Columnstore Technology Improvements in SQL Server 2016

Subtle Subtitle AlwaysOn

Niko Neugebauer

Our Sponsors

Niko Neugebauer

Microsoft Data Platform Professional

OH22 (http://www.oh22.net)

SQL Server MVP

Founder of a couple of Portuguese PASS Chapters (SQLPort, BITuga, Porto.Data)

CISL – **Open Source Columnstore Indexes Script Library**
(https://github.com/NikoNeugebauer/CISL)

MOSL – **Open Source Memory Optimised Script Library**
(https://github.com/NikoNeugebauer/MOSL)

Blog: http://www.nikoport.com

Twitter: @NikoNeugebauer

LinkedIn: http://pt.linkedin.com/in/webcaravela

SQL Konferenz 2017

http://www.sqlkonferenz.de

Our Agenda

- Operational Analytics
- Something NEW
- Data Warehousing
- High Availability
- Batch Mode
- Data Loading Improvements
- Performance Improvements
- Change Tracking
- Maintenance Improvements
- Monitoring Improvements (DMV, Extended Events, Performance Counters)
- Configuration Improvement
- New Trace Flags

Analytics

- Operational

Analytics

What is it all about:

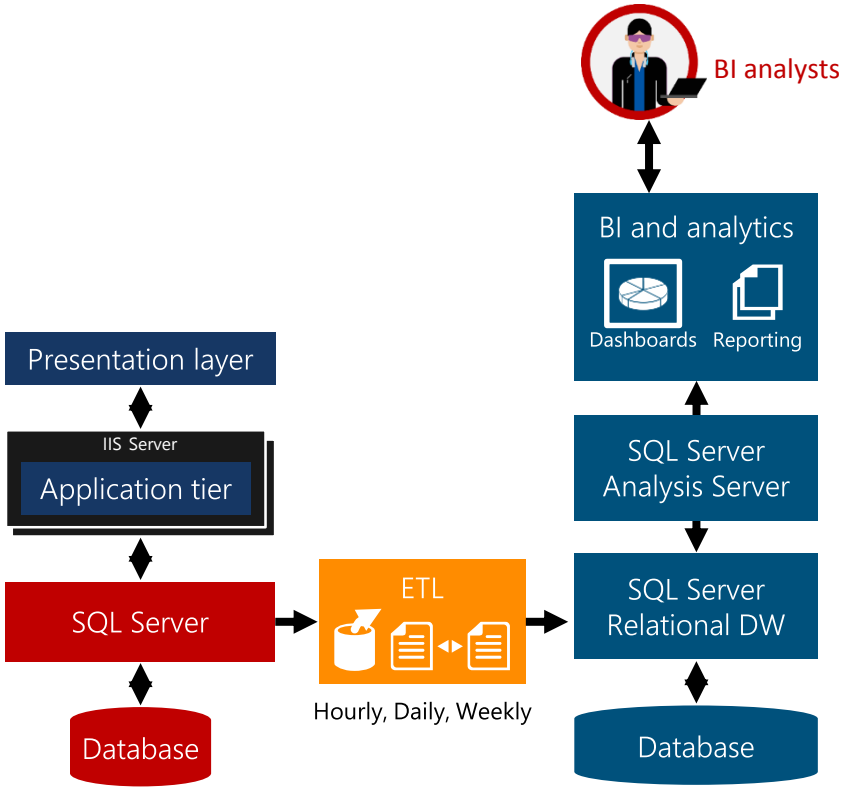
- Predictive Analytics, Prescriptive Analytics, Business Analytics, Machine Learning, Data Mining

- But technically it all goes down to the one thing we treasure the most: the DATA

Analytics Faces

- **Analytics** – is the process of discovery & communication of meaningful patterns in data.
- **Reporting** - extraction of the aggregated information for further analysis.
- **Querying** - data extraction process

Traditional operational / analytics architecture



Key issues

Complex implementation

Requires two servers (capital expenditures and operational expenditures)

