Oracle® Cloud

What's New for Oracle Big Data Cloud

E85537-18

June 2021

What’s New for Oracle Big Data Cloud

This document describes what's new for Oracle Big Data Cloud. It's organized by the date a specific feature or capability became available. When new and changed features become available, Oracle Big Data Cloud is upgraded in the data centers where Oracle Cloud services are hosted. You don’t need to request an upgrade to be able to use the new features.

June 2021 (Release 21.2.2)

This release includes security fixes.

August 2020 (Release 20.3.3)

This release includes security fixes.

September 2019 (Release 19.4.1)

Universal Credit accounts no longer use the My Services dashboard to access Oracle Big Data Cloud.

After signing in to Oracle Cloud, you use the Oracle Cloud Infrastructure Console and not the My Services dashboard to access Oracle Big Data Cloud. See Access the Service Console for Big Data Cloud in *Using Oracle Big Data Cloud*.

December 2018 (Release 18.4.4)

Feature	Description
SNAP profile	Oracle Big Data Cloud now includes Sparkline SNAP. SNAP is a terabyte-scale OLAP option on Spark. You can now provision SNAP clusters using the dedicated SNAP cluster profile. To ensure fast performance, SNAP clusters require hardware with local dense I/O storage. SNAP clusters are dedicated to interactive, fast query processing, and connect to BI tools such as Oracle Analytics Cloud and Tableau. SNAP clusters are not meant to handle general purpose Spark processing workloads and do not replace an enterprise data warehouse.
Invalid REST request URLs now return 404	The identity domain ID and the cluster ID specified in REST requests are scanned. Any request with an invalid identity domain ID or cluster ID within the request URL as defined by the API documentation now results in the request being rejected with a 404 error.

November 2018 (Release 18.4.2)

Feature	Description
Apache Spark 1.6 deprecated	Apache Spark 1.6 is deprecated and will be removed in a future release. You should immediately migrate Spark workloads to Spark 2.x.
REST request URLs	The identity domain ID and the cluster ID specified in REST requests are now scanned. Any request with an invalid identity domain ID or cluster ID within the request URL as defined by the API documentation will be honored and a warning will be logged to the log file. There's no 404 error or indication, other than the log file entry, that there's an error. Starting with release 18.4.4, any request with an invalid identity domain ID or cluster ID within the request URL will result in the request being rejected with a 404 error.

April 2018 (Release 18.2.2)

Documentation for Big Data Connectors and Oracle R Advanced Analytics for Hadoop was added. See Connecting to Oracle Database and Working with Oracle R Advanced Analytics for Hadoop (ORAAH) in *Using Oracle Big Data Cloud*.

March 2018 (Release 18.1.6)

Feature	Description
Oracle Cloud Infrastructure compute shapes	The following compute shapes are now available for Oracle Big Data Cloud clusters in Oracle Cloud Infrastructure: • VM.Standard2.2 (2 OCPUs) • VM.Standard2.4 (4 OCPUs) • VM.Standard2.8 (8 OCPUs) • VM.Standard2.16 (16 OCPUs) • VM.Standard2.24 (24 OCPUs)
Browse Oracle Cloud Infrastructure storage and upload files	You can now browse Oracle Cloud Infrastructure storage and upload files using the Big Data Cloud Console.

February 2018 (Release 18.1.4)

The base image was updated with security fixes.

January 2018 (Release 18.1.2)

This release includes hardening, bug fixes, and performance optimizations.

December 2017 (Release 17.4.6)

Feature	Description
Product name change	Oracle Big Data Cloud Service - Compute Edition has been renamed to Oracle Big Data Cloud.
Oracle Cloud Infrastructure deployment	Oracle Big Data Cloud can be deployed on Oracle Cloud Infrastructure and on Oracle Cloud Infrastructure Classic.
Improved provisioning	(Oracle Cloud Infrastructure Classic) The cluster provisioning screen in the console now validates that the correct format is being used for storage URLs. The storage URL must be a full URL. Relative URLs are not accepted.
JDK update	Java was updated to JDK 8u151. For information about the update, see the JDK 8u151 Update Release Notes.
Python 2.7	Python 2.7 is the default runtime for Python Spark.
Big Data Connectors	Oracle Loader for Hadoop and Copy to Hadoop are now preinstalled on all cluster nodes.
Oracle R Advanced Analytics for Hadoop (ORAAH)	ORAAH packages are now preinstalled on all cluster nodes if Spark 1.6 is selected during cluster creation.

October 2017 (Release 17.4.1)

Feature	Description
Oracle Identity Cloud Service integration	In addition to HTTP Basic authentication, Big Data Cloud clusters can now use Oracle Identity Cloud Service (IDCS) for cluster authentication. See Using Identity Cloud Service for Cluster Authentication in *Using Oracle Big Data Cloud*.
Oracle R	Oracle R is now included with all newly provisioned Big Data Cloud clusters. For information about Oracle R, see details about the Oracle R Distribution.
New Zeppelin tutorials	New Zeppelin tutorials are available.

September 2017 (Release 17.3.5)

Feature	Description
Notebook folders	Zeppelin notebooks can now be organized into folders. See Organizing Notes in *Using Oracle Big Data Cloud*.
HDFS file browser improvements	You can now upload and download files through the HDFS browser in the Big Data Cloud Console. See Uploading Files Into HDFS in *Using Oracle Big Data Cloud*.

