SQL Server 2016 – Die Evolution der In-Memory Technologien

Andreas Wolter
Inhaber: Sarpedon Quality Lab
Database Architect | MCSM, MCM, MVP

1

Andreas Wolter

Consultant, Trainer & Speaker
Microsoft Certified Master SQL Server 2008
+ Solutions Master Data Platform (SQL Server 2012)

- Datawarehouse & OLTP-System Architecture
- Performance Tuning
- Security

Email: a.wolter@Sarpedon.de
Web: www.andreas-wolter.com
Blog: www.insidesql.org/blogs/andreaswolter/
Facebook: www.facebook.com/SarpedonQualityLab
Xing: www.xing.com/profile/Andreas_Wolter2
Twitter: @AndreasWolter

Microsoft Certified Solutions Master Data Platform
Microsoft Certified Master SQL Server® 2008
SQL Server

Microsoft Certified IT Professional Business Intelligence Developer 2008
SSAS, SSRS, SSIS

MVP Microsoft® Most Valuable Professional

3

Die nächsten SQL Server Master-Classes ab Juni 2016

SQL SERVER MASTER-CLASS
by SARPEDON QUALITY LAB

- Workshop **Hochverfügbarkeit mit SQL Server Always On**
  - Windows Server Failover Clustering
  - SQL Server Clustered Instances
  - SQL Server Availability Groups
  - Logshipping
  - Frankfurt, 30.5.-1.6. *

15% Rabatt-Code für PASSler:
SQLPASS_de16

www.SQL-Server-Master-Class.com

- **SQL Server Performance-Analyse** Workshop
  - Frankfurt, 29.6.-1.7. *

*Termine noch unter Vorbehalt

4

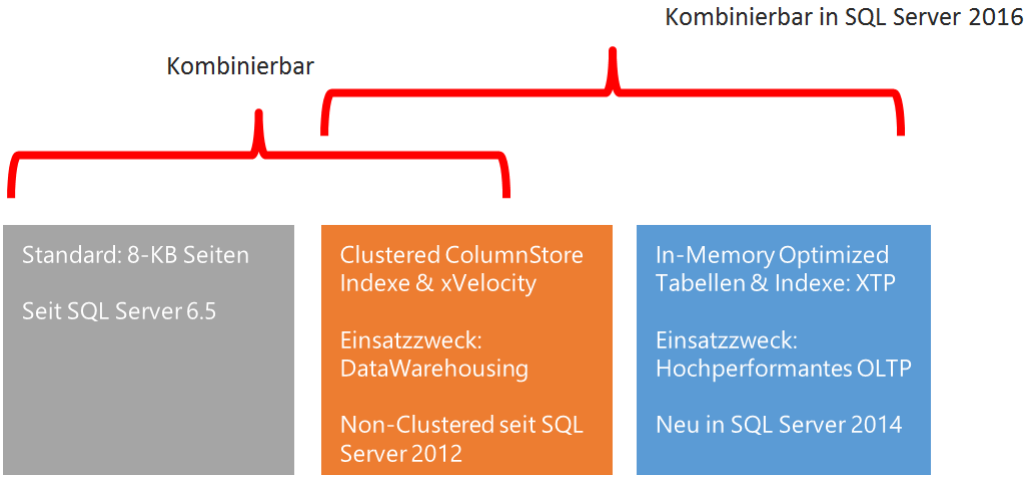
Agenda

SARPEDON QUALITY LAB

- Columnstore
- In Memory
- Real-time operational Analytics

5

SQL Server Engines seit 2014



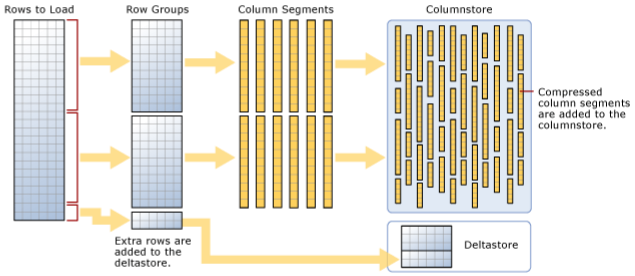
6

COLUMNSTORE

7

Columnstore, Laden via Rowgroup & Deltagroup

- Ein **Segment** einhält Werte einer Spalte
- Segmente für dieselbe Zeilenmenge bilden eine **Row Group**
  - Für ca. 1 Mio Zeilen
- Segmente sind komprimiert
- Jedes Segment wird in einer separaten BLOB-Struktur gespeichert



