Designing Columnstore Indexes for Analytic Queries

Gail Shaw
TECHNICAL LEAD

@SQLintheWild   http://sqlinthewild.co.za

Agenda

Introducing Columnstore indexes

Columnstore architecture

Clustered and nonclustered Columnstore indexes
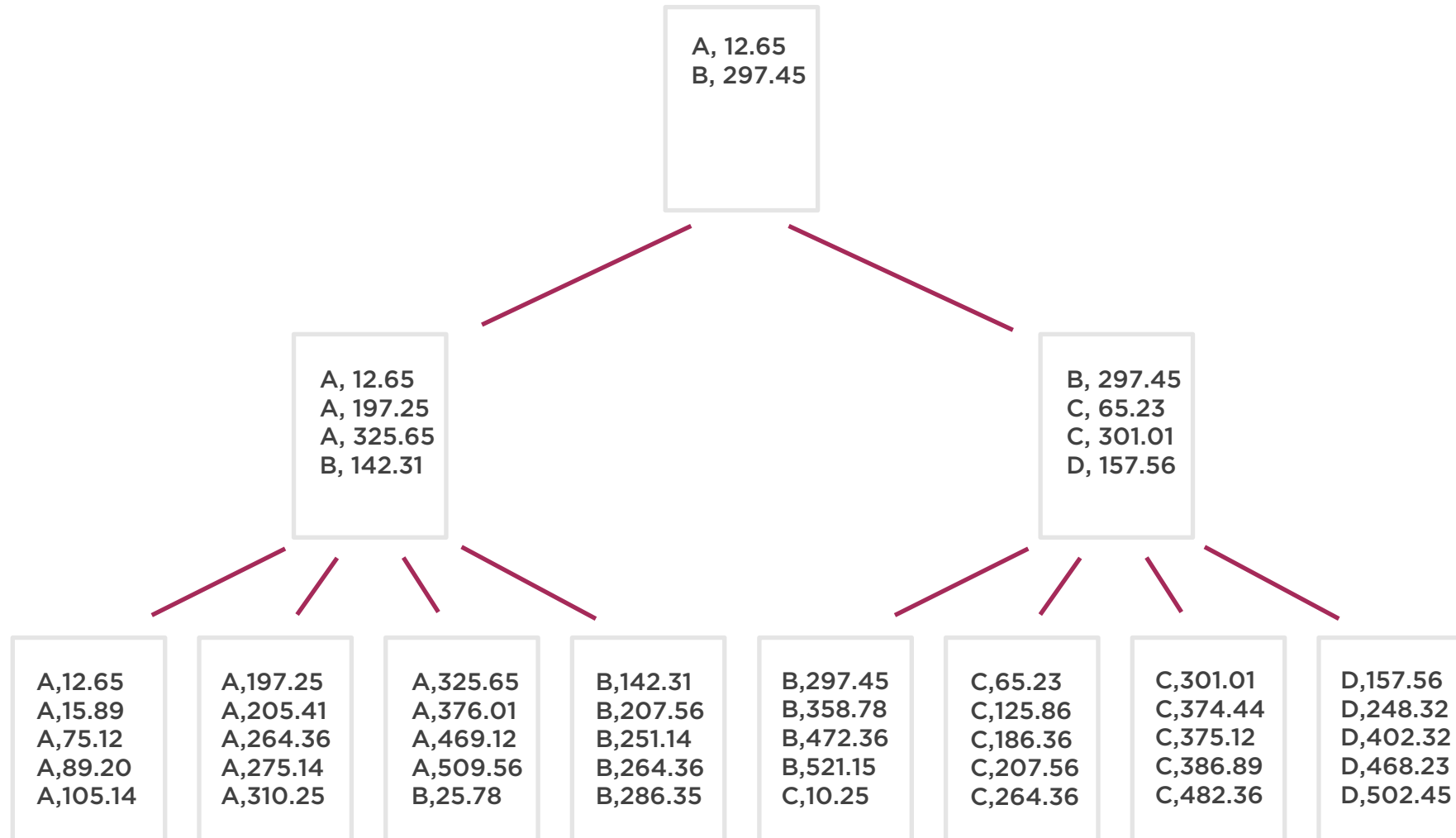
When to use Columnstore indexes

