CSE 444: Database Internals

Lectures 27

NewSQL

Slides from Andrew Pavlo
Brown University

References

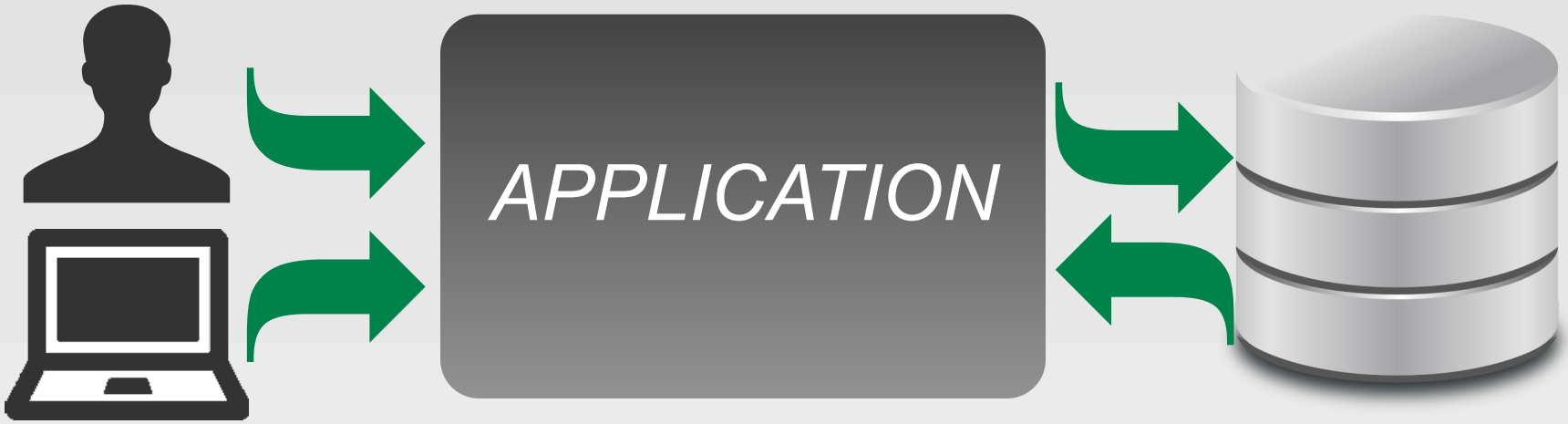
- **Scalable SQL and NoSQL Data Stores**, Rick Cattell, SIGMOD Record, December 2010 (Vol. 39, No. 4)

- **The end of an Architectural Era: (It's Time for a Complete Rewrite)**, M. Stonebraker et. al. VLDB '07

- Online documentation: H-Store

VOTER BENCHMARK

Japanese "American Idol"