Data latency in analytics

More businesses demand; requires real-time analytics

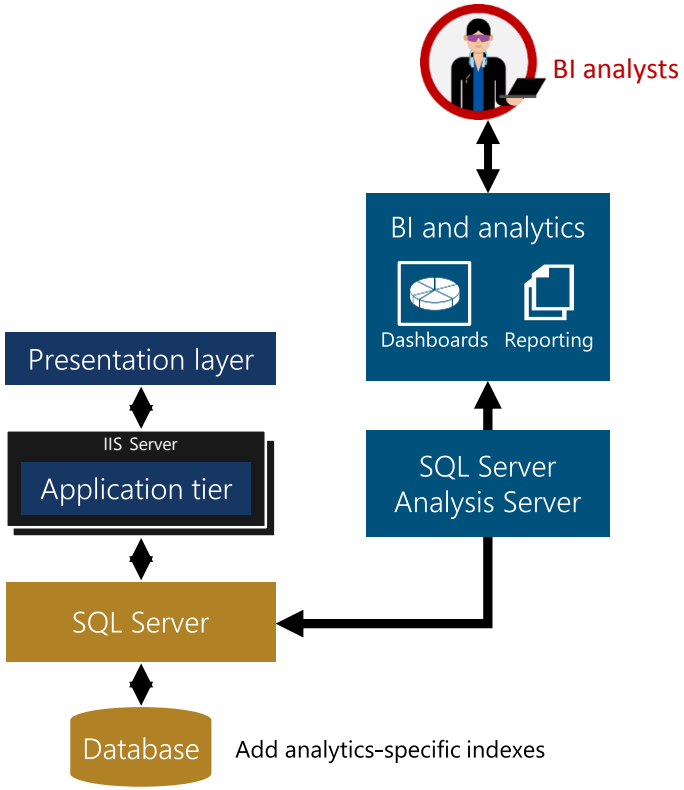
Traditional Operational Analytics Problems:

- Costs (€)
- Integration Problems (data types, constraints, network problems, etc)
- Delay for getting the actual data (minutes, hours, days, weeks)

Modern Operational Analytics notes:

- Nothing substitutes analytics queries performance,
  possible using schemas customized (Star/Snowflake) and/or pre-aggregated cubes

Minimizing data latency for analytics



Benefits

No data latency

No ETL

No separate data warehouse

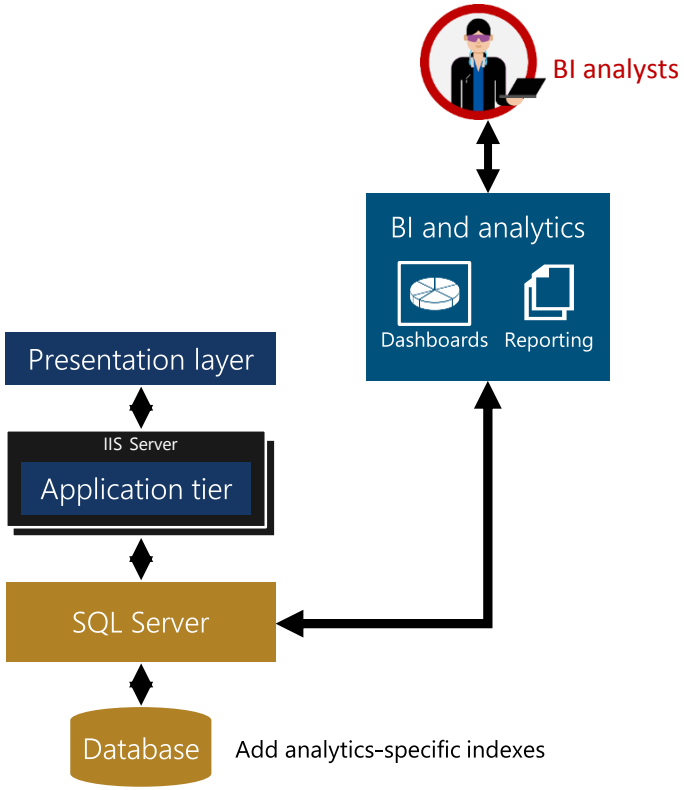
Challenges

Analytics queries are resource intensive and can cause blocking

Minimizing impact on operational workloads

Sub-optimal execution of analytics on relational schema

Optimizing data latency for analytics



What is operational analytics and what does it mean to you?