Columnstore Indexes

Indexes where the data for a column is stored together

Highly compressed

Good for analytics-type queries

Rowstore Index Architecture

Columnstore Index Architecture

**CREATE COLUMNSTORE INDEX
ON Transactions
(Amount, TransactionType,
IsPriority, ShipmentID)**

Amount	TransactionType	IsPriority	ShipmentID
10.25	C	1	16
12.65	A	0	28
15.89	A	0	46
25.78	B	1	12
65.23	C	1	15
75.12	D	0	90
80.23	C	1	8
81.24	C	0	7
82.35	A	1	42
89.20	A	1	17
105.14	A	1	36
125.86	C	0	8
142.31	B	0	96
157.56	D	1	74
186.36	C	1	23
197.25	A	1	185
205.41	A	0	25
207.41	B	1	14
207.56	C	0	89
248.32	D	1	86
251.14	B	0	112
264.36	A	1	31
264.36	B	0	75

Columnstore Index Architecture

10.25	C	1	16
12.65	A	0	28
15.89	A	0	46
25.78	B	1	12
65.23	C	1	15
75.12	D	0	90
80.23	C	1	8
81.24	C	0	7
82.35	A	1	42
89.20	A	1	17
105.14	A	1	36
125.86	C	0	8

Rowgroup ←

142.31	B	0	96
157.56	D	0	74
186.36	C	1	23
197.25	A	1	185
205.41	A	0	25
207.41	B	1	14
207.56	C	0	89
248.32	D	1	86
251.14	B	0	112
264.36	A	1	31
264.36	B	0	75

