



Microsoft Data Platform

Bringing Big Data and Data Warehousing together

Nakul Joshi
Cloud Solution Architect
nakul.joshi@microsoft.com

Agenda

- Introduction
- Modern DW - Advanced Analytics and Big Data
- Azure Big Data and Azure Data Landscape
- Common Big Data and Advanced Analytics Scenarios
- Advanced Analytics and Big Data Patterns

Objectives

Understand the technologies and tools that can be used to build an Advanced Analytics and Big Data solution

Prepare you for making architectural decisions for your solution

Partner Enterprise Architect Team (PEAT) Initiatives for FY19 (Opportunity Support)



Partner-led Proactive Proposals

32+ Modern solution scenarios based on Microsoft Cloud Platform to drive Cloud consumption and net new revenue for Partner and Microsoft



Pre-Sales Proposal support

Microsoft Architects available to help you bid and win on Microsoft Cloud solutions for RFI / RFP / Unsolicited scenarios



Depth Account Engagement

PEAT CSA to go deep on 5 identified Partner accounts to accelerate Cloud consumption



Large Transformations

Microsoft Architects available for Solution approach and Architecture for Market Making transformation wins on Microsoft Cloud



Just in Time Enablement

Microsoft Architects available for project specific Azure Services enablement (~60-80 hours) and review of implementation plan

Techs supported

Cloud :
Public and Hybrid cloud solution with Azure, M365 (E3 and E5 SKU), EMS, D365

Advanced Analytics - Hybrid solutions:
Advanced Analytics, Big Data solutions, HDInsight, Azure SQL DB

First Party Workloads :
SAP, Oracle and Open Source on Azure

Global Presence

6 Countries have physical presence of PEAT Architects. PEAT operates all opportunities out of these bases, leveraging support team in India.

US, Singapore, India, UAE, UK



60+ Solution Architects with deep expertise across Microsoft Products and other First Party products like Oracle, SAP, Linux

Qualification Criteria

Any opportunity with Enterprise customers resulting in

- > \$75k Azure Consumption **or**
- > \$300K On-Premise product + \$50k Azure for Hybrid workloads **or**
- O365 (E3 → 10K seats or E5 SKU → 1K+ seats, E5 with Enterprise Voice → 50 seats)
- > \$300K Dynamics CRM **or**
- > 300 users for CRM Online

Note: No support for Desktop products, traditional on-premise SQL, Private Cloud / Embedded Systems. HoloLens and Kinect

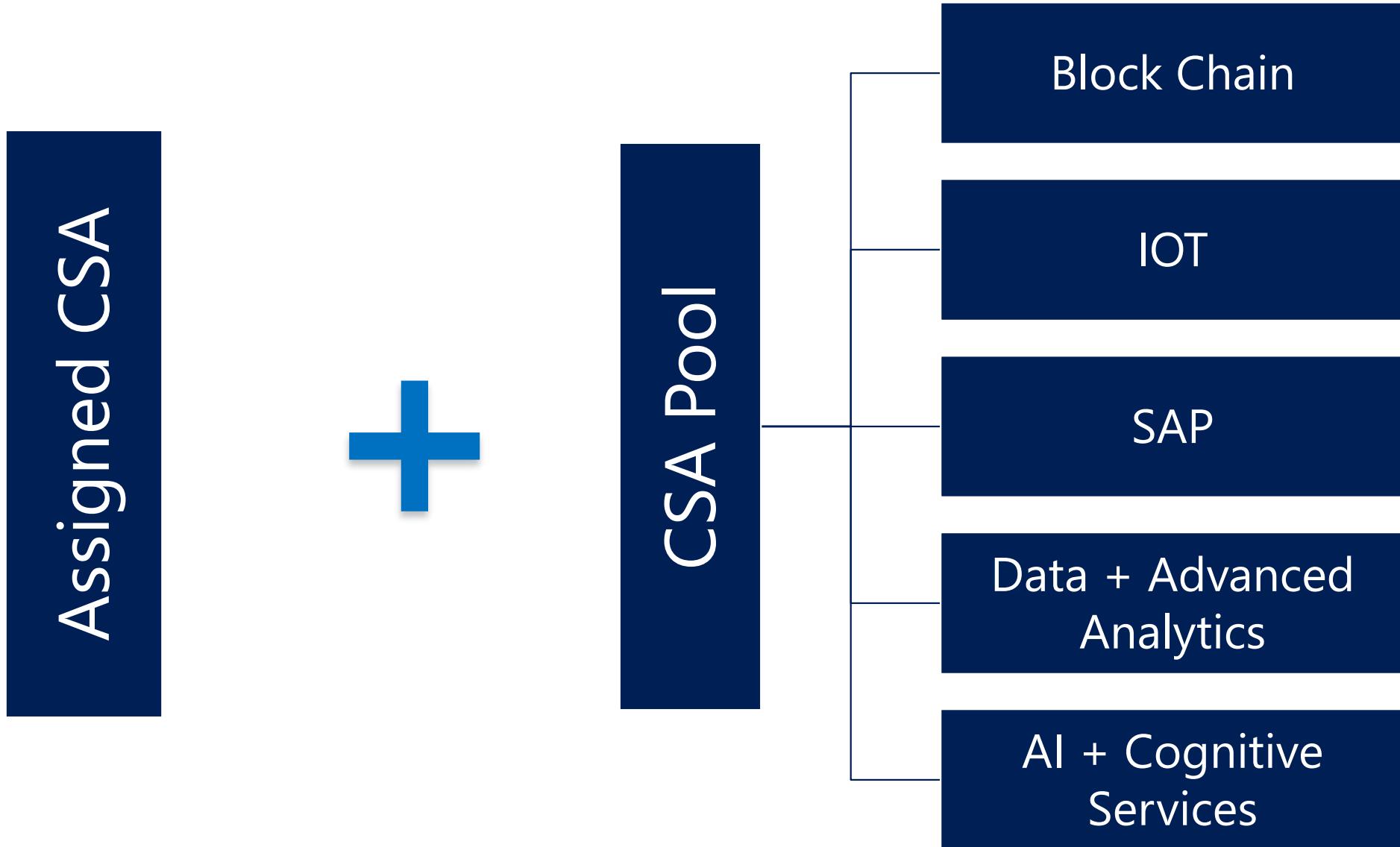
How to Engage

Engage PEAT for qualified opportunities through

- PDM / Alliance Managers
- Respective PTS/CSA
- CSA for EY – Nakul Joshi (nakul.joshi@microsoft.com)
- Global PTS – girishcj@microsoft.com



OCP – CSA support structure



Introduction and Context



Data will grow to
44 ZB in 2020

Today, 80% of organizations
adopt cloud-first strategies

AI investment
increased by
300% in 2017

There are barriers to getting value from data



Data silos



Incongruent
data types



Performance
constraints



Complexity of
solutions



Rising costs

Derive real value from your data

Data silos	Incongruent data types	Performance constraints	Complexity of solutions	Rising costs
One hub for all data	Support for diverse types of data	Unlimited data scale	Familiar tools and ecosystem	Lower TCO

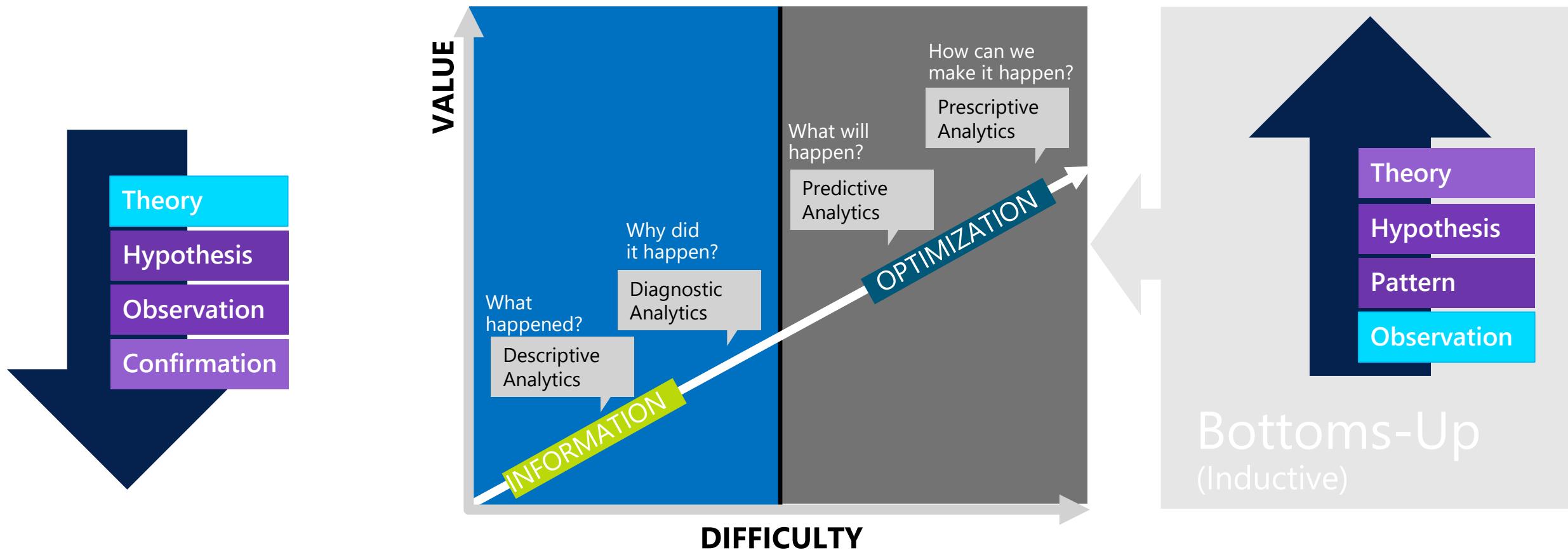
On-premises, hybrid, Azure

Data Analysis Paradigm Shift

OLD WAY: Structure -> Ingest -> Analyze

NEW WAY: Ingest -> Analyze -> Structure

Two approaches for Analytics



What is a modern data warehouse?

