Making Big Data Roar

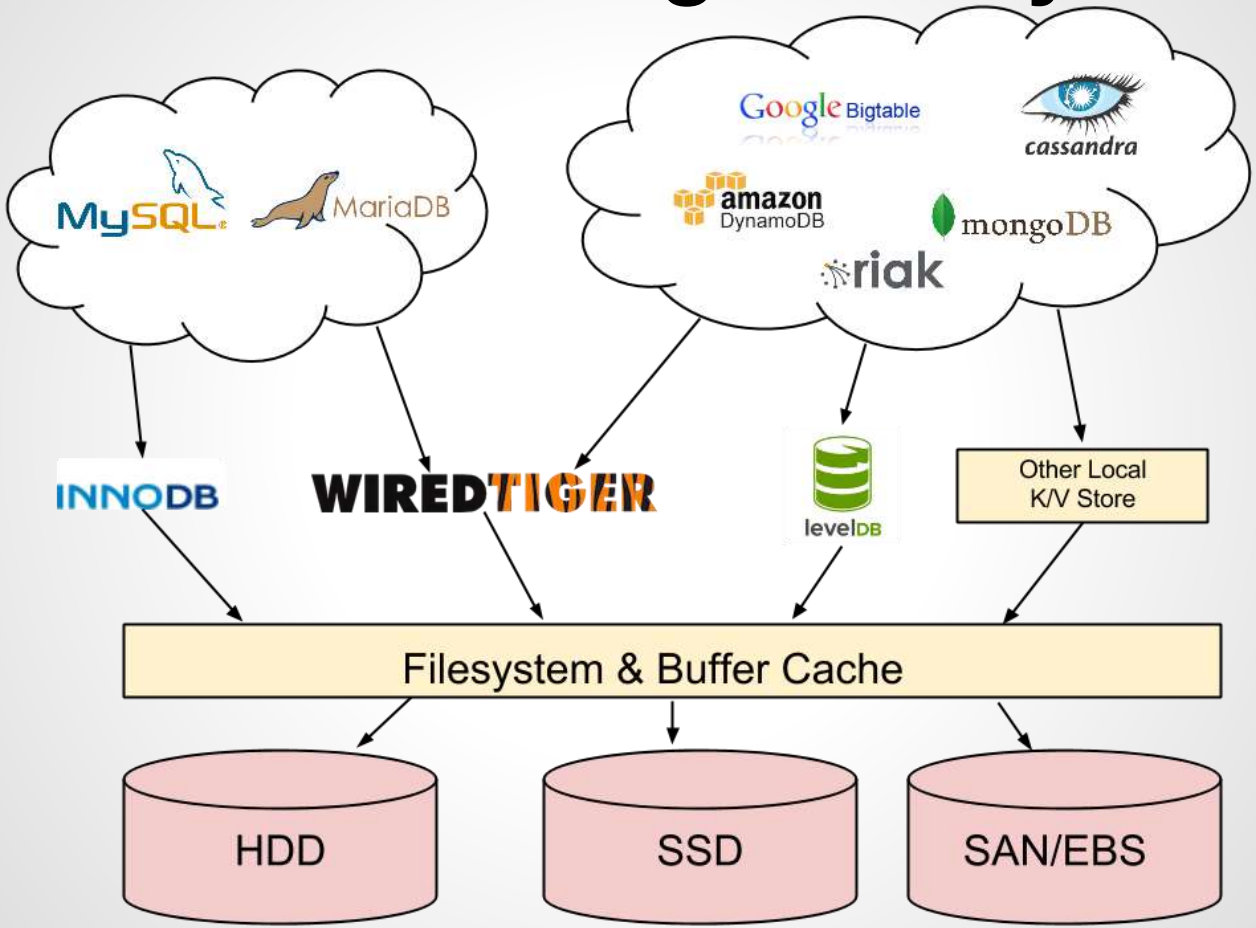
Data Centers are expensive

Company	Location	Data Center Cost	Data Center Size MW
NSA	Camp Williams, UT	$2B	133
Apple	Maiden, NC	$1B	67
Internet Villages	Annandale, Scot.	$1.6B	107
Lockerbie DC	Lockerbie, Scotland	$1.5B	100
Social Security	Baltimore, MD	$400M	27
Next Generation Data	Wales, UK	$300M	20
Facebook	Princeville, OR	$215M	15

WiredTiger Mission

WiredTiger is rethinking data management for modern hardware with a focus on multi-core scalability and maximizing the value of every byte of RAM.

