REFERENCE ARCHITECTURE FOR ALTERYX ON AZURE

White Paper

Radiant Advisors

Table of Contents

Introduction

Today's data-driven organizations are dependent upon enterprise analytics to deliver valuable insights in areas where the information will be most useful – and to do so with speed and agility. The key to competitive business and growth is to empower business users, analysts and data scientists to work with data and create analytics in a self-service manner for their dynamic, everyday business needs. The Modern Analytics Lifecycle aligns with the natural process for how people work independently with data and incorporates modern data technologies to accelerate their iterative and collaborative approach to analytics that will keep up with the changing business environment.

Radiant Advisors has studied the processes and workflows within the Modern Analytics Lifecycle and the related architecture patterns that commonly support successful analytic programs. With insights and observations from our broad network of Fortune 500 companies, we understand the people, processes and technology necessary to sustain modern analytics. This technical white paper, with assistance from research sponsor Alteryx and its partner Microsoft Azure, presents the process, platform components and architecture patterns for an analytics environment.

Specifically, this paper provides a straightforward modern reference architecture for IT architects and senior personnel interested in deploying Alteryx, the self-service data analytics platform, on Azure, along with Azure's widely adopted database, data pipeline and data science services.

The Modern Analytics Lifecycle

Many business analysts and data scientists follow an involved process of discovery, exploration, hypothesis building and vetting when building analytics to tackle a business question or challenge. When seeking to understand patterns in data and reveal informed insights to support their work, this curiosity-driven, analytical process is full of hypotheses, what-if questions and the need to validate, evolve or dismiss ideas dynamically as the secrets of the data are unlocked. Radiant Advisors calls this the Modern Analytics Lifecycle. Analytic workflow development must be fluid and benefit from self-service capabilities due to the highly iterative business and technical context and discovery-driven needs of the people working with the data. The Modern Analytics Lifecycle represents how people intuitively approach solving problems with data, and the components of the lifecycle are designed to optimize each step and deliver the agility and speed required in today's organizations.

