**Set up resources with AWS CloudFormation**

We provide three [AWS CloudFormation](https://aws.amazon.com/cloudformation/) templates in this post: for the producer account, central account, and consumer account. Deploy the CloudFormation templates in the order of producer, central, and consumer, because there are dependencies between the templates.

The CloudFormation template for the central account generates the following resources:

* Two IAM users:
  + DataMeshOwner
  + ProducerSteward
* Grant DataMeshOwner as the LakeFormation Admin
* One IAM role:
  + LFRegisterLocationServiceRole
* Two IAM policies:
  + ProducerStewardPolicy
  + S3DataLakePolicy
* Create databases “credit-card” for ProducerSteward to manage Data Catalog
* Share the data location permission for producer account to manage Data Catalog

The CloudFormation template for the producer account generates the following resources:

* Two [Amazon Simple Storage Service](http://aws.amazon.com/s3) (Amazon S3) buckets:
  + credit-card, which holds one table:
    - Credit\_Card
* Allow Amazon S3 bucket access for the central account Lake Formation service role.
* One AWS Glue crawler
* One AWS Glue crawler service role
* Grant permissions on the S3 bucket locations credit-card-lf-<ProducerAccountID>-<aws-region> to the AWS Glue crawler role
* One producer steward IAM user

The CloudFormation template for the consumer account generates the following resources:

* One S3 bucket:
  + <AWS Account ID>-<aws-region>-athena-logs
* One Athena workgroup:
  + consumer-workgroup
* Three IAM users:
  + ConsumerAdmin

**Launch the CloudFormation stack in the central account**

To create resources in the central account, complete the following steps:

1. Sign in to the central account’s AWS CloudFormation console in the target Region.
2. Choose **Launch Stack:**  
   [Launch Stack](https://console.aws.amazon.com/cloudformation/home?region=us-east-1#/stacks/new?stackName=stack-central&templateURL=https://aws-bigdata-blog.s3.amazonaws.com/artifacts/lakeformationtbac/cfn/tbac-cross-account-central.yaml)
3. Choose **Next**.
4. For **Stack name**, enter stack-central.
5. For **DataMeshOwnerUserPassword**, enter the password you want for the data lake admin IAM user in the central account.
6. For **ProducerStewardUserPassword**, enter the password you want for the producer steward IAM user in the producer account.
7. For **ProducerAWSAccount**, enter the AWS <ProducerAccountID>.
8. Choose **Next**.
9. On the next page, choose **Next**.
10. Review the details on the final page and select **I acknowledge that AWS CloudFormation might create IAM resources**.
11. Choose **Create** **stack**.
12. Collect the value for LFRegisterLocationServiceRole on the stack’s Outputs tab.

**Launch the CloudFormation stack in the producer account**

To set up resources in the producer account, complete the following steps:

1. Sign in to the producer account’s AWS CloudFormation console in the target Region.
2. Choose **Launch Stack**:  
   [](https://console.aws.amazon.com/cloudformation/home?region=us-east-1#/stacks/new?stackName=stack-producer&templateURL=https://aws-bigdata-blog.s3.amazonaws.com/artifacts/lakeformationtbac/cfn/tbac-cross-account-producer.yaml)
3. Choose **Next**.
4. For **Stack name**, enter stack-producer.
5. For **CentralAccountID**, copy and paste the value of the <CentralAccountID> .
6. For **CentralAccountLFServiceRole**, copy and paste the value of the LFRegisterLocationServiceRole collected from the stack-central.
7. For **LFDatabaseName**, keep the default value of the lf-ml database name.
8. For **ProducerStewardUserPassword**, enter the password you want for the data lake admin IAM user on the producer account.
9. Choose **Next**.
10. On the next page, choose **Next**.
11. Review the details on the final page and select **I acknowledge that AWS CloudFormation might create IAM resources**.
12. Choose **Create stack.**

**Launch the CloudFormation stack in the consumer account**

To create resources in the consumer account, complete the following steps:

1. Sign in to the consumer account’s AWS CloudFormation console in the target Region.
2. Choose **Launch Stack**:  
   [](https://console.aws.amazon.com/cloudformation/home?region=us-east-1#/stacks/new?stackName=stack-consumer&templateURL=https://aws-bigdata-blog.s3.amazonaws.com/artifacts/lakeformationtbac/cfn/tbac-cross-account-consumer.yaml)
3. Choose **Next**.
4. For **Stack name**, enter stack-consumer.
5. For **ConsumerAdminUserName** and **ConsumerAdminUserPassword**, enter the user name and password you want for the data lake admin IAM user.
6. For **ConsumerAnalyst1UserName** and **ConsumerAnalyst1UserPassword**, enter the user name and password you want for the consumeranalyst1 IAM user.
7. For **ConsumerAnalyst2UserName** and **ConsumerAnalyst2UserPassword**, enter the user name and password you want for the consumeranalyst2 IAM user.
8. Choose **Next**.
9. On the next page, choose **Next**.
10. Review the details on the final page and select **I acknowledge that AWS CloudFormation might create IAM resources**.
11. Choose **Create stack**.

**Configure Lake Formation cross-account sharing**

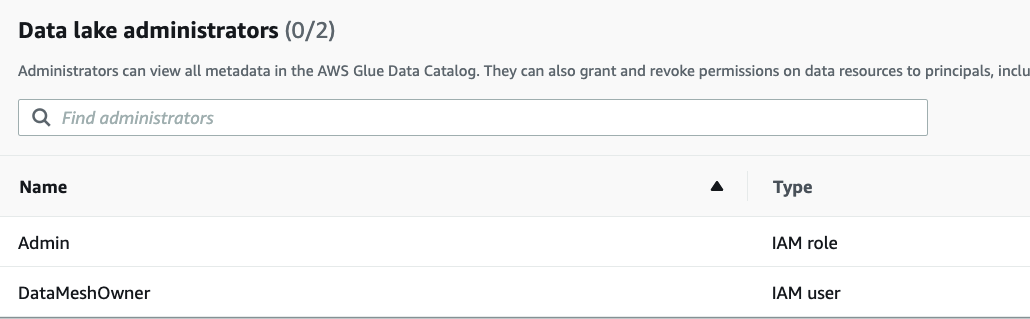
After you create your resources with AWS CloudFormation, you perform the following steps in the producer and central account to set up Lake Formation cross-account sharing.

**Central governance account**

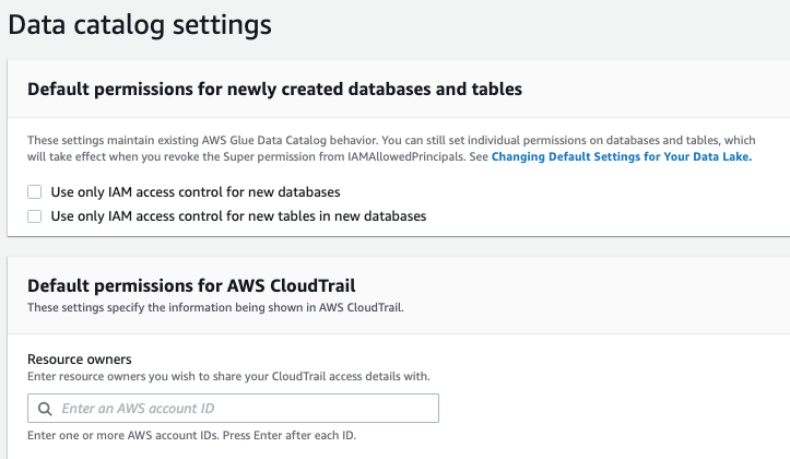
In the central account, complete the following steps:

1. Sign in to the Lake Formation console as admin.
2. In the navigation pane, choose **Permissions**, then choose **Administrative roles and tasks**.

