE, JOB_TITLE)
```

Richard Snodgrass, University of Arizona employee database

of the PostgreSQL server, such as maximum memory for sorting, were kept to default values, and no indexes were used.

We use the real world dataset *Incumben* of the University of Arizona with 83,857 entries. Each entry records a job assignment (*pcn*) for an employee (*ssn*) over a specific time interval. The data ranges over 16 years and contains 49,195 different employees. The timestamps are recorded at the granularity of days and range from 1 to 573 days with an average of approximately 180 days. Synthetic datasets used in the evaluation are described below.

#### 7.2 Database System Integration

Temporal normalization and temporal alignment fully leverage

of the PostgreSQL server, such as maximum memory for sorting, were kept to default values, and no indexes were used.

We use the real world dataset *Incumben* of the University of Arizona with 83,857 entries. Each entry records a job assignment (pcn) for an employee (ssn) over a specific time interval. The data ranges over 16 years and contains 49,195 different employees. The timestamps are recorded at the granularity of days and range from 1 to 573 days with an average of approximately 180 days. Synthetic datasets used in the evaluation are described below.

#### 7.2 Database System Integration

Temporal normalization and temporal alignment fully leverage

# BENCHBASE NÉE OLTP-BENCH

#### OLTP-Bench: An Extensible Testbed for Benchmarking Relational Databases

Djellel Eddine Difallah U. of Fribourg, Switzerland djelleleddine.difallah@unifr.ch

Carlo Curino
Microsoft Corporation, USA
ccurino@microsoft.com

Andrew Pavlo
Carnegie Mellon University, USA
pavlo@cs.cmu.edu

Philippe Cudre-Mauroux
U. of Fribourg, Switzerland
pcm@unifr.ch

#### **ABSTRACT**

Benchmarking is an essential aspect of any database management system (DBMS) effort. Despite several recent advancements, such To overcome this problem, it is imperative that application developers use a testing environment that is *stable*, *controlled* and *repeatable* [19]. In the context of DBMSs, this is achieved through the use of a *benchmark* that allows one to measure key performance



#### DDL

```
CREATE TABLE employees (
              int GENERATED BY DEFAULT AS IDENTITY NOT NULL,
   id
   valid_at daterange NOT NULL,
   name text NOT NULL,
   salary int NOT NULL,
   PRIMARY KEY (id, valid_at WITHOUT OVERLAPS)
CREATE TABLE positions (
   id
              int GENERATED BY DEFAULT AS IDENTITY NOT NULL,
   valid_at daterange NOT NULL,
   name text NOT NULL,
   employee_id int NOT NULL,
   PRIMARY KEY (id, valid_at WITHOUT OVERLAPS),
   FORFIGN KEY (employee id PFRIOD valid at) REFERENCES empl
```

# InsertPosition

```
package com.oltpbenchmark.benchmarks.temporal.procedures;
   public class InsertPosition extends Procedure {
     public final SQLStmt insertPosition =
 5
         new SQLStmt(
 6
             "INSERT INTO positions "
             + "(employee_id, valid_at, name) "
 8
             + "VALUES "
             + "(?, daterange(?, ?), concat(?, ' ', to_char(?,
             + "RETURNING id");
10
11
     public int run(Connection conn, int employeeId, String do
12
13
         throws SQLException {
       try (PreparedStatement stmt = this.getPreparedStatement
14
         stmt setInt(1 employeeId):
15
```

# InsertPosition

```
1 package com.oltpbenchmark.benchmarks.temporal.procedures;
   public class InsertPosition extends Procedure {
     public final SQLStmt insertPosition =
5
         new SQLStmt(
6
             "INSERT INTO positions "
             + "(employee_id, valid_at, name) "
8
             + "VALUES "
             + "(?, daterange(?, ?), concat(?, ' ', to_char(?,
             + "RETURNING id");
10
     public int run(Connection conn, int employeeId, String do
         throws SQLException {
       try (PreparedStatement stmt = this.getPreparedStatement
         stmt.setInt(1 emnloveeId):
```

# InsertPosition