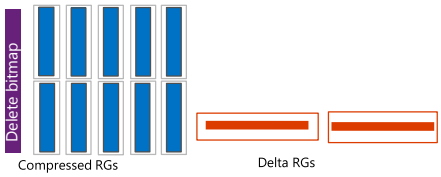
- Columnstore Indexes
  - http://msdn.microsoft.com/en-us/library/gg492088%28v=sql.120%29.aspx

8

Was ist daran neu?

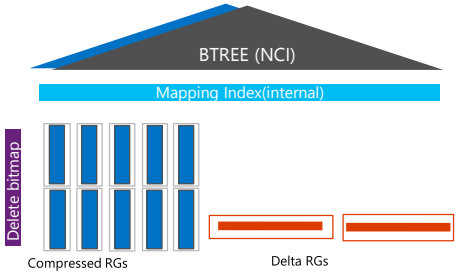
9

Columnstore Index: Clustered (CCI)



SQL 2014

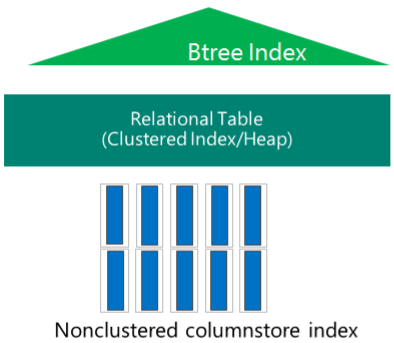
- Master copy of the data (10 x compression)
- Only 1 index supported. Simplified analytics
- No PK/FK constraints. Uniqueness can be enforced through Materialized Views
- Locking granularity for Update/Delete at rowgroup level
- DDL – ALTER, REBUILD, REORGANIZE



SQL 2016

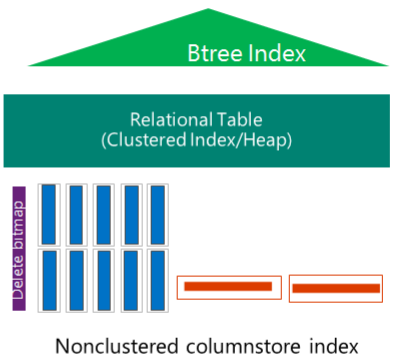
- Additional Btree indexes for efficient equality and short-range searches and PK/FK constraints.
- Can be on Memory optimized Table
- Locking granularity at row level using NCI index path
- support snapshot and read-committed snapshot isolation
- DDL – ALTER, REBUILD, REORGANIZE

Columnstore Index: NonClustered (NCCI) Updateable



SQL 2014

- Introduced in SQL Server 2012
- NCCI is read-only (No delete bitmap and delta store)
- Additionally to NCI(s) and CI or Heap



SQL 2016

- It is updateable!
- You can mix OLTP and Analytics workload
- You can create Filtered NCCI

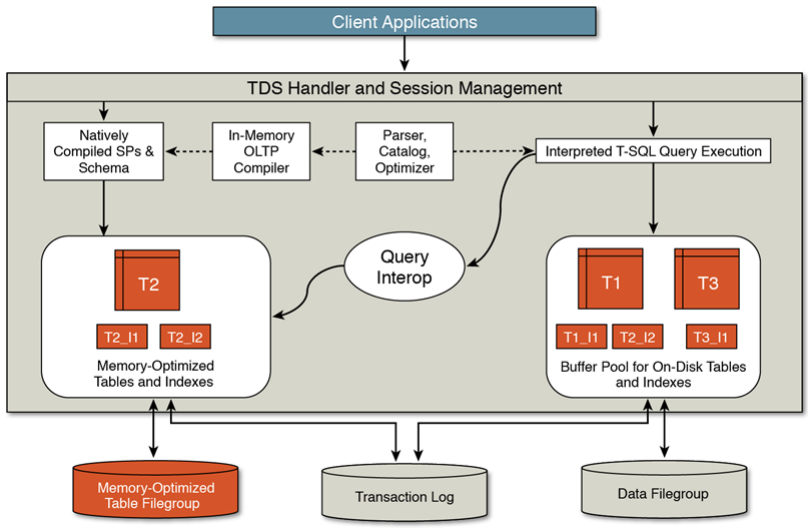
SARPEDON QUALITY LAB

IN-MEMORY OLTP

12

SARPEDON QUALITY LAB

Integration der XTP Engine



13

In-Memory OLTP Concurrency Control

SARPEDON QUALITY LAB

Multi-version	Multi-version data store	Snapshot-based transaction isolation No TempDB
Optimistic	No locks, no latches, minimal context switches No blocking	Conflict detection to ensure isolation No deadlocks