The CloudFormation template added the data mesh owner as the data lake administrator.

[](https://d2908q01vomqb2.cloudfront.net/b6692ea5df920cad691c20319a6fffd7a4a766b8/2022/04/04/BDB-2091-image013.png)

Next, we update the Data Catalog settings to use Lake Formation permissions to control catalog resources instead of IAM-based access control.

1. In the navigation pane, under **Data catalog**¸ choose **Settings**.
2. Uncheck **Use only IAM access control for new databases**.
3. Uncheck **Use only IAM access control for new tables in new databases**.
4. Choose **Save**.  
   [](https://d2908q01vomqb2.cloudfront.net/b6692ea5df920cad691c20319a6fffd7a4a766b8/2022/04/04/BDB-2091-image015.png)

Next, we need to set up the AWS Glue Data Catalog resource policy to grant cross-account access to Data Catalog resources.

1. Use the following policy, and replace the account number and Region with your own values:
2. {
3. "PolicyInJson": "{\"Version\" : \"2012-10-17\",\"Statement\" : [ {\"Effect\" : \"Allow\",\"Principal\" : {\"AWS\" : [\"arn:aws:iam::<ProducerAccountID>:root\",\"arn:aws:iam::<ConsumerAccountID>:root\"]},\"Action\" : \"glue:\*\",\"Resource\" : [ \"arn:aws:glue:<aws-region>:<CentralAccountID>:table/\*\", \"arn:aws:glue:<aws-region>:<CentralAccountID>:database/\*\", \"arn:aws:glue:<aws-region>:<CentralAccountID>:catalog\" ],\"Condition\" : {\"Bool\" : {\"glue:EvaluatedByLakeFormationTags\" : \"true\"}}}, {\"Effect\" : \"Allow\",\"Principal\" : {\"Service\" : \"ram.amazonaws.com\"},\"Action\" : \"glue:ShareResource\",\"Resource\" : [ \"arn:aws:glue:<aws-region>:<CentralAccountID>:table/\*\", \"arn:aws:glue:<aws-region>:<CentralAccountID>:database/\*\", \"arn:aws:glue:<aws-region>:<CentralAccountID>:catalog\" ]} ]}",
4. "EnableHybrid": "TRUE"

}

Replace the <aws-region>, <ProducerAccountID>, <ConsumerAccountID> and <CentralAccountID> values in the above policy as appropriate and save it in a file called policy.json.

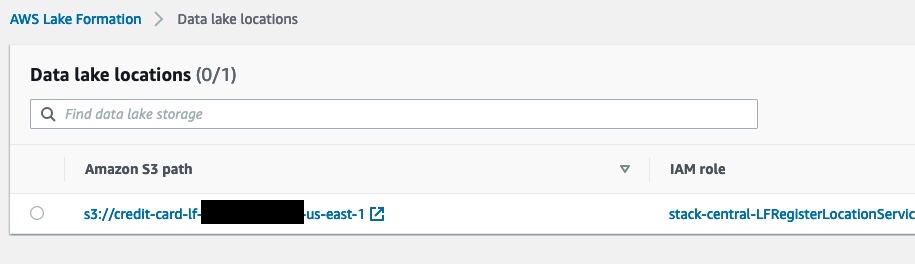
1. Next, run the following [AWS Command Line Interface](http://aws.amazon.com/cli) (AWS CLI) command on [AWS CloudShell](https://aws.amazon.com/cloudshell/).

aws glue put-resource-policy --region <aws-region> --cli-input-json file://policy.json

For more information about this policy, see [put-resource-policy](https://docs.aws.amazon.com/cli/latest/reference/glue/put-resource-policy.html).

1. Next, we verify the source data S3 bucket isregistered as data lake location in the central account. This is completed by the CloudFormation template.
2. Under **Register and ingest** in the navigation pane, choose **Data lake locations.**

You should see the S3 bucket registered under the data lake locations.



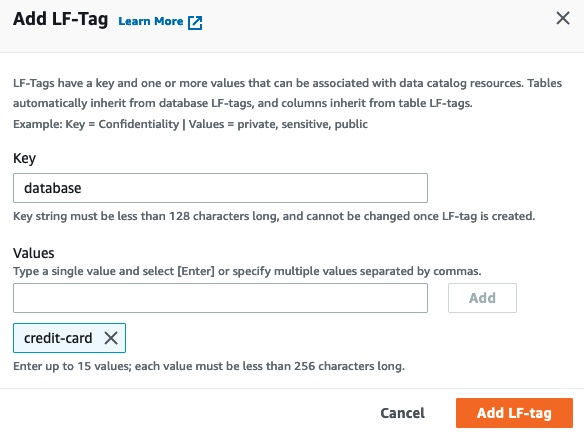
## Configure Lake Formation Data Catalog settings in the central account

After we complete all the prerequisites, we start the data mesh configuration. We log in as DataMeshOwner in the central account.

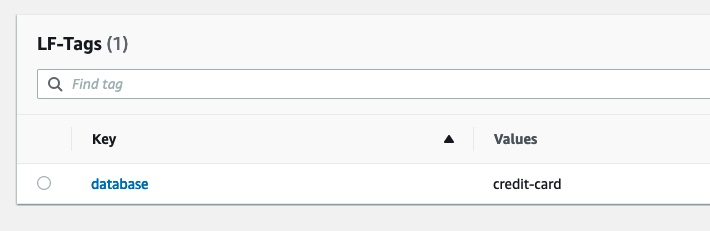
**Define LF-tags**

DataMeshOwner creates the tag ontology by defining LF-tags. Complete the following steps:

1. On the Lake Formation console, under **Permissions** in the navigation pane, under **Administrative roles and tasks**, choose **LF-Tags.**
2. Choose **Add LF-tags**.
3. For **Key**, enter database and for **Values**, choose credit-card.
4. Choose **Add** and then **Add LF-tag**.



