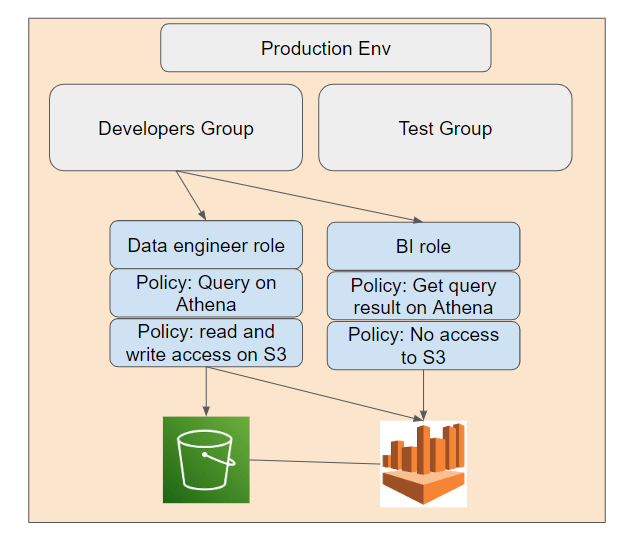
Q9) Your organization stores sensitive data in S3 and you need to restrict access to certain columns in Athena. How would you set up the necessary permissions and access controls?

Example of high-level overview of permission and access control with various groups/roles with S3 and Athena



**Step 1: Understand AWS IAM roles and groups**

- Two security concerns to be addressed: sensitive data and limited access to specific columns in Athena.In an organization, permissions are given to users on the basis of least privilege (just enough resources to do the job).

- Understand the group permissions and policies given to users (e.g. developers group)

- Roles are attached temporarily to an instance of an AWS resource such as Athena to be given/restricted to certain functions of the resource.

**Step 2: AWS Glue Data Catalog Permissions**

- Ensure that the AWS Glue Data Catalog is properly configured with the appropriate permissions. AWS Identity and Access Management (IAM) roles should be defined to control who can create, modify, or delete tables and databases in the Data Catalog.

**Step 3: AWS Athena Permissions**

- Define IAM roles and policies for AWS Athena. These roles determine who can run queries and access data in Athena.

- Create a specific IAM policy that grants access to only the necessary columns in the underlying S3 data sources. This policy should use AWS Glue Data Catalog as a resource for column-level access control.

**Step 4: Fine-Grained Column-Level Permissions**

- Utilize AWS Lake Formation to implement fine-grained column-level access control. AWS Lake Formation allows you to define and enforce data access policies at a granular level.

- Create and manage permissions at the column level, specifying which users or groups can access specific columns in Athena queries. In Lake Formation, under data filter, users could exclude the sensitive column as such student\_NRIC. With the filter set, users/roles could be applied with filter. As such, the users/role could not view the specific sensitive column.

**Step 5: IAM Role Assignment**

- Assign the IAM roles created to the users or groups who need access to Athena. Ensure that these roles have associated policies that grant permission to run Athena queries and access the relevant S3 data.

**Step 6: Data Catalog Encryption**

- Implement encryption at rest and in transit for data stored in Amazon S3. This safeguards sensitive data from unauthorized access even if someone gains access to the underlying S3 objects.

**Step 7: Regular Auditing and Monitoring**

- Set up auditing and monitoring of Athena queries and data access. AWS CloudTrail can be configured to log all API calls related to Athena, allowing you to track who accessed sensitive data and when.

**Step 9: Regularly Review and Update Permissions**

- Regularly review and update IAM policies, role assignments, and permissions as the team is constantly reshaping in a startup environment. Remove unnecessary access and permissions to minimize potential risks.

Considerations:

1. Least Privilege

Follow the principle of least privilege, granting users and roles only the minimum access required to perform their tasks. Avoid overly permissive policies. For example, a Business Intelligence (BI) role should have access to athena and quicksight, while access is restricted to S3.

2. Sensitive Data Identification depending on country’s data protection policy

Countries differ in their data protection policy. For example, Europe has strict PDPA regulation compared to Singapore or Malaysia.

3. Encryption

Ensure that data stored in S3 is encrypted both at rest and in transit. Use AWS Key Management Service (KMS) for encryption key management.

4. Data Governance

Establish data governance policies and practices to maintain data quality, security, and compliance.

5. Training and Awareness

Educate users and administrators about data security best practices and the importance of protecting sensitive information.

By following these steps and considerations, we can effectively restrict access to sensitive columns in Athena, safeguarding confidential data while enabling authorized users to perform necessary analytics and reporting tasks.