Database/Storage Ecosystem



A New Data Management Engine

- Architected for modern computer systems
- Scalable and able to handle big data
- High throughput, consistent low latency
- Row-store, column-store, log structured merge
- ACID transactions, standard isolation levels
- Checkpoint and fine-grained durability
- Supporting columns, indices, projections
- Production quality, fully supported
- NoSQL, Open Source

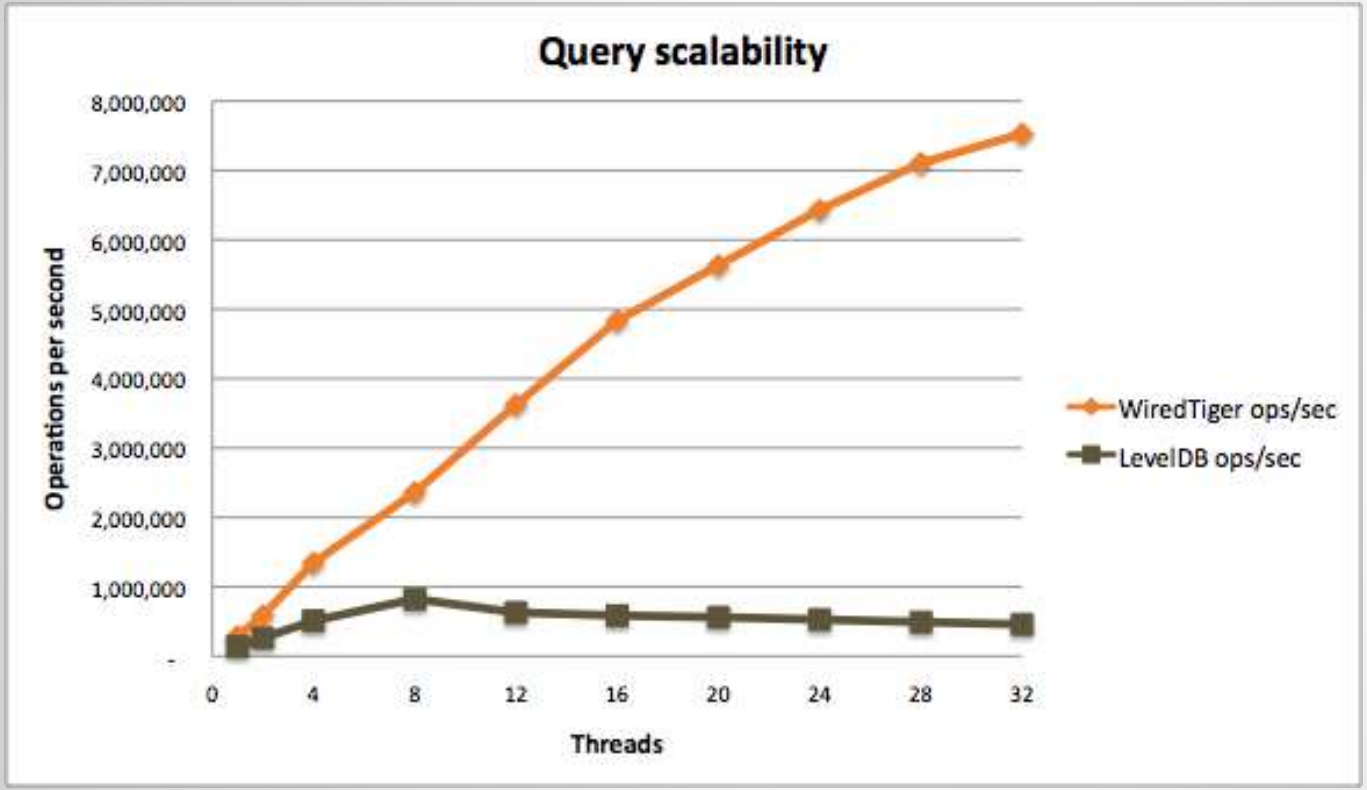
Flexible Storage

- Access methods tailored to workload
  - Row store (read mostly of all columns)
  - Column store (read mostly of some columns)
  - Log-structured merge trees (mostly random writes)
- Compact storage format
  - RLE, key-prefix, dictionary and static compression
  - Stream compression
- Adapt workload to storage (RAM, SSD, HDD)