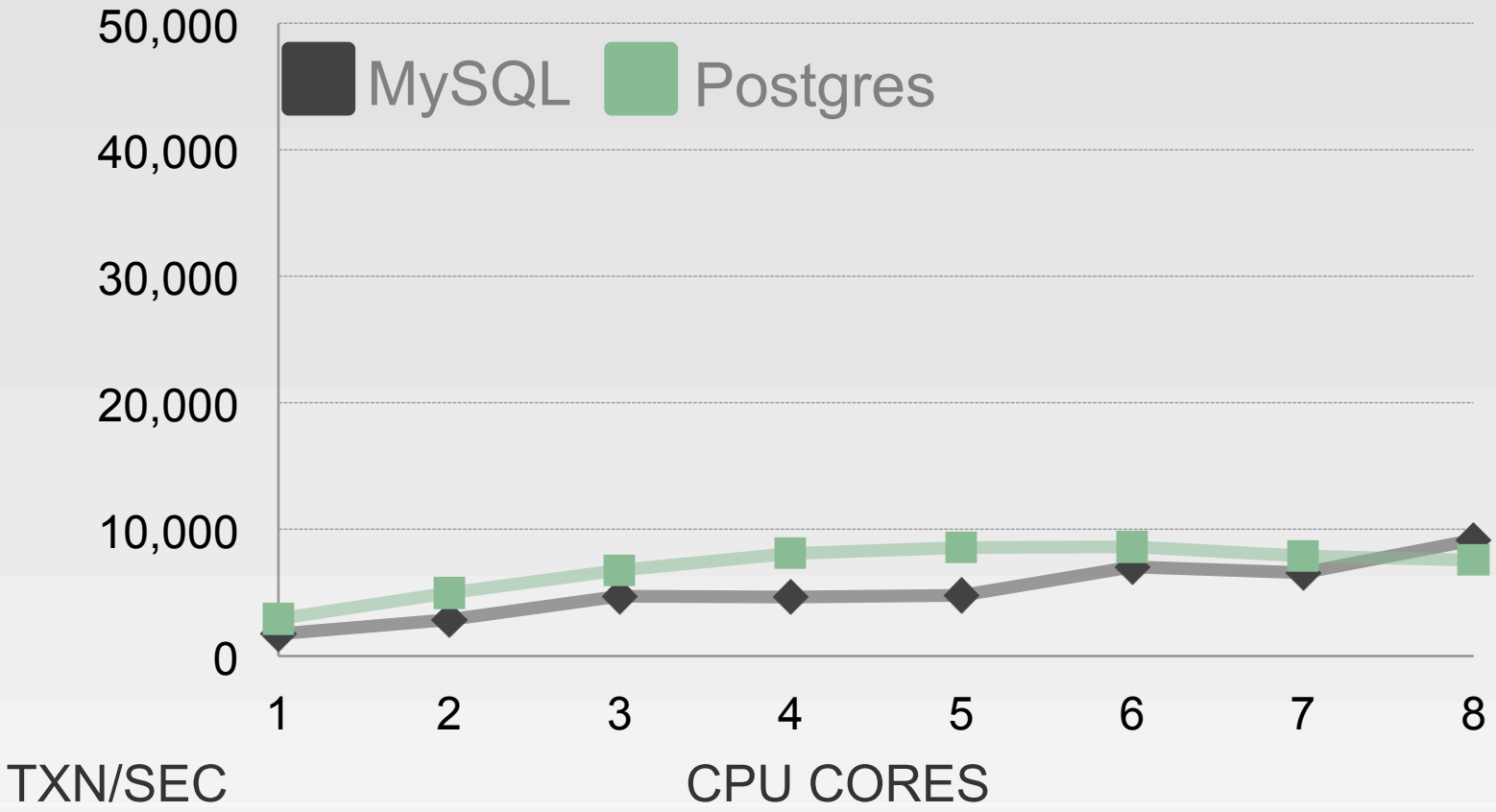


TRANSACTION

1. Check whether user has already voted.
2. Insert new vote entry.
3. Update vote count for contestant.

VOTER BENCHMARK

Japanese "American Idol"



