IBM Analytics Platform Group

The Science and Engineering Behind Super Fast Load-and-Go In-memory Analytics

Sam Lightstone

Distinguished Engineer

BLU Acceleration

1. Next generation in-memory SQL data warehousing
   - Super Fast (query performance)
   - Super Simple (load-and-go)
   - Super Small (RAM and storage savings)

2. Seamlessly integrated
   - Built seamlessly into **both** dashDB (cloud) and DB2 (software)
   - Consistent SQL, language interfaces, administration
   - Dramatic simplification

3. Hardware optimized
   - Memory optimized
   - CPU-optimized
   - I/O optimized

SEIZE THIS
MOMENT

35x-73x faster

...than traditional row-organized relational database technology, on average.

273x

"Wow…unbelievable speedup in query run times! We saw a speedup of 273x in our Vehicle Tracking report, taking a query from 10 minutes to 2.2 seconds. That adds value to our business; our end users are going to be ecstatic!"
- Ruel Gonzalez - Information Services

106x

UNIVERSITY OF TORONTO

"Compared to our current production system, DB2 10.5 with BLU Acceleration is running 106x faster for our Admissions and Enrollment workloads. We had one query that we would often cancel if it didn't finish in 30 minutes. Now it runs in 56 seconds every time. 32x faster, predictable response time, no tuning…what more could we ask for?"
- Brenda Boshoff, Sr. DBA

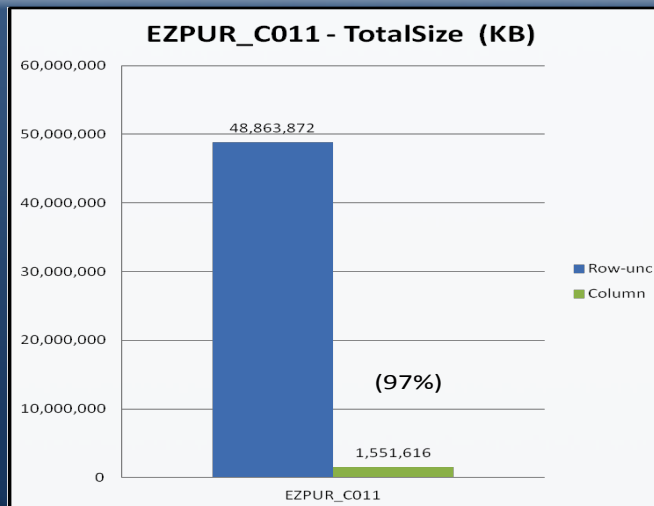SEIZE THIS MOMENT

Storage Savings

15.4x

> *"My largest row-organized, adaptive compressed table gave me 3.2x storage savings. However,* **converting this row-organized uncompressed table to a column-organized table in DB2 10.5 delivered a massive 15.4x savings!"**
>
> - Iqbal Goralwalla, Head of DB2 Managed Services, Triton

31.5x

EZPUR_C011 - TotalSize (KB)

	Value
Row-unc	48,863,872
Column	1,551,616 (97%)

Chart axis: 60,000,000 / 50,000,000 / 40,000,000 / 30,000,000 / 20,000,000 / 10,000,000 / 0

```
Records:   76M

Columns:   61

Indexes:   10

Load Time
  Row-unc 15:39:10
  Col       1:10:29
```

31.5x storage savings

(**97%** less storage required)

13.5x faster load time

Super simple

CREATE. LOAD. GO!

SEIZE THIS
MOMENT

Load-and-go simplicity

- LOAD and then… run queries

Simple.

- No indexes
- No storage reclaim (it's automated)
- No memory configuration (it's automated)
- No process model configuration (it's automated)
- No statistics collection (it's automated)
- No MDC or MQTs
- No Statistical views
- No optimizer profiles/guidelines

Coca-Cola
BOTTLING CO.
CONSOLIDATED

"The BLU Acceleration technology has some obvious benefits: … But it's when I think about **all the things I don't have to do with BLU**, it made me appreciate the technology even more: **no tuning, no partitioning, no indexes, no aggregates**."
-Andrew Juarez, Lead SAP Basis and DBA
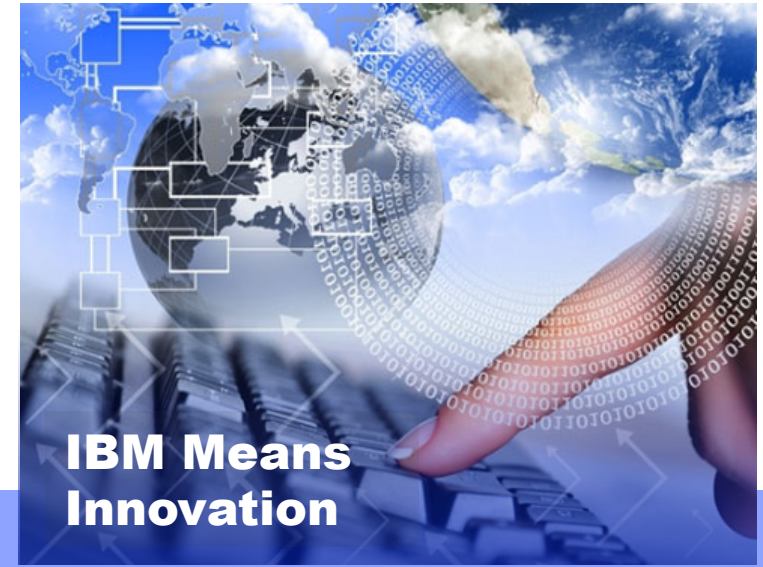
8

In-memory speeds with tiny RAM requirements

- Rule of thumb: RAM requirements are just 5% of source data size.

