Empower the data-driven organization with HPE and Hortonworks
Transform to a hybrid infrastructure

Enable workplace productivity

Protect your digital enterprise

Empower the data-driven organization

Harness 100% of your relevant data to empower people with actionable insights that drive superior business outcomes.
Yesterday’s Data-Driven

Data
- Limited business data
- Silo-ed apps

People
- Not understanding the value
- Silo-ed departments

Insights
- Hindsight analytics…reactive
- Monolithic reports, not real-time

Marginal impact and business benefit
Data-Driven in the Idea Economy

Data
- Leverages all relevant data
- Integrated across apps

People
- Executives to line employees
- Informed by all relevant data

Insights
- Predictive analytics…proactive
- Analytics-apps on every platform

Achieves superior business outcomes
The data landscape is radically changing
More connected people, apps and things generating more data in many forms

- Human data
- Machine data
- Business data

10x faster growth than traditional business data
Enterprises realize only 10-15% of the expected value on their big data investments.

Barriers:

- Silos and Lack of Alignment
- Technology Gap
- Translating Data to Value
Accelerate your path to becoming a Data-Driven Organization

Build a Data-Centric Foundation

Discover the Value of Your Data

Achieve Superior Business Outcomes with Big Data

100% of your relevant data

Achieve breakout growth

Optimize operations

Reduce risks

Human Data

Machine Data

Business Data
How to discover the value of your data

**Align** business goals and challenges with the relevant data

**Evaluate** your data and quickly test, learn, and iterate ideas to discover value

**Create** a strategic roadmap based on learnings

---

**Key HPE solutions**
- Data Discovery
- Data Driven Transformation Planning

**Business benefits**
- Agile execution to impactful projects
- Maximize alignment to value
How to build a data-centric foundation

Maximize your existing investments
Build a data-centric, flexible architecture
Choose the right platforms to power your analytics-apps
Analyze 100% of the relevant data at the speed of business
Govern your data for compliance and risk mitigation

Key HPE solutions
Enterprise-grade Hadoop
Real Time Analytics with SAP HANA

Business benefit
Faster answers for 100% of your relevant data

HPE Vertica, HPE IDOL, HPE Haven OnDemand
Workload Optimized Infrastructure
Information Governance
How to achieve superior business outcomes with big data

Uncover meaningful patterns in data applying data science

Integrate these insights and algorithms into production environments

Deliver insights across your organization through analytics-apps

Key HPE solutions

Application Solution Frameworks
- Voice of the Customer
- Warranty Analytics

Operationalized Analytics

Business benefits

Accelerate time from analytic discovery to business impact

All the relevant stakeholders empowered with insights

Insights are available at the point of action
HPE Hortonworks Joint Commitment

- Alliance partner for 2+ years
- HPE invested $50M in Hortonworks
- HPE is on the Board of Hortonworks
- Joint engineering collaboration with Hewlett Packard Labs
  - New collaboration to enhance Apache Spark
  - New class of analytic workloads that benefit from large pools of shared memory

“Hewlett Packard Enterprise is one of our top strategic partners, working closely with our engineering organization to deliver proven customer solutions for building a modern data architecture”

Rob Bearden, CEO Hortonworks
HPE and Hortonworks enable Enterprises to become data driven

<table>
<thead>
<tr>
<th>Actionable Intelligence</th>
<th>Enterprise Grade</th>
<th>Open Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hortonworks Data Platform</strong></td>
<td>Accumulate, analyze and act on all data sources</td>
<td><strong>HDP for the Enterprise</strong></td>
</tr>
<tr>
<td></td>
<td>Centralized Architecture for multi-tenancy</td>
<td>- Centralized management and monitoring</td>
</tr>
<tr>
<td></td>
<td>Enterprise Operations, Governance and Security</td>
<td>- Automated provisioning</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>HPE Servers</strong></th>
<th><strong>HPE Optimized Servers</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Balanced compute and storage for Hadoop</td>
<td>- Flexible performance</td>
</tr>
<tr>
<td>HPE DL380</td>
<td>- Density at massive scale</td>
</tr>
<tr>
<td>HPE Apollo 2000</td>
<td>- High efficiency</td>
</tr>
<tr>
<td>HPE Apollo 4200</td>
<td>- Ease of deployment</td>
</tr>
<tr>
<td>HPE Apollo 4530</td>
<td></td>
</tr>
</tbody>
</table>