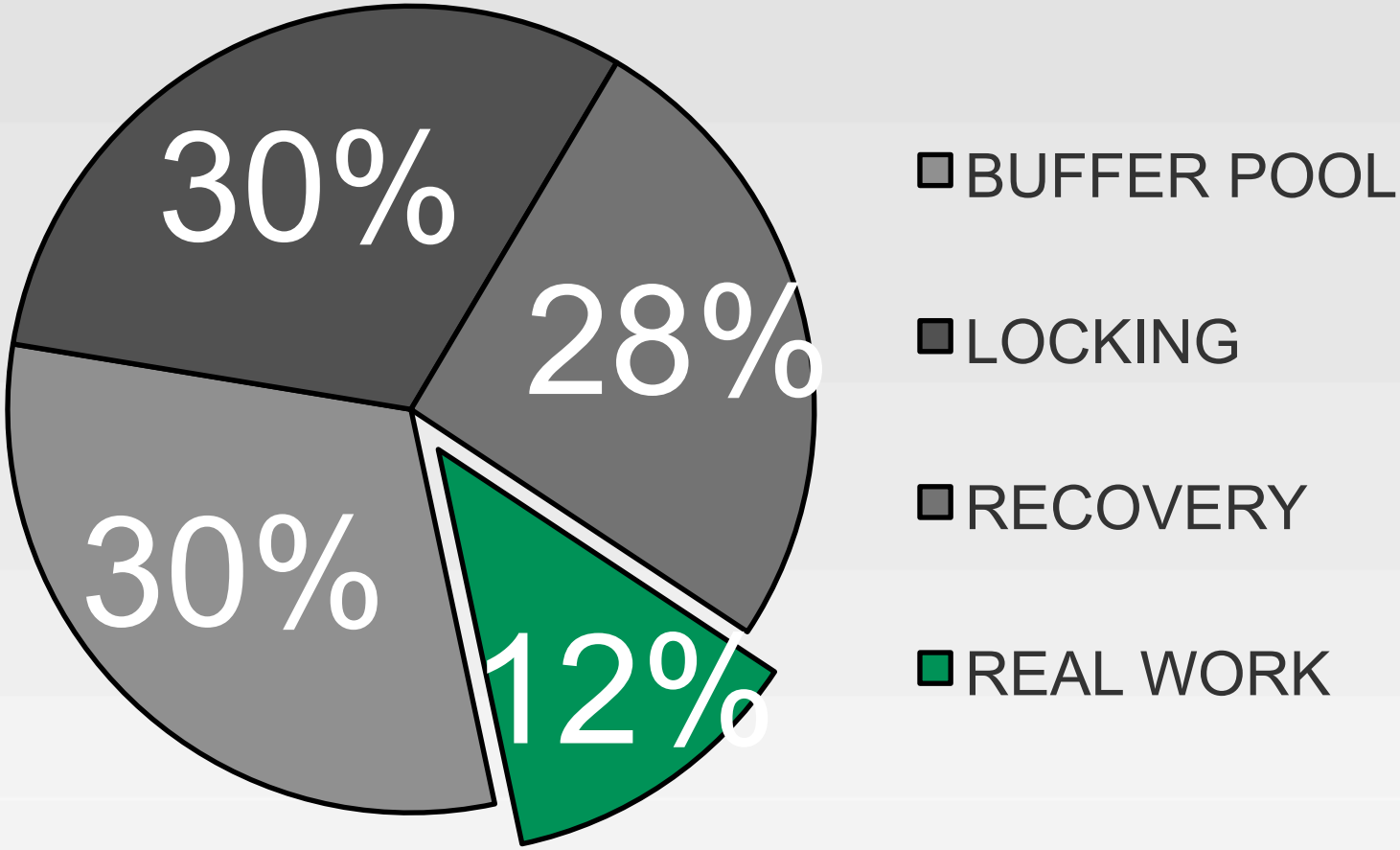
Argument from VLDB'07 paper

- Popular DBMSs based on designs from 70's
- But computer architectures are changing
- And applications have new requirements

- Past 40 years have seen extensions to DBMS design but no major re-design

TRADITIONAL DBMS

Measured CPU Cycles



OLTP THROUGH THE LOOKING GLASS,
AND WHAT WE FOUND THERE
SIGMOD, pp. 981-992, 2008.


SCALABILITY
HIGH
(Many Nodes)
NoSQL
NewSQL
LOW
(One Node)
TRADITIONAL
WEAK
(None/Limited)
GUARANTEES
STRONG
(ACID)

CAN YOU *SCALE UP* WITHOUT *GIVING UP* TRANSACTIONS?


DOG RACING

Fast *Repetitive* *Small*

Optimization

USE A LIGHTWEIGHT SYSTEM *DESIGNED* FOR OLTP TRANSACTIONS.

H-STORE: A HIGH-PERFORMANCE, DISTRIBUTED MAIN MEMORY TRANSACTION PROCESSING SYSTEM
Proc. VLDB Endow., vol. 1, iss. 2, pp. 1496-1499, 2008.

X

DISK ORIENTED

MAIN MEMORY STORAGE

i

CONCURRENT EXECUTION

SERIAL EXECUTION

/

HEAVYWEIGHT RECOVERY

COMPACT LOGGING

STORED PROCEDURE

VoteCount:

```
SELECT COUNT(*)
 FROM votes
WHERE phone_num = ?;
```

InsertVote:

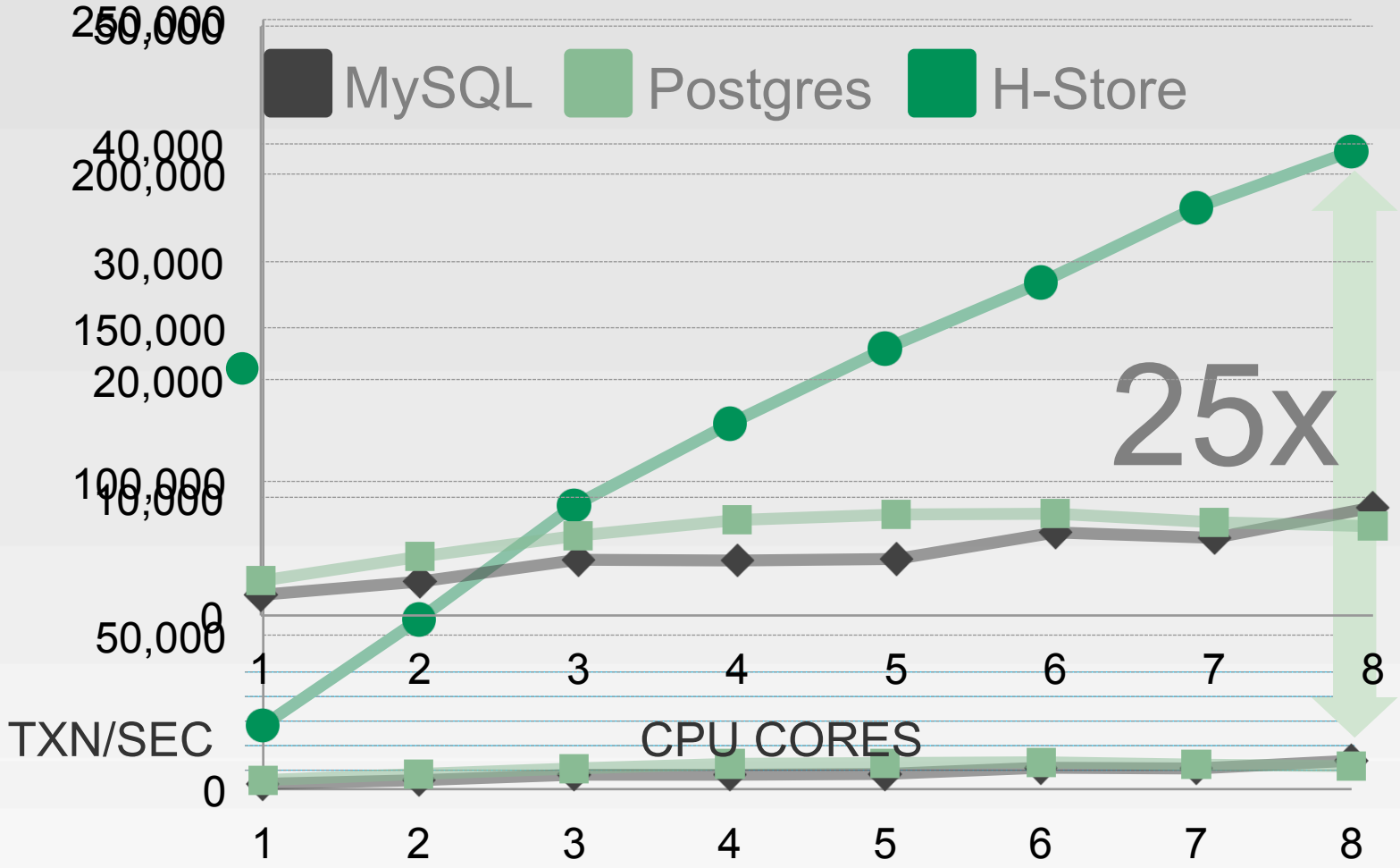
```
INSERT INTO votes
 VALUES (?, ?, ?);
```

