AWS re:Invent

ANT205-R

What's New with Amazon Athena

Sumeet Maru
Software Development Manager
Amazon Athena

Janak Agarwal
Product Manager
Amazon Athena

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws

Agenda

Abstract: As customers onboard more users to their data lakes, administrators want to balance flexibility with tighter control. In this chalk talk, we simplify authorization, authentication, security, and managing cost and auditing with Amazon Athena.

Athena Workgroups for cost controls

Athena integration with AWS Lake Formation

Athena VPCE

Related breakouts

[ANT222-R] Analytics with Amazon Athena (Workshop)

[ANT225] Access control for Amazon Athena with AWS Lake Formation

[ANT307-R] Deep dive into Amazon Athena

Cost control and audit mechanism

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws

Athena Workgroups

Workload Separation

- Permission Control
  - Workgroup is an Athena resource that supports IAM policies
  - Examples: <Specify whether User X can create Workgroups, list Workgroups, or can query using Workgroup A>
  - Query history isolation

- Overrides
  - Control where query output is stored for easy audit

- Cost Attribution
  - Activate User-Defined Cost Allocation Tags
  - Check workload’s Athena cost in AWS Billing and Cost Management Dashboard

Athena Workgroups

Cost control mechanism

- Query level controls
  - Automatically cancel query if limit reached
  - Publish metrics to Amazon CloudWatch

- Workgroup level controls
  - Configure S3 Requester Pays setting
  - Data-scan limit via CloudWatch and SNS
  - Aggregated metrics in CloudWatch

Demo

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Athena integration with Lake Formation for authorization, authentication, and security

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws

What is the problem

Permission management

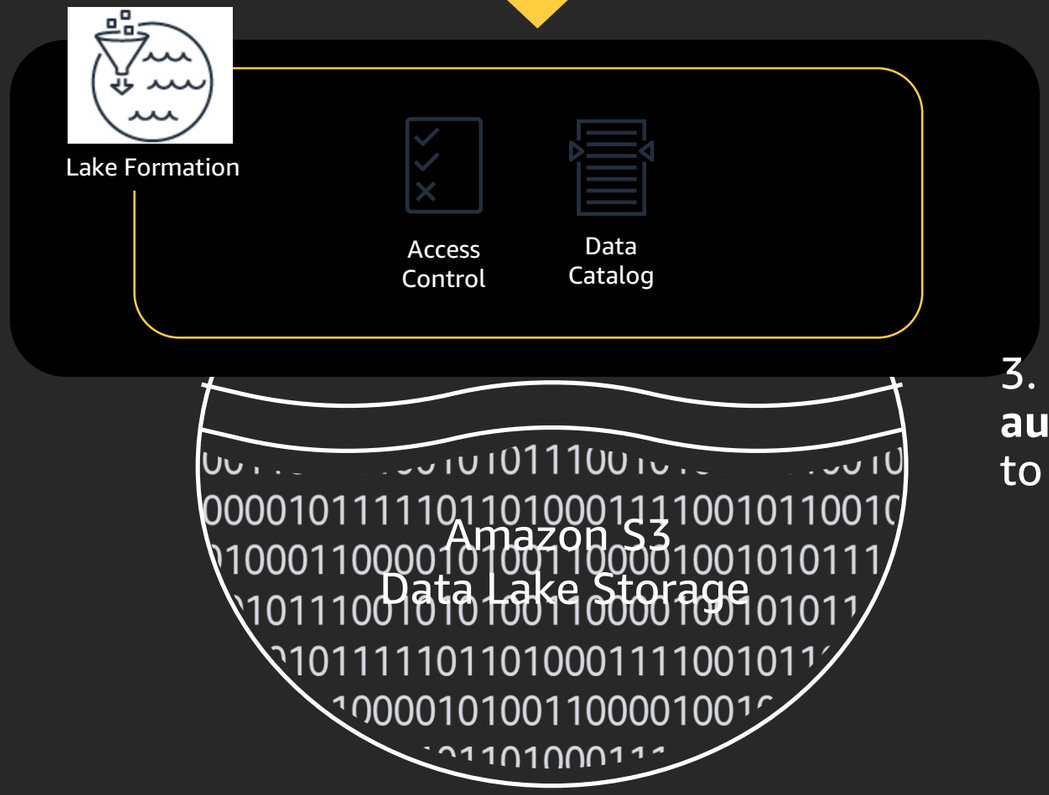
- Control access
  - Today, you need to apply access policies to 3 separate services – S3, AWS Glue, and Athena
  - Duplicate data by creating new tables with limited columns
  - Only IAM role supported in policies. Not AD Identity

