THE POWER OF INFINITE POSSIBILITIES

Traditional RDBMS Wisdom is *All* Wrong -- In Three Acts

The “Stonebraker Says” Webinar Series

The first three acts:

1. Why the elephants are toast and why main memory is the answer for OLTP
   - Today

2. Fast Data with ACID - **May 13**

3. Why main memory is still the answer - **June 10**
   - When your data doesn’t fit
   - When you are doing event processing

Our Speaker

Dr. Mike Stonebraker of
MIT, co-founder of VoltDB

Overview of Today

- Trends in the marketplace
  - Some data and fluffy marketing slides
- One size fits none
- Why Traditional RDBMS Wisdom is All Wrong

VoltDB - Background

400
Customers

Mike Stonebraker
Founder, Advisor

Gartner

Magic Quadrant
"Operational Databases"

database
TRENDS AND APPLICATIONS

2013 - One of the Top Companies that Matter Most in Data

Version 4.0 launched in January 2014

Offices in Bedford, MA & Santa Clara, CA

Partners in USA, Europe, Asia, and South America

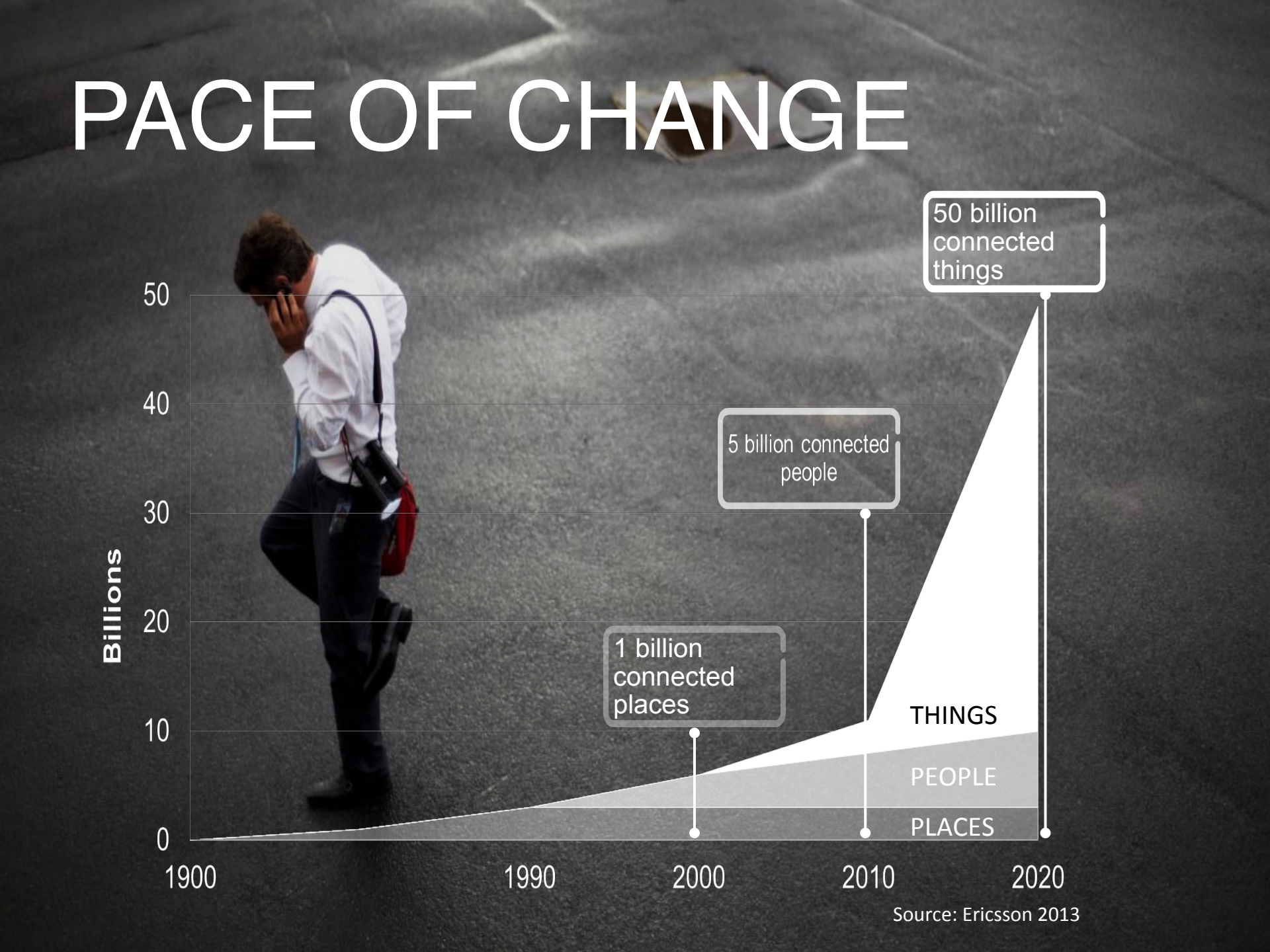
INGRES

streambase

PostgreSQL

VERTICA

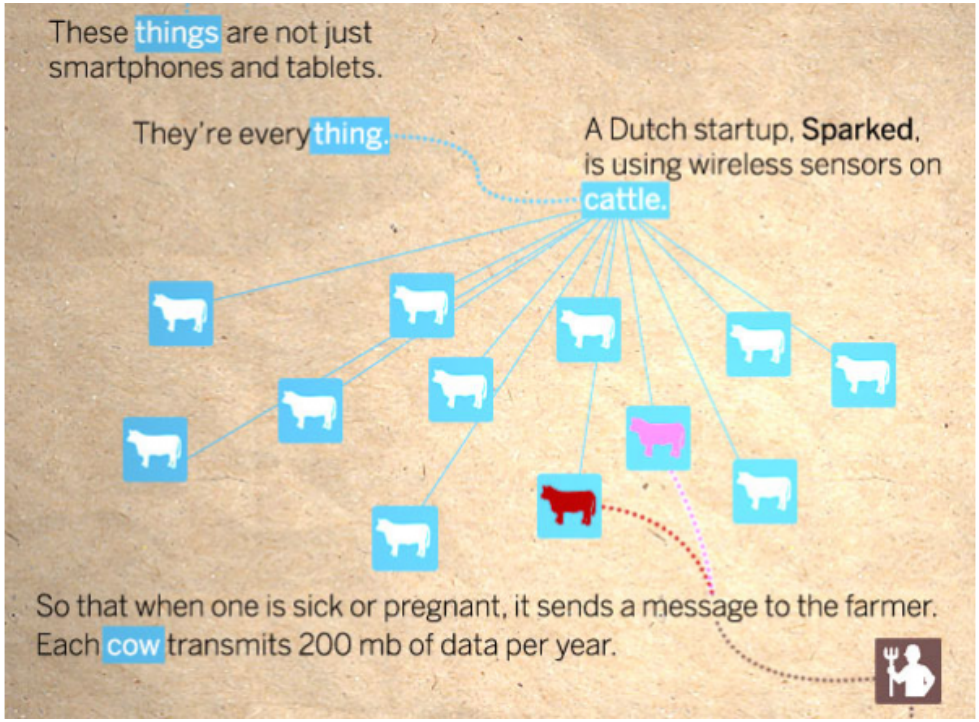
PACE OF CHANGE

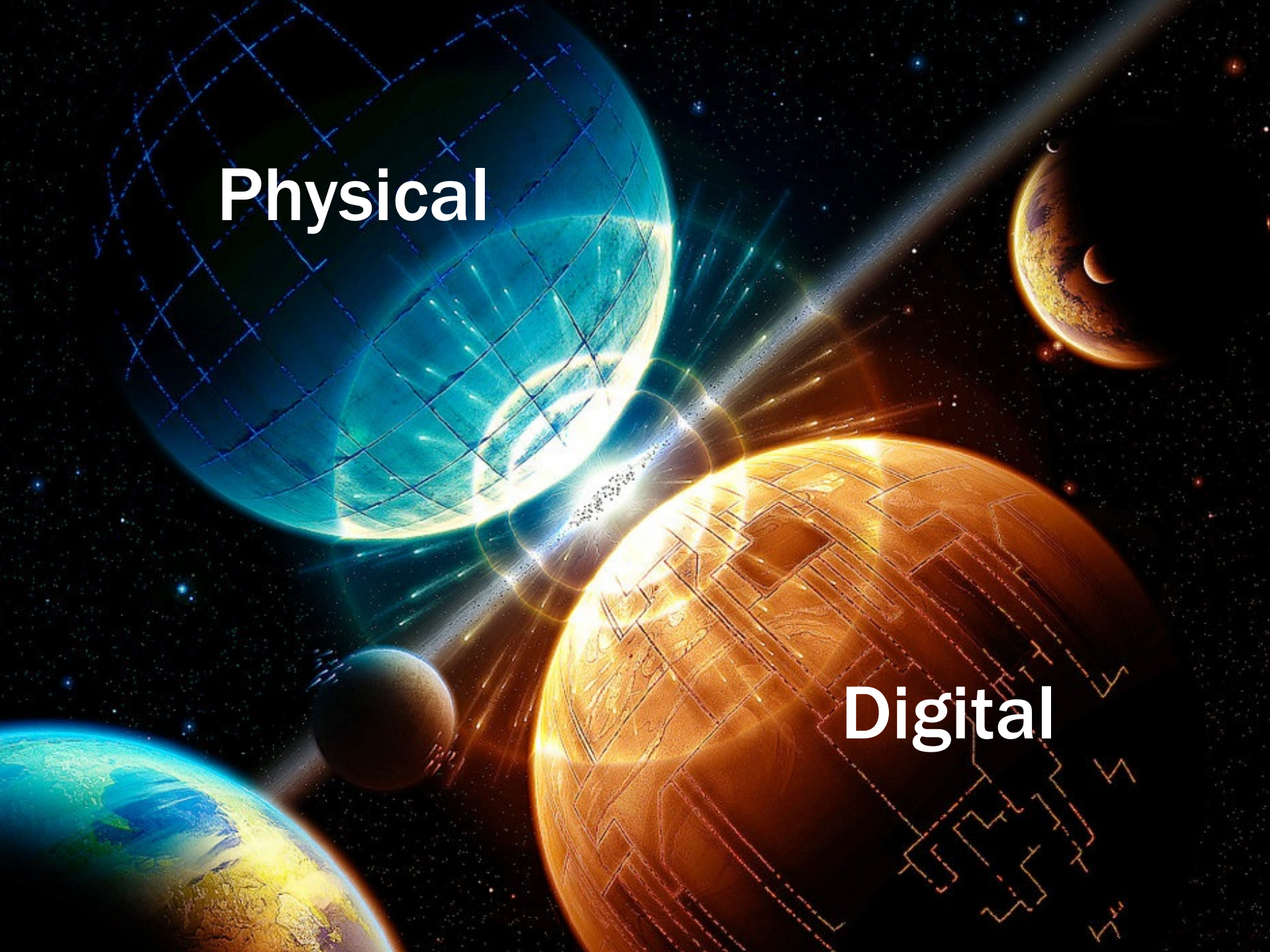


Source: Ericsson 2013

Smart Farm


JOHN DEERE
Field Connect


These things are not just smartphones and tablets.
They're everything.
A Dutch startup, Sparked, is using wireless sensors on cattle.
So that when one is sick or pregnant, it sends a message to the farmer.
Each cow transmits 200 mb of data per year.


Physical
Digital

The Potential: Smart Everything

Smart products
Smart places (cities, farms, buildings…)
Smart networks
Smart services
Smart solutions

- Physical and digital worlds collide
- But nothing happens without making things *Smart*…
- Requirements for *Smart*: Fast, Unlimited Throughput

Trends

- Internet: people + content + things
  - Digital and physical worlds collide
- Data: Big, complex, high velocity
- Technology: distributed, virtual, cloud

THE POWER OF INFINITE POSSIBILITIES

Stonebraker Says

One Size Fits None
“The elephants are toast”