1. The result looks like the image below.

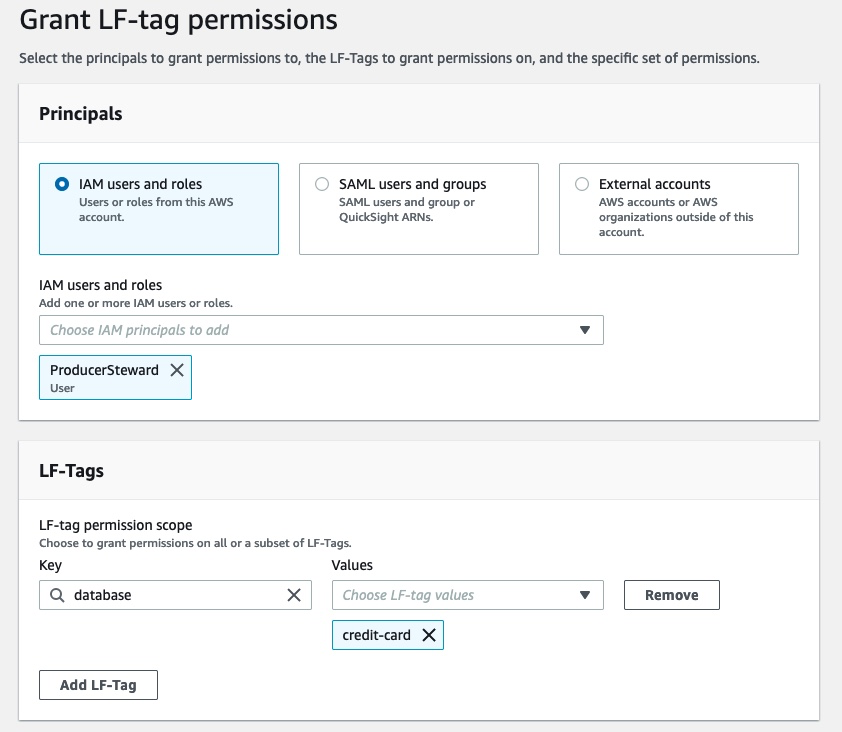


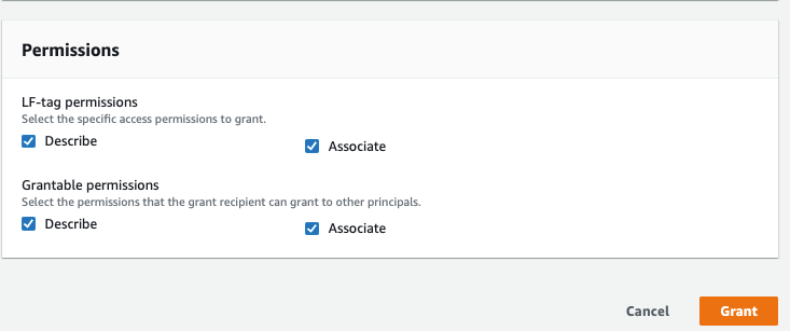
### Grant permissions

We grant ProducerSteward in the central accounts [describe and associate permissions](https://docs.aws.amazon.com/lake-formation/latest/dg/TBAC-security.html) on the preceding tag ontology. This enables ProducerSteward to view the LF-tags and assign them to Data Catalog resources (databases, tables, and columns). ProducerSteward in the central account can further grant the permission to ProducerSteward in the producer account. For more information, see [Granting, Revoking, and Listing LF-Tag Permissions.](https://docs.aws.amazon.com/lake-formation/latest/dg/TBAC-granting-tags.html) When you have multiple producers, grant the relevant tags to each steward.

In our situation, we will only have one LF tag assigned which points to the database. This could be further improved by adding extra tags for more granularity on data access. But it is out of the scope for this guide.

1. Under **Permissions** in the navigation pane, under **Administrative roles and tasks**, choose **LF-tag permissions**.
2. Choose **Grant**.
3. For **IAM users and roles**, choose the ProducerSteward user.
4. In the **LF-Tags** section, add all three key-values:
   1. Key database with values credit-card.



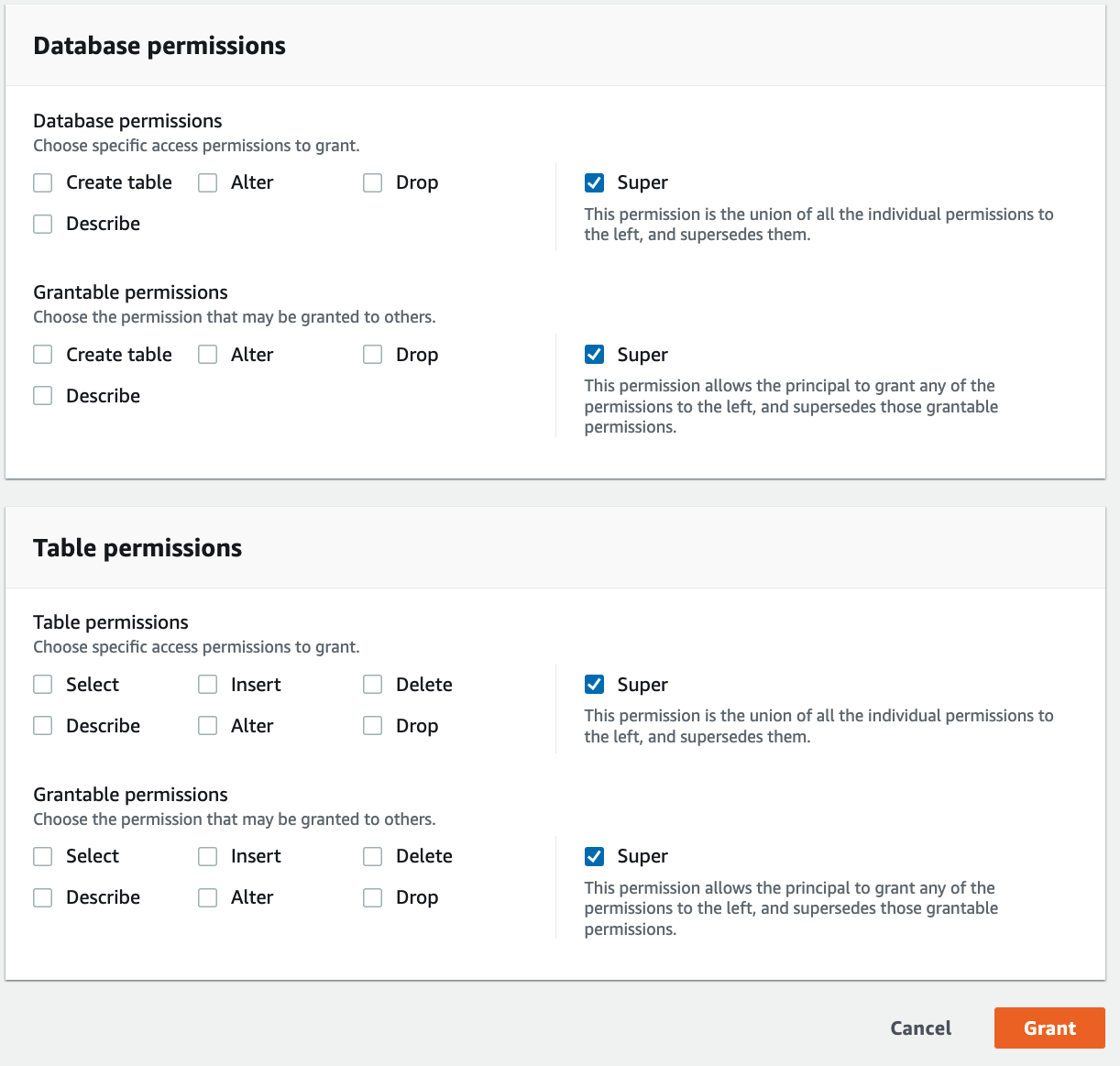
1. For **Permissions**, select **Describe** and **Associate** for both **LF-tag permissions** and **Grantable permissions**.
2. Choose **Grant.  
   [](https://d2908q01vomqb2.cloudfront.net/b6692ea5df920cad691c20319a6fffd7a4a766b8/2022/04/04/BDB-2091-image030.png)**

Next, we grant ProducerSteward tag-based data lake permissions. This enables ProducerSteward to create, alter, and drop tables in the databases with corresponding tags. ProducerSteward in the producer account can further grant the permission across accounts.

1. In the navigation pane, under **Permissions**, **Data lake permissions**, choose **Grant**.
2. For **Principals**, choose **IAM users and roles**, and choose ProducerSteward.
3. For **LF-tags or catalog resources**, select **Resources matched by LF-Tags (recommended)**.
4. Choose **Add LF-Tag**.
5. For **Key**, choose database and for **Values**, choose credit-card.
6. For **Database permissions**, select the [Super permission](https://docs.aws.amazon.com/lake-formation/latest/dg/lf-permissions-reference.html#perm-Super) because ProducerSteward owns the producer databases.