14

SARPEDON QUALITY LAB

Populating Data/Delta files

- Data file pre-allocated
- switches to new data file when the current file is full
- Transactions do not span data files
- Once a data file is closed, it becomes read-only
- Row deletes are tracked in delta file
- Files are append only

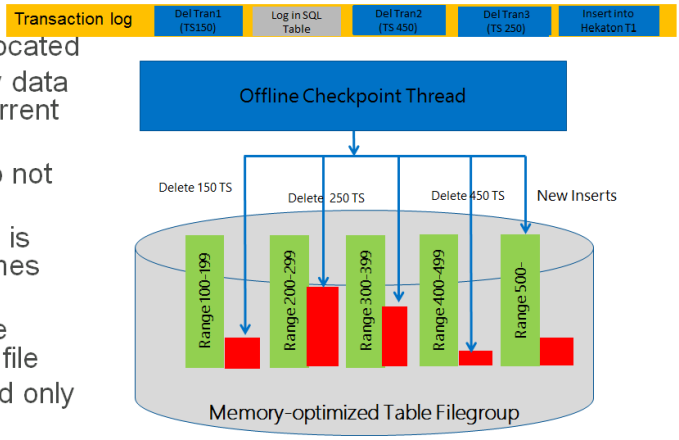


Image: Sunil Agarwal, TechEd 2013

Data file with rows generated in timestamp range

IDs of Deleted Rows (height indicates % deleted)

15

Memory-Optimized Indexes

SARPEDON QUALITY LAB

- Exist only in memory
  - Rebuilt on database startup
    - -> RTO (!)
- Do not contain data rows
  - Indexes contain memory pointers to the data rows
    - All indexes are covering
    - No duplication of data

Nonclustered Indexes

No need to specify any options

Optimized for range lookups and ordered scans
Support search on leading columns of the index key
Support point-lookups, but hash indexes are much faster

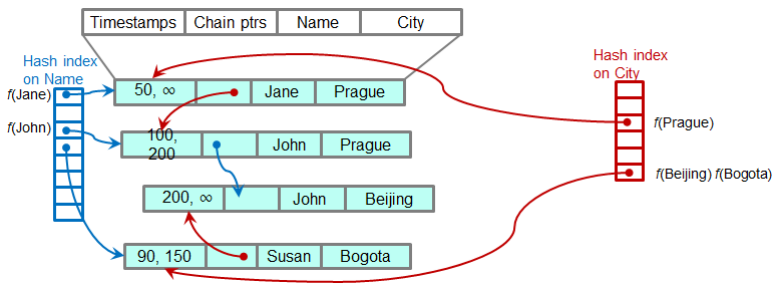
Hash Indexes

Need to specify bucket_count – usually 1-2X the number of unique index keys
Optimized for point-lookups
Cannot be used for search on subset of the index key
Cannot be used for range lookups (search on inequality predicate) or ordered scans (for ORDER BY)