```
+ "RETURNING id");
12
     public int run(Connection conn, int employeeId, String du
         throws SQLException {
13
       try (PreparedStatement stmt = this.getPreparedStatement
14
         stmt.setInt(1, employeeId);
15
16
         stmt.setDate(2, s == null ? null : Date.valueOf(s));
         stmt.setDate(3, e == null ? null : Date.valueOf(e));
17
         stmt.setString(4, duty);
18
         stmt.setInt(5, rank);
19
20
         stmt.execute();
21
         return -1;
22
```

#### **PROCEDURES**

DeleteEmployee.java
InsertPosition.java
Noop.java
SelectAllEmployees.java
SelectEmployeesWithOptionalPositions.java
SelectEmployeesWithoutPosition.java
SelectEmployeesWithPosition.java
SelectOneEmployee.java
SelectOneEmployeePositions.java
SelectOneEmployeeWithOptionalPositions.java
UpdateEmployee.java
UpdatePosition.java

# TemporalWorker

```
private RandomEmployee makeRandomEmployee(boolean gaussianEmpl
 int id = model.gaussianEmployeeId(rng());
  LocalDate s;
  LocalDate e;
 if (TemporalConstants.CHECK_FK_GAUSSIAN_RANGE) {
    s = model.today.plusDays(model.gaussianDays(rng(), (int) M
  } else {
    s = model.today.plusDays(-rng().nextInt(365 * TemporalCons
 // Pick a range from 1 day to 2 years:
 e = s.plusDays(1 + rng().nextInt(365 * 2));
 return new RandomEmployee(id, s, e)
```

#### FOREIGN KEYS

- EXISTS
- lag
- range\_agg

#### EXISTSIMPL

```
SELECT 1
-- There was a PK when the FK started:
WHERE EXISTS
 SELECT 1
  FROM [ONLY] <pktable>
 WHERE pkatt1 = $1 [AND ...]
          COALESCE(lower(pkperiodatt), '-Infinity')
  AND
       <= COALESCE(lower($n), '-Infinity')</pre>
          COALESCE(lower($n), '-Infinity')
  AND
       < COALESCE(upper(pkperiodatt), 'Infinity')
  FOR KEY SHARE OF X
. . .
```

#### EXISTSIMPL

#### **EXISTS IMPL**

#### EXISTSIMPL

```
. . .
AND
       NOT EXISTS (
 SELECT 1
 FROM [ONLY] <pktable> AS pk2
 WHERE
         pk1.pkatt1 = pk2.pkatt1 [AND ...]
         -- but skip pk1.pkperiodatt && pk2.pkperiodatt
         COALESCE(lower(pk2.pkperiodatt), '-Infinity')
 AND
      <= COALESCE(upper(pk1.pkperiodatt), 'Infinity')
          COALESCE(upper(pk1.pkperiodatt), 'Infinity')
         COALESCE(upper(pk2.pkperiodatt), 'Infinity')
 FOR KEY SHARE OF pk2
FOR KEY SHARE OF pk1
```

#### **EXISTS IMPL**

```
Result (cost=33.28..33.29 rows=1 width=4)
      One-Time Filter: ((InitPlan 1).col1 AND (InitPlan 2).col
      InitPlan 1
        -> LockRows (cost=0.28..8.32 rows=1 width=10)
 5
              -> Index Scan using employees_pkey on employees
 6
                    Index Cond: ((id = 24374) AND (valid at &&
                    Filter: ((COALESCE(lower(valid_at), '-infi
8
      InitPlan 2
        -> LockRows (cost=0.28..8.32 rows=1 width=10)
10
                  Index Scan using employees_pkey on employees
                    Index Cond: ((id = 24374) AND (valid at &&
11
                    Filter: ((COALESCE(lower(valid_at), '-infi
12
13
      InitPlan 4
14
        -> LockRows (cost=0.28..16.64 rows=1 width=10)
                  Index Scan using employees nkey on employees
15
```

#### EXISTSIMPL

```
Index Scan using employees_pkey on employees
                   Index Cond: ((id = 24374) AND (valid at &&
                   Filter: ((COALESCE(lower(valid_at), '-infi
     InitPlan 4
       -> LockRows (cost=0.28..16.64 rows=1 width=10)
             -> Index Scan using employees_pkey on employees
                   Index Cond: ((id = 24374) AND (valid at &&
                   Filter: (('2025-05-11'::date < COALESCE(up
                   SubPlan 3
                     -> LockRows (cost=0.28..8.32 rows=1 wi
                          -> Index Scan using employees_pke
                                Index Cond: (id = pk1.id)
                                Filter: ((COALESCE(lower(va)
23
   (22 rows)
```

#### EXISTSIMPL

