

Brochure

Handle big data. Today.

HP AppSystem for Apache™ Hadoop™



Delivering on the promise of big data

Did you know?

Unique design

The Hadoop Distributed File System is uniquely designed to reliably store and manage large volumes of data.

Very high capacity

Processing capacity of more than 20 petabytes of arbitrarily structured data per cluster is often achieved.

Greater support

HP already supports numerous customers running thousands of nodes in their Hadoop clusters.

Discover what's possible with an enterprise-ready, turn-key Hadoop platform that's pre-tested, factory integrated, and backed by 24x7 support.

Capitalizing on the next wave of IT investment

Today's enterprises are actively investigating innovative ways to exploit the explosion of big data—structured, unstructured, semi-structured, machine-generated, transactional, and user-generated—from within and outside their organizations. The goal is to use the information for better business insights, and to respond to and pursue customer and market opportunities more quickly.

Many organizations find the best way to handle big data is to deploy a scale-out computing model. And today, HP is the leader in scale-out computing and platforms for data management—including Apache™ Hadoop™.

Hadoop deployment made faster, simpler, less risky

Resolving your deployment challenges

An open source Linux-based platform for data storage and processing, Apache Hadoop is massively scalable, highly fault tolerant, and truly distributed. Hadoop enables enterprises to make better business decisions by processing vast amounts of semi-structured and unstructured data, and running multiple analyses on Hadoop-based clusters. Built on a proven scale-out architecture that divides workloads across multiple nodes, Hadoop has what it takes to process very complex data.

In many cases, however, enterprises lack the expertise required to deploy and configure Hadoop clusters for optimal performance and resource utilization. Adding to the challenges are the risks involved when moving from pilot Hadoop deployments to live production, as well as complex, non-integrated management and monitoring capabilities.

To address these challenges, HP designed the HP AppSystem for Apache Hadoop—an enterprise-ready, factory-assembled, rack-based cluster solution that arrives pre-configured and ready to run on day one. This pre-integrated, fully optimized solution reduces the complexity of deploying and configuring Hadoop clusters. And that means faster time-to-value for you.

State-of-the art components

HP AppSystem for Apache Hadoop includes the following components:

- **HP ProLiant DL380e Gen8 Servers** – Powered by the HP Proactive Insight architecture, HP ProLiant DL380e Gen8 Servers provide maximum data density and the optimal disk-to-core ratio. By continuously analyzing thousands of system parameters, ProLiant DL380e Gen8 Servers optimize application performance, proactively improve uptime, and provide insight into every aspect of your IT infrastructure. Built with over 150 design innovations, HP ProLiant DL380p Gen8 Servers help you eliminate common problems that can cause downtime and data loss.
- **HP Networking Switches** – Providing the deep packet buffers (1 GB) required for good networking performance. The switches are configured in an HP Intelligent Resilient Framework (IRF) configuration to accommodate network interface card (NIC) bonding for both increased bandwidth and resiliency. By changing the rules of networking through the power of convergence, HP Networking Switches offers greater security, deep packet buffers, redundancy, throughput, and speed for your Hadoop deployment.
- **HP Insight Cluster Management Utility (Insight CMU)** – An efficient, robust tool for deploying and managing Linux-based nodes in large clusters, HP Insight CMU enables push-button scale-out deployment, remote cluster management, and real-time and time-series monitoring. Insight CMU is proven cluster management software that has been in deployment in large high performance computing (HPC) clusters for more than a decade.
- **Cloudera Enterprise** – Enabling data-driven enterprises to cost-effectively run Apache Hadoop in production environments with repeatable success. As the leading distribution for Hadoop in the enterprise, Cloudera Enterprise includes a 100% open source, enterprise-ready distribution of Hadoop and related projects, a subscription to Cloudera Manager, and 24x7 production support from Cloudera.
- **Red Hat Enterprise Linux** – As the world's leading Linux distribution,¹ Red Hat Enterprise Linux powers thousands of commercial and custom applications. Built for the modern data center environment, Red Hat Enterprise Linux offers pervasive networking—a comprehensive kernel-based security system that defines and enforces the access rights of every user, application, process, and file within a Red Hat system. HP support for Linux is included in the Red Hat Enterprise Linux subscription.
- **HP Vertica Community Edition** – Broadening your solution breadth into real-time analysis of structured data, Vertica Community Edition² can load data directly from the Hadoop Data File System (HDFS) using built-in WebHDFS connectors. You get the best of both worlds with the big data file storage and processing capabilities of Apache Hadoop with targeted SQL analytics capabilities of this unique HP intellectual property for structured data.

