The background of the top section is a collage of business and technology-related icons and images. On the left, there is a large, glowing cylindrical shape resembling a data storage or a futuristic container. In the center, a list of business terms is visible: Innovation, Branding, Solution, Marketing, Analysis, Ideas, Success, and Management. To the right, a hand is shown drawing a lightbulb and other diagrams on a transparent surface. Below the hand, there are various icons including a lightbulb, a puzzle piece, a pie chart, a network diagram, and a bar chart. The entire background has a warm, orange-brown color scheme.

Intelligent Data Management with SQL Server

Session: 16

Enhancements in SQL Server 2019

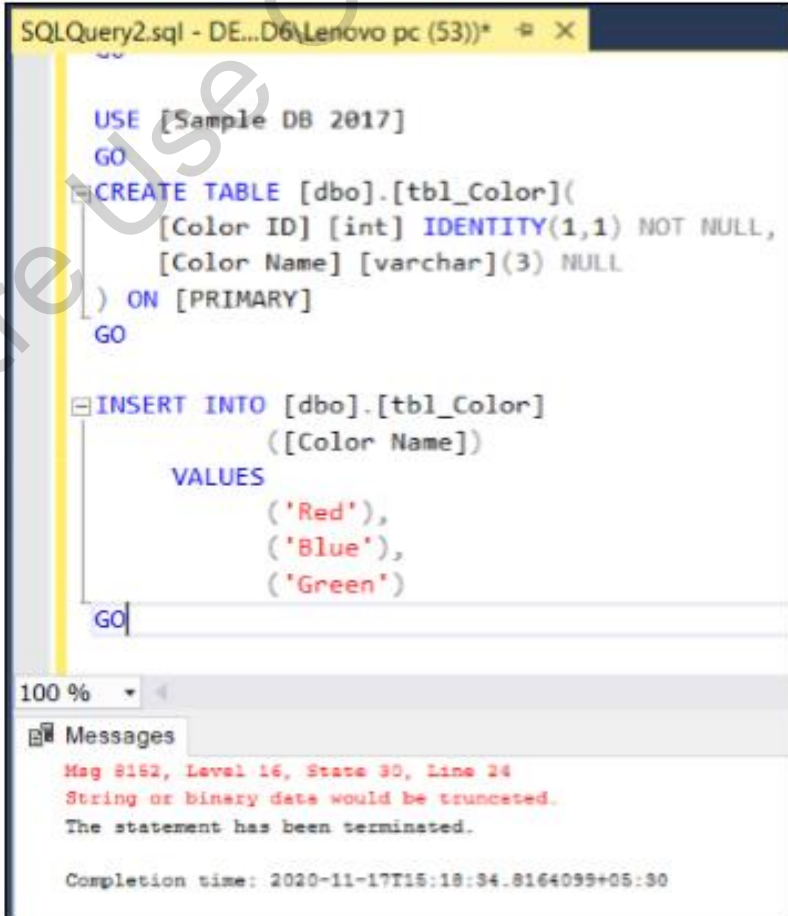
Objectives

- Describe the importance of new error messages
- Describe how to perform vulnerability assessment
- Explain Big Data clusters
- Explain how to use JSON data with SQL Server 2019

For Aptech Centre Use Only

Verbose Truncation Warnings 1-2

- Is one of the greatest features launched in SQL Server 2019.
- Saves a lot of time while importing, inserting, and updating huge amount of data.



The screenshot shows a SQL query window titled "SQLQuery2.sql - DE...D6\Lenovo pc (53))". The query contains the following SQL code:

```
USE [Sample DB 2017]
GO
CREATE TABLE [dbo].[tbl_Color](
  [Color ID] [int] IDENTITY(1,1) NOT NULL,
  [Color Name] [varchar](3) NULL
) ON [PRIMARY]
GO

INSERT INTO [dbo].[tbl_Color]
  ([Color Name])
VALUES
  ('Red'),
  ('Blue'),
  ('Green')
GO
```

Below the query window, the "Messages" pane displays the following error message:

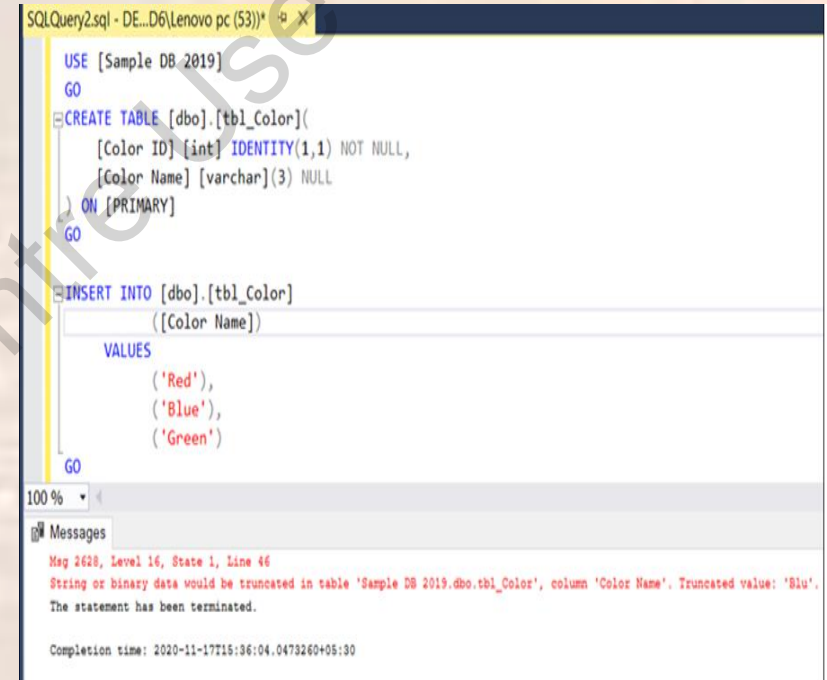
```
Msg 8162, Level 16, State 30, Line 24
String or binary data would be truncated.
The statement has been terminated.

Completion time: 2020-11-17T15:18:34.8164099+05:30
```

General Error Message

Verbose Truncation Warnings 2-2

- You can also set the database compatibility by selecting Database Properties and then, select Options and set Compatibility level.



The screenshot shows a SQL Server Enterprise Manager window titled 'SQLQuery2.sql - DE...D6\Lenovo pc (53)'. The main pane displays the following SQL script:

```
USE [Sample DB 2019]
GO
CREATE TABLE [dbo].[tbl_Color](
    [Color ID] [int] IDENTITY(1,1) NOT NULL,
    [Color Name] [varchar](3) NULL
) ON [PRIMARY]
GO
INSERT INTO [dbo].[tbl_Color]
    ([Color Name])
VALUES
    ('Red'),
    ('Blue'),
    ('Green')
GO
```

The bottom pane, titled 'Messages', displays the following error message:

```
Msg 2628, Level 16, State 1, Line 46
String or binary data would be truncated in table 'Sample DB 2019.dbo.tbl_Color', column 'Color Name'. Truncated value: 'Biu'.
The statement has been terminated.

Completion time: 2020-11-17T15:36:04.0473260+05:30
```

Detailed Error Message

Vulnerability Assessment 1-7

- SQL Vulnerability Assessment is an easy-to-configure service that can discover, track, and help you reverse or reduce potential database vulnerabilities.
 - You can use it to proactively improve your database security.
-
- Vulnerability Assessment is part of the Azure Defender for SQL offering, which is a unified package for advanced SQL security capabilities.
 - Vulnerability Assessment can be accessed and managed via central Azure Defender for SQL portal.

Vulnerability Assessment 2-7

- SQL Vulnerability Assessment is a service that provides visibility into your security state.
- Vulnerability Assessment includes actionable steps to resolve security issues and enhance your database security.

➤ It can help you to:

- Meet compliance requirements that require database scan reports
- Meet data privacy standards
- Monitor a dynamic database environment where changes are difficult to track

- The rules are based on Microsoft's best practices and focus on the security issues that present the biggest risks to your database and its valuable data.
- Results of the scan include actionable steps to resolve each issue and provide customized remediation scripts where applicable.

