

ORACLE

# Big Data Service

Bring your present, build your future





Now introducing...

## Oracle Big Data Service

---

Today, an organization's biggest asset is often its data. But organizations still fail to take full advantage of the benefits that data can offer—especially big data, which is often very complex and unwieldy to manage properly. This is why we've created Oracle Big Data Service—Oracle's cloud-native service that's purpose-built to manage big data in the best possible way.

# What is Oracle Big Data Service?

Oracle Big Data Service is a secure place to store and analyze data of different types from any source. It can be used as a data lake or a machine learning platform.

It comes with a fully integrated stack that includes both open source and Oracle value-added tools, and it's designed for enterprises that need flexible deployment options, scalability, and the ability to add tools of their choosing.

## It also provides:

An easy way to expand on premises to Oracle Cloud

Secure, reliable, and elastic Hadoop clusters in minutes

Native integration with Oracle Cloud platform services



# What Oracle Big Data Service helps you accomplish

We drew from Oracle's decades of data management experience, and then we looked to our customers' varied experiences and what they need from the changing world of data. We knew Oracle Big Data Service needed to accomplish four things to be a world-class service.



## Maximize investment

Take your existing applications and deploy to the cloud with minimal changes. Or, use cloud for test, development, and/or disaster recovery.



## Deploy rapidly

Rapidly deliver a secure, highly available platform with all the capabilities required to innovate faster.



## Be a right-size solution

Support varying workload and lifecycle requirements—from agile, purpose-built machine learning clusters to comprehensive data lakes.



## Offer native cloud integration

Offer a service that's readily integrated with a modern cloud platform and facilitates solutions encompassing a wide breadth of services.

**Today, Oracle Big Data Service has all of this and more**



# Oracle + Hadoop = Better together

With Oracle Big Data Service, we've combined the best of what Oracle and Cloudera have to offer.

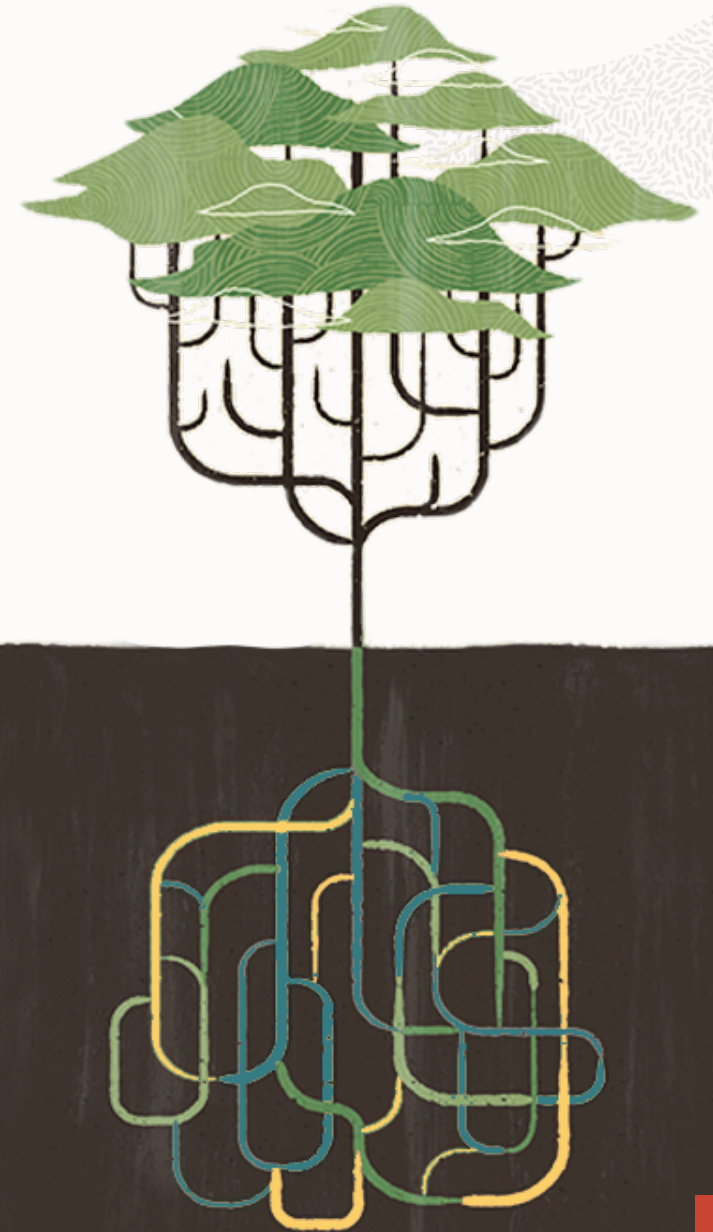
With Cloudera Enterprise Deployment, our service is vertically integrated for Hadoop, Kafka, and Spark with a best-practices, high-availability deployment.

