



AZURE ENVIRONMENT ASSESSMENT FOR INTERNATIONAL ANALYTICS FIRM

ORGANIZATION

Our client is an information, intelligence, and analytics service provider for large corporations and governments that drive the economies worldwide. Founded in 1959 and headquartered in London, our client has grown both organically and through acquisitions. Their subject-matter experts closely follow the markets and forces impacting the industry and provide a balanced perspective. Our client also offers on-demand consultation services to provide quick, actionable answers. Our client has 5,000+ analysts, data scientists, and domain experts to provide a highly integrated view of their customer's world by connecting data across different variables. Our client's information and intelligence enable its customers to identify opportunities, mitigate risks, and solve problems.

CHALLENGE

Our client used Azure Table Storage and Azure Blob Storage for querying and reporting from large volumes of infrastructure logs (time series data), including vROps, Pure performance data, WAN link performance data from SolarWinds, Lat Long information from Tangoe, backup information from Commvault, Azure/AWS path vulnerability data from Qualys, incident and asset management data from ServiceNow, and AWS inventory data from CloudCheckr. The performance of Azure Table Service for these use cases was not satisfactory. They used Power BI (Pro License) and kept all the Power BI user workspaces under 5GB due to Pro License size limitations (10 GB per user).

They moved some of the data from Azure Table Storage to Azure Data Lake Storage Gen-2 for the data queried from Power BI since they were getting time-out issues on Table Storage. There were a lot of ad hoc queries running from Azure Table Storage, which were very slow. Scheduled data refreshes into Azure happen daily (every 30 minutes, every 4 hours, or once a day). Further issues came from moving their Azure Blob storage to Azure Data Lake Analytics Gen-2 to use account-level storage.

They were prepared to bring in infrastructure performance and log data from other sources such as on-premise SCCM data, Linux server validation data from Red Hat Satellite, AWS/Azure MS SQL Server licensing data from Qualys, system coverage data from Splunk, costing data from Datadog, costing data from PagerDuty, server coverage and security information from CrowdStrike, and Oracle database information from OEM.

Our client did not have in-house Azure expertise to guide and establish a future state data architecture on Azure to accommodate analytics and reporting from large volumes of data. They needed a trusted partner to validate their Azure environment and provide recommendations.

SOLUTION

XTIVIA performed an Azure environment assessment and data analysis. The discovered root cause for poor Table Storage performance included naming conventions issues, data quality, and current architecture reliability to handle large data volumes for future use cases. XTIVIA recommended migrating the Table Storage to Cosmos, utilized Azure Analysis Services as a semantic layer, Databricks for Data Engineering and Stream Analytics, Machine Learning use cases, and provided architecture recommendation for future use cases. XTIVIA's architectural recommendations included:

- 1. Migrating Table Storage to Cosmos:** The current table storage could not handle complex ad hoc queries due to improper partition key and row key selection. We recommended moving this data and queries to Cosmos since it provides different APIs and has a query processing engine just like any other NoSQL database. Existing Table API queries can be moved over to Cosmos without any changes. In addition to this, Cosmos supports auto-indexing on columns to guarantee performance. All the Table API queries can be carried over to SQL API in the longer term for better querying capabilities. We provided additional performance improvement design patterns for Table Storage as well.

BY THE NUMBERS:

50,000 Customers in 140 countries

85% of Customers in Global 500

600+ Industry Subject-Matter Experts

KEYWORDS

Azure Environment Assessment, Azure Data Lake, Azure Storage, Azure Architecture

SOFTWARE

Azure Table Storage | Azure Data Lake Storage Gen-2 | Power BI | Azure Cosmos DB | Azure Databricks | Azure Event Hub | Azure Data Explorer



