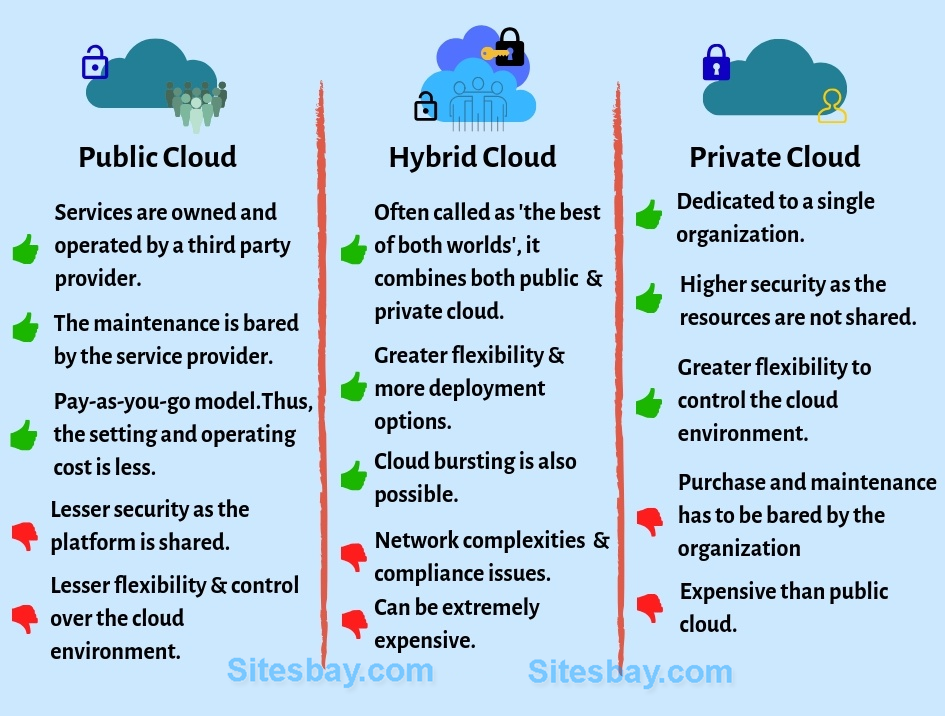
In simple words, Cloud computing is a service that lets you use any computing

service such as a server, storage, networking, databases, and intelligence,

right through your browser without owning anything.

Cloud Computing model enables **your business to communicate and share more easily outside of the traditional methods**. It allows better collaboration between employees, enabling multiple users to share and work on data and files at the same time.



Introduction (what is?)

Microsoft Azure is an ever-expanding set of cloud services that can help your organization meet its business objectives. Build the solutions you need, with the tools you want to use, on a secure and trusted platform. Azure is built to allow you to Invent with Purpose. Azure taps into the powerful combination of data and AI to create the breakthrough experiences that today's businesses require.

Enable digital transformation with Azure

In virtually every industry, new, digitally born companies threaten to disrupt industry incumbents. If you don’t adapt to stay competitive, your competitors will. The taxi industry, for instance, has been fundamentally transformed by the introduction of digitally powered, ride-on-demand services. Whether you're looking to migrate to the cloud or consider yourself a cloud-native company, Azure delivers unique value that helps keep you ahead of your competition.

Why Azure?

When it’s time to find solutions to achieve your business objectives, Azure allows you to build on a platform that you can trust, that will grow with you, and that provides the flexibility to design a solution around your specific capabilities and business objectives. This all happens regardless of your company’s approach to the cloud: on-premises, cloud only, or hybrid.

To help meet your digital transformation goals, Azure is built on four unique values (or promises):

be future ready

build on your terms

operate hybrid seamlessly

trust your cloud

Trust your cloud: Build and deploy on a platform you can trust - secure and compliant.

We’ll continue to explore the unique Azure differentiators as we dive into Azure solutions and services.

Azure solutions

Microsoft Azure offers a set of flexible solutions that evolve with your business needs. This unit is an overview of the Azure solutions that will be covered later in more depth.

Apps and infrastructure migration and modernization

Apps and infrastructure migration and modernization focuses on the various ways applications and workloads can be moved to the Azure cloud. You can move existing on-premises infrastructure strategically to the cloud, and Azure enables the modernization of applications in the cloud.

Data estate migration and modernization

Data estate migration and modernization focuses on the transformation of your data estate. Microsoft offers a comprehensive, integrated, secure, and compliant platform for data solutions.

Artificial intelligence (AI)

Artificial intelligence, combined with Microsoft tools, services, and infrastructure, can make AI real for your organization. Use it whether you're looking to unlock insights from your latent data with knowledge mining, develop your own AI models with machine learning, or build AI apps and agents.

The Internet of Things (IoT)

IoT allows your organization to create actionable business insights by taking advantage of advanced machine learning and AI capabilities. It helps you make sense of massive amounts of data. With solutions such as Azure IoT Central and Azure IoT Hub, your business can make better, data-driven decisions on a secure and trusted platform.

