New Features

Vertica 8.0.0 New Features and Changes

Read the topics in this section for information about new and changed functionality in Vertica 8.0.0.

Supported Platforms

This section contains information on updates to supported platforms for Vertica Analytics Platform 8.0.

More Details

For complete information on platform support see Vertica 8.0.x Supported Platforms

Software Download

For more information on changes to the operating system in the Red Hat 7 release, see the Red Hat Enterprise Linux 7 documentation.

FIPS Supported Platforms

HPE Vertica uses a certified OpenSSL FIPS 140-2 cryptographic module to meet the security standards set by the National Institute of Standards and Technology (NIST) for Federal Agencies in the United States or other countries.

FIPS-enabled Vertica requires the following:

- Red Hat Linux 6.6, a FIPS compliant operating system
- OpenSSL 1.0.1e
- A user-generated certificate signed by an approved Certificate Authority
- TLS 1.2 to support the server-client connection for a FIPS-enabled system

Supported Drivers

Vertica supports the following client drivers for FIPS-compliance:

- vsql
- ODBC

For more information see Federal Information Processing Standard Overview.

Upgrade and Installation

This section contains information on updates to upgrading or installing Vertica Analytics Platform 8.0.

More Details

For information see Installing Vertica.

Vertica Adds Support for SQL Server 2016 and Visual Studio 2015

The Vertica Client Drivers and Tools for Windows installer now includes support for SQL Server 2016 and Visual Studio 2015.

For more information, refer to The Microsoft Connectivity Pack for Windows.

The Vertica FIPS Client

HPE Vertica 8.0. implements FIPS. If your host is FIPS-enabled, you can communicate with a FIPS server. To do so, download and install the new FIPS client. The FIPS client supports ODBC and vsql. It is available in 64-bit only.

For information specific to the FIPS client installation, refer to Installing the FIPS Client Driver for ODBC and vsql in the Connecting to Vertica Guide.

For general information on all clients, their platforms, and compatibility, refer to Vertica 8.0.x Client Drivers in the Supported Platforms Guide.

Advanced Analytics Package - Machine Learning Package now Included in Server Install

The Advanced Analytics / Machine Learning package is now included in the server install. No package installation is required.

However, if you previously installed the Advanced Analytics package then be sure to run the machine learning upgrade script after you upgrade Vertica. For more details see Upgrading the Advanced Analytics Package 7.2.x to Machine Learning 8.0.x.

For general details about Machine Learning for Predictive Analytics see Machine Learning for Predictive Analytics.

Client Connectivity

This section contains information on updates to connection information for Vertica Analytics Platform 8.0.

More Details

For more information see Connecting to Vertica.

User-based User Client Connections

Using the parameter MAXCONNECTIONS, HPE Vertica limits the number of concurrent connections a user can have to the server. These connection limits are enforced cluster-wide by default. This allows better management as connection limits do not change as the cluster grows.

You can also implement client connection limits at the node level. This tends to increase performance as it eliminates inter-node communication.

In addition, the parameter IDLESESSIONTIMEOUT limits the amount of time a client session can remain idle before the system automatically closes the session.

For more information see Managing Client Connections.

Security and Authentication

This section contains information on updates to security and authentication features for Vertica Analytics Platform 8.0.

More Details

For more information see Security and Authentication.

Federal Information Processing Standard 140-2

HPE Vertica 8.0 is compliant with the Federal Information Processing 140-2 standard (FIPS). FIPS is a U.S government security standard that governs cryptographic modules.

FIPS implementation affects:

- Client Installation
- OpenSSL
- Password encryption (no MD5 use allowed)

Dynamically linked OpenSSL is a requirement for FIPS-enabled systems. In non-FIPS systems, this linking allows you a simple way to upgrade to new OpenSSL versions or patches in response to security bugs.

For more information see Federal Information Processing Standard Overview

Query Management

This section contains information on updates to Query Management for Vertica Analytics Platform 8.0.

More Details

For more information see Managing Queries.

New Support for Directed Queries

Directed queries now support queries that invoke:

- Analytic functions
- User-defined extensions (UDXs)

Directed queries also now support `SELECT...FOR UPDATE` queries whose predicate resolves to false.

Tables

This section contains information on updates to flex and regular tables information for Vertica Analytics Platform 8.0.

More Details

For more information see Managing Tables and Using Flex Tables.

