

Big-Data-as-a-Service with BlueData and Cisco Unified Computing System



Solution Partner

Run Hadoop and Spark on Docker Containers

SOLUTION HIGHLIGHTS

- ▶ **Platform for Big-Data-as-a-Service (BDaaS):** Offers an on-premises solution with simplified administration, on-demand scaling, and self-service operation for big data infrastructure
- ▶ **Flexibility and choice for big data:** Provides the ability to spin up instant clusters for major Hadoop distributions as well as Spark, Kafka, Cassandra, and other big data frameworks
- ▶ **Business agility and faster time to value:** Enables rapid deployment of business intelligence; extract, transform, and load (ETL); visualization; search; and other big data analytics tools
- ▶ **Containerized solution enabling bare-metal-like performance:** Delivers the elastic scaling, cost-efficiency, and utilization benefits of containers, with system performance comparable to that of bare-metal deployments

Accelerate Time-to-Value for Big Data Analytics

BlueData® provides a software platform that makes multi-tenant deployment of big data infrastructure easier, faster, and more cost effective, eliminating complexity as a barrier to adoption. With BlueData and the Cisco Unified Computing System™ (Cisco UCS®), IT teams can provide their data scientists and analysts with a Big-Data-as-a-Service experience on-premises—enabling more flexibility and higher return on investment for their big data deployments.

The BlueData EPIC™ software platform (see Figure 1) solves the infrastructure challenges that can slow down and stall big data deployments. Using Docker containers, EPIC allows you to create Hadoop and Spark clusters in minutes rather than months, with the analytical tools that your data scientists need. EPIC incorporates innovations for running big data workloads in a containerized environment without performance degradation. EPIC also allows you to reduce data movement and data duplication, using shared infrastructure for your big data deployment.

BlueData EPIC delivers agility, speed, and scalability for big data deployments through:

- **ElasticPlane™** Enables users to deploy self-service clusters on-demand in a secure, multi-tenant environment
- **IOBoost™** Provides application-aware caching for extreme performance and scalability using Docker containers
- **DataTap™** Offers the capability to scale computing and storage resources independently, improving efficiency and resource utilization

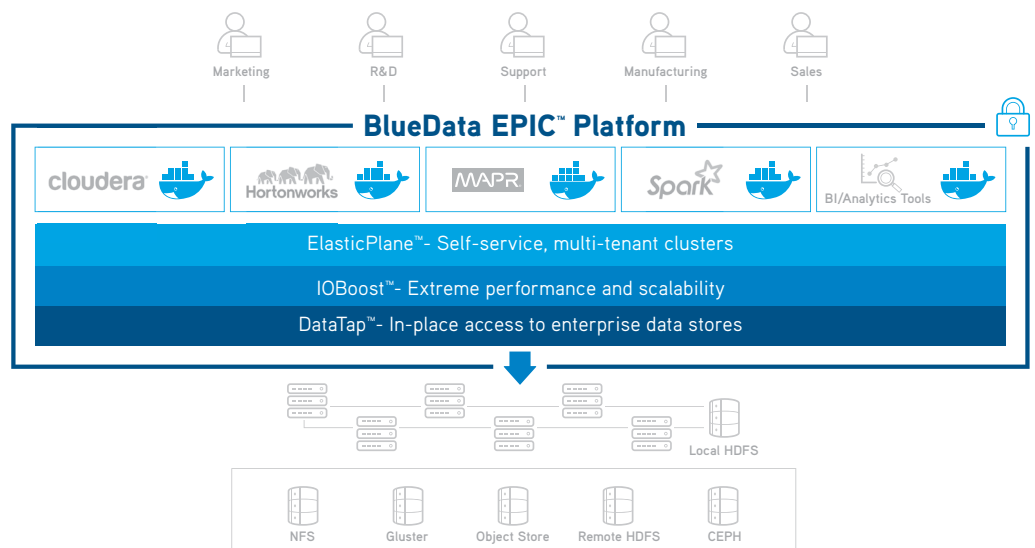


Figure 1: Big Data Infrastructure Made Easy with the BlueData EPIC Platform

BlueData on Cisco UCS for Big Data

Today, most enterprise deployments for Hadoop and big data analytics meet the needs of multiple teams of users (or tenants) with multiple, dedicated bare-metal servers on-premises. This approach results in low hardware utilization, high management overhead, and slow adoption.

