

IT-Tage 2015

Schwerpunkt: Datenbanken
14.12. - 18.12.2015
Frankfurt am Main

Stefan Hummel:
DB2 In-Memory - Eine Technologie nicht
nur für typisches OLAP

DB2 In-Memory Acceleration

Eine Technologie nicht nur für typisches OLAP

www.ibmBLUhub.com



Stefan Hummel
IT Specialist | IBM Germany

Disclaimer

© Copyright IBM Corporation 2015. All rights reserved.

U.S. Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

THE INFORMATION CONTAINED IN THIS PRESENTATION IS PROVIDED FOR INFORMATIONAL PURPOSES ONLY. WHILE EFFORTS WERE MADE TO VERIFY THE COMPLETENESS AND ACCURACY OF THE INFORMATION CONTAINED IN THIS PRESENTATION, IT IS PROVIDED “AS IS” WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. IN ADDITION, THIS INFORMATION IS BASED ON IBM'S CURRENT PRODUCT PLANS AND STRATEGY, WHICH ARE SUBJECT TO CHANGE BY IBM WITHOUT NOTICE. IBM SHALL NOT BE RESPONSIBLE FOR ANY DAMAGES ARISING OUT OF THE USE OF, OR OTHERWISE RELATED TO, THIS PRESENTATION OR ANY OTHER DOCUMENTATION. NOTHING CONTAINED IN THIS PRESENTATION IS INTENDED TO, NOR SHALL HAVE THE EFFECT OF, CREATING ANY WARRANTIES OR REPRESENTATIONS FROM IBM (OR ITS SUPPLIERS OR LICENSORS), OR ALTERING THE TERMS AND CONDITIONS OF ANY AGREEMENT OR LICENSE GOVERNING THE USE OF IBM PRODUCTS AND/OR SOFTWARE.

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion. Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision. The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

IBM, the IBM logo, ibm.com, Information Management, DB2, DB2 Connect, DB2 OLAP Server, pureScale, System Z, Cognos, solidDB, Informix, Optim, InfoSphere, and z/OS are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at www.ibm.com/legal/copytrade.shtml

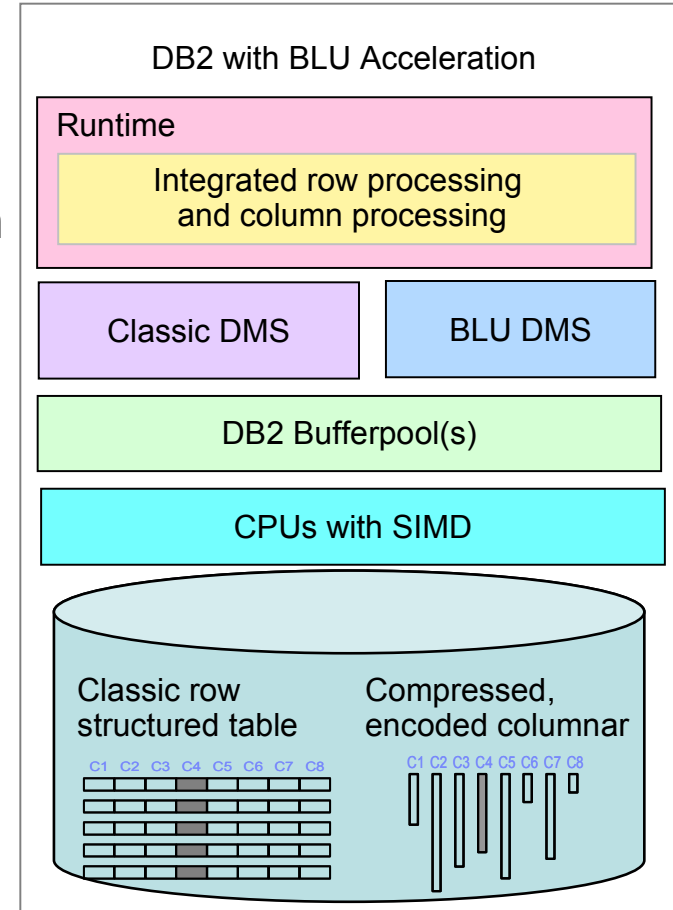
Other company, product, or service names may be trademarks or service marks of others.

What is DB2 with BLU Acceleration?

- **New technology in DB2 10.5, delivering large gains in performance, compression, and simplicity for analytic workloads**

- **Much more than just a columnar data store**
 - Columnar storage, vector processing
 - Unique data encoding for speed & compression
 - CPU-optimized SQL runtime processing
 - “*Better than in-memory*”
 - All built directly into the DB2 kernel !

- **Revolution or evolution**
 - Enable BLU at the table or database level
 - BLU tables coexists with traditional row tables in same schema, storage, and memory
 - Easy conversion



Using BLU is Easy !

- Example:

```
CREATE TABLE sales_col (  
    c1 INTEGER NOT NULL,  
    c2 INTEGER,  
  
    ...  
    PRIMARY KEY (c1) ) ORGANIZE BY COLUMN;
```

Columnar tables are always compressed by default.