## With Oracle, you also get:

Highly secure, highly available clusters provisioned in minutes

Ability to expand on-premises Hadoop, which enables you to deploy, test, develop, and/or move data lakes to cloud

Flexibility to scale as you wish using high-performance bare metal or cost-effective virtual machine shapes



# The best of Oracle's Data Management Platform

Your organization spends time and effort creating, attaining, and storing data and you want to be able to use it. With Oracle, you can reduce the time, cost, and effort of getting data from wherever it originates to all the places it's needed across the enterprise.

Oracle has spent decades building and expanding its data management platform.

With Oracle's end-to-end data management, you get easy connection to:

- **Oracle Autonomous Database**
- **Oracle Analytics Cloud**
- **Oracle Cloud Infrastructure Streaming**
- **Oracle Cloud Infrastructure Data Catalog**
- **Oracle Cloud Infrastructure Data Science**
- **Oracle Cloud Infrastructure Data Flow**

The list goes on ...

And with unified query with Oracle Cloud SQL, you'll be able to correlate information from a variety of sources using Oracle SQL. In addition, you will gain a host of Oracle analytic and connectivity options, including:

- **Oracle Machine Learning**
- **Oracle Big Data Spatial and Graph**
- **Oracle Big Data Connectors**
- **Oracle Data Integrator Enterprise Edition**

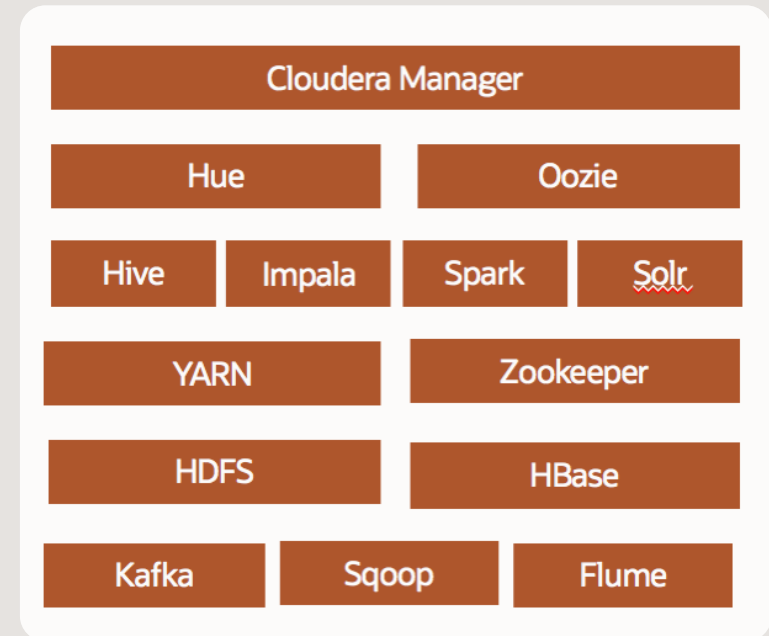
# The best of the Hadoop ecosystem

There's no way around it. Hadoop is complicated. But Cloudera Enterprise Data Hub makes it easier to take advantage of Hadoop. With Oracle Big Data Service and Cloudera, you can:

- Use all of Cloudera Enterprise Data Hub
- Have vertically integrated Spark, Kafka, Impala, Hive, Solr, and more
- Have automatically deployed security and management features

You also can choose your Cloudera version, giving you the ability to:

- **Match your current deployment**  
which is important for test and dev environments
- **Deploy new versions**  
allowing you to take advantage of the distribution's latest features



- Authentication: Kerberos
- Authorization: Sentry
- Auditing: Navigator
- Data at-rest encryption: HDFS Transparent Encryption



# Oracle Big Data Service features

---

We built Oracle Big Data Service to be your go-to big data and data lake solution, one that's specifically designed for a diverse set of big data use cases and workloads. From short-lived clusters used to tackle specific tasks to long-lived clusters that manage large data lakes, Oracle Big Data Service scales to meet an organization's requirements at a low cost and with the highest levels of security.

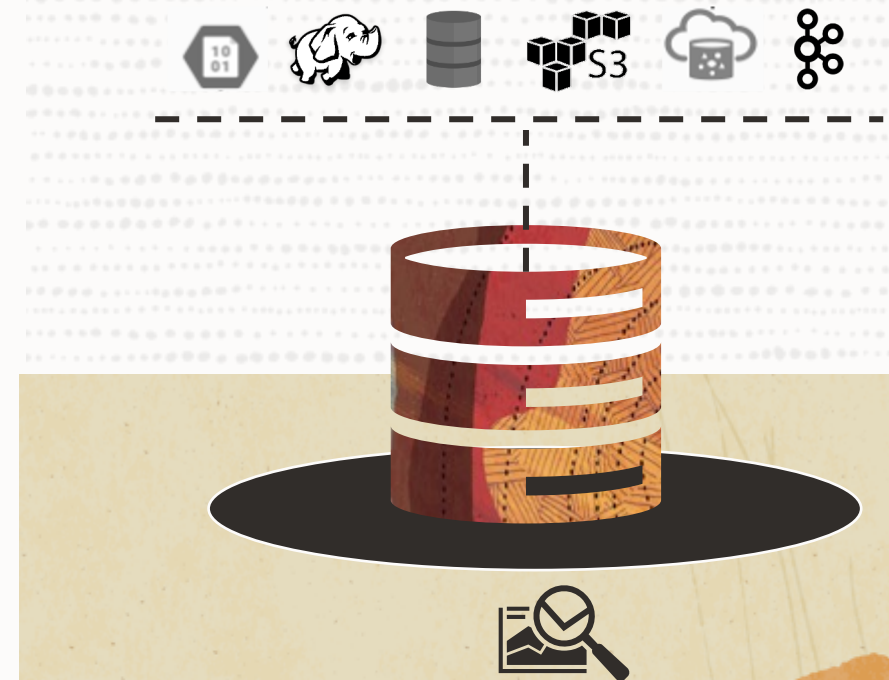
**Let's explore just how  
Oracle Big Data Service  
does this**



# Oracle Big Data Service offers **Oracle Cloud SQL**

Use Oracle SQL to query across big data sources with Oracle Cloud SQL, including the Hadoop Distributed File System (HDFS), Hive, object stores, Kafka, and NoSQL.

You can accomplish all of this with simple administration, because Oracle Cloud SQL uses existing Hive metadata and security, and offers fast, scale-out processing using Oracle Cloud SQL compute.



# Oracle Big Data Service offers **big data analytics**

What use is managing and accessing your data if you can't run analytics to find real results? We offer support in the areas of machine learning, spatial analysis, and graph analysis to help you get the information your organization needs to gain better business results and improved metrics.