Figure 1:
The Modern Analytics Lifecycle

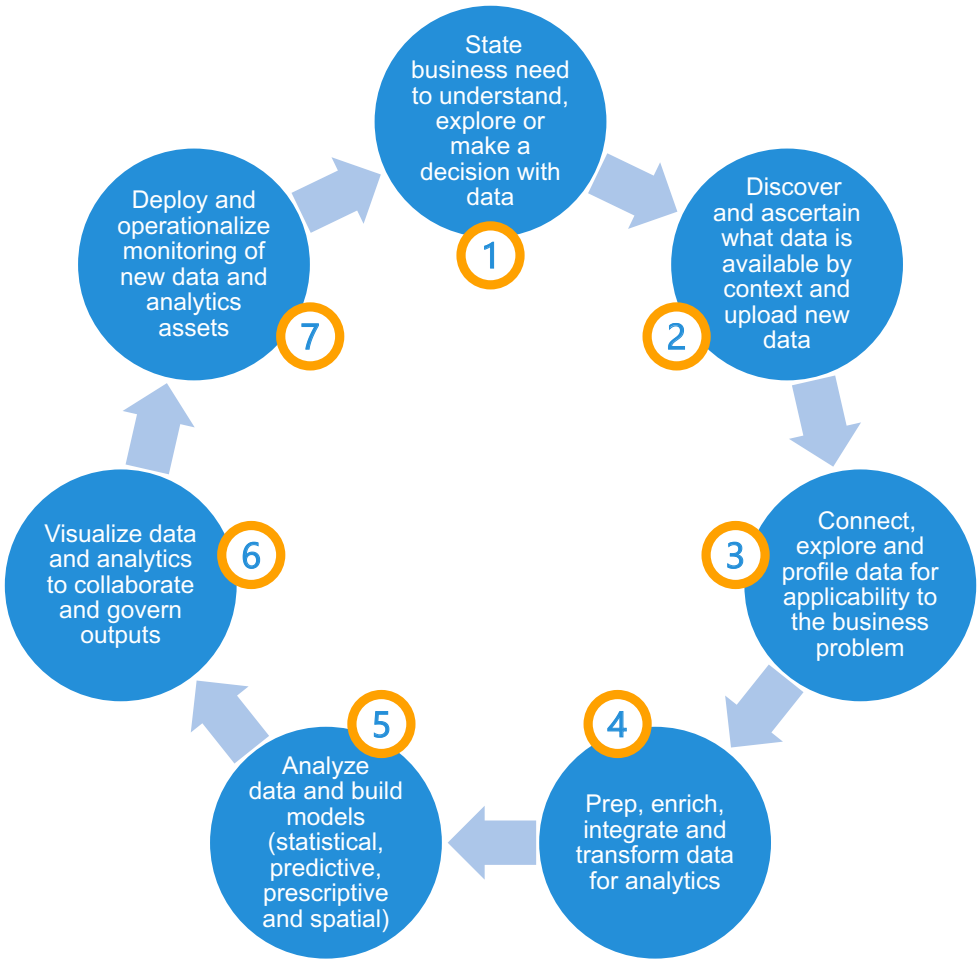


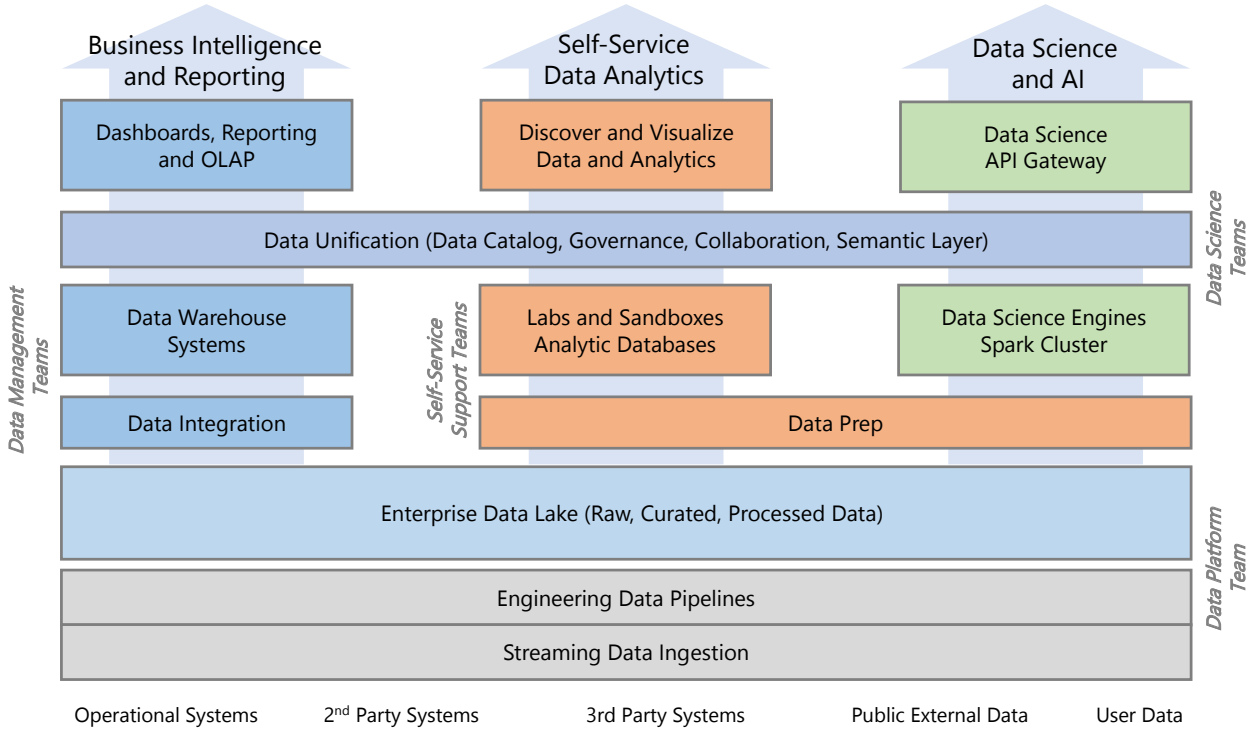
Figure 1 outlines seven data-driven steps of the Modern Analytics Lifecycle that business analysts and data scientists work through when solving challenges and advancing business objectives.

The beginning steps of the lifecycle apply broadly to power users, business analysts and data scientists, while the latter steps specifically benefit data scientists and statisticians who depend on derived data sets for their analytic model development. When data identification and data prep is optimized with predominantly no-code intuitive visual development tools, data science teams and business analysts can build broad data sets for statistical analysis that drive predictive, prescriptive and spatial analytics to iteratively identify correlations and patterns in the data. The last two steps focus on visual communication and collaboration as well as easily deploying the analytics for optimum business impact. The time-to-analytic-value is further accelerated when previously curated and proven data sets can be reused.

The Modern Data Platform Driving Enterprise Analytics

The Modern Data Platform is a data and technology framework that supports the Modern Analytic Lifecycle. Designed by Radiant Advisors to deliver a single optimized architecture for enterprise information and analytics, the Modern Data Platform is an applied ecosystem of vendor technologies that enables organizations to efficiently access and produce enterprise data assets with the Modern Analytics Lifecycle for increased agility, efficiency from reusability and time-to-insight from data.

Figure 2:
Common Platform to Deliver Three Forms of Enterprise Analytic



The Modern Data Platform is optimized to meet the requirements of the three domains of modern enterprise analytics: business intelligence and reporting, self-service data analytics, and data science and artificial intelligence. (See Figure 2.) To do so, the functional layers and components in the framework support data movement, data persistence and collaborative data access. As a conceptual framework, the subsequent logical and physical architectures can be applied on-premises in data centers or with public cloud service providers, such as Azure, incorporating various vendors and products.

Conceptual Architecture and Data Flow

The **data integration architecture** facilitates data acquisition (streaming or batch-oriented) and engineering data pipelines to make data available for self-service data analytics. Data ingestion and data processing are distinctly separated in the data integration architecture. (See Figure 3.) The centralized data streaming hub simplifies all data source ingestion and scales for many independent and asynchronistic data consumers (data stores or pipelines). The data pipeline platform subscribes to data streams and performs developer functions such as cleansing, integration and enrichment prior to delivering the data to data stores and analytic applications for access.