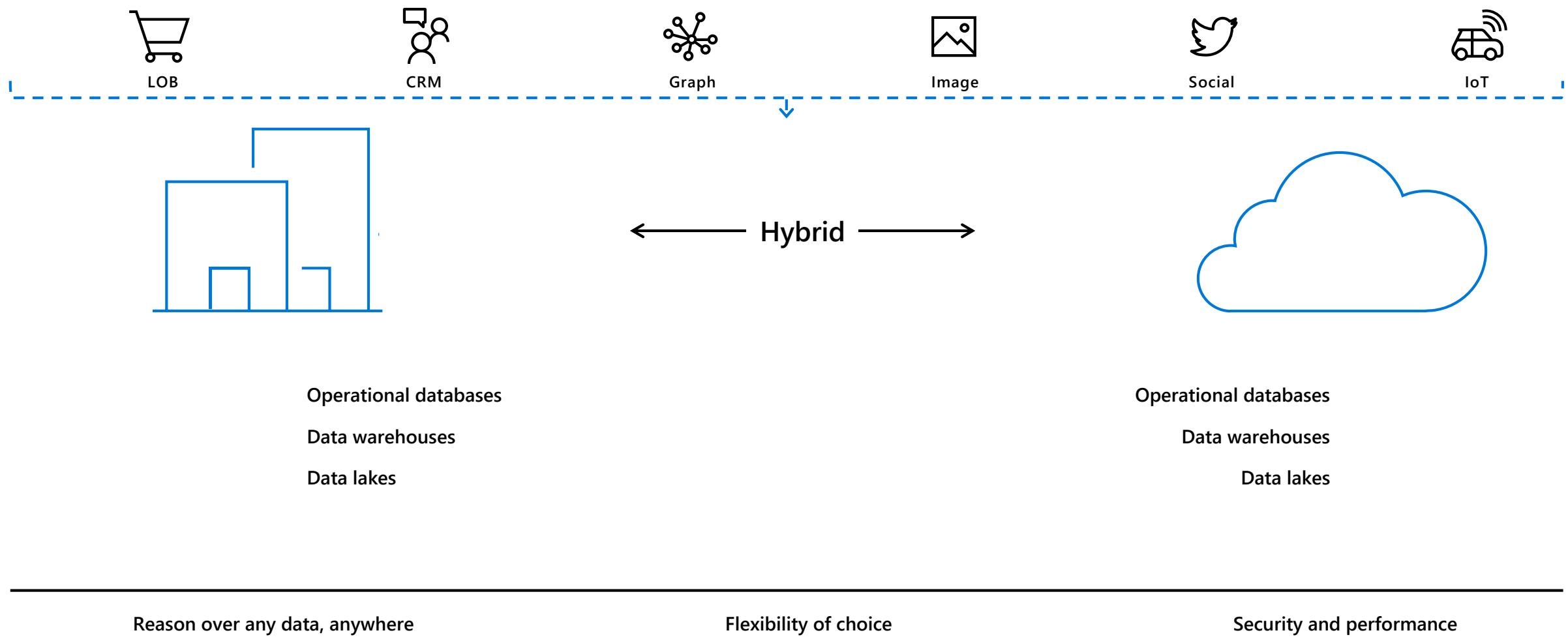
- Integrated Big Data Platform
- Near Real-Time
- Advanced Analytics
- Multi-Structured Data
- Performance
- Scalable (Growth Demand)



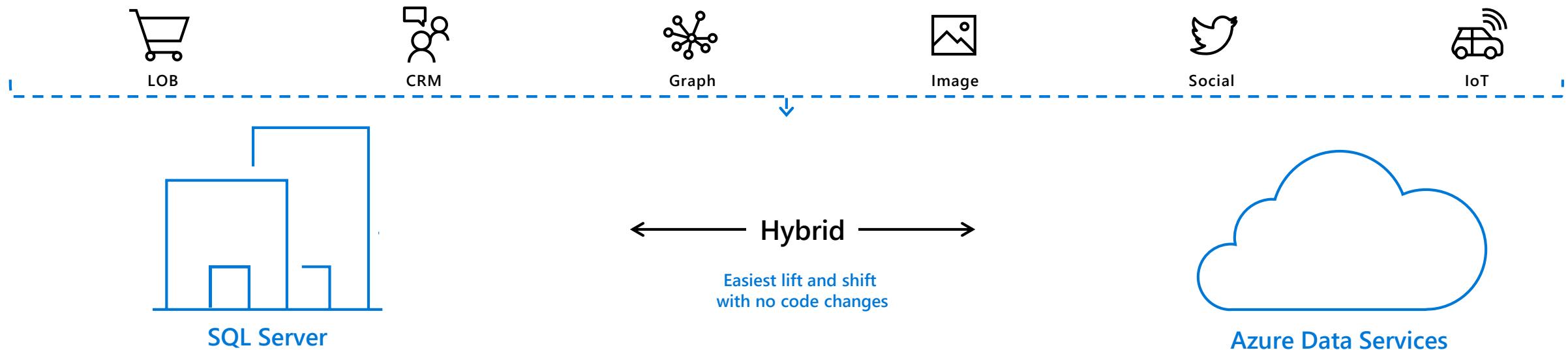
Source: Russom, P. (2013) *The Modern Data Warehouse: What Enterprises Must Have Today and What They'll Need in the Future*, TWDI

Understanding the customer landscape for Big Data and advanced analytics

The modern data estate



The Microsoft offering



Industry leader 4 years in a row Operational databases

#1 TPC-H performance Data warehouses

T-SQL query over any data Data lakes

Operational databases 70% faster
Data warehouses 2x the global reach
Data lakes 99.9% SLA

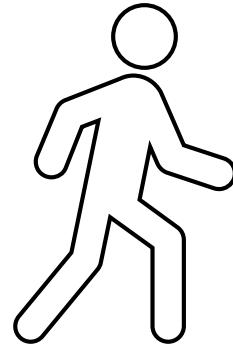
AI built-in | Most secure | Lowest TCO

Reason over any data, anywhere

Flexibility of choice

Security and performance

Our customers



Traditionalists

Have strong processes and practices, need prescriptive guidance, mature stack players



Early adopters

Have lean structures and are agile and flexible, usually on the latest-and-greatest tech

Understanding the Azure portfolio for Big Data and advanced analytics

The Azure data landscape



Azure Data Factory



Azure Import/Export service



Azure CLI



Azure SDK



Azure SQL DB



Azure Cosmos DB



Azure SQL data warehouse



Azure Analysis Services



Power BI



Azure IoT Hub



Azure event hubs



Kafka on Azure HDInsight



Azure Blob Storage



Azure Data Lake Store



Azure Data Lake Analytics



Azure HDInsight



Azure Databricks



Azure ML



ML Server



Azure Databricks



Azure Search



Azure Data Catalog



Azure Stream Analytics



Azure HDInsight



Azure Databricks



Bot service



Cognitive services



Azure ExpressRoute



Azure Active Directory



Azure network security groups



Azure key management service



Operations Management Suite

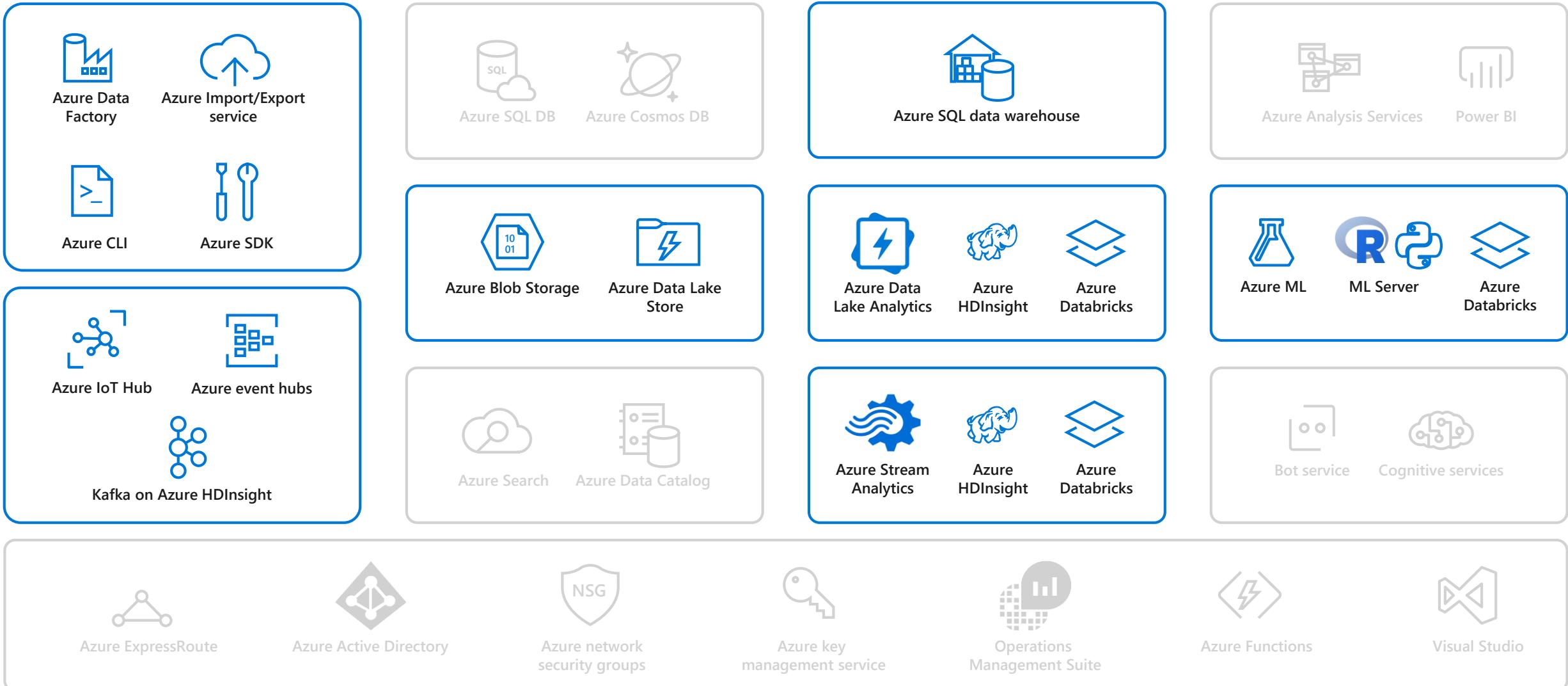


Azure Functions



Visual Studio

The Azure Advanced Analytics and Big Data landscape



Common Big Data and advanced analytics scenarios

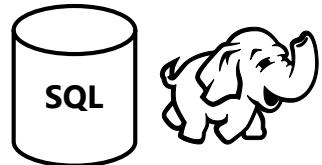
Things to note



- There are no right or wrong solutions, only optimal solutions
- We lead with certain solutions and customize based on customer scenarios
- Customer voice and product and service maturity govern lead solutions
- Consider price and performance, ease of use, and ecosystem acceptance as factors
- Everything is fluid - a lead solution today might be non-optimal tomorrow, based on the factors above and new releases

