SQL Server 2012 (V2) Parallel Data Warehouse Insights (Level 200)

Henk.vanderValk@Microsoft.com
Agenda

• Introducing
• Why SQL PDW
  – PDW Hardware Architecture
  – PDW New Database Features

• SQL Server PDW Live demo’s
  Comparing SQL2012/SQL2014 (SMP) vs PDW (MPP)
  ✓ Data loading
  ✓ Query Performance - Data loading
  ✓ Query Concurrency
  ✓ Polybase / Hadoop integration

• Q&A
About Henk

• 10 years of Unisys-EMEA Performance Center
• 2002- Largest SQL DWH in the world (SQL2000)
• Project Real – (SQL 2005)
• ETL WR - loading 1TB within 30 mins (SQL 2008)
• Contributed to several SQL whitepapers

• Schuberg Philis-100% uptime for mission critical applications
• Since april 1st, 2011 – Microsoft SQL PDW -  Western Europe
1999 - How it all started with Unisys & High End SQL Server: the ES7000

- 32 Way / 64 GB RAM
- 32bit PIII - 550 MHz
- 64bit Itanium 800 MHz
  - SQL Server (64-bit) released in April 2002.
- SQL Server 2000 EE, SP1
Unisys ES7000 …
Big Windows / Wintel Mainframe Architecture

➤ Up to 32 CPU’s
➤ 64/128 GB RAM
➤ 96 PCI slots

Example: Sequential Read
➤ ES7000/100
➤ 1 x OS: Windows 2000 Datacenter
➤ 42 x 1Gbit HBA’s
➤ EMC Clariion’s with 250 physical drives
PDW: Querying 1 Peta byte of data in 1 second / 40 node

Querying 294 billion rows in seconds
SQL Server 2012
PDW
Introduction
Limitations: Performance and Scale today

Diminishing performance

Scale UP

Diminishing Scale as requirements grow

Rowstore

Existing Tables (Partitions)

Non-optimal performance for many DW queries
SQL Server 2012 Parallel Data Warehouse (PDW)

Insights on any data of any size

Scale-Out Architecture on Industry Standard Hardware

Next-generation Performance At Scale

Built For Big Data
What is Parallel Data Warehouse?

- **Shared-nothing parallel database system**
  - Massively parallel processing (MPP)
  - A “Control” server that accepts user queries, generates a plan, and distributes operations in parallel to compute nodes
  - Multiple “Compute” servers running SQL Server
  - A “Management” server for administering the system
  - A “Data Movement Service” that facilitates parallel SQL operations

- **Delivered as an appliance**
  - Balanced and pre-configured software and industry standard hardware from Dell or HP
  - Single Call Support
  - Fastest Time to Market
  - Scales from 2 to 56 Nodes
Key Design Elements

• Modular Design
• High Density
• Leverage latest Microsoft software features
  
  » Windows Server 2012 Storage Spaces
  » Windows Server 2012 Hyper-V
  » SQL Server 2012 PDW xVelocity ColumnStore

![Works as Designed!](image)
Parallel or not parallel...

**Scale up (SMP)**
- Build for specific requirement
- Build HA etc. additionally
- Maintain and Tune (Load/File Distribution)
- Unknown Future workloads
- Still a very good data mart solution in a Hub and Spoke architecture with SQL Server PDW

**Scale out (MPP)**
- Resilient & Predictable
- Big data / DW Best Practices in a box
- Deploy Fast and Drive Value
- Built-in HA
- Scalable (start small/grow when needed)
Ultra Shared Nothing architecture:

Larger Fact Table is Hash Distributed Across All Compute Nodes

- **Time Dim**
  - Date Dim ID
  - Calendar Year
  - Calendar Qtr
  - Calendar Mo
  - Calendar Day

- **Product Dim**
  - Prod Dim ID
  - Prod Category
  - Prod Sub Cat
  - Prod Desc

- **Sales Facts**
  - Date Dim ID
  - Store Dim ID
  - Prod Dim ID
  - Mktg Camp ID
  - Qty Sold
  - Dollars Sold

- **Store Dim**
  - Store Dim ID
  - Store Name
  - Store Mgr
  - Store Size

- **Mktg Campaign Dim**
  - Mktg Camp ID
  - Camp Name
  - Camp Mgr
  - Camp Start
  - Camp End

Diagram showing the distribution of the fact table across multiple compute nodes.
Ultra Shared Nothing architecture:

Smaller “Dimension Tables” are Replicated on each server/node

- **Time Dim**
  - Date Dim ID
  - Calendar Year
  - Calendar Qtr
  - Calendar Mo
  - Calendar Day

- **Product Dim**
  - Prod Dim ID
  - Prod Category
  - Prod Sub Cat
  - Prod Desc

- **Sales Facts**
  - Date Dim ID
  - Store Dim ID
  - Prod Dim ID
  - Mktg Camp ID
  - Qty Sold
  - Dollars Sold

