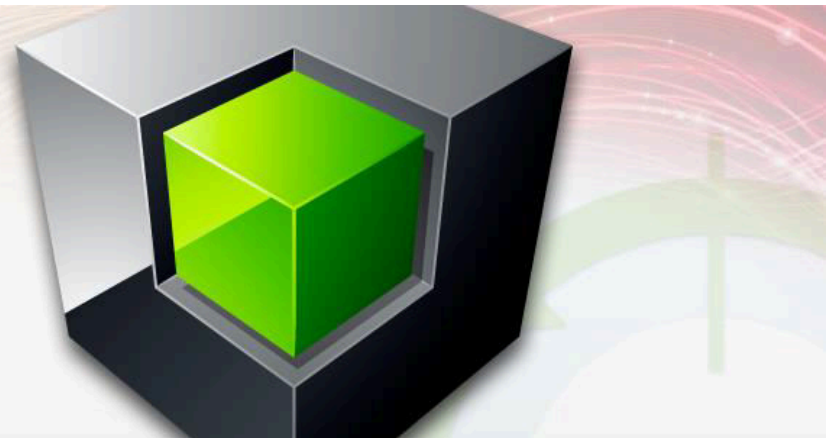# DB2® 10.5

with BLU Acceleration

Les King
Director, Information Management
lking@ca.ibm.com
August, 2013

Customer Video
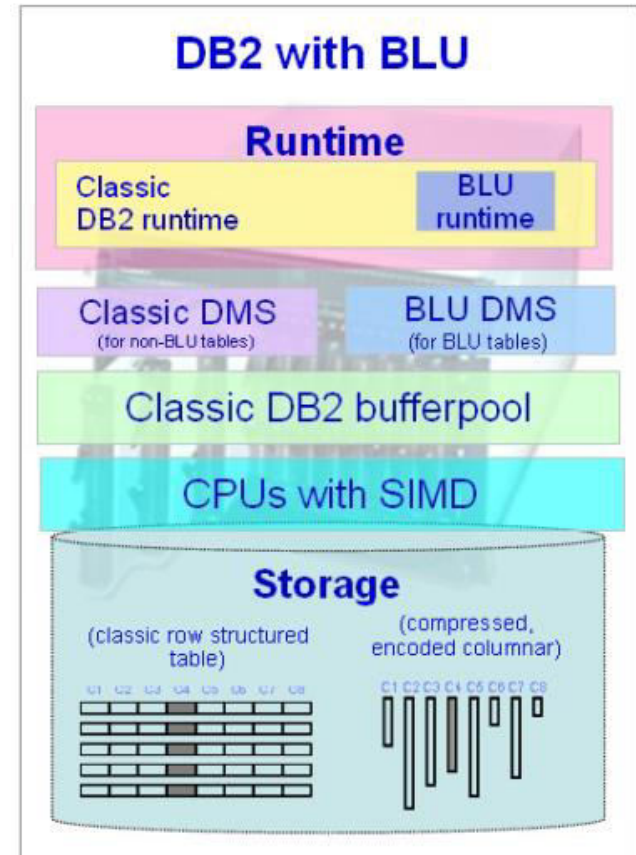
http://www.youtube.com/watch?v=5T6f74gYu1Y&noredirect=1

What is DB2 with BLU Acceleration?

- **New technology** **for analytic queries in DB2 LUW**
  - DB2 column-organized tables add columnar capabilities to DB2 databases
    - Table data is stored column organized rather than row organized
    - Using a vector processing engine
    - Using this table format with star schema data marts provides significant improvements to storage, query performance, ease of use, and time-to-value
  - New unique runtime technology which leverages the CPU architecture and is built directly into the DB2 kernel
  - New unique encoding for speed and compression
    - This new capability is both main-memory optimized, CPU optimized, and I/O optimized

DB2 with BLU

Runtime

| Classic DB2 runtime | BLU runtime |

| Classic DMS (for non-BLU tables) | BLU DMS (for BLU tables) |

Classic DB2 bufferpool

CPUs with SIMD

Storage

(classic row structured table) | (compressed, encoded columnar)

Target Use Cases for BLU Acceleration

1. Analytics and Reporting
- Single Server implementation
- Targeting environments with <20TB of active data
- Some DPF environments can now be single partition

2. Mixed Workload Environments
- 40%-50% of OLTP environments have analytics and reporting activity
- Accelerate the analytic and reporting workloads
- Reduce pressure on the OLTP workload

3. Cognos, OLAP, Dynamic Cubes
- Would provide same value in any other similar type environment

4. SAP BW
- In beta now
- Expect certification by end of 3Q

How Fast Is BLU Acceleration?

