

Aas Trailblazers

Unleashing insights

Index Options

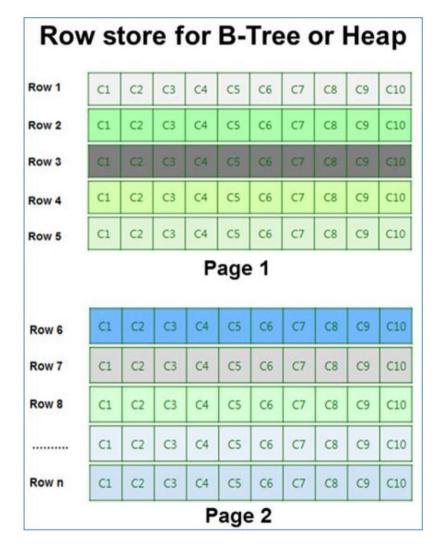
Row Store

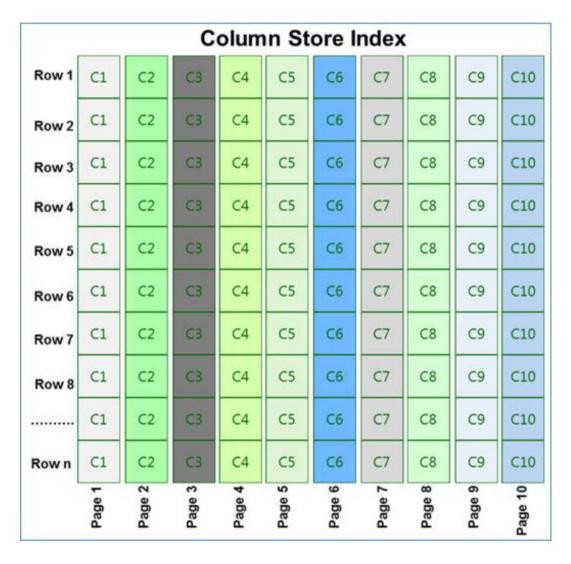
- Heap
- Clustered Index
- Non-Clustered Index

Clustered Columnstore Index

Default if not defined in CREATE TABLE

Row store vs. Column store





https://www.databasejournal.com/features/mssql/understanding-new-column-store-index-of-sql-server-2012.html

Row store: Heaps

- Just a collection of pages
 - Append only
 - No ordering
 - Minimal fragmentation
- Fast target for loading
- Much slower when used for small range scans & single selects

Understanding Heaps

Deletes are not physically removed from a heap until it is rebuilt. As data is added SQL allocates more pages to the table.

Row store: Indexes (clustered and non-clustered)

- Use a Balanced Tree (b-tree) to organize data
- Get fragmented over time
 - Require more maintenance
- Leaf level
 - Clustered index: data
 - Non-clustered index: Row ID (RID) may require lookup of data
- Ideal for limited range scans & singleton selects
- Slower for table scans / partition scans / loading

Columnstore: Index

Group rows into batches of ~1M rows

This is called a row group

Compress/Encode each column

- Lower selectivity compresses more efficiently
- Primitive data types (int, date, short string, etc.)
 compresses efficiently
- Long string fields do NOT compress well

Typically, up to 10x compression and query performance up to 10x*

^{* &}lt;a href="https://docs.microsoft.com/en-us/sql/relational-databases/indexes/columnstore-indexes-overview?view=sql-server-ver15">https://docs.microsoft.com/en-us/sql/relational-databases/indexes/columnstore-indexes-overview?view=sql-server-ver15

Columnstore: Design and storage

Data

2034857,23552534,26262569085923458958294582342-52935-2385349085295-25894-589245-285928592-5845829582-58258295849058-28592-