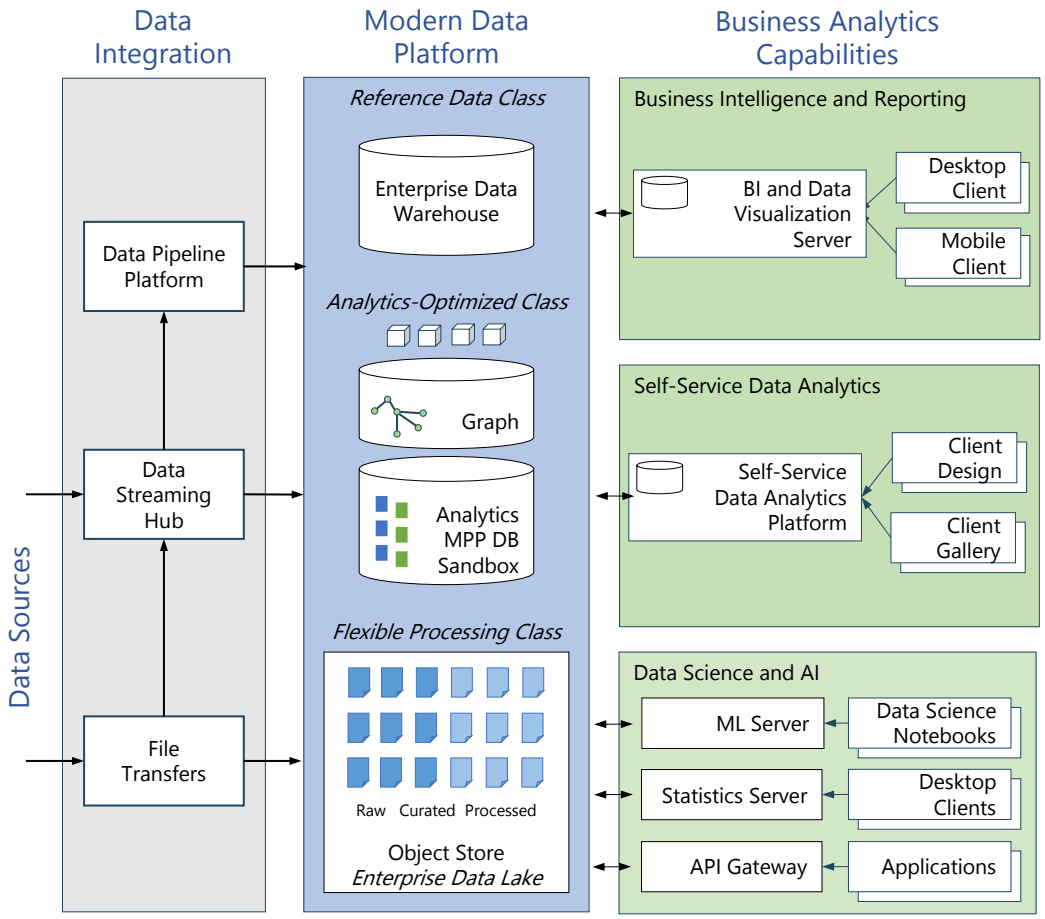
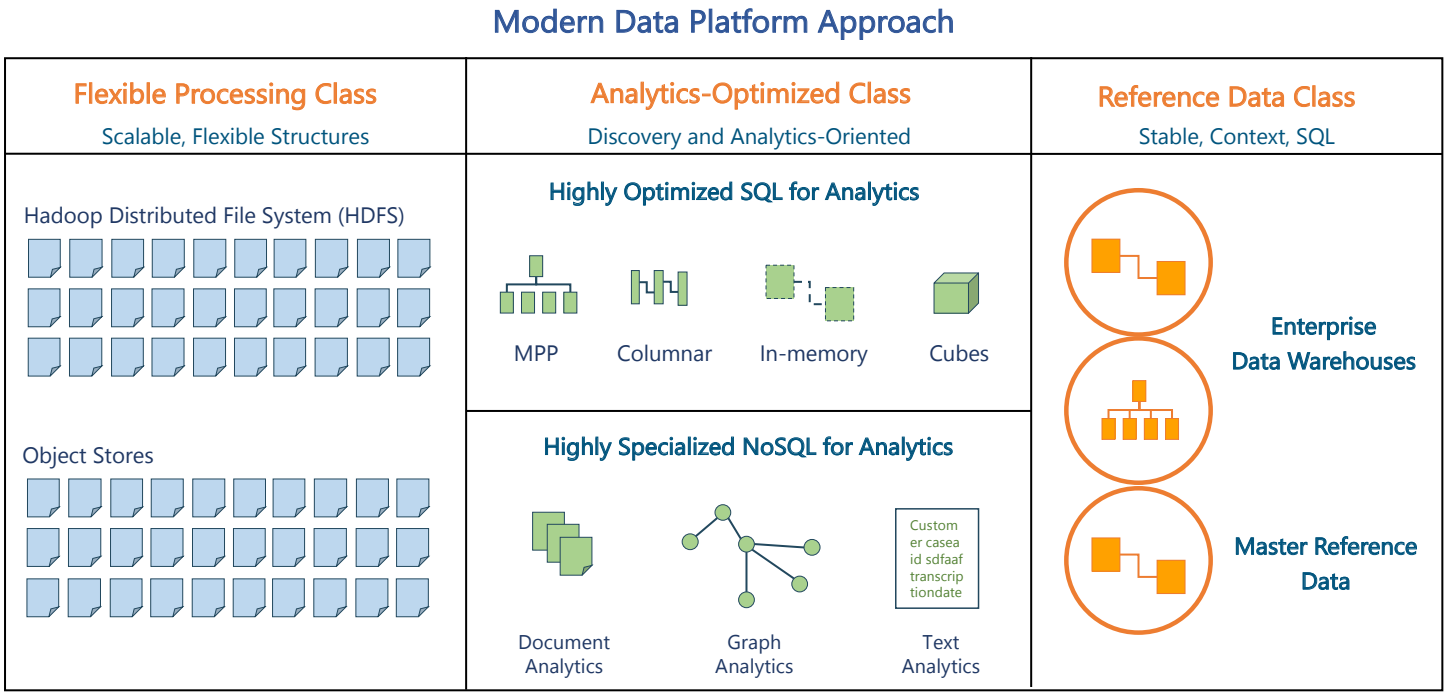


Figure 3:
Conceptual Architecture for Modern Data Platform

Figure 4:
Modern Data Platform for Analytics Based on Polyglot Persistence Principle

The **data architecture** and technology design principle of the Modern Data Platform is to persist data in the data store that is most optimal to execute its primary workload. (This concept is also known as polyglot persistence or best-of-breed design.) The Modern Data Platform framework adopts this principle for data management and analytic data use cases. Three data workload classifications relate to available data technologies and demonstrate the distinct strengths over each other in a complementary way. Each of the three workload classifications and corresponding analytic data technologies is shown in Figure 4.