Customer	Performance Gains
BNSF	Up to 137x
Handelsbanken	7x – 100x
Triton Consulting	46x
Yonyou	40x
Coca-Cola Bottling	4x - 15x

~25x
speedup
is average

BNSF *RAILWAY*

"It was amazing to see the faster query times compared to the performance results with our row-organized tables. **The performance of four of our queries improved by over 100-fold! The best outcome was a query that finished 137x faster by using BLU Acceleration."**
- Kent Collins, Database Solutions Architect, BNSF Railway

BLU Acceleration Performance

New

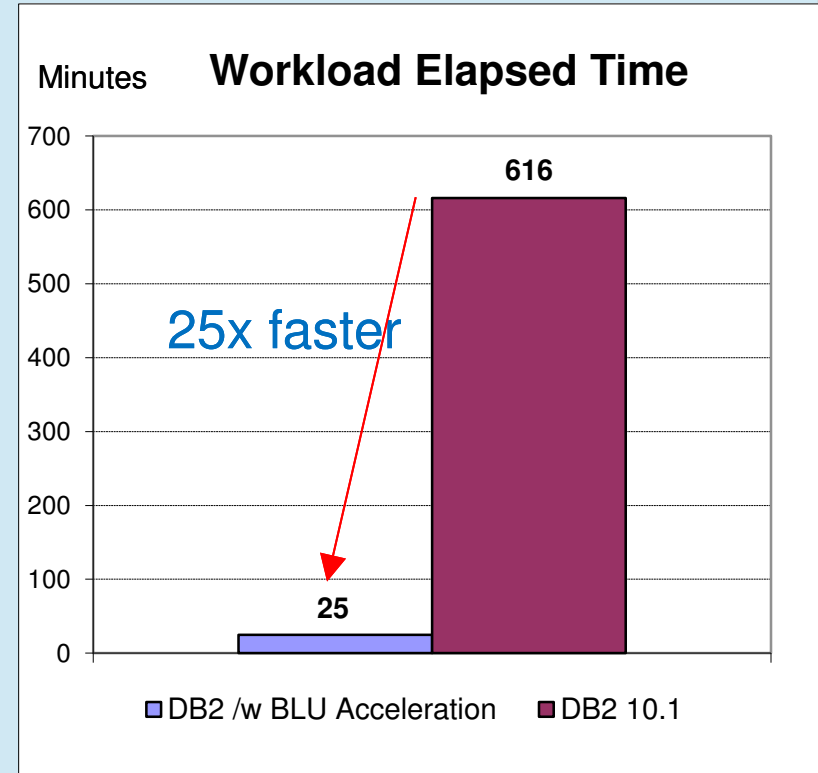- ▪ **POPS (Proof of Performance and Scalability)**
  - Derived from Redbrick performance test
  - Classic sales analytics

  - 5.5years of data (2000 days) for 63 stores
    - ~4TB of raw data
    - 2 fact tables
    - 5 dimension tables