This permission allows a principal to perform every supported Lake Formation operation on the database. Use this admin permission when a principal is trusted with all operations.

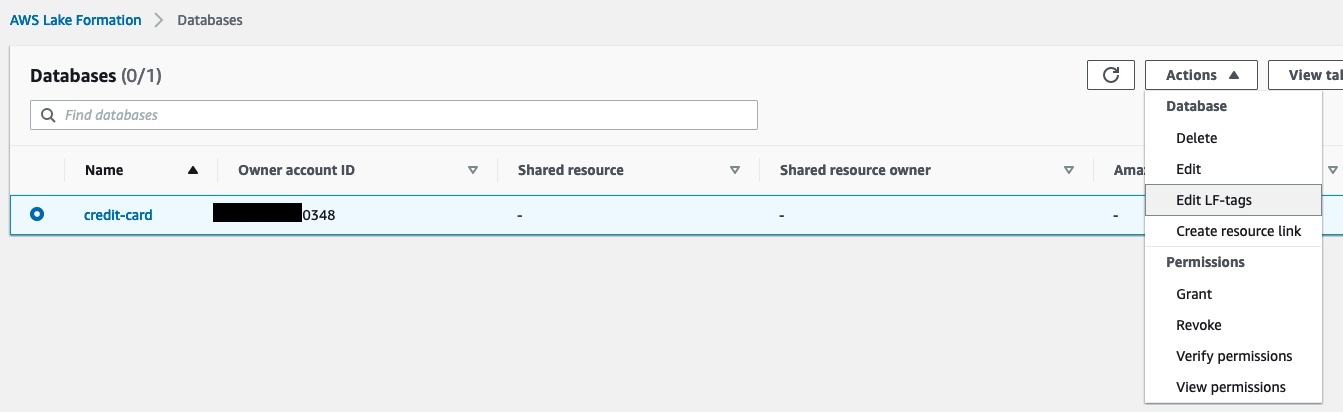
1. Select **Super** under **Grantable permissions** so the ProducerSteward user can grant database-level permissions to the producer and consumer accounts.
2. For **Table permissions**, select **Super**.
3. Select **Super** permission under **Grantable permissions**.
4. Choose **Grant**.

[](https://d2908q01vomqb2.cloudfront.net/b6692ea5df920cad691c20319a6fffd7a4a766b8/2022/04/04/BDB-2091-image032.png)

## Producer data steward actions in the central account

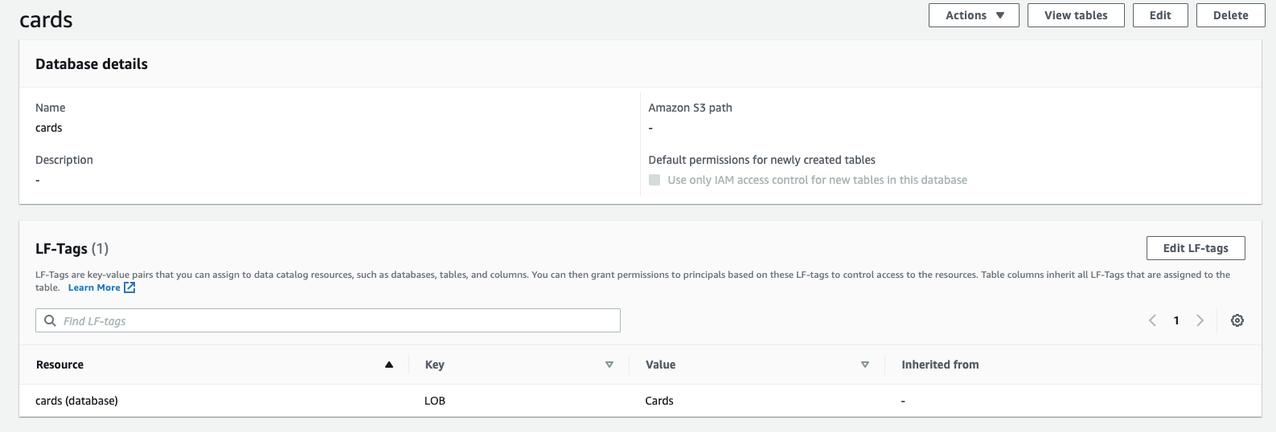
Next, we log in as the ProducerSteward user in the central account and create skeleton databases.

1. Sign in to the Lake Formation console as ProducerSteward.
2. In the navigation pane, under **Data catalog,** select **Databases**.
3. Choose the credit-card database.
4. On the **Actions** menu, choose **Edit LF-tags**