Operational analytics with disk-based tables

Operational analytics with In-Memory OLTP

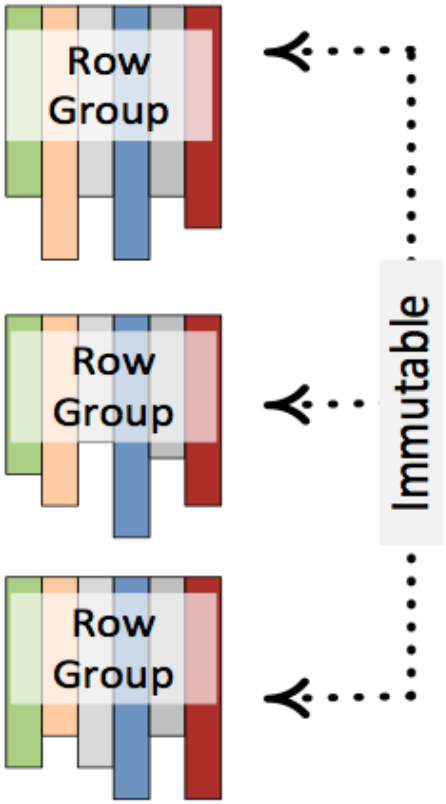
Operational Analytics: SQL Server 2016

- There are now 2 types of Operational Analytics:
- Operational Analytics for Disk-Based Tables
- Operational Analytics for InMemory

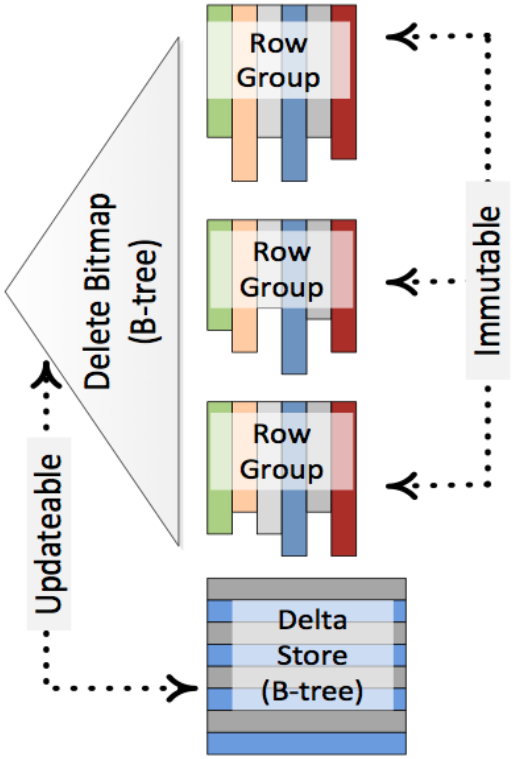
Columnstore

Updatable? Updatable!