  - Broad range of queries with varying selectivity / aggregation

- ▪ **Substantial Storage Savings with BLU Acceleration**
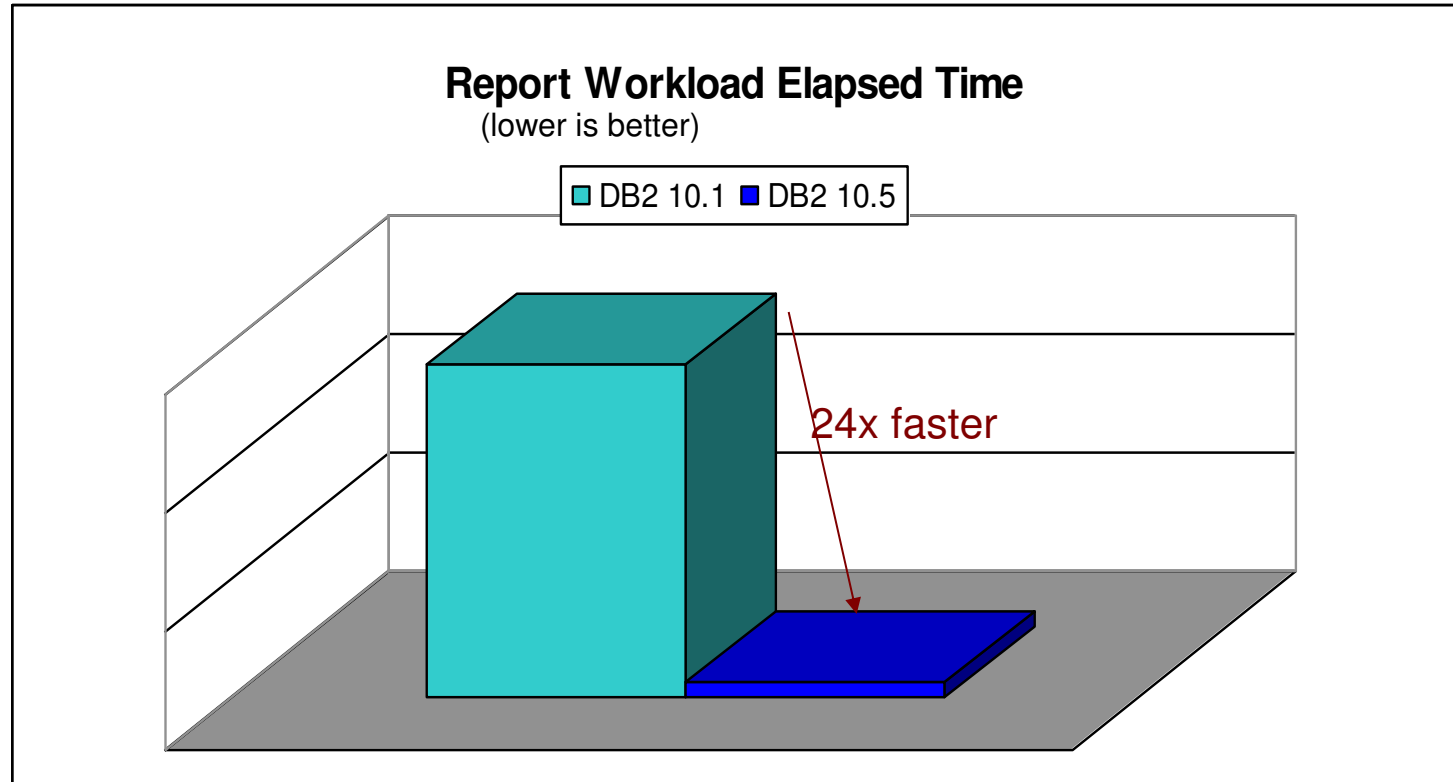 - 2.5x less space than DB2 10.1

- ▪ **Massive Performance Gains**
  - 25x speedup over DB2 10.1

  - Maximum query speed up over 400x

Workload Elapsed Time

Minutes

25x faster

616

25

☐ DB2 /w BLU Acceleration   ☐ DB2 10.1

Intel® Xeon® Processor E5-4650

(intel)

