RFC - Addition of Access Management roles

ExternalRoles

Summary

Every organization can maintain their own access management system for managing the data assets. Users in the organisation will be able to access the dataset only if a particular role is provisioned under the org's access management system.

The aim of this feature is to enable users to see the required access & roles of an organisation's access management system for accessing a dataset. Also request for the appropriate roles from the Datahub frontend to the organisation if the role is not provisioned for the user in organisation's access management system.

Motivation

Currently entities such as Dataset can have properties attached as aspects. We do not want to attach the org's roles as properties to the dataset since the external roles fetching and dataset fetching are two different processes.

Roles under the organisation's scope, can be created at any point of time. Creating new entity would help us add the about the role names and properties of the external role. Also we can attach the provisioned users of a role to the entity. These users can be tied to corpuser in Datahub. This will enable the users to know if a role is already provisioned or yet to be requested to access the dataset.

A dataset can have one or more roles for access. READ, WRITE, ADMIN roles are the three different roles.

The external roles can be ingested in two different ways as below,

Ways of ingesting access roles

1. During ingestion of datasets, if access roles can be discovered, roles and access information can be added
2. There can be separate batch process which can be run offline to load IAM roles for all the datasets in the system

Requirements

1. Ability to associate ExternalRoles entity to Dataset
2. Same ExternalRole can be associated to multiple datasets

Detailed Design

Metadata model changes

1. Create new entity (**externalRole**) for external role access management
2. Add required roleproperties as aspect to externalRole
3. Also add required provisioned users as aspect to externalRole
4. Link the externalRole entity with Dataset entity with one or more relationship
5. Update the existing getDataset graphql api with the addition of new entity

In entity-registry.yml, add the new entity **externalRole**

```
  - name: externalRole
    category: core
    keyAspect: externalRoleKey
    aspects:
      - externalRoleProperties
      - externalRoleProvisionedUsers
```

ExternalRoleProperties

```
record ExternalRoleProperties {

  /**
   * Display name of the External Role in an organisation
   */
  @Searchable = {
    "fieldType": "TEXT_PARTIAL",
    "enableAutocomplete": true,
    "boostScore": 10.0
  }
  name: string

  /**
   * Description of the external Role
   */
  description: optional string

  /**
   * Can be READ, ADMIN, WRITE
   */
  type: optional string

  /**
   * Link to request access in external access management
   */
  requestlink: optional string

  /**
   * Created Audit stamp
   */
   @Searchable = {
     "/time": {
      "fieldName": "createdTime",
      "fieldType": "DATETIME"
     }
   }
   created: optional AuditStamp
}
```

ExternalRoleProvisionedUsers

```
record ExtRoleProvisionedUsers {

  /**
   * List of provisioned users of a role
   */
  provisionedUsers: array[ExtRoleProvisionedUser]

}
```

ExternalRoleProvisionedUser

```
record ExtRoleProvisionedUser {

    /**
     * Link provisioned corp user of datahub
     */
    @Relationship = {
      "name": "Has",
      "entityTypes": [ "corpuser,corpGroup" ]
    }
    provisionedUser: Urn
}
```

Non Requirement

A new **AccessManagement** tab will be added as part of the dataset details as below