**Oracle Big Data Service customers are licensed for these options and can deploy at no extra cost.**

## Machine learning with Oracle

- Scalable, distributed analytics for big data
- Statistics and advanced matrix computation
- Regression models
- Classification models

## Spatial analysis with Oracle

- Proximity and containment analysis
- Location data enrichment
- Vector and raster data preparation

## Graph analysis with Oracle

- Property graphs
- Social media relationships analysis
- Churn analysis
- Cybersecurity fraud detection
- Object recognition

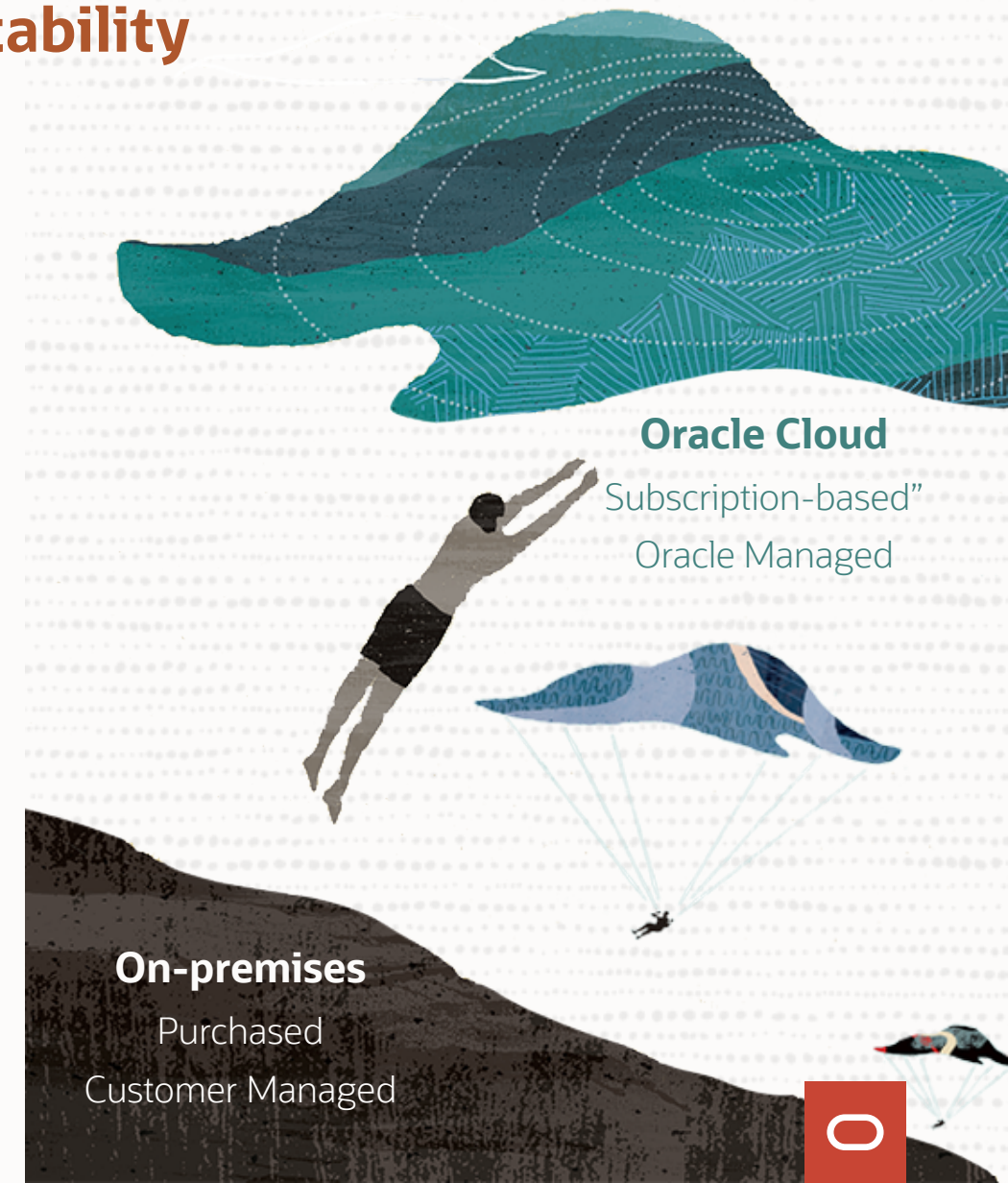
It's also easy to connect to Oracle Cloud services such as Oracle Analytics Cloud, Oracle Cloud Infrastructure Data Science, or Oracle Autonomous Database. Or you can use any Cloudera-certified application for a wide range of analytic tools and applications.