Application

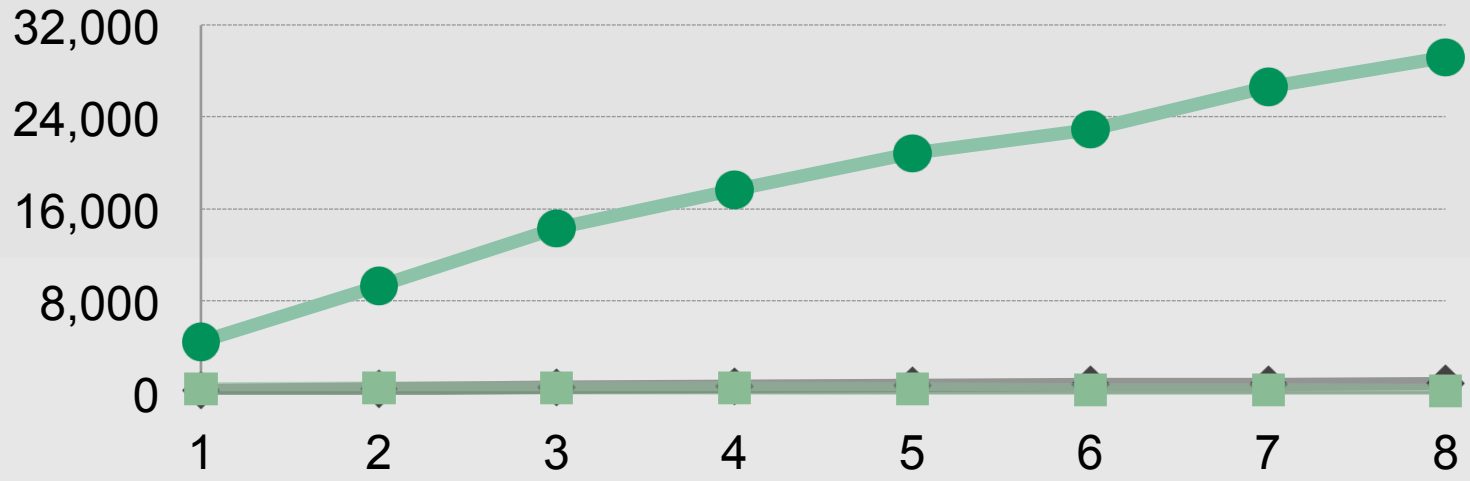
```
run(phoneNum, contestantId, currentTime) {
        result = execute(VoteCount, phoneNum);
        if (result > MAX_VOTES) {
                return (ERROR);
        }
        execute(InsertVote, phoneNum,
                        contestantId,
                        currentTime);
        return (SUCCESS);
}
```