- Example:
  - 10 Terabytes of raw user data
  - 500 GB of RAM

SEIZE *THIS*
MOMENT

Magic: How did they do that?

# BLU Acceleration Easiest. Fastest. Smallest.

IBM Means Innovation

BLU Acceleration includes over **30 new patents and patents pending** from IBM Research & Development Laboratories.

SEIZE THIS MOMENT

Looking at Disk Improvements

- **Incredibly inexpensive drives (& processors) have made it possible to collect, store, and analyze huge quantities of data**
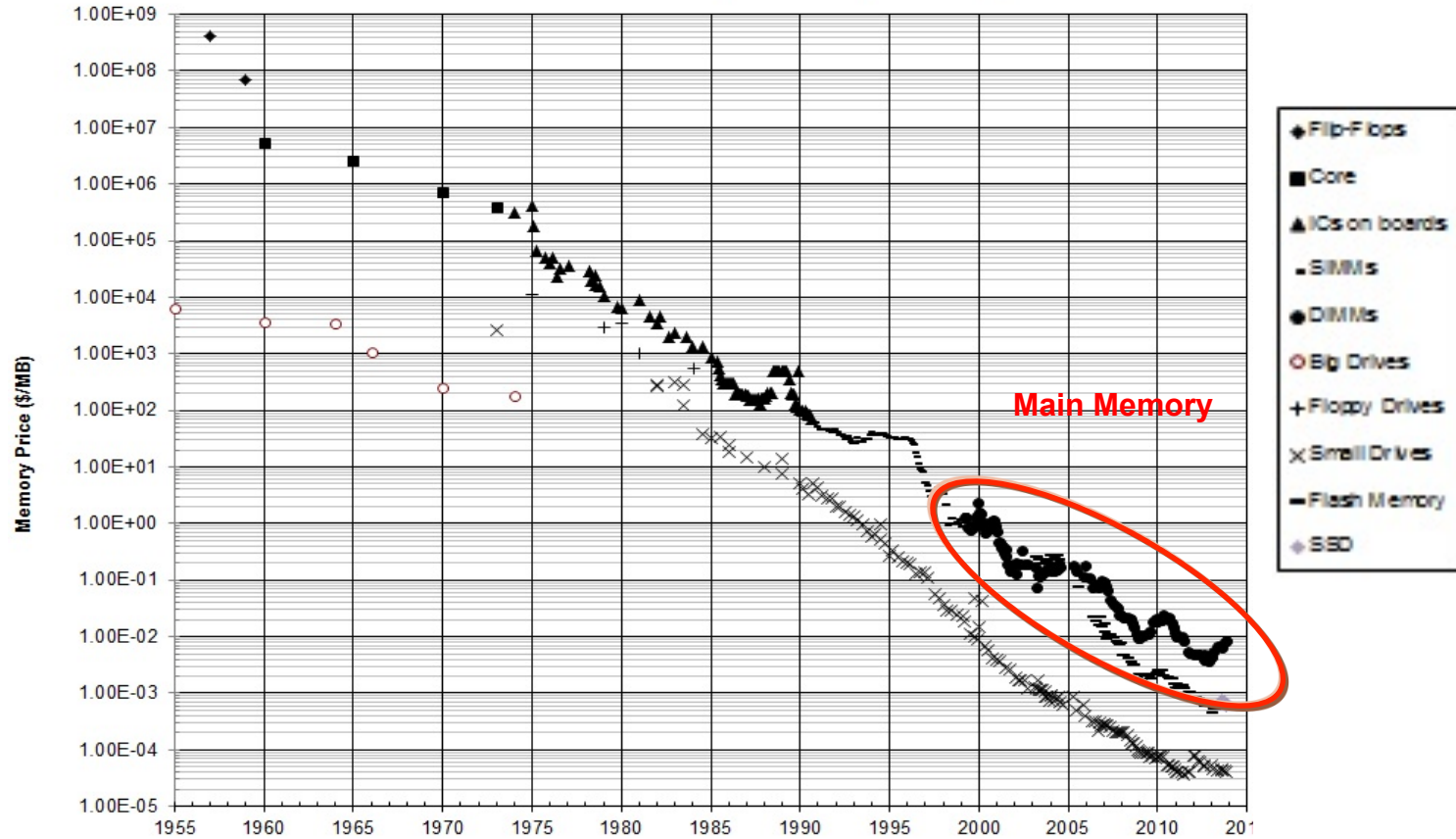
Over the last 30 years