August 2017 (Release 17.3.3)

Feature	Description
pip	pip is the recommended tool for installing Python packages and is now made available from the command line of each node in the cluster. For information about pip and other tools, see https://packaging.python.org/guides/tool-recommendations/
Automatic notification for object store credentials	If the object store (Cloud Storage) credentials are out of sync a warning message is displayed in the Big Data Cloud Console. If you get this message you'll need to update the password as described in Updating Cloud Storage Credentials in *Using Oracle Big Data Cloud*.
Settings tab displays Thrift URLs	The Settings tab in the Big Data Cloud Console has a new JDBC URLs page that lists the exact URLs to be used for Hive and Spark Thrift. For information about accessing Thrift, see About Accessing Thrift in *Using Oracle Big Data Cloud*.

Feature	Description
Status tab	A new Status tab in the Big Data Cloud Console shows the cluster topology and the current state of each service and component within the cluster. See Viewing Cluster Component Status in *Using Oracle Big Data Cloud*.
BDFS write semantics	The Big Data File System (BDFS) write semantics have been updated to automatically persist data written to BDFS to Oracle Cloud Infrastructure Object Storage Classic.
BDFS memory allocation	The amount of memory allocated has been changed from the previous default of 1 GB per BDFS master and slave to instead be proportional to the shape selected when a cluster is created.
Zeppelin shell interpreter path additions	The Alluxio executable has been added to Zeppelin's shell interpreter path.
Object store browser improvements	You can now upload and delete files from the Cloud Storage browser in the Big Data Cloud Console.

July 2017 (Release 17.3.1)

Feature	Description
Spark 2.1	Spark 2.1 is now supported. You can select Spark 1.6 or Spark 2.1 when you're creating a cluster and that version of Spark is deployed on the cluster.
Zeppelin 0.7	Big Data Cloud now uses Zeppelin 0.7.
Enhanced Cloud Storage browsing	Cloud Storage browsing on the cluster Data Stores page has been improved. There's an improved layout and folder/file browsing structure, plus you can upload files into Cloud Storage (up to 5 GB), see details for a file or directory (including the Swift URL), and refresh the page.

June 2017 (Release 17.2.5)

Feature	Description
MapReduce	The MapReduce feature is now in production and is no longer experimental. See About MapReduce Jobs in *Using Oracle Big Data Cloud*.
Cluster bootstrap script	Advanced users can use the cluster bootstrap script to customize clusters. See Customize Clusters in *Using Oracle Big Data Cloud*.

Feature	Description
Big Data File System	Big Data Cloud includes the Oracle Big Data File System (BDFS), an in-memory file system that accelerates access to data stored in multiple locations and enables Spark jobs to run much faster. See About the Big Data File System (BDFS) in *Using Oracle Big Data Cloud*.

May 2017 (Release 17.2.3)

Feature	Description
Deployment profiles	Specify the type of cluster you want to create based on its intended use. Deployment profiles are predefined sets of services optimized for specific uses. You can choose from the Full profile, which includes all services, or the Basic profile, which includes just some of them. See **Deployment Profile** in Creating a Cluster in *Using Oracle Big Data Cloud*.
Cluster topology	Documentation now includes information about cluster topologies. See About Cluster Topology in *Using Oracle Big Data Cloud*.
Experimental Feature: MapReduce	You can experiment with creating and running MapReduce jobs. The **MapReduce (Experimental)** option is available as a job type when you create a job, but is for experimental use only and is not supported for production use. This option will be fully supported in a future release.

April 2017 (Release 17.2.1)

Feature	Description
High performance storage	Use high performance storage for performance-critical workloads. This option is available when you create a cluster. With this option the storage attached to nodes uses SSDs (solid state drives) instead of HDDs (hard disk drives). See Creating a Cluster in *Using Oracle Big Data Cloud*.
Jobs and notes displayed in list or table view	Use list or table view to see jobs and notes. See Viewing Jobs and Job Details and Viewing and Editing a Note in *Using Oracle Big Data Cloud*.

March 2017 (Release 17.1.5)

Feature	Description
Storage password can be updated from the web-based console	Use the cluster console to update the storage password associated with a cluster when the cluster was created. See Updating Cloud Storage Credentials in *Using Oracle Big Data Cloud*.
Cluster credential store	Create and store credentials in the credential store for a cluster, so they're not passed in clear text in command line parameters or job code. See Using the Cluster Credential Store in *Using Oracle Big Data Cloud*.
Hive interpreter	Use the Hive interpreter for your notebook. For the list of supported interpreters, see Interpreters Available for Big Data Cloud in *Using Oracle Big Data Cloud*.

January 2017 (Release 17.1.3)

Oracle Big Data Cloud was released. The service combines open source technologies such as Apache Spark and Apache Hadoop with unique innovations from Oracle to provide a complete Big Data platform for running and managing Big Data Analytics applications.

See Oracle Big Data Cloud online for documentation, videos, tutorials, and other resources.

Documentation Accessibility

For information about Oracle's commitment to accessibility, visit the Oracle Accessibility Program website at http://www.oracle.com/pls/topic/lookup?ctx=acc&id=docacc.

Access to Oracle Support

Oracle customers that have purchased support have access to electronic support through My Oracle Support. For information, visit http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info or visit http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs if you are hearing impaired.

Oracle® Cloud What's New for Oracle Big Data Cloud,
E85537-18