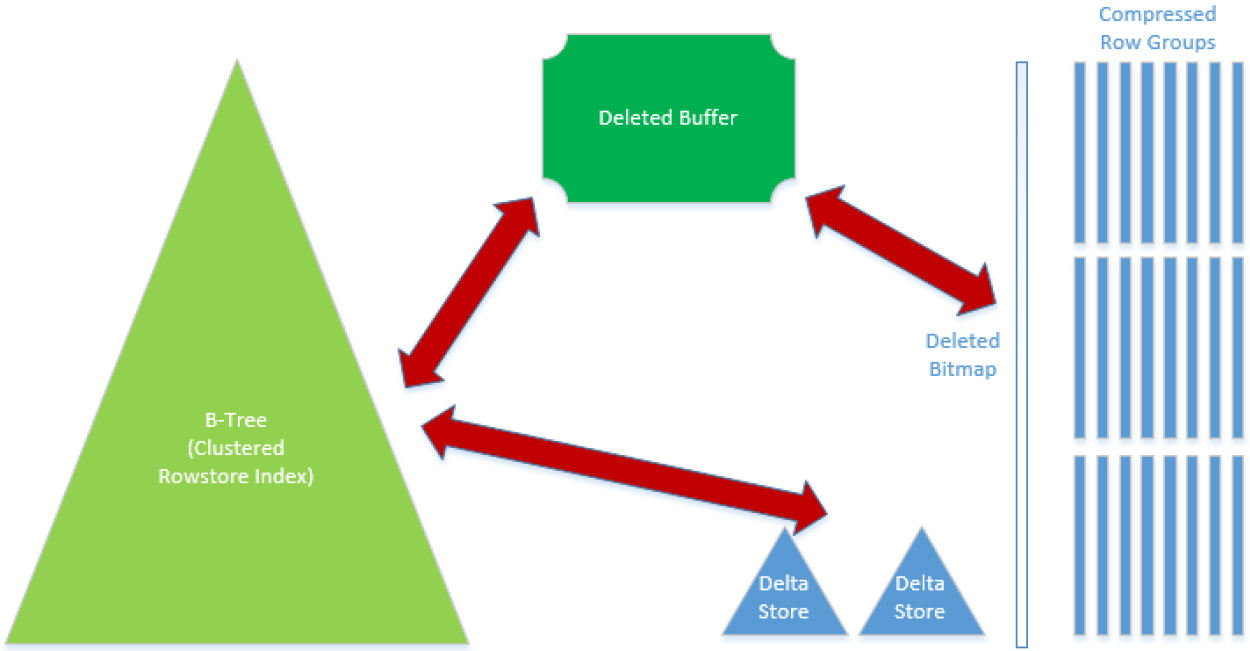
NonClustered Columnstore 2012/2014



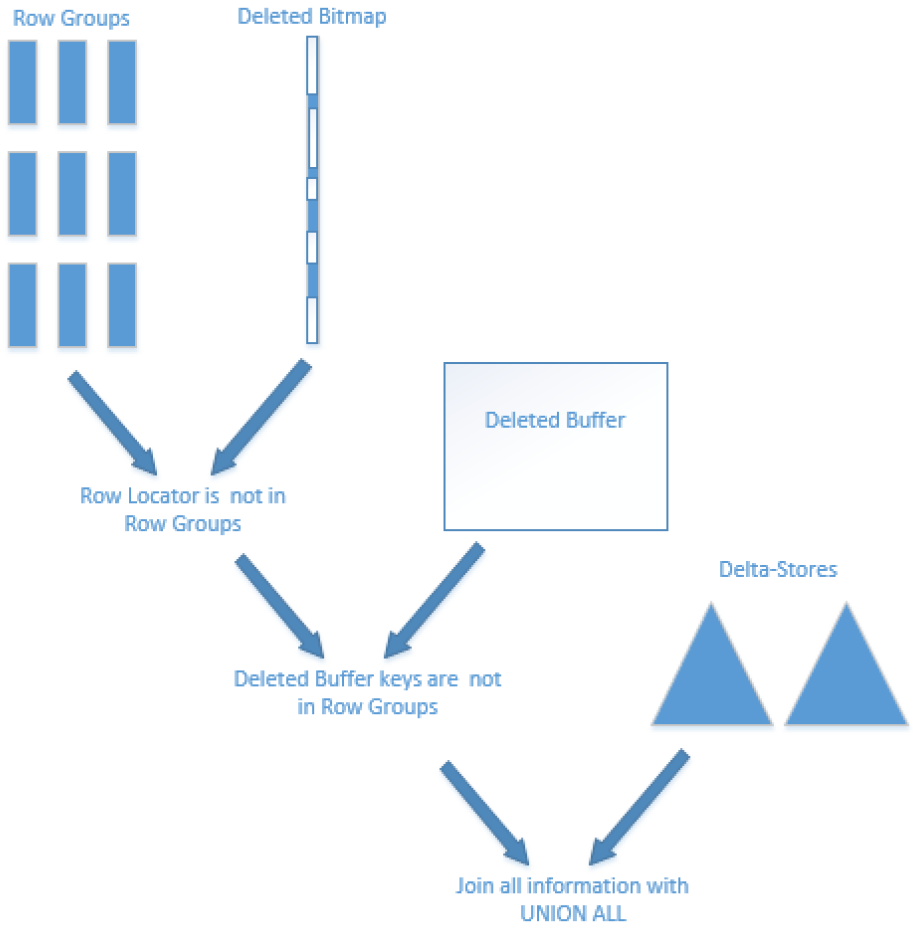
Updatable Columnstore SQL Server 2014



Nonclustered Columnstore 2016 (aka Disk-Based Operational Analytics)