Drives

Transfer Rates:
1.2MB/sec → 80MB/sec
65x

Capacity:
80MB → 800GB
10,000x
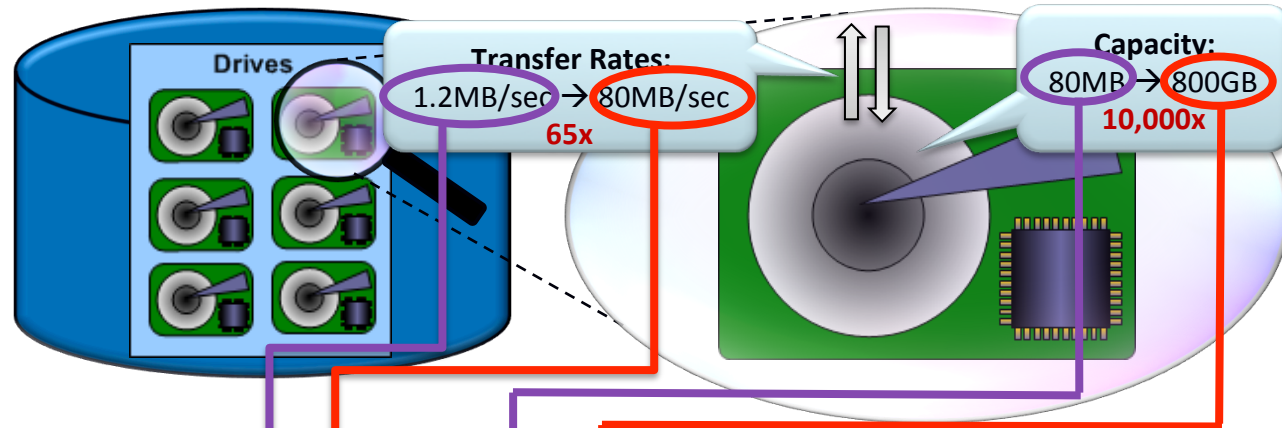
But, consider the metric, **transfer bandwidth/byte**

- **1980: 1.2 MB/sec / 80 MB = 0.015**
- **2009:   80 MB/sec / 800,000 MB =.0001**

When relative capacities are factored in, drives are 150X slower today!!!

We've been snookered!

## Historical Cost of Computer Memory and Storage

Main Memory

Legend:
- Flip-Flops
- Core
- ICs on boards
- SIMMs
- DIMMs
- Big Drives
- Floppy Drives
- Small Drives
- Flash Memory
- SSD

Source: http://www.jcmit.com/mem2013.htm

Big Idea: RAM is Too Slow

CPU cache optimized

- RAM is at the bottom of the memory hierarchy. It is the slowest non-persistent memory in a server.

- CPU cache is many times faster than RAM.

- Extreme re-engineering of database algorithms to be CPU cache optimized

- BLU algorithms adapt automatically to hardware cache size.