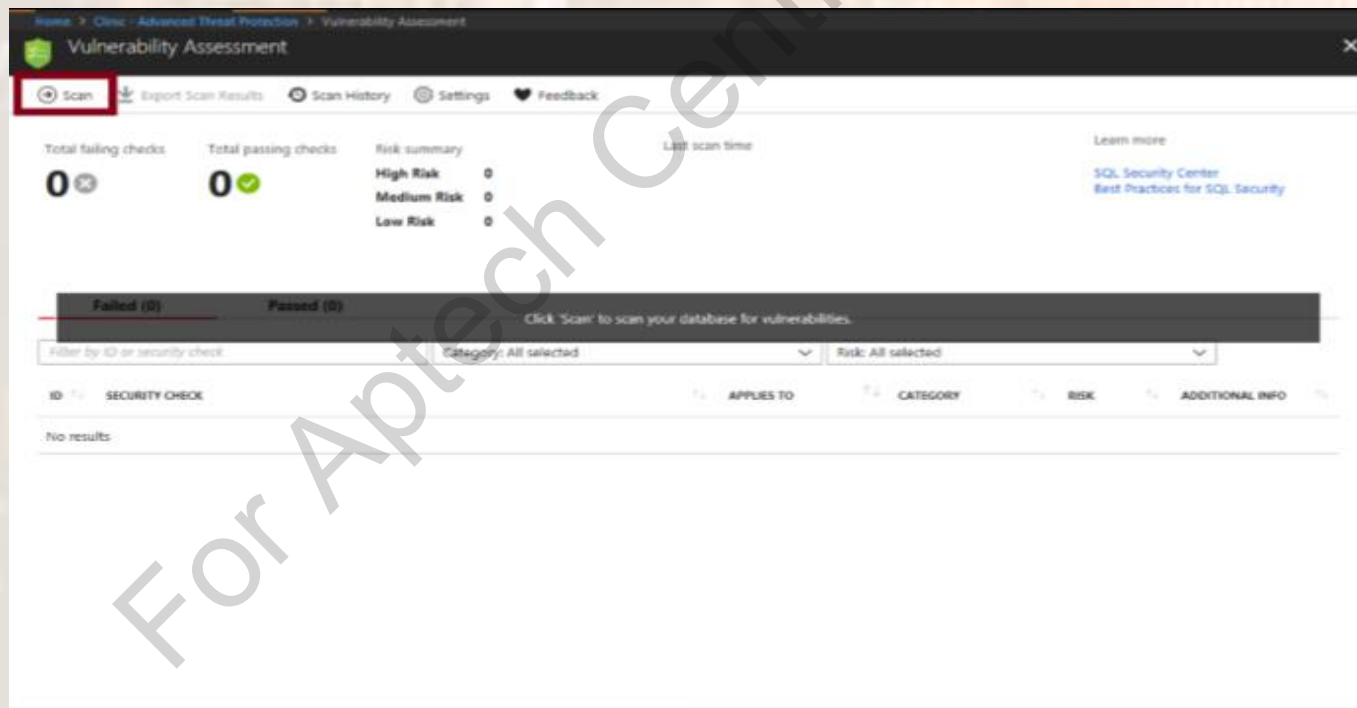
Vulnerability Assessment 3-7

You can customize an assessment report for your environment by setting an acceptable baseline for:

- ✓ Permission configurations
- ✓ Feature configurations
- ✓ Database settings

Following steps are used to implement vulnerability assessment:

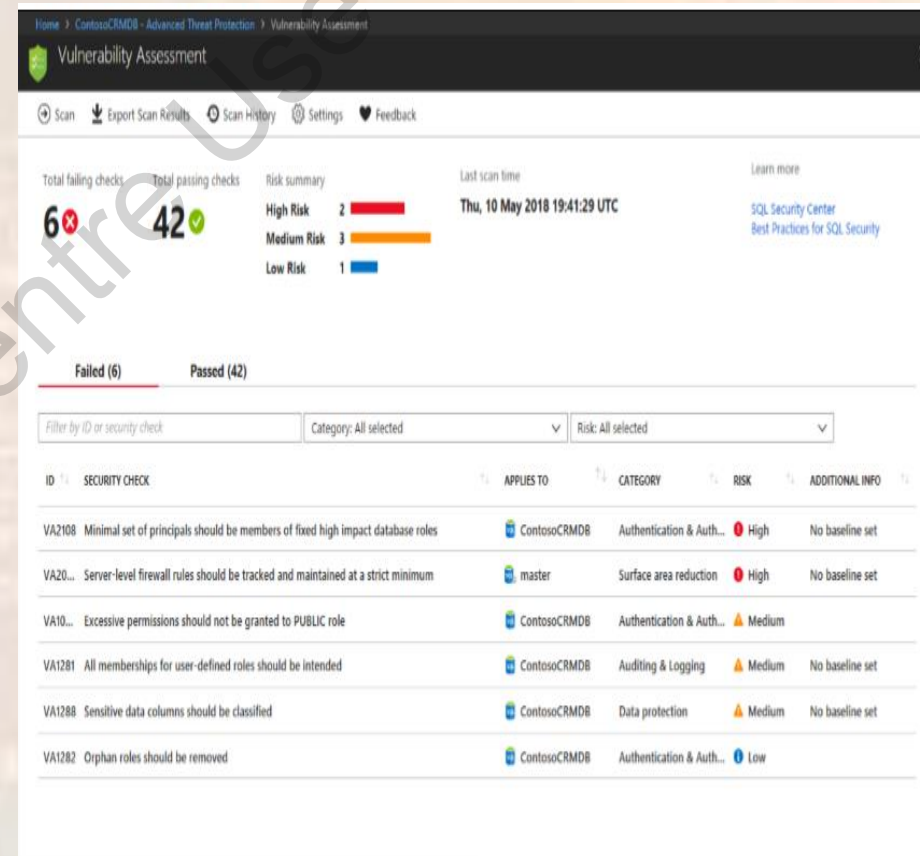
Step 1: Run a scan.