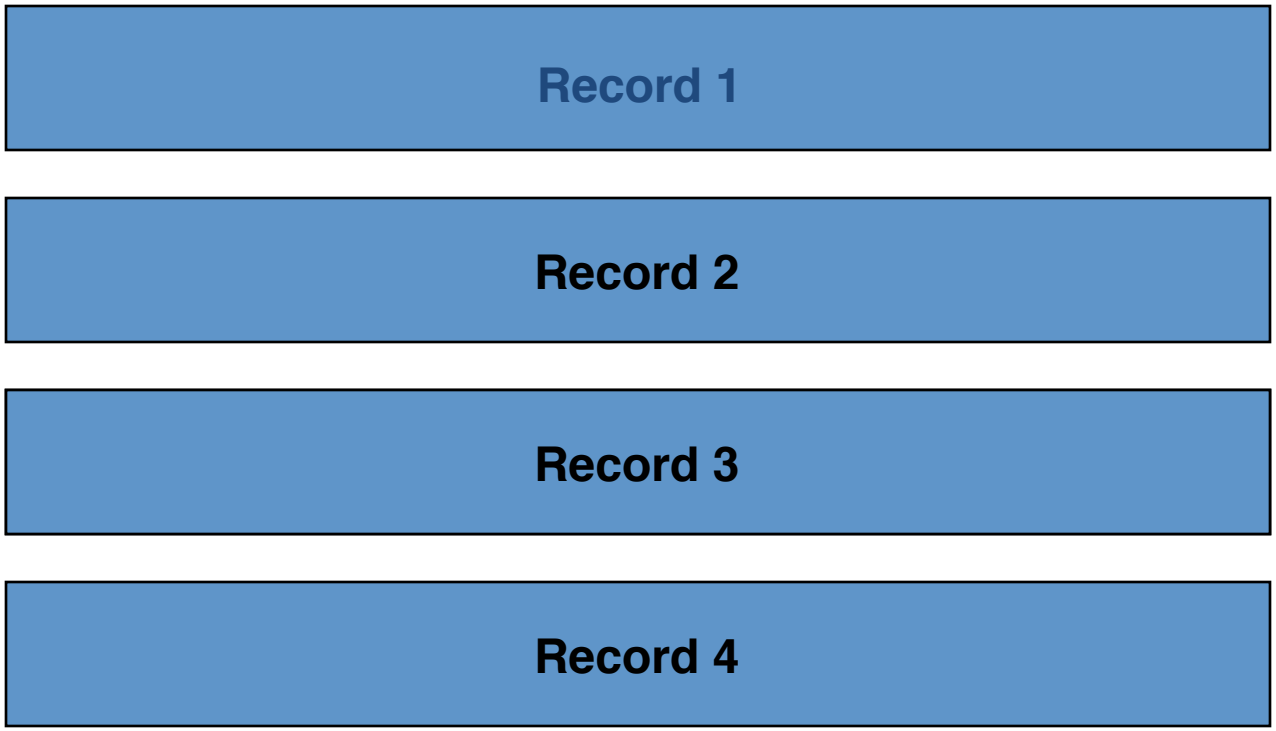
Traditional RDBMS: The Elephants

- Sell code lines that date from the 1970's
  - Legacy code
  - Built for very different hardware configurations
  - And some cannot adapt to grids….
- That was designed for business data processing (OLTP)
  - Only market back then
  - Now warehouses, science, real time, embedded, ..

Current DBMS Gold Standard

- Store fields in one record contiguously on disk
- Use B-tree indexing
- Use small (e.g., 4K) disk blocks; heavily encoded
- Align fields on byte or word boundaries
- Conventional (row-oriented) query optimizer and executor
- Write-ahead log
- Row-level dynamic locking

Terminology -- “Row Store”



E.g. DB2, Oracle, Sybase, SQLServer, Postgres, MySQL, Netezza, Teradata,…

At This Point, RDBMS is “long in the tooth”