Speed-up versus RAM

SEIZE THIS
MOMENT
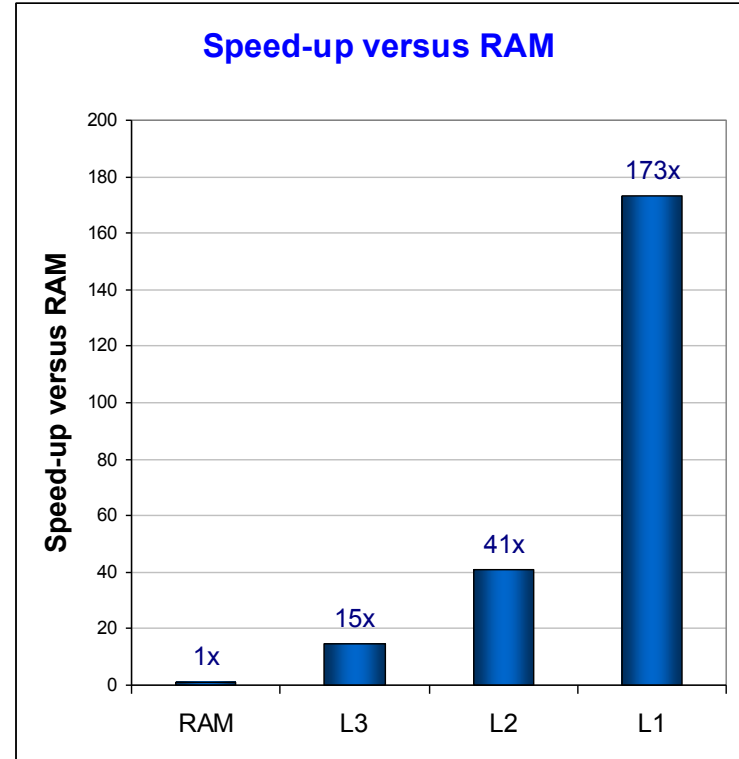
More Evil Than Ever

1. Human intervention

2. Voluminous I/O

3. Random I/O

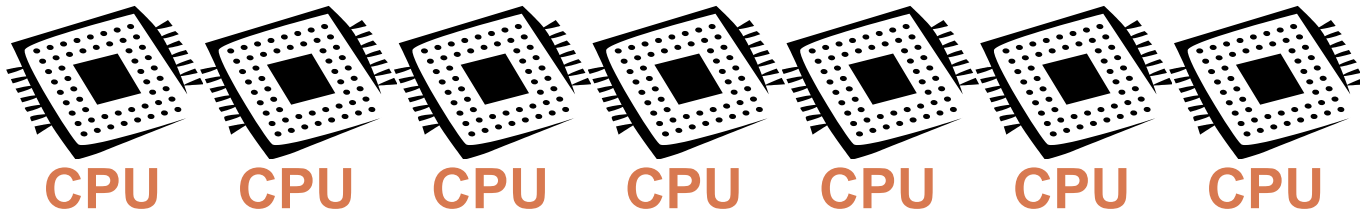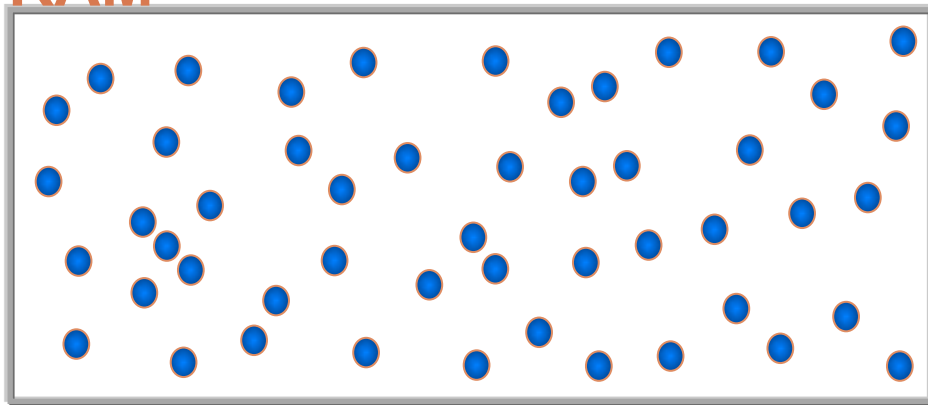4. Memory stalls

5. Single core processing

6. The cost of RAM

Load-and–go simplicity: System resources

1. Auto-detect and adapt to available RAM

2. Auto detect and adapt to core number and type

3. Auto detect and adapt to CPU cache size
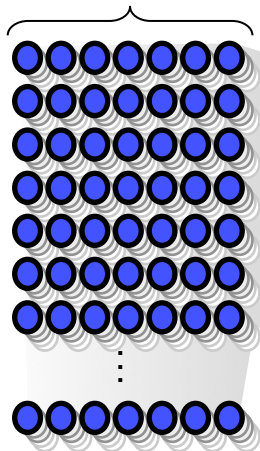
RAM

CPU CPU CPU CPU CPU CPU CPU

Load-and-go simplicity: Automatic Workload Management

- Built-in and automated query resource consumption control
- Many queries can be submitted, but effective concurrency, per query RAM, and CPU are automatically and dynamically controlled
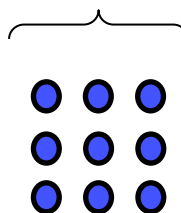
Applications and Users

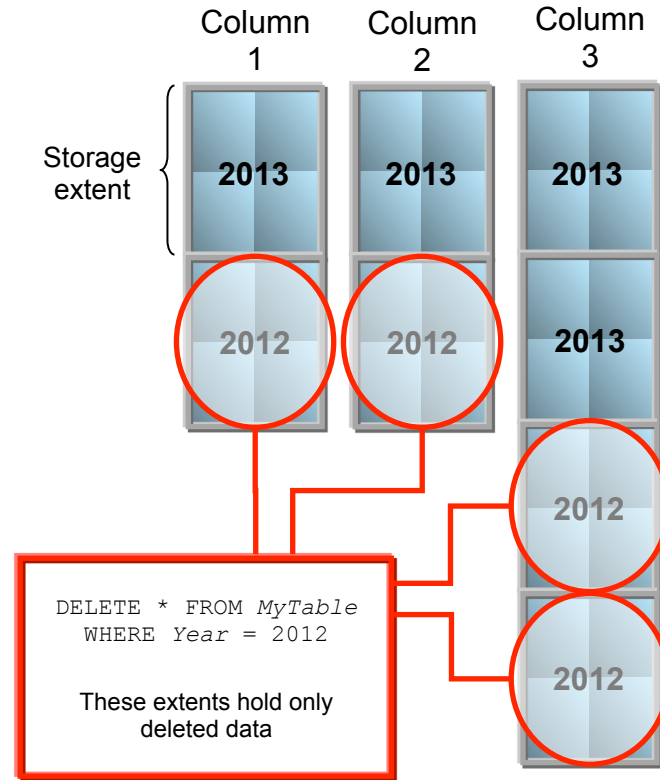Up to tens of thousands of SQL queries at once