Vulnerability Assessment 4-7

Step 2: View the report

- When vulnerability scan is finished, scan report is automatically displayed in the Azure portal.
- Results include warnings on deviations from best practices and a snapshot of your security-related settings, such as database principals and roles and their associated permissions.



Vulnerability Report

Vulnerability Assessment 5-7

Step 3: Analyze the results and resolve issues

- Review your results and determine the findings in the report that are true security issues in your environment.
- Drill down to each failed result to understand the impact of the finding and why each security check failed.
- Use the actionable remediation information provided by the report to resolve the issue.

VA2108 - Minimal set of principals should be members of fixed high impact database roles

✓ Approve As Baseline ✕ Clear Baseline

NAME: VA2108 - Minimal set of principals should be members of fixed high impact database roles

RISK: High

STATUS: ✕ FAIL

DESCRIPTION: SQL Server provides roles to help manage the permissions. Roles are security principals that group other principals. Database-level roles are database-wide in their permission scope. This rule checks that a minimal set of principals are members of the fixed database roles.

IMPACT: Fixed database roles may have administrative permissions on the system. Following the principle of least privilege, it is important to minimize membership in fixed database roles and keep a baseline of these memberships. See <https://docs.microsoft.com/en-us/sql/relational-databases/security/authentication-access/database-level-roles> for additional information on database roles.

RULE QUERY: `SELECT user_name(sr.member_principal_id) as [Principal], user_name(sr.role_principal_id) as [Role], type_desc as [Principal]`

Run in Query Editor

MICROSOFT RECOMMENDATION: Empty Set

RESULTS:

IN BASELINE	PRINCIPAL	ROLE	PRINCIPAL TYPE	AUTHENTICATION TYPE
✕	michaelB	db_ddladmin	SQL_USER	NONE
✕	test1	db_ddladmin	DATABASE_ROLE	NONE

REMEDIATION: Remove members who should not have access to the database role

REMEDIATION SCRIPT: `ALTER ROLE [db_ddladmin] DROP MEMBER [michaelB]
ALTER ROLE [db_ddladmin] DROP MEMBER [test1]`

Run in Query Editor

Security Issues

Vulnerability Assessment 6-7

Step 4: Set your baseline

- As you review your assessment results, you can mark specific results as being an acceptable baseline.
- Results that match the baseline are considered as passing in subsequent scans.
- After you have established your baseline security state, Vulnerability Assessment only reports on deviations from the baseline.

VA2108 - Minimal set of principals should be members of fixed high impact database roles

☒ Approve As Baseline ☐ Clear Baseline

NAME: VA2108 - Minimal set of principals should be members of fixed high impact database roles

RISK: High

STATUS: ✗ FAIL

DESCRIPTION: SQL Server provides roles to help manage the permissions. Roles are security principals that group other principals. Database-level roles are database-wide in their permission scope. This rule checks that a minimal set of principals are members of the fixed database roles.

IMPACT: Fixed database roles may have administrative permissions on the system. Following the principle of least privilege, it is important to minimize membership in fixed database roles and keep a baseline of these memberships. See <https://docs.microsoft.com/en-us/sql/relational-databases/security/authentication-access/database-level-roles> for additional information on database roles.