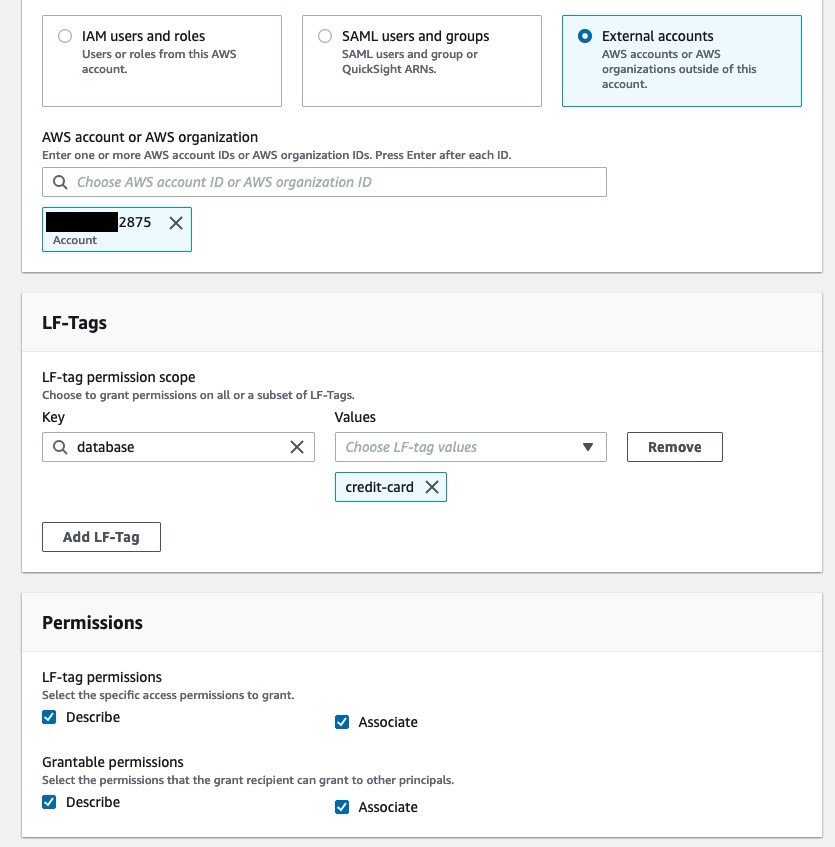
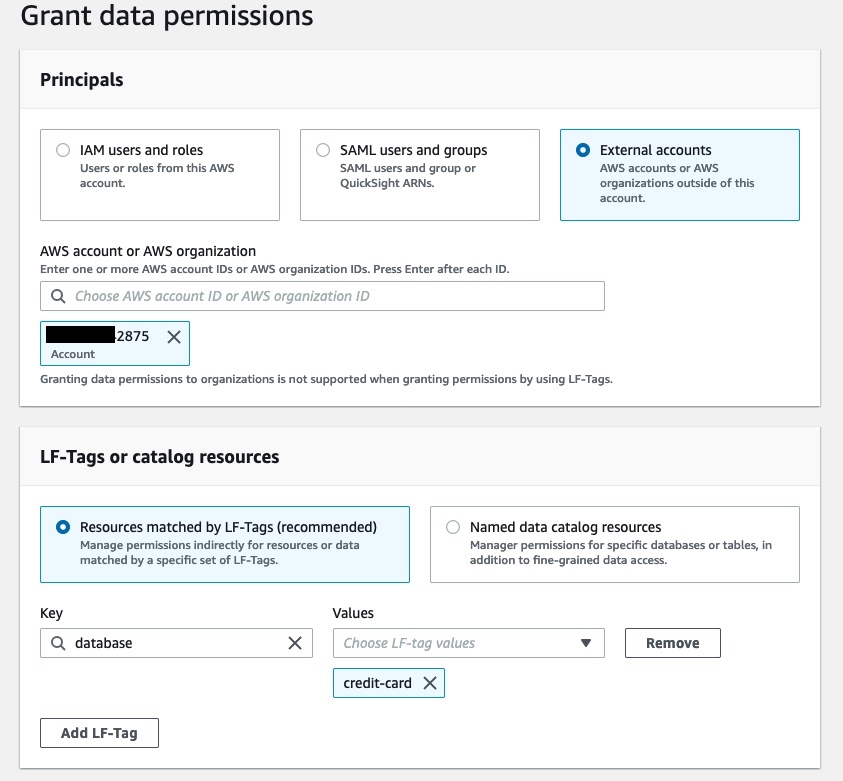
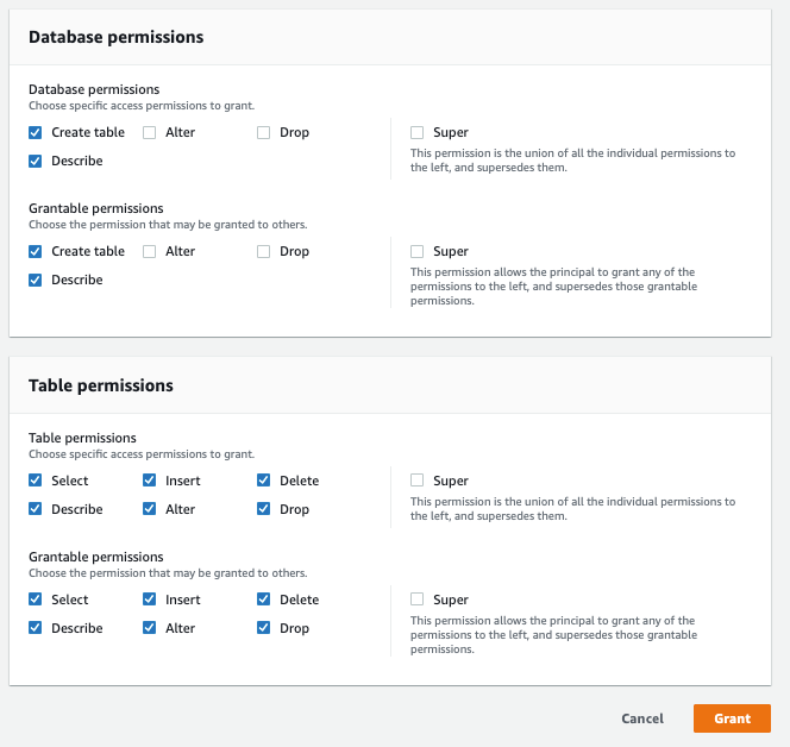


1. Choose **Assign new LF-tag**.
2. For **Assigned** **Keys**, enter database and for **Values**, choose credit-card.
3. Choose **Save.**

This assigns the database=credit-card tag to the credit-card database.

[](https://d2908q01vomqb2.cloudfront.net/b6692ea5df920cad691c20319a6fffd7a4a766b8/2022/04/04/BDB-2091-image046.png)

Next, we share the LF-tags and data lake permissions with the producer account so that ProducerSteward in the producer account can run AWS Glue crawlers and generate tables in the preceding skeleton databases.

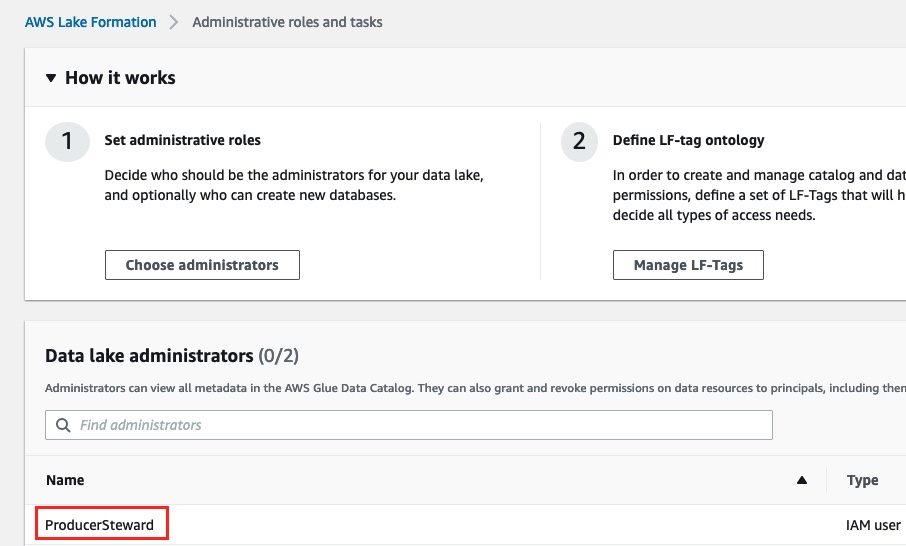
1. Under **Permissions** in the navigation pane, under **Administrative roles and tasks**, choose **LF-tag permissions**.
2. Choose **Grant**.
3. For **Principals**, select **External accounts**.
4. For **AWS account or AWS organization**, enter the account ID for the producer account.
5. In the **LF-Tags** section, we only need to add database-level tags.
6. For **Key**, enter database and for **Values**, choose credit-card.
7. For **Permissions**, choose **Describe** and **Associate** for both **LF-tag permissions** and **Grantable permissions**.
8. Choose **Grant**.  
   
9. In the navigation pane, under **Permissions**, **Data lake permissions**, choose **Grant**.
10. For **Principals**, select **External accounts**.
11. For **AWS account or AWS organization**, enter the account ID for the producer account.
12. For **LF-tags or catalog resources**, select **Resources matched by LF-Tags (recommended)**.
13. Choose **Add LF-Tag**.
14. Choose the key database and value credit-card.  
    
15. For **Database permissions**, select **Create table** and **Describe** because the ProducerSteward user in the producer account will add tables in the database.
16. Select **Create table** and **Describe** under **Grantable permissions** so the ProducerSteward user can further grant the permission to the AWS Glue crawler.
17. For **Table permissions**, select all the permissions.
18. Select all the permissions under **Grantable permissions.**
19. Choose **Grant**.  
    [](https://d2908q01vomqb2.cloudfront.net/b6692ea5df920cad691c20319a6fffd7a4a766b8/2022/04/04/BDB-2091-image056.png)

Now the Lake Formation administrators on the producer account side has the right permissions to add tables.

