

Microsoft Big Data

Key Benefits

- **Insights to everyone**, on any data with familiar Microsoft Office and business intelligence (BI) tools.
- **Enrich with the world's data** to unlock hidden patterns through a combination of internal and publicly available data and services, including social media sites.
- **Manage any data, any size, anywhere** through a modern data management platform that supports structured, unstructured, and real-time data with the simplicity of Microsoft Windows and Microsoft SQL Server and the elastic scalability of the cloud.

Why Microsoft?

Microsoft has been doing Big Data long before it was popular. For example, Microsoft Bing analyzes over 100 petabytes of data to deliver high-quality search results.

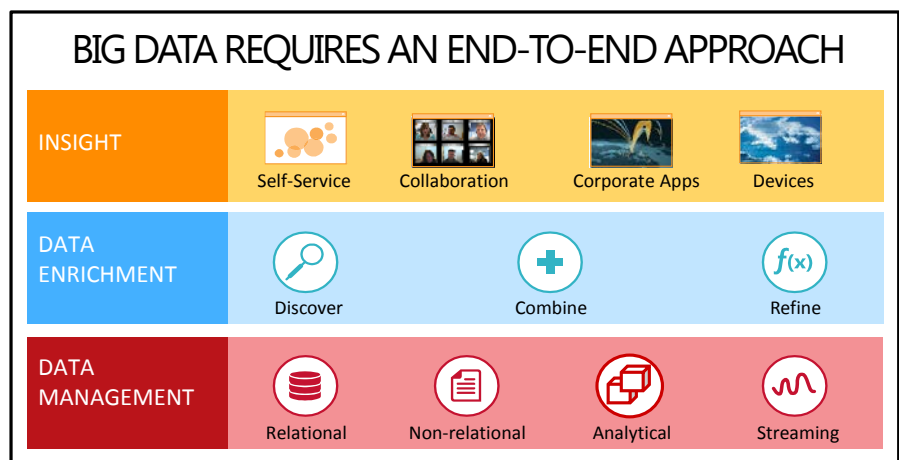
Organizations can use the Microsoft Big Data to unleash actionable insights from any data through familiar tools like Microsoft Office and Microsoft SharePoint. **HDInsight** is Microsoft's new Hadoop-based distribution, built on the Hortonworks Data Platform (HDP). It combines the simplicity of Windows with a fully Apache Hadoop compatible distribution to unlock insights from all data. It also enables customers to uncover new value by connecting to the world's data and services.

Microsoft has a comprehensive Big Data strategy that delivers:

- **A modern data management layer** that supports all data types — structured, semi-structured, and unstructured data at rest or in motion.
- **An enrichment layer** that enhances your data through discovery, combining with the world's data and by refining with advanced analytics.
- **An insights layer** that provides insights to all users through familiar tools like Microsoft Office.

Customer Challenges

Today, organizations are struggling to gain business insight from the unprecedented volume of data they are capturing from internal and external sources, including social media sites. They want to easily mine, process, and analyze these huge amounts of unstructured data (such as files, images, videos, blogs, clickstreams, and geo-spatial data) in real time, without worrying about the required setup. Organizations are looking for an open and flexible platform that enables them to deploy their solutions on-premises or in the cloud.



Solution Advantages

Microsoft Big Data offers the following benefits:

Insights to Everyone

Get actionable insights from all types of data using familiar Microsoft Office and BI tools.

- **Analyze Hadoop data with familiar tools:** Enables analysts and business users to interact with and gain valuable insight from Hadoop data from the very familiar Microsoft Excel interface with the Hive add-in for Excel.
- **Get immersive insights from any data:** Use familiar BI tools like Microsoft SQL Server Analysis Services (SSAS), Microsoft PowerPivot, and Power View through the Hive Open Database Connectivity (ODBC) Driver to analyze unstructured data in Hadoop. Enable self-service BI on relational data using PowerPivot and Power View in SQL Server 2012 and Excel 2013.
- **Drive insights through simplified programming:** Simplify programming on Hadoop through integration with .NET and new JavaScript libraries. Developers can use the new JavaScript libraries to easily write MapReduce programs in JavaScript, and then deploy their JavaScript code from a simple browser.

Enrich with the World's Data

Unleash breakthrough discoveries by combining your data with publicly available data and services, including social media sites. Uncover hidden patterns using the applications and mining algorithms on Windows Azure Marketplace.

- **Discover data:** Unique tools to facilitate discovery of data both within and outside an organization. Discover relevant datasets through automatic recommendations and create a private Data Market to facilitate discovery and sharing of data and analytical models.

- **Combine with the world's data:** Connect to data, smart mining algorithms and people outside firewalls and integrate social media data into business applications. Windows Azure Marketplace offers hundreds of datasets from trusted third party providers.
- **Refine with external data:** Convert raw data into credible consistent data by enriching data through enterprise information management tools and refine your data with advanced analytics using Mahout libraries. Develop bespoke mining algorithms in C++, C#, Python, Ruby and Pearl through Hadoop streaming.

Manage Any Data, Any Size, Anywhere

Manage data of all types through a modern data platform. Enjoy the simplicity of Windows for Hadoop, extend your data warehouses with Hadoop, and take advantage of the elastic scalability of the cloud with Big Data.

- **Enterprise-ready Hadoop with HDInsight:** Get the most reliable, innovative and trusted distribution through Microsoft's HDInsight, an Enterprise-ready Hadoop service built on HDP for Windows Server and Windows Azure. Integration with Active Directory enables IT to secure their Hadoop cluster using enterprise-based security policies. Integration with Microsoft System Center allows IT to easily manage their Hadoop clusters and effectively meet SLAs.
- **Bringing the simplicity and manageability of Windows to Hadoop:** Smart packaging from Microsoft enables simple and straightforward installation of your Hadoop clusters. Simplify the provisioning and management of Hadoop clusters on-premises using Hyper-V and System Center Virtual Machine Manager. Accelerate setup in the cloud by deploying a Hadoop

cluster on Windows Azure in just 10 minutes.

- **Seamlessly extend your data warehouse with HDInsight and PolyBase:** SQL Server 2012 Parallel Data Warehouse with PolyBase provides a smart query processor that seamlessly combines structured and unstructured data using a standard Transact-SQL query. Hadoop connectors for SQL Server and Parallel Data Warehouse appliance enable easy integration of Hadoop with Microsoft Enterprise Data Warehouses and BI solutions. In addition, you can also integrate Hadoop with your relational Data Warehouses using HCatalog.
- **Seamless scale and elasticity of the cloud:** Microsoft offers two options for deploying Hadoop - HDInsight in the cloud or on-premises. Windows Azure HDInsight Service offers elastic peta-scale analytics on Microsoft's cloud platform. It also offers seamless migration to Microsoft HDInsight Server that runs on premise.
- **Open Big Data platform:** Built on HDP, HDInsight is 100% compatible with Apache Hadoop. Microsoft has already submitted proposals to Apache including new JavaScript libraries for Hadoop (developed by Microsoft), as well as the Hive ODBC Driver.

Additional Information

Get more information on Microsoft's Big Data solution:

- www.microsoft.com/bigdata
- www.windowsazure.com/en-us/home/scenarios/big-data/

Download Developer Community Technology Preview (CTP) of Microsoft HDInsight Server:

- www.microsoft.com/bigdata

Download the Microsoft SQL Server Connector for Apache Hadoop:

- www.microsoft.com/download/en/details.aspx?id=27584