IDUG
Leading the DB2 User Community since 1988

IDUG DB2 Tech Conference
Prague, Czech Republic | November 2014

#IDUG

IBM Data Retrieval Technologies: RDBMS, BLU, IBM Netezza, and Hadoop

Frank C. Fillmore, Jr.
The Fillmore Group, Inc.

Session Code: E7
Tuesday, 11th November 2014 15:15 – 16:15 | Platform: z/OS, LUW, Netezza, Hadoop

THE FILLMORE GROUP
Relational Database Solutions

Agenda

- The Fillmore Group
- Relational Database Management Systems (RDBMS)
- Netezza - Field-programmable Gate Array (FPGA)
- DB2 10.5 with BLU Acceleration - Columnar
- Hadoop - MapReduce
- Wrap-up, Q&A, Evaluations

"Rarely affirm, seldom deny, always distinguish."
St. Thomas Aquinas

Premier Business Partner IBM

2

THE FILLMORE GROUP
Relational Database Solutions

The Fillmore Group, Inc.

- Frank Fillmore, DB2 Gold Consultant, IBM Champion
- Kim May, IBM Champion
- Dozens of experienced, certified consultants
- Information Management Consulting Services
- IBM Authorized Training Partner
- IBM Information Management Software Reseller

IBM CHAMPION

3

Relational Database Management System

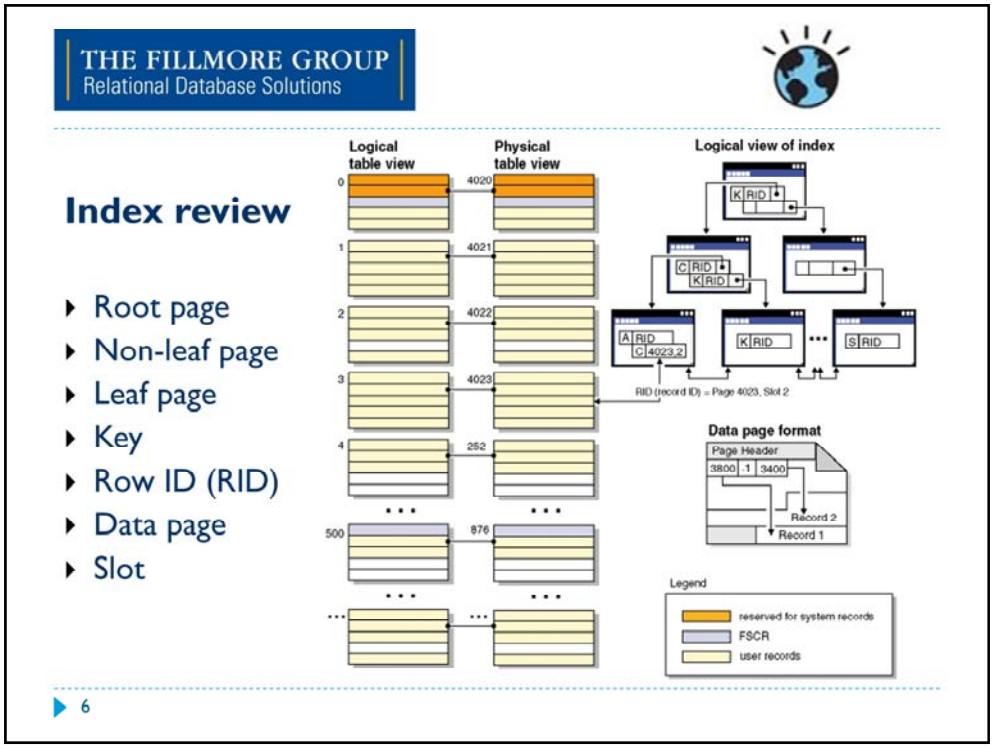
THE FILLMORE GROUP
Relational Database Solutions

Relational Database Management System

- The "foundation" characteristics
  - Data stored in ***tabular*** form (relations)
  - Schema stored in centralized repository
  - Strong datatyping
  - B-tree ***indexes*** used to create alternate access paths
  - Cost-based optimizer chooses access path
- Examples: DB2 for z/OS, LUW

Premier Business Partner IBM

5



http://pic.dhe.ibm.com/infocenter/db2luw/v9r7/index.jsp?topic=%2Fcom.ibm.db2.luw.admin.perf.doc%2Fdoc%2Fc0005424.html

THE FILLMORE GROUP
Relational Database Solutions

Relational Database Management System

- Typical use cases
  - Online transaction processing (OLTP)
  - E-Commerce
- Extended use cases
  - Data warehousing
  - Online analytical processing (OLAP)

Premier Business Partner IBM

7

THE FILLMORE GROUP
Relational Database Solutions

Issues and limitations

- Rigid schema and strong datatyping diminish flexibility and ***responsiveness*** to change
- Data is transferred in units of ***pages*** from disk to memory for processing
- At billions of rows in a table, B-tree index I/O becomes expensive, impacting ***response*** times