BLU Runtime

Moderate number of queries consume resources

SQL Queries

SEIZE THIS
MOMENT

Load-and-go simplicity: Automatic Space Reclaim

- Automatic space reclamation
  - Frees extents with no active values
  - The storage can be subsequently reused by any table in the table space

- No need for costly DBA space management and `REORG` utility

- Space is freed online while work continues

- Regular space management can result in increased performance of `RUNSTATS` and some queries

Column 1   Column 2   Column 3

Storage extent

2013	2013	2013
2012	2012	2013
		2012
		2012

```
DELETE * FROM MyTable
  WHERE Year = 2012
```

These extents hold only deleted data

Super Compress

- **10-20X smaller is common**

- **Compress as small as 1 bit**

- **Compress the most frequent data the smallest**
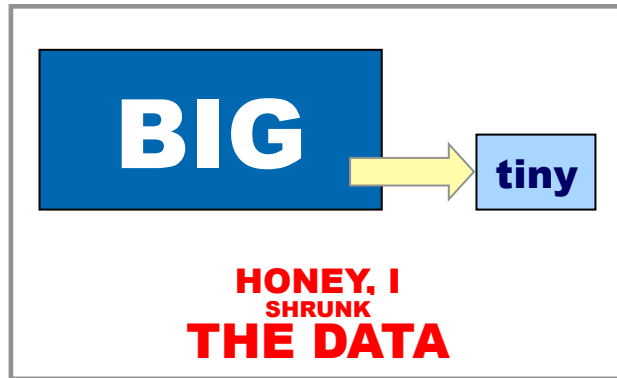
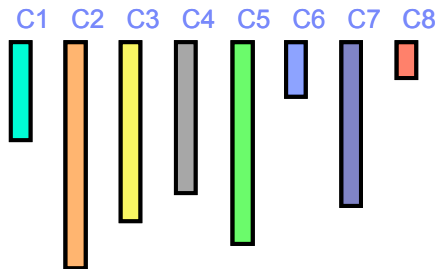# Query on Compressed Data

- **No CPU for decompression**

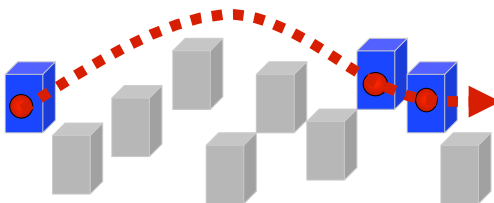- **Data flows through memory and CPU at compressed size**

- **"Actionable Compression"**

Columnar Everywhere

- **Reduce I/O**

- **Increase data density in RAM**
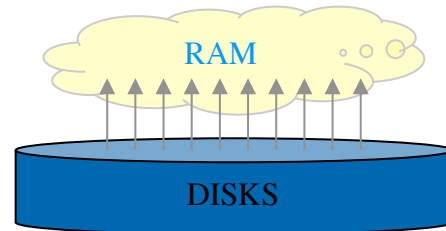
- **Increase CPU efficiency**

Skip Boring Data

- **Queries skip uninteresting data**

- **Synopses on every column, automatically.**

- **"Data Skipping"**

Rethink Memory

- **Cache intelligently for analytics**

- **Predictive I/O with "Dynamic List Prefetching"**

- **Massive I/O reduction**

Multiply the Power of the CPU

- **Use modern SIMD instructions to do multiple operations with a single instruction.**

- **Use 1 instruction instead of 8.**

- **Massive CPU acceleration**

- **Store data as vectors**

- **"Parallel Vector Processing"**
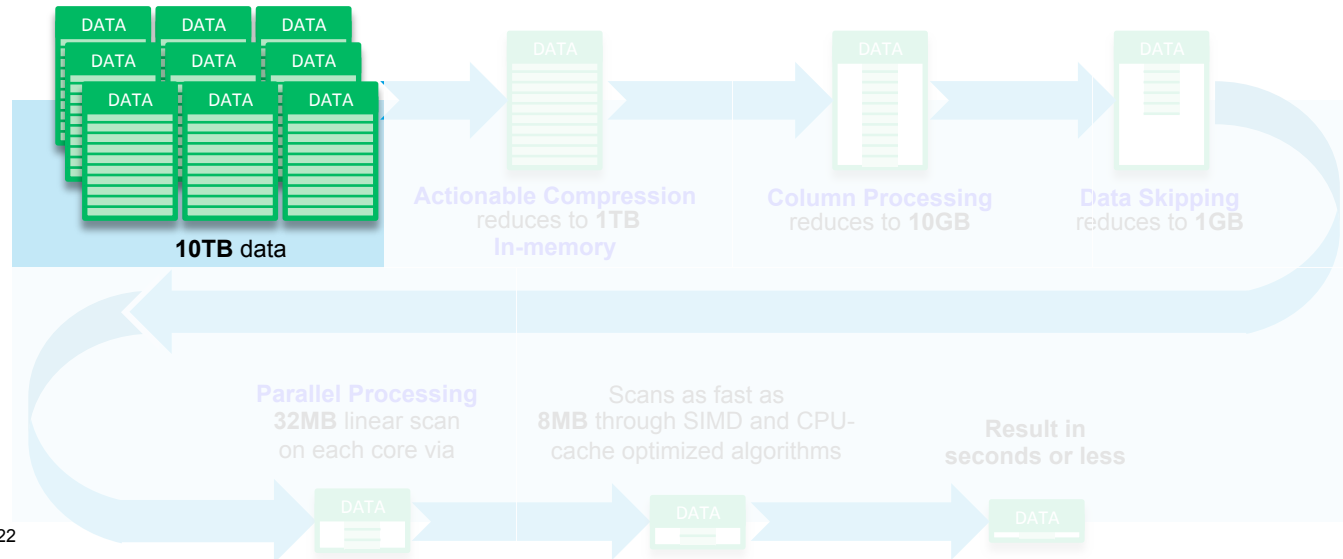
Faster than RAM!