16

Multiple Hash Indexes & Versioning

SARPEDON QUALITY LAB



17

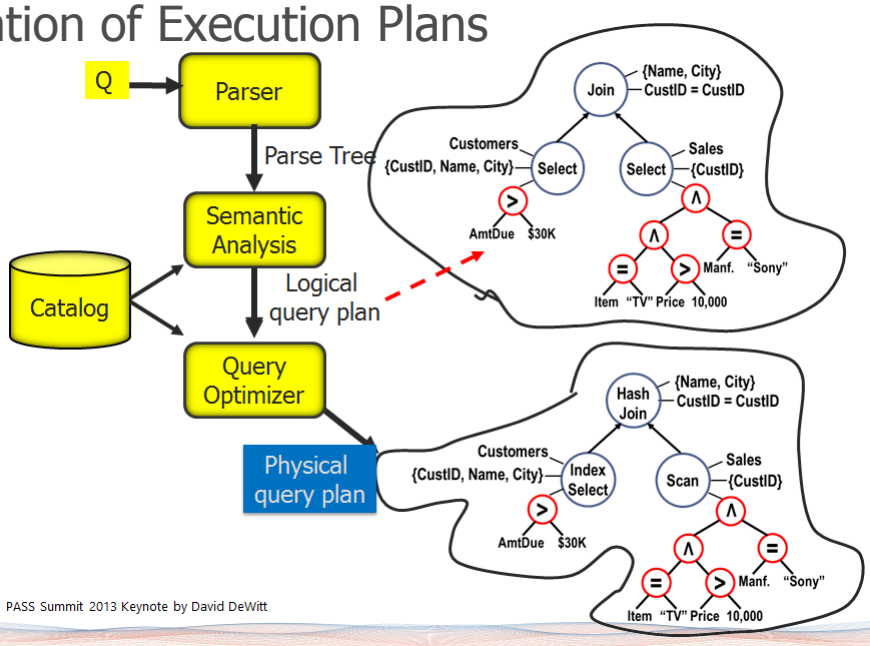
SARPEDON QUALITY LAB

NATIVELY COMPILED PROCEDURES

18

SARPEDON QUALITY LAB

Generation of Execution Plans

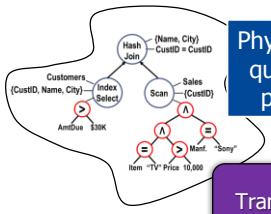


PASS Summit 2013 Keynote by David DeWitt

19

DLL Generation

SARPEDON QUALITY LAB



Physical query plan

Translator

C Code

C Compiler

DLL

DLL Loaded & Invoked

- DLL Generation at 1st execution of procedure
  - Aggressively optimized
  - No function calls

Areas of (CPU) Performance-Gains

SARPEDON QUALITY LAB

- Shared data structures
  - Latches
- Pessimistic Concurency Control
  - Locking
- Query Execution
  - Interpretation

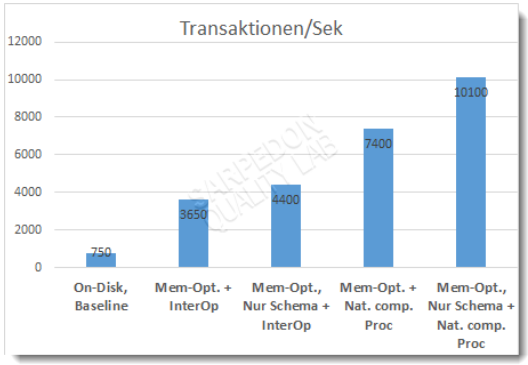
2-5x

10-30x

- Shared data structures
  - Latch-free data structures
- Optimistic Concurency Control
  - Versioning with Timestamps
- Query Execution
  - Compilation in DDL

Performance-Comparison TicketReservation App

SARPEDON QUALITY LAB



- http://www.insidesql.org/blogs/andreaswolter/2014/05/artikel-sql-server-2014-richtigstellungen

22

In-Memory OLTP enhancements

```
ALTER TABLE Sales.SalesOrderDetail
   ALTER INDEX PK_SalesOrderID
   REBUILD
      WITH (BUCKET_COUNT=100000000)
```

```
T-SQL surface area: New

   {LEFT|RIGHT} OUTER JOIN
   Disjunction (OR, NOT)
   UNION [ALL]
   SELECT DISTINCT
   Subqueries (EXISTS, IN, scalar)
```

ALTER support

Full schema change support: add/alter/drop column/constraint
Add/drop index supported
`ALTER PROCEDURE`
Use `sp_recompile` to recompile stored procedures on the next execution

Surface area improvements

Almost full T-SQL coverage including scaler user-defined functions

Improved scaling

Increased size allowed for durable tables; more sockets

Other improvements

Support for Transparent Data Encryption
MARS support

Improved scaling

- Other enhancements include:
  - Multiple threads to persist memory optimized tables
  - Multi-threaded recovery
  - MERGE operation
  - Dynamic management view improvements to `sys.dm_db_xtp_checkpoint_stats` and `sys.dm_db_xtp_checkpoint_files`
- In-memory OLTP engine has been enhanced to scale linearly on servers up to 4 sockets