Premier Business Partner IBM

8

Netezza

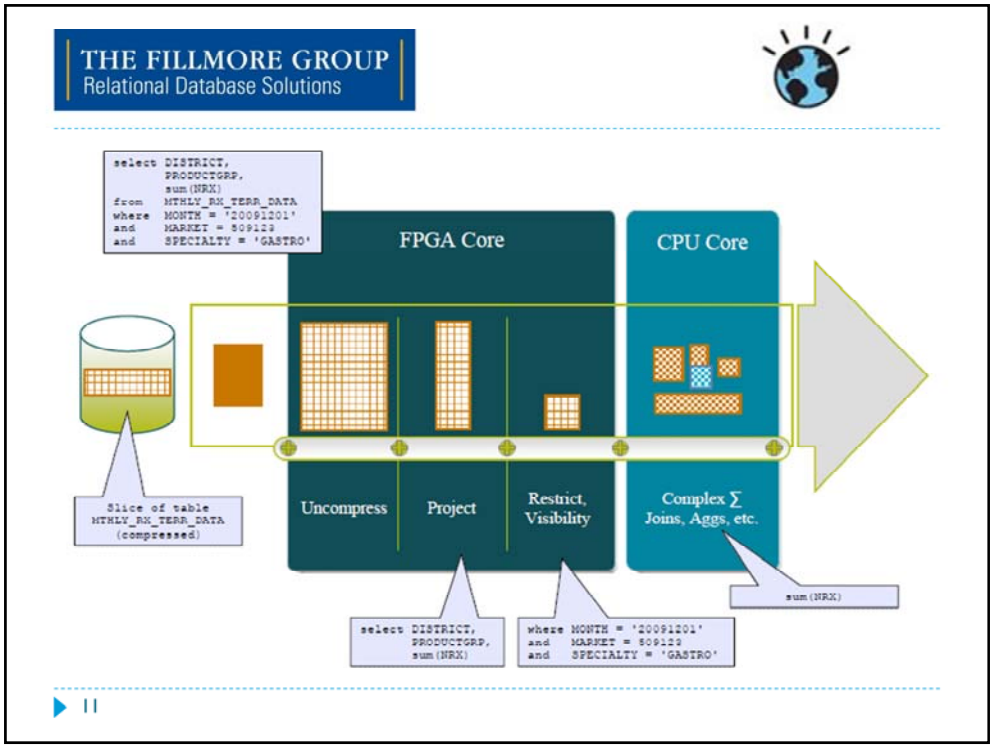
THE FILLMORE GROUP
Relational Database Solutions

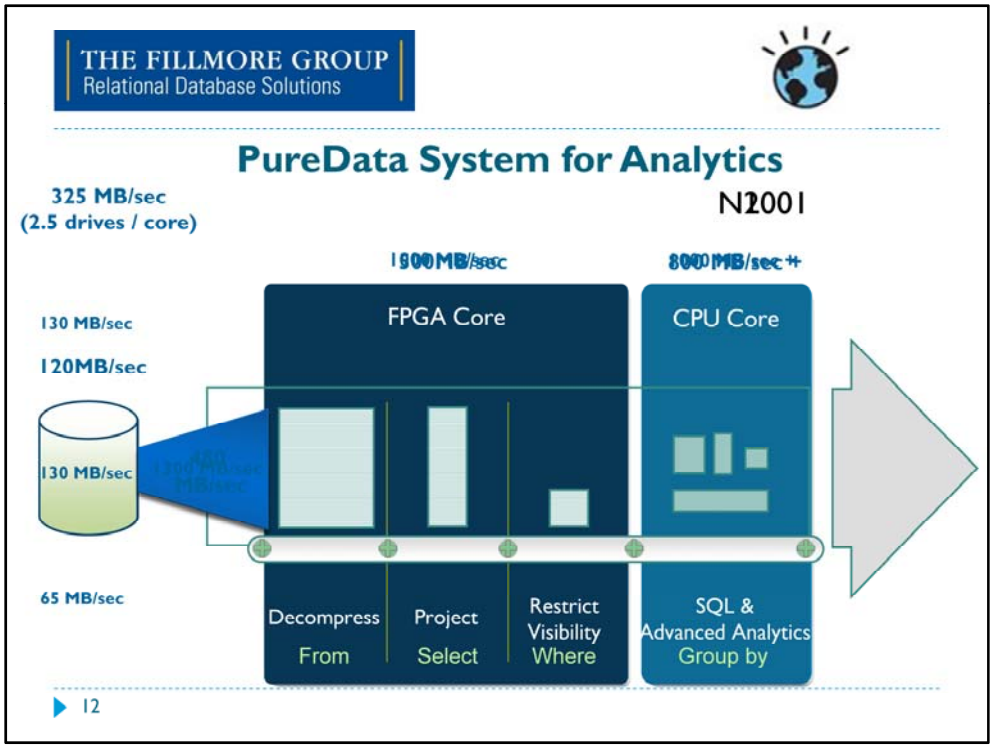
Netezza – aka PureData System for Analytics

- Appliance: integrated database, server, storage
- Key concepts
  - Field-programmable Gate Array (FPGA)
  - Data slice
  - Distribution Key
  - Zone Map

Premier Business Partner

10


THE FILLMORE GROUP
Relational Database Solutions
select DISTRICT,
PRODUCTGRP,
sum(NRX)
from MTHLY_RX_TERR_DATA
where MONTH = '200912O1'
and MARKET = 509123
and SPECIALTY = 'GASTRO'
FPGA Core
CPU Core
Slice of table
MTHLY_RX_TERR_DATA
(compressed)
Uncompress
Project
Restrict,
Visibility
Complex ∑
Joins, Aggs, etc.
sum(NRX)
select DISTRICT,
PRODUCTGRP,
sum(NRX)
where MONTH = '200912O1'
and MARKET = 509123
and SPECIALTY = 'GASTRO'



Netezza_under_the_hood Feinsmith (page 8)

What we see here is how we increased the scan speeds. Basically about each drive in the N2001 is capable of delivering about 130 megabytes per second of throughput (compared to approximately 120 MB/sec in the N1001).

What we had before with the 1 to 1 to 1 ratio, one drive to one FPGA core to one CPU core, the speed of the drive was the limiting factor as far as how fast the FPGA core could process the data - because the FPGA core could handle way more than that and the CPU core even more than the FPGA core. So the speed of the drive was a limiting factor.

So we now have more than one drive per FPGA core and per CPU core. Using basic math, we have about 2 1/2 drives per FPGA core and per CPU core. So that drives up the speed that data can be scanned and delivered to the FPGA for processing up to about 325 megabytes per second. If you add in the 4x compression that's going to get up to around 1300 megabytes per second.

IN the N2001 we now have faster FPGA cores that can process about 1000 megabytes per second. So we are no longer I/O starved.

So we're now delivering both 2 1/2 times as much data, you know, per second to the FPGA , to the CPU core and that is how we fundamentally and how we increased the scan speed and how we increase the performance of this system.

