Deploying Microsoft SQL Server
Big Data Clusters with HPE Synergy and Nimble storage

Maurice De Vidts
SQL Engineer

November 5, 2019
Session agenda

– What are SQL Server Big Data Clusters?
  – Key functionality and use cases

– HPE Design
  – HPE Synergy
  – HPE Nimble

– Deployment steps

– Scaling the solution

– Resources and call to action
What are SQL Server Big Data clusters?

New feature of SQL Server 2019

- Native, scalable HDFS storage and spark processing
- Virtual data hub for external data source integration (Polybase)
- Ability to manage and integrate both relational and big data in one platform
- AI and ML tools
SQL Server Big Data cluster scenarios

Data virtualization

- Analytics
- T-SQL
- Apps
- SQL Server External Tables
- Compute pools and data pools
  - Open database connectivity
  - NoSQL
  - Relational databases
  - HDFS

Managed SQL Server, Spark, and data lake

- Admin portal and management services
- Integrated AD-based security
- SQL Server
- Spark
- Scalable, shared storage (HDFS)

Complete AI platform

- REST API containers for models
- SQL Server ML Services
- Spark & Spark ML
- External data sources
- HDFS
SQL Server Big Data Cluster Architecture – at a glance

**Cluster**

- **Compute pool**
  - SQL Compute Node
  - SQL Compute Node

- **Storage pool**
  - Spark
  - SQL Server
  - HDFS Data Node

**External data sources**

- Microsoft SQL Server
- Teradata
- MongoDB
- Oracle

**IoT data**

**Persistent storage**

- Kubernetes pod

**Control Plane**

- Controller Svc
- Azure FSM Engine
- Kibana
- Grafana
- Configuration Store (SQL Server)
- Elastic Search
- InfluxDB

**Directly read from HDFS**

**SQL Server master instance**

**Custom apps**

**BI**

**Analytics**

**Hewlett Packard Enterprise**
HPE Design - Synergy

- Software Defined Infrastructure
- Flexible
- Easy to manage
- Scalable

- Ideal for any workload
- Fast deployment
- OneView management
- Automates infrastructure operations
HPE Design – Nimble storage

- Reliable
- Efficient storage
- Predictive
- Scalable

- Six Nines
- Store more TB
- Eliminates App Gap
Node sizing and mapping

— Rough node sizing

<table>
<thead>
<tr>
<th></th>
<th>Standard</th>
<th>Medium</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Instances</td>
<td>Cores</td>
</tr>
<tr>
<td>SQL Server master</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>Compute pool</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Data pool</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Storage pool w/ Spark</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Spark pool</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Storage pool w/o Spark</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Control/Management services</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>
High level deployment steps

HPE -> OpenShift Container Platform -> MS SQL BDC = Resilient, scalable and manageable data platform

Nimble CSI dynamic driver

Key deployment steps

1) Image Servers
2) Setup prerequisites (DNS, NTP etc)
3) Install docker
4) Install container management <Kubernetes, open shift>
5) Create the storage classes for each kid of storage device (SAN, Local, etc)
6) Edit BDC configuration files to reflect node count, storage sizes
7) Annotate and label nodes
8) Run AZDATA utility to install SQL Server Big Data Cluster
Resources and call to action!

– Tested HPE reference architectures and reference configurations

– Come visit our booth and talk to us about HPE, Big Data clusters or SQL performance in general
HPE Data and Analytics Reference Architectures

- HPE Reference Configuration for Red Hat OpenShift Container Platform on HPE Synergy Composable Infrastructure
- HPE Reference Architecture for Microsoft SQL Server 2017 on Linux with HPE 3PAR StoreServ Storage using compression
- HPE Reference Architecture for Microsoft SQL Server 2014 on HPE SimpliVity 380
- HPE Reference Architecture for Microsoft SQL Server 2016 Operational Analytics on the HPE ProLiant DL580 Gen9
- HPE Reference Architecture for Microsoft SQL Server 2016 on HPE Synergy
- HPE Reference Configuration for Microsoft SQL Server 2016 on HPE ProLiant for Microsoft Azure Stack
- Microsoft SQL Server 2017 OLTP Workloads on HPE MSA 2052 Storage

Check out all Data and Analytics Reference Architectures: www.hpe.com/info/dm-RA