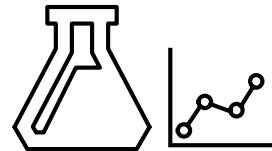
Solution scenarios

Big Data and advanced analytics



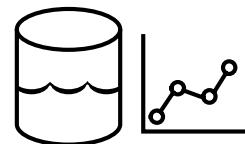
Modern data warehousing

"We want to integrate all our data—including Big Data—with our data warehouse"



Advanced analytics

"We're trying to predict when our customers churn"

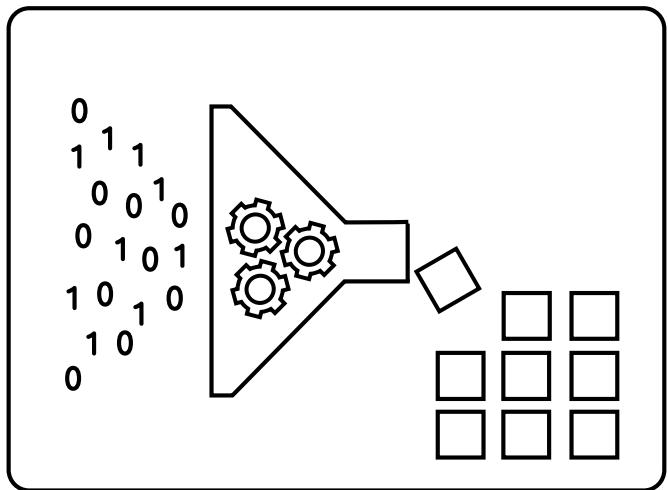


Real-time analytics

"We're trying to get insights from our devices in real-time"

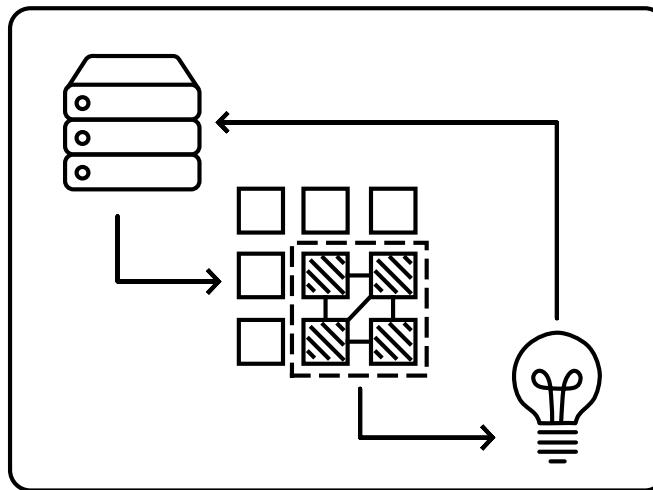
Canonical operations

Load and ingest



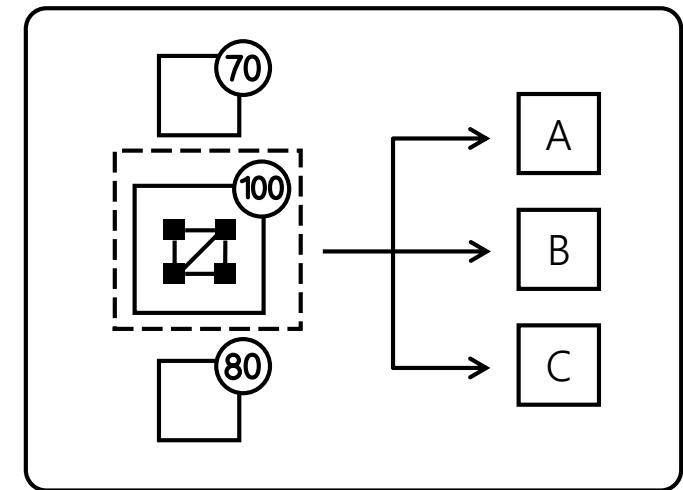
Transfer and store

Process



Process and clean

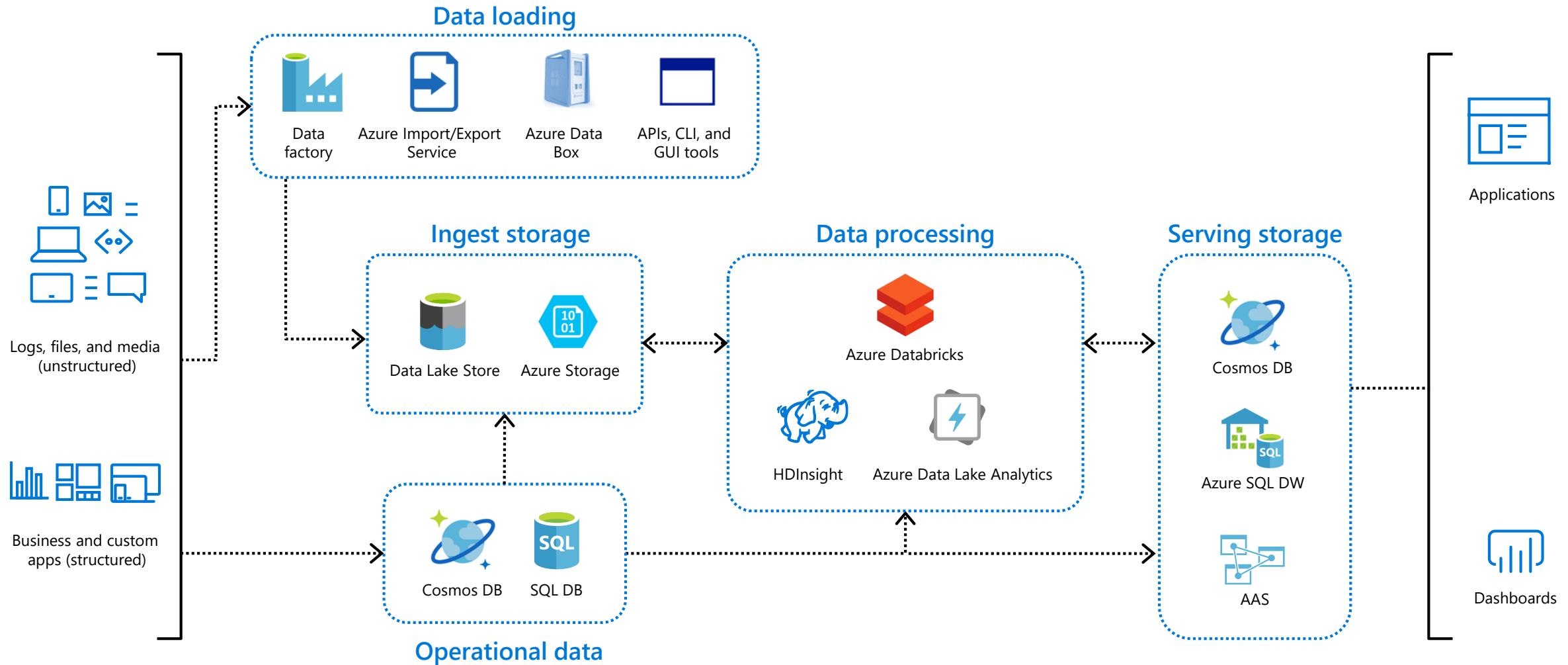
Serve



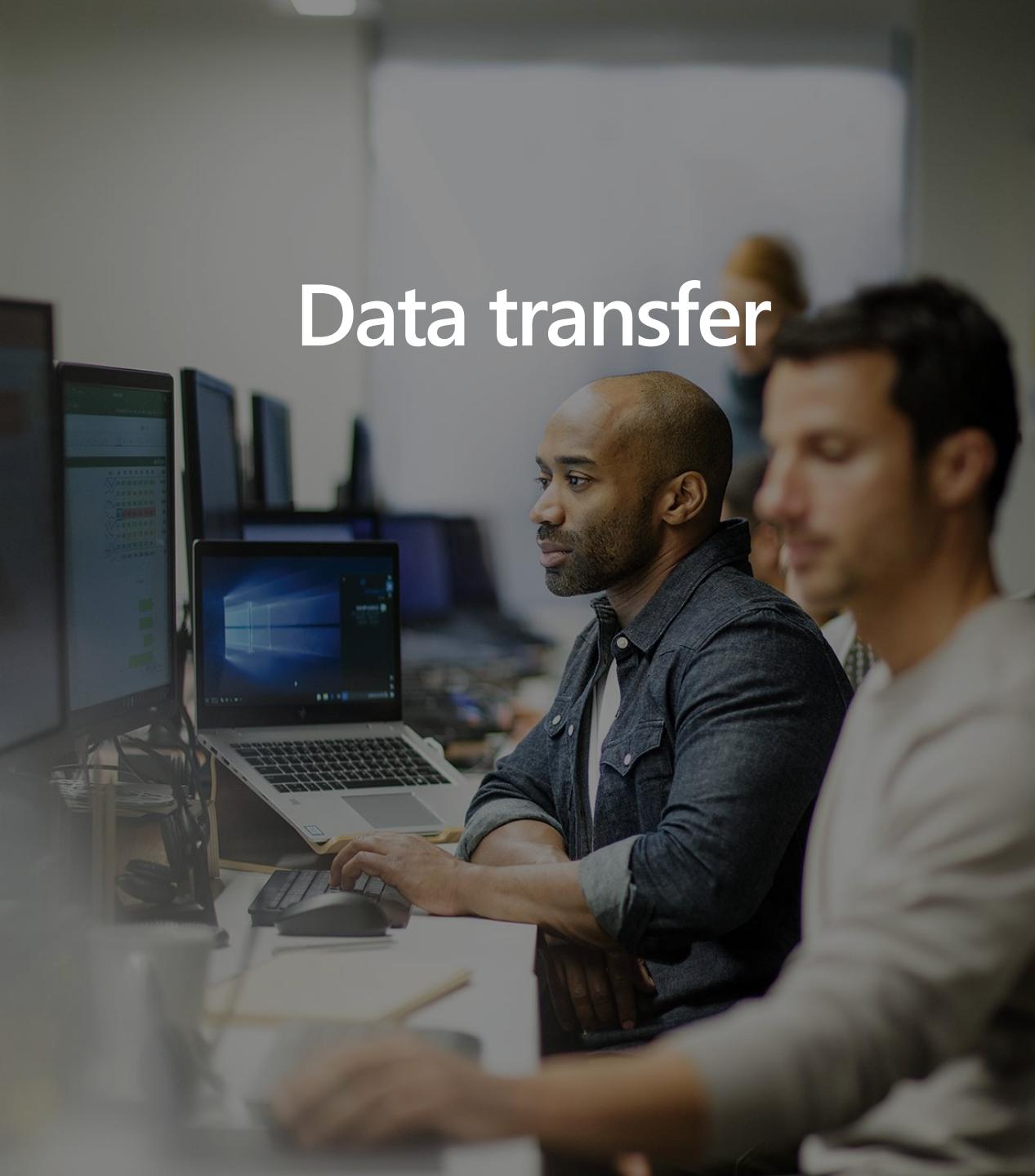
Serve and analyze

Advanced Analytics/Data warehousing pattern in Azure

Loading and preparing data for analysis with a data warehouse



Data Transfer

A photograph showing two men in an office environment. One man, a Black man with a beard, is in the foreground, focused on a laptop screen. Another man, a white man, is partially visible behind him, also working at a desk. They are surrounded by multiple computer monitors displaying various data and software interfaces.