- 5829458240598294852905840958958459028590285920458294582598205895829 0582945082905825-2502-45905-93245,vitoortkgldkggjwov j4o534585-
- 0348565920345234059=3405943=-5923405=23950345923=559235=239560235932=46942306496046940693=460436
- 93b069,hb05,b6905869347 87-987g89-988g-89-89 89-89-89mg89wert8t9et8-t-=8349652-
- =856=8=98t0e=t9e0t=e9t09=90=39560=659450693=565096=35695=69305=69,v w06=62=96069,b]si5=96292500000-
- 2034857,23552534,26262569085923458958294582342-52935-2385349085295-25894-589245-285928592-5845829582-58258295849058-28592-
- 5829458240598294852905840958958459028590285920458294582598205895829 0582945082905825-2502-45905-93245, vitoortkgldkg vlgjwov j4o534585-0348565920345234059-3405943-6
- 5923405=23950345923=509235=239560235932=46942306496046940693=460436 93b069,hb05,b6905869347 87-987g89-9s8g-89-89 89-89-89mg89wer-
- t8t9et8-t-=8349652-=856-8=98t0e=t9e0t=e9t09=90=39560=659450693=565096=35695=69305=69,v
- w06=62=96069,b]si5=96292500000-2034857.23552534.26262569085923458958294582342-52935-2385349085295-
- 25894-589245-285928592-5845829582-58258295849058-28592-582945824059829485290584095892034857,23552534,26265690859234589582 94582342-52935-2385349085295-25894-589245-285928592-5845829582-58258295849058-28592-
- 5829458240598294852905840958958459028590285920458294582598205895829 0582945082905825-2502-45905-93245,vitoortkgldkggjwov j4o534585-
- 0582945082905825-2502-45905-93245,vitoortkgldkggjwov j4o534585-0348565920345234059-3405943--5923405-23950345923-509235-239560235932-46942306496046940693-460436
- 93b069,hb05,b6905869347 87-987g89-9s8g-89-89 89-89-89mg89wert8t9et8-t-=8349652-
- =856=8=98t0e=t9e0t=e9t09=90=39560=659450693=565096=35695=69305=69,v w06=62=96069,b]si5=96292500000-
- 2034857,23552534,26262569085923458958294582342-52935-2385349085295-25894-589245-285928592-5845829582-58258295849058-28592-
- 5829458240598294852905840958958459028590285920458224582598205895829 0582945082905825-2502-45905-93245, vitoortkgldkg vlgjwov j4o534585-0348565920345234059-3405943--
- 5923405=23950345923=509235=239560235932=46942306496046940693=460436 93b069,hb05,b6905869347 87-987g89-9s8g-89-89 89-89-89mg89wer-
- t8t9et8-t-=8349652-=856=8=98t0e=t9e0t=e9t09=90=39560=659450693=565096=35695=69305=69,v
- w06=62=96069,b]si5=96292500000-2034857,23552534,26262569085923458958294582342-52935-2385349085295-
- 25894-589245-285928592-5845829582-58258295849058-28592-58294582405982948529058409589

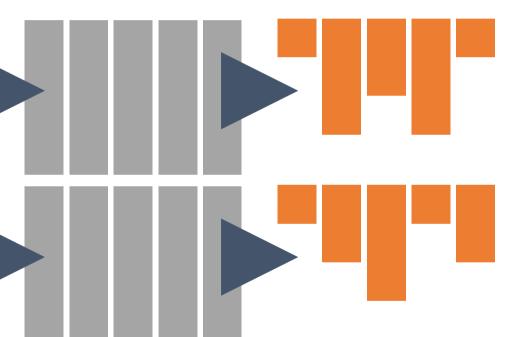
Row Group

Segments Columnstore

- 2034857,23552534,26262569085923458958294582342-52935-2385349085295-25894-589245-285928592-5845829582-58258295849058-28592-
- 58294582405982948529058409589584590285902859204582945825982058958290582945082905825-2502-45905-93245,vitoortkgldkggjwov j4o534585-
- 5923405-23950345923-509235-239560235932-46942306496046940693-46043693b06 b05,b6905869347-87-987g89-988g-89-89-89-89mg89wer-t819et8-t--8349652-
- -856-8-950-et-t9e0t-e9t09-90-39560-659450693-555096-35695-69305-69, vw06-6
 -95,b)si5-96292500000-2034857,23552534,2626256908592345895829458234252935-2385349085295-25894-589245-285928592-5845829582-58258295849058-285
 5829458240598294852905840958958459028590285902458294582598205859582
- 82905825-2502-45905-93245, vitoortkgldkg vlgjwov j4o534585-0348565920345234059-3405943--
- 0348565920345234059=3405943=-5923405=23950345923-559235=239560235932=46942306496046940693=46043693b06
- b05,b6905869347 87-987g99-988g-89-89 89-89-887g89wer-t8t9et8-t--82349652 ---8668-98t0e=t9e0t-e9t09-90-39560-659450693-5655096-35695-69305-69, vw06-62-9 -6669,b]si5-96292500000-2034857,23552534,26262569085923458958294582342 -
- $52935 2385349085295 25894 589245 285928592 5845829582 58258295849058 28592 582945824059829485290584095892034857, \\ 23552534, \\ 26262569085923458958294$

2034857,23552534,26262569085923458958294582342-52935-2385349085295-25894-589245-285928592-582582958295829582-589580-589582-5895800-589580-589580-589580-589580-589580-589580-589580-589580-58958