VOTER BENCHMARK

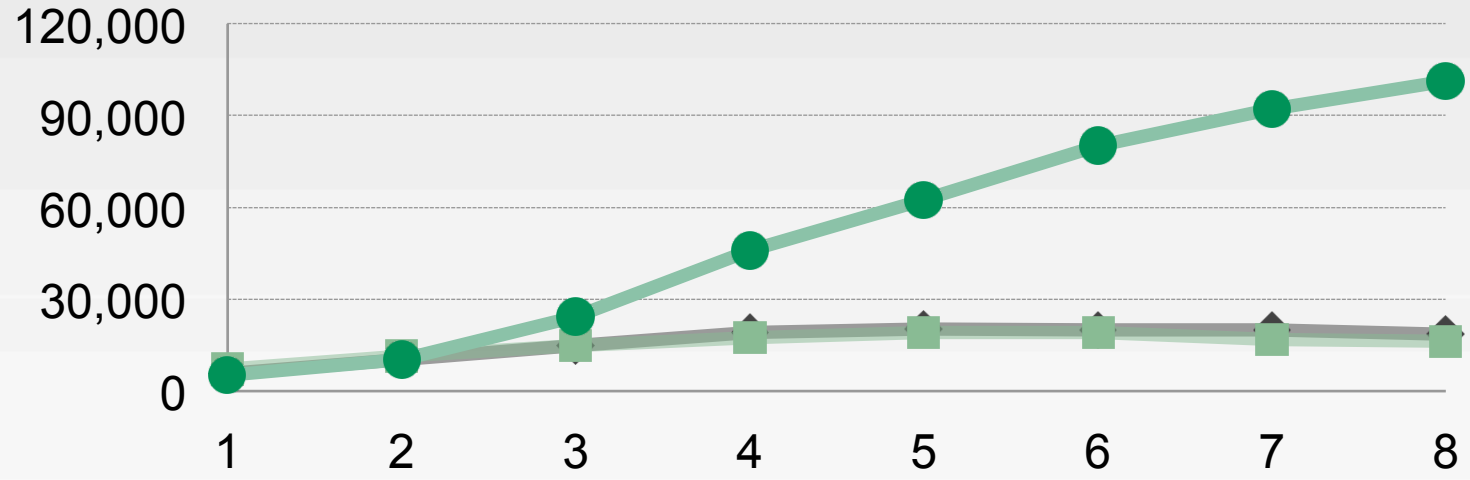
Japanese "American Idol"