Data transfer

The process and tools used to move data from the source to the initial destination for processing

Network data transfer

A side-by-side comparison of the capabilities and features

Requirement	Dedicated low latency (5 to 10 ms)	Secure transfer	Reliability
Public internet		X	Good
VPN		X	Good
ExpressRoute	X	X	Best

Physical data transfer

A side-by-side comparison of the capabilities and features

	Azure Import/Export service	Azure Data Box
Form factor	Internal SATA HDDs or SSDs	Secure, tamper-proof, single hardware appliance
Microsoft manages shipping logistics	No	Yes
Integrates with partner products	No	Yes
Reduces administrative overhead of purchasing, preparing, and copying data to multiple drives	No	Yes

Command line tools and APIs

A side-by-side comparison of the capabilities and features

	Azure CLI	AzCopy	PowerShell	AdlCopy	Distcp	Sqoop	PolyBase	Hadoop command line
Compatible platforms	Linux, OS X, Windows	Linux, Windows	Windows	Linux, OS X, Windows	Hadoop/ HDInsight *	Hadoop/ HDInsight *	Windows with SQL Server instance, Azure SQL Data Warehouse	Hadoop/ HDInsight *
Copy to and from relational database	No	No	No	No	No	Yes	Yes	No
Copy to blob storage	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Copy from blob storage	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Copy to data lake store	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Copy from data lake store	No	No	Yes	Yes	Yes	Yes	Yes	No
Optimized for Big Data	No	No	No	Yes **	Yes	Yes	Yes ***	Yes

* Can be run from a command line or shell session invoked from Linux, OS X, or Windows

** AdlCopy is optimized for transferring Big Data when used with a Azure Data Lake Analytics account

*** PolyBase [performance can be increased](#) by pushing computation to Hadoop and using [PolyBase scale-out groups](#) to enable parallel data transfer between SQL Server instances and Hadoop nodes

Graphical user interface tools

A side-by-side comparison of the capabilities and features

	Azure Storage Explorer	Azure Portal*	Azure Data Factory
Copy to and from relational database	No	No	Yes
Copy to blob storage	Yes	No	Yes
Copy from blob storage	Yes	No	Yes
Copy to data lake store	No	No	Yes
Copy from data lake store	No	No	Yes
Upload to blob storage	Yes	Yes	Yes
Upload to data lake store	Yes	Yes	Yes
Orchestrate data transfers	No	No	Yes
Custom data transformations	No	No	Yes

* Azure Portal in this case means using the web-based exploration tools for blob storage and data lake store. This excludes using the portal for other services, such as Azure Data Factory.

Data warehousing pattern in Azure

Loading data into ingest storage



Data Storage

A photograph of a man in a blue baseball cap and a dark zip-up hoodie standing in a data center aisle. He is reaching up towards a row of server racks, his right hand near a handle or connector. The racks are filled with various components and cables. The lighting is dramatic, coming from the side and above, creating strong highlights and shadows. The overall atmosphere is technical and industrial.

Data storage and data ingest

The storage that persists the transferred data and is consumed by subsequent processing

File storage requirements

Requirement	Comment
Capacity	Should be able to store terabytes or petabytes of data economically. File storage should be able to store any number of objects and associated metadata.
Performance	Should be able to store the incoming data as fast as it arrives. Should support high bandwidth, high throughput, and low-latency writes.
Multiple tiers	Should support storing data for extended periods of time—that is months or years—economically. Should have multiple storage tiers (hot, cold, and archival).
Multiple object sizes	Objects can be stored individually—could be as small as few tens of bytes—or as data sets that create large objects—from several gigabytes to terabytes.
Replication	Should provide data replication that suits your needs for a combination of durability, bandwidth, and data governance requirements. Locally redundant storage provides the highest maximum bandwidth, with the least durability, whereas geo-redundancy options provide higher durability with possible asynchronous replication delay.

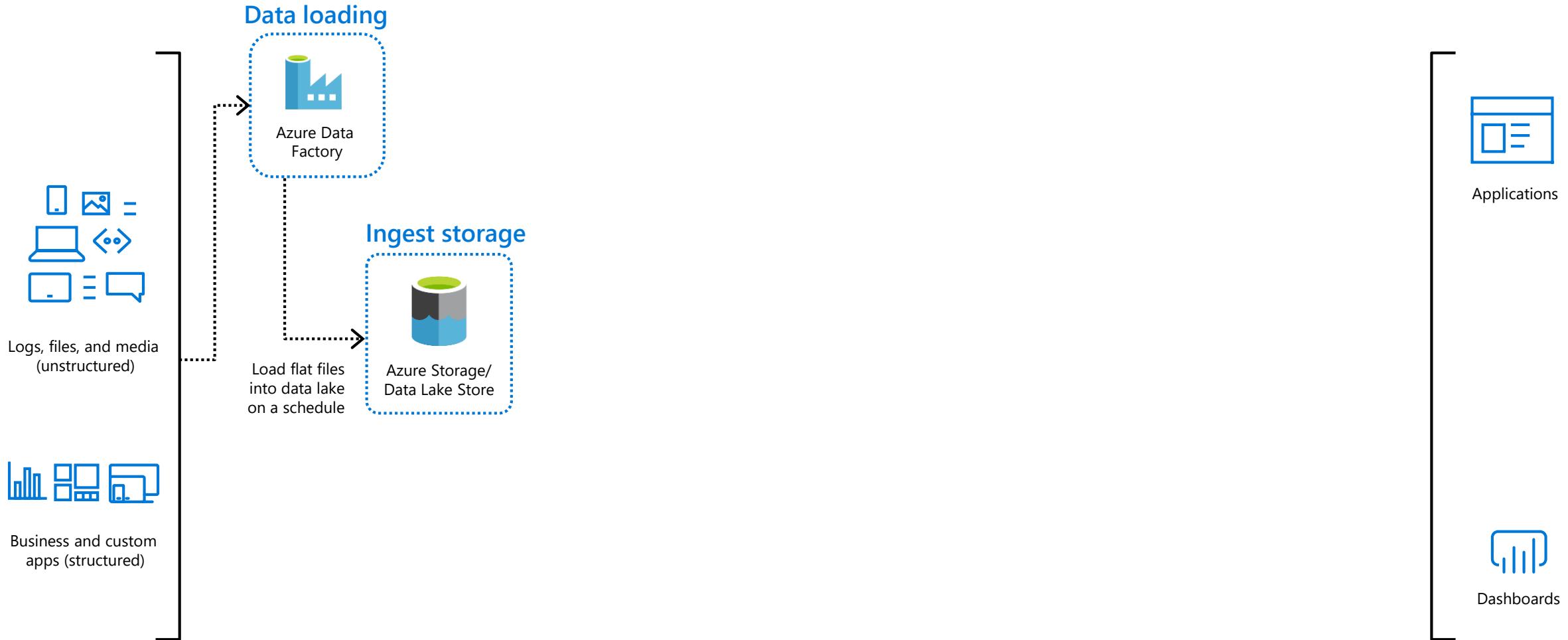
File storage

A side-by-side comparison of the capabilities and features

	Azure Data Lake Store	Azure Blob Storage containers
Purpose	Optimized storage for Big Data analytics workloads	General purpose object store for a wide variety of storage scenarios
Structure	Hierarchical file system	Object store with flat namespace
API	REST API over HTTPS	REST API over HTTP/HTTPS
Analytics workload performance	Optimized performance for parallel analytics workloads, high throughput and IOPS	Not optimized for analytics workloads
Size limits	No limits on account sizes, file sizes, or number of files	Max 500 TB per account and 4.75 TB per file
Geo-redundancy	Locally-redundant (multiple copies of data in one Azure region)	Locally redundant (LRS), globally redundant (GRS), and read-access globally redundant (RA-GRS). See Azure Storage replication for more information
Service state	Generally available	Generally available
Regional availability	Some regions	All regions