- **RAM is too slow for BLU !**

- **Redesign the query engine to operate at CPU cache speeds instead**

- **CPU cache is 10-75X faster than RAM access.**

BLU Acceleration Illustration
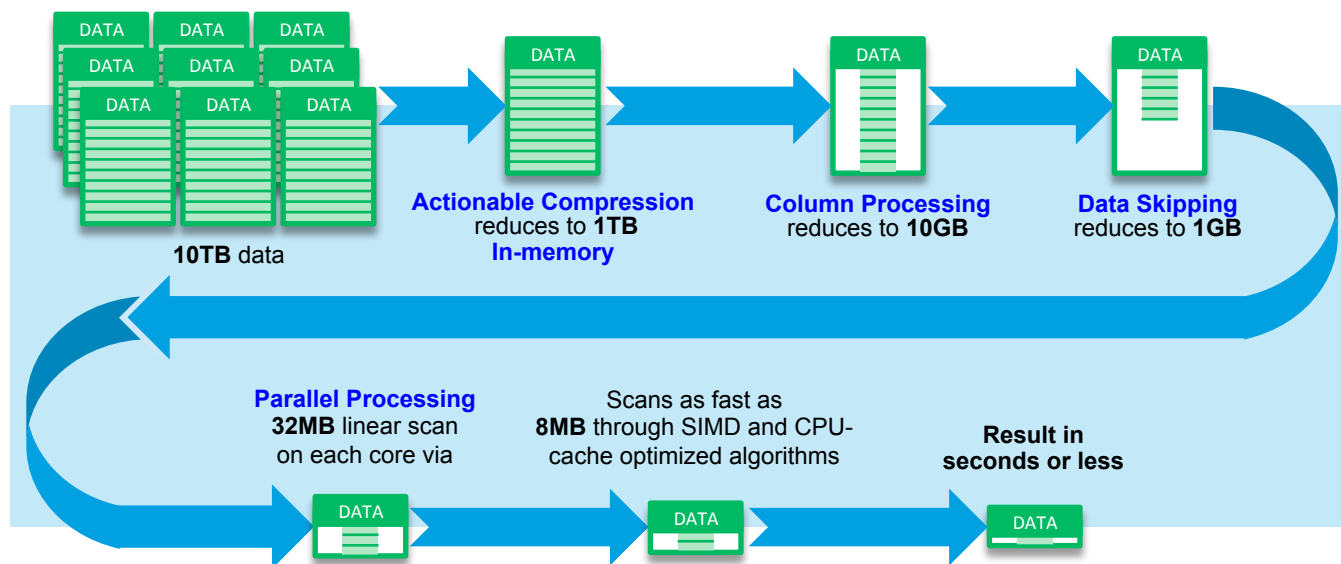## 10TB query in seconds or less

- The System: 32 cores, 1TB memory, 10TB table with 100 columns and 10 years of data

- The Query: How many "sales" did we have in 2010?
  - `SELECT COUNT(*) from MYTABLE where YEAR = '2010'`
- The Result: In seconds or less as each CPU core examines the equivalent of just 8MB of data

10TB data

Actionable Compression
reduces to **1TB**
In-memory

Column Processing
reduces to **10GB**

Data Skipping
reduces to **1GB**

Parallel Processing
32MB linear scan
on each core via

Scans as fast as
8MB through SIMD and CPU-
cache optimized algorithms

**Result in
seconds or less**

BLU Acceleration Illustration
10TB query in seconds or less

- 1GB RAM required to cache all data for query over 10TB of data.
- **10,000 times reduction** in memory requirements to achieve in-memory speeds

10TB data

Actionable Compression
reduces to **1TB**
In-memory

Column Processing
reduces to **10GB**

Data Skipping
reduces to **1GB**

Parallel Processing
32MB linear scan
on each core via

Scans as fast as
8MB through SIMD and CPU-cache optimized algorithms

**Result in
seconds or less**

IBM dashDB – Data Warehousing as a Service

Terabytes of data ready to analyze
within minutes

- **Fast querying**
  - In-memory
  - Columnar
  - SIMD hardware acceleration
  - Actionable compression