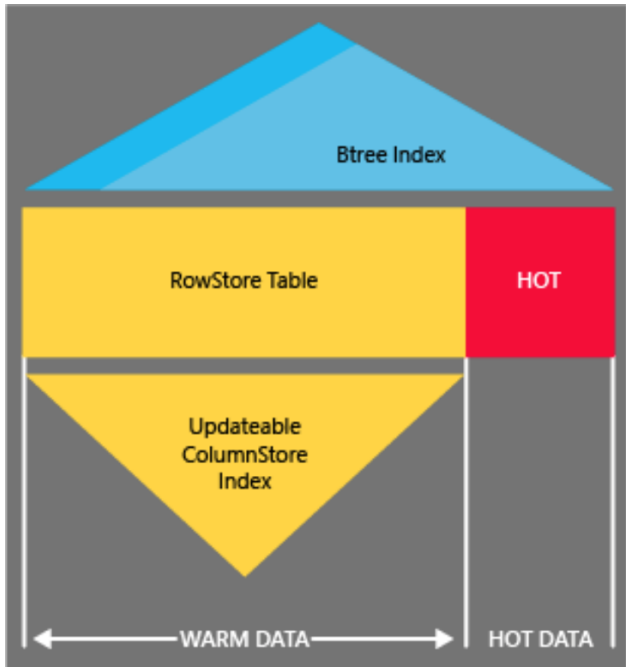
Operational Analytics Query Processing



Operational Analytics

- Demo

Filtered Operational Analytics

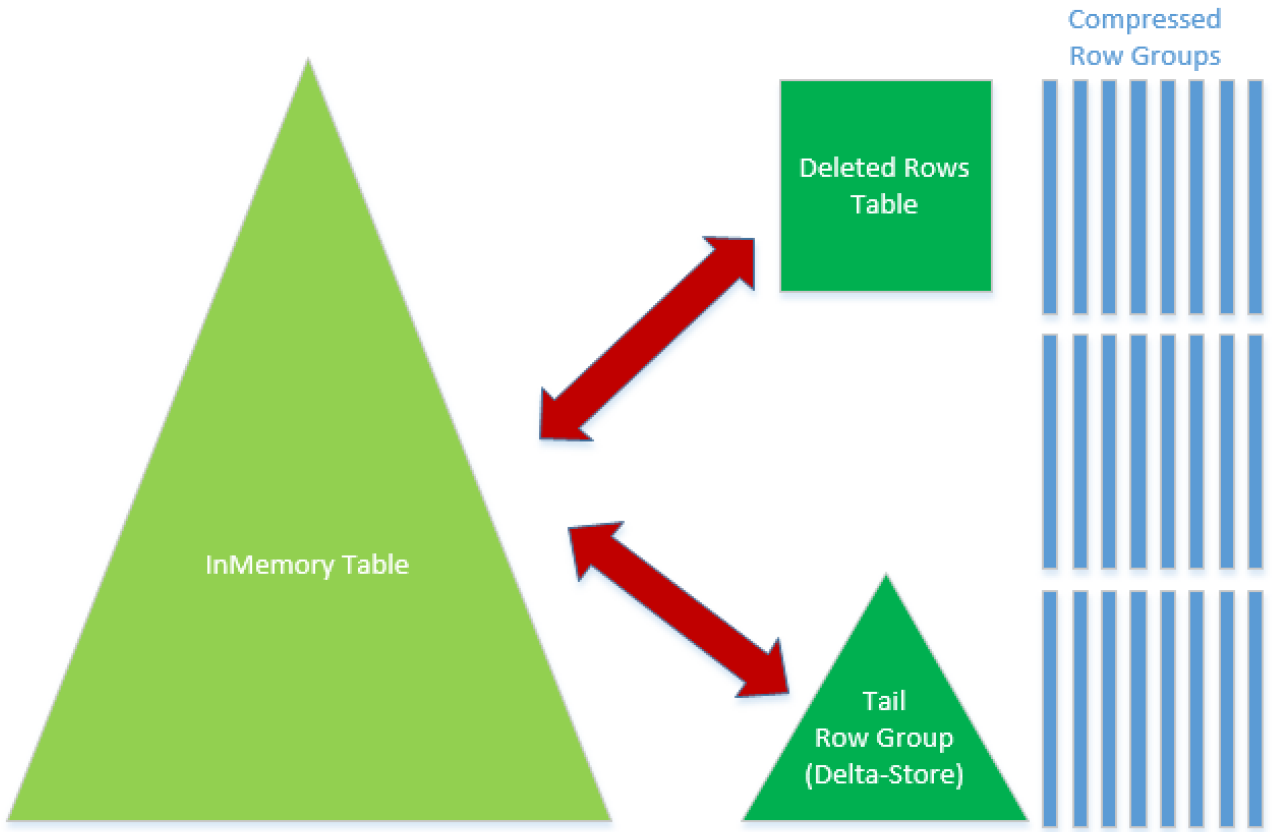


- Create a **filtered** Nonclustered Columnstore Index!