Flex Table Keys Functions

The default for the database or session-based configuration parameter, `EnableBetterFlexTypeGuessing`, is now set to 1 (ON). This changes the default behavior for flex tables. When you use the COMPUTE_FLEXTABLE_KEYS or COMPUTE_FLEXTABLE_KEYS_AND_BUILD_VIEW functions, Vertica determines the type for each non-string type key. All data types are listed in the `data_type_guess` column of the flex `_keys` table. See Computing Flex Table Keysin the Using Flex Tables guide.

To resume the previous behavior, change the `EnableBetterFlexTypeGuessing` configuration parameter to 0 (OFF). The functions then determine all flex table keys as string types (`[LONG]VARCHAR`) or (`[LONG] VARBINARY`). For more information, see General Parameters and Setting Configuration Parameter Values in the Administrator's Guide.

Constraints

This section contains information on updates to constraints for Vertica Analytics Platform 8.0.

More Details

For more information see About Constraints.

Enhanced Reporting of Constraint Violations

Prior to this release, for enabled constraints, Vertica would report one constraint violation before rolling back the SQL statement that caused the violation.

For enabled *primary key* and *unique* constraints, Vertica now reports multiple violations before rolling back.

For enabled *check* constraints, Vertica continues to report one check constraint violation before roll back, but there are exceptions.

> **Note:** Vertica continues to roll back the offending statement upon finding one or more constraint violations of an enabled primary key, unique, or check constraint.

For more information on this new enhanced error reporting, see Reporting Constraint Violations.

Text Search

This section contains information on updates to Text Search in Vertica Analytics Platform 8.0.

More Details

For information see Using Text Search.

Multilanguage Tokenizer for Text Search

The new ICU Tokenizer for text search supports multiple languages. When creating text indexes, you can use this tokenizer to identify word boundaries in languages other than English, including Asian languages that do not separate words by whitespace.

For more information, see ICU Tokenizer.

Geospatial Analytics

This section contains information on updates to Vertica Geospatial Analytics for Vertica Analytics Platform.

More Details

For more information see Geospatial Analytics.

Geospatial Analytics Included in Vertica Installer

The Vertica installation package now installs the Geospatial Analytics package automatically.

STV_Intersect Accepts Two Separate Columns

You can now perform spatial joins without transforming your spatial data to GEOMETRY or GEOGRAPHY data types. STV_Intersect now accepts two columns (x and y) to perform spatial joins, allowing you to do exploratory data analysis more quickly.

For more information, see STV_Intersect Scalar Function and STV_Intersect Transform Function.

Transforming Your Spatial Data Between SRIDs

You can now transform your spatial data between spatial reference identifiers (SRID) WGS84 and Web Mercator using ST_Transform. With ST_Transform, you can convert your spatial data from one SRID to another.

For more information, see ST_Transform.

Loading Data

This section contains information about loading data Vertica Analytics Platform 8.0.

More Details

For more information see:

Vertica Library for Amazon Web Services

Using Flex Tables

COPY

Expanded Load Method Support

Vertica now offers expanded support for specifying load methods AUTO, DIRECT, and TRICKLE. You can specify load methods at two levels:

- Table metadata: `CREATE TABLE` and `ALTER TABLE` can now specify default loading behavior for all DML operations.

- DML operations: `COPY`/`COPY FROM VERTICA`, `INSERT`, `MERGE`, and `UPDATE` can specify the load method for a given operation.

For details, see Choosing a Load Method in the Administrator's Guide.

Apportioned Load Expansion

This release expands support for apportioned load to parallelize load operations. The built-in delimited parser can now apportion a load automatically, if the source is available on multiple nodes and can be divided. Using apportioned load can improve the performance of loads using globs (multiple files with wildcards) or containing large files. For more information, see Using Parallel Load Streams.

A new COPY parameter, `ON (nodeset)`, allows you to designate particular nodes to share a load operation. This parameter provides finer control when performing concurrent load operations.

Insert Data into Flex Tables

You can load data into a Vertica flex table using a standard INSERT statement, specifying data for one or more columns. When you use INSERT, Vertica populates any materialized columns and stores the VMap data in the __raw__ column .

Vertica provides two ways to use INSERT with flex tables:

- INSERT ... VALUES
- INSERT ... SELECT

For details, see Inserting Data into Flex Tables in Using Flex Tables.