THE FILLMORE GROUP
Relational Database Solutions

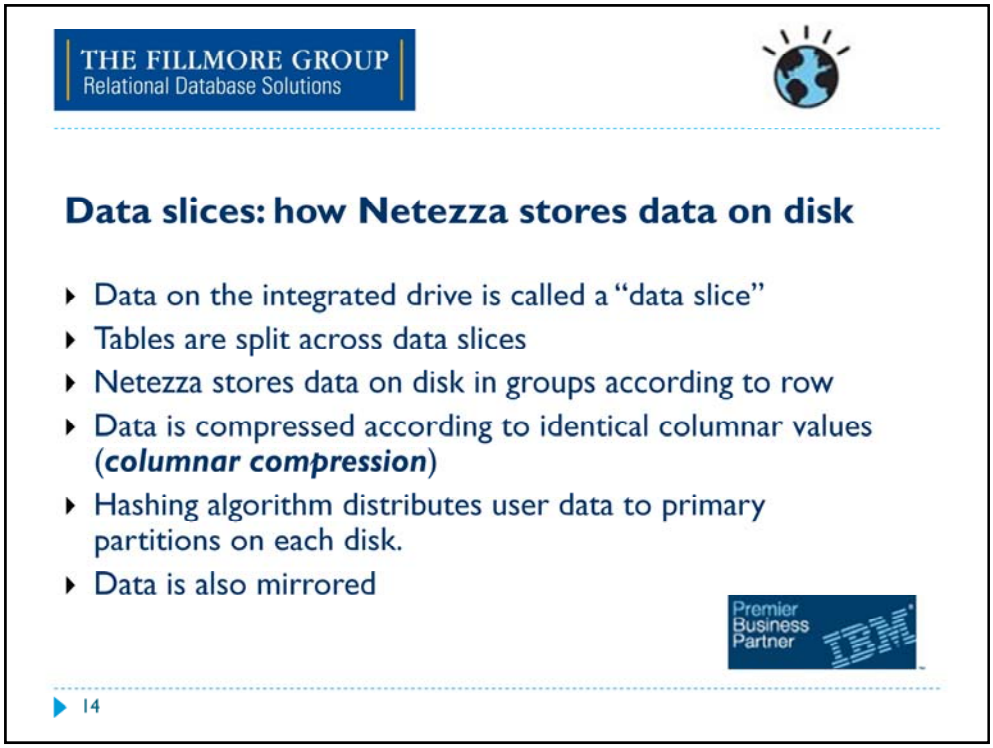
Just say 'No'tezza

- No data modeling
- No ***indexes***/MQTs/MDCs
- No dbspace/tablespace sizing and configuration
- No page/block/extent sizing and configuration for tables
- No redo/physical/journaling/logical log sizing and configuration

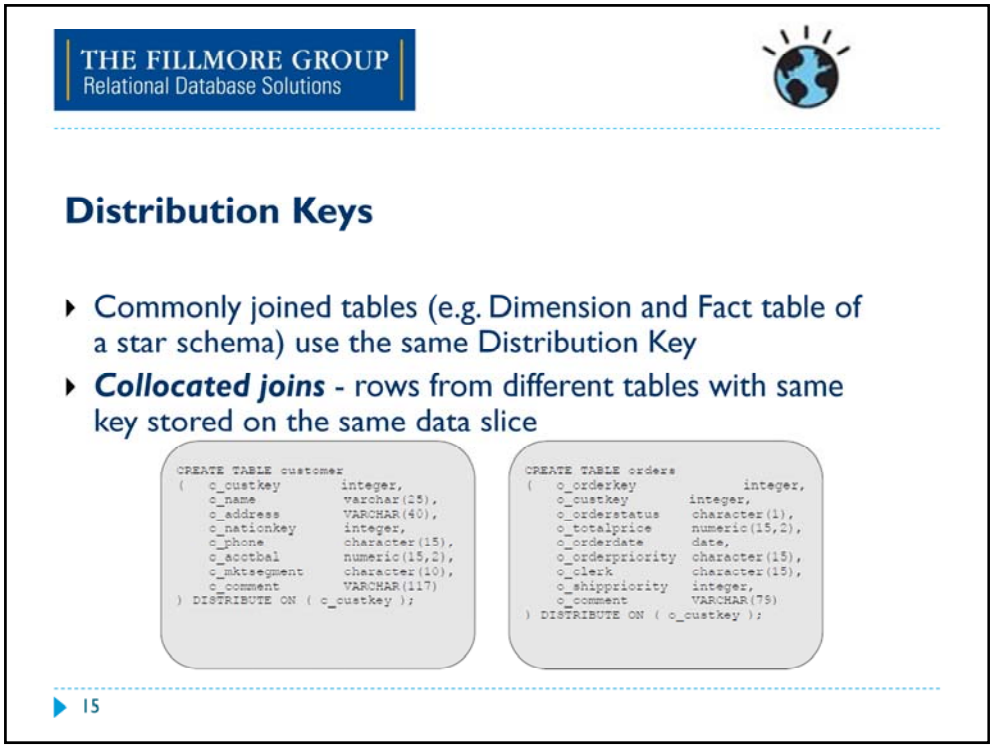
- One partitioning strategy: ***hashing algorithm***