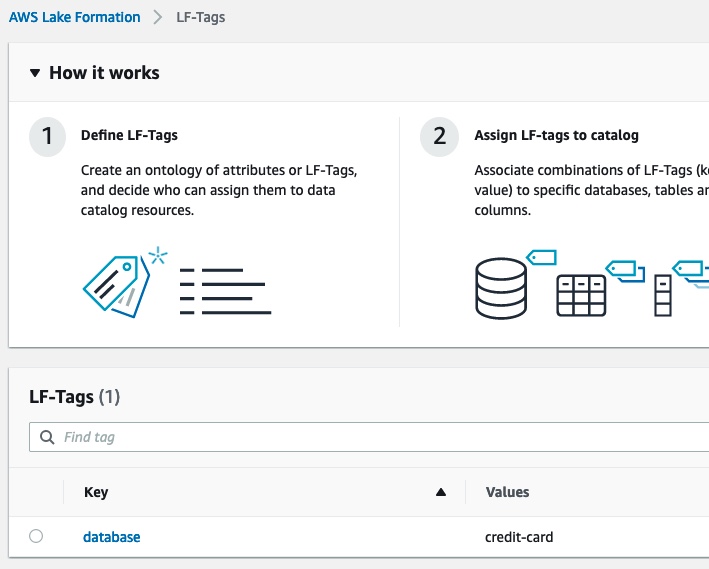
### Crawl source tables in the producer account

Next, we log in as the ProducerSteward user in the producer account to crawl the source tables for the Cards and Retail databases.

1. Sign in to the Lake Formation console as ProducerSteward.
2. In the navigation pane, under **Administrative Roles and Tasks**, verify that ProducerSteward is configured as the data lake administrator.

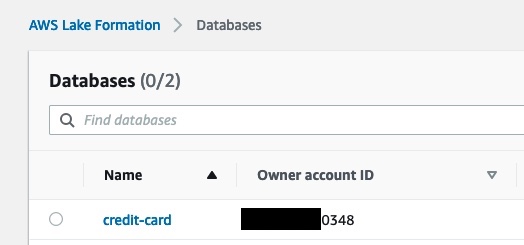


1. In the navigation pane, under **Permissions**, then choose **Administrative roles and tasks**, choose **LF-Tags**.

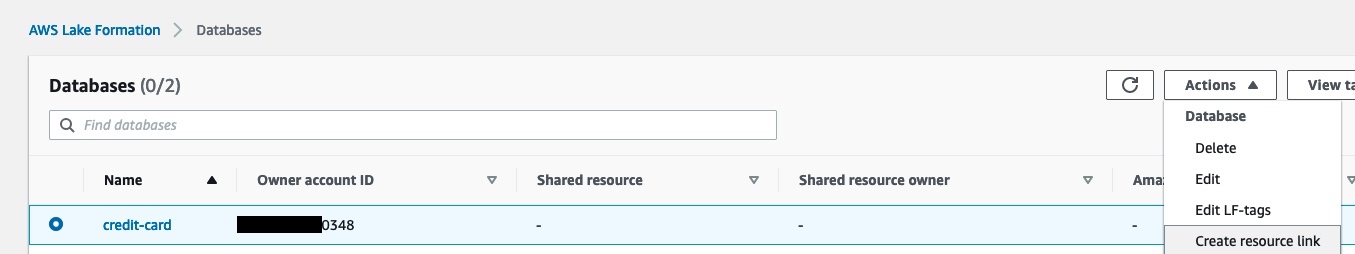
You can verify the database tag that was shared with the producer account.  


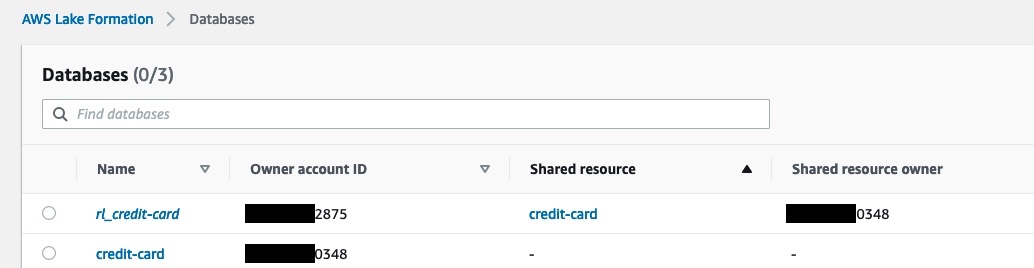
1. In the navigation pane, under **Data catalog**, select **Databases.**

You can verify the two databases cards and retail that were shared with the producer account from the previous step.



Now, we create a [resource link](https://docs.aws.amazon.com/lake-formation/latest/dg/resource-links-about.html) in the producer account for this database. This link point at the shared database and is used by AWS Glue crawler to create the tables. First, we create a resource link for the credit-card database.

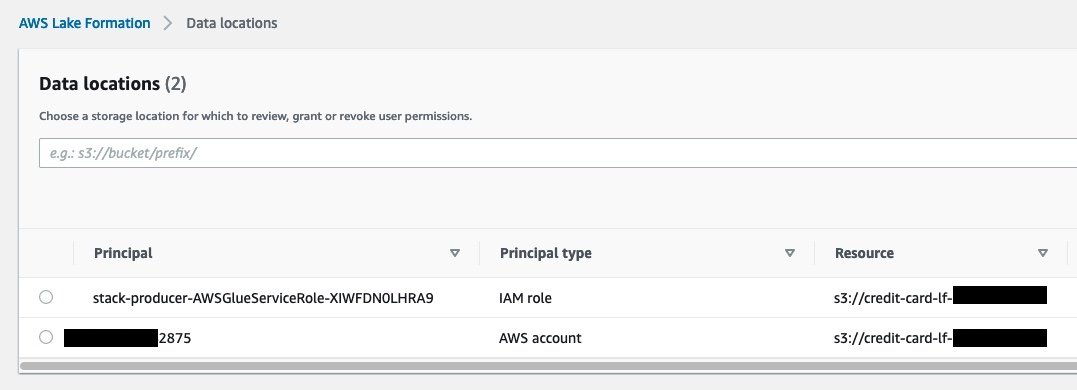
1. Select the cards database and on the **Actions** menu, choose **Create resource link**.  
   
2. For **Resource link name**, enter rl\_credit-card.
3. Choose **Create**.

After the resource link creation, you should see both the resource link databases as shown in the following screenshot.  


Next, we need to grant permissions to the AWS Glue crawler role so that the crawler can crawl the source bucket and create the tables.

1. Select the rl\_credit-card database and on the **Actions** menu, choose **Grant**.
2. In the **Grant data permissions** section, select **IAM users and roles**, and choose the AWS Glue crawler role that was created by the CloudFormation template (for example, stack-producer-AWSGlueServiceRoleDefault-xxxxxx).
3. For **Databases**, choose rl\_credit-card.
4. For **Resource link permissions**, select **Describe**.
5. Choose **Grant**.  
   [](https://d2908q01vomqb2.cloudfront.net/b6692ea5df920cad691c20319a6fffd7a4a766b8/2022/04/04/BDB-2091-image072.png)
6. Next, in the navigation pane, choose **Data lake Permissions** and choose **Grant**.
7. For **IAM users and roles**, choose the role stack-producer-AWSGlueServiceRoleDefault-XXXX.
8. For **LF-Tags or catalog resources**, select **Resources matched by LF-Tags**.
9. Enter the key database and values credit-card.
10. For **Database permissions**, select **Create table** and **Describe**.
11. For **Table permissions**, choose **Select**, **Describe**, and **Alter**.
12. Choose **Grant**.