Operational Analytics

- Demo with a Filtered Index

In-Memory Operational Analytics



In-Memory Operational Analytics

- **sys.sp_memory_optimized_cs_migration** – compresses data from InMemory OLTP Table into In-Memory Columnstore

Operational Analytics

- Demo Memory-Optimized

The new thing that Niko twitted yesterday ☺

FINALLY!

Clustered Columnstore (DWH)

- Nonclustered Indexes
- Primary & Foreign Keys
- Nonclustered Secondary Rowstore Locking

Further Columnstore Improvements

- Data Loading Improvements
- High Availability
- Batch Mode
- Performance Improvements
- Change Tracking
- Maintenance Improvements
- Monitoring Improvements (DMV, Extended Events, Performance Counters)

Data Loading Improvements

- Parallel Data Loading finally enabled
- SIMD support
- Delta-Stores are not Page-Compressed!!!
- Under NCCI Delta-Stores are automatically increasing their max size with each iteration (1,2,4,8,16,32 Million Rows)

Data Loading Landing Improvements

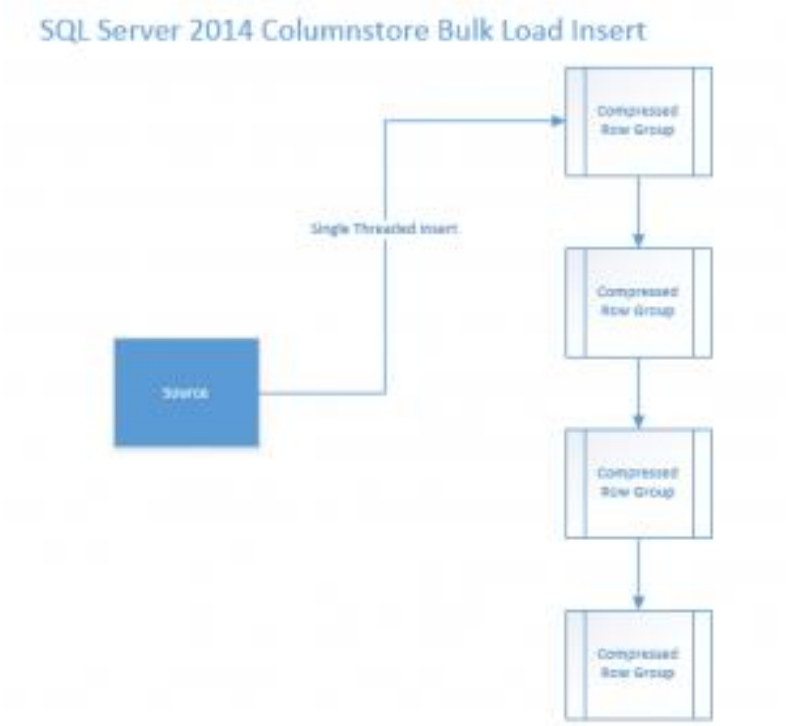
Compression Delay:

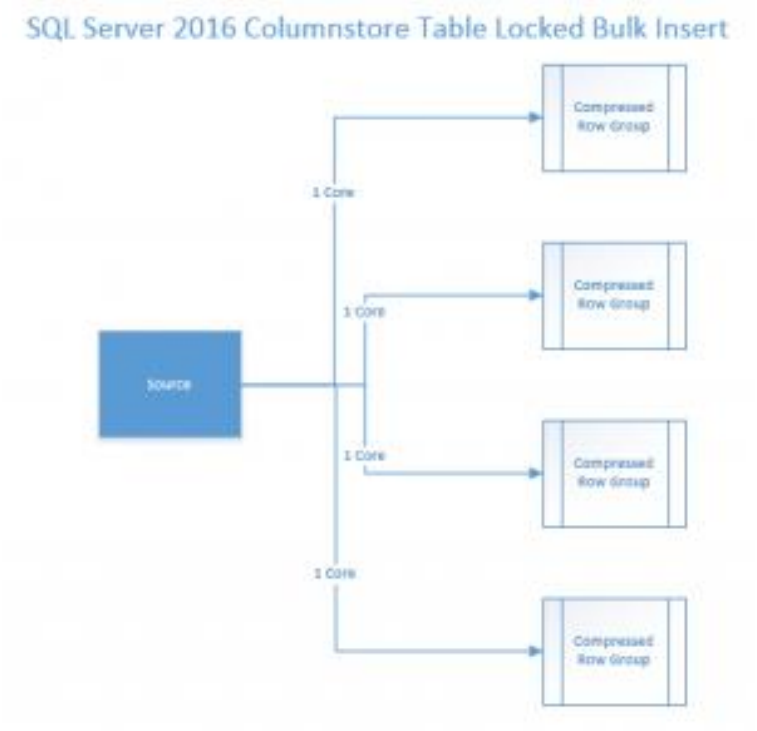
```
alter index PK_NCCi_test_inline
    on dbo.ncci_test_inline
    set (COMPRESSION_DELAY = 60 Minutes);
```

SSIS 2016 Improvements

- AutoAdjustBufferSize - True
- *Warning: It's not enough, you will need to set up the maximum number of rows per buffer, otherwise your performance will be extremely slow!*
- **DefaultBufferMaxRows = 1048576**

Parallel Data Loading





High Availability

- Readable Secondaries for Availability Groups – through
- Snapshot &
- Read Committed Snapshot
- Isolation Levels support

Batch Mode Improvements

- Batch Mode support for 1 core execution plan operators
- Batch Mode support for the Sort operator
- Batch Mode support for the Multiple Distinct Count operations
- Batch Mode support for the Left Anti-Semi Join operators

Batch Mode

- Demo

Batch Mode for Windowing Functions

Window Batch Mode

- Demo

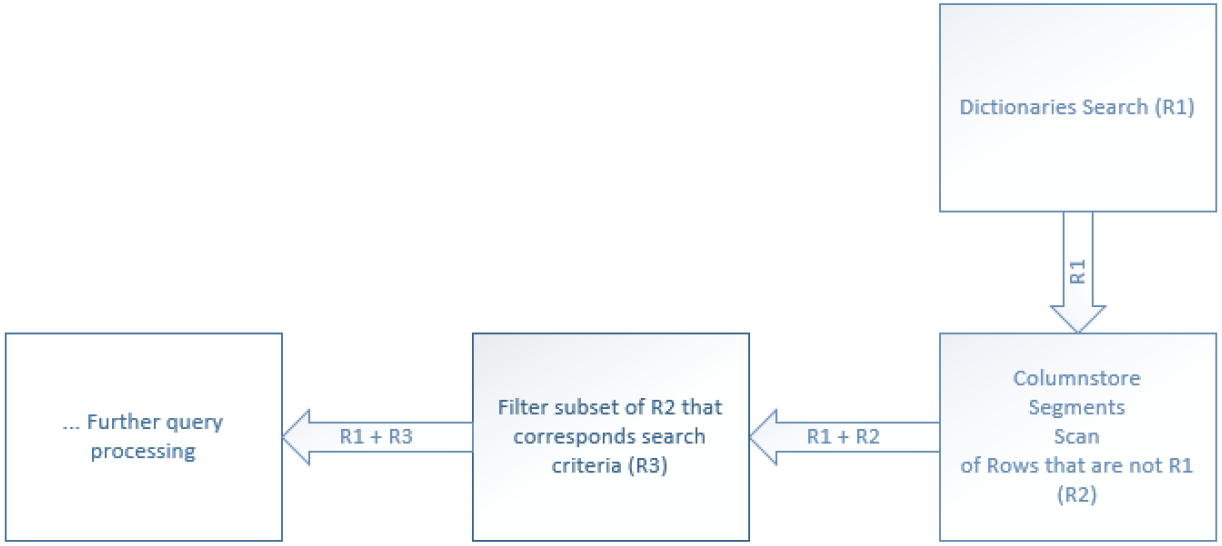
Further Batch Improvements

- Simple Aggregate Predicate Pushdown
- String Predicate Pushdown for Index Scan operator in Batch Mode