Premier Business Partner IBM

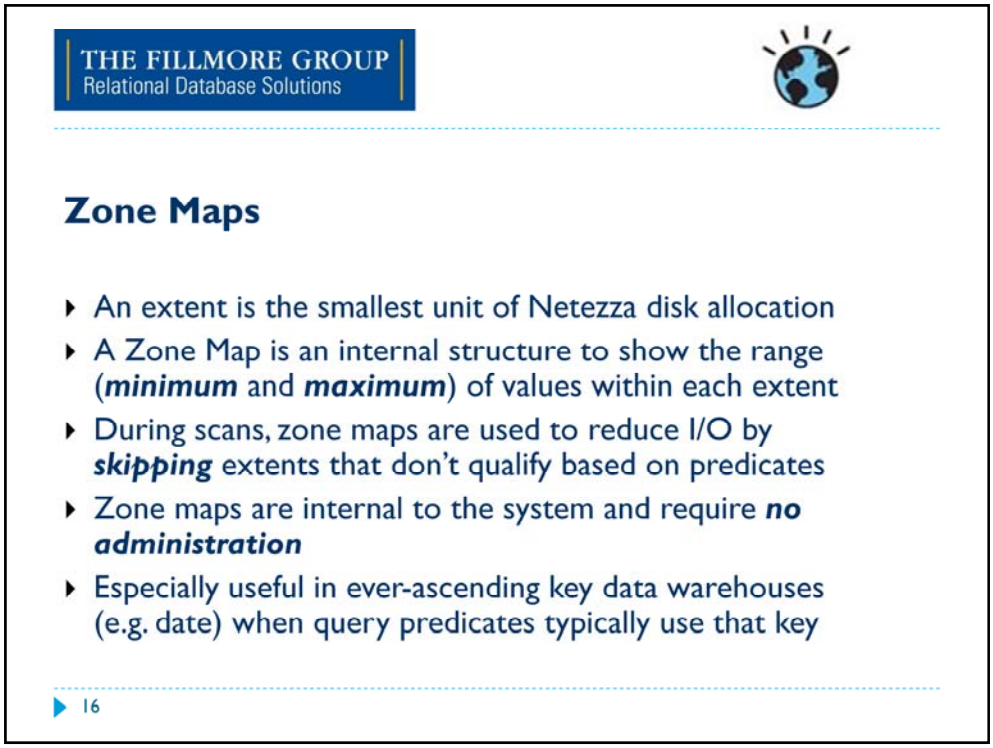
13



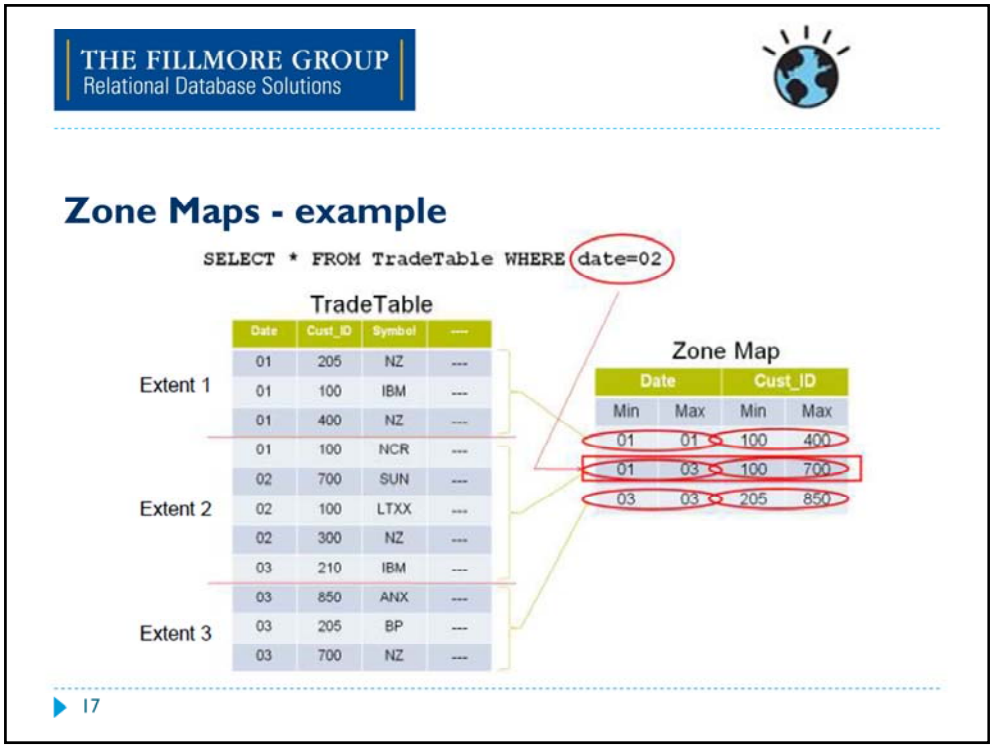
Netezza Bootcamp (page 62).



Netezza Bootcamp (page 133).



Netezza Bootcamp (page 227).



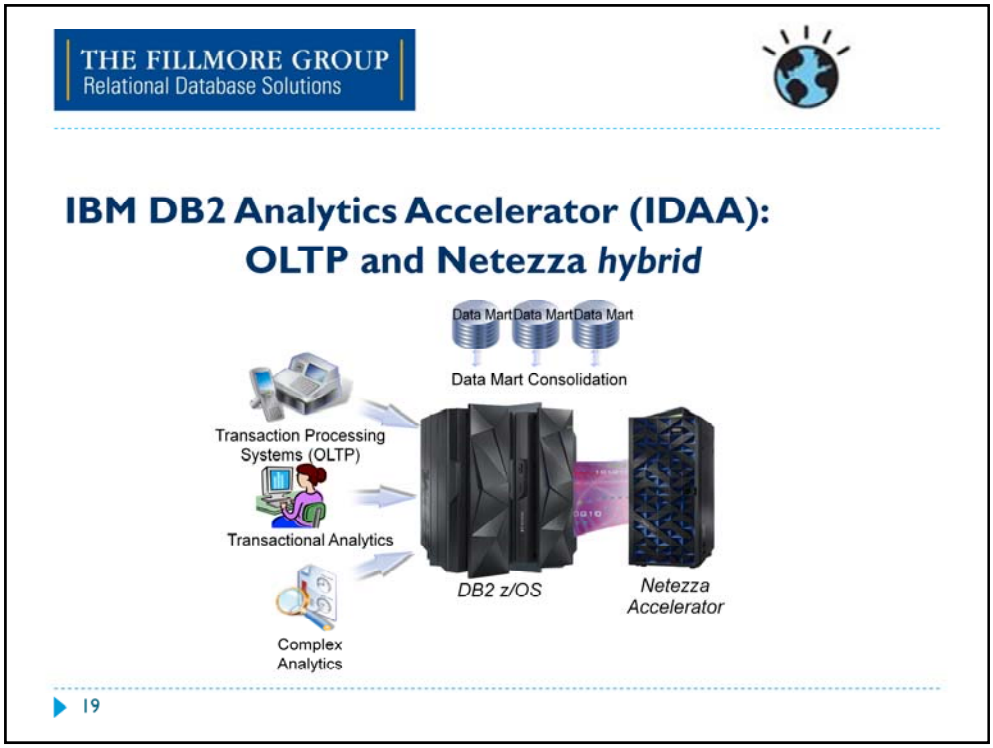
Netezza Bootcamp (page 228).