Polyglot Persistence Methodology

Reference Data Class databases are used for logically modeled data that predefines the database tables in need of support from row-based record processing and high user concurrency performance. These are typically existing enterprise data warehouses (EDWs) and master data management (MDM) hubs that rely on relational databases deployed on a single server, cluster or massively parallel processing (MPP) architecture. Reference or subject area data is relatively smaller than other enterprise transactional or event data sets and contains slowly changing historical data attributes that provide the analytic insights for the event-based, high-volume data in which the subjects participate.

- Database strengths: Row-based operations, mature SQL and high concurrency
- Database weaknesses: Cost-prohibitive to scale

The Analytics-Optimized Class of databases can be SQL or NoSQL databases that have been optimized specifically for analytics. The SQL databases include newer-generation columnar databases, MPP databases, in-memory databases as well as traditional OLAP cubes. NoSQL data stores include document databases and graph databases to support other forms of analytical questions. Event-based or big data sets are typically an order of magnitude larger in volume over the reference data class data and capably operate with hundreds of terabytes data. The analytical SQL databases also provide business analysts and data scientists with dedicated analytic sandbox environments to explore and work with data on their own.

- Database strengths: High-performance analytical queries on big data
- Database weaknesses: High concurrency and row-based operations

The Flexible Processing Class of data stores are highly scalable distributed file systems, such as object stores or a Hadoop Distributed File System (HDFS). They're proven to leverage built-in redundancy to leverage low-cost storage devices to affordably scale to many petabytes. Additionally, the flexible processing class data stores persist data without predefining a data structure. (Structure can be added later as metadata for structured SQL access.) This is where enterprise data lakes are typically built, and all data ingestion should be persisted, organized and governed for use independently by the business.

- Data store strengths: Highest affordable scalability and flexible unstructured data support
- Data store weaknesses: Lower query performance, programmed access or emerging SQL

Self-service data analytics platforms serve the role of being a centralized platform that enables business analysts and data scientists to work and follow the Modern Analytics Lifecycle. Data workers are empowered to identify data from any accessible data stores, explore and transform the data in their own analytic sandboxes, then deploy the data and analytics workflow back into an enterprise data lake, analytic sandbox, or deploy analytic models via REST APIs for authorized applications. There is also the possibility that derived data sets can be made reusable for other analytics and data science projects in the future.

Azure Reference Architecture for Alteryx

Microsoft Azure is a market leader in global cloud services for enterprises intending to modernize their data and analytics environment to capitalize on cloud benefits, such as agility, robust elasticity, lower total cost of ownership and reliable scalability. The Azure global data centers provide a combination of cloud services for infrastructure (IaaS), platform (PaaS) and software-as-a-service (SaaS) for meeting enterprise analytics needs. Available in over 40 regions to represent its Marketplace in 140 countries, more than 100 Azure products enable an ecosystem of services, technologies and infrastructure to develop and operate all types of enterprise applications in both hybrid and pure cloud architectures.

Most Modern Data Platform components are preferable on Azure as PaaS or SaaS services for data ingestion, data processing, database services, data science and data visualization. The Modern Data Platform can be architected with Azure cloud services, but Azure may not offer all of the components required. For instance, robust self-service data analytics is not available in the Azure ecosystem. In instances such as this when Azure does not have a specific PaaS or SaaS component, it does have IaaS so that you can create your own virtual network, provision a virtual machine and install a self-service data analytics platform, such as Alteryx Server. The installed platform will then be able to capitalize on cloud-based benefits, such as on-demand resources, easy administration or auto-scaling of its virtual machines for increased workloads with load balancers and built-in fault tolerance. Once Alteryx is running on Azure, it connects and works with all the other Azure services within the Modern Data Platform. Alteryx can also execute algorithms in a cloud-based Apache Spark cluster provided by Azure Spark for HDInsight (PaaS). Figure 5 (on page 9) shows the logical architecture for how Azure services and Alteryx can fulfill the Modern Data Platform framework for enabling the Modern Analytics Lifecycle.

As an enterprise-class cloud provider, Azure ensures service level agreements (SLAs) are met with the proper industry-standard compliance and security requirements. Azure is built upon the basic cloud computing concept of Regions, Availability Zones and Geographies to provide fault-tolerance and data protection throughout its global platform.

- An Azure region is a set of data centers with a dedicated low-latency network. Companies will need to validate that the Azure products they need are available for a given region.
- Every region has two or more Availability Zones – physically separate data centers with their own independent power, environmental controls (such as cooling) and networking – to support high availability and low-latency replication of applications.
- Azure geographies are defined as Azure markets that preserve data residency and compliance boundaries with two or more regions. A geography can support a complete region failure with its dedicated high-capacity networking infrastructure. *(Reference: https://azure.microsoft.com/en-us/global-infrastructure/regions/)*

Azure Components for Data Ingestion and Data Processing

Azure connectivity from on-premises corporate data centers is best accomplished by Azure ExpressRoute. This private connection between on-premises data centers and Azure data centers ensures reliable high-speed, high-bandwidth, low-latency network connectivity for both on-premises system data transfers as well as administrators and end-user access. Most enterprise IT and InfoSec departments opt to set up private addressing and routing with an Azure ExpressRoute provider to establish dedicated connectivity between corporate and Azure data centers and regions. This establishes a private link connection for higher bandwidths, lower latency and less complexity than over the public internet with Internet Protocol Security (IPSec) virtual private network (VPN) access. This also avoids the public Internet for consistent and isolated data transfer performance. Company security policies may require remote users to connect to the corporate data center with their established corporate VPN and then utilize the ExpressRoute so that only the company data center communicates with Azure. Azure does offer an IPSec VPN connection service that allows users to connect securely over the public Internet to any IP addresses within their Azure Virtual Network.