DB2 10.5 and Cognos BI 10.2 Dynamic Cubes on Power 7+

Report Workload Elapsed Time
(lower is better)
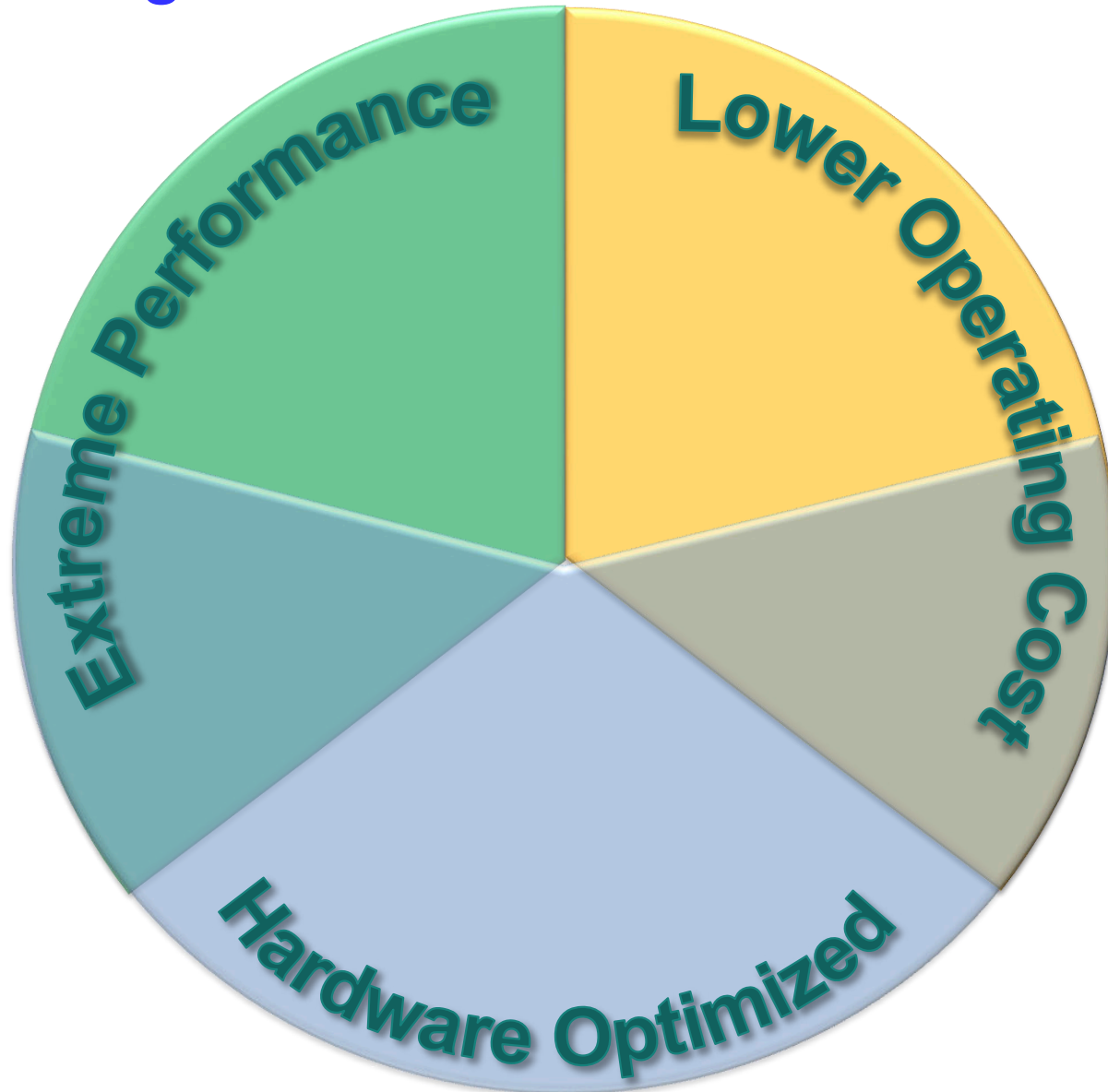
■ DB2 10.1   ■ DB2 10.5