RULE QUERY: `SELECT user_name(sr.member_principal_id) as [Principal], user_name(sr.role_principal_id) as [Role], type_desc as [Principal]`

MICROSOFT RECOMMENDATION: Empty Set

RESULTS:

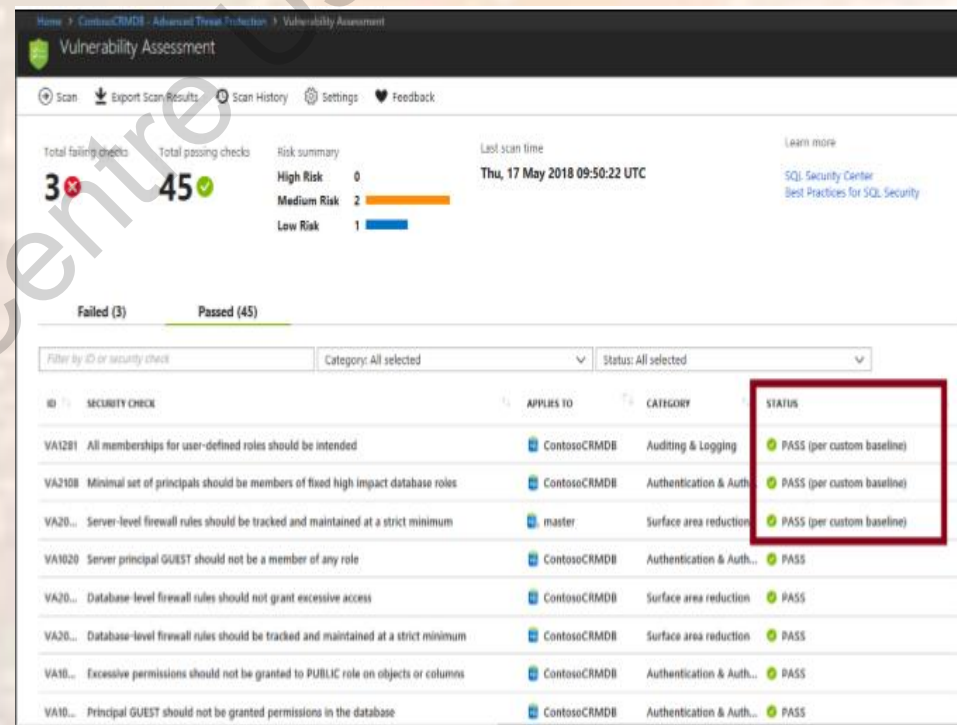
IN BASELINE	PRINCIPAL	ROLE	PRINCIPAL TYPE	AUTHENTICATION TYPE
✗	michaelB	db_ddladmin	SQL_USER	NONE
✗	test1	db_ddladmin	DATABASE_ROLE	NONE

Approve as Baseline

Vulnerability Assessment 7-7

Step 5: Run a new scan to see your customized tracking report

- After you finish setting up your Rule Baselines, run a new scan to view the customized report.
- Vulnerability Assessment now reports only the security issues that deviate from your approved baseline state.



Customize Scan Report

Big Data Clusters 1-6

- In SQL Server 2019 Big Data Clusters allow to deploy scalable clusters of SQL Server, Spark, and Hadoop Distributed File System (HDFS) containers running on Kubernetes.

Popular uses of Big Data Clusters:

Deploy scalable clusters of SQL Server, Spark, and HDFS containers running on Kubernetes

Read, write, and process big data from Transact-SQL or Spark

Easily combine and analyze high-value relational data with high-volume big data

Query external data sources

Store big data in HDFS managed by SQL Server

Query data from multiple external data sources through the cluster

Use the data for AI, machine learning, and other analysis tasks

Deploy and run applications in Big Data Clusters • Virtualize data with PolyBase

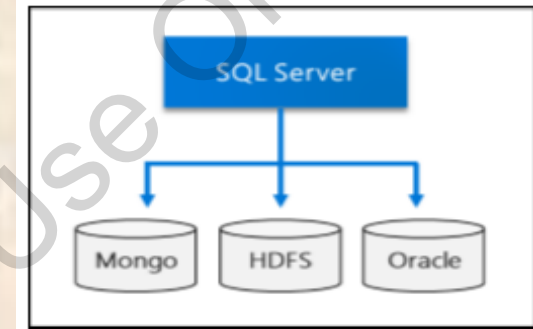
Query data from external SQL Server, Oracle, Teradata, MongoDB, and ODBC data sources with external tables

Provide high availability for the SQL Server master instance and all databases by using Always On availability group technology

Big Data Clusters 2-6

Data Virtualization

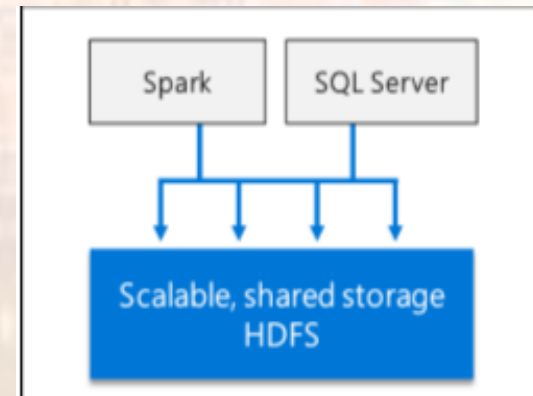
- SQL Server Big Data Clusters can query external data sources without moving or copying the data.