THE FILLMORE GROUP
Relational Database Solutions

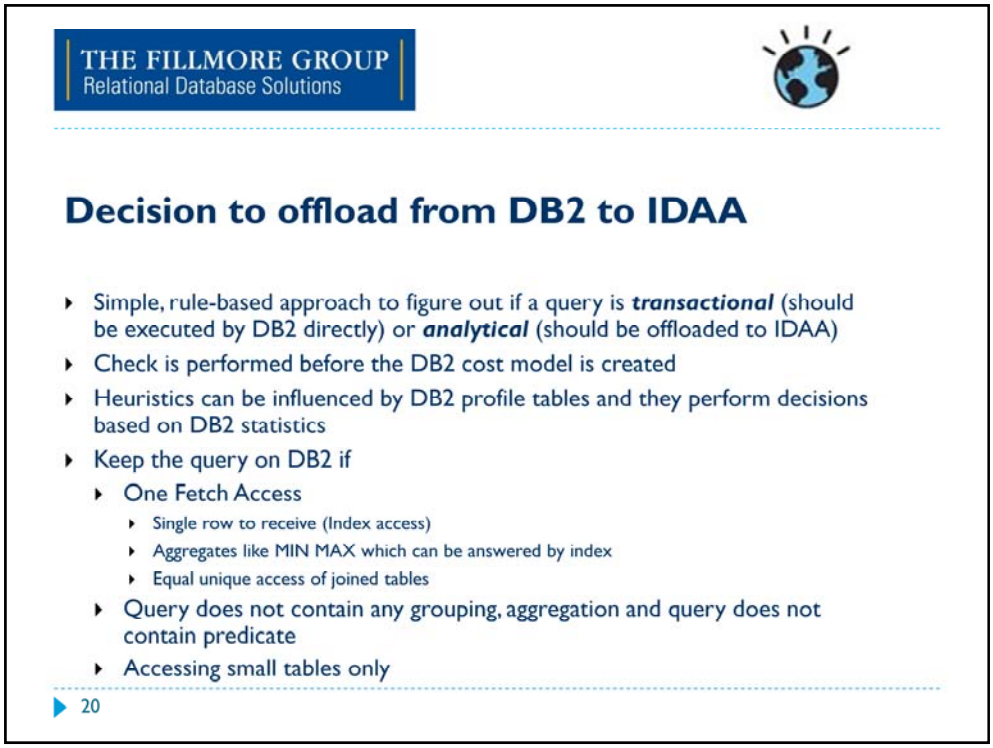
Issues and limitations

- Rigid schema and strong datatyping diminish flexibility and ***responsiveness*** to change
- Netezza is not optimized to process SQL Data Manipulation Language (DML)
  - INSERTs
  - UPDATEs
  - DELETEs

  i.e. OLTP transactions



IZAS_zEnterprise_Analytics Favero, et al (page 21)



Favero, et al (page 69)

THE FILLMORE GROUP
Relational Database Solutions

Checkpoint

- Q: Hashed Distribution Keys and collocation of data is similar to what DB2 for LUW technology?
- A: Data Partitioning Feature (DPF)
- Q: Zone Maps are *somewhat* similar to what DB2 index component?
- A: Leaf pages

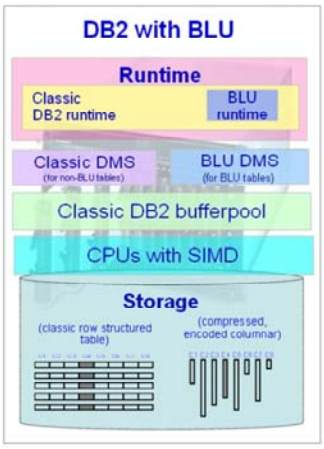
"There's no such thing as a free lunch."
Milton Friedman

