

A photograph of a man and a woman in business attire. The man, on the left, is wearing a dark blue suit and a light blue patterned tie. He is sitting in a chair and looking at a tablet held by the woman. The woman, on the right, is wearing a dark blazer and has her hair pulled back. She is pointing at the tablet with a pen. The background is a bright, modern office space with large windows.

Best Practice Guide

Transforming Big Data into Profitable Business Insight

Best Practices for the Modern Data Centre



Hewlett Packard
Enterprise



A 2015 study by IDG found that 80% of enterprises and 63% of small to mid-sized businesses (SMBs) already have deployed or were planning to deploy Big Data projects. Enterprises were planning to invest an average of \$13.8 million USD in 2015 on Big Data.¹



Big Data Trends

In recent years Big Data has changed in three critical ways:

- **Volume**—New applications, multiple endpoints, and new inputs like social media and the Internet-of-Things (IoT) are increasing the volume of data.
- **Variety**—Structured data such as sales records, manufacturing reports, and HR information is generally tied to dedicated applications housed in silos and now needs to be shared and correlated across functional areas and disparate applications, both within and outside of the company.
- **Velocity**—Businesses need real-time decision support, (i.e., predictive data that will help them make better decisions now).



400 large companies who have already adopted Big Data analytics "...gained a significant lead over the rest of the corporate world."—Bain & Co. Report²

There is more data available today than ever before, but many organisations do not get the full value and business insight they need to run their businesses and stay competitive. Business executives continually try to answer two basic questions to stay competitive:

- **How do I provide my customers with a better experience than my competitors?**
- **How do I do this in a way that is cost-effective?**

Many would be shocked to know that researchers analyse and gather insights from only 1 per cent of the world's data, according to experts.

The good news is that business leaders have an opportunity, to answer these questions. Today's Big Data market is filled with opportunity and progressive organisations on the competitive edge are taking strides to extract valuable insights from that data.

Critical components of success in today's Idea Economy are a data-driven mentality and the ability to collect, process, analyse, and manage data to generate outcomes that improve business results.

Understanding Big Data

The systems of the past can no longer handle the barrage of inputs that companies collect and analyse in order to respond quickly to their business needs. Web-scale companies like Google™ and Amazon actively utilise Big Data and analytics through their own proprietary infrastructure, but few companies have the scale to solve the problem in that manner. IT needs to think differently about the systems they deploy for Big Data and analytics to meet new business demands in a more efficient way.

Leveraging Big Data for better business results

Deriving actionable insight from Big Data assets requires rethinking many things about your business, including the infrastructure that you use to store and process all of that data.

IoT-created data

An aircraft manufacturer is driving innovation by embedding sensors on its engines to track performance and improve engine design for maximum efficiency. As a result, one newly designed aircraft engine is 20% more fuel-efficient than the engines it will replace. The improved design is the direct result of analysing and modelling petabyte-scale data running in a massive, parallel-processing environment.

Simulation-generated data

Oil and gas companies use compute-intensive software applications to turn billions of complex, unstructured data points into interpretable 2D, 3D, and now 4D simulations of potential reservoirs. These visualisations are rendered and further analytics are performed to help make adjustments to current models and create detailed operational guidelines for the drilling process.

New use of historical data

Financial institutions are using predictive analytics to make better decisions with their trading desk applications. Historical data on past performance of financial vehicles allows traders to make better predictive decisions for investment purposes.

Research and patient data

The Technical University of Denmark harnesses huge amounts of research and patient data, making it available at the right time to authorised health professionals to revolutionise diagnosis and treatment based on real information. **[Get more details.](#)**

¹ 2015 Big Data and Analytics Insights into Initiatives and Strategies Driving Data Investments, IDG Enterprise [scribd.com/doc/258158270/2015-Big-Data-and-Analytics-Survey](https://www.scribd.com/doc/258158270/2015-Big-Data-and-Analytics-Survey)

² Big Data: The organizational challenge [bain.com/publications/articles/big_data_the_organizational_challenge.aspx](https://www.bain.com/publications/articles/big_data_the_organizational_challenge.aspx)



Getting the most from data resources requires rethinking the way an enterprise collects, processes, stores, manages, and analyses data.

Using Best Practices for Infrastructure Modernisation

Providing a better customer experience comes from gaining a deeper understanding of the customer. The more data that is available, the better the insights. The same is true when it comes to understanding how to serve each customer more cost-effectively. To better support the business, IT teams need to analyse more data faster and answer these questions:

- **How do I store and manage that much data at scale?**
- **How do I store and manage non-traditional data sources like social media, clickstreams, machine and sensor data, video, etc.?**
- **How do I scale processing to analyse that much data?**

Thinking a bit differently about technologies like storage and compute should provide you with the ability to leverage server-based storage, optimise computing, and intelligently scale your infrastructure.

Leveraging server-based storage—Designed to affordably scale-out to petabyte scale.