```
Result (cost=33.28..33.29 rows=1 width=4)
      One-Time Filter: ((InitPlan 1).col1 AND (InitPlan 2).col
 3
      InitPlan 1
        -> LockRows (cost=0.28..8.32 rows=1 width=10)
              -> Index Scan using employees_pkey on employees
                    Index Cond: ((id = 24374) AND (valid at &&
                    Filter: ((COALESCE(lower(valid_at), '-infi
8
      InitPlan 2
        -> LockRows (cost=0.28..8.32 rows=1 width=10)
              -> Index Scan using employees_pkey on employees
                    Index Cond: ((id = 24374) AND (valid at &&
                    Filter: ((COALESCE(lower(valid_at), '-infi
13
      InitPlan 4
        -> LockRows (cost=0.28..16.64 rows=1 width=10)
                  Index Scan using employees_pkey on employees
```

```
1 SELECT 1
 2 FROM
     SELECT
             uk.uk start value,
             uk.uk end value,
             NULLIF(LAG(uk.uk_end_value) OVER (ORDER BY uk.uk_start_value), uk.uk_sta
     FROM
               coalesce(lower(x.pkperiodatt), '-Infinity') AS uk_start_value,
       SELECT
               coalesce(upper(x.pkperiodatt), 'Infinity') AS uk end value
 8
       FROM
               pktable AS x
10
       WHERE
               pkatt1 = $1 [AND ...]
11
               uk.pkperiodatt && $n
12
       FOR KEY SHARE OF X
13
     ) AS uk
14 ) AS uk
15 WHERE
           uk.uk start value < upper($n)
16 AND
           uk.uk end value >= lower($n)
           MIN(uk.uk start value) <= lower($n)</pre>
17 HAVING
           MAX(uk.uk_end_value) >= upper($n)
18 AND
19 AND
           array agg(uk.x) FILTER (WHERE uk.x IS NOT NULL) IS NULL
```

```
uk.uk end value,
               coalesce(upper(x.pkperiodatt), 'Infinity') AS uk_end_value
               pktable AS x
               pkatt1 = $1 [AND ...]
               uk.pkperiodatt && $n
11
       FOR KEY SHARE OF X
           uk.uk start value < upper($n)
           uk.uk end value >= lower($n)
           MIN(uk.uk start value) <= lower($n)</pre>
```

```
uk.uk end value,
               coalesce(upper(x.pkperiodatt), 'Infinity') AS uk_end_value
               pktable AS x
               pkatt1 = $1 [AND ...]
               uk.pkperiodatt && $n
       FOR KEY SHARE OF X
           uk.uk start value < upper($n)
           uk.uk end value >= lower($n)
17 HAVING
           MIN(uk.uk start value) <= lower($n)</pre>
```

```
uk.uk end value,
               coalesce(upper(x.pkperiodatt), 'Infinity') AS uk_end_value
               pktable AS x
               pkatt1 = $1 [AND ...]
               uk.pkperiodatt && $n
       FOR KEY SHARE OF X
           uk.uk start value < upper($n)
           uk.uk end value >= lower($n)
17 HAVING
           MIN(uk.uk start value) <= lower($n)</pre>
18 AND
           MAX(uk.uk_end_value) >= upper($n)
```

```
uk.uk end value,
             NULLIF(LAG(uk.uk_end_value) OVER (ORDER BY uk.uk_start_value), uk.uk_sta
               coalesce(upper(x.pkperiodatt), 'Infinity') AS uk end value
               pktable AS x
               pkatt1 = $1 [AND ...]
               uk.pkperiodatt && $n
       FOR KEY SHARE OF X
           uk.uk start value < upper($n)
           uk.uk end value >= lower($n)
17 HAVING
           MIN(uk.uk start value) <= lower($n)</pre>
           MAX(uk.uk_end_value) >= upper($n)
18 AND
           array_agg(uk.x) FILTER (WHERE uk.x IS NOT NULL) IS NULL
19 AND
```