Data ingestion and data processing are two distinct functions of modern data integration that are represented in Azure architecture diagrams as separate columns or areas of Azure services that the data flow passes through. Data ingestion includes incoming data with a centralized data streaming hub that consolidates all incoming data from batch-oriented and streaming data sources as well as data acquisition techniques, such as database SQL queries, data streams, data file transfers or application APIs. Data processing focuses on a platform for engineering data pipelines that consume data from the data streaming hub to further prepare, integrate and transform the data for use in Azure databases, data stores or analytic applications.

Data ingestion focuses on providing reliability and scalability needed to ensure that all incoming data can serve all independent data consumers to receive and process the data they need. The Azure Event Hub PaaS is the first choice for the centralized data streaming hub for its fully managed, multi-tenant data streaming platform. Azure Event Hub is based on the Advanced Messaging Queueing Protocol (AMQP) for ingesting data streams with serverless workflows in real-time or batch-oriented pipelines. If needed, the Azure IoT Hub is further specialized for ingesting IoT device data (including telemetry data) while providing the bi-directional secure communication and registry of billions of devices.

Many companies prefer Apache Kafka as their centralized data streaming hub for its proven reliability and scalability. Azure offers the HDInsight PaaS as an ideal way to quickly deploy and manage Hadoop-based projects such as Kafka, Spark, Storm and HBase. The Azure HDInsight service is based on open source Apache, allowing for portability from on-premise or other cloud platforms if that is a requirement. Azure Kafka for HDInsight simplifies the provisioning of a cluster of servers to several clicks online; however, this data streaming platform is considered Azure IaaS rather than PaaS.

Data processing or engineering data pipelines are best handled by the Azure HDInsight service for deploying a separate cluster of Azure Virtual Machines to run Apache Spark and centralize data pipeline processing with the Apache open source framework for large-scale data analytics applications. The Azure Spark for HDInsight service runs on a cluster of Azure Virtual Machines with the Spark core engine and includes Spark SQL, Spark Streaming, MLlib (for machine learning) and GraphX from Apache.

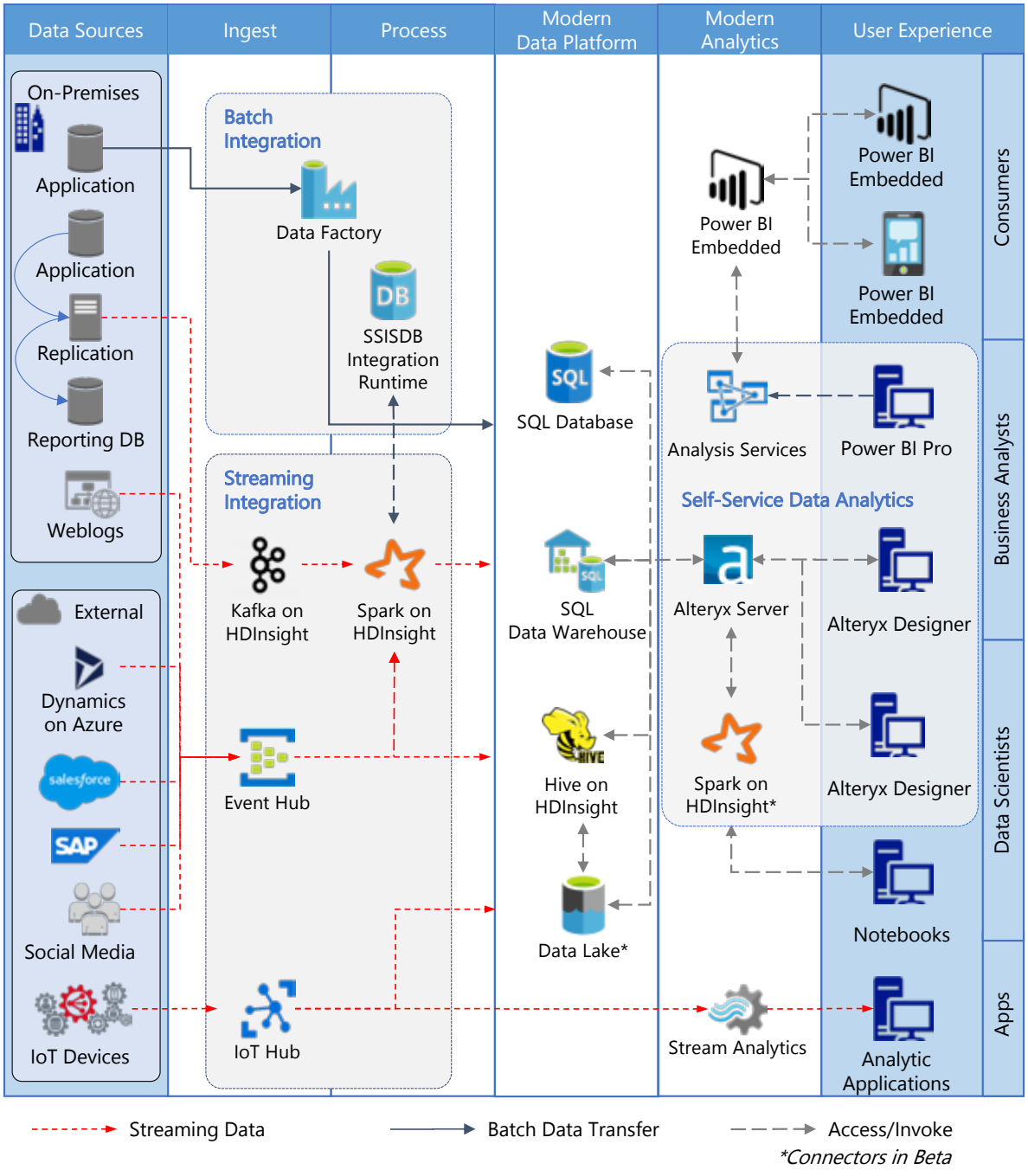


Figure 5: Logical Architecture for Modern Data Platform on Azure

The Azure Spark for HDInsight service simplifies provisioning the servers needed with no software to install and includes Jupyter (iPython) notebooks and R Server. The data pipelines executed in Spark for HDInsight subscribe to the data streaming hub, process the data, then output data to Azure databases and data storage, and also directly to Power BI, Tableau, Qlik and SAP (as well as others) for publishing real-time events to dashboards.

