



Introduction to Azure Data Factory



Session Objectives And Takeaways

- Learn Azure Data Factory Concepts
- Learn how you can use Azure Data Factory to build an end to end solution
 - Take raw telco call data
 - Use HDInsight to enrich the data
 - Score enriched dataset against an Azure ML model
 - Push results into an Azure SQL DB
 - Use Power BI to visualize the results

The need for trusted information production



Marketing
campaign analysis

Interactive Entertainment



User and
product profiling

Interactive
Entertainment/Retail



Customer
sentiment analysis

Interactive
Entertainment/Retail



Personalized product
recommendation

Retail



Customer shopping
behavior analysis

Retail



Pricing optimization

Retail



Corrective and predictive
maintenance and repairs

Manufacturing (IOT)



Operational telemetry
and health reporting

Online Services



Actuarial modelling and
reporting automation

Financial Services



Financial risk modelling
and analysis

Financial Services





Data Sources

Raw Materials

Ingest

Acquire Raw
Materials

Transform & Analyze


Transform raw materials into
“finished goods”

Publish


Deliver

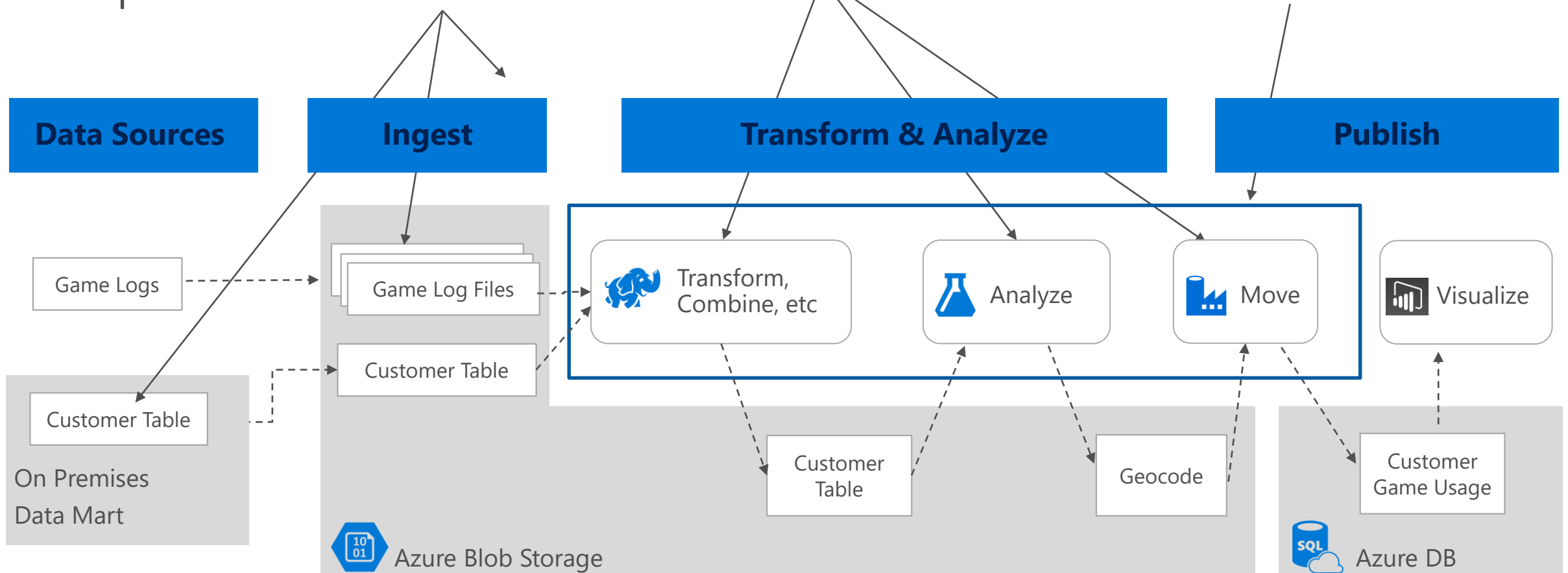


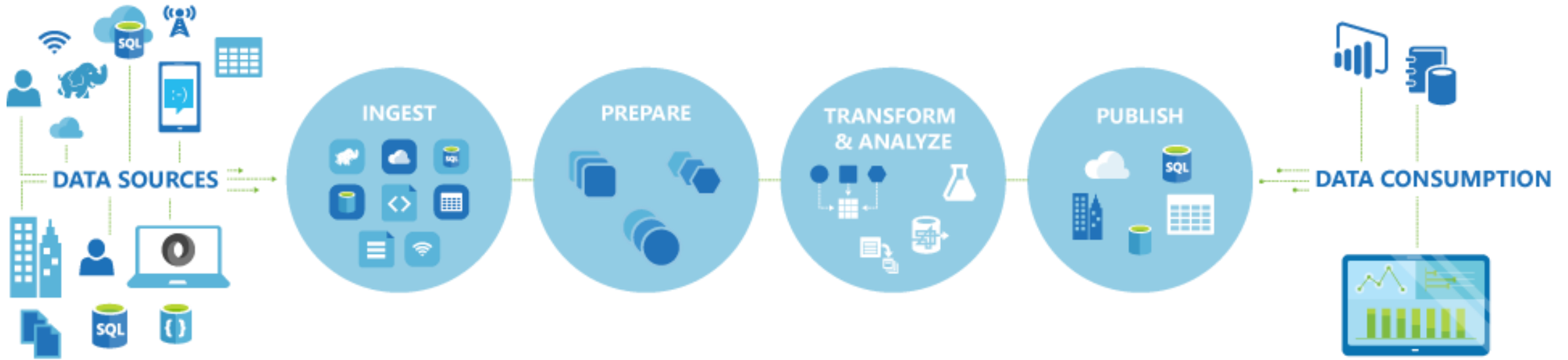
Data Factory Concepts

 **Data Set**
(Collection of files, DB table, etc)

 **Activity:** a processing step
(Hadoop job, custom code, ML model, etc)

 **Pipeline:** a sequence of activities (logical group)

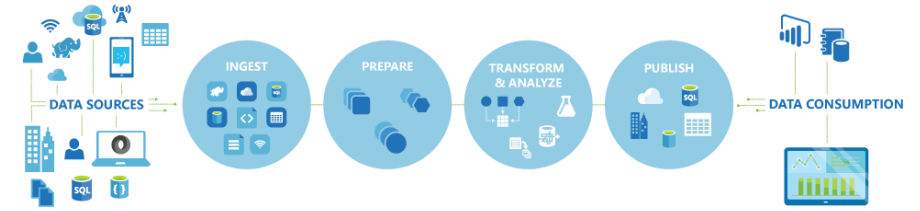