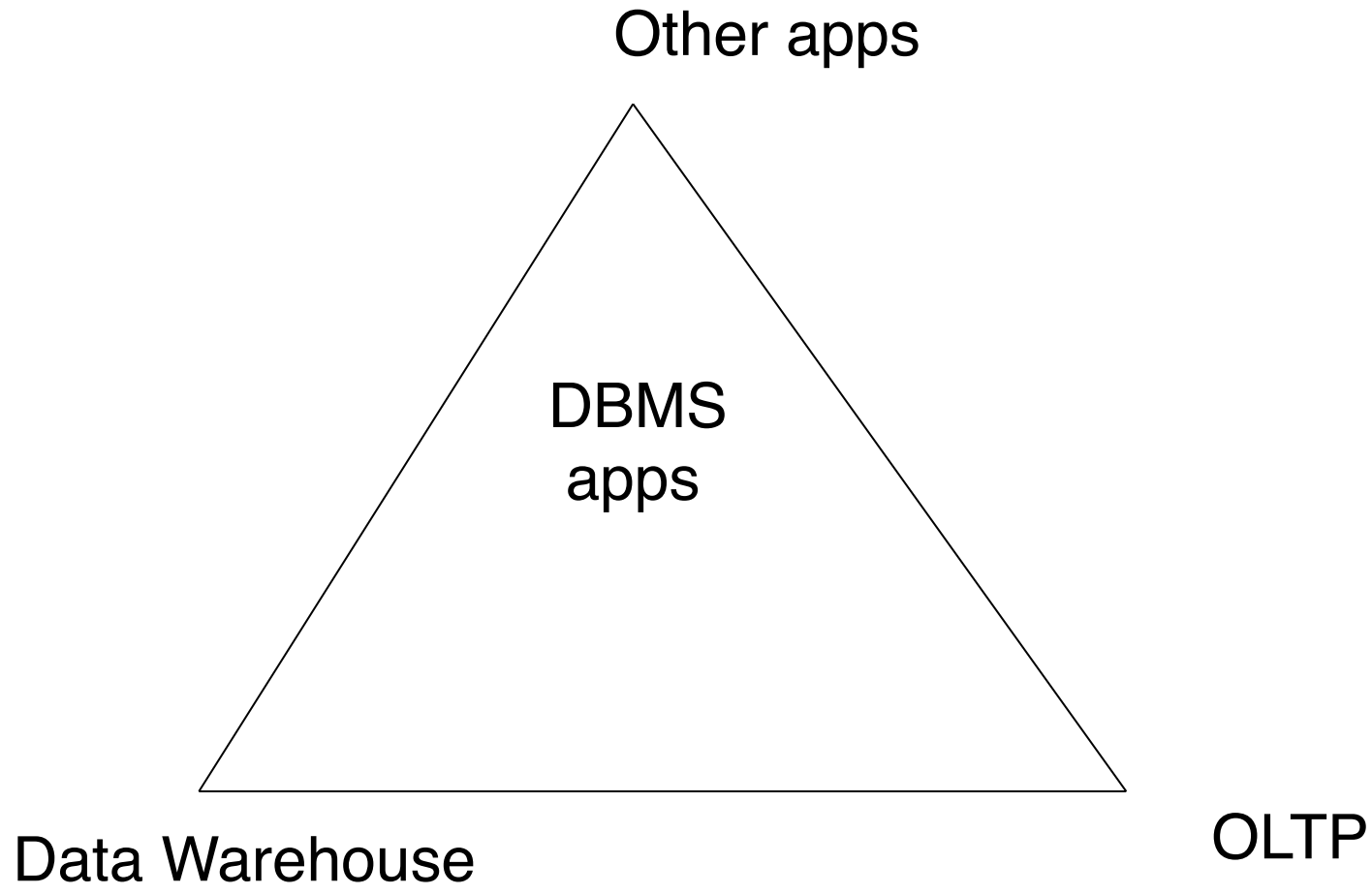
There are at least 6 (non trivial) markets where a row store can be clobbered by a specialized architecture

- Warehouses (Vertica, Red Shift, Sybase IQ)
- OLTP (VoltDB, Hana, Hekaton)
- RDF (Vertica, et. al.)
- Text (Google, Yahoo, …)
- Scientific data (R, MatLab, SciDB)
- Streaming data (coming in part 3)

