

Use Case:

Migrating big data to Azure with Azure Data Box



Big data means more data

A byproduct of the digital transformation is exploding data growth, largely driven by big data, taking the form of anything from the sequencing of millions of genomes from a growing biotech company to the gathering of millions of customer transactions from an online retailer.

Big data can be used to identify trends that help companies make predictive decisions versus reactive decisions to gain a competitive advantage, in the above examples and beyond. Big data benefits don't come free; instead they require massive amounts of IT storage and compute as well as special analytics tools.

Adding to the IT burden is the fact that organizations must provision storage and compute for peak demands, which by definition will create over-provisioning during off-peak periods. This means that a substantial portion of the infrastructure is underutilized and, by definition, wastes capital funds. Moreover, on-premises datacenters may not offer advanced analytics tools, infrastructure, and expertise that help organizations get the most out of big data.

So while big data promises big payoffs, it requires the right platform and tools to fully realize its potential.

Azure offers storage and tools for big data

Azure can help you realize the full potential of your data. The Azure platform provides a pay-as-you-go, Opex-based pricing model along with available analytics and cognitive tools.

Azure is optimized for big data, enabling you to do more:

- Eliminate the need to maintain local systems and facilities, enabling system admins to focus on more proactive responsibilities.
- Avoid compute and storage overprovisioning and underutilization, with the ability to scale up and down with Azure infrastructure.
- Take advantage of flexible storage tiers to store big data securely and at a price point based on access time, cost, and other factors, such as for Hot, Cool, and Archive storage tiers.

USE CASE

Migrating big data to Azure with Azure Data Box



Small: Data Box Disk – 8 TB



Medium: Data Box – 100 TB



Large: Data Box Heavy – 1 PB

- Transform big data into actionable insights with big data analytics solutions from Microsoft, with tools such as HDInsight, Data Lake Analytics—even in real-time with products like Stream Analytics.
- Deliver intelligent decisions, provide better forecasting of personnel and supply chains, increase revenues, and reduce costs.

Your options for data migration

Migrating archive data over a network is always the best first choice. Yet migrating large amounts of data into Azure can be a challenge, especially for organizations with limited network bandwidth. This is especially true with the typical big chunk of data that can be tens or hundreds of terabytes—or even petabytes—in size.

For example, moving 1 PB of archive data over a network with 100 Mbps of available bandwidth will take over three years to complete. That's obviously not a workable migration scenario for archive data.

Devices from the Data Box family are a solid alternative to address this challenge. Data Box devices are designed to move large amounts of data from your datacenter, colo, or vault, into Azure. They are simple to order and use, extremely secure, and very efficient at moving data into the cloud. Data transfer, including filling the Data Box, shipping, and uploading archive data into Azure, is fast.

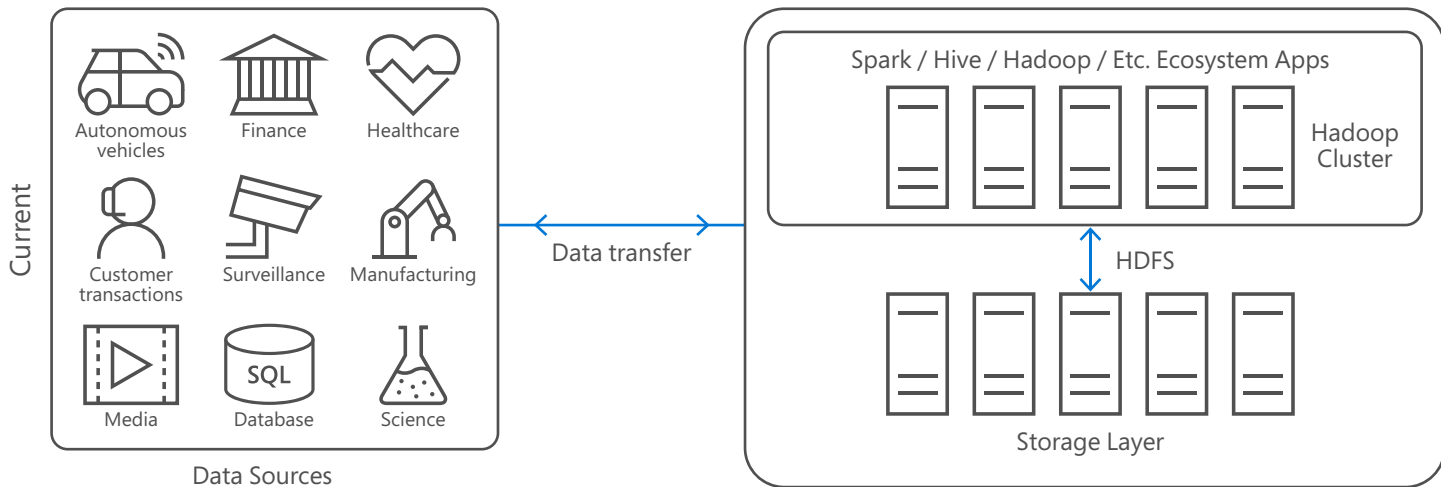
Data Boxes come in three sizes; one is perfect for your needs:

- Small: Data Box Disk—a small form factor 8 TB SSD, orderable in sets of five
- Medium: Data Box—a ruggedized, highly secure 100 TB portable device
- Large: Data Box Heavy—a 1 PB device for lifting massive data payloads to Azure

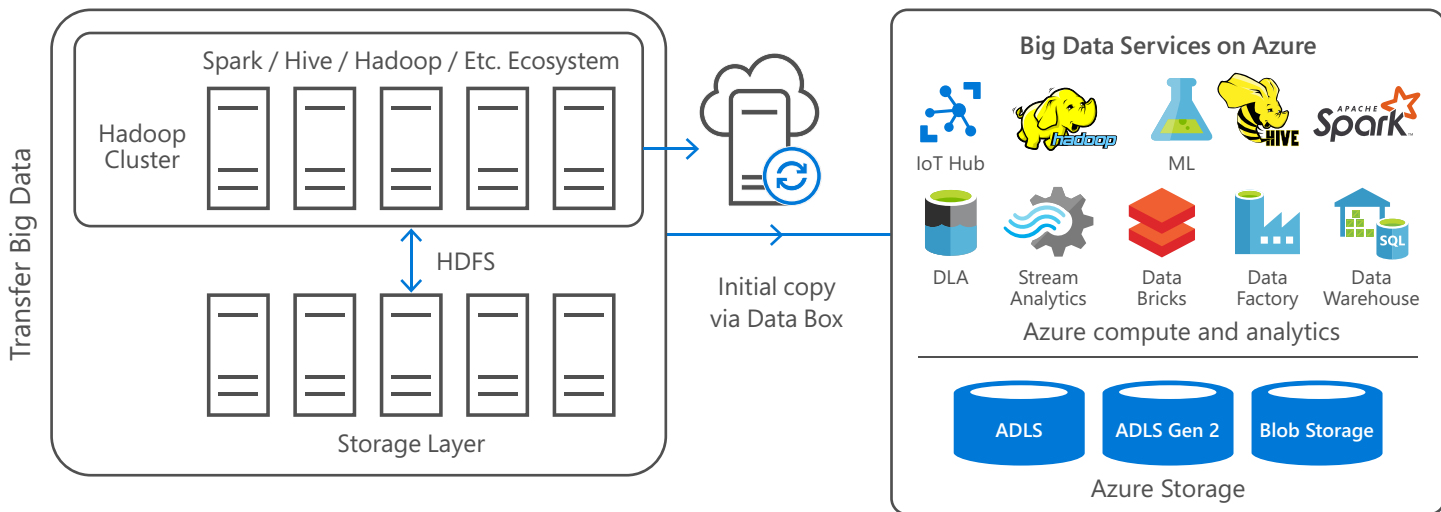
Once the initial bulk archive data has been uploaded into Azure via a Data Box, the ongoing archives will use the network to transfer incremental updates to the cloud, as illustrated in the diagrams on the following page.