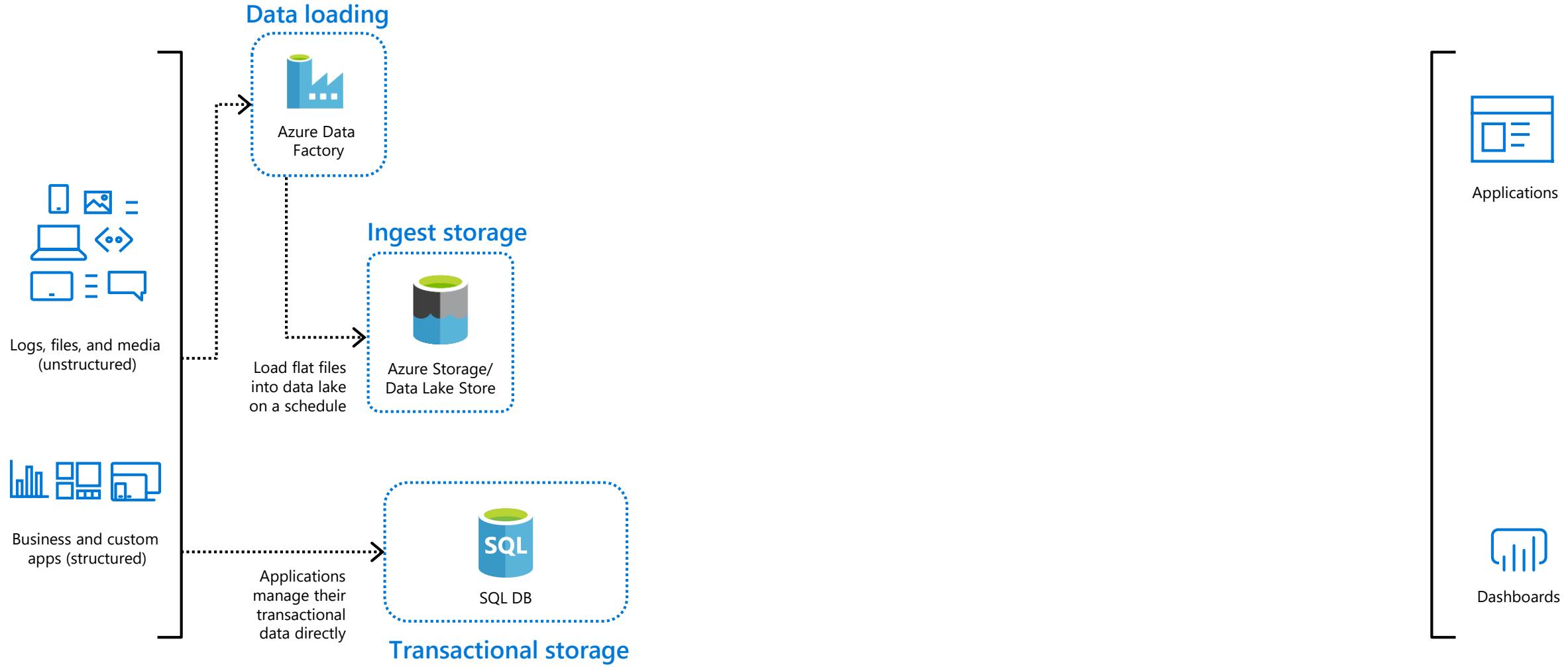
Data warehousing pattern in Azure

Loading data into ingest storage

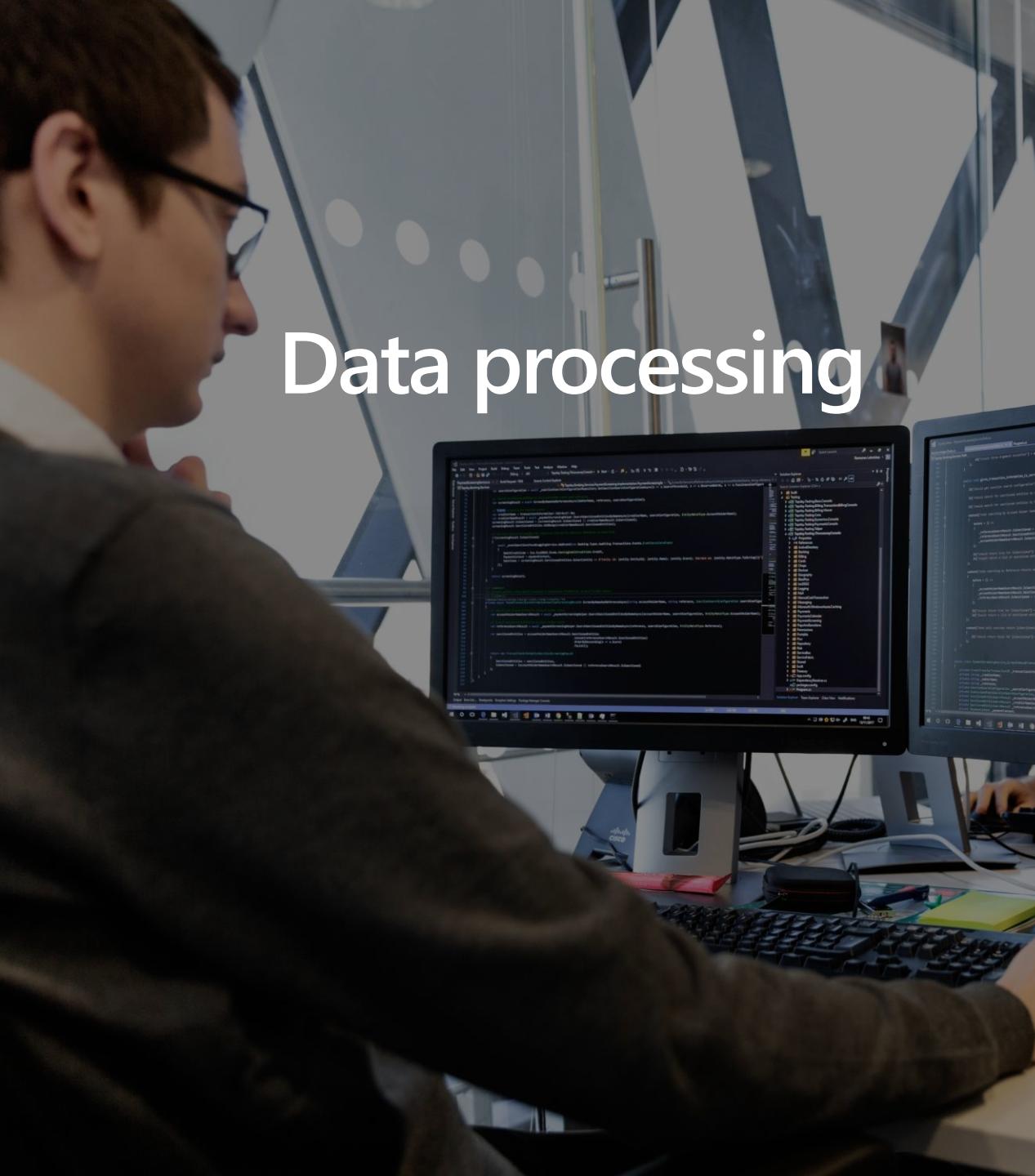


Data warehousing pattern in Azure

Load data into multiple source data stores



Data Processing

A photograph of a man with glasses, seen from the side and slightly from behind, looking intently at a computer monitor. The monitor displays two windows of code, possibly Python or Java, showing complex data structures and logic. The background is blurred, suggesting an office environment.

Data processing

Is data cleansing, structuring, curation, and aggregation

In data warehousing, the data is batch processed in preparation for loading into a data warehouse

Data processing requirements

Requirement	Comment
Scalability	The amount of data that needs to be processed at once can vary widely on any given day, or can grow over time. The batch data processing technology should scale to meet your needs with the level of granularity and within an acceptable time range for your solution.
Choice of language	The batch data processing technology should provide a choice of languages with which to create batch operations, including Python, Java, U-SQL, HiveQL, and R.
Integration choices	The batch data processing technology should allow you to choose a cloud-based data source from which to query, such as Azure Storage or Azure Data Lake Store. Some options allow additional integration options by also allowing you to query external relational data stores, such as SQL Data Warehouse.

Batch data processing

A side-by-side comparison of general capabilities and features

	Azure Data Lake Analytics	HDInsight with Spark	HDInsight with Hive	HDInsight with Hive LLAP	SQL Data Warehouse	Azure Databricks
Is a managed service	Yes	Yes	Yes	Yes	Yes	Yes
Auto-scaling	No	No	No	No	No	Yes
Supports pausing compute	No	No	No	No	Yes	Yes
Programmability	U-SQL	Python, Scala, Java, R, SQL	HiveQL	HiveQL	T-SQL	Python, Scala, Java, SQL, R
Programming paradigm	Mixture of declarative and imperative	Mixture of declarative and imperative	Declarative	Declarative	Declarative	Mixture of declarative and imperative
Pricing model	Per job (by job run per hour times analytics unit used)	By cluster hour	By cluster hour	By cluster hour	By cluster hour	By cluster hour

Batch data processing

A side-by-side comparison of integration capabilities

	Azure Data Lake Analytics	HDInsight with Spark	HDInsight with Hive	HDInsight with Hive LLAP	SQL Data Warehouse	Azure Databricks
Access Azure Data Lake Store	Yes	Yes	Yes	Yes	Yes	Yes
Query Azure Storage	Yes	Yes	Yes	Yes	Yes	Yes
Query external relational stores (like Azure SQL Database, SQL Server in virtual machine, or Azure SQL Data Warehouse)	Yes	Yes	Yes	No	Yes	Yes