- 589245-285928592-5845829582-58258295849058-28592-5829458240598294852905840958958459028590285920458259820589582905829450 82905825-2502-45905-93245.viroortkalikaariwov i4o534585-
- 82905825-2502-45905-93245, vitoortkgidkggjwov j4o5345: 0348565920345234059=3405943=-
- 5923405=23950345923=509235=239560235932=46942306496046940693=46043693b06 bb5, b6905869347 87-987g89-988g-89-89 89-89-89mg99wer-tbt9et8-t-=8349652-2856=8980et96bte=9069=0039606560=559450693=56509636569569305=69, ww06-6
- 9, b) s1596292500000-2034857, 23552534, 26262569085923458958294582342-5235-238349085295-25894-589245-25852659-5645825862-56526258695055-265 5829458240598294882905840958958459028590285902859294582598205895829058
- 82905825-2502-45905-93245,vitoortkgldkg vlgjwov j4o534585-0348565920345234059=3405943--
- 0348565920345234059-3405943-5923405-23950345923-509235-239560235932-46942306496046940693-46043693b06
 bbs h60f886331 87-087/89-0687-89-80 89-89-89-89-89-89-88-80-88-1-8830665-
- b05, b6905869347 87-987g9-98g-989-89 89-89-89mg99wg-t8t5et8-t--8349652-=856-8-98tbe-t9et-e950-99-09-3565-659540693-56506-35659-69305-69; www.6e62-9 6069, bjsi5-96292500000-2034857,2355254,2626256908592458958294582342-57935-238343985795-75840-589745-787958795-787857879587857878398589582
- 582945824059829485290584095892034857,23552534,262662569085923458958294



Demo

Columnstore: Design and storage

Data

2034857,23552534,26262569085923458958294582342-52935-2385349085295-25894-589245-285928592-5845829582-58258295849058-28592-

- 5829458240598294852905840958958459028590285920458294582598205895829 0582945082905825-2502-45905-93245,vitoortkgldkggjwov j4o534585-
- 0348565920345234059=3405943=-5923405=23950345923=559235=239560235932=46942306496046940693=460436
- 93b069,hb05,b6905869347 87-987g89-988g-89-89 89-89-89mg89wert8t9et8-t-=8349652-
- =856=8=98t0e=t9e0t=e9t09=90=39560=659450693=565096=35695=69305=69,v w06=62=96069,b]si5=96292500000-
- 2034857,23552534,26262569085923458958294582342-52935-2385349085295-25894-589245-285928592-5845829582-58258295849058-28592-
- 5829458240598294852905840958958459028590285920458294582598205895829 0582945082905825-2502-45905-93245, vitoortkgldkg vlgjwov j4o534585-0348565920345234059-3405943-6
- 5923405=23950345923=509235=239560235932=46942306496046940693=460436 93b069,hb05,b6905869347 87-987g89-9s8g-89-89 89-89-89mg89wer-
- t8t9et8-t-=8349652-=856-8=98t0e=t9e0t=e9t09=90=39560=659450693=565096=35695=69305=69,v
- w06=62=96069,b]si5=96292500000-2034857.23552534.26262569085923458958294582342-52935-2385349085295-
- 25894-589245-285928592-5845829582-58258295849058-28592-582945824059829485290584095892034857,23552534,26265690859234589582 94582342-52935-2385349085295-25894-589245-285928592-5845829582-58258295849058-28592-
- 5829458240598294852905840958958459028590285920458294582598205895829 0582945082905825-2502-45905-93245,vitoortkgldkggjwov j4o534585-
- 0582945082905825-2502-45905-93245,vitoortkgldkggjwov j4o534585-0348565920345234059-3405943--5923405-23950345923-509235-239560235932-46942306496046940693-460436
- 93b069,hb05,b6905869347 87-987g89-9s8g-89-89 89-89-89mg89wert8t9et8-t-=8349652-
- =856=8=98t0e=t9e0t=e9t09=90=39560=659450693=565096=35695=69305=69,v w06=62=96069,b]si5=96292500000-
- 2034857,23552534,26262569085923458958294582342-52935-2385349085295-25894-589245-285928592-5845829582-58258295849058-28592-
- 5829458240598294852905840958958459028590285920458224582598205895829 0582945082905825-2502-45905-93245, vitoortkgldkg vlgjwov j4o534585-0348565920345234059-3405943--
- 5923405=23950345923=509235=239560235932=46942306496046940693=460436 93b069,hb05,b6905869347 87-987g89-9s8g-89-89 89-89-89mg89wer-
- t8t9et8-t-=8349652-=856=8=98t0e=t9e0t=e9t09=90=39560=659450693=565096=35695=69305=69,v
- w06=62=96069,b]si5=96292500000-2034857,23552534,26262569085923458958294582342-52935-2385349085295-
- 25894-589245-285928592-5845829582-58258295849058-28592-58294582405982948529058409589

