



# Accelerate Big Data insights

## HPE Apollo 4000 with Cloudera Enterprise Hadoop

Generate opportunities from rapidly increasing data volumes and varieties with HPE Apollo 4000, density-optimized compute and storage servers built to withstand the rigorous demands of Big Data analytics environments

### Hadoop use cases by industry

- **Finance:** Fraud detection, anti-money laundering, compliance, active archive, market data, and risk management
- **Government:** Fraud detection, anti-money laundering, border control, tax optimization, cyber threat, taxpayer 360, and risk or intelligence management
- **Telecom:** Customer 360, broadcast monitoring, media streaming, EDW modernization, data monetization, churn prediction or prevention, and advertising optimization
- **Manufacturing:** Supply chain optimization, defect tracking, radio-frequency identification (RFID) correlation, warranty management, predictive maintenance, IoT-enabled data “smart services,” diagnostics, and connected car, plane, train, or equipment
- **Energy:** Weather forecasting, natural resource exploration, predictive maintenance, IoT-enabled data analytics, supply chain optimization, and staffing or inventory
- **Healthcare:** Drug development, scientific research, evidence-based medicine, healthcare outcomes analysis, clinical trials, patient care 360, and diagnostics

### The Big Data challenge

As companies strive to identify and realize the value in Big Data, many now seek more agile and capable analytic systems. Some of the business drivers are to improve customer retention, increasing operational efficiencies, influence product development and quality, and to gain a competitive advantage.

Information-driven organizations are required to leverage an unprecedented volume and variety of data. These organizations must also cost-effectively store and deliver insights from data in a number of ways, to a wide range of audiences and stakeholders—each with their own diverse requirements. Apart from the volume and variety of data, data security is a major challenge many organizations are facing.

These organizations need data platforms to augment existing data warehouses, and scale out to handle exploding data volumes from both existing and new data sources, while providing the flexibility to archive historical data cost effectively. Apache Hadoop is a software framework that provides measurable savings and value for storage and data processing at large scale, and has been adopted by many enterprises as a solution to meet those needs.

Hadoop began its journey as a highly scalable data platform for reporting and analytics on the Web, and the platform is now increasingly put to use in digital organizations. The journey is now taking Hadoop into a wider range of industries, use cases, and types of organization. This

time around, Hadoop must meet more stringent requirements and capabilities that IT departments demand of their platforms for enterprise adoption. As Hadoop broadens across the enterprise, its ownership is shifting from departments and application teams to central IT on a shared enterprise platform.

### Get big value from Big Data

Leveraging the Apache Foundation’s Hadoop open-source Big Data platform, Cloudera and HPE Apollo give you fast and reliable analysis of both structured and unstructured data at enterprise scale. An open source, Linux®-based platform for data storage and processing, Cloudera Enterprise Hadoop is massively scalable, highly fault tolerant, and truly distributed. Hewlett Packard Enterprise and Cloudera helps you make better business decisions by laying the foundation for the analysis of growing structured and unstructured data and accelerating time to value.

### HPE and Cloudera are your partners on the path to Big Data

Cloudera Enterprise and our density optimized Apollo servers help you become more information-driven by leveraging the best of the open source community with the ultra-dense and ultra-scalable enterprise servers you need in order suppress your competition, know your customers, and manage your growing data set. Designed

## Solution brief

specifically for mission-critical environments, Cloudera Enterprise includes Cloudera Data Hub (CDH), the world's most popular open source Hadoop platform that has been proven and tested on the HPE Apollo 4200 and HPE Apollo 4530, as well as advanced system management and data management tools, plus dedicated support and community advocacy from our world-class team of Hadoop developers and experts.

Deploying Cloudera Enterprise Hadoop on optimized HPE infrastructure gives you the compute power and storage required to respond, with flexibility, to the demands of growing Big Data.

### Rethink Big Data with HPE and Cloudera

With the HPE Apollo 4000 family of density optimized servers, Hewlett Packard Enterprise is helping you reconsider how to leverage and manage Big Data more effectively than before. This innovative family of high-density servers is purpose-built for Big Data with a unique focus on both density and scalability of compute and storage that Hadoop requires.