Data Virtualization

Data Lake

- Data Lake is a storage repository that holds a huge amount of raw data in its native format.
- It is a scalable HDFS storage pool.

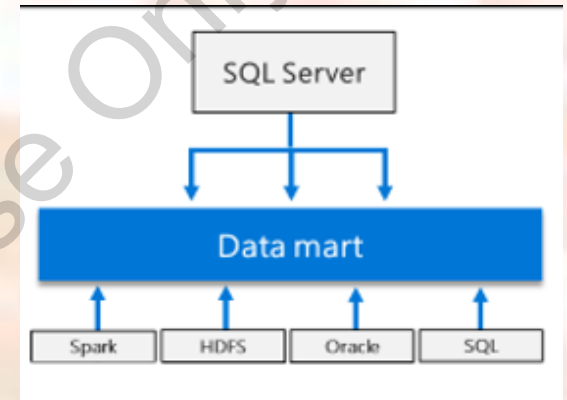


Data Lake

Big Data Clusters 3-6

Scale-out data mart

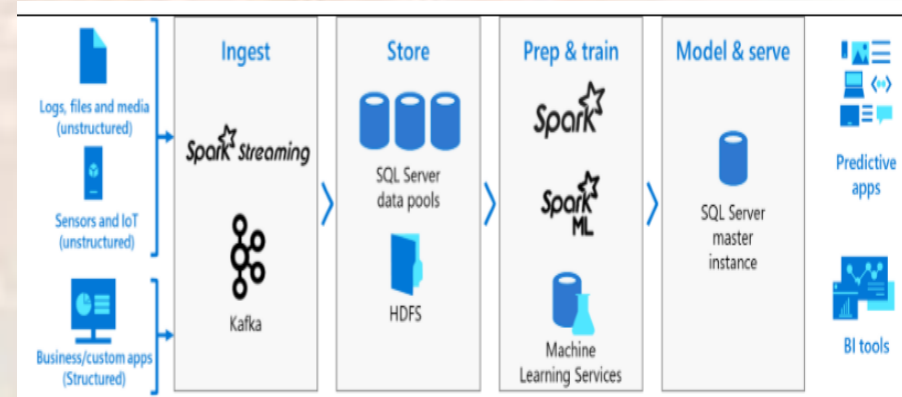
- Provides scale-out compute and storage to improve the performance of analyzing any data.



Data Mart

Integrated AI and Machine Learning

- Enables AI and machine learning tasks on the data stored in HDFS storage pools and the data pools.