Modern business intelligence (BI)

Modern BI solutions and tools help you understand how to transform your company’s data to support informed decision-making, gain deeper insight into your data to stay in the know, and recognize trends as they happen.

What is Azure Data Factory?

In the world of big data, raw, unorganized data is often stored in relational, non-relational, and other storage systems. However, on its own, raw data doesn't have the proper context or meaning to provide meaningful insights to analysts, data scientists, or business decision makers.

Azure Data Factory is a managed cloud service that's built for these complex hybrid extract-transform-load (ETL), extract-load-transform (ELT), and data integration projects.

For example, imagine a gaming company that collects petabytes of game logs that are produced by games in the cloud. The company wants to analyze these logs to gain insights into customer preferences, demographics, and usage behavior. It also wants to identify up-sell and cross-sell opportunities, develop compelling new features, drive business growth, and provide a better experience to its customers.

To analyze these logs, the company needs to use reference data such as customer information, game information, and marketing campaign information that is in an on-premises data store. The company wants to utilize this data from the on-premises data store, combining it with additional log data that it has in a cloud data store.

To extract insights, it hopes to process the joined data by using a Spark cluster in the cloud (Azure HDInsight), and publish the transformed data into a cloud data warehouse such as Azure Synapse Analytics to easily build a report on top of it. They want to automate this workflow, and monitor and manage it on a daily schedule. They also want to execute it when files land in a blob store container.

Azure Data Factory is the platform that solves such data scenarios. Using Azure Data Factory, you can create and schedule data-driven workflows (called pipelines) that can ingest data from disparate data stores. You can build complex ETL processes that transform data visually with data flows or by using compute services such as Azure HDInsight Hadoop, Azure Databricks, and Azure SQL Database.

Additionally, you can publish your transformed data to data stores such as Azure Synapse Analytics for business intelligence (BI) applications to consume. Ultimately, through Azure Data Factory, raw data can be organized into meaningful data stores and data lakes for better business decisions.

Top-level concepts

An Azure subscription might have one or more Azure Data Factory instances (or data factories). Azure Data Factory is composed of below key components.

* Pipelines
* Activities
* Datasets
* Linked services
* Data Flows
* Integration runtimes

Pipeline

A data factory might have one or more pipelines. A pipeline is a logical grouping of activities that performs a unit of work. Together, the activities in a pipeline perform a task. For example, a pipeline can contain a group of activities that ingests data from an Azure blob, and then runs a Hive query on an HDInsight cluster to partition the data.

The benefit of this is that the pipeline allows you to manage the activities as a set instead of managing each one individually. The activities in a pipeline can be chained together to operate sequentially, or they can operate independently in parallel.

Activity

Activities represent a processing step in a pipeline. For example, you might use a copy activity to copy data from one data store to another data store. Similarly, you might use a Hive activity, which runs a Hive query on an Azure HDInsight cluster, to transform or analyze your data. Data Factory supports three types of activities: data movement activities, data transformation activities, and control activities.

Datasets

Datasets represent data structures within the data stores, which simply point to or reference the data you want to use in your activities as inputs or outputs.

Linked services

Linked services are much like connection strings, which define the connection information that's needed for Data Factory to connect to external resources. Think of it this way: a linked service defines the connection to the data source, and a dataset represents the structure of the data. For example, an Azure Storage-linked service specifies a connection string to connect to the Azure Storage account. Additionally, an Azure blob dataset specifies the blob container and the folder that contains the data.

Linked services are used for two purposes in Data Factory:

To represent a **data store** that includes, but isn't limited to, a SQL Server database, Oracle database, file share, or Azure blob storage account. For a list of supported data stores, see the [copy activity](https://docs.microsoft.com/en-us/azure/data-factory/copy-activity-overview) article.

To represent a **compute resource** that can host the execution of an activity. For example, the HDInsightHive activity runs on an HDInsight Hadoop cluster. For a list of transformation activities and supported compute environments, see the [transform data](https://docs.microsoft.com/en-us/azure/data-factory/transform-data) article.

data flows

Create and manage graphs of data transformation logic that you can use to transform any-sized data. You can build-up a reusable library of data transformation routines and execute those processes in a scaled-out manner from your ADF pipelines. Data Factory will execute your logic on a Spark cluster that spins-up and spins-down when you need it. You won't ever have to manage or maintain clusters.

Integration Runtime

In Data Factory, an activity defines the action to be performed. A linked service defines a target data store or a compute service. An integration runtime provides the bridge between the activity and linked Services. It's referenced by the linked service or activity and provides the compute environment where the activity either runs on or gets dispatched from. This way, the activity can be performed in the region closest possible to the target data store or compute service in the most performant way while meeting security and compliance needs.

[Data lakes](https://www.talend.com/resources/what-is-data-lake/) and [data warehouses](https://www.talend.com/resources/what-is-data-warehouse/) are both widely used for storing [big data](https://www.talend.com/products/big-data/), but they are not interchangeable terms. A data lake is a vast pool of raw data, the purpose for which is not yet defined. A data warehouse is a repository for structured, filtered data that has already been processed for a specific purpose.

.