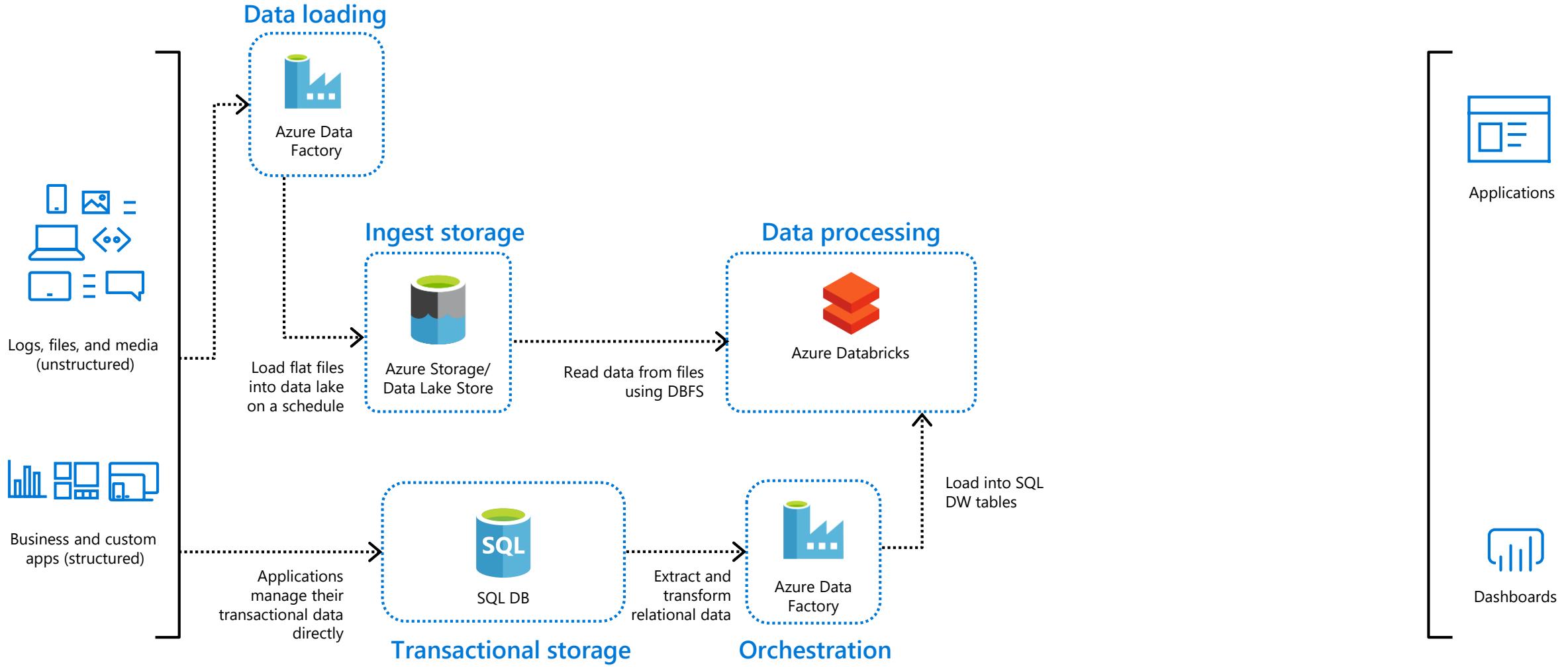
Batch data processing

A side-by-side comparison of scalability capabilities

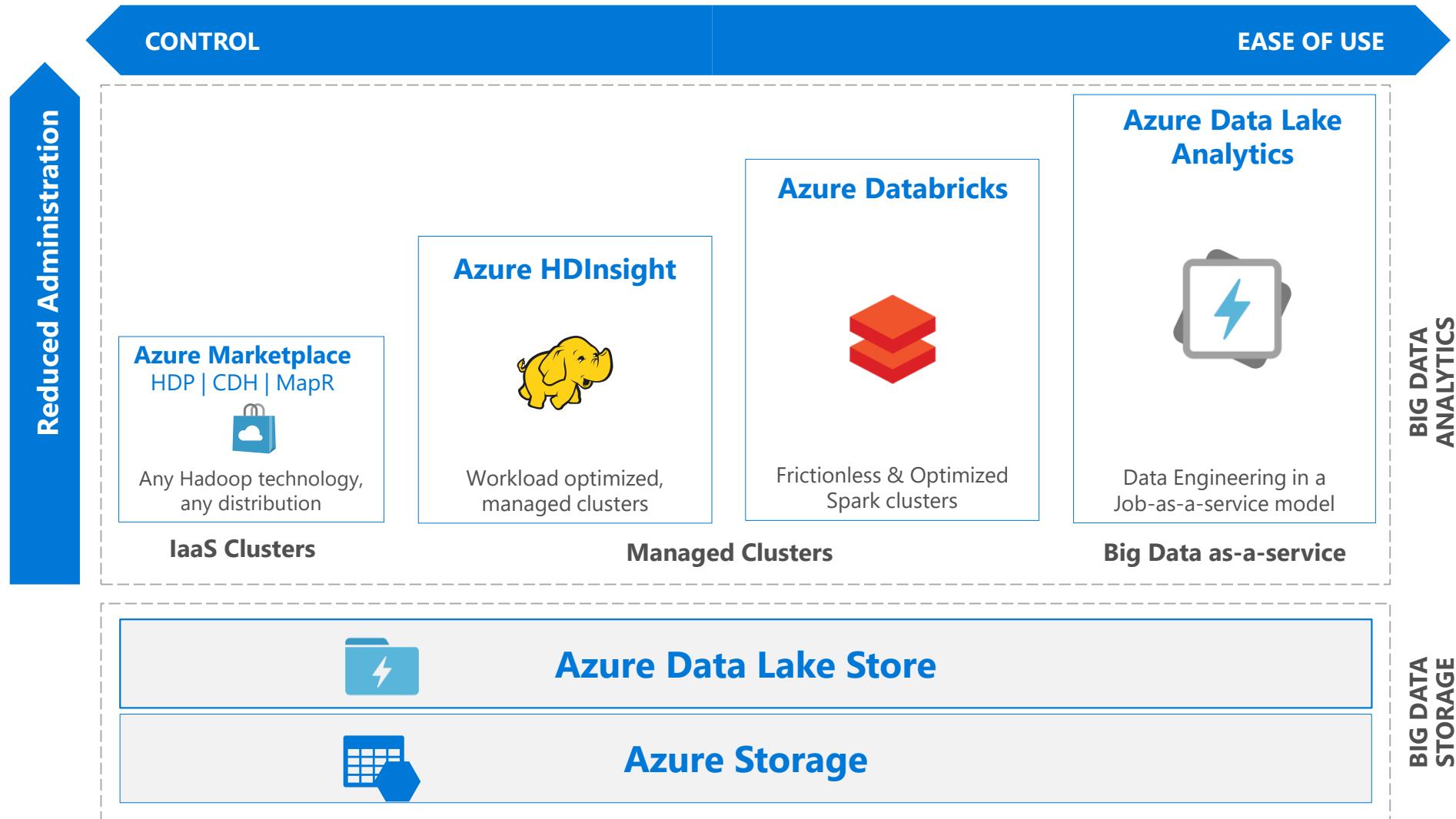
	Azure Data Lake Analytics	HDInsight with Spark	HDInsight with Hive	HDInsight with Hive LLAP	SQL Data Warehouse	Azure Databricks
Scale-out granularity	Per job	Per cluster	Per cluster	Per cluster	Scale out by compute units (DWU)	Per cluster
Supports fast scale out (less than 1 minute)	Yes	No	No	No	No	Yes
Supports in-memory caching of data	No	Yes	No	Yes	Yes	Yes

Data warehousing pattern in Azure

Data processing with Azure Databricks



Knowing the various big data solutions



Azure HDInsight or Databricks

Azure HDInsight

Azure Databricks

What It Is

- Hortonworks distribution as a first party service on Azure
- Big Data engines support – Hadoop Projects, Hive on Tez, Hive LLAP, Spark, HBase, Storm, Kafka, R Server
- Best-in-class developer tooling and monitoring capabilities
- **Enterprise Features**
 - VNET support (join existing VNets)
 - Ranger support (Kerberos based Security)
 - Log Analytics via OMS
 - Available in most Azure Regions (27) including Gov Cloud and Federal Clouds

Guidance

- Customer needs Hadoop technologies other than, or in addition to Spark
- Customer prefers Hortonworks Spark distribution to stay closer to OSS codebase and/or ‘Lift and Shift’ from on-premises deployments
- Customer has specific project requirements that are only available on HDInsight

What It Is

- Databricks’ Spark service as a first party service on Azure
- Single engine for Batch, Streaming, ML and Graph
- Best-in-class notebooks experience for optimal productivity and collaboration
- **Enterprise Features**
 - Native Integration with Azure for Security via AAD (OAuth)
 - Optimized engine for better performance and scalability
 - RBAC for Notebooks and APIs
 - Auto-scaling and cluster termination capabilities
 - Native integration with SQL DW and other Azure services
 - Serverless pools for easier management of resources

Guidance

- Customer needs the best option for Spark on Azure
- Customer teams are comfortable with notebooks and Spark
- Customers need Auto-scaling and Auto-termination
- Customer needs to build integrated and performant data pipelines

Data Modeling and Serving

Data serving



Processed data served by a data warehouse to analytic clients and reporting tools

The data warehouse provides increased query flexibility and reduced query latency in comparison to batch data processing options

Data serving

A side-by-side comparison of general capabilities and features

	SQL Database	SQL Data Warehouse	Azure Analysis Services
Is a managed service	Yes (Azure SQL Database)	Yes	Yes
Primary database model	Relational (columnar format when using columnstore indexes)	Relational tables with columnar storage	Tabular and MOLAP semantic models
SQL language support	Yes	Yes	No
Optimized for speed serving layer	Yes, using memory-optimized tables and hash or nonclustered indexes	No	No