# Oracle Big Data Service offers **workload portability**

Cloud may be the future of enterprise computing, which is why we've built the newest, best cloud infrastructure out there with Oracle Cloud Infrastructure. But it's not everything—at least, not yet. You still need to maintain a mix of public cloud, local cloud, and traditional on-premises computing for the foreseeable future.

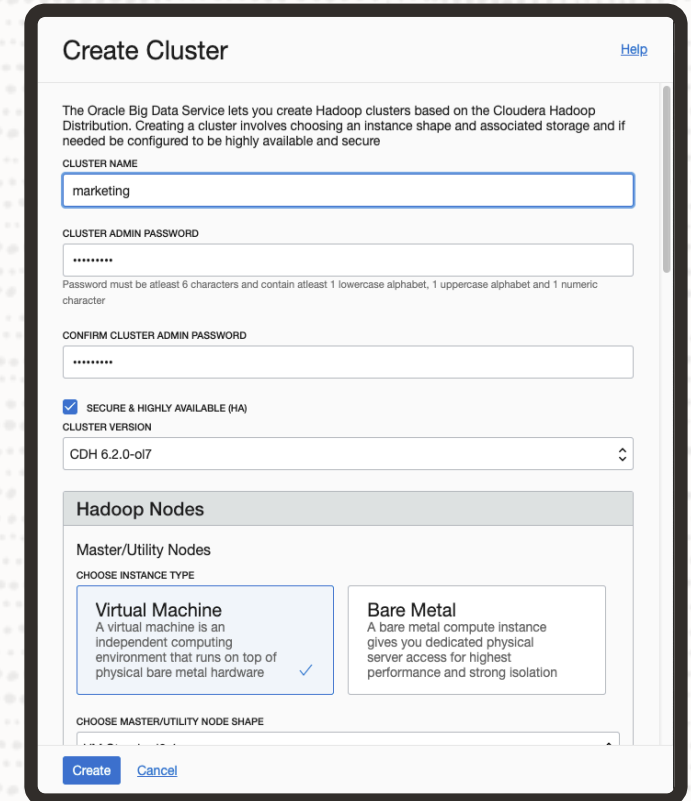
With Oracle Big Data Service, deploy where it makes sense. With Oracle, if you develop something on premises, it's easy to move that to the cloud and vice versa.



# Oracle Big Data Service offers **secure, high-availability clusters**

With Oracle Big Data Service, expect easy deployment when creating your clusters. Specify minimal settings to create the cluster, then use just one click to create a cluster with highly available Hadoop services.

You also get a choice of Cloudera versions, enabling “Cloud Also” deployments to match for on-premises compatibility, or you can choose newer versions to take advantage of latest features.



The screenshot shows the 'Create Cluster' interface in the Oracle Big Data Service console. The interface is titled 'Create Cluster' with a 'Help' link in the top right corner. Below the title, there is a descriptive paragraph: 'The Oracle Big Data Service lets you create Hadoop clusters based on the Cloudera Hadoop Distribution. Creating a cluster involves choosing an instance shape and associated storage and if needed be configured to be highly available and secure'. The form includes several input fields and options:

- CLUSTER NAME:** A text input field containing the value 'marketing'.
- CLUSTER ADMIN PASSWORD:** A password input field with a masked password '\*\*\*\*\*'. Below it, a note states: 'Password must be atleast 6 characters and contain atleast 1 lowercase alphabet, 1 uppercase alphabet and 1 numeric character'.
- CONFIRM CLUSTER ADMIN PASSWORD:** A second password input field with a masked password '\*\*\*\*\*'.
- SECURE & HIGHLY AVAILABLE (HA):** A checkbox that is checked.
- CLUSTER VERSION:** A dropdown menu showing 'CDH 6.2.0-oi7'.
- Hadoop Nodes:** A section with two options:
  - Master/Utility Nodes:** A sub-section with the label 'CHOOSE INSTANCE TYPE'. It contains two cards:
    - Virtual Machine:** A card with the description 'A virtual machine is an independent computing environment that runs on top of physical bare metal hardware'. It has a blue border and a checkmark icon.
    - Bare Metal:** A card with the description 'A bare metal compute instance gives you dedicated physical server access for highest performance and strong isolation'.
  - CHOOSE MASTER/UTILITY NODE SHAPE:** A dropdown menu.

