Simplifying Data Governance and Classification in Banking and Financial Organizations with Azure Cloud.

Problem Statement

Let us deep dive into something very critical in today's age as everyone after it, no matter they are in BFSI or any kind of domain. Everyone wants to safeguard their data at personal to large organization level. We know now DATA is new "Oil" and with this data we can manipulate every decision. This becomes even mission critical for health care and banking organization as risk of this data is paramount.

So, Let's discuss how i have been working with biggest Bank in world. This involves managing vast amounts of sensitive information while adhering to stringent regulatory requirements. We will explore how Azure empowers BFSI organizations to tackle this challenge effectively.

Understanding Data Governance and Classification

Data governance involves the management of data assets, including data quality, security, and compliance. For BFSI, where data privacy and regulatory compliance are paramount, a robust data governance strategy is non-negotiable. Classification, on the other hand, is the process of labelling data to identify its type, sensitivity, and importance. In general, we would classify data based on its use and compliance requirement. for example, if we are dealing with PDI (personal identifiable information) then we would tag it as strictly confidential data, another classification would be Confidential data, limited access, and public data (which lowest level of classification). Public data also means this already available on internet and does not need to protect utmost care at very high level.

Key thing not only to use azure offering to safe guard your data but also understand your data flow and at each level how you are monitoring and safe guarding leakages, Let us see data workflow in general.

Data Ingestion: Data is ingested from multiple sources, including customer transactions, market data feeds, and internal operations.

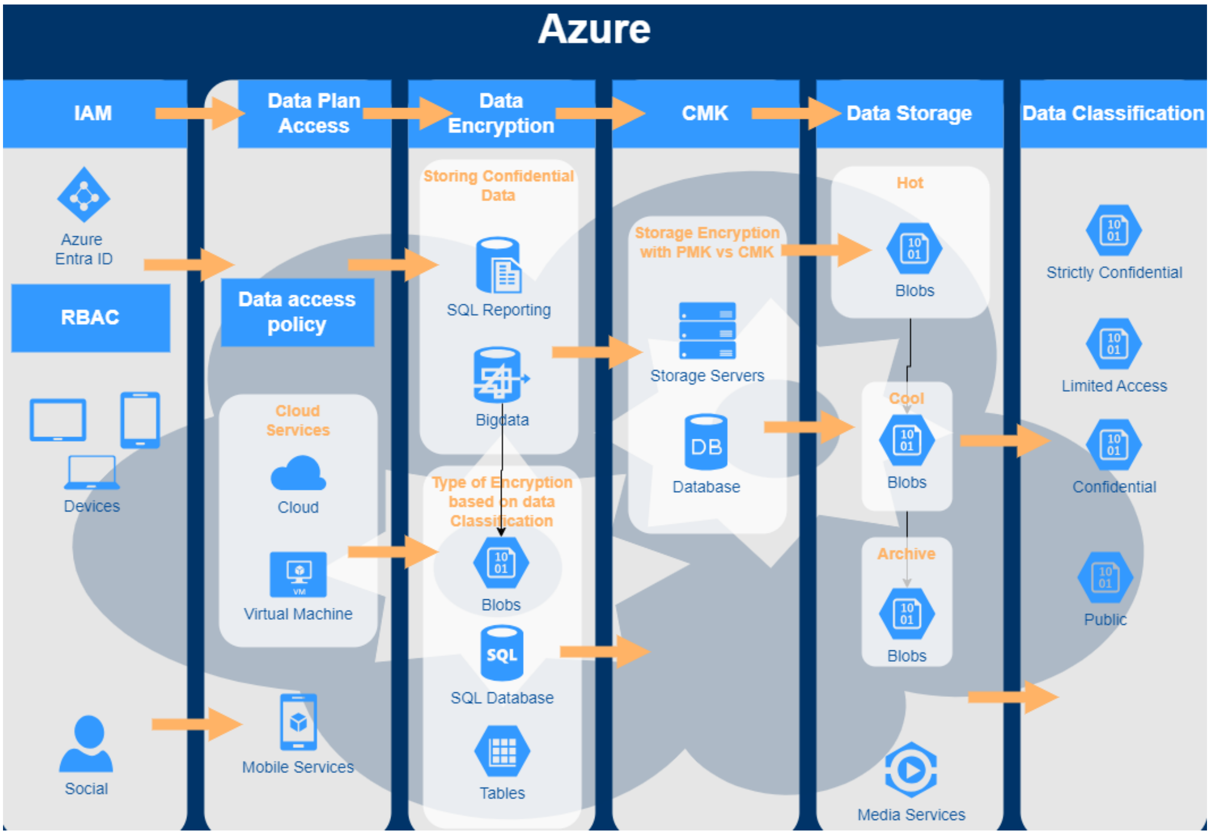
Data Classification: Azure Information Protection and Azure Purview classify the data based on predefined policies and sensitivity levels.