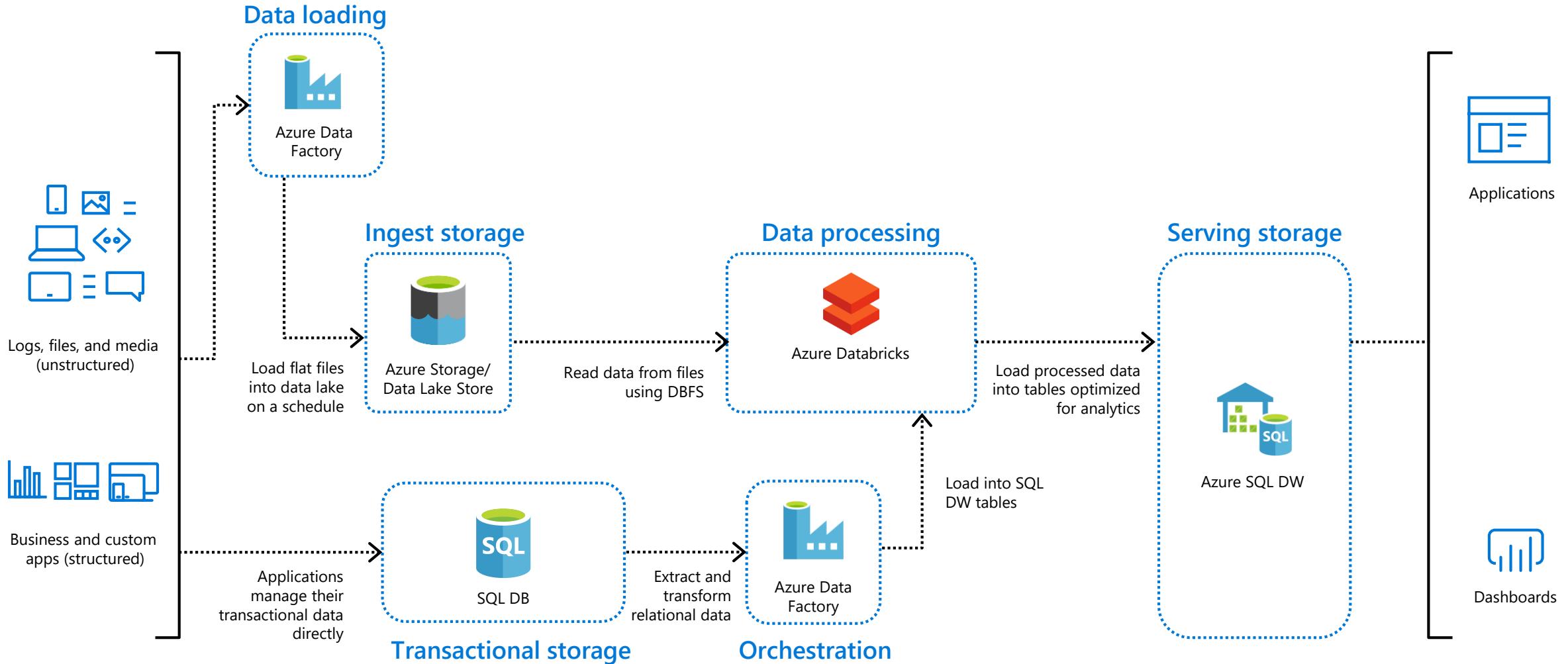
Data serving

A side-by-side comparison of scalability capabilities

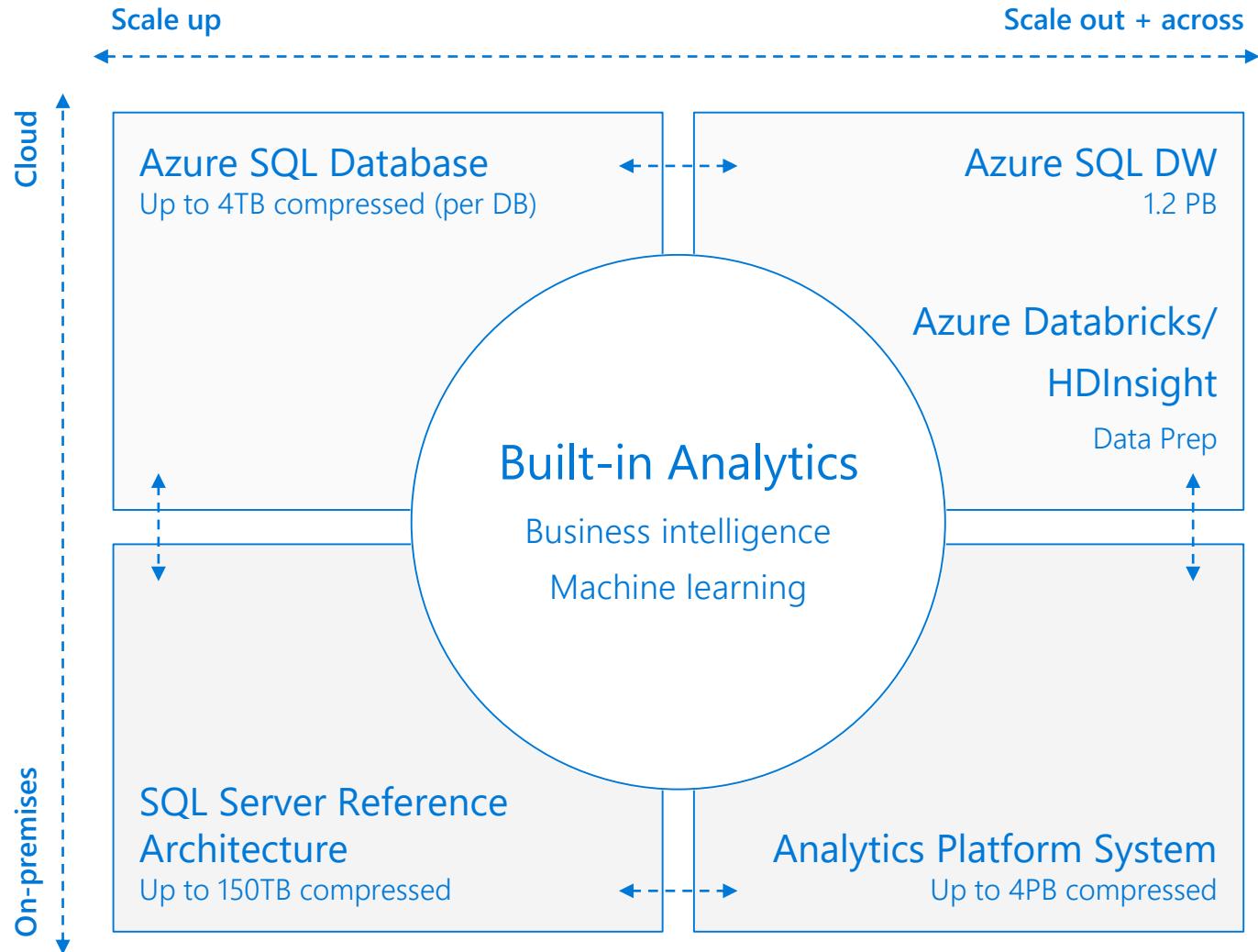
	SQL Database	SQL Data Warehouse	Azure Analysis Services
Redundant regional servers for high availability	Yes (Azure SQL Database)	Yes	No
Supports query scale out	No	Yes	Yes
Dynamic scalability (scale up)	Yes (Azure SQL Database)	Yes	Yes
Supports in-memory caching of data	Yes	Yes	Yes

Data warehousing pattern in Azure

Data serving with Azure SQL DW



What to use when



SSAS/Azure Analysis Services Models

Reasons to report off cubes instead of the data warehouse:

- Semantic layer
- Handle many concurrent users
- Aggregating data for performance
- Multidimensional analysis
- No joins or relationships
- Hierarchies, KPI's
- Security
- Advanced time-calculations
- Slowly Changing Dimensions (SCD)
- Required for some reporting tools

Hybrid Architecture

A dark, atmospheric photograph of a server room. On the left, several tall, black server racks are lined up, their front panels showing various ports and status lights. Above the racks, a large, bright white cloud graphic rises from behind them, symbolizing the integration of on-premises hardware with cloud computing.

Hybrid architectures

Enable the data storage, processing, and serving to span on-premises and cloud environments

Choosing Azure data factory

When Azure Data Factory can be a good option for your hybrid data pipelines

When you want...	Description
To orchestrate your data pipeline on-premises and in the cloud	Use Data Management Gateway and Azure Data Factory to move data between both cloud and self-hosted environments. Data is compressed and transferred in parallel and resilient to intermittent network issues through auto retry logic. You can connect on-premises data to cloud services to benefit from cloud services while keeping the business running with on-premises data.
To execute your SQL Server Integration Services (SSIS) packages in the cloud	When you provision an Azure and SSIS integration runtime (IR) in Azure Data Factory, you can deploy your SSIS packages to the runtime in Azure. Azure Data Factory orchestrates the SSIS package execution, which creates new opportunities for shifting existing on-premises data workflows to Azure.
To move your non-relational data to Azure for processing and transformation	Create and schedule data-driven workflows—pipelines—in Azure Data Factory that move your non-relational and unstructured data to Azure, then process and transform the data using compute services such as Azure HDInsight Hadoop, Spark, Azure Data Lake Analytics, and Azure Machine Learning.

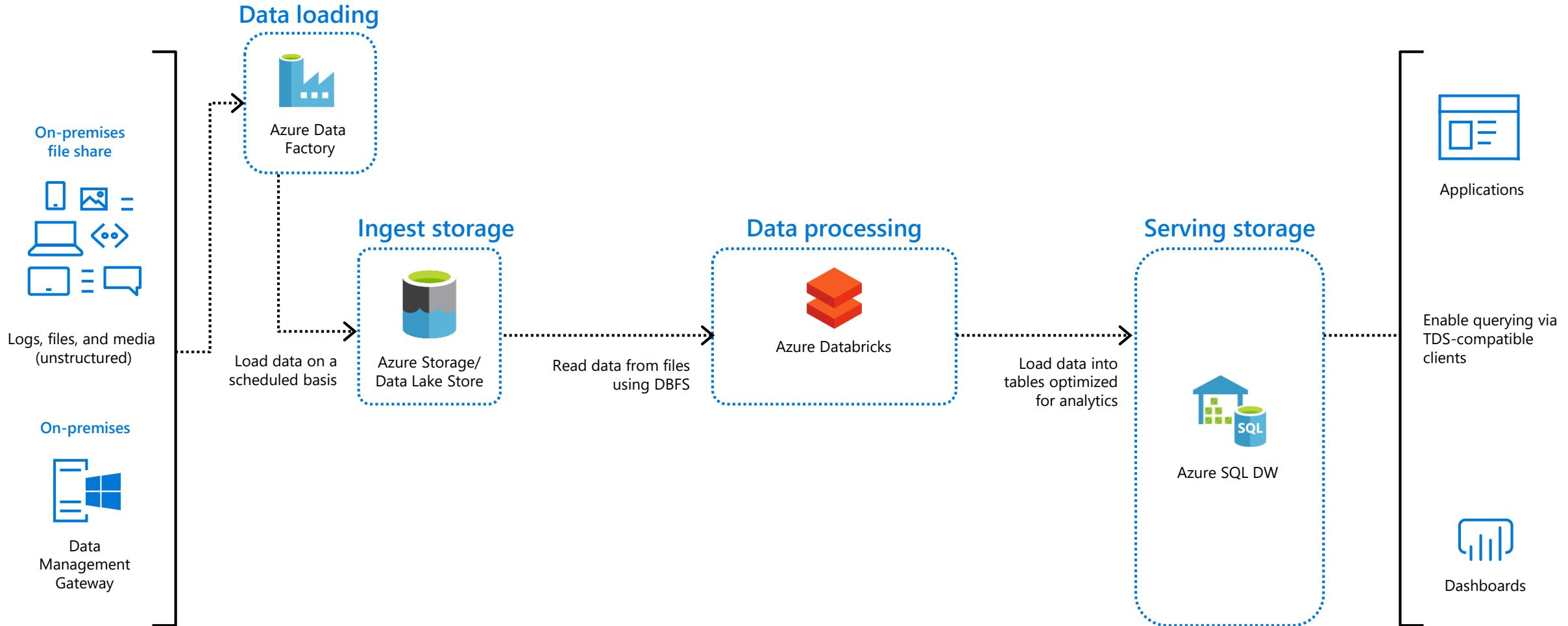
Choosing SQL server stretch database

When SQL Server Stretch Database can be a good option for your hybrid architecture

When you want...	Description
Cost-effective availability for cold data	Stretch warm and cold transactional data dynamically from your on-premises SQL Server to Microsoft Azure with Stretch Database. Unlike typical cold data storage, your data is always online and available to query. You can provide longer data retention timelines without breaking the bank for large tables like customer order history. Benefit from the low cost of Azure rather than scaling expensive, on-premises storage. You choose the pricing tier and configure settings in the Azure Portal to maintain control over price and costs.
Access to your SQL data, regardless of location, without changes to your queries or applications	Access your SQL Server data seamlessly regardless of whether it's on-premises or stretched to the cloud. You set the policy that determines where data is stored, and SQL Server handles the data movement in the background. The entire table is always online and query-able. Stretch Database doesn't require any changes to existing queries or applications—the location of the data is completely transparent to the application.
Streamlined on-premises data maintenance	Reduce on-premises maintenance and storage for your data. Backups for your on-premises data run faster and finish within the maintenance window. Backups for the cloud portion of your data run automatically. Your on-premises storage needs are greatly reduced. Azure storage can be 80 percent less expensive than adding to on-premises SSD.