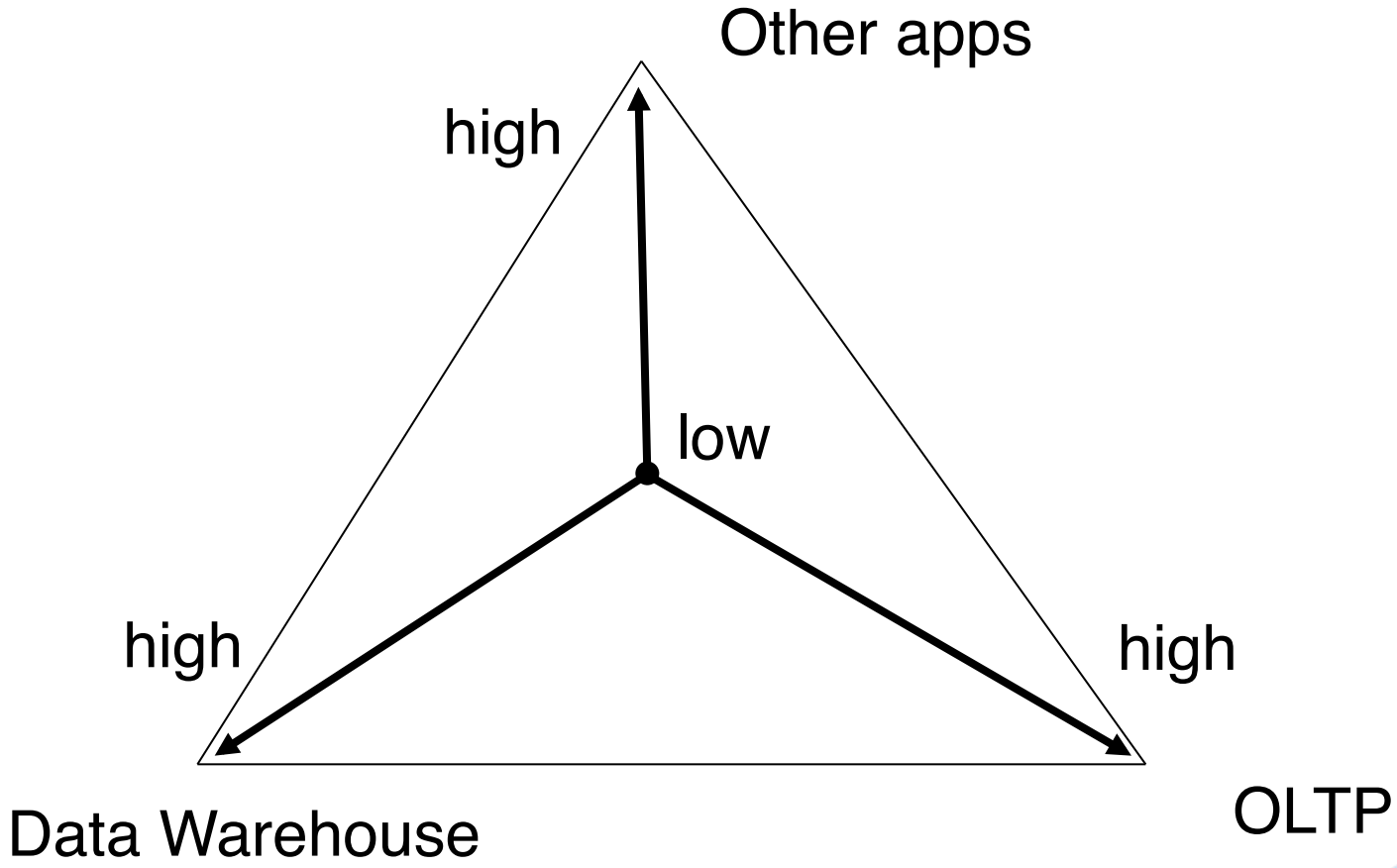
Definition of “Clobbered”