Tested and proven on HPE Platforms
Enterprise-Grade Hadoop

Optimize the Hadoop Data Lake for More Business Value

- Unlock the most value and performance from Hadoop
- Scale without compromising data security, reliability, and ROI
- Enterprise-Grade, Trusted, and Proven HPE solution

Flexible, Optimized Infrastructure for Hadoop

High-Performing Analytics Engines for Hadoop

Data Security for Hadoop

Consulting & Implementation Services for Hadoop
HPE Optimized Compute Portfolio for the Data centric Foundation
Innovative, workload optimized platforms for all data intensive applications and use cases

Big Data Analytics  |  Object Storage  |  High Performance Compute  |  Real-time Analytics  |  Traditional Data Warehousing

- HPE Apollo + Moonshot
  Scale, Flexibility, Efficiency, Density, Performance

- HPE Integrity SuperdomeX
  In-memory Performance

- HPE ProLiant + BladeSystem
  Database Processing

Scale-Out Compute  |  Scale-Up Compute
# HPE platforms for Big Data

Helping customers get the best economics for their Big Data

Tested & Proven with Hortonworks HDP

<table>
<thead>
<tr>
<th>Traditional</th>
<th>Optimized</th>
<th>Converged</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Most common deployment</td>
<td>- Optimized for Big Data workloads and storage</td>
<td>- MPP DBMS approach + open source</td>
</tr>
<tr>
<td>- Small to large deployments (very often ~20 nodes)</td>
<td>- Mid-size to large deployments</td>
<td>- Mid-size to large deployments</td>
</tr>
<tr>
<td>- Linear growth of balanced workloads</td>
<td>- Single, resource-intensive workload</td>
<td>- Multiple workloads, latency deployments</td>
</tr>
<tr>
<td>- Smaller fault zones</td>
<td>- Non-linear storage growth</td>
<td>- Isolate workload hot spots</td>
</tr>
<tr>
<td>- Risk averse, majority adopter customers</td>
<td>- Multi-temperate storage</td>
<td>- Compute and storage scale independently</td>
</tr>
<tr>
<td>- Mainstream platform benefits</td>
<td>- “Optimized traditional”</td>
<td>- TCO-driven</td>
</tr>
<tr>
<td></td>
<td>- Higher density, lower TCO</td>
<td>- A Reference Architecture using HPE Server options – ProLiant, Apollo, and Moonshot</td>
</tr>
<tr>
<td></td>
<td>- Very high density</td>
<td></td>
</tr>
</tbody>
</table>
Building a Data-Centric Foundation
Infrastructure Innovation purpose built for Big Data demands

**Traditional Big Data Approach**
- Processing / storage always collocated
- All identical servers
- Data partitioned across servers on direct-attached storage (DAS)

**HPE Big Data Reference Architecture**
- Separate processing / storage tiers connected by Ethernet networking
- Standard Hadoop installed asymmetrically
  - storage components on storage servers
  - yarn apps on processing servers

### Flexibility to scale
Scale compute and storage independently

### Cluster consolidation
Multiple big data environments can directly access a shared pool of data

### Maximum elasticity
Rapidly provision compute without affecting storage

### Breakthrough economics
Significantly better density, cost and power through workload optimized components
Apollo 4200 – Bringing Big Data Storage Density to Enterprise
The Enterprise Bridge to Big Data

- **Storage Density**: Leadership storage density - 224 TB in a 2U server
- **Plug & Play**: Enterprise bridge - Fits traditional enterprise/SME rack server data centers – deploy today, no cost of change
- **Performance Efficiency**: Configuration flexibility - Balanced capacity, performance & throughput with flexible options - Disks, CPUs, I/O and interconnects

Highest storage density in a traditional 2U rack server – 224 TB
Apollo 4530 – Optimized Scale-out Hadoop Server
Balanced compute and storage, triple-node density

Hadoop Analytics Server

![Image of Apollo 4530 server]

- **Workload Optimized**
  - Scale-out server density
  - Three independent, high-storage-capacity nodes in a space saving 4U form factor

- **Lower TCO**
  - Cost effective
  - Balanced processing and storage in compact 4U’s of rack space for low-cost, power & space efficient solutions

- **Flexibility**
  - Easy deployment
  - Great fit for three-copy Hadoop designs. Broad set of choices for 2P processor, memory, I/O, and storage.