Greater Transact-SQL coverage

- CREATE PROCEDURE (Transact-SQL)
- DROP PROCEDURE (Transact-SQL)
- ALTER PROCEDURE (Transact-SQL)
- SELECT (Transact-SQL) and INSERT SELECT statements
- SCHEMABINDING and BEGIN ATOMIC (required for natively compiled stored procedures)
- NATIVE_COMPILATION
- Parameters and variables can be declared as NOT NULL
- Table-valued parameters
- EXECUTE AS OWNER, SELF, and user.
- GRANT and DENY permissions on tables and procedures
- Nesting natively compiled stored procedures
- RIGHT OUTER JOIN, LEFT OUTER JOIN, INNER JOIN, and CROSS JOIN in SELECT statements
- NOT, OR, and IN operators in SELECT, UPDATE and DELETE statement
- UNION ALL and UNION
- SELECT DISTINCT
- GROUP BY clause with no aggregate functions in the SELECT clause (<select> list).
- COLUMNSTORE
- COLLATE

Case Studies from Microsoft Customers

SARPEDON QUALITY LAB

Bwin – large online betting company — **17X**

Application: session state
Read and updated for every web interaction
Current max throughput: 15,000 requests/sec
Throughput with Hekaton: 250,000 requests/sec

SBI Liquidity Market – foreign exchange broker — **2X**

Application: online calculation of currency prices from aggregate trading data
Current throughput: 2,812 TPS with 4 sec latency
Hekaton throughput: 5,313 TPS with <1 sec latency

EdgeNet – provides up-to-date inventory status information — **17X**

Application: rapid ingestion of inventory data from retailers
Current max ingestion rate: 7,450 rows/
Hekaton ingestion rate: 126,665 rows/
Allows them to move to continuous, online
once-a-day batch ingestion

Ferranti Computer Systems — **~200 Mio rows/15 Min**

Analysis of Smart Meter data for power supply

In-Memory OLTP, Reactive Extensions und HDInsight technologies
Standard Hardware

www.microsoft.com/casestudies/Microsoft-SQL-Server-2012/Ferranti-Computer-Systems/Utilities-ISV-Scales-to-Meet-Customer-Needs-for-Storage-and-Analysis-of-Big-Data/710000000701

26

SARPEDON QUALITY LAB

REAL-TIME OPERATIONAL ANALYTICS & IN-MEMORY OLTP

27

Operational Analytics

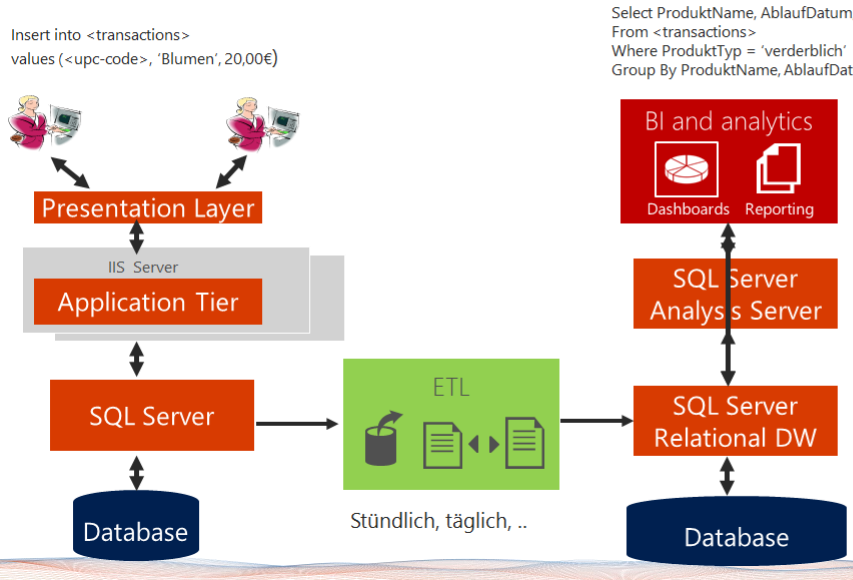
SARPEDON QUALITY LAB

"Ability to run Analytics Queries concurrently with Operational workload using the same schema."