24x faster

Storage Savings

- **Multiple examples of data requiring substantially less storage**
  - 5% of the uncompressed size
  - Fewer objects required
- **Multiple compression techniques**
  - Combined to create a near optimal compression strategy
- **Compression algorithm adapts to the data**
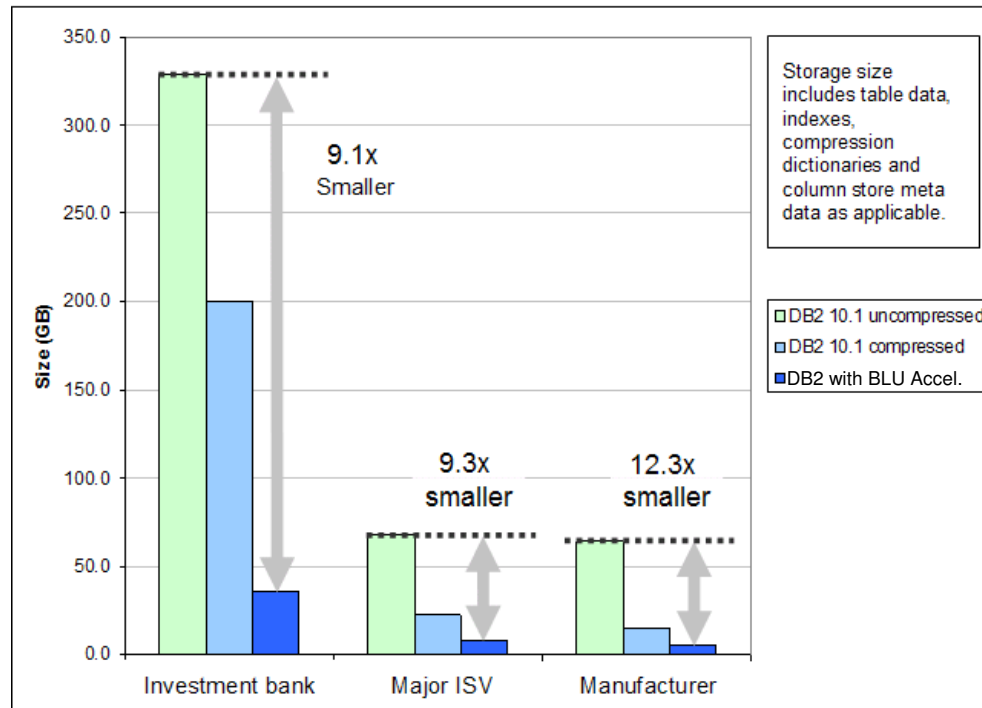
The Seven Big Ideas of DB2 with BLU Acceleration

7 Big Ideas: Our secret sauce

1

- **Compute-friendly encoding & compression**
  - Massive compression with approximate Huffman encoding
  - Evaluation while compressed!
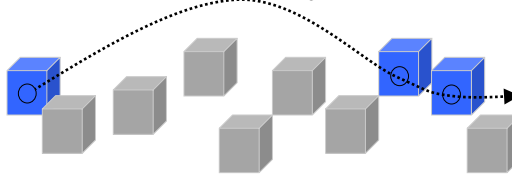 - Register-friendly encoding dramatically improves efficiency ~~of I/O,~~

2

- **Automatic Strata Maps**
 - Synopses enable automatic detection of large sections of data that can be ignored by a query
  - Order of magnitude savings in <u>all</u> of I/O, RAM and CPU
  - No DBA action to define or use – truly invisible.

3

- **Scan-friendly memory caching**
  - New algorithms cache in RAM effectively
  - High percent of interesting data fits in memory
  - Data can be larger than RAM

RAM
Near optimal caching

DISKS
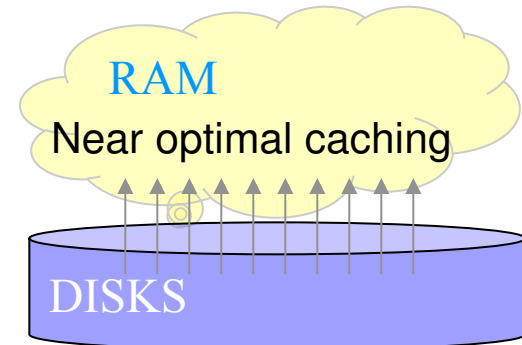
7 Big Ideas: Our secret sauce

4

- **Core-friendly parallelism**
  - KIWI: Kill It With Iron. Uniquely leverage multi-core
  - Careful attention maximizes CPU cache, cacheline efficiency.
  - Many have tried, few have succeeded

5

- **Multiply the power of the CPU**
  - Single Instruction Multiple Data (SIMD)
  - Compute with 1 instruction what may have taken >4.
  - CPU vector processing
  - Scanning, joins

Instructions   Data

SIMD

Results
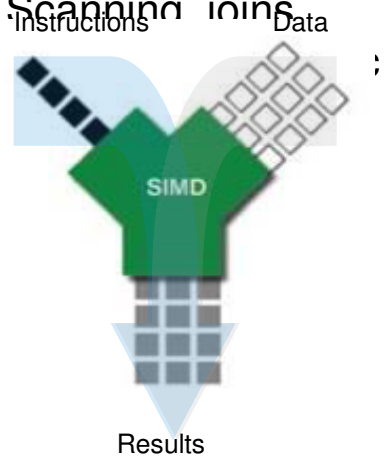
6

- **Column Store**
  - Minimal I/O
  - Improved compression
  - Favors scan based processing
  - L2 efficiency