- No need to define indexes, compression, etc.
- No need to change load, backup/restore, etc.
- Can set database parameter `dft_table_org = COLUMN` so that all tables are column-organized by default
- Automatic workload management
- Automatic space reclamation in case of deletions

What makes BLU Acceleration different?

Unmatched innovations from IBM Research & Development labs

Next Generation In-Memory

In-memory columnar processing with dynamic movement of data from storage



Analyze Compressed Data

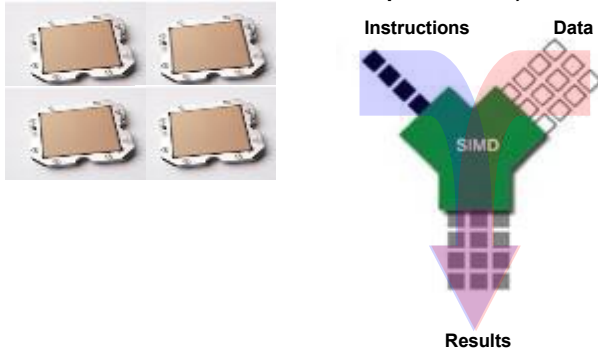
Patented compression technique that preserves order so data can be used without decompressing



Encoded

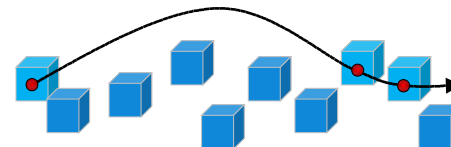
CPU Acceleration

Multi-core and SIMD parallelism
(Single Instruction Multiple Data)



Data Skipping

Skips unnecessary processing of irrelevant data



BLU Acceleration illustration

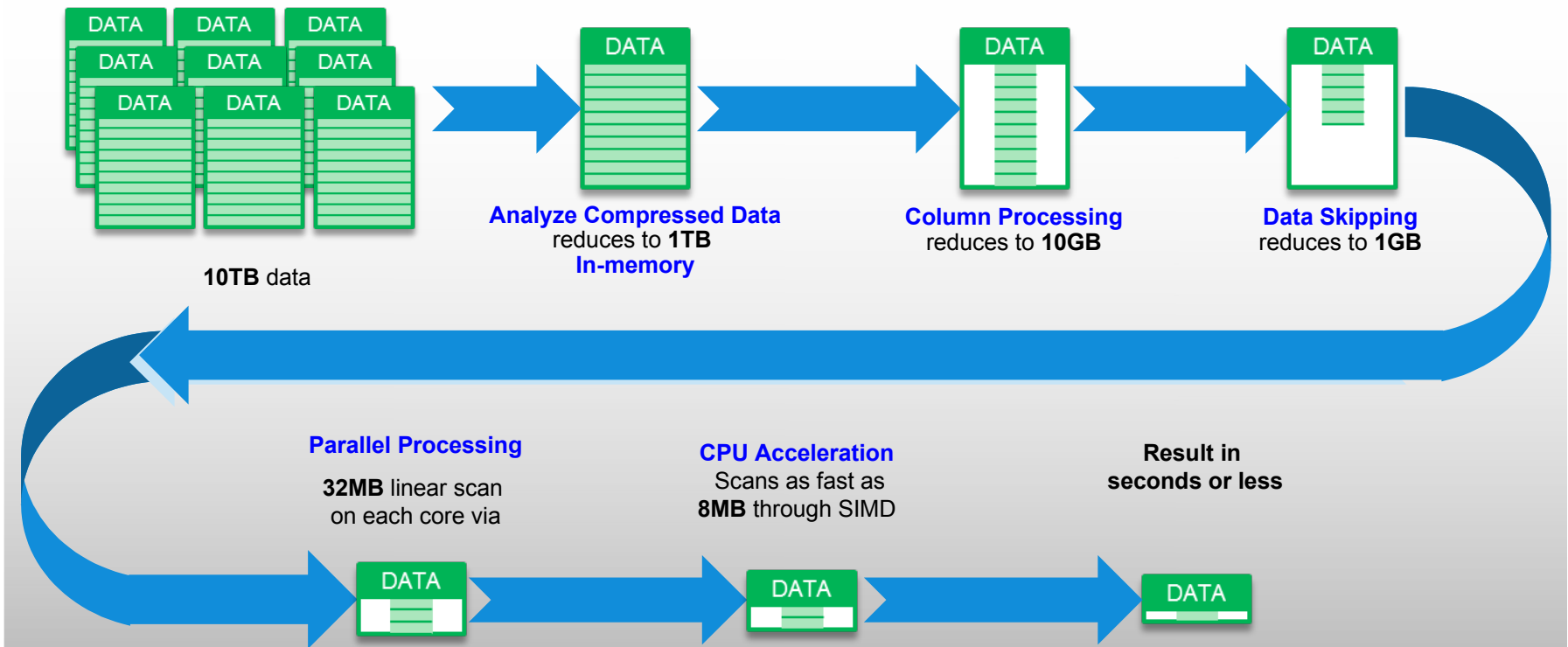
10TB query in seconds or less

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

Query: How many “sales” did we have in 2010?

```
- SELECT COUNT(*) from MYTABLE where YEAR = '2010'
```

Result: In seconds or less as each CPU core examines the equivalent of just 8MB of data



General Sizing Guideline for DB2 BLU

- Assumptions:
 - 20-30 active concurrent users
 - 80% simple, 20% complex queries
 - 30% active rows, 50% active columns, 7.5x compression (raw to BLU)
 - Goal: all or most of the active data in memory
- **Sizing:** Use 8 cores (or IFLs) and 128GB RAM (16GB/core) for each 3 TB of uncompressed raw data

Raw Data	#Cores	Main Memory
3TB	8	128 GB
6TB	16	256 GB
9TB	24	384 GB
12TB	32	512 GB
...