28

Klassische Real-time Analytics Architektur

SARPEDON QUALITY LAB



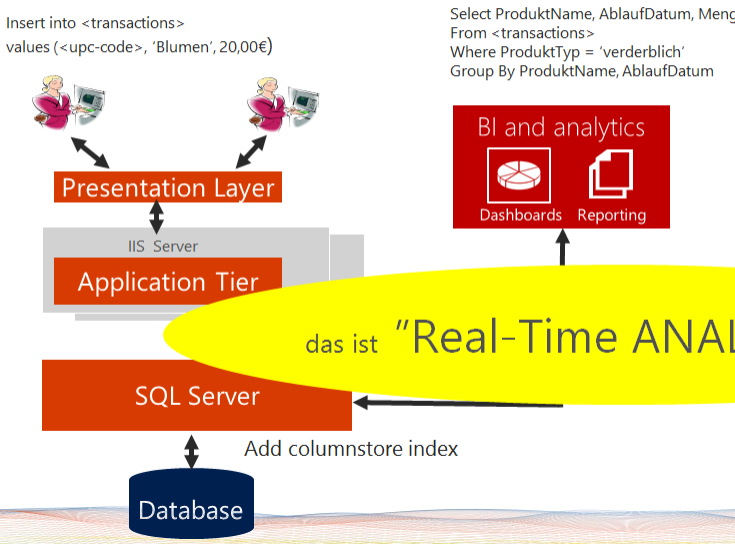
Herausforderungen:

- Komplexe Umsetzung
- Min. 2 Server benötigt (CapEx and OpEx)
- Daten-Latenz während der Analysen
- Fachbereich benötigt immer häufiger "Real-time Analytics"

29

SARPEDON QUALITY LAB

Daten-Latenz Minimierung für Analysen



Vorteile:

- Keine Daten-Latenz
- Kein ETL notwendig
- Kein zusätzliches DW

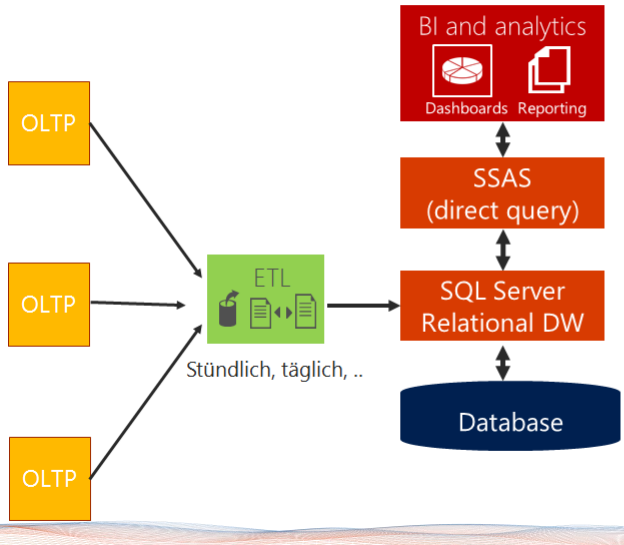
g:

- Minimierter Einfluss auf den OLTP Workload
- Performante Analysen

30

SARPEDON QUALITY LAB

Real-Time Analytics – Wofür NICHT geeignet?

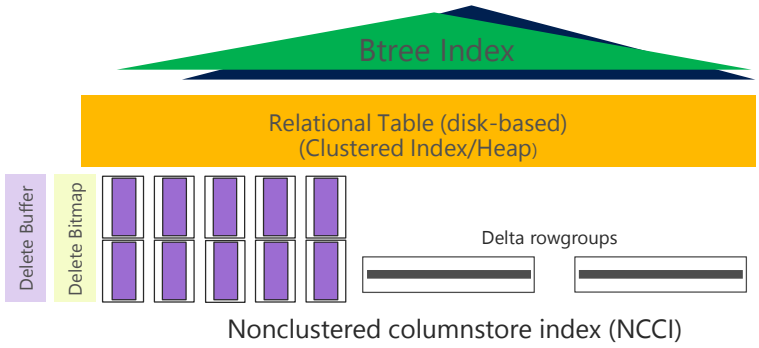


- Verwendung/ Aggregation mehrere OLTP Systeme
- Extreme Analysen
  - Vor-aggregierte CUBEs notwendig
  - Star-Schema
- Schwierigkeiten beim Zugriff auf das OLTP Schema
  - Normalisierte Daten
  - "Multi-table Joins" notwendig