- **Store Dim**
  - Store Dim ID
  - Store Name
  - Store Mgr
  - Store Size

- **Mktg Campaign Dim**
  - Mktg Camp ID
  - Camp Name
  - Camp Mgr
  - Camp Start
  - Camp End

Result: Fact - Dimension Joins can be performed locally = Co-location!
The Logical View on the base unit

General Details

- All hosts run Windows Server 2012 Standard
- All VMs run Windows Server 2012 Standard as a guest OS
- All Fabric and workload activity happens in Hyper-V virtual machines, with Fabric VM’s sharing 1 server
- Failover is handled by Hyper-V
- PDW Agent runs on all hosts and all VMs, collects appliance health data on fabric and workload
- Windows Storage Spaces handles mirroring and spares

PDW Workload Details

- SQL Server 2012 Parallel Data Warehouse is used on control node and compute nodes for PDW workload

Specs

- 32 cores on 2 active compute nodes
- 512 GB of RAM on compute
- Up to 225TB of user data
Seamlessly add capacity

Start small with a few TB and linearly Scale Out

Smallest (0TB) To Largest (6PB)

• Start small with a few Terabyte warehouse
• Scale Out: Incrementally add HW for Near Linear Scale
• Shared Nothing
• Add capacity up to 6 Petabytes

Start Small And Grow
Largest Warehouse
Minimal Downtime
xVelocity Gives Next-Gen Performance
Lighting Fast Data Query Processing

Columnstore Provides Dramatic Performance

• Updateable and clustered xVelocity Columnstore
• Stores data in columnar format
• Updateable to support bulk and/or trickle loading
• Full DML (Insert / Update / Delete / Select) supported directly on ColumnStore

• Memory-optimized for next-generation performance

Up to 50X Faster
Up to 15x compression
Save Time and Costs
Real-Time DW

Updateable Columnstore Already available NOW!
SQL 2012 PDW - Clustered Columnstore

• Clustered column store index (CCI) as primary storage for a table
  • Support both row-oriented and column-oriented storage (choice per table)
  • Support conversions between row-oriented and column-oriented storage
  • CCI automatically includes all columns in the table
  • CCI cannot coexist with nonclustered indexes
  • CCIs can be created for both distributed and replicated tables
  • All PDW data types are supported

• CCI is **updateable** (can directly insert, update, delete)
  • Delete is a logical delete
  • Update is a delete followed by an insert
  • Insert operations execute against “delta store”
  • CCI **reorganize** is an online operation
    • Move inserts from delta store into columnar storage
  • CCI **rebuild** is an offline operation
    • Reclaims space (since delete is logical)
    • Can choose to rebuild partition by partition
SMP vs MPP

Sweetspot #1 - Data Loading
SQL on SMP vs MPP

SMP: DL580 / 256GB
With High-End
EMC VNX 7500 - flash drives + 100x SAS

VS

8 node SQL PDW
35 spindles / compute note
Data loading SQL SMP 2012/2014/PDW

Single 75 GB / 600 Million row LineItem flat file

**SQL 2012 SMP**

```
USE DemoDb_SMP
GO
BULK INSERT demoDB_SMP.dbo.LINEITEM
FROM 'M:\tpch\lineitem.tbl'
WITH ( FIELDTERMINATOR = '|',
      ROWTERMINATOR = '\n', CODEPAGE = 'RAW',
      TABLOCK)
```

*(500037902 rows) affected*

**SQL 2014 SMP**

```
BULK INSERT demoDB_SMP_CCI.dbo.LINEITEM_SQL14_HEAP
FROM 'M:\tpch\lineitem.tbl'
WITH ( FIELDTERMINATOR = '|', ROWTERMINATOR = '\n', TABLELOCK)
```

*(500037902 rows) affected*

**SQL PDW**

```
dlloader.exe -i D:\TPCH\100\lineItem.tbl -M fastappend -s 1/2.16.49.15 -m -E -c -b 1048576 -rt value -rv 100 -R D:\TPCH\100\lineItem.tbl 30.rejects -e ascii -t "|" -r \r\n-U hevandv -T SQLRALLY2013.dbo.lineitem_CCI_SQLRALLY2013
```

*(500037902 rows) affected*
Comparing throughput

- Impression Bulk inserting a single flatfile on SMP SQL2012 vs PDW

Avg. 21 MB/sec vs 270 MB/sec
SMP SQL201x
Bulk insert uses a single core per task

PDW Bulk insert is writing into all distributions in parallel
Data loading

Single 75 GB / 600 Million row LineItem flat file

<table>
<thead>
<tr>
<th>SQL Server Build</th>
<th>Table Type</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>SQL2012</td>
<td>HEAP</td>
<td>1 hour 35 min</td>
</tr>
<tr>
<td>SQL2014</td>
<td>HEAP</td>
<td>1 hour 14 min</td>
</tr>
<tr>
<td>SQL2014</td>
<td>Create CCI on Heap</td>
<td>6 min 35 sec</td>
</tr>
<tr>
<td>SQL2014</td>
<td>CCI</td>
<td>2 hour 59 min</td>
</tr>
<tr>
<td>SQL2012 PDW</td>
<td>HEAP</td>
<td>6 min 5 sec</td>
</tr>
<tr>
<td>SQL2012 PDW</td>
<td>CCI</td>
<td>5 min 25 sec</td>
</tr>
</tbody>
</table>