Data Storage: Azure Data Lake and Azure SQL Database store the classified data securely.

Data Access: Azure Security Center ensures that only authorized users and applications can access the data.

Monitoring and Compliance: Azure Policy and Blueprints continuously monitor data access and usage, ensuring compliance with industry regulations.

Solution/ Architecture:



Now, Lets move what and which Azure Services we can leverage for Data Governance and Classification.

Technical Details and Implementation of solution:

Microsoft Azure offers a suite of services that align perfectly with the needs of BFSI: Azure Information Protection: This service allows you to classify and label data based on its sensitivity. It ensures that data is protected, tracked, and controlled even when shared outside the organization.

Azure Purview: Formerly known as Azure Data Catalog, this service provides a holistic view of your data estate. It enables data discovery, classification, and cataloging, simplifying data governance.

Azure Policy and Blueprints: These tools enable you to define and enforce data governance policies across your organization. You can ensure compliance with industry regulations and internal standards.

Azure Security Centre: With advanced threat protection, this service helps safeguard your data from security threats, ensuring compliance with industry-specific regulations like GDPR and HIPAA.

Security is critical for BFSI and for better reason as their stake is on their name, they dont want to be in limelight for security reasons. so when you store you data in azure storage you have to ensure you are not just selecting default options but explore many other below options.

Private link and private endpoints this can be used in azure to keep your data on your network and not exposed to public internet at any level.

Customer managed Key : This will build confidence of your security team if you manage CMK (customer managed key) and store in key vault then manage its lifecycle.

Encryptions for data at motion and data at rest with highest level for encryption.

Data Plan and Control plan access will define who and what can be accessed by whom for how long.

Data Lifecycle for move data across hot, cool and archive tier for ensure you data is available at lowest cost.

Using above azure offering we have ensured that **Data plane and control plan** access is managed controlled and this was enforced by Azure policy and implemented using IaC with terraform. Also based on data classification we have ensured all strictly confidential, confidential, and limited access data stays on our **own network all the time using Azure private link.**

Once we have access and private link configured for storage, we used **customer managed key** to encryption of data and stored and managed key lifecycle using **azure key vault**. Later we setup storage life cycle based on data usage so we could move data from **Hot, Cool then Archive** tier for cost management. We have implemented **azure policy in audit mode initially** to understand pattern and then implemented in deny mode for any non-compliance storage account creation which do not follow proper CMK and Encryption at the time for storage account creation, this helped to reduced non-compliance storage getting provisioned.

Challenges in implementing the solution

While implementing strict rules or enforcing azure policy we got lots of push back from many business teams, also key challenge was to educate dev team why we are implementing and how they can follow the correct process to ensure there solution in align with requirements.

We had challenge in CMK and Key Vault integration for encryption key management, how keys can be seamlessly managed across lifecycle. This was quick fix as Storage and Azure key vault on same platform.

Leveraging private link for azure blob and dfs group was new to us then managing DNS entry and automating approval took sometime for implementation.

Business Benefit

In the BFSI sector, data governance and classification are vital for maintaining trust and regulatory compliance. Implementation for this model helped us to migrate and move new data on azure cloud with correct tagging and security level with optimize cost. This implementation was also trailblazer for many other business team to move there data on cloud with confidence.