Apache Spark Connector

HPE Vertica 8.0 introduces a new connector to transfer data between Vertica and Apache Spark. The Vertica Connector for Apache Spark allows you to :

- Transfer data from Spark DataFrames to Vertica Tables (using HDFS as an intermediate staging location).
- Transfer data from Vertica tables to Spark DataFrames.
- Transfer data from Vertica tables to Spark Resilient Distributed Datasets (RDDs).

The Vertica Connector for Apache Spark is a JAR library available for download from my.vertica.com.

For more information see Integrating with Apache Spark

Configuration Parameters

This section contains information on updates to Configuration Parameters in Vertica Analytics Platform8.0.

More Details

For more information, see Configuration Parameters.

EnableBetterFlexTypeGuessing

The `EnableBetterFlexTypeGuessing` configuration parameter now has a default value of 1 (ON). The new default value changes the default behavior when using the flex table COMPUTE_FLEXTABLE_KEYS or COMPUTE_FLEXTABLE_KEYS_AND_BUILD_VIEW functions. You can set the parameter at the database or session level.

See also General Parameters and the Administrator's Guide.

EnableTMOnRecoveringNode

The EnableTMOnRecoveringNode parameter allows Tuple Mover to perform moveout and mergeout activities on nodes with a node state of RECOVERING. Enabling Tuple Mover reduces the number of ROS containers generated during recovery. Having fewer than 1024 ROS containers per projection allows Vertica to maintain optimal recovery performance.

For more information, see Tuple Mover Parameters.

Setting Numeric Precision

Vertica allows you to add implicit digits for numeric data type fields. Adding implicit digits helps you avoid overflow. Overflow can occur when you use the function SUM, SUM_FLOAT, or AVG in your queries, or the functions SUM or SUM_FLOAT in your LAPs. To add implicit digits, you set two Vertica parameters.

- The parameter AllowNumericOverflow, by default, is set to 1 (true). The default allows silent numeric overflow. Set to 0 (false) to turn off silent numeric overflow.
- The parameter NumericSumExtraPrecisionDigits has a default value of 6. This parameter applies only when the AllowNumericOverflow parameter is set to 0 (false). You can set this parameter to a value between 0 and 20. When set, the implicit digits extend your NUMERIC field, allowing Vertica to store larger numbers without overflow.

Hewlett Packard Enterprise recommends that you turn off silent numeric overflow if you expect to exceed the precision specified in your DDL. Adding implicit digits can affect performance, so consider adding only what you need to store expected results.

For more information on how Vertica processes NUMERIC data types when you use the functions SUM, SUM_FLOAT, and AVG, see:

Numeric Data Type Overflow with SUM, SUM_FLOAT, and AVG.

For more information about the new parameters, see Numeric Precision Parameters.

MaxTieredPoolScale

New configuration variable `MaxTieredPoolScale` specifies the threshold for allocating chunks of memory from the system or dedicated pools for Vertica system components—-for example, for WOS containers. Above this threshold, Vertica allocates memory from the system (malloc), and returns it when no longer needed. Below this threshold, Vertica allocates memory from dedicated pools (talloc), which manage smaller memory allocations more efficiently.

For details, see General Parameters in the Administrator's Guide.

LDAPLinkStopIfZeroUsers

The LDAPLinkStopIfZeroUsers parameter enables or disables the shutdown of LDAPLink synchronization if no users are found in LDAP.

Setting this parameter to true (1) prevents inadvertent dropping of Vertica users if no corresponding LDAP users are found in the LDAP database during an LDAPLink session.

For more information see Troubleshooting LDAP Link Issues.

Management Console

This section contains information on updates to the Management Console for Vertica Analytics Platform 8.0.

More Details

For more information see Using Management Console.

Viewing Continuous Load Jobs in MC

Management Console (MC) introduces the ability to view a history of continuous Kafka loading jobs in MC. The Data Load Activity page now displays all loading jobs,

including those initiated in Management Console with Amazon S3, and those created using the MC extended monitoring feature.

For more about viewing data loads in MC, see Viewing Load History.