Rack-density for Hadoop – 3.6 PB per Rack!
ProLiant DL380 Gen9 – General Purpose multi use server
Versatile, industry-leading platform

Latest Processor and NVMe Options
Latest Intel Xeon E5-2600 processors and HPE Persistent Memory, offering unprecedented levels of analytic performance

Lower TCO
Easy to repurpose for use with different workloads

Flexible Design
Choice of SFF, LFF and NVMe drive options. Broad choice of processor, memory, I/O, and storage options. Universal Media Bay and Graphic card options for deep learning

Industry-leading server with flexible choices for Big Data workloads
Big Data Reference Architecture
Scale compute and storage independently on workload-optimized platforms

- **Purpose-Built Clusters**
  Build a Big Data cluster with storage and compute optimized nodes to flexibly run multiple workloads at the same time

- **Workload Optimized**
  Elastically scale resources to meet the varying demands of Big Data / Big Analytics workloads

- **Consolidation**
  Create a multi-temperate Data Lake and consolidate multiple Big Data / Big Analytics workloads

- **Flexibility**
  Independently Scale Compute and Storage
  Elastically scale resources to meet the varying demands of Big Data / Big Analytics workloads

**Apollo 4200 Storage Server**

**Apollo 2000 Compute Nodes**

**2x the density, half the wattage of a traditional Hadoop cluster**
The HPE Vertica portfolio
All built on the same trusted and proven HPE Vertica Core SQL Engine

Core HPE Vertica SQL Engine
• Advanced Analytics
• Open ANSI SQL Standards ++
• R, Python, Java, Spark, Scala

HPE Vertica AMI
• Get up and running quickly in the cloud
• Virtualization
• Flexible, enterprise-class cloud deployment options

HPE Vertica Enterprise
• Columnar storage and advanced compression
• Maximum performance and scalability
• Flex Tables for schema-on-read

HPE Vertica for SQL on Hadoop
• Native support for ORC, more
• Support for industry-leading distributions
• No helper node or single point of failure
HPE Vertica for SQL on Hadoop features and benefits

Query data, no matter where it is located

- Install HP Vertica
  - directly on your Hadoop infrastructure, supporting YARN
  - beside your Hadoop infrastructure, accessing data in the Hadoop cluster

- ORC, Parquet, Avro, Vertica ROS and JSON supported

- Full-functionality ANSI SQL

- 100% of TPC-DS queries
  - No helper node or single point of failure

- Competitive price point
**HP Vertica for SQL on Hadoop**

- Same Vertica MPP Columnar Architecture
- Base ANSI SQL
- Co-Located with Hadoop
- Data Query Across parquet, ORC, JSON, and many other format
- Hadoop Agnostic

```plaintext
<table>
<thead>
<tr>
<th>ANSI SQL ENGINE</th>
<th>Open Formats + ROS/Flex</th>
<th>HDFS</th>
<th>HPE Big Data Reference Architectures</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANSI SQL ENGINE</td>
<td>Open Formats + ROS/Flex</td>
<td>HDFS</td>
<td>HPE Big Data Reference Architectures</td>
</tr>
<tr>
<td>ANSI SQL ENGINE</td>
<td>Open Formats + ROS/Flex</td>
<td>HDFS</td>
<td>HPE Big Data Reference Architectures</td>
</tr>
<tr>
<td>ANSI SQL ENGINE</td>
<td>Open Formats + ROS/Flex</td>
<td>HDFS</td>
<td>HPE Big Data Reference Architectures</td>
</tr>
</tbody>
</table>
```