- Audit access
  - Which user accessed what catalog, database, S3 path, table is not easy to get

Lake Formation: centralized authorization



1. Administrator sets up user permissions on lake resources: databases, tables, and columns



3. Lake Formation **authorizes** access to resources





2. User access data via integrated services

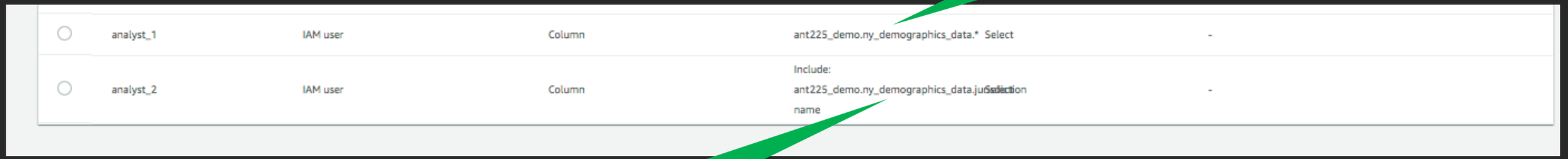
Easier security controls via Lake Formation

Permission management

- Control access
  - With Lake Formation integration, centralize access controls, no need to set controls on 3 separate services
  - Specify column level access controls
  - AD Identity support coming soon …

- Real-time audit and monitoring
  - Use Lake Formation APIs or Lake Formation console to easily audit
    - Who granted/revoked access to a resource
    - What did the query access (S3 location, user ARN)

Grant users access to table data

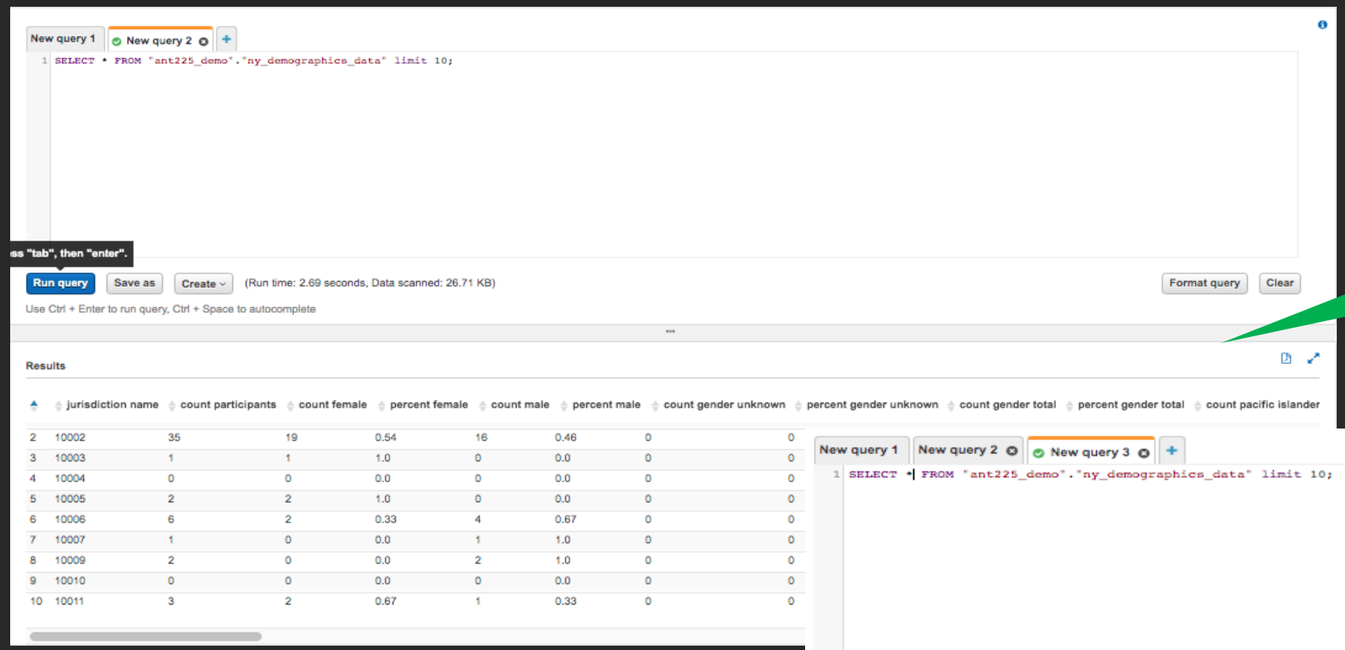


○	analyst_1	IAM user	Column	ant225_demo.ny_demographics_data.* Select	-
○	analyst_2	IAM user	Column	Include: ant225_demo.ny_demographics_data.jurisdiction name	-