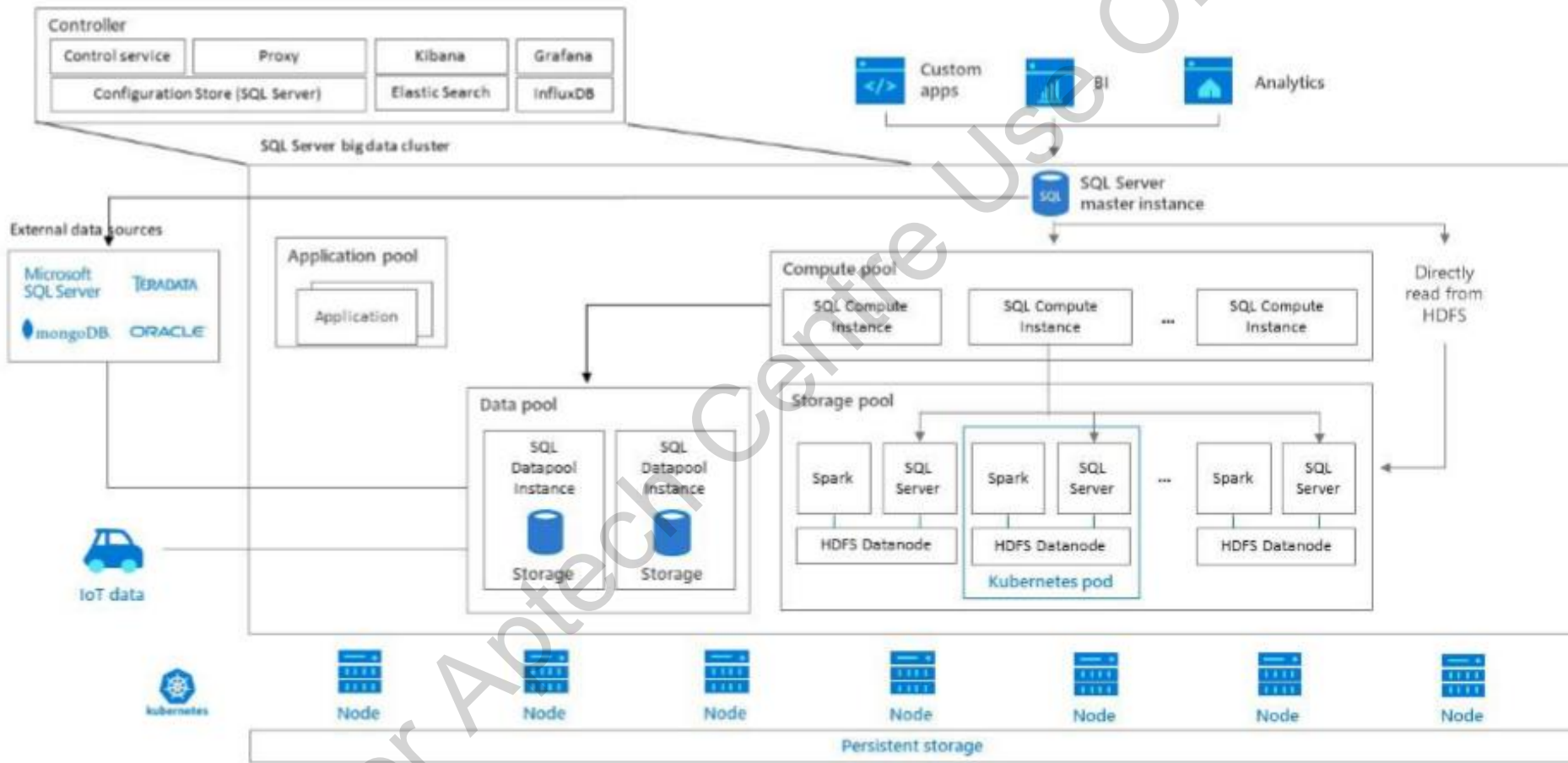


Big Data Clusters 4-6

Kubernetes Terms

Term	Description
Cluster	A Kubernetes cluster is a set of machines, known as nodes. One node controls the cluster and is designated the master node; the remaining nodes are worker nodes. The Kubernetes master is responsible for distributing work between the workers and for monitoring the health of the cluster.
Node	A node runs containerized applications. It can be either a physical machine or a virtual machine. A Kubernetes cluster can contain a mixture of physical machine and virtual machine nodes.
Pod	A pod is the atomic deployment unit of Kubernetes. A pod is a logical group of one or more containers-and associated resources-required to run an application. Each pod runs on a node; a node can run one or more pods. The Kubernetes master automatically assigns pods to nodes in the cluster.

Big Data Clusters 5-6



SQL Server Big Data Cluster

Big Data Clusters 6-6

Controller

Controller provides management and security for the cluster. It contains the control service, the configuration store, and other cluster-level services such as Kibana, Grafana, and Elastic Search.

Compute Pool

Compute pool provides computational resources to the cluster. It contains nodes running SQL Server on Linux pods. The pods in the compute pool are divided into SQL Compute instances for specific processing tasks.

Data Pool

Data pool is used for data persistence and caching. The data pool consists of one or more pods running SQL Server on Linux. It is used to ingest data from SQL queries or Spark jobs. SQL Server big data cluster data marts are persisted in the data pool.

Storage Pool

Storage pool consists of storage pool pods comprised of SQL Server on Linux, Spark, and HDFS. All the storage nodes in a SQL Server big data cluster are members of an HDFS cluster.

JSON data in SQL Server 1-4

- JSON is a textual data format that is used for exchanging data in modern Web and Mobile applications.
- JSON is also used to store unstructured data in log files or NoSQL databases such as Microsoft Azure Cosmos DB.
- Many REST Web services return results that are formatted as JSON text or accept data that is formatted as JSON.

Through SQL Server built-in functions and operators, you can:

- Parse JSON text and read or modify values
- Transform arrays of JSON objects into table format
- Run any Transact-SQL query on the converted JSON objects
- Format the results of Transact-SQL queries in JSON format

JSON data in SQL Server 2-4

Modify JSON Values

- To modify parts of JSON text, JSON_MODIFY (Transact-SQL) function is used to update the value of a property in a JSON string and return the updated JSON string.

```
{"info":{"address":[{"town":"Belgrade"}, {"town":"London"}, {"town":"Madrid"}]}}
```

OPENJSON to Convert JSON to rowset

- Custom query language is not required to query JSON in SQL Server. To query JSON data, standard TSQL can be used.