```
Aggregate (cost=8.38..8.40 rows=1 width=4)
  Filter: ((array_agg(uk.x) FILTER (WHERE (uk.x IS NOT NULL))
      Subquery Scan on uk (cost=8.33..8.37 rows=1 width=12)
        Filter: ((uk.uk_start_value < '2026-05-26'::date) AND
        -> WindowAgg (cost=8.33..8.36 rows=1 width=12)
              Window: w1 AS (ORDER BY uk_1.uk_start_value)
              -> Sort (cost=8.33..8.34 rows=1 width=8)
                    Sort Key: uk_1.uk_start_value
                    -> Subquery Scan on uk_1 (cost=0.28.8.
                          -> LockRows (cost=0.28..8.31 rows
                                -> Index Scan using employee
                                      Index Cond: ((id = 2437)
(12 rows)
```

```
1 SELECT 1
2 FROM (
3   SELECT pkperiodatt AS r
4   FROM [ONLY] pktable x
5   WHERE pkatt1 = $1 [AND ...]
6   AND   pkperiodatt && $n
7   FOR KEY SHARE OF x
8 ) x1
9 HAVING $n <@ range_agg(x1.r)</pre>
```

```
1 SELECT 1
2 FROM (
3   SELECT pkperiodatt AS r
4   FROM [ONLY] pktable x
5   WHERE pkatt1 = $1 [AND ...]
6   AND   pkperiodatt && $n
7   FOR KEY SHARE OF x
8 ) x1
9 HAVING $n <@ range_agg(x1.r)</pre>
```

```
1 SELECT 1
2 FROM (
3   SELECT pkperiodatt AS r
4   FROM [ONLY] pktable x
5   WHERE pkatt1 = $1 [AND ...]
6   AND   pkperiodatt && $n
7   FOR KEY SHARE OF x
8 ) x1
9 HAVING $n <@ range_agg(x1.r)</pre>
```

```
1 SELECT 1
2 FROM (
3   SELECT pkperiodatt AS r
4   FROM [ONLY] pktable x
5   WHERE pkatt1 = $1 [AND ...]
6   AND   pkperiodatt && $n
7   FOR KEY SHARE OF x
8 ) x1
9 HAVING $n <@ range_agg(x1.r)</pre>
```

```
Aggregate (cost=8.32..8.34 rows=1 width=4)
Filter: ('[2025-05-11,2026-05-26)'::daterange <@ range_agg(
-> Subquery Scan on x1 (cost=0.28..8.32 rows=1 width=10)
-> LockRows (cost=0.28..8.31 rows=1 width=16)
-> Index Scan using employees_pkey on employee
Index Cond: ((id = 24374) AND (valid_at & come)
```

### HYPOTHESIS

- range\_agg fastest
- lag nearly as fast
- EXISTS much slower

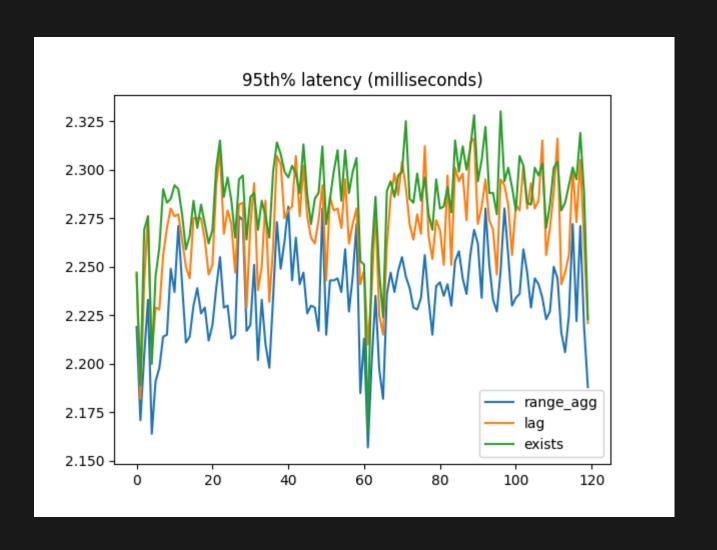
### WORKLOAD