DB2 10.5 BLU Acceleration

THE FILLMORE GROUP
Relational Database Solutions

DB2 10.5 BLU Acceleration

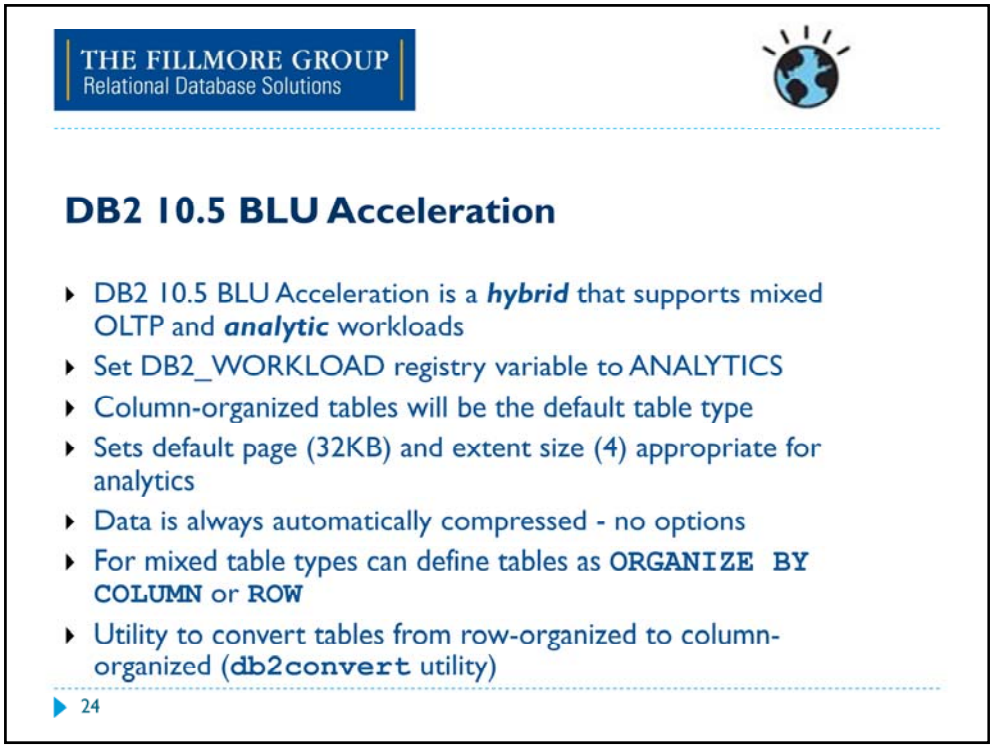
- Memory optimized
  - In-memory ***columnar*** processing
  - Dynamic data movement from storage (no LRU)
- Actionable Compression
  - Patented ***compression*** technique that preserves order so that the data can be used without decompressing (column cardinality)
- Parallel Vector Processing
  - Multi-core and Single Instruction Multiple Data (SIMD) parallelism
- Data Skipping
  - Skips unnecessary processing of irrelevant data



Premier Business Partner IBM

23

Schiefer (page 4)

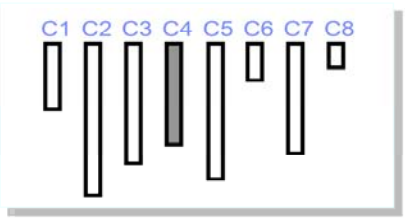


Schiefer (page 10)

THE FILLMORE GROUP
Relational Database Solutions

Column-oriented storage

Columns stored separately and packed in different buffers in memory



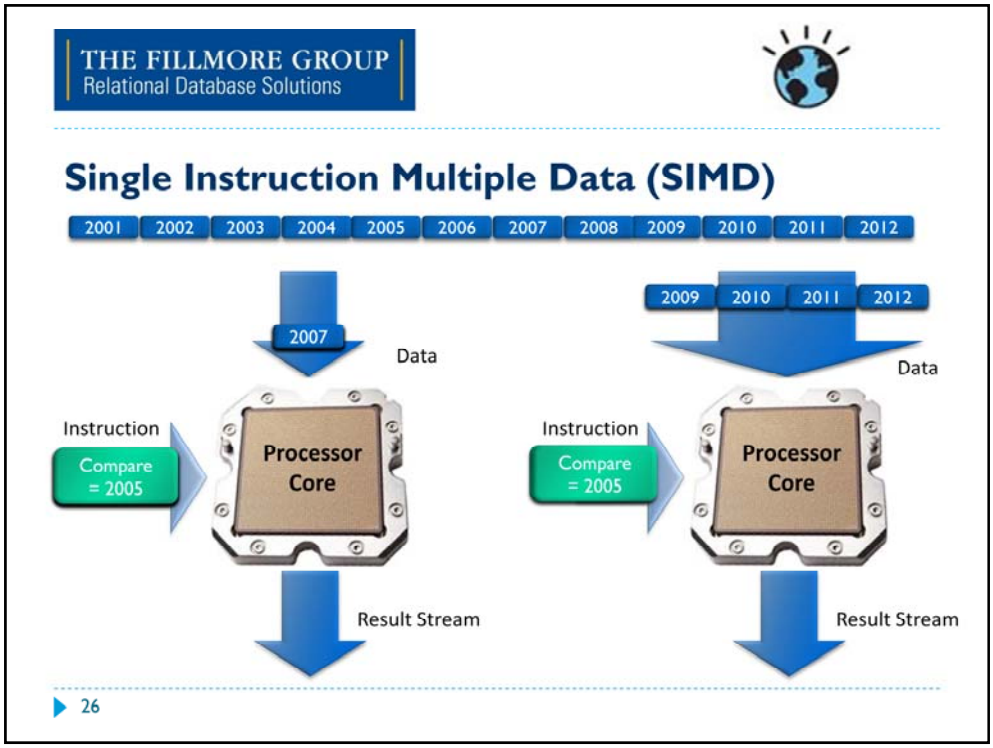
SELECT C4 ... WHERE C4=X

Consumes I/O bandwidth memory buffers and memory bandwidth only for C4

Premier Business Partner IBM

25

Schiefer (page 16)

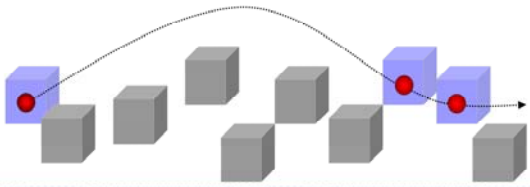


Schiefer (page 14)

THE FILLMORE GROUP
Relational Database Solutions

Data skipping

- Retrieve only data that meets predicate criteria
- "Synopsis" automatically created and maintained as data is LOADed or INSERTed (***no administration***)
- Persistent storage of ***minimum*** and ***maximum*** values for sections of data values



27

Schiefer (page 18)

THE FILLMORE GROUP
Relational Database Solutions

Issues and limitations

- Rigid schema and strong datatyping diminish flexibility and ***responsiveness*** to change
- Columnar tables optimized for read-mostly