DB2 for Big Data

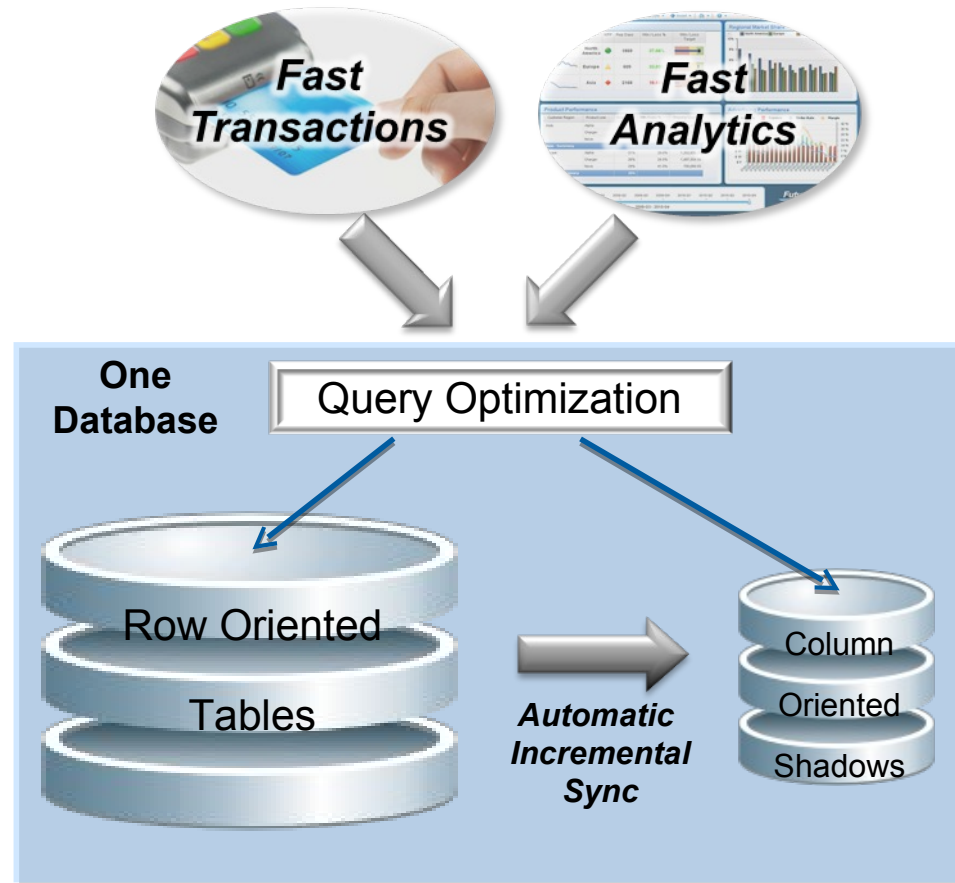
- Combination of DB2 Advanced Enterprise Server Edition (DB2 AESE) and BigInsights for Apache Hadoop
- Change mix of deployment at any time
- SQL language for DB2 and Hadoop (BigSQL)
- Use Cases
 - Verknüpfung von strukturierten mit unstrukturierten Daten über BigSQL
 - Sandboxing for Power User / Advanced Analytics / Performance Offloading
 - ETL Processing
 - Archivierung sowie Zugriff über BigSQL

BLU Shadow Tables

- Instant insight into operational data without compromising transaction performance
- DB2 creates column-based ‘Shadow Table’ versions of row-based operational data
- Analytic queries are seamlessly routed to Shadow Tables to take advantage of BLU Acceleration analytics performance in the transaction processing environment
- With BLU Shadow Tables, the performance of analytical queries can improve by 10x or more, with equal or greater transactional performance*. In one instance, the removal of secondary analytical indexes improved transactional performance by 2x**

* - Based on internal IBM testing of sample transactional and analytic workloads by replacing 4 secondary analytical indexes in the transactional environment with BLU Shadow Tables. Performance improvement figures are cumulative of all queries in the workload. Individual results will vary depending on individual workloads, configurations and conditions.

** - Based on internal IBM testing of sample transactional and analytic workloads by replacing 20 secondary analytical indexes in the transactional environment with BLU Shadow Tables. Performance improvement figures are cumulative of all queries in the workload. Individual results will vary depending on individual workloads, configurations and conditions.



Reporting and Transactions in the same continuously available system