C1 C2 C3 C4 C5 C6 C7 C8

7 Big Ideas: Our secret sauce

7

- **Single <u>mature</u> database system**
  - **BLU is part of DB2**
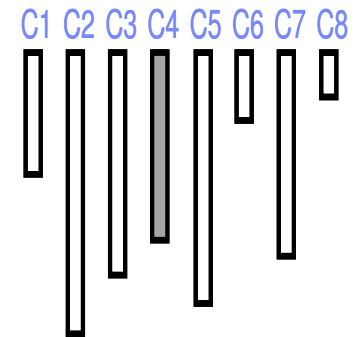   - Reuse DB2 process model, storage
  - **Rich function and admin tools**
    - All core SQL capabilities
      - DB2 classic runtime compensates for any missing pieces. (e.g. some scalar functions, sorting, etc)
    - Rich existing DB2 admin utilities and tools
  - **Reliable**
    - Majority of the system is mature DB2 code
    - New code is validated with **hundreds of thousands** of existing DB2 test cases, stress systems, and performance measurements.

DB2 with BLU

Runtime

Classic DB2 runtime | BLU runtime

Classic DMS (for non-BLU tables) | BLU DMS (for BLU tables) | Classic IXM

Classic DB2 bufferpool

CPUs with SIMD

Storage

(classic row structured table)

C1 C2 C3 C4 C5 C6 C7 C8

(compressed, encoded columnar)

C1 C2 C3 C4 C5 C6 C7 C8

BLU Acceleration : Memory optimized. CPU optimized. I/O optimized

Memory Optimized

– Memory latency optimized for
- Scans
- Joins
- Aggregation

– More useful data in memory
- Data stays compressed
- Scan friendly caching

– Less to put in memory
- Columnar access
- Late materialization
- Data skipping

CPU Optimized

–CPU acceleration
- SIMD processing for
  - Scans
  - Joins
  - Grouping
  - Arithmetic

–Keeping the CPUs busy
- Core friendly parallelism

–Less CPU processing
- Operate on compressed data
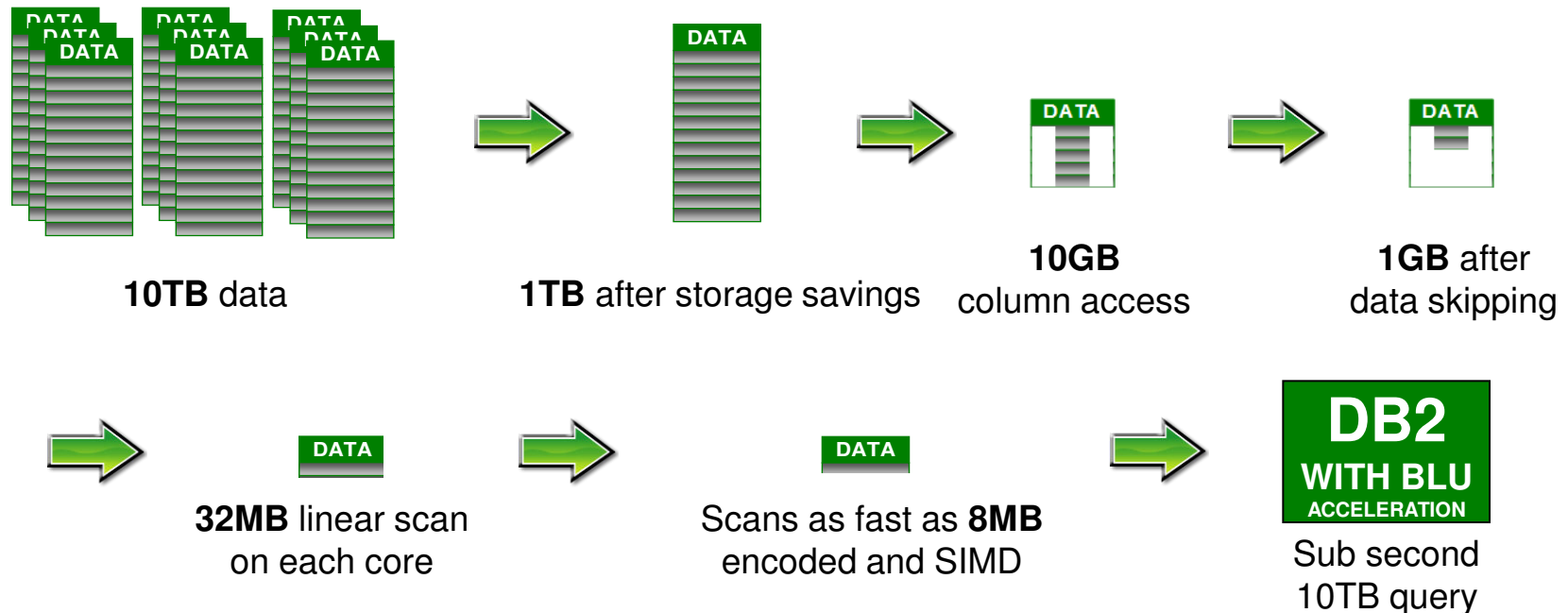- Late materialization
- Data skipping