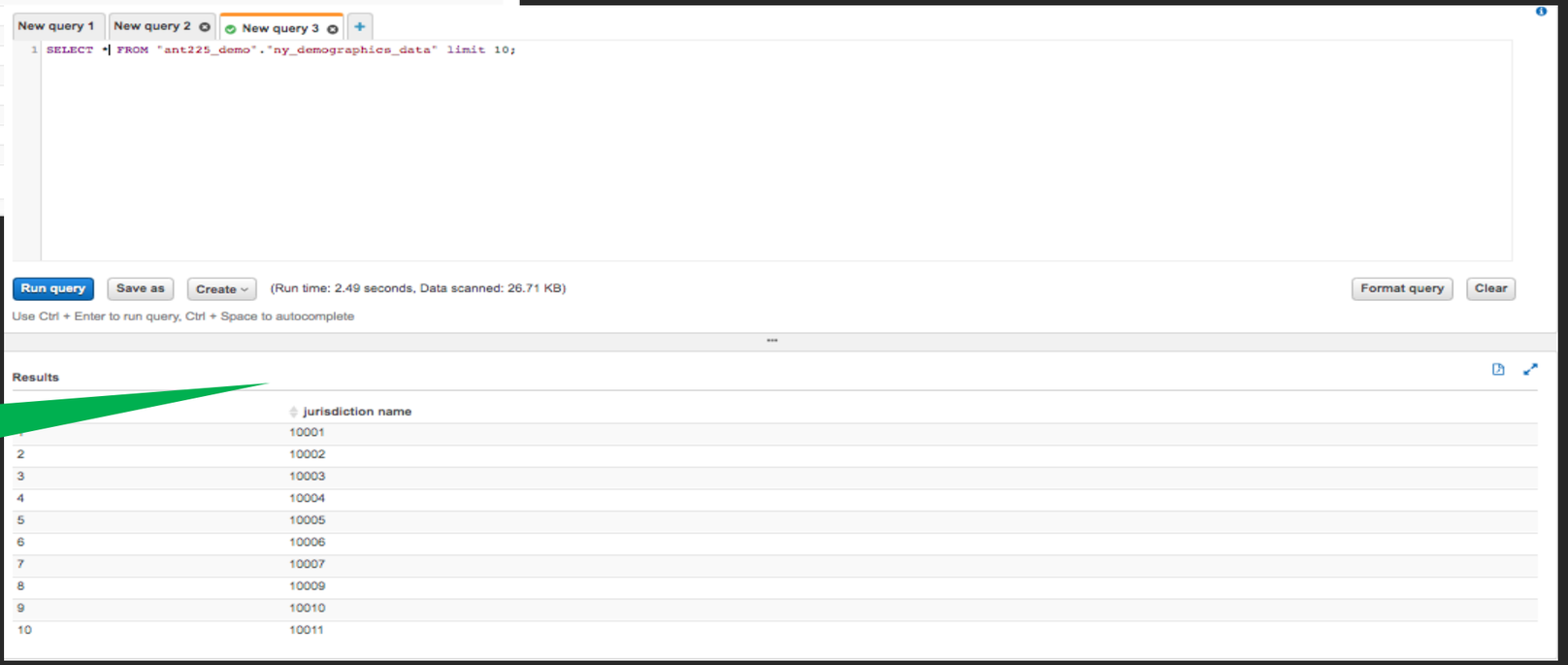
analyst_1 is granted access to all columns in Lake Formation