Azure Data Factory

A managed cloud service for building & operating data pipelines

Azure Data Factory

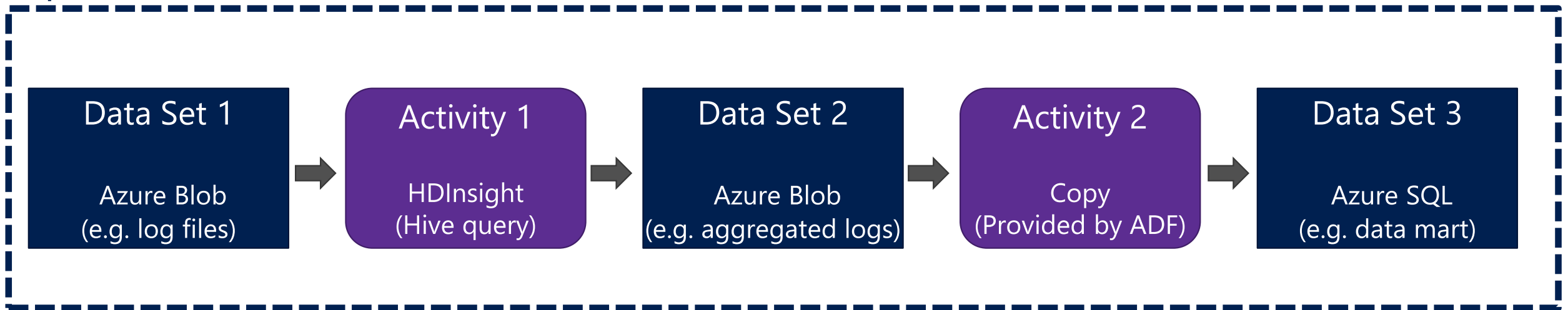


- **Orchestrate, monitor & schedule** - Compose data processing, storage & movement services (on premises & cloud)
- **Automatic infrastructure management**
 - Combine pipeline intent w/ resource allocation & mgmt
 - Globally available data movement capability
- **Single pane of glass** - One place to manage and monitor your data pipelines

Example

Process log files with Hadoop & load the results into my Azure SQL Database

Pipeline (Active Period: July 2015 to July 2016)



Additional Common Use Cases

- **Customer Profiling:** Product Recommendations, customized offers, customer usage tracking, customized marketing, etc
- **Operational Excellence:** Telemetry systems, usage optimization, etc
 - Load your Data Lake
 - Load your Data warehouse/mart

Get Started!

Visit www.azure.com/df

Github Code Sample Repository

<https://github.com/Azure/Azure-DataFactory>

Microsoft Case Studies

[Milliman - Actuarial Automation](#)

[Rockwell Automation - Operational Excellence](#)

Visit Azure Data Factory Feedback Forums on MSDN and Azure UserVoice

The screenshot shows the Microsoft Azure Data Factory website. The header includes the Microsoft Azure logo, navigation links (Why Azure, Features, Documentation, Pricing, Downloads, Partners, Blog, Community, Support), and a 'FREE