At the bottom of the form, there are two buttons: 'Create' and 'Cancel'.

## Oracle Big Data Service offers **security**

If you're using off-box virtualization, Oracle can't see customer data and customers can't see Oracle management code.

That's a problem. And in most first-generation clouds, the network and tenant environments are coupled, only abstracted by the hypervisor.

Oracle follows a Least Trust Design principle. We don't trust the hardware, the customer (think rogue employees), or the hypervisor. That's why we've separated our network and tenant environments. Isolating that network virtualization helps prevent the spread and lateral movement of attacks.

In addition, with Oracle Big Data Service, all Cloudera security features are enabled with strong authentication, role-based authorization, auditing, and encryption.





# Oracle Big Data Service offers the **compute and storage you want**

Development ➤ Test ➤ Data Science ➤ Data Lakes



## **VM Standard**

1-24 OCPUs  
15-320 GB RAM  
Up to 1PB Block



## **Dense IO**

8-24 OCPUs  
120-320 GB RAM  
6.4-25.6TB NVMe



## **Bare Metal Standard**

52 OCPUs  
762 GB RAM  
Up to 1PB Block



## **Bare Metal HPC Dense IO**

36 OCPUs (3.7GHz)  
384 GB RAM  
6.7 TB NVMe



## **Bare Metal Dense IO**

52 OCPUs  
768 GB RAM  
51.2 TB NVMe

**Performance**

## **Flexibility**

You can use Oracle Big Data Cloud for development, test, data science, or data lakes. In all of these cases, we provide the compute offerings you need for your use case. You can achieve ultimate flexibility by combining bare metal and virtual machines with different storage options, which optimizes your deployment based on your resource needs.

# Oracle Big Data Service offers **superior networking**

With Oracle Big Data Service, you can expect high fidelity, virtual networks, and connectivity.

## Our networking is:

### Customizable

- Fully configurable IP addresses, subnets, routing, and firewalls to support new or existing private networks

### High performance and consistent

- High bandwidth, microsecond latency network
- Private access without traversing internet

### Capable of connecting to corporate networks

- FastConnect—dedicated, private connectivity
- VPN Connect—simple and secure internet connectivity



# Use cases for Oracle Big Data Service

---

Use Oracle Big Data Service for all your big data needs. Here's just some of what you can accomplish.

## **Analytics on Hadoop**

Combine Oracle Analytics Cloud and Oracle Cloud Infrastructure Data Science with Oracle Big Data Service to deliver a powerful, modern analytics platform.

## **End-to-end data management**

Connect Oracle Autonomous Database and Oracle Big Data Service with Oracle Analytics Cloud to create a differentiated data management platform.

## **Hadoop for on-premises and “cloud-also”**

Keep or expand existing on-premises systems with Big Data Service in Oracle Cloud Infrastructure.

## **Flexible machine learning**

Execute machine learning workloads in Spark and Hadoop. Scale for different-size workloads and add your own third-party tools.

## **Oracle SQL on Hadoop**

Use Oracle SQL with familiar tools for data exploration with Query Server and Cloud SQL. Combine data sources for discovery and innovation.







## Oracle Big Data Service for all your big data needs

---

From enabling machine learning to storing and analyzing data, Oracle Big Data Service is a scalable, secure service that meets your requirements at a low cost and the highest levels of security.

It allows you to worry less about managing and storing data. And it empowers you to start analyzing your data in a way that makes the future of your organization more successful than ever before.

# Thank you

---

