Follow the CloudTrail

Learn what happens in your AWS account when you're not watching

Keith Gregory
AWS Practice Lead, Chariot Solutions

CHARIOT SOLUTIONS

It's the Access Log for AWS

All AWS tasks are performed via GET, PUT, or POST

CloudTrail captures most (but not all!) API calls, with request and response bodies

Two types of events:

- Management events: anything that creates/updates/deletes service
- Data events: S3 operations, Lambda invocations

CloudTrail Insights is a layer on top of these events

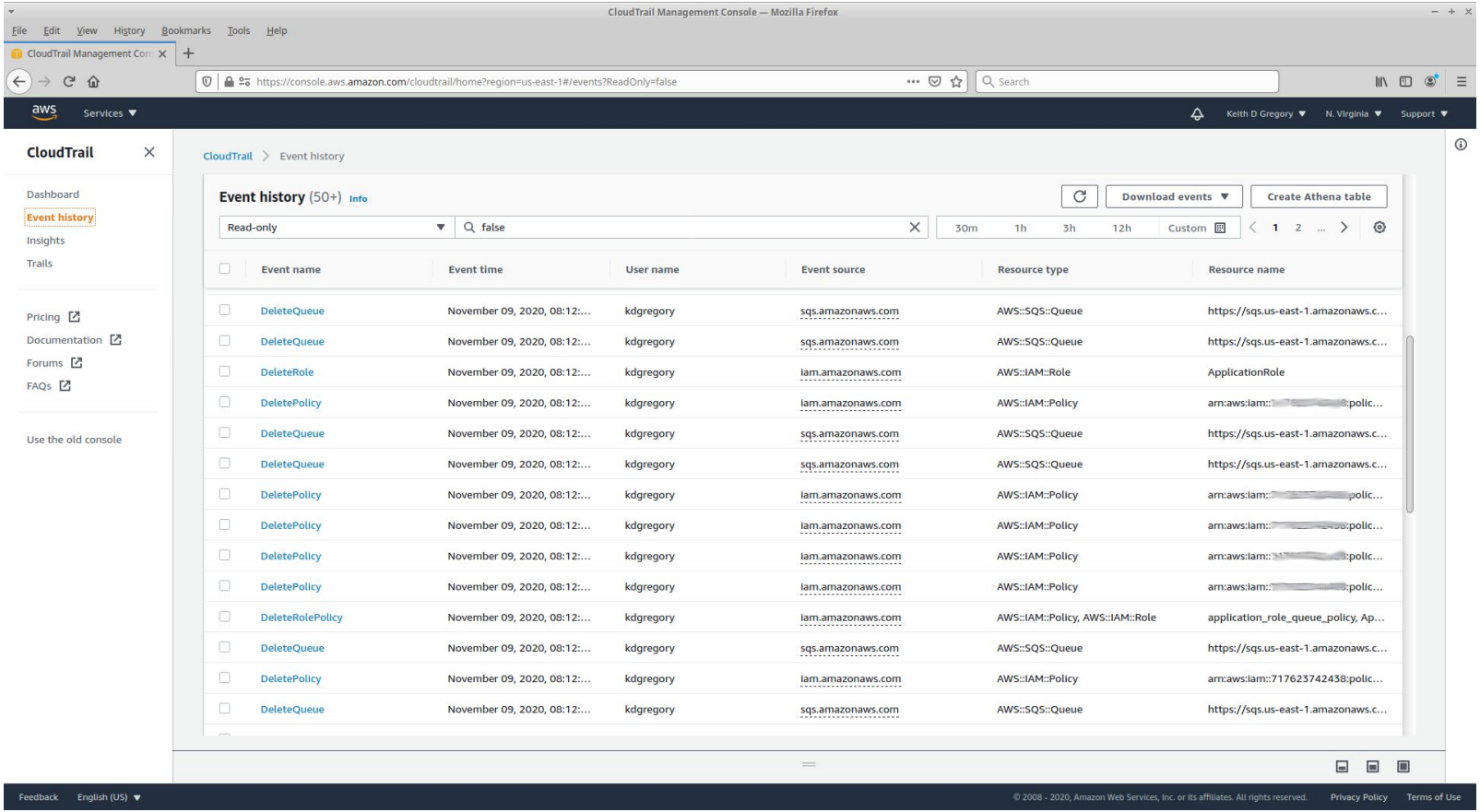
CHARIOT SOLUTIONS

Example Event


```
{
    "eventVersion": "1.05",
    "userIdentity": {
        "type": "IAMUser",
        "principalId": "REDACTED",
        "arn": "arn:aws:iam::123456789012:user/kdgregory",
        "accountId": "123456789012",
        "accessKeyId": "REDACTED",
        "userName": "kdgregory"
    },
    "eventTime": "2020-11-09T13:12:00Z",
    "eventSource": "sqs.amazonaws.com",
    "eventName": "DeleteQueue",
    "awsRegion": "us-east-1",
    "sourceIPAddress": "1.2.3.4",
    "userAgent": "aws-sdk-go/1.35.19 (go1.14.5; linux; amd64) APN/1.0 HashiCorp/1.0 Terraform/0.12.29...",
    "requestParameters": {
        "queueUrl": "https://sqs.us-east-1.amazonaws.com/123456789012/ReportGeneration"
    },
    "responseElements": null,
    "requestID": "d274e87d-860c-54b1-b49d-b05ec46ceae6",
    "eventID": "0222214f-8887-460c-9958-47ca41a3177e",
    "eventType": "AwsApiCall",
    "recipientAccountId": "123456789012"
}
```


CHARIOT SOLUTIONS

Console View



Creating a Trail

First trail is free!*

* management events only

All events stored in S3

Limit access to this bucket: it may contain sensitive data

Use S3 Object Lock to prevent deletion (accidental or otherwise)

Can also send to CloudWatch Logs

Enable cross-region and cross-organization capture!

CHARIOT SOLUTIONS

How to Query Events

CloudTrail Console