31

SARPEDON QUALITY LAB

Real-time Analytics: Nonclustered Columnstore Index (NCCI)



Nonclustered columnstore index (NCCI)

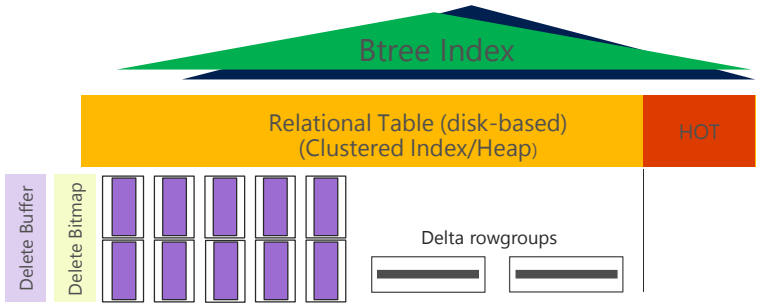
Vorgehen

- Erstellung eines "updateable non-clustered columnstore index" (NCCI) für analytische Abfragen
- Löschung aller weiteren Analysebedingten Indexe
- KEINE Änderungen an der OLTP Anwendung notwendig!
- ColumnStore Index wird automatisch über DML Operationen aktualisiert
- Query Optimizer verwendet den NCCI sofern notwendig

32

SARPEDON QUALITY LAB

Real-time Analytics: Minimierung der Belastung durch den Columnstore Index



Nonclustered columnstore index (NCCI) – filtered index

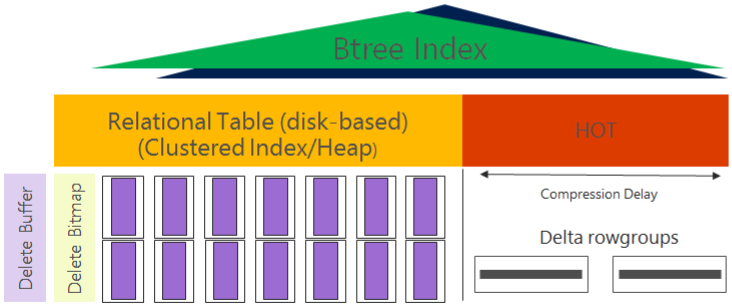
Vorgehen

- Erstellung des Columnstore Index NUR auf "kalten" Daten – Filterkriterium minimiert den Wartungsaufwand
- Analytische Abfragen greifen auf heiße & kalte Daten transparent zu
- Syntax:
  Create nonclustered columnstore index <name> on <table> (<columns>)
  where <filter_col> = <value>

33

SARPEDON QUALITY LAB

Real-time Analytics: Minimierung der Belastung durch den Columnstore Index



Nonclustered columnstore index (NCCI) – filtered index

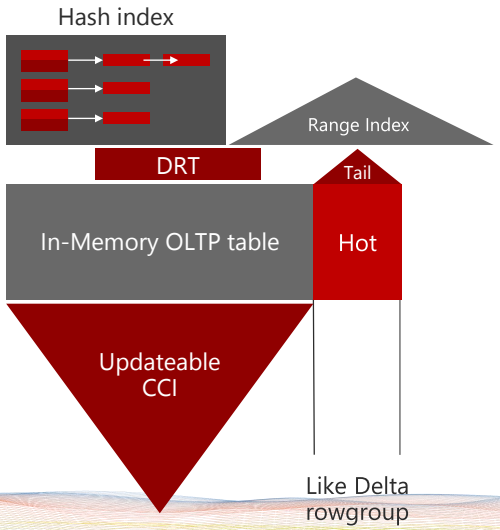
Vorgehen

- Delta rowgroup wird erst nach Ablauf des "compression delay" Zeitraums komprimiert
- Minimierung/ Vermeidung von Index Fragmentierung
- Syntax:
  Create nonclustered columnstore index <name> on <table> (<columns>) with (compression_delay = 30 Minutes)

34

SARPEDON QUALITY LAB

Operational analytics: Columnstore on in-memory tables



Columnstore data fully resident in memory
Persisted together with operational data

