In-memory Column Databases

Trend & Concepts in the Software Industry -
Lecture June 2011

Memory Hierarchy

Lower Price / Higher Latency

CPU Registers

CPU Caches

Main Memory

Flash

Hard Disk

Higher Performance
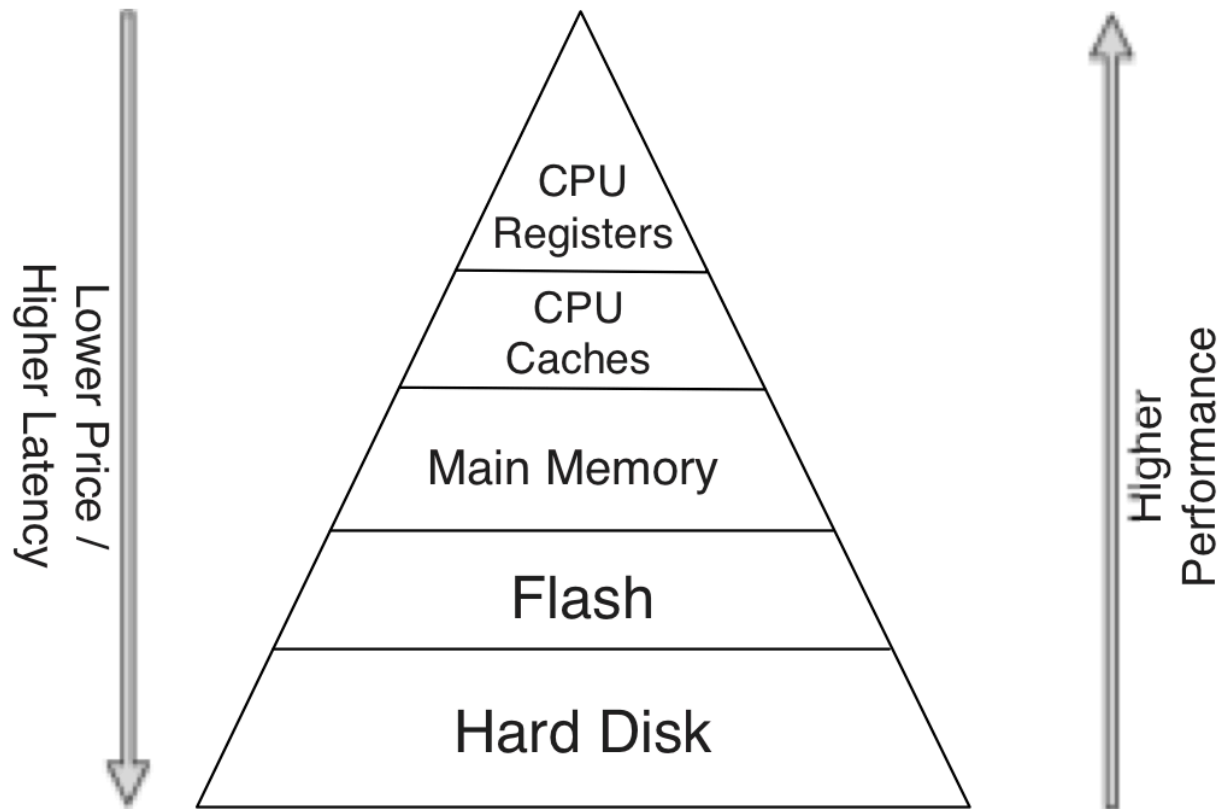
Row and Column-oriented Data Layout