The Azure Data Factory service is a batch-oriented, codeless graphical user interface (GUI) environment that straddles both data ingestion and data processing functions. This service is known as an orchestration tool, and its placement in the architecture is due to the fact that it is a batch-oriented ETL service. Its scheduler starts data pipelines with any of its 60+ data connectors and then implements a data pipeline of ETL/ELT drag-and-drop functions, SQL Server Integration Services (SSIS) packages on Azure or jobs in Azure Spark for HDInsight before sending the output to Azure databases and data stores. The Azure Data Factory also has the benefit of being PaaS with no infrastructure to manage.

Azure Components for the Modern Data Platform

Azure has several database and data storage services that align with the Modern Data Platform principles to store data, from file and object formats to relational databases, analytic-oriented databases and NoSQL databases. The Modern Data Platform's conceptual architecture identifies three classifications of database engines that have different strengths in analytics for Azure database and data storage services. (Refer back to Figure 4.) From a data architecture perspective, Microsoft follows a hub and spoke architecture with the Azure SQL Data Warehouse being the hub of the architecture and Azure SQL Database being the spoke data marts for users to connect to. Power BI Embedded is also considered to be a spoke of the data warehouse with data caching for Power BI visualizations and dashboards.

Reference Data Class refers to subject-oriented reference data or mastered data that require the functionality of transaction performance and high user concurrency. For this purpose, the Azure SQL Database is the fully managed cloud database service based on the relational SQL Server database. The Azure SQL Database can scale easily as needed with minimal disruption and provides the in-memory OLTP capability for the faster updates, queries and concurrency needed by analytics. Azure SQL Database currently supports up to 4TB of data storage and 32GB of memory for reference class data such as customers, products or locations.

Analytics-Optimized Class is important for its analytic workload capabilities and queries that need to support many operational and ad hoc analytic sandboxes with the Modern Analytics Lifecycle. The Azure SQL Data Warehouse is a managed cloud service and column-oriented massively parallel processing (MPP) relational database that can independently scale compute and storage for analytics and data warehouse workloads.

Flexible Processing Class enables the enterprise data lake to become the foundational component of the Modern Data Platform, thus serving as the expansive repository for all enterprise data available for analytics. The Azure Data Lake is comprised of the Data Lake Store and Data Lake Analytics services. The Azure Data Lake Store service is a cloud-enabled version of the Hadoop distributed file system and YARN that leverages the open HDFS standard for deploying enterprise data lakes to store trillions of files and run massively parallel analytics. Alteryx SQL access to structured data in the Azure Data Lake will be via Hive on HDInsight along with the Hive metastore or native access to CSV, JSON, and Avro file formats. Data stored in the Azure Data Lake is always encrypted at rest using service or user-managed keys in Azure Key Vault. Data in transit is protected with Secure Socket Layer (SSL). Meanwhile, the Azure Data Lake Analytics service is a cloud service that uses massively parallel processing in U-SQL, R, Python and .NET without requiring any infrastructure to manage query access to the Azure Data Lake Store. The service is billed in Azure Data Lake Analytics Units (AU) to only charge for the processing used per job.

Azure Blob Storage is a massively scalable object storage service for exabytes of unstructured data as billions of objects organized in hot, cool or archive tiers. Azure Table Storage is a massive NoSQL key-value store for petabytes of semi-structured data or flexible data schema.

Azure Components for the Modern Analytics Lifecycle

The Modern Analytics Lifecycle focuses solving business analytics challenges through a self-service approach for finding, exploring and preparing data for use in analytic model development that can be deployed in operational business applications. The business analytics capabilities enabled by the Modern Data Platform include enterprise BI with reporting and data visualization, self-service data analytics, and enterprise data science and AI. While Azure's Power BI and Analysis Services can take care of enterprise BI, and Azure Spark for HDInsight fulfills enterprise data science and AI needs, Alteryx is better-suited for the self-service data analytics capabilities while integrating with Power BI and Apache Spark implementation as needed.

The **Alteryx Server** provides the self-service data analytics platform where data and analytic workflows are deployed and managed after development in Alteryx Designer. The Alteryx Server is comprised of a Controller component that coordinates all job activities, a Worker Node that executes data and analytics processing, a Database component for storing any data necessary for job execution and metadata, and a Gallery component for making data sets and data workflows available to others in the enterprise. The Alteryx Server (with all of its components) can be deployed on Azure as IaaS with a single Azure Virtual Machine with the Windows Operating System and Managed Disk. A basic single virtual machine installation is recommended with a minimum recommended size

of 8 vCPUs and 1TB of Managed Disk. The four server components can also be distributed over many dedicated virtual machines in the virtual network for increased scalability with data processing or user access. In either case, the size of the virtual machines and their number of vCPUs will need to align with the purchased Alteryx Server license. Alteryx Server is also available in the Azure Marketplace for a quick-to-deploy, pay-per-use model if desired.