TRIAL' button. The main content area features the 'Data Factory' title with a 'PREVIEW' badge, followed by the tagline 'Compose and orchestrate data services at scale'. Below this are two columns of bullet points: 'Compose storage, movement, and processing pipelines', 'Transform structured and unstructured data', 'Work with on-premises and cloud data', 'Monitor data pipeline health', 'Visualize data lineage', and 'Feed applications with trusted data'. A 'Get started' button is prominently displayed. Below the main content, there is a section titled 'Customers using Azure Data Factory' with logos for Milliman, Pier1 imports, Rockwell Automation, and Ziosk. The bottom section is titled 'Use diverse data storage and processing services' and includes a diagram showing data flow from various sources (SQL, Blob, Table) into a Data Factory pipeline. The final section is titled 'Transform data into information' and discusses the challenges of data transformation and how Data Factory can help.

Microsoft Azure

SALES 1-800-867-1389 MY ACCOUNT PORTAL Search

Why Azure Features Documentation Pricing Downloads Partners Blog Community Support FREE TRIAL

Data Factory ^{PREVIEW}

Compose and orchestrate data services at scale

- ✓ Compose storage, movement, and processing pipelines
- ✓ Transform structured and unstructured data
- ✓ Work with on-premises and cloud data
- ✓ Monitor data pipeline health
- ✓ Visualize data lineage
- ✓ Feed applications with trusted data