analyst_2 is granted access to only jurisdiction name column

Query the data using Amazon Athena



analyst_1 can view and query all columns



analyst_2 can view and query only ‘jurisdiction name’ column

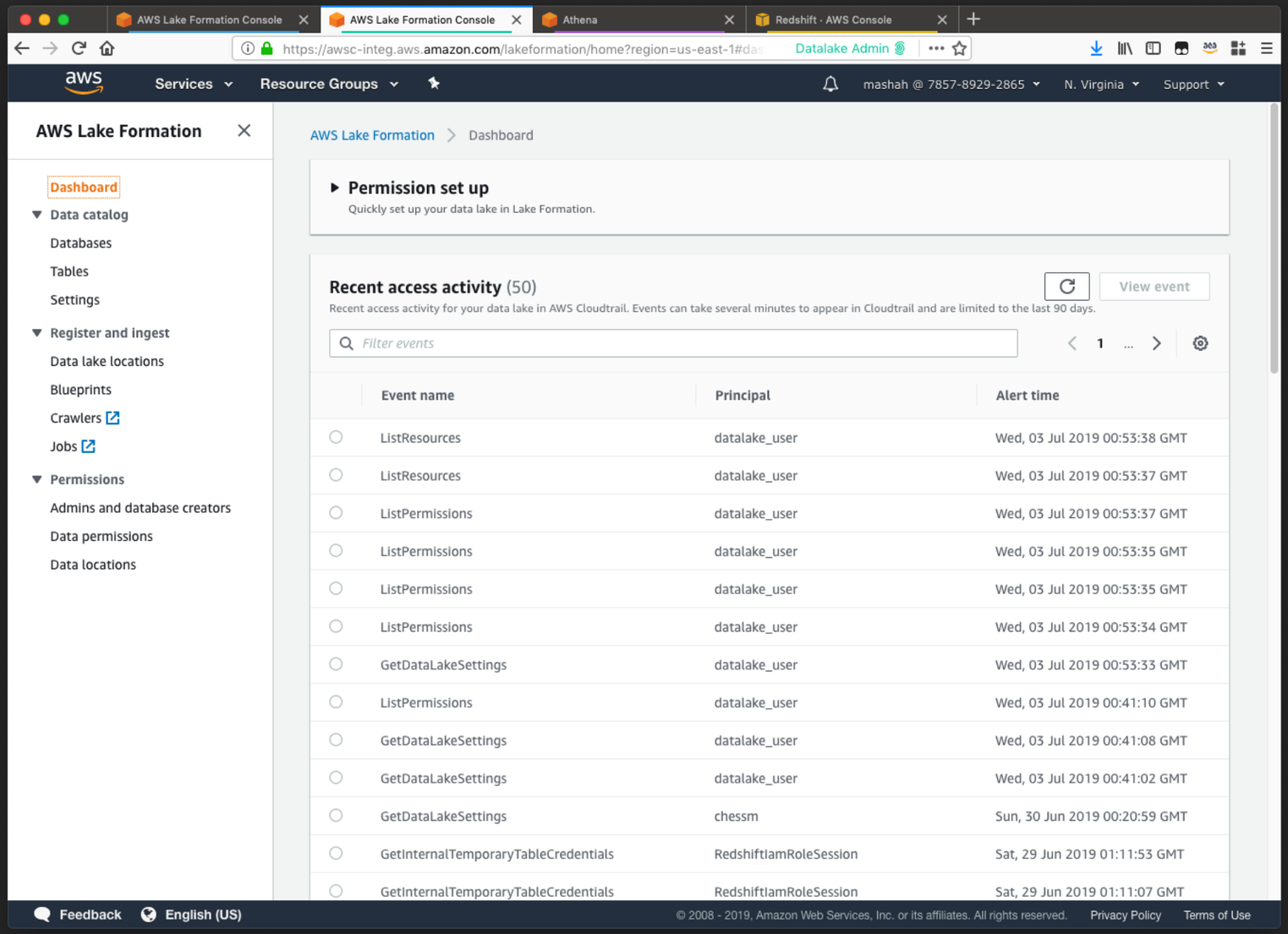
Audit and monitor in real time



See detailed alerts in the console

Download audit logs for further analytics

Data ingest and catalog notifications also published to Amazon CloudWatch events



Athena support for interface endpoint (PrivateLink)

Submit queries securely

- No internet gateway required in your VPC
  - Secure communication between your VPC and Athena APIs
- Set VPC endpoint policies
  - Example endpoint policy

```
{
  "Statement": [{
    "Principal": "*",
    "Effect": "Allow",
    "Action": [
      "athena:StartQueryExecution",
      "athena:RunQuery",
      "athena:GetQueryExecution",
      "athena:GetQueryResults",
      "athena:CancelQueryExecution",
      "athena:ListWorkGroups",
      "athena:GetWorkGroup",
      "athena:TagResource"
    ],
    "Resource": [
      "arn:aws:athena:us-west-1:AWSAccountId:workgroup/workgroupA"
    ]
  }]
}
```

Resources

Athena Workgroups

https://docs.aws.amazon.com/athena/latest/ug/manage-queries-control-costs-with-workgroups.html

Athena integration with Lake Formation

https://docs.aws.amazon.com/athena/latest/ug/lf-athena-access.html

Athena infrastructure security and (PrivateLink) support

https://docs.aws.amazon.com/athena/latest/ug/infrastructure-security.html

Thank you!

Sumeet Maru

linkedin.com/in/sumeet-maru-44ba925

Janak Agarwal

linkedin.com/in/janakagarwal
@JanakAgarwal

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws



Please complete the session survey in the mobile app.

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.