Results		Messages			
	id	firstName	lastName	age	dateOfBirth
1	2	John	Smith	25	NULL
2	5	Jane	Smith	NULL	2005-11-04 12:00:00.0000000

JSON Variable to ROWSET

JSON data in SQL Server 3-4

- JSON documents may have sub-elements and hierarchical data that cannot be directly mapped into the standard relational columns. In this case, you can flatten JSON hierarchy by joining parent entity with sub-arrays.

Results

Messages

	id	firstName	lastName	age	dateOfBirth	skills	skill
1	2	John	Smith	25	NULL	NULL	NULL
2	5	Jane	Smith	NULL	2005-11-04 12:00:00.0000000	["SQL", "C#", "Azure"]	SQL
3	5	Jane	Smith	NULL	2005-11-04 12:00:00.0000000	["SQL", "C#", "Azure"]	C#
4	5	Jane	Smith	NULL	2005-11-04 12:00:00.0000000	["SQL", "C#", "Azure"]	Azure

Result of OPENJSON function calls

JSON data in SQL Server 4-4

Export SQL Server Data to JSON

- You can format SQL Server data or results of SQL queries as JSON by adding the FOR JSON clause to a SELECT statement. The FOR JSON delegates the formatting of JSON output from your client applications to SQL Server.



```
[
  {
    "BusinessEntityId":1,
    "info":{"name":"Ken","surname":"Sánchez"},
    "dob":"2009-01-07T00:00:00"
  },
  {
    "BusinessEntityId":2,
    "info":{"name":"Terri","surname":"Duffy"},
    "dob":"2008-01-24T00:00:00"
  },
  {
    "BusinessEntityId":3,
    "info":{"name":"Roberto","surname":"Tamburello"},
    "dob":"2007-11-04T00:00:00"
  },
  {
    "BusinessEntityId":4,
    "info":{"name":"Rob","surname":"Walters"},
    "dob":"2007-11-28T00:00:00"
  },
  {
    "BusinessEntityId":5,
    "info":{"name":"Gail","surname":"Erickson"},
    "dob":"2007-12-30T00:00:00"
  },
  {"BusinessEntityId":6,"info":{"name":"Jossef","surname":"Goldberg"},"dob":"2013-12-16T00:00:00"}, {"BusinessEntityId":7,"info":{"name":"Annette","surname":"Baker"},"dob":"2006-09-19T00:00:00"}, {"BusinessEntityId":8,"info":{"name":"Diana","surname":"Adams"},"dob":"2006-07-15T00:00:00"}, {"BusinessEntityId":9,"info":{"name":"Luis","surname":"Gonzalez"},"dob":"2006-04-08T00:00:00"}, {"BusinessEntityId":10,"info":{"name":"Margaret","surname":"Smith"},"dob":"2005-04-08T00:00:00"}, {"BusinessEntityId":11,"info":{"name":"Michael","surname":"Baker"},"dob":"2005-02-15T00:00:00"}, {"BusinessEntityId":12,"info":{"name":"Patricia","surname":"Johnson"},"dob":"2005-02-04T00:00:00"}, {"BusinessEntityId":13,"info":{"name":"Robert","surname":"Johnson"},"dob":"2005-02-04T00:00:00"}, {"BusinessEntityId":14,"info":{"name":"Susan","surname":"Patten"},"dob":"2005-02-04T00:00:00"}, {"BusinessEntityId":15,"info":{"name":"Tina","surname":"Crawford"},"dob":"2005-02-04T00:00:00"}, {"BusinessEntityId":16,"info":{"name":"Yvonne","surname":"Jensen"},"dob":"2005-02-04T00:00:00"}
]
```

Exported JSON Data

Summary

- Verbose Truncation Warnings makes troubleshooting easy by displaying additional and exact details in error messages.
- SQL Vulnerability Assessment is a service that provides visibility into your security state.
- Vulnerability report lists how many issues were found and their respective severities.
- SQL Server Big Data Clusters provide flexibility in interacting with Big Data.
- Kubernetes is an open source container orchestrator.
- Kubernetes is responsible for the state of the SQL Server Big Data Clusters.
- JSON is used to store unstructured data in log files or NoSQL databases such as Microsoft Azure Cosmos DB.