- A factor of 50 in performance

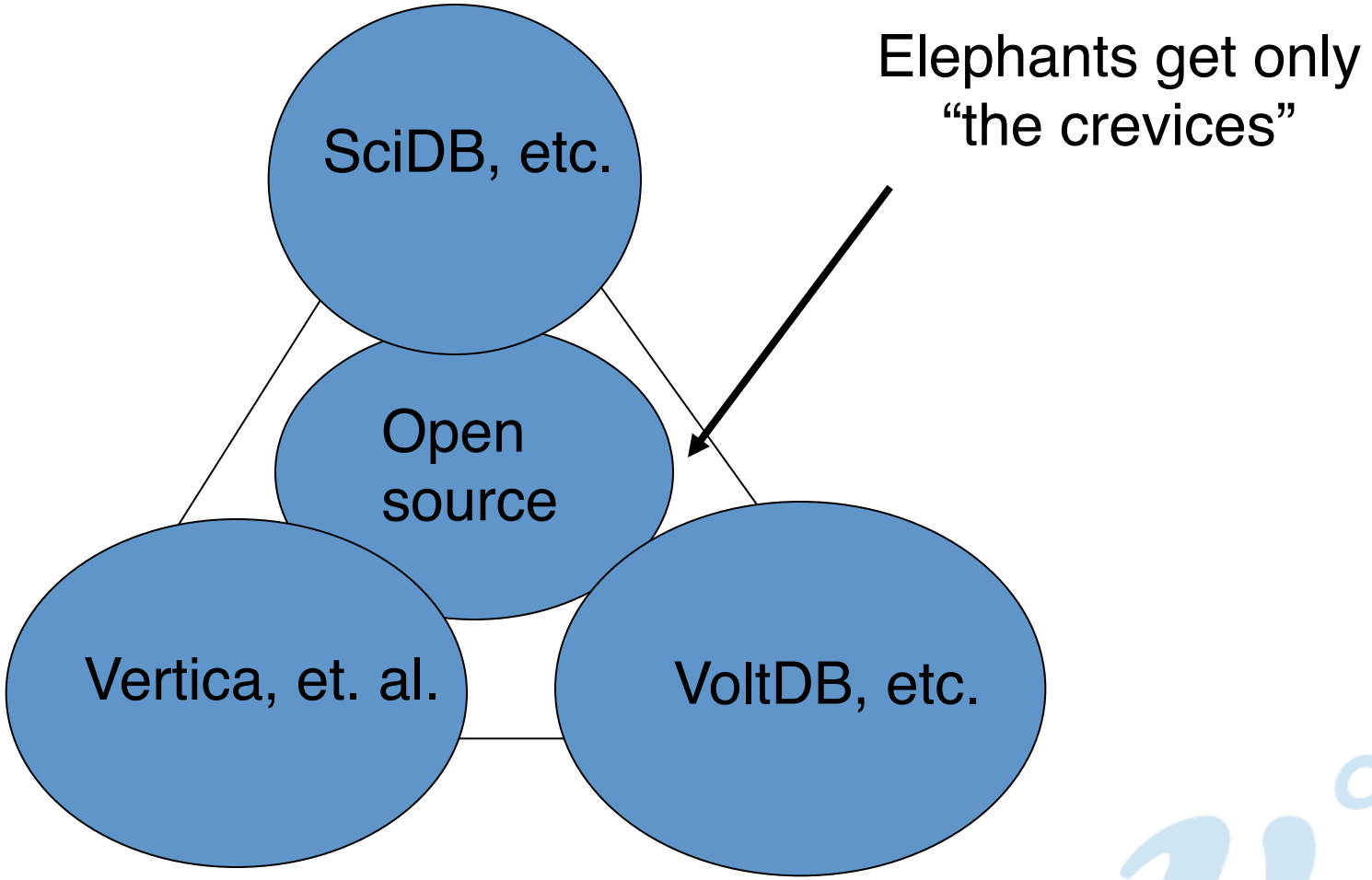
Pictorially:



The DBMS Landscape – Performance Needs



One Size Does Not Fit All -- Pictorially



Stonebraker's Prediction: Coming True

- The DBMS market will move over the next decade
  - To specialized (market-specific) architectures
  - And open source systems