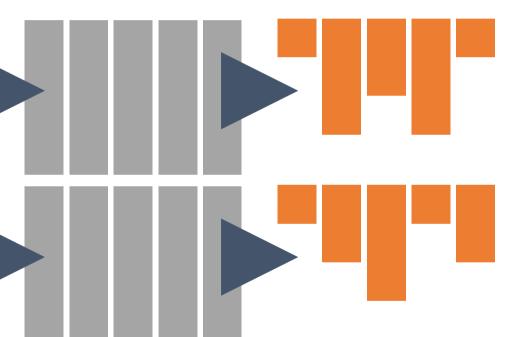
Row Group

Segments Columnstore

- 2034857,23552534,26262569085923458958294582342-52935-2385349085295-25894-589245-285928592-5845829582-58258295849058-28592-
- 58294582405982948529058409589584590285902859204582945825982058958290582945082905825-2502-45905-93245,vitoortkgldkggjwov j4o534585-
- 5923405-23950345923-509235-239560235932-46942306496046940693-46043693b06 b05,b6905869347-87-987g89-988g-89-89-89-89mg89wer-t819et8-t--8349652-
- -856-8-950-et-t9e0t-e9t09-90-39560-659450693-555096-35695-69305-69, vw06-6
 -95,b)si5-96292500000-2034857,23552534,2626256908592345895829458234252935-2385349085295-25894-589245-285928592-5845829582-58258295849058-285
 5829458240598294852905840958958459028590285902458294582598205859582
- 82905825-2502-45905-93245, vitoortkgldkg vlgjwov j4o534585-0348565920345234059-3405943--
- 0348565920345234059=3405943=-5923405=23950345923-559235=239560235932=46942306496046940693=46043693b06
- b05,b6905869347 87-987g99-988g-89-89 89-89-887g89wer-t8t9et8-t--82349652 ---8668-98t0e=t9e0t-e9t09-90-39560-659450693-5655096-35695-69305-69, vw06-62-9 -6669,b]si5-96292500000-2034857,23552534,26262569085923458958294582342 -
- $52935 2385349085295 25894 589245 285928592 5845829582 58258295849058 28592 582945824059829485290584095892034857, \\ 23552534, \\ 26262569085923458958294$

2034857,23552534,26262569085923458958294582342-52935-2385349085295-25894-589245-285928592-582582958295829582-589580-589582-5895800-589580-589580-589580-589580-589580-589580-589580-589580-58958

- 589245-285928592-5845829582-58258295849058-28592-5829458240598294852905840958958459028590285920458259820589582905829450 82905825-2502-45905-93245.viroortkalikaariwov i4o534585-
- 82905825-2502-45905-93245, vitoortkgidkggjwov j4o5345: 0348565920345234059=3405943=-
- 5923405=23950345923=509235=239560235932=46942306496046940693=46043693b06 bb5, b6905869347 87-987g89-988g-89-89 89-89-89mg99wer-tbt9et8-t-=8349652-2856=8980et96bte=9069=0039606560=559450693=56509636569569305=69, ww06-6
- 9, b) s1596292500000-2034857, 23552534, 26262569085923458958294582342-5235-238349085295-25894-589245-25852659-5645825862-56526258695055-265 5829458240598294882905840958958459028590285902859294582598205895829058
- 82905825-2502-45905-93245,vitoortkgldkg vlgjwov j4o534585-0348565920345234059=3405943--
- 0348565920345234059-3405943-5923405-23950345923-509235-239560235932-46942306496046940693-46043693b06
 bbs h60f886331 87-087/89-0687-89-80 89-89-89-89-89-89-88-80-88-1-8830665-
- b05, b6905869347 87-987g9-98g-989-89 89-89-89mg99wg-t8t5et8-t--8349652-=856-8-98tbe-t9et-e950-99-09-3565-659540693-56506-35659-69305-69; www.6e62-9 6069, bjsi5-96292500000-2034857,2355254,2626256908592458958294582342-57935-238343985795-75840-589745-787958795-787857879587857878398589582
- 582945824059829485290584095892034857,23552534,262662569085923458958294



Columnstore: Row Groups

Exist in one of these states

- Open: In delta store; accepting new rows, data is not compresses or indexed
- Closed: In delta store; not accepting new rows
- Compressed: In Column Store format
- Tombstone: a rowgroup with all the data deleted