[Logo of Hortonworks, MapR, and Cloudera]
Data storage options and performance

Query Engine
- Vertica
  - ANSI SQL-99

Format
- Vertica ROS

File System
- EXT4

HPE Vertica SQL on Hadoop
- Vertica ROS
  - HDFS
- Hadoop Format
  - HDFS
- Flex Tables
  - HDFS
- Flat Files
  - HDFS

Performance
- Fastest
  - Analytics
  - Structured
- Slowest
  - Discovery
  - Semi-Structured
HPE Security - Data Security
We protect the world’s most sensitive data

– Protect the world’s largest brands & neutralize breach impact by securing sensitive data-at-rest, in-use and in-motion.
– Over 80 patents & 51 years of expertise

Our Solutions:
provide advanced encryption, tokenization & key management

Market leadership:
– Data-centric security solutions used by six of the eight top U.S. payment processors & seven of the 10 top U.S. banks.
– Thousands of enterprise customers across all industries including transportation, retail, financial services, payment processing, banking, insurance, high tech, healthcare, telecom & public sector.
– Email solution used by millions of users and thousands of enterprise & mid-sized businesses including healthcare organizations, regional banks & insurance providers.
– Contribute technology to multiple standards organizations.
Why is securing Hadoop difficult?

Rapid innovation in a well-funded open source community

Multiple feeds of data in real time from different sources with different protection needs

Multiple types of data combined in a Hadoop “Data Lake”
Why is securing Hadoop difficult?

Access by many different users with varying analytic needs

Automatic replication of data across multiple nodes once entered into the HDFS data store

Reduced control if Hadoop clusters are deployed in a cloud environment
HPE Format-Preserving Encryption (FPE)

- Supports data of any format: name, address, dates, numbers, etc.
- Preserves referential integrity
- Only applications that need the original value need change
- Used for production protection and data masking
- Currently in the NIST standardization process

| FPE       | 253-67-2356 | First Name: Gunther Last Name: Robertson  
| SSN: 253-67-2356  
| DOB: 18-06-1972 |
| AES       | 8juYE%Uks&dDFa2345^WFLERG  
| Tax ID: 934-72-2356 |
| Tax ID: 934-72-2356  
| SSN: 934-72-2356  
| DOB: 20-07-1966 |

First Name: Uywjlqo Last Name: Muwruwwbp  
SSN: 253-67-2356  
DOB: 18-06-1972
HPE SecureData

– **HPE Stateless Key Management**
  – No key database to store or manage
  – High performance, unlimited scalability

– **Both encryption and tokenization technologies**
  – Customize solution to meet exact requirements

– **Broad platform support**
  – On-premise / Cloud / Big Data
  – Structured / Unstructured
  – Linux, Hadoop, Windows, AWS, IBM z/OS, HPE NonStop, Teradata, etc.

– **Quick time-to-value**
  – Complete end-to-end protection within a common platform
  – Format-preservation dramatically reduces implementation effort
Options for securing data in Hadoop with HPE SecureData

Legend:
- Unprotected Data
- Application with HPE SecureData Interface Point
- De-Identified Data
- Standard Application

1. Applications and data → HPE SecureData
2. Applications and data → ETL and batch → HPE SecureData
3. Applications and data → Hadoop jobs → HPE SecureData
4. Hadoop Cluster
5. Hadoop jobs and analytics → Hadoop jobs and analytics
6. Applications, analytics and data
7. Egress Zone

HPE SecureStorage

HDFS

Hadoop jobs and analytics

ETL and batch

Landing Zone

Standard Application

Applications, analytics and data

BI Tools and Downstream Applications
## Why Hewlett Packard Enterprise?

Empowering the Data-Driven Organization

<table>
<thead>
<tr>
<th>Experience and expertise</th>
<th>Solution leadership</th>
<th>Market leadership</th>
<th>Flexible and Open</th>
</tr>
</thead>
</table>
| 3000+ global analytics and data management professionals | Proven analytics and compute platforms for all data, environments, and analytics | Gartner’s Magic Quadrant leader for:  
  — Enterprise Data Warehouse and Data Management Solutions for Analytics (2015)  
  — eDiscovery (2015) | Solutions built on open-standards, offering choice and flexibility  
  Strong strategic alliances complementing HPE solutions |