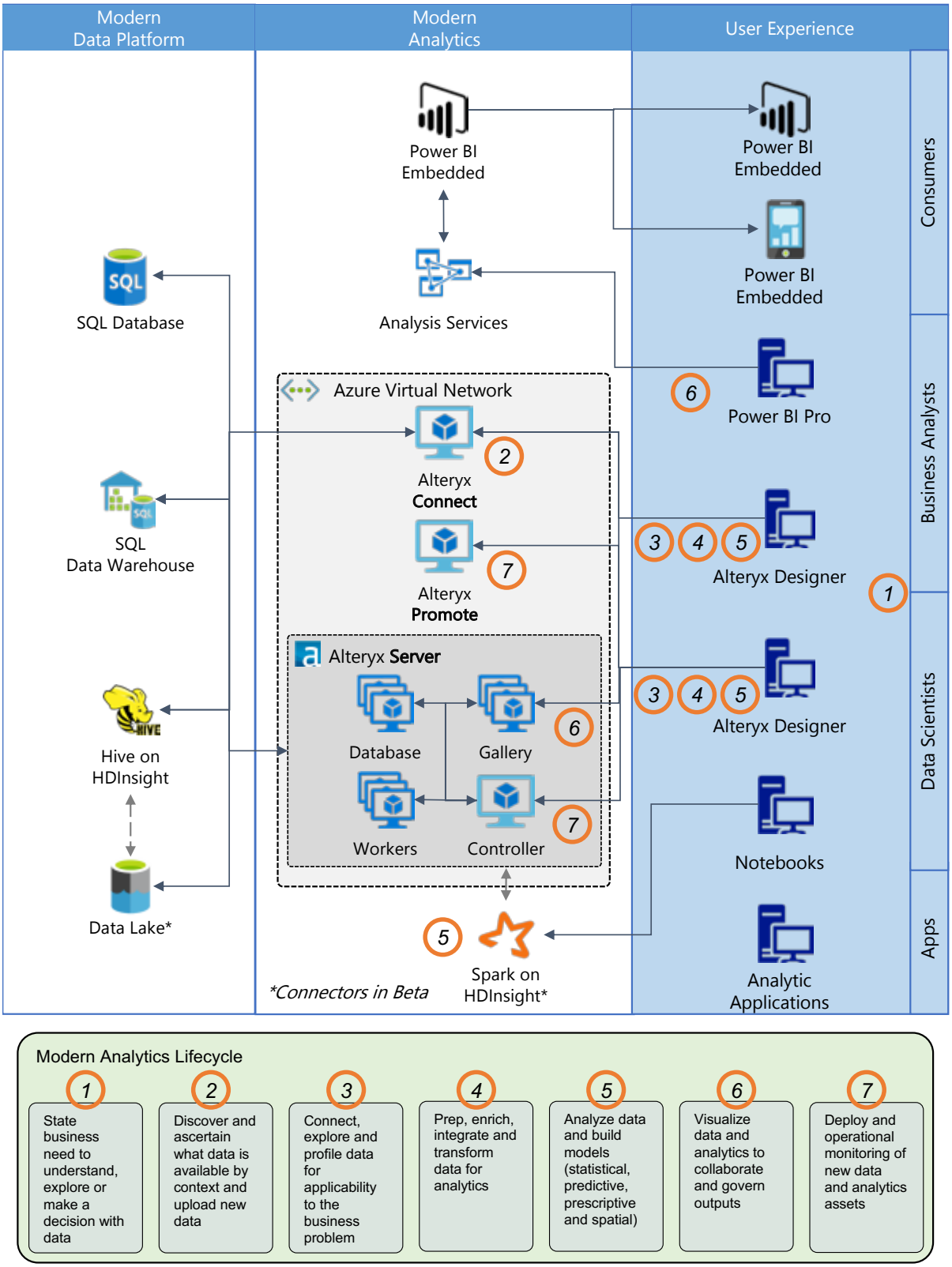
When the Alteryx Server is combined with Alteryx Connect and Alteryx Promote servers, the platform delivers the complete Modern Analytics Lifecycle on Azure while working with other Azure services in the Modern Data Platform. Alteryx Connect is dedicated to assisting the business analyst or data scientist in finding the data that is available on Azure or on-premises. Further, Alteryx Promote is dedicated to operationalizing model deployment and management so that business applications can embed real-time or batch-oriented analytics for widespread use and business impact.

Modern Analytics Lifecycle in Action with Alteryx on Azure

When Alteryx is deployed on the Azure cloud environment, the Modern Analytics Lifecycle becomes fully realized with Azure components fulfilling the modern data platform and Azure services supporting the Alteryx platform. (See Figure 6 on page 13.)

Business analysts and power users in the enterprise are faced every day with the immediate needs of business questions and challenges to solve. For the empowered business analyst, the Modern Analytics Lifecycle begins in Alteryx Connect to identify what data is available on-premises and in the Azure databases and data stores. Business analysts can read definitions, metadata and comments about the data they're interested in to evaluate whether the data fits their business need. With the Modern Data Platform on Azure, the connections for their Azure Data Lake Store, Azure SQL Database or Azure SQL Data Warehouse are all available for access. These data connections allow them to explore and profile the data in Alteryx Designer to determine how the data sources could be incorporated into their project. Depending on the project complexity and size, the business users may request a temporary analytic sandbox schema to be created for them in Azure SQL Data Warehouse or Azure Data Lake to store their working data or output for further verification and analysis. The output data set could be written back into the Azure SQL Data Warehouse or Azure Data Lake as an analytic sandbox for validation and data visualization development. Operationalizing data preparation and analytic models follows a collaborative governance and validation phase, which governs the output, and data management determines where the newly created analytic data should live in the architecture in order to remove it from the temporary sandbox. Deploying the analytic models can be accomplished with Alteryx Promote for making the analytic model available to business applications as a REST API or application code. Alteryx Promote will facilitate version control, unit testing and monitor model performance.

