



# **Developing Your Indexing Strategy**

*For the BI Professional*

Ben DeBow

*E-mail: [bdebow@sqlha.com](mailto:bdebow@sqlha.com)*

*Twitter: @BBQSQL*

*Web/Blog: <http://www.sqlha.com>*

# About Me - @BBQSQL

- Partner SQLHA
- Working with SQL Server for 18yrs
- Author, speaker, and consultant focusing on Performance, Scalability, and Availability



# Just Tables and Columns

- Lowest level, just tables and columns
  - How are the tables accessed?
  - Critical processes
  - Data types
  - Frequency of requests
  - Read/write ratios
  - Size of data

# Type of Indexes in SQL 2012

- Clustered
- Non-clustered
- Columnstore
  - SQL 2014 – clustered and updatable
- Spatial
- XML
- Full-Text
- Variants – Unique, Filtered, Computed, INCLUDED columns

# Index and Statistics Properties

- Indexes
  - Data compression (PAGE and ROW)
  - Enable and disable indexes
  - Enable and disable ROW and PAGE locks
  - File group placement
  - FILLFACTOR and PAD\_INDEX
- Statistics
  - Filtered Statistics
  - FULLSCAN or SAMPLE
  - NORECOMPUTE

# Before we start...

- Type of system
  - OLTP
  - OLAP
  - Hybrid
  - Staging
- Understand the high level processes/application
- Review the data model
- Capture runtime and metadata

# What is an Indexing Strategy

- Optimal path to the data
  - More indexes != strategy
- Balancing act
  - Competing priorities – DML versus SELECTs
- Lower risk, high reward
- Reduce contention
- Increase concurrency

# Typical Access Paths

WideTable									
PK - C1	C2	C3	C4	C5	C6	C7	C8	C59	C60

Index1  
(Primary Key)

Index2  
(Alternate Key)

Index3

15 more indexes to cover unique access paths

NarrowTable					
PK - C1	C2	C3	C4	C5	C6

Index1  
(Primary Key)

Index2  
(Alternate Key)

Index3

More normalized design, fewer entry points



# Analysis

- Review schema
  - Type of data model
- Efficiency of existing indexing strategy
  - Leverage SQL DMVs
    - `sys.dm_db_index_usage_stats`
- Review processes
  - `sys.dm_exec_cached_plans`
  - Extended events or Server side traces

## DMVs and Schema

# BI Indexing Tips

- Depends on type of design
  - Hybrid – OLTP with reporting
  - Ad-hoc reporting
  - Staging
  - Star schema
  - Dimensional model
- Add clustered index to surrogate keys
- Add non-clustered index to FKs, AKs and key dim columns

# BI Indexing Tips

- Add clustered index to fact table
- Leverage table partitioning
  - By time, partition elimination
  - Switch in/out partitions
- Disable indexes during loads into staging
- Enable data compression
- Add FKs with NOCHECK

# Columnstore Index Tips

- Add to the fact table, every column
- Do not add to dims unless there are several million rows
- To benefit from batch processing
  - Avoid OUTER joins, NOT IN, UNION ALL
- Limitations:
  - Read only
  - No clustered columnstore index
- New features in SQL 2014
- See: <http://tinyurl.com/n2bu84a>

# Summary

- New columnstore features in SQL Server 2014
- Gather high level application information
- Implement a couple indexes at a time
- Capture utilization data, if it is not as expected then remove the indexes
- Keep before and after statistics
  - Use this data for a raise
- Make data available to developers
  - Eventually you want them to tune the applications



## **Questions and Discussion**