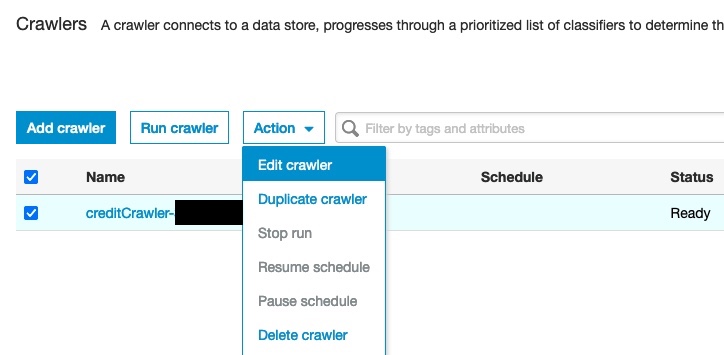
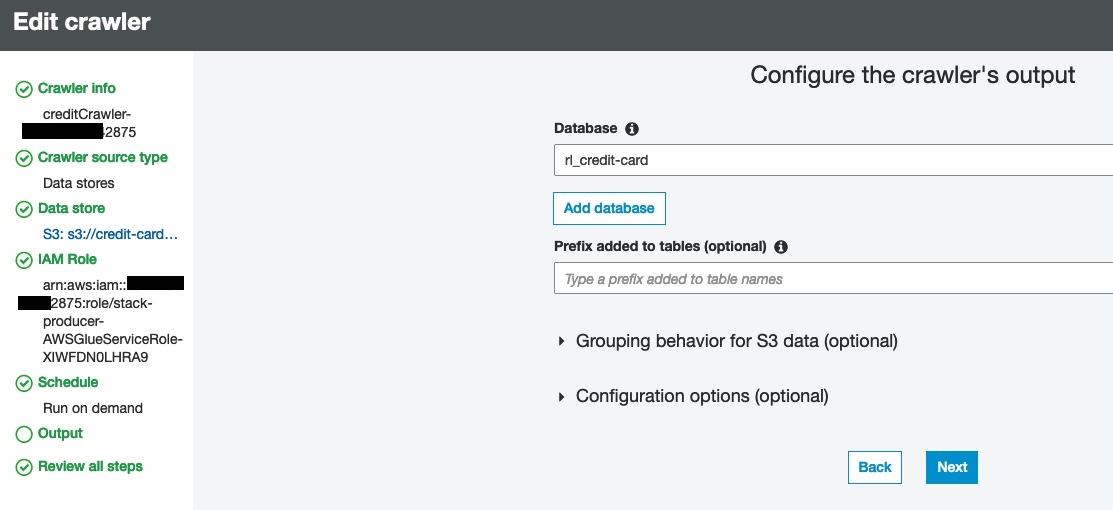
Next, we will verify grant permissions on the S3 bucket locations corresponding to credit-card producer to the AWS Glue crawler role. This is completed by the CloudFormation template.

In the navigation pane, under **Permissions**, on the **Data Locations**, you should see the locations.  


Now we’re ready to run the crawler. We configure the crawler that the CloudFormation template created, to point it to the resource link database.

1. On the AWS Glue console, under **Data catalog** in the navigation pane, choose **Crawlers**.

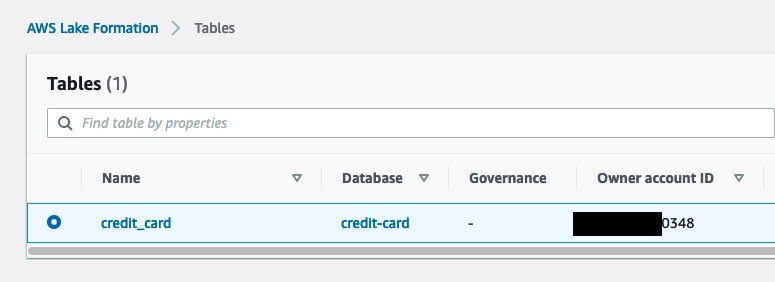
The crawler you created should be listed.

1. Select the crawler for the cards database CardsCrawler-xxxxxxxxxxxx and on the **Action** menu, choose **Edit crawler**.  
   
2. For the input data store, choose the S3 bucket for the credit-card producer.
3. For **IAM role**, choose the AWS Glue service role created by the CloudFormation template.
4. For **Schedule**, choose **Run on demand**.
5. For the output database, choose the resource link database rl\_credit-card corresponding to the cards database.  
   
6. Verify all the information and choose **Save**.
7. Select the crawler and choose **Run crawler**.

When the crawler finish, it creates table(s) corresponding to the producer in its respective resource link database. The table schemas are present in the shared database in the central account.

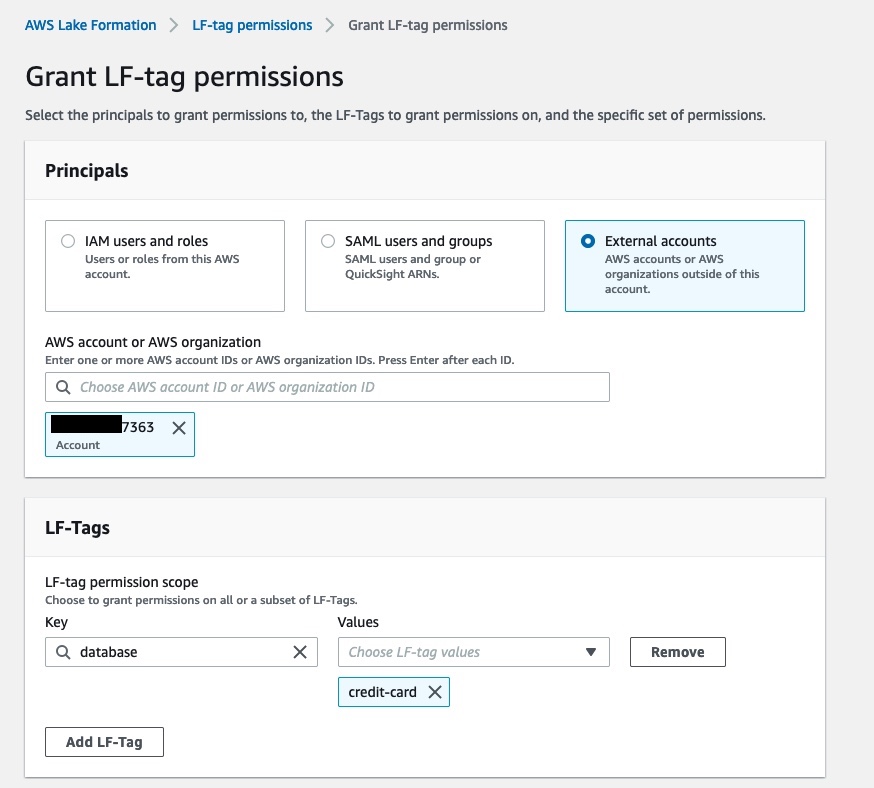
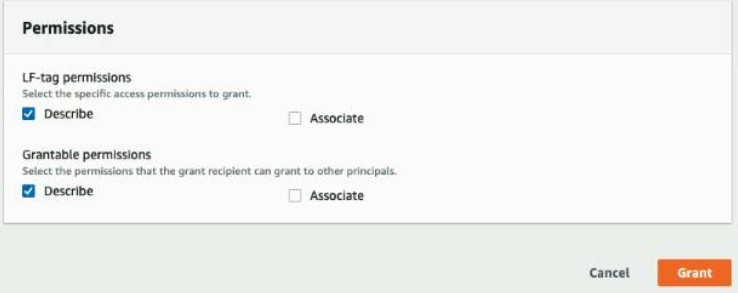
### Configure Lake Formation tags in the central account

1. Log in to central account as IAM user ProducerSteward.
2. On the Lake Formation console, in the navigation pane, choose **Data catalog** and then choose **Tables.**

You should see the credit\_card table corresponding to credit-card database.  