THE FILLMORE GROUP
Relational Database Solutions

Checkpoint

- Q: Which two technologies surveyed so far focus on analytics?
- A: Netezza and DB2 10.5 BLU Acceleration
- Q: Why do the following matter?
  - number of columns in a row
  - ratio of columns retrieved to columns in the row
- A: If there are many columns in a row and most of them are retrieved most of the time, `ORGANIZE BY ROW` might be more efficient

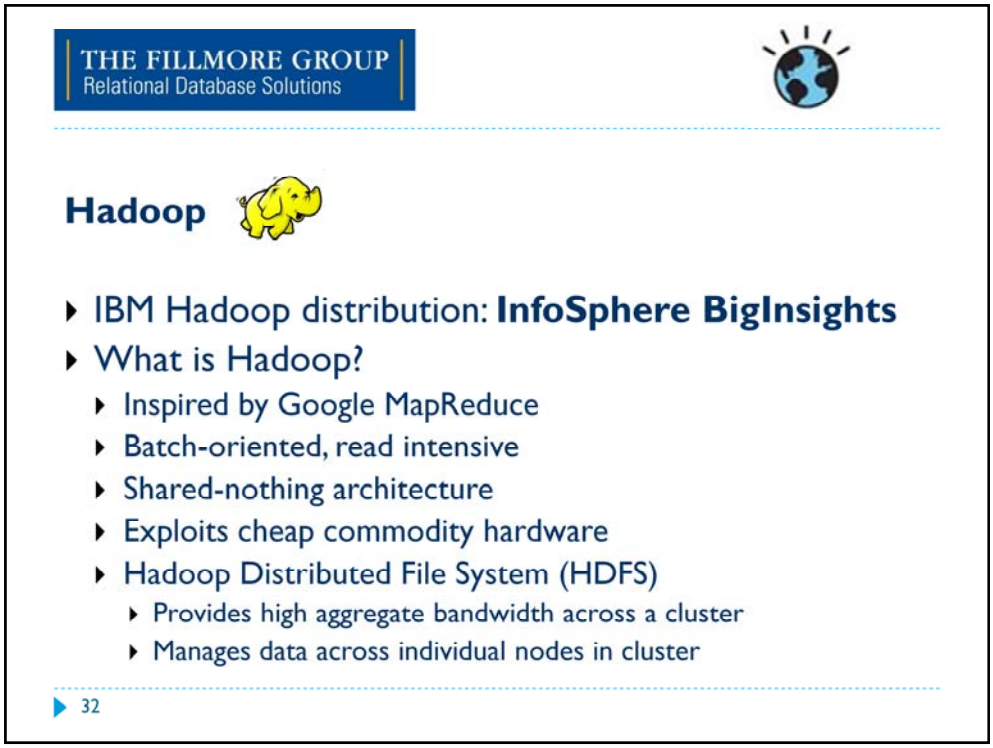
THE FILLMORE GROUP
Relational Database Solutions

Checkpoint

- DB2 10.5 is *similar* to IDAA as a ***hybrid*** platform
- Factors to consider when choosing between standalone Netezza and integrated DB2 10.5 BLU Acceleration
  - Business SLAs for ETL processing between OLTP and analytics databases
  - Existing skills
  - Appliance (Netezza) vs. hardware/software integration (DB2)
  - Relative performance?
- Q: Data skipping sounds like which Netezza feature?
- A: Zone Maps

30

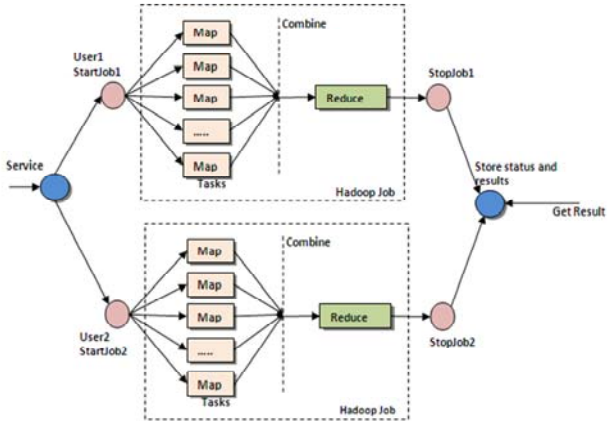
Hadoop



DW611 (page 2-7)

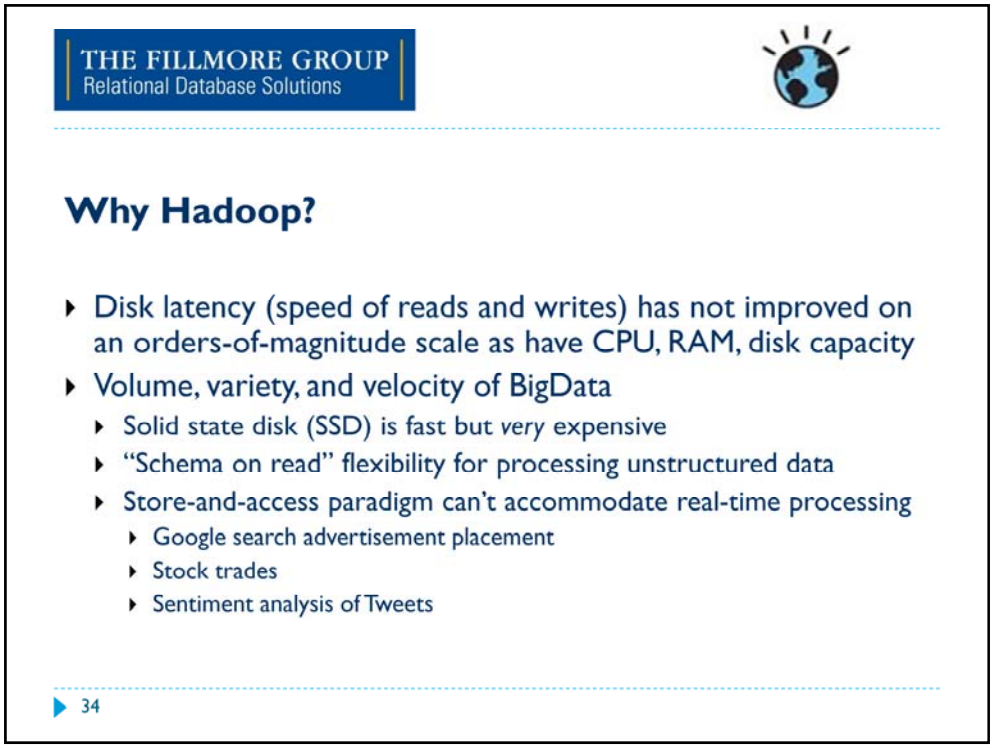
THE FILLMORE GROUP
Relational Database Solutions