Extended Monitoring in Management Console

Management Console (MC) introduces the ability to gather more long-term metrics in an external MC storage database. Enabling extended monitoring:

- Allows MC charts to display longer time frames
- Reduces the effect of monitoring on your database's performance
- Allows you to view your MC charts while the monitored database is down

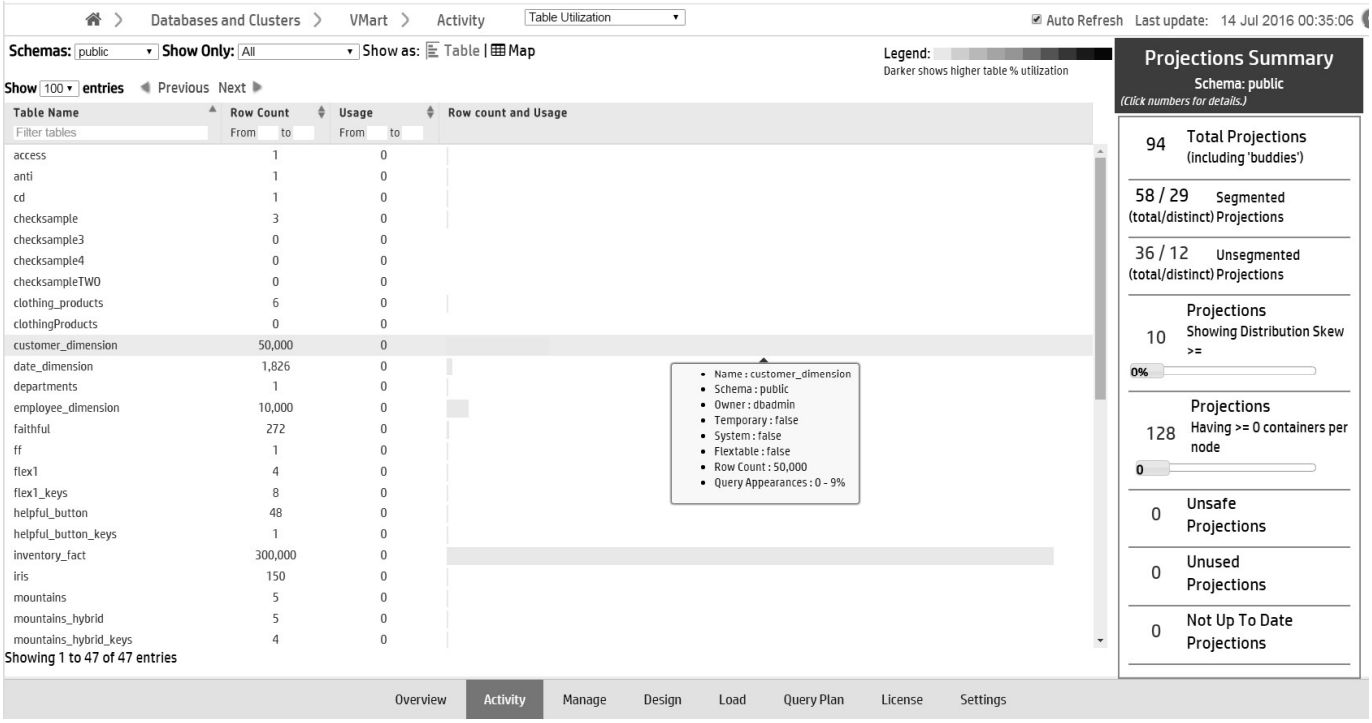
MC uses Kafka to stream your database's monitoring data to the dedicated storage database.

For more details, see Extended Monitoring.

Viewing Table Utilization in Management Console

Management Console (MC) introduces a new way to visualize tables in your database. As an alternative to the Table Treemap chart, MC can now also display table information in a sortable table chart on the Table Utilization activity page.

Use this chart if your database has a large number of tables that crowd the graphical Treemap visualization, or to more easily search for tables you are interested in.



For details, see Monitoring Table Utilization and Projections.

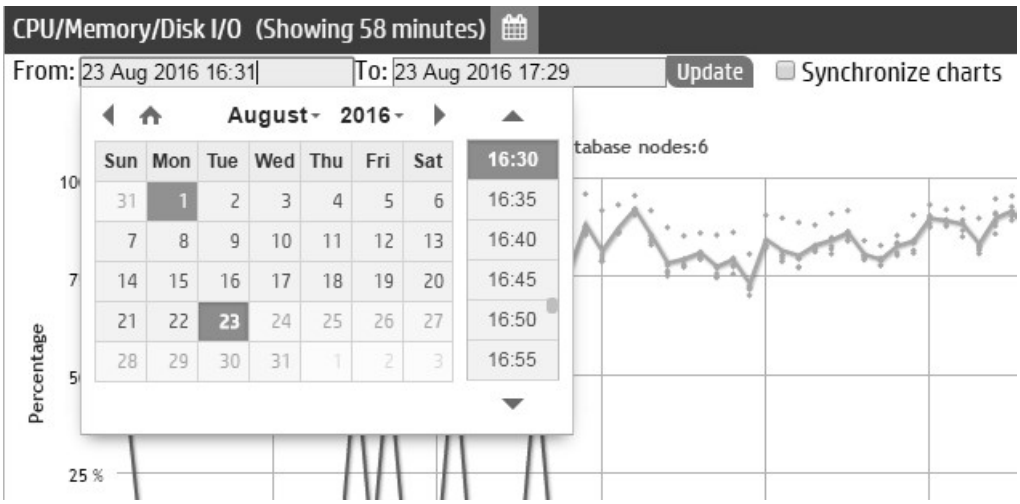
Chart Enhancements in Management Console