```
<transactiontypes>
   <transactiontype>
       <name>InsertPosition</name>
   </transactiontype>
   <transactiontype>
       <name>UpdatePosition
   </transactiontype>
   <transactiontype>
       <name>UpdateEmployee</name>
   </transactiontype>
   <transactiontype>
       <name>DeleteEmployee
   </transactiontype>
</transactiontypes>
```

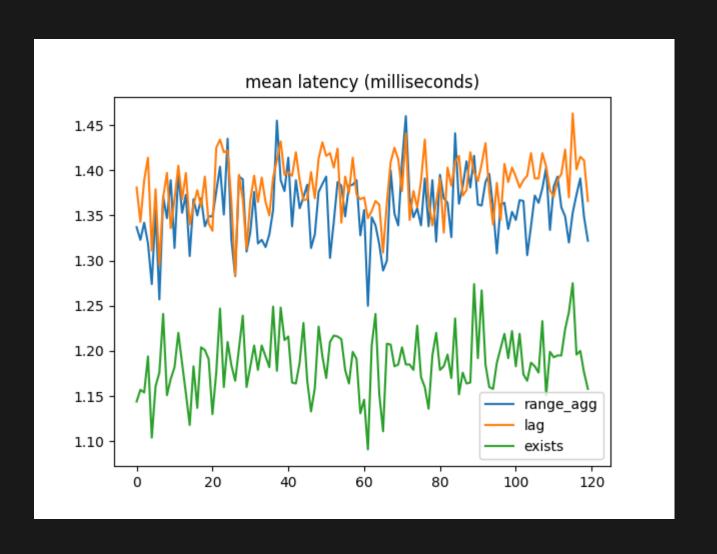
### PROCEDURES

CheckForeignKeyExists.java CheckForeignKeyLag.java CheckForeignKeyRangeAgg.java

## 95% LATENCY



### MEAN LATENCY



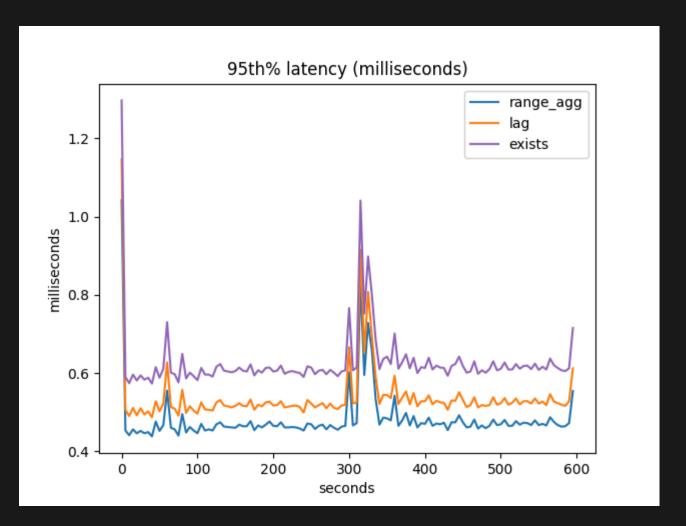
# **EXISTS IMPL**

```
Result (actual time=0.020..0.020 rows=0.00 loops=1)
      One-Time Filter: ((InitPlan 1).col1 AND (InitPlan 2).col
      InitPlan 1
        -> LockRows (actual time=0.018..0.018 rows=0.00 loops
 5
              -> Index Scan using employees_pkey on employees
 6
                     Index Cond: ((id = 24374000) AND (valid_at)
 7
                     Filter: ((COALESCE(lower(valid_at), '-infi
 8
                     Index Searches: 1
      InitPlan 2
        -> LockRows (never executed)
10
11
              -> Index Scan using employees_pkey on employees
                     Index Cond: ((id = 24374000)) AND (valid at)
12
13
                     Filter: ((COALESCE(lower(valid_at), '-infi
14
                     Index Searches: 0
15
      InitPlan /
```

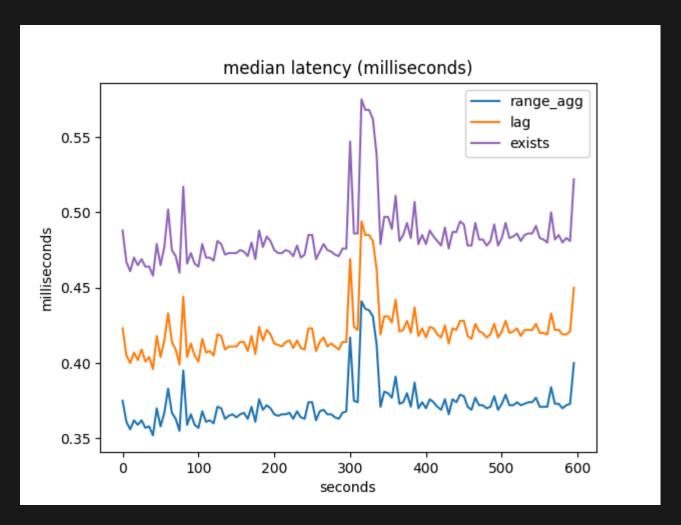
# **EXISTS IMPL**