Deleted Rows Table (DRT) – Tracks deleted rows

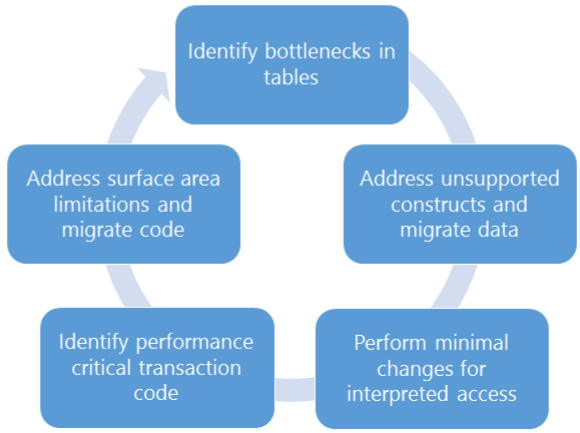
No application changes required

- Rows (tail) not in columnstore stay in in-Memory OLTP table
- No columnstore index overhead when operating on tail
- Background task migrates rows from tail to columnstore in chunks of 1 million rows not changed in last 1 hour – as of CTP3

35

SARPEDON QUALITY LAB

Iterative Approach to Migration



36

SARPEDON QUALITY LAB

Vielen Dank für Eure Zeit

Andreas Wolter

Microsoft CERTIFIED Solutions Master Data Platform
Microsoft CERTIFIED Master SQL Server® 2008

Contact: andreas.wolter@SarpedonQualityLab.com
Blog: www.insidesql.org/blogs/andreaswolter/
Xing: www.xing.com/profile/Andreas_Wolter2
Twitter: @AndreasWolter

37

Connect Items, which love to get your vote ☺

SARPEDON QUALITY LAB

- Extended Events UI Export Display Settings: include grouping
  - To improve the XEvent GUI in terms of saving view settings incl. grouping
- Support SQL Broker Service to be a target of Extended Events
- Providing a special Server principal for Database Ownership
  - To solve the problem with the database owner, which unfortunately is sa in 80% of all servers – with all the security risks – explained in my blog:
    - http://www.insidesql.org/blogs/andreaswolter/2013/12/survey-sql-server-database-ownership-datenbankbesitzer
- Allow the use of saved Credentials/Proxy Accounts for Reporting Services Subscriptions
  - Bekommen wir für File-Subscriptions in SSRS 2016 ☺
- Shared Datsets with spatial data - no preview in map pane and as well as in spatial data + analytical dat…
- Group Managed Service Accounts Support for SQL Server Failover Clusters
  - Bekommen wir nun in SQL Server 2016 ☺

38

Die nächsten SQL Server Master-Classes ab Juni 2016

SQL SERVER MASTER-CLASS
by SARPEDON QUALITY LAB

- Workshop **Hochverfügbarkeit mit SQL Server Always On**
  - Windows Server Failover Clustering
  - SQL Server Clustered Instances
  - SQL Server Availability Groups
  - Logshipping
  - Frankfurt, 30.5.-1.6. *
- **SQL Server Performance-Analyse** Workshop
  - Frankfurt, 29.6.-1.7. *

15% Rabatt-Code für PASSler:
SQLPASS_de16

www.SQL-Server-Master-Class.com

*Termine noch unter Vorbehalt

39

Sarpedon Quality Lab:
Ihr Spezialist für Datenbank-Systeme basierend auf SQL Server Technologien

Wir sind **Deutschlandweit die einzige Firma,**
die die **höchsten technischen Zertifizierungen** von Microsoft für sowohl **SQL Server 2008** als auch **SQL Server 2012** erreicht hat.

Wir unterstützen und setzen unser Know-How gern zu Ihrem Vorteil ein.

Unsere **Dienste** umfassen:

- *SQL Server Supportverträge & Hotline*
- *SQL Server Health Checks*
- *Architektur-Planung, Beratung und Implementierung*
- *Performance Analyse & Tuning*
- *Disaster-Recovery & SLA-Compliance-Checks*
- *Security-Checks*
- *Datenrettung bei Korruption*

- ***Training***: SQL SERVER MASTER-CLASS

Fragen Sie uns: info@Sarpedon.de

40