Management Console (MC) introduces improvements to the Queries and Queries Statistics modules on the Overview page. In addition, MC introduces the ability to specify longer time frames on select graphs on your database's Overview and Activity pages.

To read more about viewing charts in MC, see Viewing the Overview Page and Monitoring System Resources.

Longer Time Frames

You can specify longer time frames for charts in Management Console that display a calendar icon in their title bars. Click the calendar icon to specify the time frame for that module.



Expandable Overview Charts

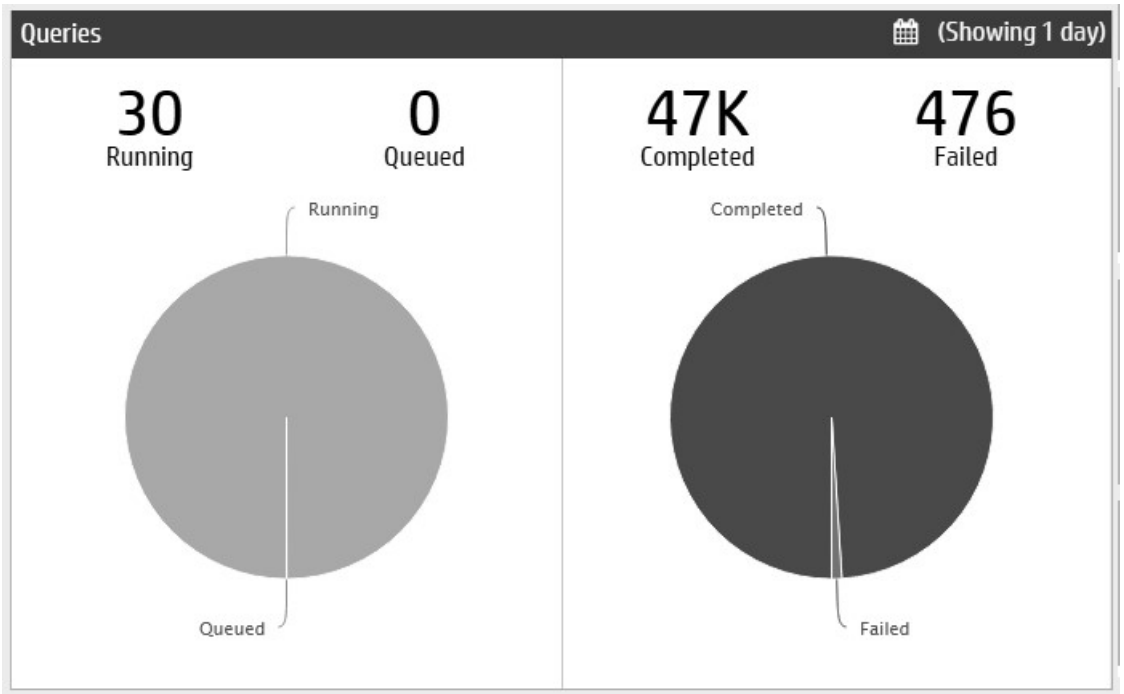
You can expand select charts on the Overview page to view them in larger windows. Use this feature to view large amounts of data across longer time frames.