I/O Optimized

– Less to read
- Columnar I/O
- Data skipping
- Late materialization

– Read less often
- Scan friendly caching

– Efficient I/O
- Specialized columnar prefetching algorithm

7 Big Ideas: How DB2 with BLU Acceleration Helps
~*Sub second 10TB query – An Optimistic Illustration*

- **The system – 32 cores, 10TB table with 100 columns, 10 years of data**
- **The query:** `SELECT COUNT(*) from MYTABLE where YEAR = '2010'`
- **The optimistic result: sub second 10TB query! Each CPU core examines the equivalent of just 8MB of data**

10TB data → **1TB** after storage savings → **10GB** column access → **1GB** after data skipping

→ **32MB** linear scan on each core → Scans as fast as **8MB** encoded and SIMD → **DB2 WITH BLU ACCELERATION** Sub second 10TB query

More Customer Experience

"It was amazing to see the faster query times compared to the performance results with our row-organized tables. The performance of four of our queries improved by over 100-fold! The best outcome was a query that finished 137x faster by using BLU Acceleration."
- Kent Collins, Database Solutions Architect, BNSF Railway

"We were very impressed with the performance and simplicity of BLU. We found that some queries achieved an almost 100x speed up with literally no tuning!"
- Lennart Henäng, IT Architect, Handelsbanken

"I have now reviewed IBM's new Big Data effort, BLU Acceleration, and my opinion is this: yes, it will deliver major performance enhancements in a wide variety for specific Big Data cases, and yes, I do view their claim of 1000x acceleration in some cases as credible."
– Wayne Kernochan, President

The performance of DB2 10.5 with BLU Acceleration is quite amazing. We ran our tests on a system that is about 16x less powerful than our current production system. And yet BLU was able to outperform our production system in every aspect. We were truly running our analytics on the DB2 10.5 column organized tables at the speed of thought."
–Mohankumar Saraswatipura, Lead DBA.Reckitt Benckiser

"With my analytic query workload running 45x times faster with BLU Acceleration in DB2 10.5, I no longer have an excuse for my usual coffee run!"
- Iqbal Goralwalla, Head of DB2 Managed Services, Triton

DB2 with BLU Acceleration Early Customer Quotes

"Using DB2 10.5 with BLU Acceleration, our storage consumption went down by about 10x compared to our storage requirements for uncompressed tables and indexes. In fact, I was surprised to find a 3x increase in storage savings compared to the great compression that we already observed with Adaptive Compression on the DB2 10.5 server."
- Kent Collins, Database Solutions Architect, BNSF Railway

*"One of the things I really like about BLU Acceleration is that it enables me to put **column-organized tables beside row-organized tables in the same database.** In our mixed environment, we realized an amazing **10-25x reduction in the storage** requirements for the database when taking into account the compression ratios, along with all the things **I no longer need to worry about: indexes, aggregates, and so on**."*
-Andrew Juarez, Lead SAP Basis and DBA

Thank You!