Segment ←

Compressed
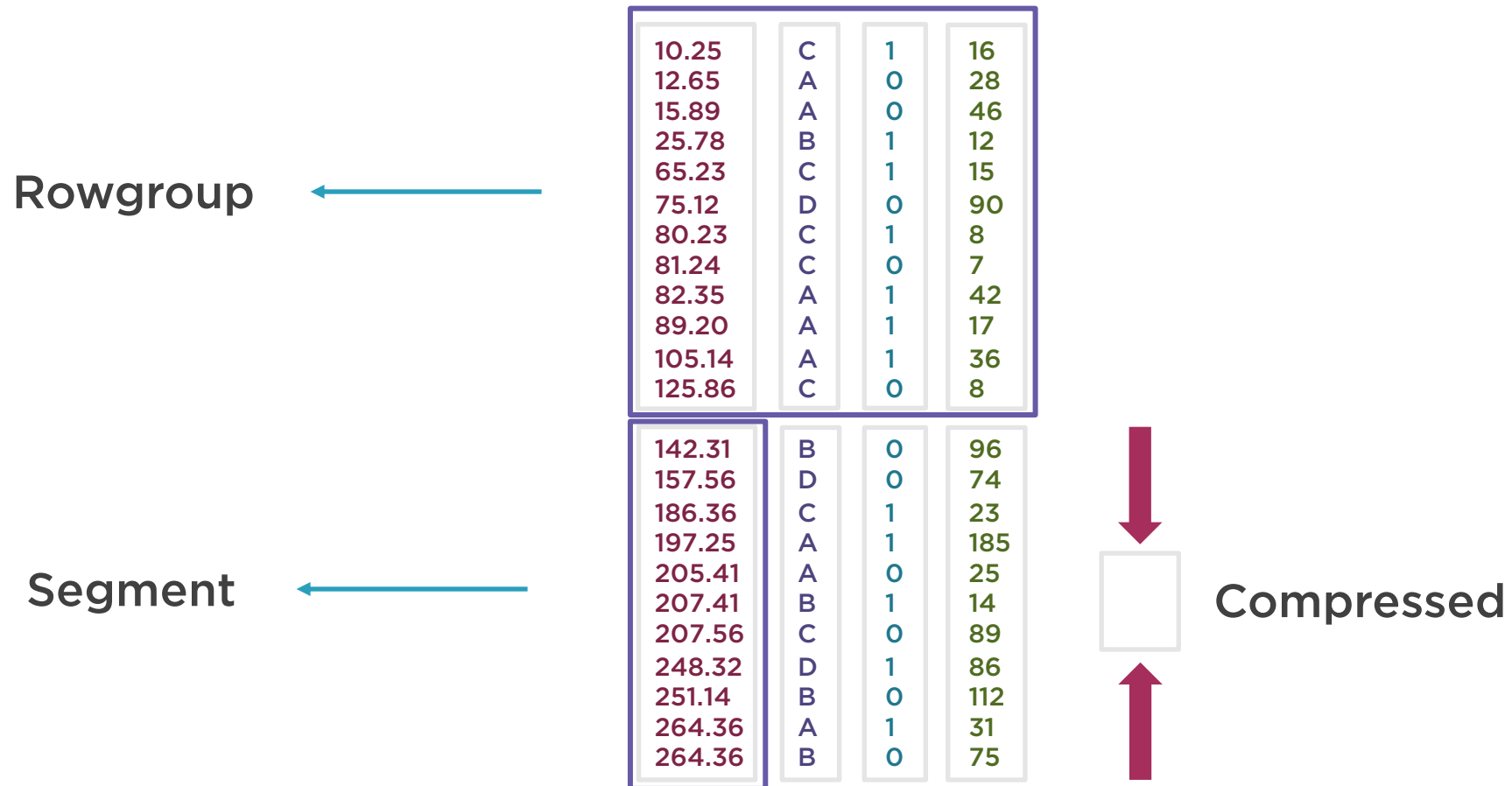
Modifying Data in a ColumnStore Index
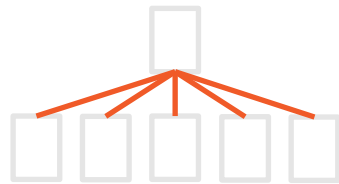
Deleted List
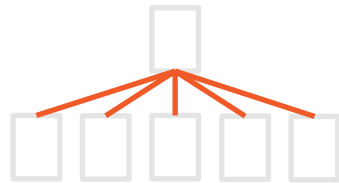
Deltastore

10.25	C	1	16
12.65	A	0	28
15.89	A	0	46
25.78	B	1	12
65.23	C	1	15
75.12	D	0	90
80.23	C	1	8
81.24	C	0	7
82.35	A	1	42
89.20	A	1	17
105.14	A	1	36
125.86	C	0	8

142.31	B	0	96
157.56	D	0	74
186.36	C	1	23
197.25	A	1	185
205.41	A	0	25
207.41	B	1	14
207.56	C	0	89
248.32	D	1	86
251.14	B	0	112
264.36	A	1	31
264.36	B	0	75

Types of Columnstore Indexes

Clustered columnstore

Organises the table, consists of all the columns in the table

Nonclustered columnstore

Secondary structure, contains just the columns specified

Demo

Creating and examining columnstore Indexes

Columnstore Guidelines

Clustered columnstore indexes on large fact tables

Can add nonclustered rowstore indexes to help with searches

Nonclustered columnstore indexes can help with analytics in OLTP systems

Considerations for Columnstore Indexes

Not ideal for systems with lots of single-row modifications

Not efficient on small tables. Rowgroups should ideally have ~ 1 million rows

Compression works best on columns with repeated patterns in the data

Can't be used for seek operations

Can't be used to enforce uniqueness

Order of Columns

Not of major importance with a Columnstore

No 'left-based subset of the key'

For nonclustered Columnstores, ensure they contain the columns to be aggregated and grouped

Demo

Query Performance with Columnstore Indexes

Summary

Columnstore index architecture

Guidelines for creating columnstore indexes