Click the expand icon in a chart's title bar to do so:

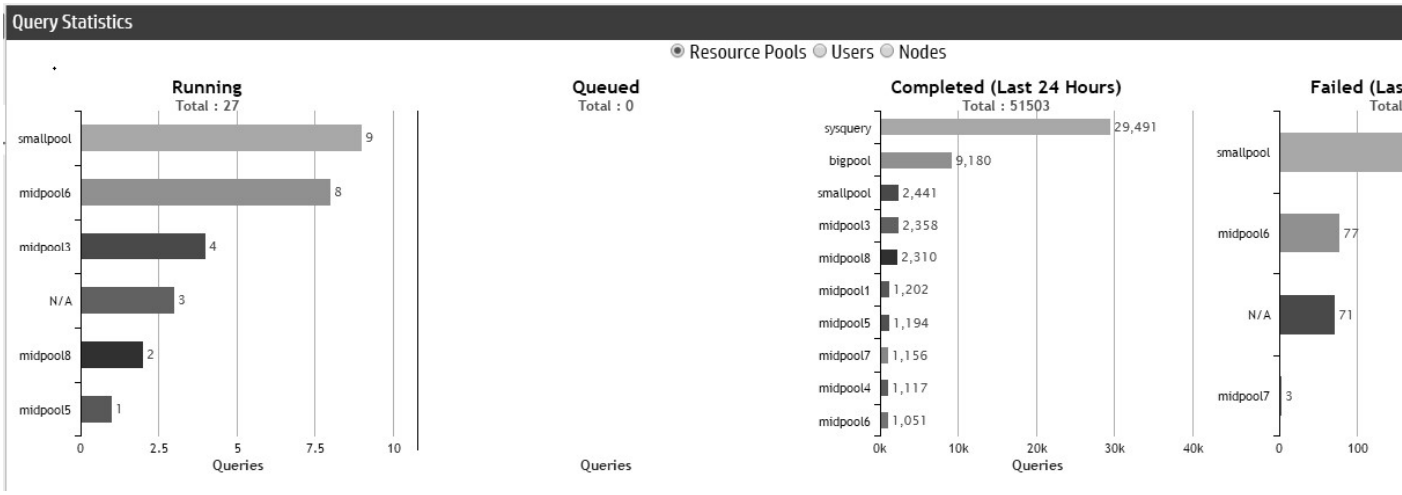
Queries and Query Statistics Modules

Two charts on your database's Overview page have been improved:

- The **Queries** module (on the Status Summary tab) now displays its data in pie charts. These charts represent your running, queued, completed, and failed queries.



- The **Query Statistics** module (on the Query Synopsis tab) now displays your data in bar graphs. This chart groups your queries by resource pools, users, or nodes.



System Table Updates

This section contains information on updates to System Tables for Vertica Analytics Platform 8.0.

More Details

For more information see Vertica System Tables.

New System Tables

Vertica Analytics Platform added the following Monitor system tables:

- LDAP_LINK_EVENTS
- LOAD_SOURCES
- MODELS
- REBALANCE_OPERATIONS
- REPARENTED_ON_DROP

SQL Functions and Statements

This section contains information on updates to SQL Functions and Statements for Vertica Analytics Platform 8.0.

More Details

For more information see the SQL Reference Manual.

Machine Learning Functionality

In Vertica 7.2.2, Machine Learning for Predictive Analytics was introduced as a separate add-on package for Vertica. As of Vertica 8.0.0, the Machine Learning functionality introduced in the Machine Learning for Predictive Analytics package comes standard with the Vertica Analytics Platform. You no longer need to install a separate package to use the Machine Learning functionality.

If you installed the Advanced Analytics package in 7.2.x and are upgrading to 8.0.x, then make sure to run the backwards compatibility script. For more information see Upgrading the Advanced Analytics Package 7.2.x to Machine Learning 8.0.x.

Function Name Changes

The following functions have new names:

7.2.x Function Name	8.0.x Function Name
v_ml.logisticReg	LOGISTIC_REG

7.2.x Function Name	8.0.x Function Name
v_ml.linearReg	LINEAR_REG
v_ml.kmeans	KMEANS
v_ml.deleteModel	DELETE_MODEL
v_ml.renameModel	RENAME_MODEL
v_ml.normalize	NORMALIZE
v_ml.summaryLinearReg	SUMMARIZE_MODEL
v_ml.summaryLogisticReg	SUMMARIZE_MODEL
v_ml.summaryKmeans	SUMMARIZE_MODEL
v_ml.predictLogisticReg	PREDICT_LOGISTIC_REG
v_ml.predictLinearReg	PREDICT_LINEAR_REG
v_ml.kmeansApply	APPLY_KMEANS
v_ml.confusionMatrix	CONFUSION_MATRIX
v_ml.ErrorRate	ERROR_RATE
v_ml.liftTable	LIFT_TABLE
v_ml.ROC	ROC
v_ml.MSE	MSE
v_ml.rSquared	RSQUARED

v_ml Schema

The v_ml schema has been deprecated. All of the Machine Learning functions are now accessible in the public schema.