For detailed steps on how to migrate big data workloads using the Data Box family of products, visit <http://aka.ms/databoxADLSmigrate>

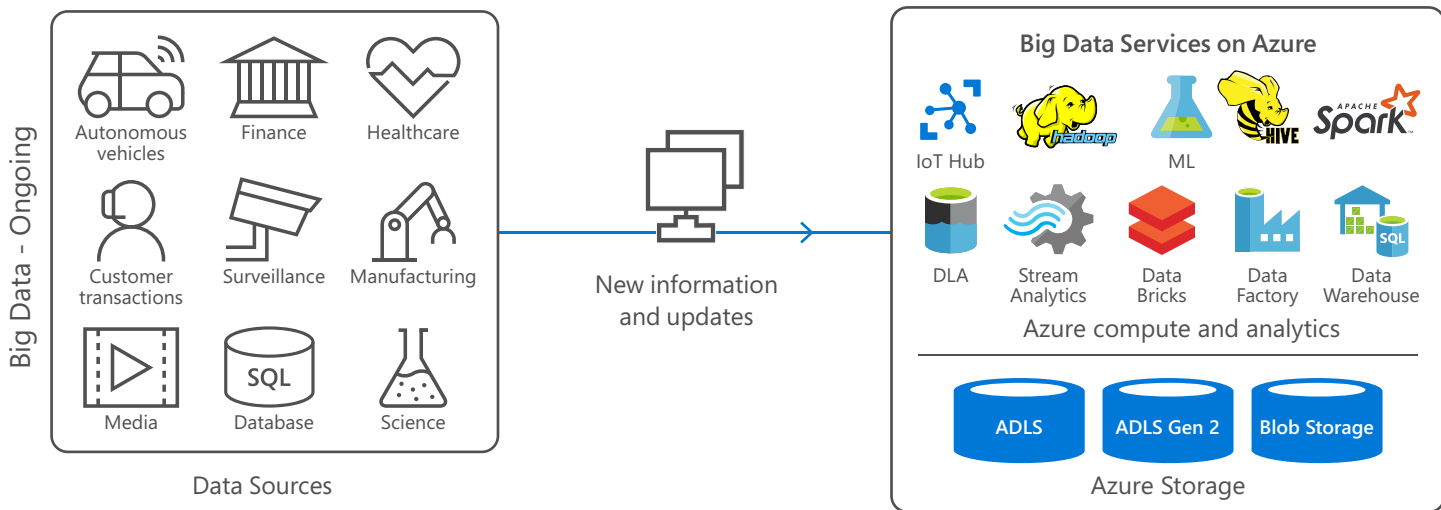
How it Works: Data Box facilitates big data transfer into Azure



Current State: Maintaining big data can create overprovisioning or underutilization scenarios for your datacenters.



Transfer state: Data Box facilitates big data migration into Azure. Users can now take advantage of Azure services.



Transfer state: Data Box facilitates big data migration into Azure. Users can now take advantage of Azure services.

Azure partners integrate with Data Box

A variety of Azure partners offer products and services integrated with Data Box to make big data management easier, faster, and better. The [Data Box partner page](#) contains a list of Azure Data Box partners who specialize in big data and other Azure-related products and services.

Summary

The Azure cloud is an excellent platform for your big data. It provides a safe, cost-effective storage and compute platform for your big data, and also offers tools to analyze and extract maximum value.

However, moving large amounts of big data into Azure can present a big challenge. The data – which can be terabytes and even petabytes in size—can take an unacceptably long time to move across the network, especially in limited bandwidth scenarios. In those cases, when the network isn't an option, the Data Box offers a fast, simple and secure way to transport large amounts of big data into Azure.

Learn more about Azure Data Box at <http://azure.com/DataBox>.

Learn more about optimizing big data in Azure
<http://azure.microsoft.com/solutions/big-data>.