```
9
      InitPlan 2
        -> LockRows (never executed)
10
              -> Index Scan using employees_pkey on employees
                     Index Cond: ((id = 24374000)) AND (valid at)
                     Filter: ((COALESCE(lower(valid_at), '-infi
                     Index Searches: 0
15
      InitPlan 4
        -> LockRows (never executed)
16
              -> Index Scan using employees_pkey on employees
                     Index Cond: ((id = 24374000)) AND (valid at)
                     Filter: (('2025-05-11'::date < COALESCE(up
                     Index Searches: 0
                     SubPlan 3
21
22
                       -> LockRows (never executed)
                                 Inday Scan using amployees pkg
```

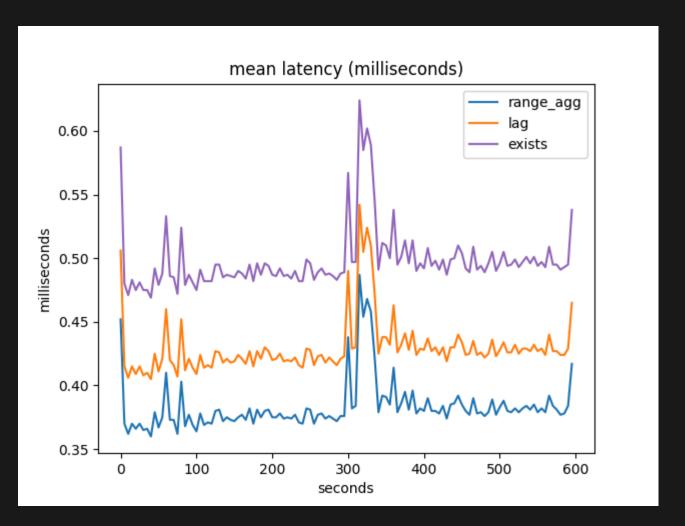
## 95% LATENCY



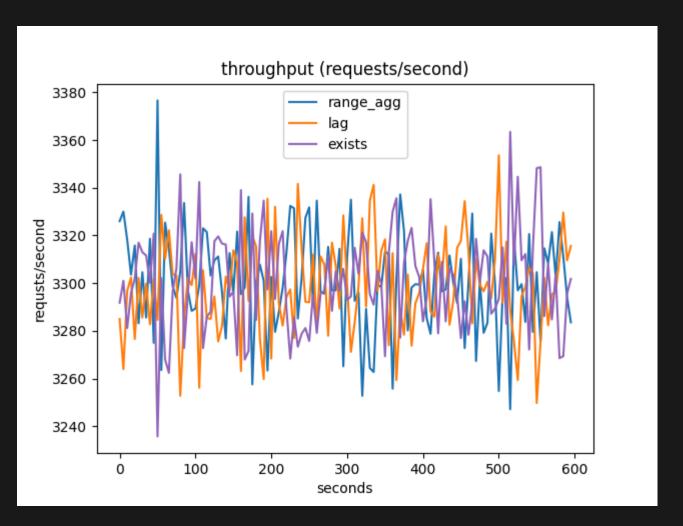