Did you know?

HP Converged Infrastructure powers:

- Three of the most popular social media properties in the U.S.
- Eight of the top 10 websites
- Four of five of the world's top search engines³

Best-in-class partner for Hadoop management

To deliver a solution that provides the greatest value to our customers, HP partners with Cloudera—the leader in Apache Hadoop-based software and services. With Cloudera Enterprise running within your HP AppSystem, you get the leading solution for Apache Hadoop in the enterprise. It includes not only CDH (Cloudera Distribution for Hadoop)—the 100% open source, enterprise-ready distribution of Hadoop and related projects from Cloudera—but also a subscription to Cloudera Manager and technical support for the core components of CDH.

Hadoop Management

As the industry's first end-to-end management application for Apache Hadoop, Cloudera Manager sets the standard for enterprise deployment by delivering granular visibility into and control over every part of CDH—empowering operators to improve cluster performance, enhance quality of service, increase compliance, and reduce administrative cost.

Hadoop Support

Cloudera offers a dedicated team of 24x7 support engineers comprised of contributors and committers for every component of CDH. With Cloudera Support, you'll experience more uptime, faster issue resolution, and better performance to support your mission-critical applications.

¹ IDC Worldwide Linux Operating Environment Forecast 2011–2015

² Vertica is a SQL database from HP that provides real-time structured data analytics. Scaling to petabytes of data, Vertica extends your analytics capabilities through a wide ecosystem of analytic tools

³ From alexa.com/topsites

HP solution advantage

Rather than offer you an assortment of piece-parts, HP AppSystem for Apache Hadoop arrives as a factory-integrated, pre-configured, pre-tested system that's ready to deploy on day one. Pre-loaded with Red Hat Enterprise Linux and Cloudera Enterprise, and designed according to HP best practices, HP AppSystem for Hadoop translates into quick, production-quality roll-outs at low risk.

We deliver a total solution through HP Technology Consulting and Support Services and our partners, including Cloudera Enterprise production support offered directly by Cloudera. And HP Technology Support Services delivers 24x7 global infrastructure support tailored to your specific requirements. This HP solution advantage helps you:

- **Get ready for big data** – Leveraging the advanced capabilities of the latest HP ProLiant Gen8 servers, market-leading Hadoop software from Cloudera, proven HP Insight CMU cluster management, and 11-plus years of hyperscale customer experience.
- **Overcome the complexity of Hadoop cluster operations and scarcity of resources** – By assembling and configuring HP AppSystem for Hadoop in the HP factory, you can focus on generating greater business value rather than tracking down able bodies to sift through complex documentation, configure and test all the components, and optimize them for the right configuration.
- **Update software and maintain your Hadoop cluster** – HP Insight CMU helps you provision, control, and monitor your Hadoop cluster. In fact, you can deploy more than 800 nodes in 30 minutes. Plus, HP Insight CMU is integrated with Cloudera Manager and Ganglia—the scalable, distributed, open source monitoring system for high-performance computing systems. With HP Insight CMU, you can optimize and monitor your end-to-end Hadoop hardware and software platform.

In addition, HP offers a 1–3 day big data workshop for needs definition and use case analysis to help you evaluate options, as well as our “Roadmap Service for Hadoop” to help you effectively plan for a successful implementation of Hadoop in your organization.

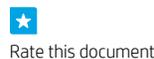
HP—bringing it all together

For HP AppSystem for Apache Hadoop, one hardware platform offers all the flexibility, scalability, and resiliency you need—HP Converged Infrastructure—where all IT resources are pooled and shared enterprise-wide. By overcoming the inflexibility and high costs created by IT sprawl, HP Converged Infrastructure solutions enable you to shift more resources to innovation and strategic initiatives. The HP approach and solutions have been designed to help make your data center simpler, more efficient, and less expensive to operate. This way, your IT infrastructure is positioned to deliver any workload, anywhere, any time, to accelerate the time to application and business value.

For more information visit:

hp.com/go/hadoop
vertica.com
cloudera.com

Sign up for updates
hp.com/go/getupdated



© Copyright 2012 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP 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. HP shall not be liable for technical or editorial errors or omissions contained herein.

Apache and Hadoop are either registered trademarks or trademarks of the Apache Software Foundation in the United States and/or other countries.

4AA4-4425ENW, December 2012