- To the detriment of the elephants

THE POWER OF INFINITE POSSIBILITIES

Stonebraker Says

Traditional RDBMS Wisdom is All Wrong for OLTP

Reality Check #1 on OLTP Data Bases

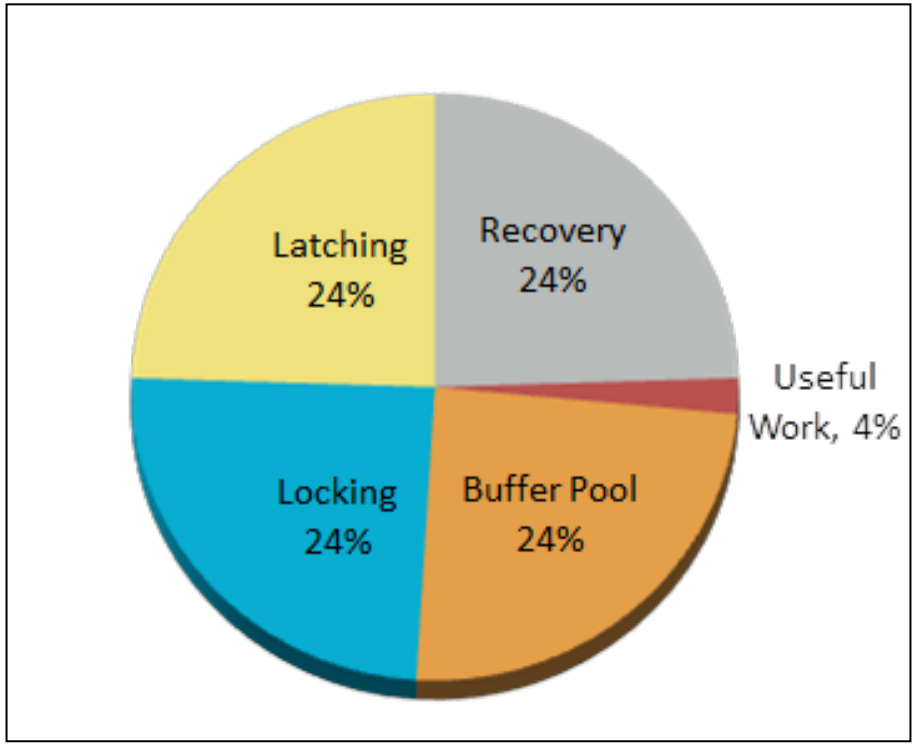
- TP data base size grows at the rate transactions increase
- 1 Tbyte is a really big TP data base
- 1 Tbyte of main memory buyable for around $30K (or less)
  - (say) 64 Gbytes per server in 16 servers
- If your data doesn't fit in main memory now, then wait a couple of years and it will…..
- Facebook is an outlier

Reality Check #2: Client Interactions

- ODBC/JDBC is way too heavy to go fast
- Need a lightweight protocol
  - Asynchronous
  - Connection pooling
  - Stored procedure interface
- Otherwise “the high pole in the tent” is the client interface
  - Hekaton problem!

Reality Check #3: Main Memory Performance

- TPC-C CPU cycles
- On the Shore DBMS prototype
- “Elephants” should be similar (or worse)



To Go Fast….

- Better search data structures
  - E.g., fractal trees, better B-trees, …
- Affect the 4% piece of the pie
  - Irrelevant
  - Except in things like bulk load

To Go Fast….

- Have to get rid of ALL FOUR pieces of the pie
- Consider Oracle's TimesTen
  - Main memory DBMS (one slice gone)
  - Row level locking
  - WAL
  - Multi-threaded
- 3 slices remain!

Implications….

Slices that remain	Relative performance
4	1 (elephants)
3	1.3 (TimesTen)
2	2
1	4
0	25 (Obvious goal of VoltDB)

Removing Slice #1: Buffer Pool

- Main memory DBMS
  - Suppose my data is too big?
  - Stay tuned for part 3
- Any disk-based system will lose!
  - All the NoSQL guys
  - All the elephants

Removing Slice #2: Latches

- A DBMS has a LOT of shared data structures
  - B-trees
  - Lock table
  - Head of the WAL
  - Sorting buffers
  - Malloc-able memory
  - Admission control data structures
  - …

Removing Slice #2: Latches

- In a multi-core world, you need latches to avoid corruption
- In a many core world, this will be a worse disaster
  - All the traditional concurrency control systems degrade after (at most 32-64) cores

Solutions

- Single-threaded code
  - VoltDB
- Latch-free data structures
  - Lots of research in this area!
  - So far, focused mainly on indexing and concurrency control
    - What about all the other shared data structures?

So Far….

- Elephants must convert to
  - Main memory DBMS
  - Single-threaded or latch-free or …
- A complete rewrite
  - See Clayton Christianson “The Innovator’s Dilemma”
- But it gets worse…

Next Time – Fast ACID

- Recovery strategy
  - Aries WAL is dead; long live transaction logging
- Replication strategy
  - Active-active is the answer
  - In contrast to the elephant's active-passive
- Concurrency control strategy
  - Determinism wins; nobody uses row level locking
- More headaches for the elephants!

Questions?

- Use the chat window to type in your questions.
- Try VoltDB yourself:
  - Free trial of the Enterprise Edition:
    - www.voltdb.com/download
  - Open source version is available on github.com
- Register for the May 13th webinar

THE POWER OF INFINITE POSSIBILITIES

THANK YOU!