#### Grant tag permissions

Next, grant LF-tag permissions to the external consumer account.

1. On the Lake Formation console, in the navigation pane, choose **Permissions**, then choose **Administrative roles and tasks** and choose **LF-tag permissions**.
2. Choose **Grant**.
3. For **Principals**, select **External accounts**.
4. For **AWS account or AWS organization**, enter the AWS account number corresponding to the consumer account.
5. For **LF-Tags**, choose **Add LF-Tag**.
6. For **Key**, choose database and for **Values**, choose credit-card.  
   
7. For **Permissions**, choose **Describe**.
8. For **Grantable permissions**, choose **Describe**.
9. Choose **Grant**.  
   [](https://d2908q01vomqb2.cloudfront.net/b6692ea5df920cad691c20319a6fffd7a4a766b8/2022/04/04/BDB-2091-image096.png)

Next, we grant Lake Formation policy tag expression permissions to the external consumer account.

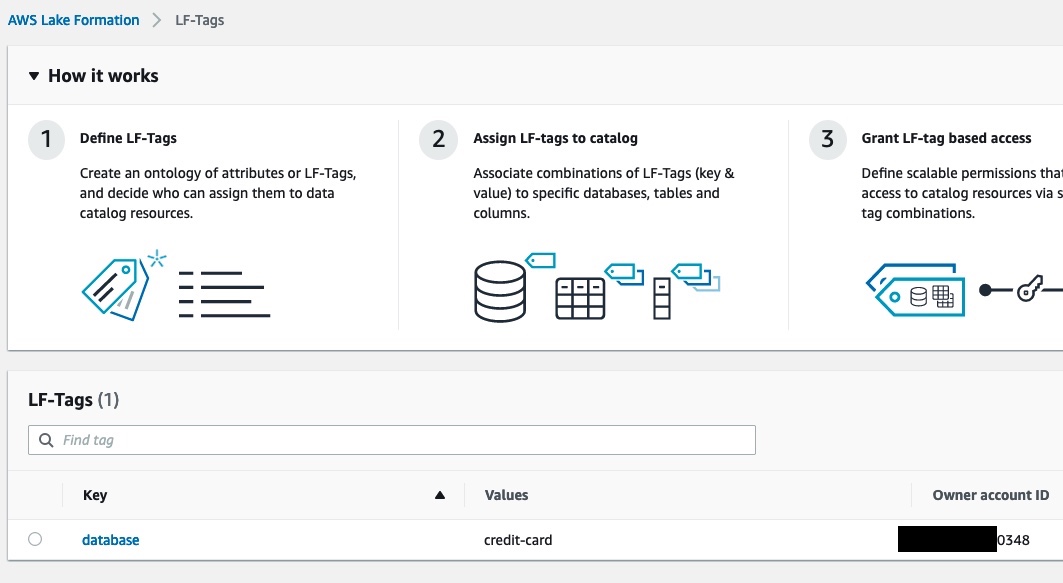
1. In the navigation pane, choose **Data lake permissions** and choose **Grant**.
2. In the **Principals** section, select **External accounts**.
3. For **AWS account or AWS organization**, enter the AWS account number corresponding to the consumer account.
4. For **LF-Tags or catalog resources**, select **Resources matched by LF-Tags**.
5. Choose **Add LF-Tag**.
6. For **Key**, choose database and for **Values**¸ choose credit-card.
7. For **Database permissions**, select **Describe**.
8. For **Grantable permissions**, select **Describe**.
9. Choose **Grant**.

Next, we grant table permissions.

1. In the navigation pane, choose **Data lake permissions** and choose **Grant**.
2. For **Principals**, select **External accounts**.
3. For **AWS account or AWS organization**, enter the AWS account number corresponding to the consumer account.
4. For **LF-Tags or catalog resources**, select **Resources matched by LF-Tags**.
5. Add key database with value credit-card
6. For **Table Permissions**, select **Select** and **Describe**.
7. For **Grantable permissions**, select **Select** and **Describe**.
8. Choose **Grant**.

## Share and consume tables in the consumer account

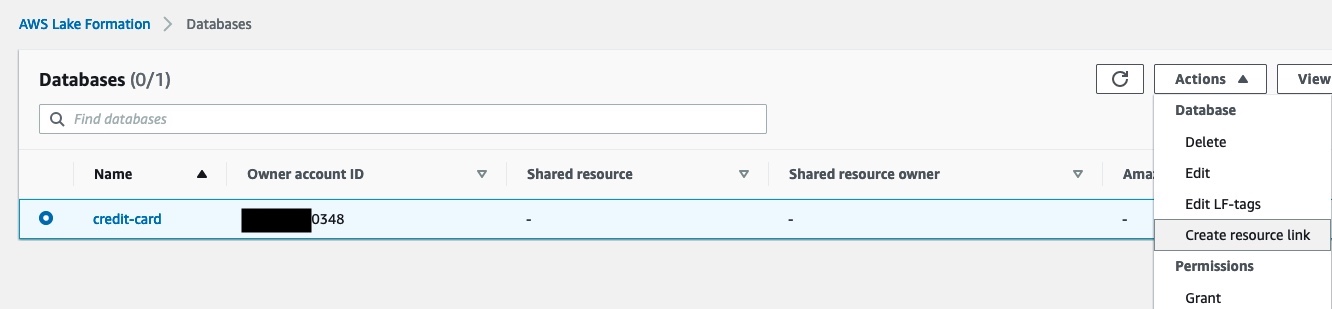
When you sign in to the Lake Formation console in the consumer account as ConsumerAdmin, you can see the tags and the corresponding value that was shared by the producer.



In these next steps, we share and consume the table in the consumer account.

### Create a resource link to the shared database

On the **Databases** page on the Lake Formation console, you can see all the databases that were shared to the consumer account. To create a resource link, complete the following steps:

1. On the **Databases** page, select the credit-card database and on the **Actions** menu, choose **Create resource link**.  
   
2. Enter the resource link name as rl\_credit-card.
3. Leave the shared database and shared database’s owner ID as default.
4. Choose **Create**.

### Grant Describe permission to SageMaker (SM) role used by the SageMaker Studio user

To grant Describe permissions on resource link databases to SM Studio user, complete the following steps:

1. On the **Databases** page, select the resource database rl\_credit-card and on the **Actions** menu, choose **Grant**.
2. In the **Grant data permissions** section, select **IAM users and roles**.
3. Choose the role corresponding to the SageMaker Studio user.
4. In the **Resource link permissions** section, select **Describe**.
5. Choose **Grant**.  
   Grant Tag permissions to ConsumerAnalyst1

To grant Tag permissions on the database:credit-card tag to SM Studio user to access the credit\_card table, complete the following steps:

1. On the Lake Formation console, on the **Data permission** page, choose **Grant**.
2. In the **Grant data permissions** section, select **IAM users and roles**.
3. Choose the role corresponding to the SageMaker Studio user
4. For **LF-Tags or catalog resources**, select **Resources matched by LF-Tags**.
5. Add the key database with value credit-card
6. For **Table permissions**, select **Select** and **Describe**.
7. Choose **Grant**.

The last step for the SageMaker Studio user to be able to access the data is to update the Studio user role by adding the following policy to the SageMaker role used by the Studio user

{

"Version": "2012-10-17",

"Statement": [

{

"Sid": "VisualEditor0",

"Effect": "Allow",

"Action": [

"lakeformation:GetDataAccess",

"glue:GetDatabase"

],

"Resource": "\*"

}

]

}