- **Advanced analytics**
  - Support for OLAP SQL extensions
  - In-database analytics & R for predictive modeling
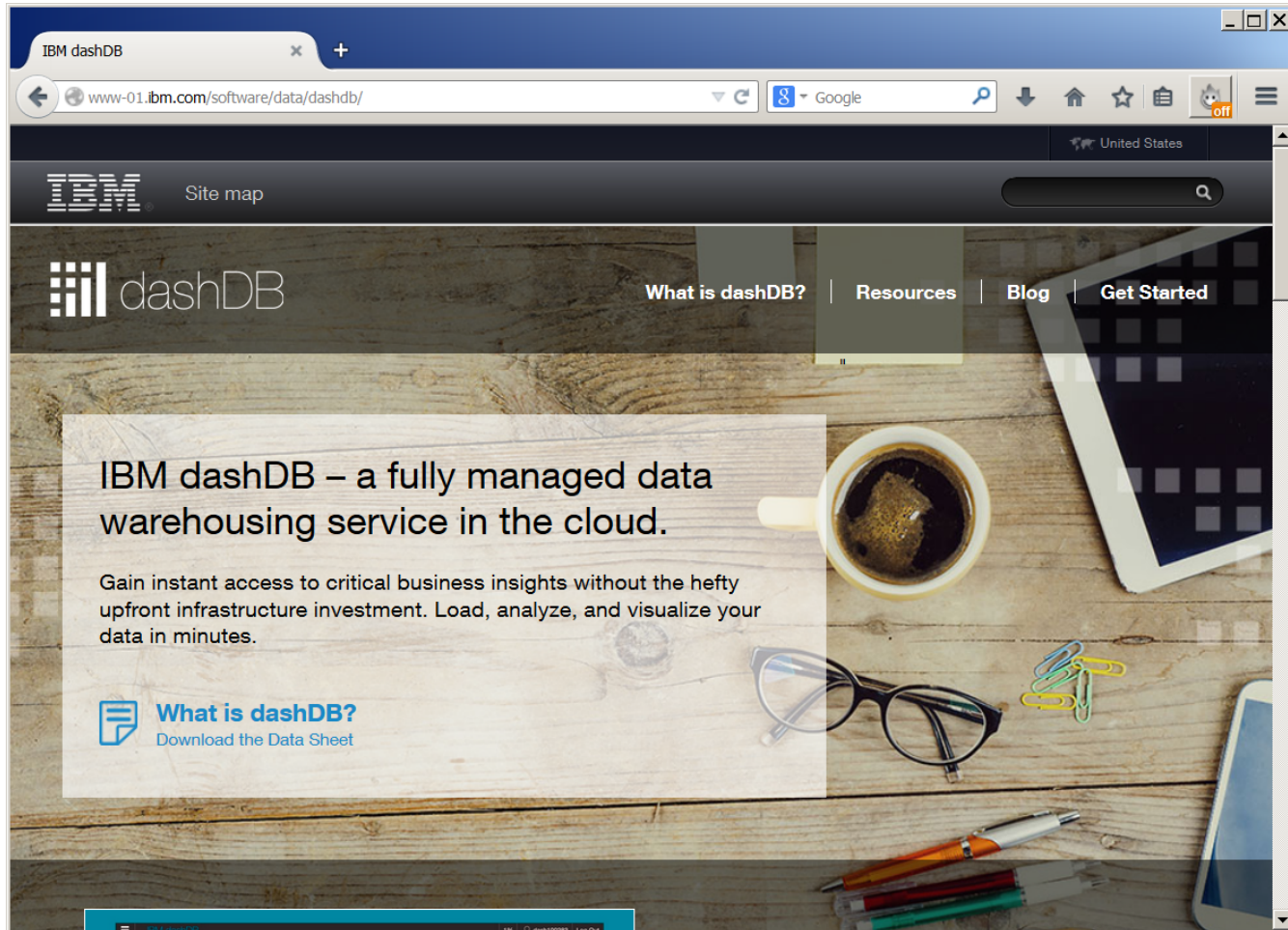 - Spatial analytics

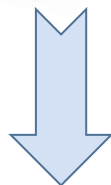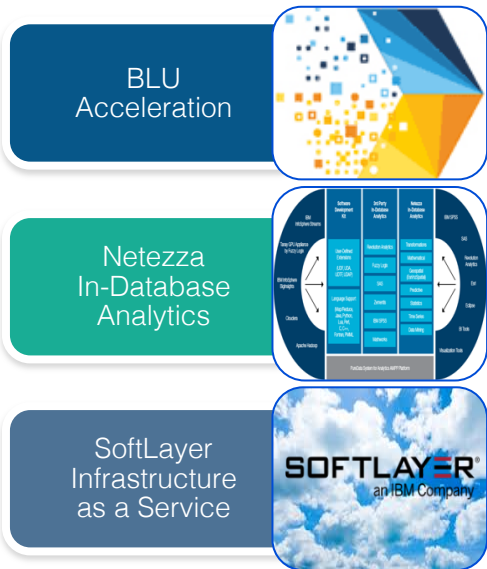- **Easy to integrate**
  - Connect common 3rd party BI tools

dashDB

*keeps data warehouse
infrastructure out of your way*

The dashDB Cocktail: Three Parts IBM, Shaken…

BLU
Acceleration

Netezza
In-Database
Analytics

SoftLayer
Infrastructure
as a Service

dashDB

*Focus on the business,
not the business of data warehousing!*

Build More

Grow More

Know More

Partnership with IBM Cloudant and dashDB

- Cloudant is a fully managed distributed NoSQL Database as a Service (DBaaS)