TPC-C



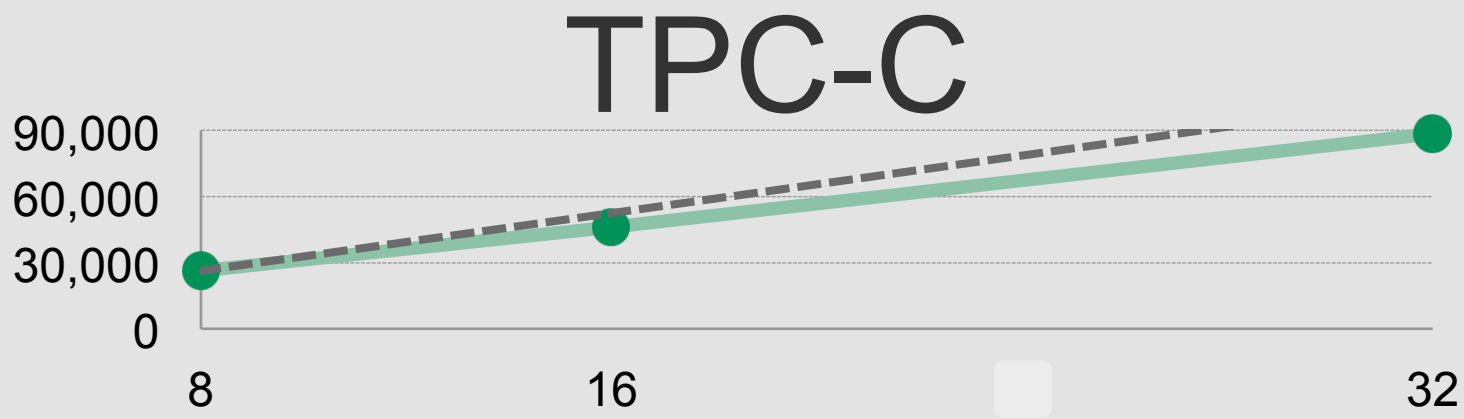
TELECOM 1

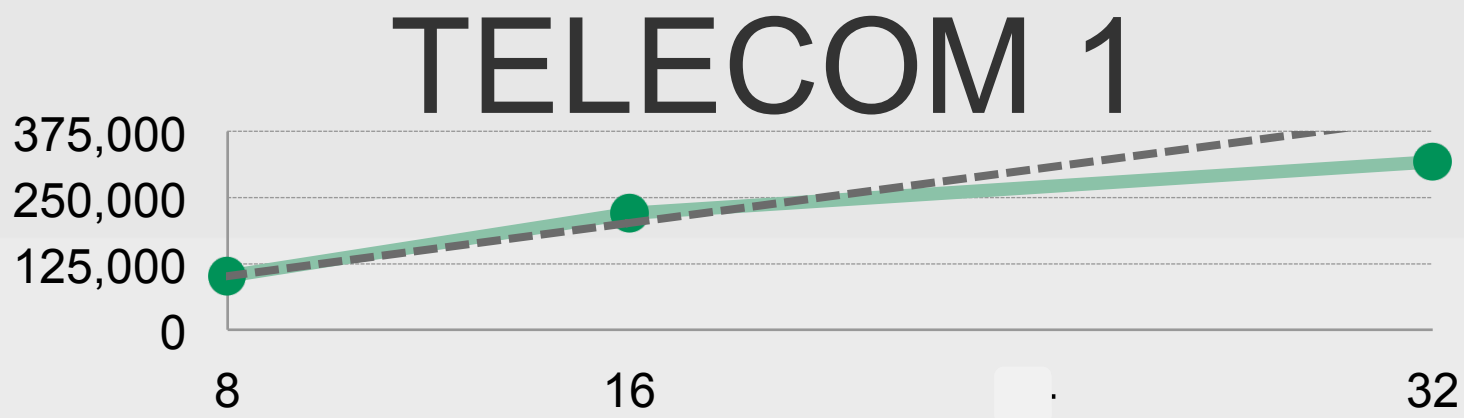


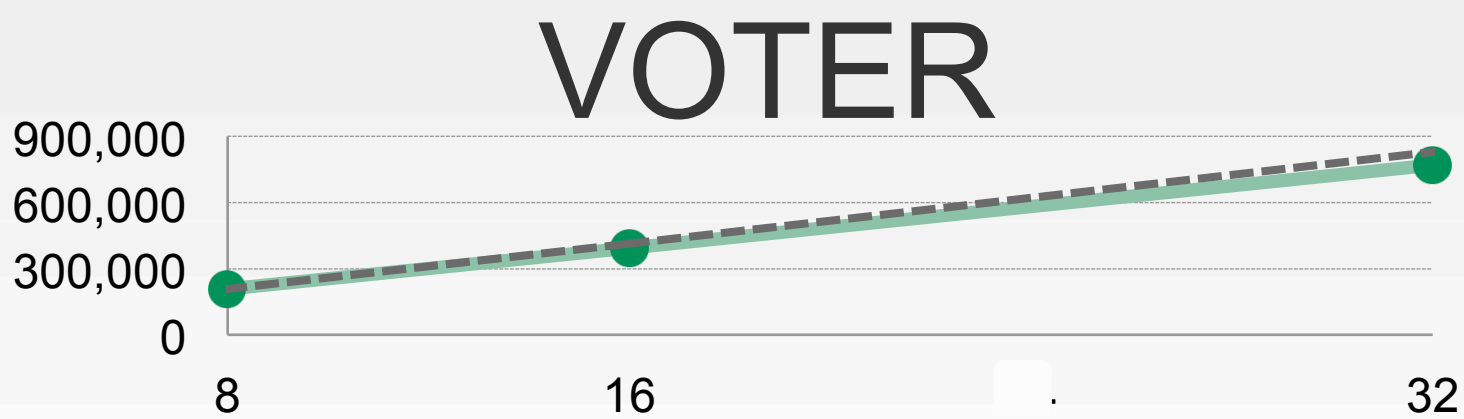
H-Store
Postgres
MySQL

TXN/SEC

CPU CORES







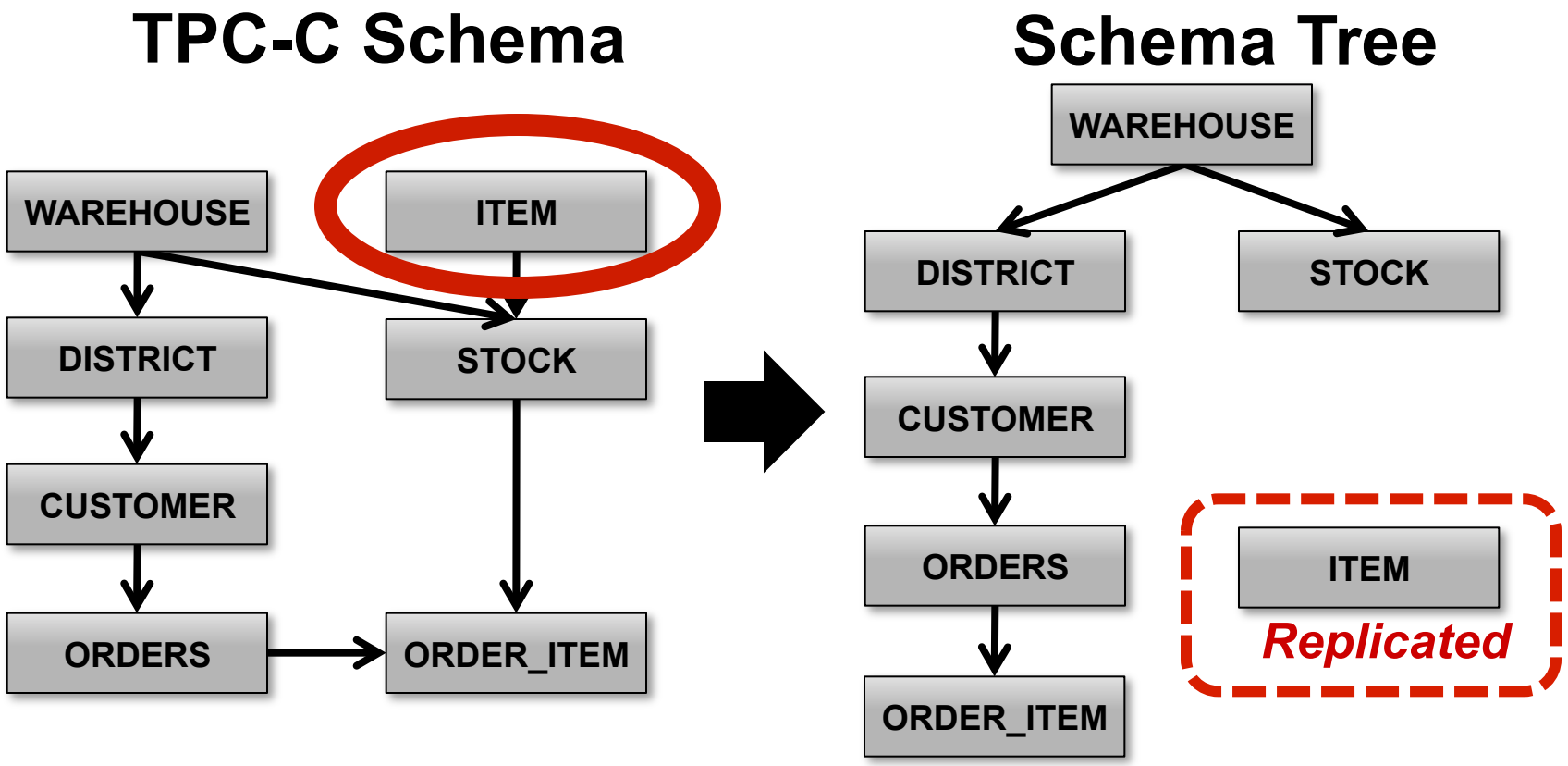