Models Table

Vertica no longer creates the v.ml.models table by default. This table has been replace by v_catalog.models. For more information about the v_catalog.models table, see MODELS.

Documentation Changes

You can find the SQL functions in the SQL Reference Manual, see Machine Learning Functions.

You can find task-focused documentation in Analyzing Data, see Machine Learning for Predictive Analytics for more information.

Lightweight Copy Table

Vertica now includes the COPY_TABLE function.

This lightweight, in-memory copy increases performance by initially sharing the same storage between two tables. Once copied, the source and copy tables are independent of each other. Users can perform operations on one table without impacting the other. These operations can increase the overall storage required for both tables.

Table Load Options

`CREATE TABLE` and `ALTER TABLE`, and copy operations `COPY`/`COPY FROM VERTICA` now support the following load options:

- `AUTO` (default): Initially loads data into WOS, suitable for smaller bulk loads.
- `DIRECT`: Loads data directly into ROS containers, suitable for large (>100 MB) bulk loads.
- `TRICKLE`: Loads data only into WOS, suitable for frequent incremental loads.

If a table specifies a load option, Vertica uses it for all DML operations unless a DML statement specifies otherwise.

Vertica also supports three load hints: `/*+AUTO*/`, `/*+DIRECT*/`, and `/*+TRICKLE*/`. These hints let you control how individual `INSERT`, `MERGE`, and `UPDATE` operations load table data, overriding the target table's load setting, if any.

For more information on using load options and hints, see Choosing a Load Method in the Administrator's Guide

The system table `CATALOG.TABLES` includes a new column, `storage_mode`, which specifies the table's load option setting.

.

Sampling Data in Vertica

You can now use the TABLESAMPLE clause in your SELECT statements to return a sample of your data. When you use the clause, it returns a sample of the underlying table. The value passed to the function determines the size of the sample.

For more information, see FROM Clause.

STV_Intersect Accepts Two Separate Columns

You can now perform spatial joins without transforming your spatial data to GEOMETRY or GEOGRAPHY data types. STV_Intersect now accepts two columns (x and y) to perform spatial joins, allowing you to do exploratory data analysis more quickly.

For more information, see STV_Intersect Scalar Function and STV_Intersect Transform Function.

Transforming Your Spatial Data Between SRIDs

You can now transform your spatial data between spatial reference identifiers (SRID) WGS84 and Web Mercator using ST_Transform. With ST_Transform, you can convert your spatial data from one SRID to another.

For more information, see ST_Transform.

SDK Updates

This section contains information on updates to the SDK for Vertica Analytics Platform 8.0.

More Details

For more information see Extending Vertica.

Python SDK

Vertica 8.0.0 introduces a new Software Development Kit (SDK) for the Python programming language. This SDK allows you to use Vertica Python APIs to add additional functionality to your Vertica cluster.

For more information, see the following topics:

- Introduction to the Python SDK
- Python Example: currency_convert

SDK Additions to Support Concurrent Load

The API for User-Defined Load (UDL) now supports concurrent load. You can write a `SourceFactory` that creates multi-threaded `UDSource` instances. A new example, FilePortionSource, demonstrates use of this API for both C++ and Java.

For more information, see the following topics in Extending Vertica:

- SourceFactory Class
- Example: Concurrent Load

Apache Kafka Integration

This section contains information on updates to Kafka-integration information for Vertica Analytics Platform8.0.

More Details

For more information see Integrating with Apache Kafka .

Apache Kafka Integration Utility Options

Vertica Analytics Platform has updated and added the following Kafka configuration utility options:

- Shared Utility Options
- Cluster Utility Options
- Scheduler Utility Options
- Load Spec Utility Options
- Microbatch Utility Options
- Source Utility Options
- Target Utility Options
- Launch Utility Options

Apache Kafka System Tables

Vertica Analytics Platform has added the following new and redesigned Kafka integration tables:

- stream_clusters
- stream_events
- stream_load_specs
- stream_microbatch_history

- stream_microbatch_source_map
- stream_microbatches
- stream_scheduler
- stream_scheduler_history
- stream_sources
- stream_targets

Hadoop Integration

This section contains information on updates to Hadoop-integration information for Vertica Analytics Platform 8.0.

More Details

For more information see Integrating with Apache Hadoop.