Data warehousing pattern in Azure

Loading data from on-premises sources



Security

A photograph of a server room. On the left, there are several tall, dark server racks. In the center, a person wearing a black shirt and blue jeans is standing in front of a row of server racks, looking down at something in their hands. The ceiling is white with various pipes and cables running across it.

Security

Enables the data warehouse to control access in order to protect sensitive data and maintain desired compliance

Data storage security

A side-by-side comparison of the capabilities and features

	Azure Data Lake Store	Azure Blob Storage containers
API	REST API over HTTPS	REST API over HTTP/HTTPS
Data operations: Authentication	Based on Azure Active Directory Identities	Based on shared secrets account access keys and shared access signature keys , and role-based access control (RBAC)
Data operations: Authorization	POSIX access control lists (ACLs). ACLs based on Azure Active Directory identities can be set at file and folder level	For account-level authorization use account access keys . For account, container, or blob authorization use shared access signature keys
Encryption data at rest	Transparent, server side With service-managed keys With customer-managed keys in Azure Key Vault	Transparent, server side With service-managed keys With customer-managed keys in Azure Key Vault (coming soon) Client-side encryption
Management operations (for example, account create)	Role-based access control (RBAC) provided by Azure for account management	Role-based access control (RBAC) provided by Azure for account management

Batch data processing security

A side-by-side comparison of the capabilities and features

	Azure Data Lake Analytics	HDInsight with Spark	Apache Hive on HDInsight	Hive LLAP on HDInsight	Azure Databricks
Authentication	Azure Active Directory	No	Local/Azure Active Directory *	Local/Azure Active Directory *	Azure Active Directory (native/built-in)
Authorization	Yes	No	Yes *	Yes *	Yes
Auditing	Yes	No	Yes *	Yes *	Yes
Data encryption at rest	Yes	Yes	Yes	Yes	Yes
Row-level security	No	No	Yes *	Yes *	No
Supports firewalls	Yes	Yes	Yes ***	Yes ***	Coming soon
Dynamic data masking	No	No	Yes *	Yes *	No

* Requires using a [domain-joined HDInsight cluster](#)

** Requires using Transparent Data Encryption (TDE) to encrypt and decrypt your data at rest

*** Supported when [used within an Azure virtual network](#)

Data serving security

A side-by-side comparison of the capabilities and features

	SQL Database	SQL Data Warehouse	Azure Analysis Services	Azure Cosmos DB
Authentication	SQL/Azure Active Directory	SQL/Azure Active Directory	Azure Active Directory	Database users and Azure Active Directory via access control (IAM)
Authorization	Yes	Yes	Yes	Yes (hash-based message authentication code (HMAC))
Auditing	Yes	Yes	Yes (when integrated with Azure Monitor resource diagnostic logs)	Yes (through audit logging and activity logs)
Data encryption at rest	Yes **	Yes **	Yes	Yes
Row-level security	Yes	No	Yes (through object-level security in model)	No
Supports firewalls	Yes	Yes	Yes	Yes
Dynamic data masking	Yes	No	No	No

** Requires using transparent data encryption (TDE) to encrypt and decrypt your data at rest

*** Supported when [used within an Azure virtual network](#)

Automation

A photograph of three people in a meeting room. A man in a grey sweater is standing and writing on a whiteboard with a blue marker. A woman with short curly hair, wearing a light-colored cardigan, sits at a round wooden table with two laptops. She is looking towards the man. Another man with a beard, wearing a dark t-shirt, sits across from her, gesturing with his hands as if speaking. The room has bookshelves in the background.

Automation

Enables all components of the data warehouse solution to be controlled, deployed, and monitored programmatically

Choosing Azure automation

When Azure Automation can be a good option for cloud-based automation

When you want...	Description
Process automation	Automate frequent, time-consuming, and error-prone cloud management tasks by authoring runbooks in a graphical UI, in PowerShell, or in Python.
Configuration management	Manage your desired state configuration (DSC) resources and apply configurations to virtual or physical machines in Azure. Monitor and automatically update machine configuration across physical and virtual machines, Windows or Linux, in the cloud or on-premises. Collect inventory about in-guest resources and track changes across services, daemons, software, registry, and files.
Update management	Update Windows and Linux systems across hybrid environments. Gain visibility of update compliance across Azure, on-premises and in other clouds. Schedule deployments to orchestrate installation of updates within a defined maintenance window.
Build and deploy resources	Deploy Azure resources using Runbooks and Azure Resource Manager (ARM) templates. Integrate into development tools like Jenkins and Visual Studio Team Services, ensuring continuous delivery and operations automation.

Choosing Azure resource manager templates

When Azure Resource Manager (ARM) templates can be a good option for cloud-based automation

When you want...	Description
To consistently and repeatedly deploy resources	ARM templates are composed of a JavaScript Object Notation (JSON) file that defines one or more resources, including any dependencies between them. This adds the benefit of treating your resources for a solution as a single unit, rather than independent components, making it easier to consistently deploy and manage the resources to development, test, staging, and production environments.
To manage your infrastructure through declarative templates rather than scripts	Declarative templates make it easier to define your resource parameters, dependencies, and infrastructure, compared to executing a series of scripts. Furthermore, you can apply tags to resources to logically organize all the resources in your subscription.
To include your infrastructure definition as part of your app source code	The template can become part of the source code for your app. Check it in to your source code repository and update it as your app evolves. Simply edit the template through Visual Studio or your favorite IDE.
To ensure your resources are deployed in the correct order	Resource dependencies are declaratively expressed within the ARM template. This ensures that components add their dependencies when they are provisioned, and that the resources within the template are created in the proper order. For instance, an ARM template that creates a VM and a VNet to which it is added, first creates the VNet, then creates the VM and associates the two.

Monitoring

A photograph showing three people in a meeting room. A man in a brown jacket is leaning forward, looking intently at a laptop screen. A woman with curly hair, wearing a red blouse, is seated next to him, also looking at the screen. Another person's arm is visible on the left, pointing towards the laptop. The background shows a window with blinds.

Monitoring

Provides insights into the status and health of the data warehouse solution

Choosing Azure for monitoring

When Azure Monitor can be a good option for your monitoring solution

When you want...	Description
To access base-level metrics and logs	Azure Monitor provides base-level infrastructure metrics and logs for most services in Microsoft Azure. Azure services that do not yet put their data into Azure Monitor will put it there in the future.
To discover, configure, and on-board Azure Monitor features	Provides a landing page that helps you understand the monitoring capabilities offered by Azure. This starting point for on-boarding platform and premium monitoring capabilities shows curated notable issues from different services, allowing you to navigate to them in context.
To view important monitoring events across a given subscription	<p>In Azure Monitor, select a subscription and view the following across the components of the subscription:</p> <p>Triggered alerts and alert sources Activity log errors Azure Service Health data and alerts Application Insights KPIs (key performance indicators)</p> <p>If Log Analytics, Azure Alerts, or Application Insights haven't been configured, the page provides links to begin your on-boarding process.</p>

Choosing Azure application insights

When Azure Application Insights can be a good option for your monitoring solution

When you want...	Description
To monitor your live web application	Azure Application Insights provides a rich monitoring solution optimized for collecting and visualizing performance metrics from both the server and client, identifying exceptions in the application, and monitoring application usage.
To track exceptions down to actual failing code	When you receive an alert or discover a problem, you can assess how many users are affected. Correlate failures with exceptions, dependency calls, and traces. To dig deeper, examine profiler, snapshots, stack dumps, and trace logs.
To write custom telemetry in your code	Use the Azure Application Insights core telemetry API to send custom events and metrics, and your own versions of standard telemetry. Create custom events that are relevant to your application, and that can provide more custom monitoring options than standard out-of-the-box telemetry. For example, your e-commerce site can send events like <i>item added to cart</i> and <i>coupon applied</i> to Azure Application Insights, where you can use the built-in visualization tools to aggregate and compare these events over a given timeframe.
To monitor web site availability and responsiveness	Create availability tests for any HTTP or HTTPS endpoint that is accessible from the public internet. Web requests will be sent to your application at regular intervals from points around the world and alert you if your application doesn't respond, or responds slowly. Use Visual Studio to record a multi-step web test scenario for more advanced test automation.

Choosing Azure Log Analytics

When Azure Log Analytics can be a good option for your monitoring solution

When you want...	Description
To be able to collect data generated by your cloud resources, on-premises environments, and other monitoring tools	Azure Log Analytics can collect data from multiple sources, including information sent directly from agents running on VMs, and other monitoring tools such as Azure Monitor, System Center Operations Manager, and Azure Application Insights. This allows you to correlate this data and have a single pane of glass through which to query and view logs pertinent to your cloud and on-premises environments.
Service and application-specific monitoring	Solutions are available for a variety of functions and additional solutions are constantly being added. You can easily browse available solutions and add them to your workspace from the Azure Marketplace. Many will be automatically deployed and start working immediately while others will require moderate configuration. Some examples include Logic Apps Management, Azure Search, HDInsight, SQL Health Check, and Azure Active Directory.
To create alert rules on log data	Alerts can be created through alert rules that automatically run log searches at regular intervals. If results of the log search match particular criteria, an alert record is created and can be configured to perform an automated response.

All Up

Big Data & Advanced Analytics