Row-Oriented Database

Column-Oriented Database
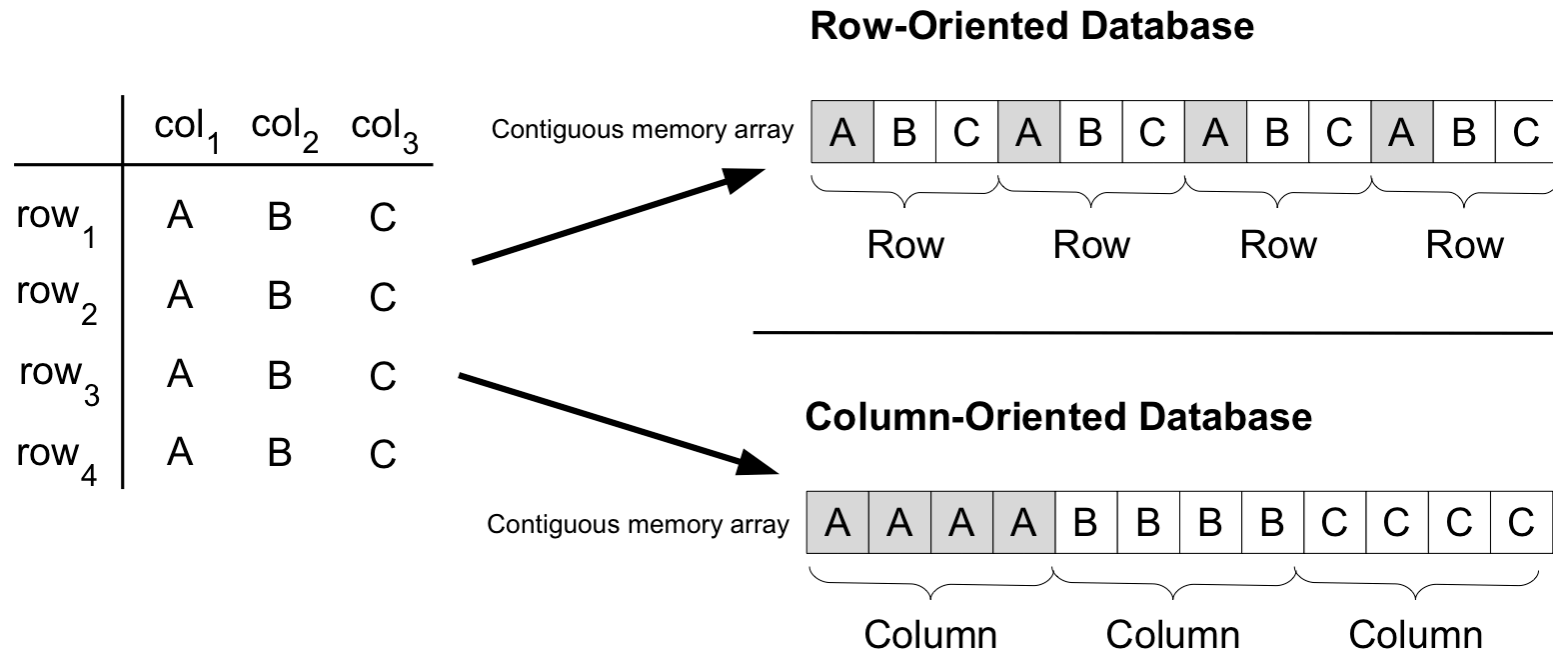
Read and Write-optimized Storage

Write/Change   Read

Compaction Process

Write-optimized Storage

Read-optimized Storage
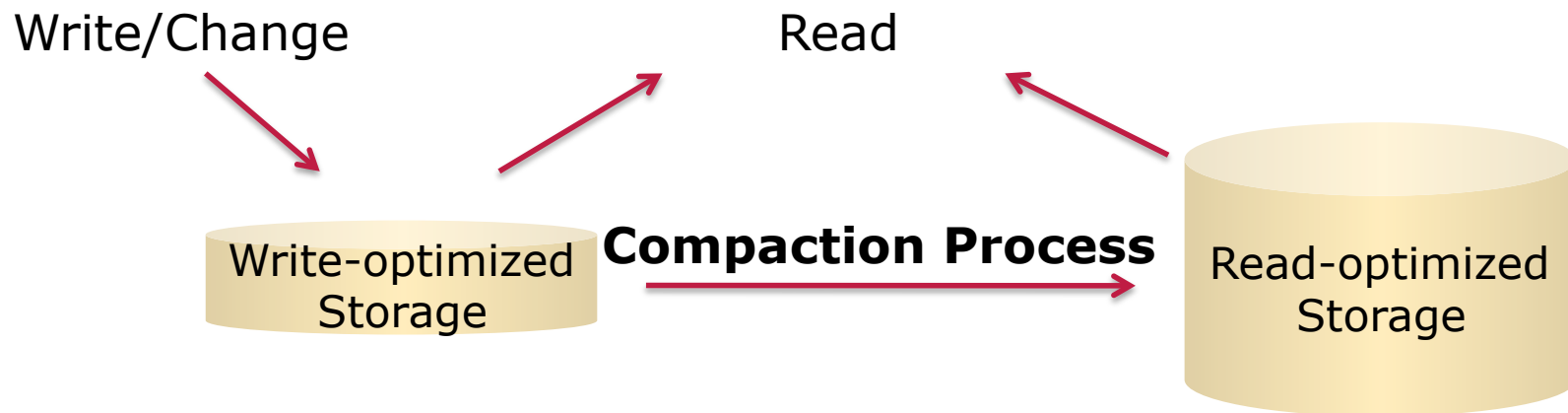
Compression Techniques

	Few distinct values	Many distinct values
Ordered	Run-length encoding	Delta representation
Unordered	Bit vector encoding	Dictionary compression

Dictionary Compression

Document Number	Document Date	Sold-To Party	Order Value
95769214	2009-10-01	584	10.24
95769215	2009-10-01	1215	124.35
95779216	2009-10-21	584	47.11
95779217	2009-10-21	454	21.20

Document Number	
0	95769214
1	95769215
2	95779216
3	95779217

Order Value	
0	10.24
1	21.20
2	47.11
3	124.35

Document Number	Document Date	Sold-To Party	Order Value
0	0	1	0
1	0	2	3
2	1	1	2
3	1	0	1

Document Date	
0	2009-10-01
1	2009-10-21

Sold-To Party	
0	454
1	584
2	1215

Operating on Compressed Data