System View: sys.pdw_nodes_column_store_row_groups

Columnstore: Column Segments

Example of a segment from sys.pdw_nodes_column_store_segments

segment_id	row_count	base_id	min_data_id	max_data_id
0	1048576	19568138	19568141	30426164
1	1048576	19757017	19757020	31243669
2	1048576	19788588	19788591	31341348
3	1048576	19778863	19778866	31349845
4	1048576	19574032	19574035	31365043
E	1040E7C	10772022	10772020	21220102

Demo

Columnstore: Updating

CCIs perform large block compression

- Maximises compression
- Expensive to update

CCIs use two mechanisms to handle updates

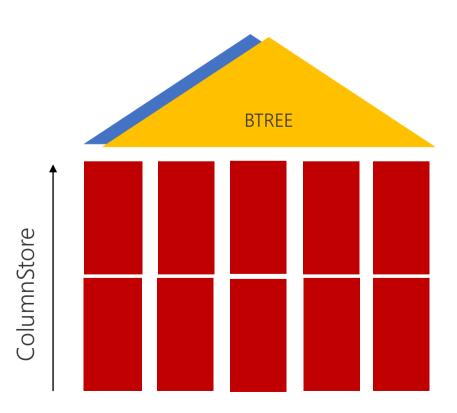
- Delete bitmap
- Delta stores

Columnstore: Delta Store

Page compressed b-tree i.e. a row store

- B-tree on unique integer row ID
- Matches user columns defined in the CCI
- Aligned to underlying CCI partition
- Created only if needed

Max # rows delta store 1,048,576



Columnstore: Delta Store to Columnstore

Once a row group has closed it can be converted to column store format:

As an **online operation**

- Tuple Mover: a background process
- ALTER INDEX..REORGANIZE WITH (COMPRESS_ALL_ROW_GROUPS = ON);

As an offline operation

ALTER INDEX..REBUILD

Columnstore: Small DML

Inserts

- Small inserts are written first to the delta store
- Bulk-load based on some internal threshold writes as column store

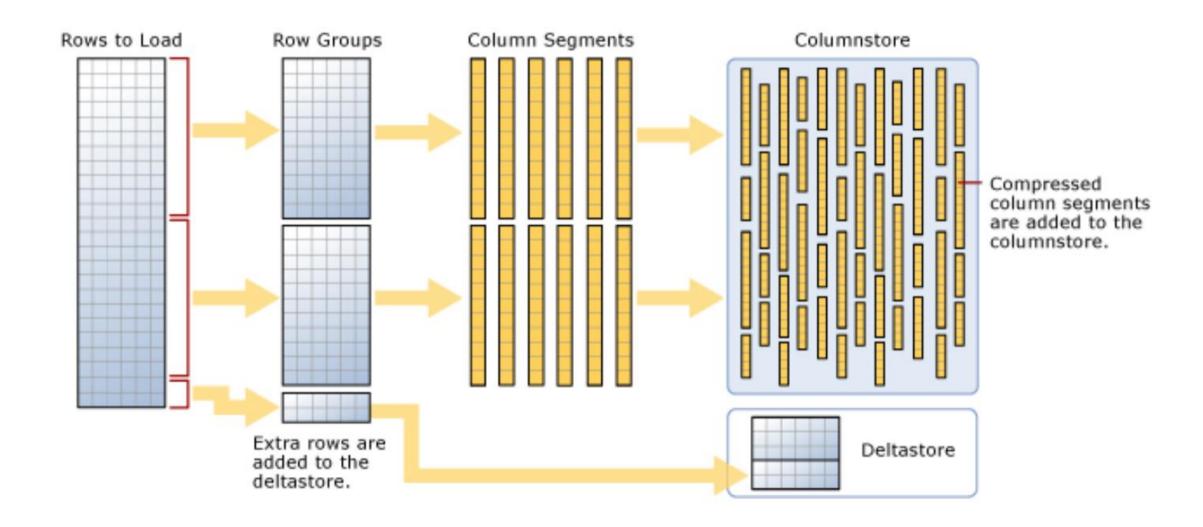
Deletes

- Logical against rows in column store
- Physically against rows in the delta store

Updates

Converted to a logical delete and an insert

Columnstore: Design and storage



Demo

Optimizing with Indexes: Choose right type

Clustered Columnstore (Default)

- Optimal choice for large tables
- Limits scans to columns in the query
- Optimal compression
- Slower to load than Heap
- Keep partitions large enough to compress (> 1 million rows)

Heap

- Optimal choice for temporary or staging tables
- Fastest load performance

Clustered Index

- Optimal for tables < 60M rows
- Sorting operation slowsdown load

Non-clustered Index

- Use sparingly
- Optimize single row lookups
- Will slow-down load