MapReduce process flow



33

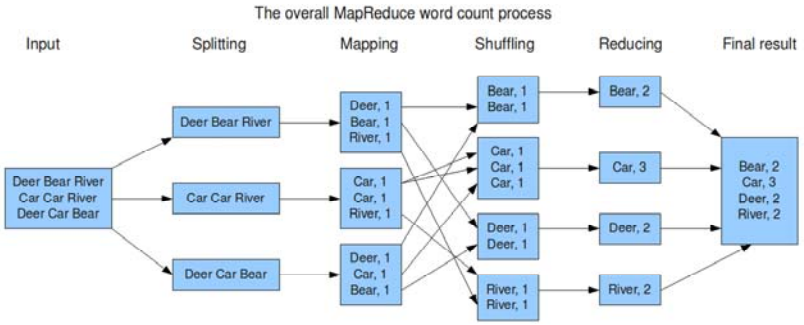
DW611 (chapter 3)



DW611 (chapter 3)

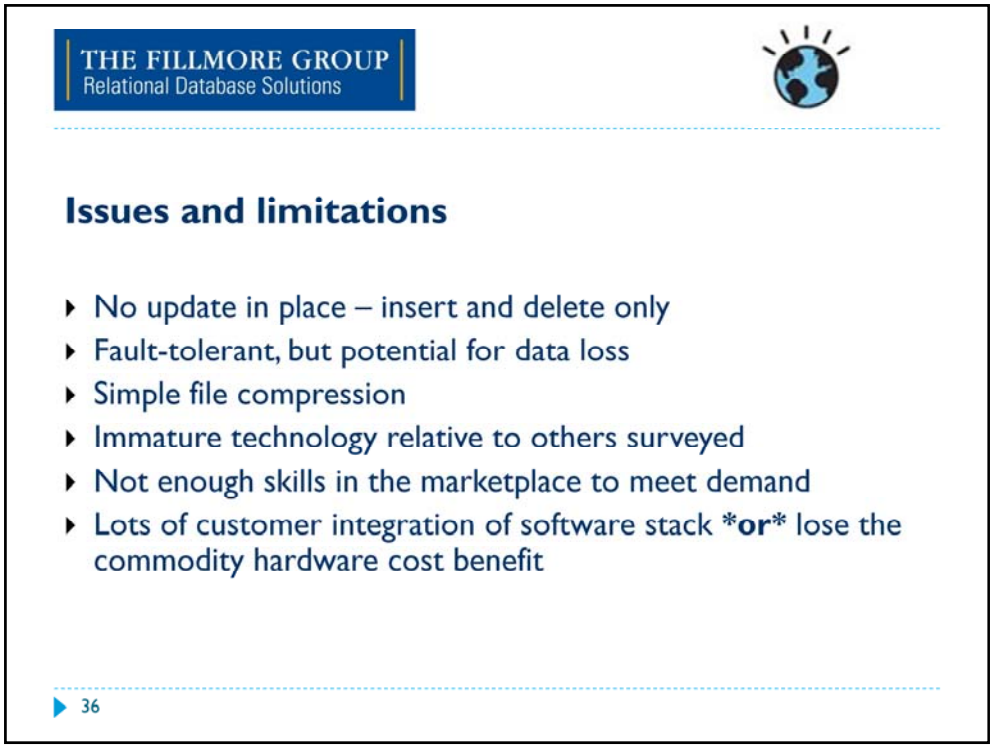
THE FILLMORE GROUP
Relational Database Solutions

Hadoop conceptual example



35

DW611 (chapter 3)



DW611 (chapter 3)

THE FILLMORE GROUP
Relational Database Solutions

Checkpoint

- Q: MapReduce distribution of portions of the data to different nodes (Map) and then reassembling the answer set (Reduce) is similar to what DB2 for LUW technology?
- A: Data Partitioning Feature (DPF)
  - DPF distributes data to nodes at load, then the data is queried repeatedly – favors response time over throughput
  - MapReduce distributes the data at query run time – favors throughput over response time

"When a farmers buys another 40 acres, he doesn't get a bigger ox, he gets a second ox." Admiral Grace Hopper

37

THE FILLMORE GROUP
Relational Database Solutions

Audience Use Case Open Discussion

- What data management requirement do you need to support?
- Which technology of the four surveyed would you choose and why?

38

THE FILLMORE GROUP
Relational Database Solutions

Contacts

- Kim May
  - kim.may@thefillmoregroup.com
  - twitter.com/KimMayTFG
  - www.linkedin.com/pub/kim-may/4/462/84
- Frank Fillmore
  - frank.fillmore@thefillmoregroup.com
  - twitter.com/ffillmorejr
  - www.linkedin.com/pub/frank-fillmore/6/597/9a6/
  - tinyurl.com/ChannelDB2
  - Flipboard for iPad, iPhone, Android: "BigData"

Premier Business Partner IBM

39

THE FILLMORE GROUP
Relational Database Solutions

Resources

- IBM DB2 Express-C
- How to Handle the "V" for Volume in Big Data on DB2 for z/OS
- DB2 10.5 with BLU Acceleration Flashbook
- Optimizing analytic workloads using DB2 10.5 with BLU Acceleration
- BigData – Hadoop
- IBM InfoSphere BigInsights

Premier Business Partner IBM

40

Thank You

Your feedback is important!