Get started >


Pricing details > Documentation >

Customers using Azure Data Factory

Milliman Pier1 imports Rockwell Automation Ziosk

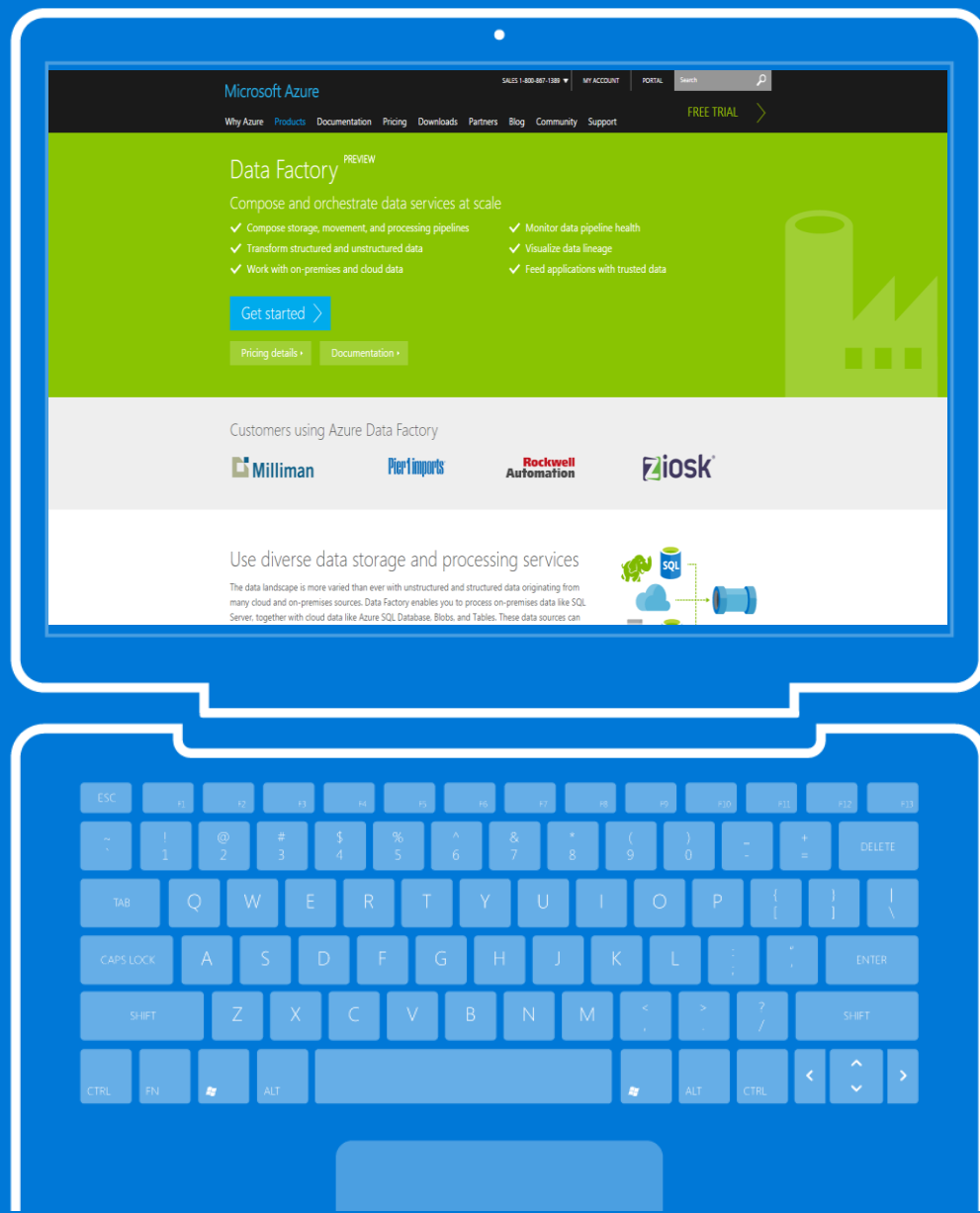
Use diverse data storage and processing services

The data landscape is more varied than ever with unstructured and structured data originating from many cloud and on-premises sources. Data Factory enables you to process on-premises data like SQL Server, together with cloud data like Azure SQL Database, Blobs, and Tables. These data sources can be composed, processed, and monitored through simple, highly available, fault-tolerant data pipelines.



Transform data into information

Combining and shaping complex data can take more than one try to get right, and changing data models can be costly and time consuming. Using Data Factory you can focus on transformative



Let's Get started!

Deploying your first Azure Data Factory