    - Multi-tenant and single (dedicated) tenants
    - Managed 24x7 by Cloudant expert engineers
    - Service Level Agreement
    - Operational data store

- dashDB's partnership with Cloudant presents two opportunities:
    1. A **launching point** for new NoSQL customers looking for low-risk, low-cost avenues for getting started with data warehousing & analytics in a multi-tenant environment on the cloud
    2. A **new point of entry** for existing Cloudant developers to access industry-leading INZA analytics and BLU warehousing
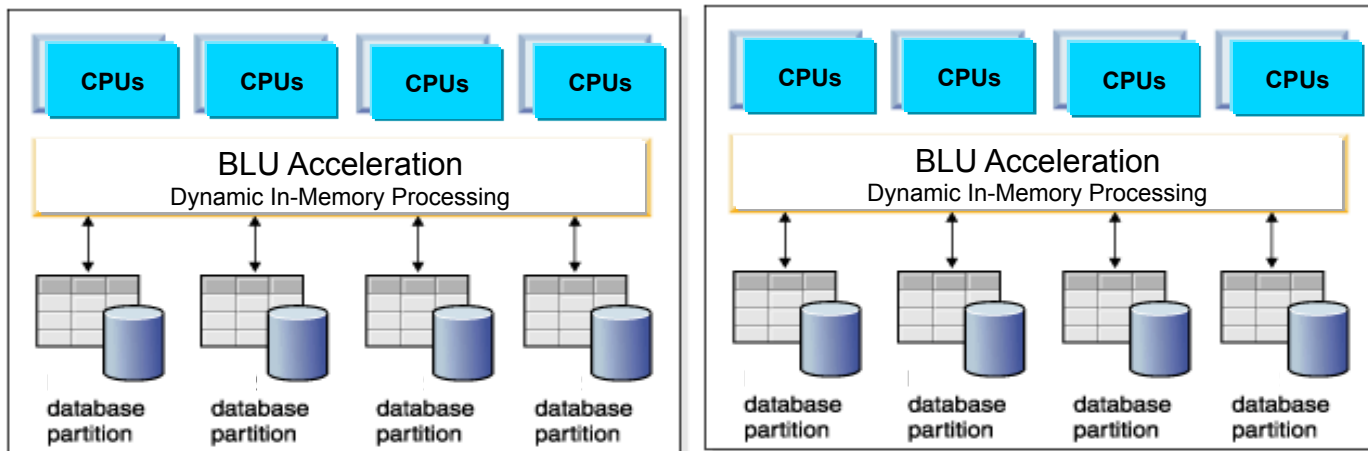   3. Automatic schema discovery from theJSON data. One button push to create warehouse.

dashDB MPP
~elastic growth

1. 4TB building blocks (subject to change)
2. Min 3 servers.
3. Elastic growth without re-hashing all row data
4. Webscale clustering
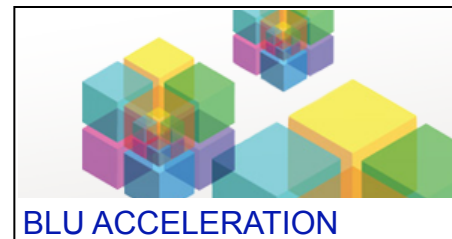
BLU ACCELERATION

MPP Scale-out of dashDB with BLU Acceleration

Now in Closed Beta!

The what & the wow

BALLUFF

sensors worldwide

1. Fast: Superior Performance 35-73X
2. Small RAM requirements: Typically just 5% of the original uncompressed user data.
3. Small: Superior Compression
4. Simple: Create. Load. Go.
5. Introducing dashDB for cloud
   - Fully managed service on IBM BlueMix
   - JSON NoSQL integration, Cloudant.com
   - R and Spatial analytics
6. SAP BW certified
7. Oracle compatibility @98% +

*"We **cut report runtimes by up to 98%** thanks to IBM DB2 with **BLU Acceleration technology – without changing** operations processes or investing in new hardware or software. We were impressed **how easy boosting database performance can be**."*

-Bernhard Herzog, Team Manager Information Technology SAP, Balluff

Where to find more information:

- dashDB.com

- ibmbluhub.com

- Sam Lightstone's BLOG: SoftwareTradecraft.com

- Sam Lightstone's papers on DBLP:

  - http://dblp.uni-trier.de/pers/hd/l/Lightstone:Sam

SEIZE THIS
MOMENT