### MEDIAN LATENCY



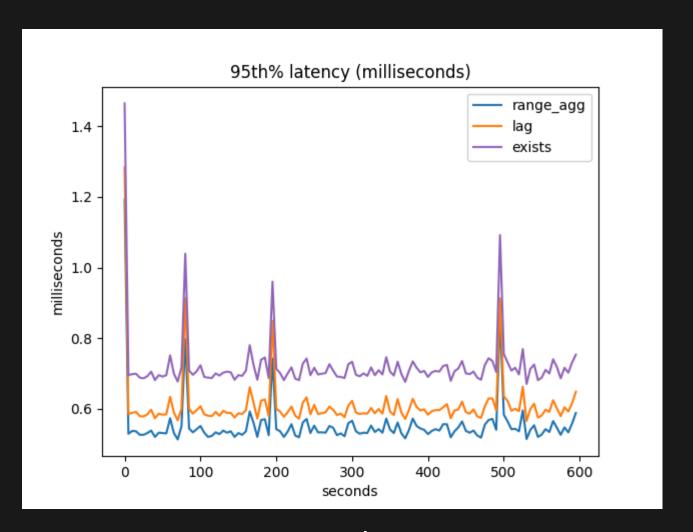
### MEAN LATENCY



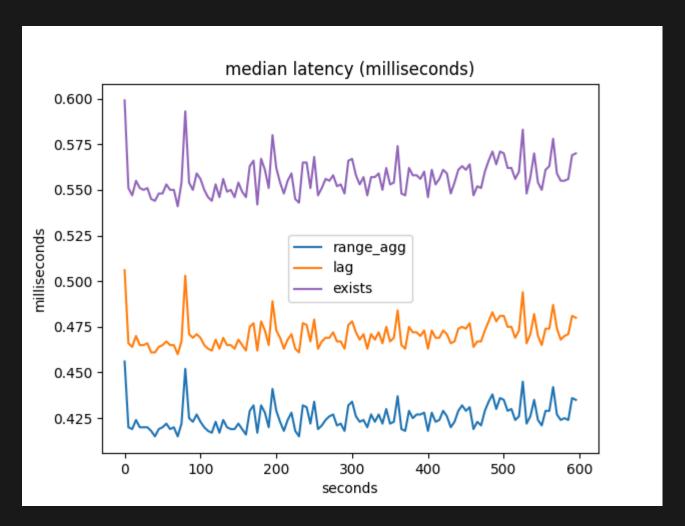
# THROUGHPUT



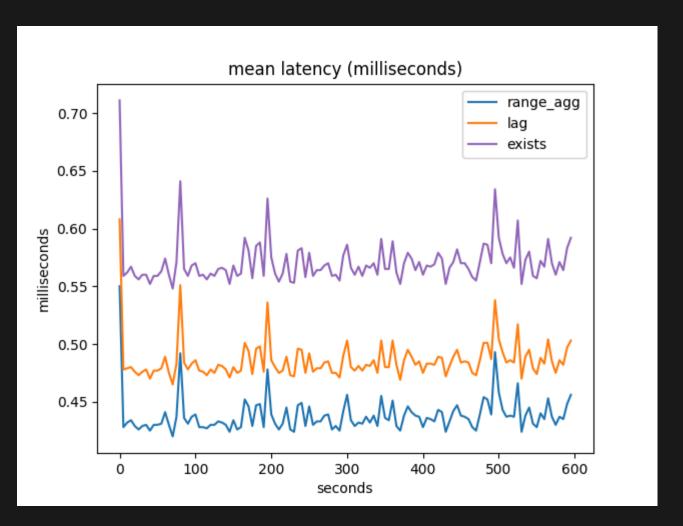
## 95% LATENCY



### MEDIAN LATENCY



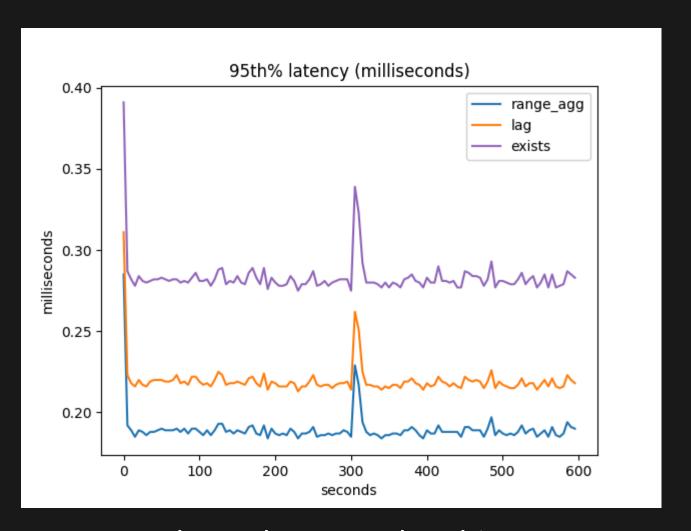
### MEAN LATENCY



### WITH HISTORY

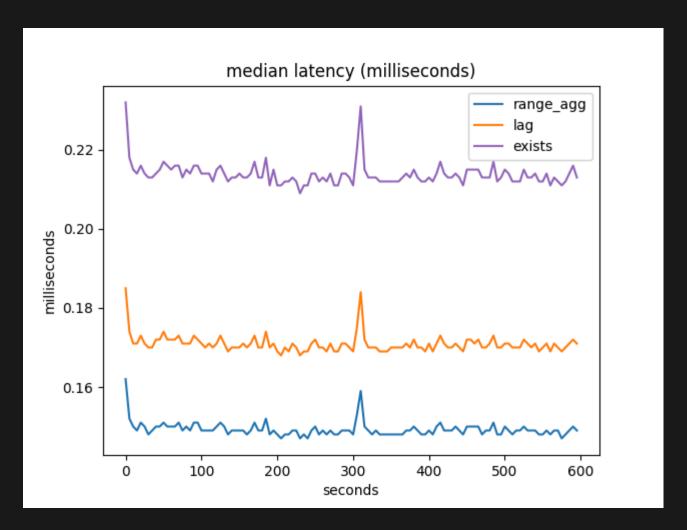
```
protected void updateEmployees(Connection conn, int lo, int hi
 String sql =
     "UPDATE employees FOR PORTION OF valid_at FROM ? TO ? "
     "SET salary = salary * 1.01 " +
     "WHERE id = ?";
 RandomDistribution.Gaussian =
     new RandomDistribution.Gaussian(this.rng(), 0, config.ge
 try (PreparedStatement employeeUpdate = conn.prepareStatemen
   // For each employee:
    for (int i = lo; i <= hi; i++) {</pre>
     int raises = gaussian.nextInt();
      LocalDate s = this.model.today;
     LocalDate e;
```

### 95% LATENCY



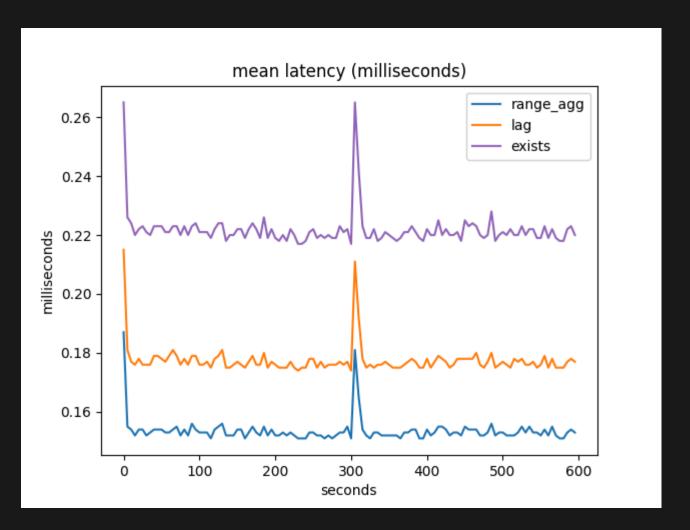
100k employees, salary history

### MEDIAN LATENCY



100k employees, salary history

### MEAN LATENCY



100k employees, salary history

### FUTURE WORK

- More experiments
- More system stats
- PERIODs
- Compare to other RDBMSes

### REFERENCES

- https://github.com/pjungwir/introducing-a-temporal-benchmarkin-benchbase
- https://dl.acm.org/doi/10.1145/2213836.2213886
- https://www.vldb.org/pvldb/vol7/p277-difallah.pdf
- https://github.com/cmu-db/benchbase
- https://github.com/pjungwir/benchbase/tree/temporal
- https://github.com/pjungwir/temporal\_ops
- https://illuminatedcomputing.com/pages/pdxpug2024benchbase-and-temporal-foreign-keys/
- https://github.com/pjungwir/postgresql/tree/valid-time
- https://github.com/pjungwir/benchmarking-temporal-tables

### THANK YOU

https://github.com/pjungwir/introducing-a-temporal-benchmark-in-benchbase