Flexible Configuration

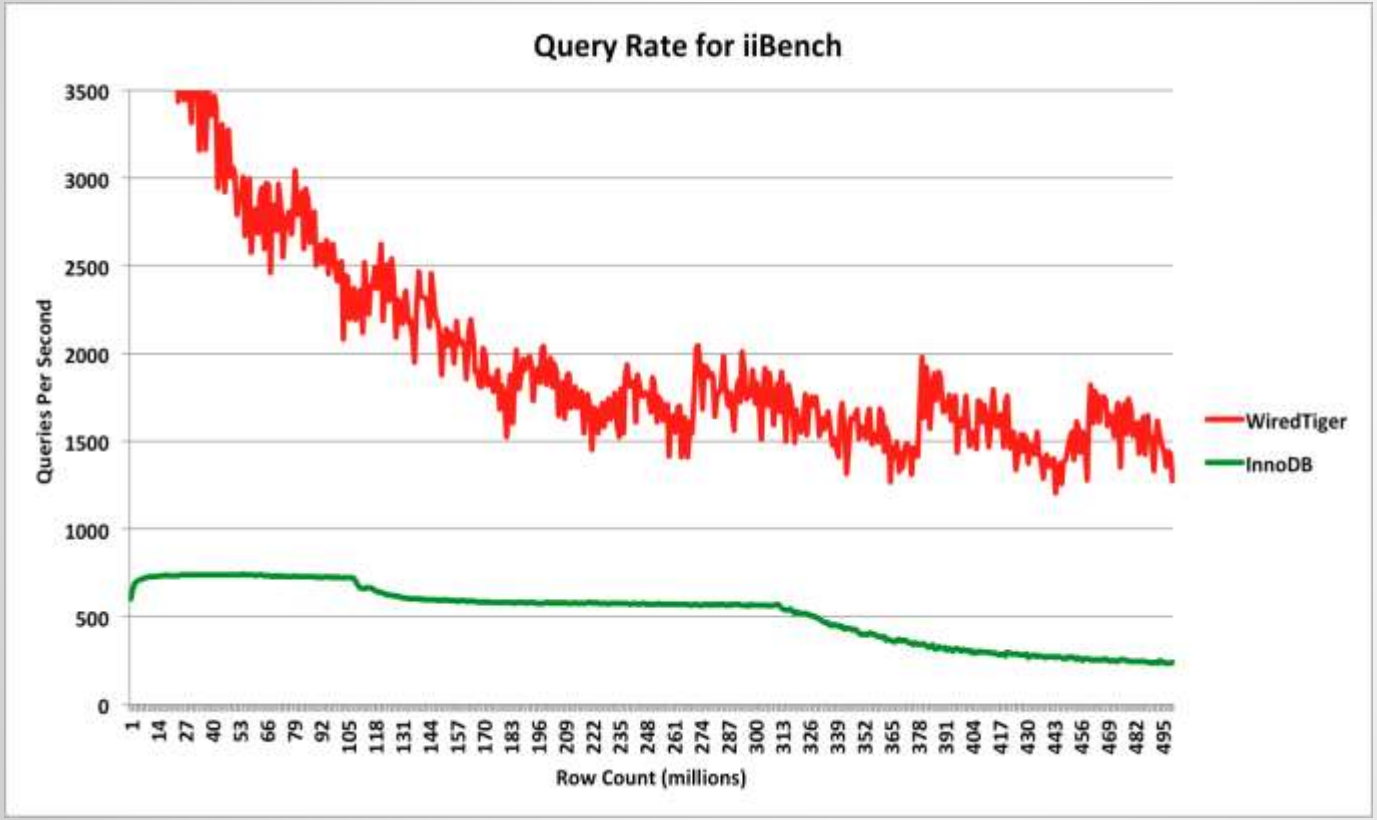
- API offers a simple key/value store, or
- A complete schema layer
  - Specify data types
  - Map columns to files
  - Automatically maintain indices
  - Queries only read required columns
  - Projections, index-only scans
- Checkpoint or fine-grained durability

Improved Efficiency



- Higher CPU Utilization
  - Multi-core scalability
  - Minimize contention between threads
  - Non-locking algorithms
  - Hazard pointers
- Lower Power Costs
- Flash Optimized Block Layout

Consistent High Performance



- In-cache or I/O bound
- Workload Configuration
  - Efficient sparse data (column-store)
  - Bounded queries and updates (row-store)
  - Write-optimized (LSM)
- Data structures for access at RAM speed

Consistent Low Latency

- Non-locking algorithms
- Multi-versioned data
- Optimistic concurrency control
- Deadlock-free transactions
- I/O shifted to background threads

Cost Effective

Metric	WIREDTIGER	INNODB
iiBench run cost	$6.44	$12.88
Cost per Billion inserts*	$20.30	$40.60

- WiredTiger provides a 50% cost savings for the same AWS workload
- More details on this benchmark are available here.

Customers


amazon
web services™


sTec
HGST
a Western Digital company


facebook

Management Team

Keith Bostic is a founder and architect at WiredTiger. He was a founder of Sleepycat Software, (acquired by Oracle Corp. in 2006), and one of the architects of the Berkeley DB, the most widely-used embedded data management software in the world.

Mr. Bostic was one of architects of the University of California, Berkeley, 2.10BSD and 4BSD releases, where he lead the 4BSD release Open Source effort. He is the recipient of a USENIX Association Lifetime Achievement Award (The Flame), which recognizes singular contributions to the UNIX community.

Dr. Michael Cahill is a founder and architect at WiredTiger. He was an architect of Berkeley DB at Sleepycat Software and Oracle Corp., responsible for design and implementation of multiversion concurrency control, as well as SQL interfaces and programming language APIs. Previously, Dr. Cahill was CTO at Bullant Technology, which grew tenfold and raised over US$30 million from investors including Intel Capital and JP Morgan during his three year tenure.

Dr. Cahill's PhD from the University of Sydney is in the area of transaction processing and concurrency control. His work on a new algorithm for implementing serializable isolation received an ACM SIGMOD Best Paper award and was added to PostgreSQL 9.1.

Summary and Next Steps

We’d like to discuss how we could help you with your solution.

Thanks! Questions? info@wiredtiger.com