Any changes or updates to these traditional big data environments are complex and challenging: whether expanding the existing deployment, adding more tenants, upgrading to the latest Hadoop version, or trying other big data frameworks such as Spark, Kafka, and Cassandra.

With BlueData and Cisco UCS, enterprises can use the power of Docker containers and a common unified infrastructure to deliver exceptional simplicity, speed, efficiency, and performance:

- Quickly implement a container-based solution with big data analytics on Cisco UCS for faster deployments and faster time to business value.
- Create containerized big data clusters in minutes, delivering a Big-Data-as-a-Service solution with a self-service user experience similar to that of a public cloud.
- Provide the flexibility for different tenants to use their big data tools of choice (and quickly try out new tools or upgrade to new versions), with secure multi-tenancy.
- Consolidate big data server, storage, and networking resources in a unified on-premises platform to improve data center utilization and reduce costs.

Together, Cisco and BlueData provide a next-generation multi-tenant infrastructure platform for Big-Data-as-a-Service in an on-premises deployment model. The result is greater agility, increased flexibility, faster deployment, optimized performance, and higher ROI for enterprise big data initiatives.

BlueData works with each of the major Hadoop distributions—including Cloudera, Hortonworks, and MapR—as well as Spark, Kafka, Cassandra, and other big data tools (see Figure 2). Administrators can easily add, change, or upgrade these tools and make them available for self-service cluster creation.

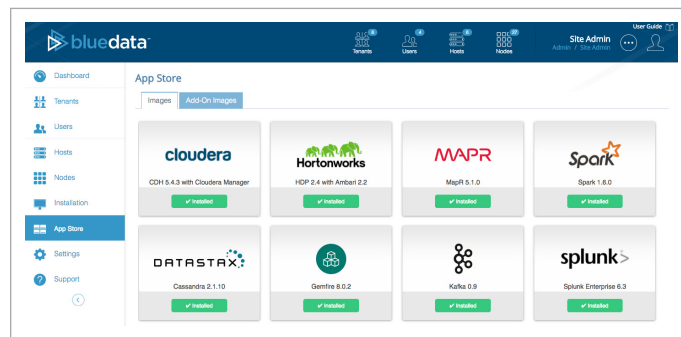


Figure 2: BlueData EPIC Software Platform with Sample Docker Images

With BlueData and Cisco UCS, different teams or tenants can use different big data distributions and applications on shared infrastructure with shared data sets (see Figure 3).

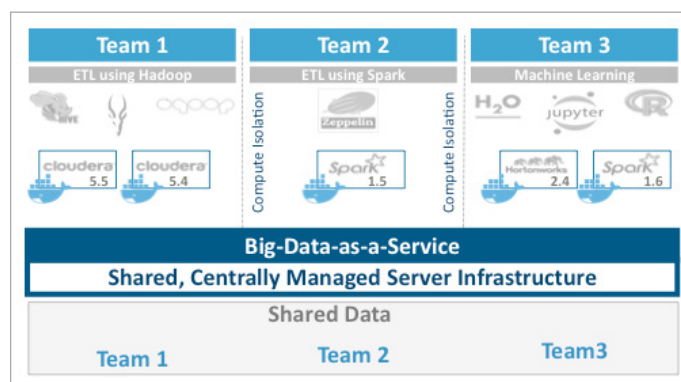


Figure 3: Multi-Tenant Architecture for Big-Data-as-a-Service

BENEFITS

Rapid deployment: BlueData and Cisco together ensure faster deployment with self-service, elasticity, agility, and performance for big data workloads.

Easy, no-touch configuration: BlueData uses service profiles in Cisco UCS Manager to select and configure the hardware transparently.

Simple, point-and-click scalability: BlueData is designed to scale easily on Cisco UCS in just a few mouse clicks, without requiring downtime or complex procedures.

FEATURES

Easy installation: A guided installation process identifies available servers based on criteria for choosing the most appropriate Cisco UCS configuration.

BlueData service profiles for Cisco UCS: The installation associates relevant Cisco UCS servers with the correct service profiles and transparently configures the computing, storage, networking, OS, and software resources.

Web-based GUI to scale out: Expanding your BlueData deployment is as easy as getting a list of IP addresses for the newly available servers in Cisco UCS and adding them to the web-based GUI.

For more information, visit www.bluedata.com

Solution Brief: BlueData on Cisco UCS