TXN/SEC

CPU CORES

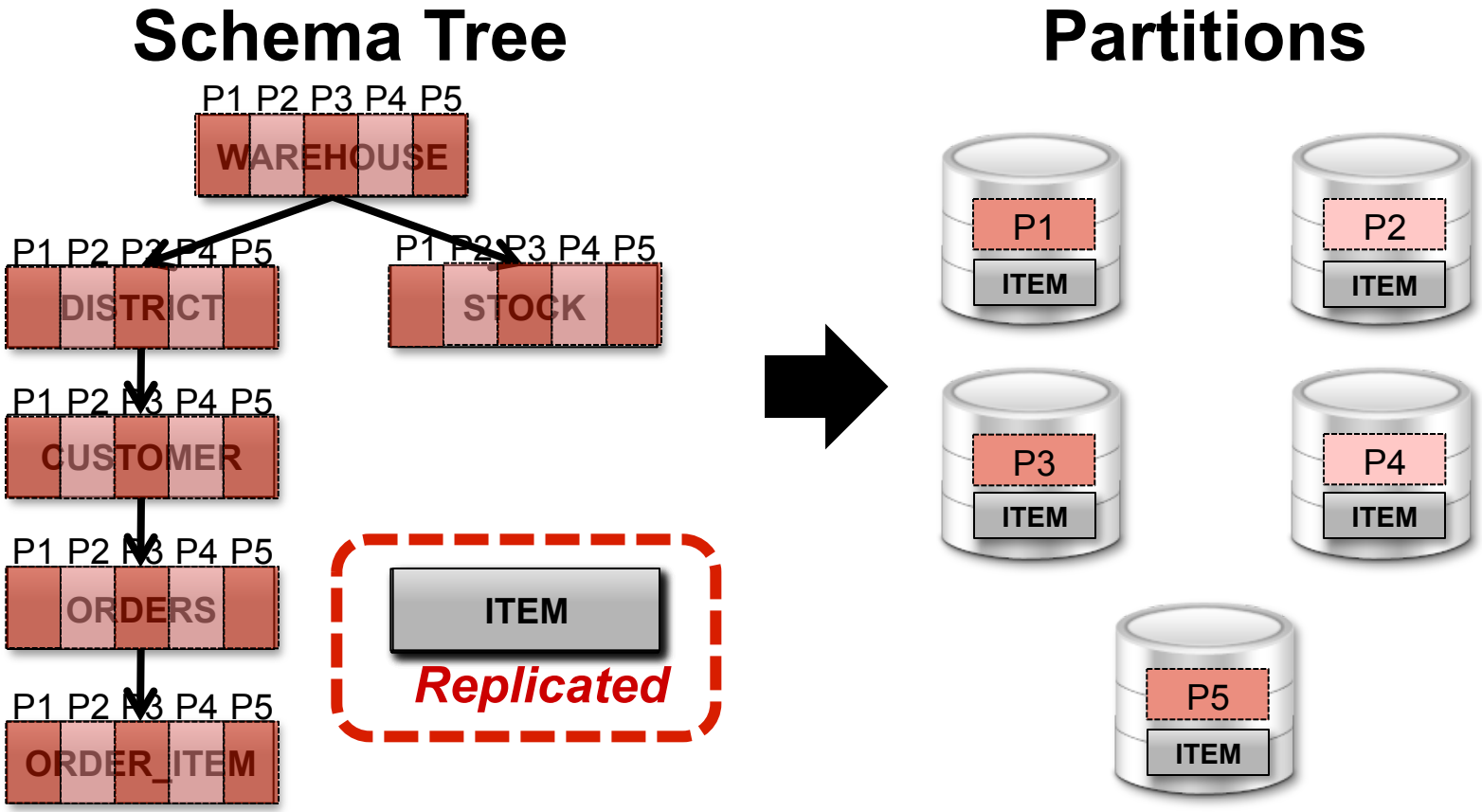
Distributed Transactions

- Discussion based on VLDB'07 paper

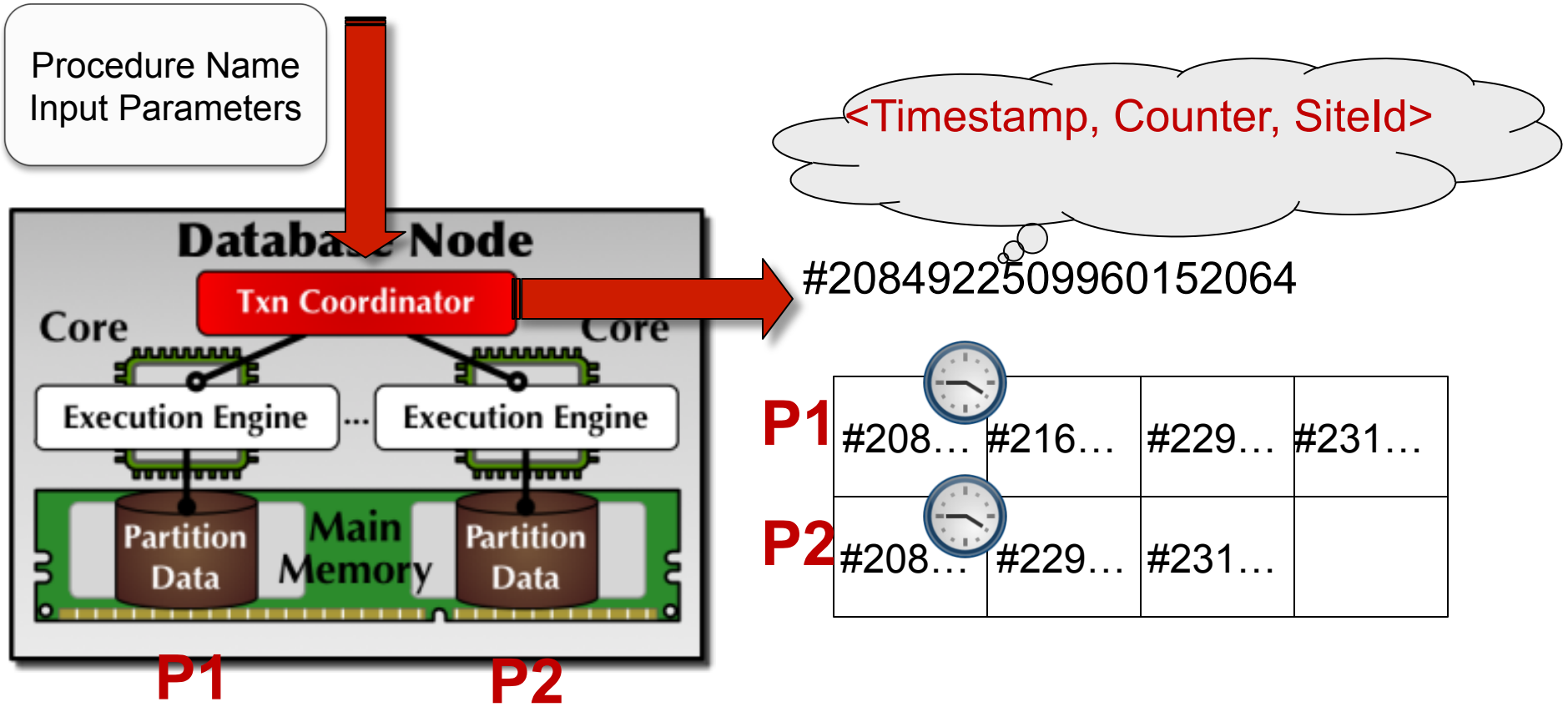
Database Partitioning



Database Partitioning



Distributed Transaction Protocol



Distributed Transaction Protocol

Two-Phase Commit

- TransactionPrepare Request
- TransactionPrepare Response
- [illegible]
- #20849225099601 52064
- [illegible]