Hadoop Performance Improvements

When reading data from ORC and Parquet files, you can now use the `hdfs` scheme instead of the `webhdfs` scheme. The `hdfs` scheme is backed by a native C++ library for improved performance. This library supports both Kerberized and non-Kerberized HDFS clusters and also integrates with High Availabiltiy Name Nodes (HA NN).

For information about how to use the new scheme, see Reading Native Hadoop File Formats. Some additional configuration is required to use this scheme; see Configuring hdfs:/// Access.

For information about using this scheme when reading ORC and Parquet files with the HCatalog Connector, see Configuring Vertica for HCatalog.

Vertica on the Cloud

This section contains information on updates for Vertica on cloud platforms.

More Details

For more information, see Using Vertica on the Cloud.

Vertica on Amazon Web Services

The Vertica on Amazon Web Services Guide has moved into the main documentation set. You can find it under the new Using Vertica on the Cloud section.

Vertica on Microsoft Azure

You can create Vertica clusters on the Microsoft Azure cloud platform. You can use the Azure Marketplace to deploy a standard 1-5 node Vertica cluster, or you can manually create a cluster specific to your needs.

For more information, see Using Vertica on Microsoft Azure.

Documentation Updates

This section contains information on updates to the product documentation for Vertica Analytics Platform 8.0.

More Details

For complete product documentation see ® Documentation.

Documentation Changes

The following changes are effective for Vertica 8.0.x.

Document Additions and Revisions

The following additions and substantial revisions have been made to the Vertica product documentation:

- Upgrade documentation is reorganized to more clearly differentiate between upgrade prerequisites and post-upgrade recommendations.
- Documentation on using partition pruning and predicate pushdown to improve query performance for ORC and Parquet files was added in Improving Query Performance.
- Integrations with the MapR Hadoop distribution were clarified in Integrating Vertica with the MapR Distribution of Hadoop.
- The Vertica Place documentation has been moved to Analyzing Data and renamed to Geospatial Analytics.

Removed from Documentation

The following documentation elements were removed from the Vertica 8.0.x product documentation:

- Documentation on the Vertica Connector for Apache Hadoop MapReduce was removed. This connector was removed from the product and is now available on GitHub.

Deprecated and Retired Functionality

This section describes the two phases HPE follows to retire Vertica functionality:

- **Deprecated.** Vertica announces deprecated features and functionality in a major or minor release. Deprecated features remain in the product and are functional. Documentation is included in the published release documentation. Accessing the feature can result in informational messages noting that the feature will be removed in the following major or minor release. Vertica identifies deprecated features in this document.
- **Removed.** HPE removes a feature in the major or minor release immediately following the deprecation announcement. Users can no longer access the functionality. Vertica announces all feature removal in this document. Documentation describing the retired functionality is removed, but remains in previous documentation versions.

Deprecated Functionality in This Release

In version 8.0. the following Vertica functionality was deprecated:

- `DURATION_SEC` column in system table REBALANCE_PROJECTION_STATUS and REBALANCE_TABLE_STATUS was deprecated. This column is now populated with NULL values.
- The configuration parameter `SnapshotRetentionTime` has been deprecated.
- Pre-join projections are deprecated.
- The v_ml schema used by the Advanced Analytics package has been deprecated.
- The v_ml.models table used by the Advanced Analytics package has been deprecated.

See Also

For a description of how Vertica deprecates features and functionality, see Deprecated and Retired Functionality.

Retired Functionality History

The following functionality has been deprecated or removed in the indicated versions:

Functionality	Component	Deprecated Version	Removed Version
Prejoin projections	Server	8.0	
`PROJECTIONS_USED` column in system table `V_MONITOR.QUERY_PROFILES`	Optimizer	7.2.2	7.2.3
Ext3 File System	Documentation	7.2	8.0
Version 6.0 vbr configuration mapping	Server	7.0	7.2
Backup and restore `overwrite` configuration parameter	Server	7.2	7.2
Buddy projections with different sort order	Server	7.2	7.2
verticaConfig vbr configuration option	Server	7.1	7.2
JavaClassPathForUDx configuration parameter	Server	7.1	
ADD_LOCATION()	Server	7.1	
bwlimit	Server	7.1	
Geospatial Package SQL Functions • `BB_WITHIN` • `BEARING` • `CHORD_TO_ARC` • `DWITHIN` • `ECEF_CHORD` • `ECEF_x` • `ECEF_y`	Server	7.1	7.2