Batch

- Aggregation Demo

String Predicate Pushdown



Change Tracking / CDC in SQL 2016

Columnstore	Change Tracking	Change Data Capture	Temporal (2016+)
Clustered	**No**	**No**	Yes
Nonclustered	Yes	Yes	Yes
Memory-Optimized	**No**	**No**	Yes

Maintenance Improvements

- Better ALTER INDEX … REORGANIZE (removes deleted rows, less memory pressure)

New DMVs:

- sys.dm_column_store_object_pool
- sys.dm_db_column_store_row_group_physical_stats
- sys.dm_db_column_store_row_group_operational_stats
- sys.internal_partitions

Enhanced DMVs:

- sys.dm_db_index_operational_stats
- sys.dm_db_index_physical_stats

Configuration Improvement

- Large Pages (Trace Flag 834)
  should be supported in SQL Server 2016,
  well, at least the Connect Item for it is closed as fixed. ☺

New Trace Flags

Trace Flag	Description
9347	Disables batch mode sort operator
9349	Disables batch mode top sort operator
9358	Disable batch mode sort operations in a complex parallel query in SQL Server 2016
9389	Enables dynamic memory grant for batch mode operators
10204	Disables merge/recompress during columnstore index reorganization

Resources:

My Columnstore Blogpost Series (70+):

http://www.nikoport.com/columnstore

CISL – Open Source Columnstore Library:

https://github.com/NikoNeugebauer/CISL

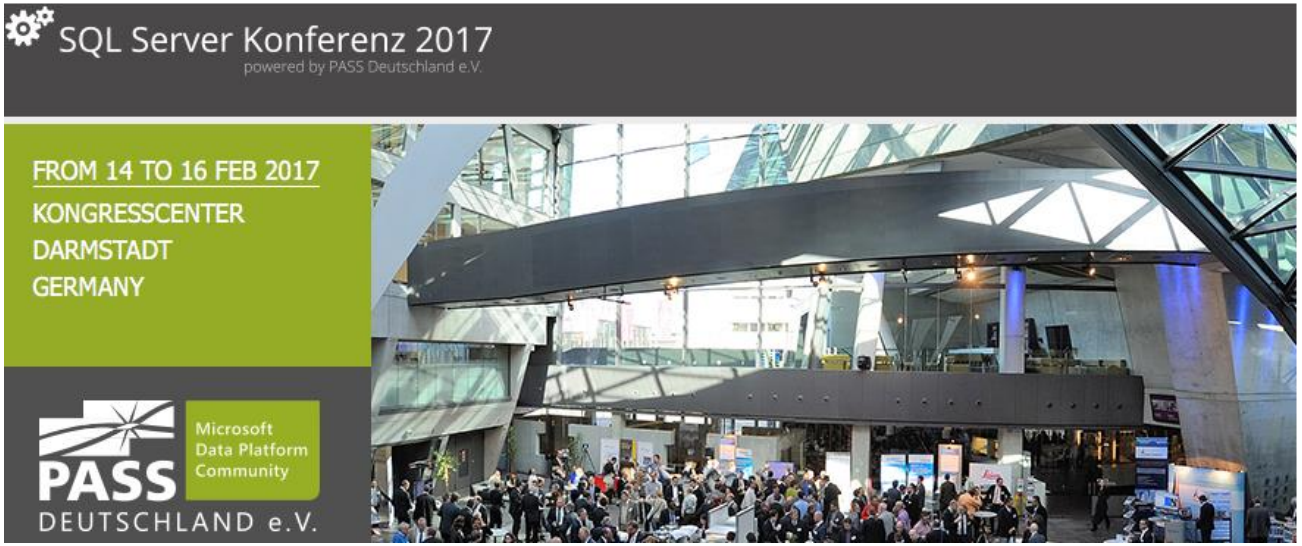
Title Goes Here, 36 pt.

- Main Text / Bullets Here, Gray, 30 pt.
- Main Text / Bullets Here, Gray, 30 pt.
  - Bullet Points, Line 2, 26 pt.
    - Bullet Points, Line 3, 22 pt.
      - Bullet Points, Line 4, 20 pt.

Save the Dates!

PASS Camp 2016 - 06. to 09. December 2016

SQL Konferenz 2017 - 14. to 16. February 2017



How did you like it?

Please give feedback

- to the event:
- http://goo.gl/5tStsd

- to me as a speaker:
- http://goo.gl/TgViTX