The hardware market is inundated with storage and data solutions. Many of these are SAN and array-based solutions, well designed and executed for enterprise data storage. These solutions were developed for enterprise use-cases, such as relational databases, structured data, traditional applications, and smaller data sets.

Most traditional storage systems were designed for data as we thought about it years ago, structured data at terabyte scale. Today, instead of terabytes, we're often talking about tens or even hundreds of petabytes of mostly unstructured data, with exabyte-scale coming in the not-too-distant future. The linear cost growth of scaling traditional storage arrays to multi-petabyte scale and beyond quickly becomes unaffordable.

The alternative to traditional storage is to move data from proprietary storage platforms to a software-defined, server-based approach using technologies like object storage. Direct-attached storage allows you to scale storage by adding commodity-type servers, which costs considerably less over time than scaling traditional storage arrays. Using object storage on server platforms allows IT to easily scale to hundreds of petabytes today and provides a clear and cost-effective path to expand to exabyte or even zettabyte-scale in the future.

Optimising compute—Densely clustered and tuned specifically to the needs of Hadoop and other tools. Collecting, storing, and managing your data is only part of the solution. If you can't make use of the data because it's not easily accessible or is not properly integrated with the analysis element, then you haven't really solved anything. Most businesses have enterprise data warehouse products to analyse structured data that exists in relational databases (RDBMS), but in today's data environment, structured data represents a small fraction of all enterprise data. Unstructured data like social media, sensor data, and video files don't fit within the RDBMS model.

Many enterprises use an open-source project called Hadoop to manage unstructured data through a scalable, server-based storage model. A Hadoop cluster includes many servers with data stored on local disks, with open source tools running on top of the cluster to analyse unstructured data. Hadoop is designed so that it can run on commodity servers, but high-performance compute infrastructures designed specifically for the needs of Big Data can dramatically boost analytic performance while driving down infrastructure and operating costs.

Intelligently scale infrastructure—Utilise reference architectures built on best practices together with flexible appliances to get just the right scale for storage and processing resources. While Hadoop can run on commodity servers, enterprises can drive down costs and optimise performance by moving to high-performance computing appliances that are purpose-built for Hadoop applications with a need for parallel processing, moving high-performance computing out of the laboratory and into the mainstream.

Each Hadoop node includes compute and storage as part of the standard Hadoop architecture. These nodes are the basic building blocks for Hadoop clusters.

If a cluster runs into a storage problem, you simply add another node, but this can become inefficient since each node includes both storage and compute. Using traditional Hadoop methods, you solve a storage backlog with a compute solution, adding a new server node complete with new software licensing costs in order to add new storage. If there is a compute bottleneck, the solution involves adding new storage to the cluster (i.e., a new server node with its attached storage). Having an architecture that splits storage and compute components allows better performance, better density, better power savings, and allows IT to intelligently scale Hadoop clusters. If there's a storage problem, it can be solved by replacing the storage without adding unnecessary compute resources. If there's a compute bottleneck, it can be solved by replacing the compute nodes without adding additional direct attached storage.

The concept of separating out compute and storage is a key tenet of the Hadoop community, with community leaders like Hortonworks and Cloudera supporting this new architecture. By designing infrastructure with the goal of eliminating suboptimal data and technology silos, applications can drive better performance and efficiency while turning data into insights on these purpose-built and optimised HPC infrastructure platforms.

Bottom line

Today's Big Data technologies allow businesses to extract valuable insight from rapidly growing volumes of data. However, getting the most from data resources requires rethinking the way an enterprise collects, processes, stores, manages, and analyses data.

How HPE can help with Big Data

Hewlett Packard Enterprise is one of the few vendors truly innovating by finding new technologies and making wise partnership investments around Big Data. HPE owns all of the hardware touch points—processing, storage, and networking—as well as critical software components like Vertica, IDOL, and Haven to implement your core functionality and analytic tooling. HPE also innovates through its collaboration with ISVs like Hadoop and the open source community. Purpose-built products like the **HPE Apollo System family** are designed around the needs of Big Data (HPE Apollo 4000 Systems), and high-performance computing (HPE Apollo 6000 Systems) and up into supercomputing (HPE Apollo 8000 Systems), bringing tremendous scalability and reliability to the infrastructure-level. As the undisputed leader in the HPC market with more than 34% market share, Hewlett Packard Enterprise delivers solutions that are easy to implement, manage, and support.

How to get started

Whether you want pure open-source or well-established, proprietary solutions, HPE can help you start building your data-centric foundation and power your analytical applications with a one-day **HPE transformation workshop**. This workshop can help crystallise your Big Data vision and identify the impact Big Data will have on your IT infrastructure. HPE integration and implementation services are also available to achieve the desired business results.

Learn more at
hpe.com/info/apollo
hpe.com/info/bigdata



Sign up for updates

★ Rate this document