PDW loads single flatfile data 15-36x faster!
SSIS PDW V2 Destination Adapter

Single 75 GB / 600 Million row LineItem flat file

SSIS Dataload

2.5x Faster
SSIS PDW V2 Destination Adapter

- Append
- Reload
- Upsert
- Fastappend
SMP vs MPP

Sweetspot #2 – Query Performance
SMP vs MPP

Sweetspot #3 – Exporting Data
Exporting Data

• Bulk Write data out:
  – Remote Table Copy Into SMP (RTC)
  – SQLCMD

• Polybase (Hadoop integration)
Remote table copy from PDW to SMP

- Creating a remote table on an regular SQL SMP Server utilizing all available cores 😊
- Export data from all distributions in Parallel

```sql
CREATE REMOTE TABLE
SMP_DB.dbo..LineItem_test AT
( 'Data Source = {Destination},1433 ;
 ...' )
AS SELECT
* FROM SQLRally2013.dbo.lineitem
```
SQLCMD

Writing data out to flatfile fast:

```sql
C:\Users\Administrator>sqlcmd -S "10.19.252.1,17001" -U hevandy -P [REDACTED] -I

2> :out b:\output.txt
2> go
1> select * from time
2> :out b:\time.txt
2> go
```
SMP vs MPP

Sweetspot #4 – Hadoop integration
PolyBase

Enables integrated query across Hadoop and relational data

- Single query for Relational and Hadoop data:
  - Parallel copy data into pdw/ out of pdw
  - Query the hadoop data seamlessly
  - Excel/ PowerBI

- Future expansion to other data sources
Parallel Data Transfers

This is PDW!
Familiar Tools To Analyze Structured/Unstructured Data

• Native Microsoft BI Integration to PDW
• Structured and unstructured data in same spreadsheet
• Widely adopted and familiar user tools
SMP vs MPP

Sweetspot #5 – ROLAP
Run SSAS Distinct Count queries directly on PDW

- new Msmdsrv.ini SSAS 2012 SP1 setting
- Set EnableRolapDistinctCountOnDataSource to ‘1’

```xml
<ProcessPlan>
  <MemoryLimit>65</MemoryLimit>
  <MemoryLimitErrorEnabled>1</MemoryLimitErrorEnabled>
  <MemoryAdjustConst>1000000</MemoryAdjustConst>
  <MemoryAdjustFactor>1</MemoryAdjustFactor>
  <CacheRowsSetRows>65536</CacheRowsSetRows>
  <CacheRowsSetToDisk>1</CacheRowsSetToDisk>
  <ForceMultiPass>0</ForceMultiPass>
  <DistinctBuffer>0x10000</DistinctBuffer>
  <OptimizeSchema>0x1</OptimizeSchema>
  <MaxTableDepth>8x100</MaxTableDepth>
  <EnableTableGrouping>1</EnableTableGrouping>
  <EnableRolapDimQueryTableGrouping>1</EnableRolapDimQueryTableGrouping>
  <EnableRolapDistinctCountOnDataSource>1</EnableRolapDistinctCountOnDataSource>
</ProcessPlan>
```

C:\Program Files\Microsoft SQL Server\MSAS11.MSSQLSERVER\OLAP\Config
SMP vs MPP

Sweetspot #6 – Ad-Hoc Query Concurrency
PDW V2 - Improved Workload management

- 10 compute node under load / 200 concurrent users
- PDW V1 completes 400 complex queries in 10 minutes
PDW V2 default resource class

- **2 node V2** Base unit under load / 200 concurrent users
- PDW V2 completes 1135 complex queries in 10 minutes
Additional Resources

» SQL Server Parallel Data Warehouse (PDW) Landing Page: www.microsoft.com/PDW


PDW is...

- A DWH appliance delivering a new type of value at lower cost
- All-In-One MPP Hub for Data Warehousing
- Deals easily with Complex Calcs, near realtime DW & unpredictable growth

- Steep learning curve
- ‘Out-of-the-box performance’ – MS Stamped’
- Easy to manage, integrate and highly available

- Complex query response from hours -> minutes/secs
Wrap up

... What would YOU do if you could query all your data & more within seconds?
THANK YOU!
For attending this session and
PASS SQLRally Nordic 2013, Stockholm

Henk.vanderValk@microsoft.com
Titles are set to 34 pt, Arial

Click to edit Master title style

- Level 1 text is 28 pt Arial
  - Level 2 text is 24 pt Arial
    - Level 3 text is 20 pt Arial
      - Level 4 text is 20 pt Arial
        - Level 5 text is 20 pt Arial
Notes (hidden)

• Some speakers may use this slide for hidden notes
• Please delete if you prefer not to use
• Please note you are also able to use notes section for each slide