#### HPE Apollo 4200 Density Optimized Server The enterprise bridge to Big Data solutions

The HPE Apollo 4200 Gen9 Server offers revolutionary storage density in a 2U form factor, designed for traditional enterprise data centers. It provides a low initial investment, and allows you to scale in increments of one 2U server at a time, up to 224 TB per server with 28 large form factors (LFFs) 8 TB hard disk drives (HDDs).

#### HPE Apollo 4530 Density Optimized Server Right-sized for Hadoop analytics

The HPE Apollo 4530 Server is built for the wide variety of Big Data analytics workloads based on parallel Hadoop-based data mining.

This 4U, three-node chassis boasts of a one-to-one core-to-spindle ratio. It provides the performance and storage density that enables you to develop a 360-degree view of customers to improve marketing cost effectiveness, boost online sales, and enhance customer retention and satisfaction.

Each of the three servers in the chassis has up to 120 TB capacity—providing building blocks for efficient implementations at scale with up to 30 nodes and 3.6 PB capacity in a 42U rack.

#### HPE Big Data Reference Architecture

Most Hadoop users start with the traditional, symmetric Hadoop architecture. However, as their needs grow, they often end up with more, sometimes many, separate Hadoop clusters, with substantial amount of the same data, all representing a sizable investment.

Implementing the HPE Big Data Reference Architecture (BDRA) allows you to migrate to this new architecture, reducing your overall cost and allowing for future growth, while protecting your current investment. BDRA enables you to separate compute and storage and scale either one based on your organizational needs.

#### HPE ProLiant DL380 Gen9 Server platform

Best-selling 2U servers create the ideal platform for smaller Hadoop deployments or for customers that have standardized on HPE ProLiant DL servers.

### Solution benefits

#### Time to insight

The speed of the HPE platform combined with the power of Cloudera Enterprise Hadoop enables organizations to accelerate collection and real-time analysis of Big Data from multiple sources and in multiple formats. Drive faster time to insight and increase value with low latency and streaming analytics using new technologies like Vertica, Spark, and Impala.

#### Scale efficiently and cost effectively

Hewlett Packard Enterprise and Cloudera solutions are architected to enable you to achieve hyperscale performance at petabyte scale. Density-optimized servers have

evolved to deliver operating expenditure (OPEX) savings (power, cooling and space) while boosting system performance.

#### Ease of management

HPE Insight Cluster Management Utility (CMU) is a product that helps you rapidly provision, manage, and monitor scale-out clusters. This solution scales along with your Hadoop clusters.

By combining the power of CMU, Cloudera Manager, and Cloudera Director, quickly provision, deploy, and extend a Hadoop cluster into hybrid cloud.

#### Build a scalable data lake

Effectively manage your data silos by consolidating data on HPE Apollo Density Optimized Servers. Deploy workload optimized HPE BDRA to drive diverse analytic use cases. Make better use of available capacity by scheduling cluster resources where needed using Hewlett Packard Enterprise's contribution to YARN labels.

### The support you need to succeed

Implementing Hadoop requires a skill set that encompasses not only knowledge of the Hadoop environment itself but also the operating system, network, automation, and operations of the Hadoop cluster. The HPE Reference Architecture Implementation Service for Hadoop offers deployment services where HPE Big Data consultants install, configure, deploy, and test your specific Hadoop environment.

This is just one of many offerings from HPE Consulting Services developed to help you define, architect, build, and safeguard an integrated IT strategy for Big Data. Our experienced team of Big Data technology consultants can devise the most efficient and effective means to capture, consolidate, manage, and protect business-aligned information.

Learn more at  
[hpe.com/info/hadoop](http://hpe.com/info/hadoop)



Sign up for updates

  
**Hewlett Packard  
Enterprise**

© Copyright 2016 Hewlett Packard Enterprise Development LP. The information contained herein is subject to change without notice. The only warranties for Hewlett Packard Enterprise products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Hewlett Packard Enterprise shall not be liable for technical or editorial errors or omissions contained herein.

Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.

4AA6-6030ENN, June 2016