Figure 6:
Workflow for Business Analysts and Data Scientists



Data science teams follow the Modern Analytics Lifecycle similarly but with more emphasis on different types of transformations required for data science, workflow steps that involve more statistical and analytic processing including R and Python, and publishing of analytic APIs for applications. Alteryx Connect is a working environment that dramatically reduces the time and effort to identify data that is available in the Azure databases and data stores, explore it and integrate it for use. Data scientists seek out clusters and patterns within the data by analyzing as many features or attributes about event data such as transactions and behavior. Driving finer-grain segmentation will allow for A/B testing of predictive routines for optimizing next-best actions and recommendations. Along the way of this entire process, data sets can be written to Azure SQL Database, SQL Data Warehouse or Data Lake so that data visualization can communicate these segmentations, predictions and outcomes. Additionally, reusable derived data sets can be stored in Azure SQL Data Warehouse or Azure Data Lake for use by other analytic workflows, thereby accelerating their development time and data quality. Analytic models in the Alteryx workflows can incorporate Python and R code or access an Apache Spark cluster for execution. The analytic workflows are created in Alteryx Designer, are scheduled and executed in Alteryx Server and can be made available with Alteryx Promote to applications, which would call a published REST APIs.

Operationalizing Analytics with Alteryx on Azure

Performance and Scalability

For most production deployments, the Alteryx Server as a single Azure Virtual Machine with 8 vCPUs and 1TB of Managed Disk is ideal to get up and running quickly with your enterprise. Keep in mind that Azure Virtual Machines have vCPUs that are usually based on two hyper-threads per CPU core when maintaining your enterprise license agreement with Alteryx. However, it is reassuring to know your analytic lifecycle processes can scale as needed very easily within the Azure environment.

Two of the four Alteryx Server components – Worker and Gallery – are where horizontal scale-out will likely be needed. In order to do so, when moving from a single Azure Virtual Machine in the Azure Virtual Network to many Azure Virtual Machines dedicated to different Alteryx components, a private subnet should be created in the Virtual Network for each Azure Availability Zone to allow the Alteryx Server components to securely communicate with each other. This private subnet can be a handy security domain that will ensure that Azure Virtual Machines are not inadvertently accessed, as only the Alteryx components need to communicate with each other.

The Controller component orchestrates and distributes the Alteryx workflows across the Worker nodes as needed. There are two approaches to implement Worker node scalability. First, you can manually configure Alteryx Worker nodes as dedicated Azure Virtual Machines that are selected for the specific workload characteristics. Manually configuring the Azure Virtual Machine sizes will give you further control of different workloads from the Alteryx Controller. The Alteryx Controller will configure and administrate the Worker nodes as Azure Virtual Machine IP addresses on the private subnet. Alternatively, you can work with Azure's Virtual Machine Scale Sets to dynamically add and remove Azure Virtual Machines for Worker nodes by leveraging Azure Insights Autoscale.

The scalability process is slightly different for the Gallery component, which does not rely on the Controller to coordinate workloads. Increasing the number of Azure Virtual Machines used for the Gallery component is dependent on the number of users accessing the Gallery. Usually, separating Gallery component to another Azure Virtual Machine is sufficient since it isolates the Gallery workload from the other Alteryx Server components on a shared single Azure Virtual Machine. However, the Gallery can be further scaled out by replicating it to other Azure Virtual Machines and implementing the Azure Load Balancer as a connection point that will distribute the users over many copies of the Gallery Azure Virtual Machines.

Environments, Fault Tolerance and Disaster Recovery

The recommended best practice for deploying the Alteryx Server, Connect Server and Promote Server is to maintain at least two Availability Zones for fault isolation of services and applications deployed. Some Azure regions offer up to four Availability Zones if that is a preference or requirement for regional fault isolation. Azure Virtual Machines are deployed in Availability Zones by specifying the correct subnets; Azure Virtual Machine Scale Sets can do this automatically if desired. The Azure Data Lake and Blob Storage preserve data residency and compliance boundaries with two or more regions through Azure geographies.

The isolation provided by Availability Zones within a region does not protect against some classes of wide-scale geographic disasters, despite not sharing power, transit or physical features like floodplains or vulnerability to fire. An Azure geography can withstand a complete region failure with its dedicated high-capacity networking infrastructure to meet most corporate disaster recovery (DR) policies. Data replication and configurations can then be maintained between regions leveraging the required high-speed network. In some cases, companies may be able to configure each region as an active-active approach that offers full load balancing capabilities. Alteryx Server does not support multi-region active-active in the same deployment but can be configured to support active-passive or failover from replication.

Conclusion

Competitive businesses today are pressured for time and must optimize insights in data to continually transform in the face of digital disruption and continue to meet progressively higher customer expectations for how their products or services are valued by tech-savvy consumers. The Modern Analytics Lifecycle transforms organizations and competitive analytics programs by empowering users to work with data and deliver analytics to every part of the business. When barriers and inefficiencies are eliminated at each step in the Modern Analytics Lifecycle, business users, analysts and data scientists can confidently work in their self-service data and analytics environment, and organizations can meet modern business demands for speed and agility.

The Modern Data Platform is carefully architected to deliver the essential capabilities and efficiencies that the Modern Analytic Lifecycle can take advantage of fully. Combined with Azure and leveraging its elasticity and reliable scalability, this is a proven combination for companies to adopt.

This edition published July 2018.

About the Author

John O'Brien is Principal Advisor and CEO of Radiant Advisors. A recognized thought leader in data strategy and analytics, John provides research, strategic advisory services and mentoring that guide companies in data strategy, architecture, analytics and emerging technologies.

About Radiant Advisors

Radiant Advisors is an independent research and advisory firm that delivers innovative, cutting-edge research and thought-leadership to transform today's organizations into tomorrow's data-centric industry leaders.

Radiant Advisors
Boulder, CO USA
Email: info@radiantadvisors.com

To learn more, visit www.radiantadvisors.com