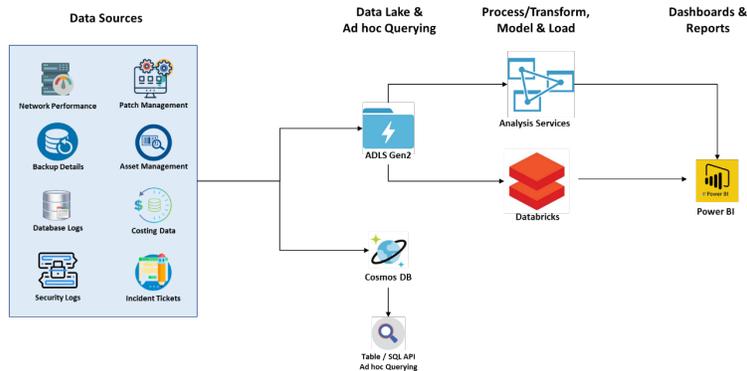


AZURE ENVIRONMENT ASSESSMENT FOR INTERNATIONAL ANALYTICS FIRM

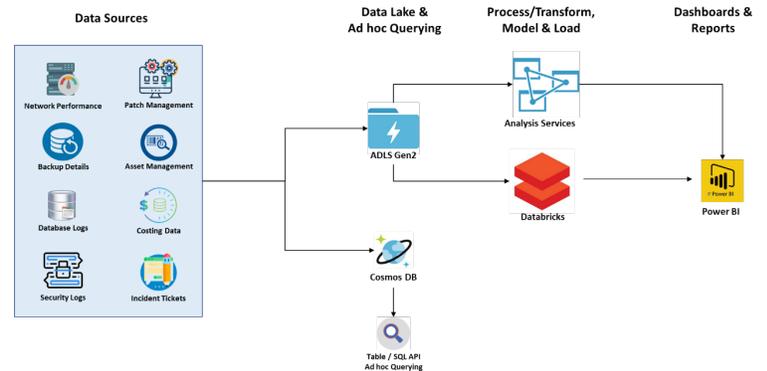
SOLUTION (CONTINUED)

- Use of Naming Conventions:** XTIVIA recommended creating naming conventions for Azure resources, data asset names, attribute names, etc., and publish them across the organization to enforce standards and consistency. The recommendations were also made in blob naming as the combination of account, container, and blob acts as a partition key used to partition the blob data into different ranges.
- Analytics Enablement:** XTIVIA recommended using Azure Databricks for data engineering and machine learning. Azure Databricks provides a collaborative workspace for data engineers, data scientists, and business analysts. It can be linked to Azure AutoML to automate and manage the end-to-end data science lifecycle from data preparation till model deployment. Stream analytics use cases involving large volumes of infrastructure logs that can be handled using Azure Databricks and Event Hub.
- Ad hoc Capabilities on Streaming Data:** XTIVIA recommended using Azure Data Explorer for exploring and mining terabytes of data. Azure Data Explorer provides multiple querying options (Kusto query language, Power BI, Tableau, ODBC, Excel, etc.) along with indexing and joining capabilities.
- Enterprise Semantic Layer (Optional):** XTIVIA recommended building a semantic layer by integrating multiple time-series datasets with common business keys for reporting on millions of rows. We provided this as an option for a possible future use case of ad hoc analysis on large data volumes by a large population of end-users (with limited infrastructure domain knowledge).

Phase One | Short Term



Phase Two | Long Term



BUSINESS RESULT

The detailed analysis and future-state architecture provided our client with a better understanding of Azure environment best practices, tactical, strategic objectives, and solution tradeoffs. The deliverables from the engagement provided the basis for cost-benefit analysis, budgeting, staffing, and risk management in implementing the solution. The client also gained an understanding of the processes and data required for machine learning use cases relevant to infrastructure logs and performance data.

XTIVIA OVERVIEW

XTIVIA is a proven technology integration and innovation company known for delivering leading-edge technology solutions to our clients' specific requirements, regardless of project complexity. We bring next-level business operations to reality through Application Development, Business Intelligence, Data Warehousing, Database Support & Management, Enterprise Information Management, Digital Experience Solutions and Enterprise Resource Planning. Our success stems from a demonstrated ability to deliver deep expertise via professional services, empowering clients to leverage their chosen technology successfully, competitively and profitably. XTIVIA has offices in Colorado, New York, New Jersey, Missouri, Texas, Virginia, and India. www.xtivia.com