Copy files from S3, use jq

Amazon Athena

CloudWatch Logs Insights

Elasticsearch / other search engines

CHARIOT SOLUTIONS

Problems

Why you don’t want to use the Console, JQ, or Athena

CHARIOT SOLUTIONS

#1: Every event is different

Some fields are common

eventTime, eventSource, eventName, sourceIPAddress, …

Most of the “interesting” information is event-specific

requestParameters: information the caller provides

responseElements: information about what AWS did

User identity changes depending on how API invoked

IAM user versus IAM role versus AWS internal

CHARIOT SOLUTIONS

#2: Who invoked that API?

- Most API calls are made under assumed role
  - Or at least should be: Lambda execution role, EC2 instance role, ...
  - If you explicitly assume a role in your program, use a good session ID!
- AWS services *usually* set `userIdentity.invokedBy`
  - Typically to the service name, sometimes to “AWS Internal”
- Some events are tied to others
  - Example: `SharedSnapshotVolumeCreated`, from EC2 `RunInstances`

CHARIOT SOLUTIONS

#3: This data is *noisy*!

- API calls are granular
  - Starting an EC2 instance from the Console involves a dozen calls
- The Console does a lot that you don't know about
- Internal AWS operations recorded alongside user actions
  - Load balancers regularly check their instances
  - Trusted Advisor, other monitoring services run regularly, cross regions
  - Lambda invocations may not be recorded, but their assume role and KMS decrypt operations are

CHARIOT SOLUTIONS

#4: There are a lot of files

Each file contains approximately 15 minutes of events

For one account, and one region

CHARIOT SOLUTIONS

#5: Not all API calls are tracked

- Lambda invocations tracked as “data events”
- Most services only track “infrequent” operations
  - Example: DynamoDB tracks `CreateTable` but not `Scan`
- Some services don’t expose a public API
  - Example: DeepLens

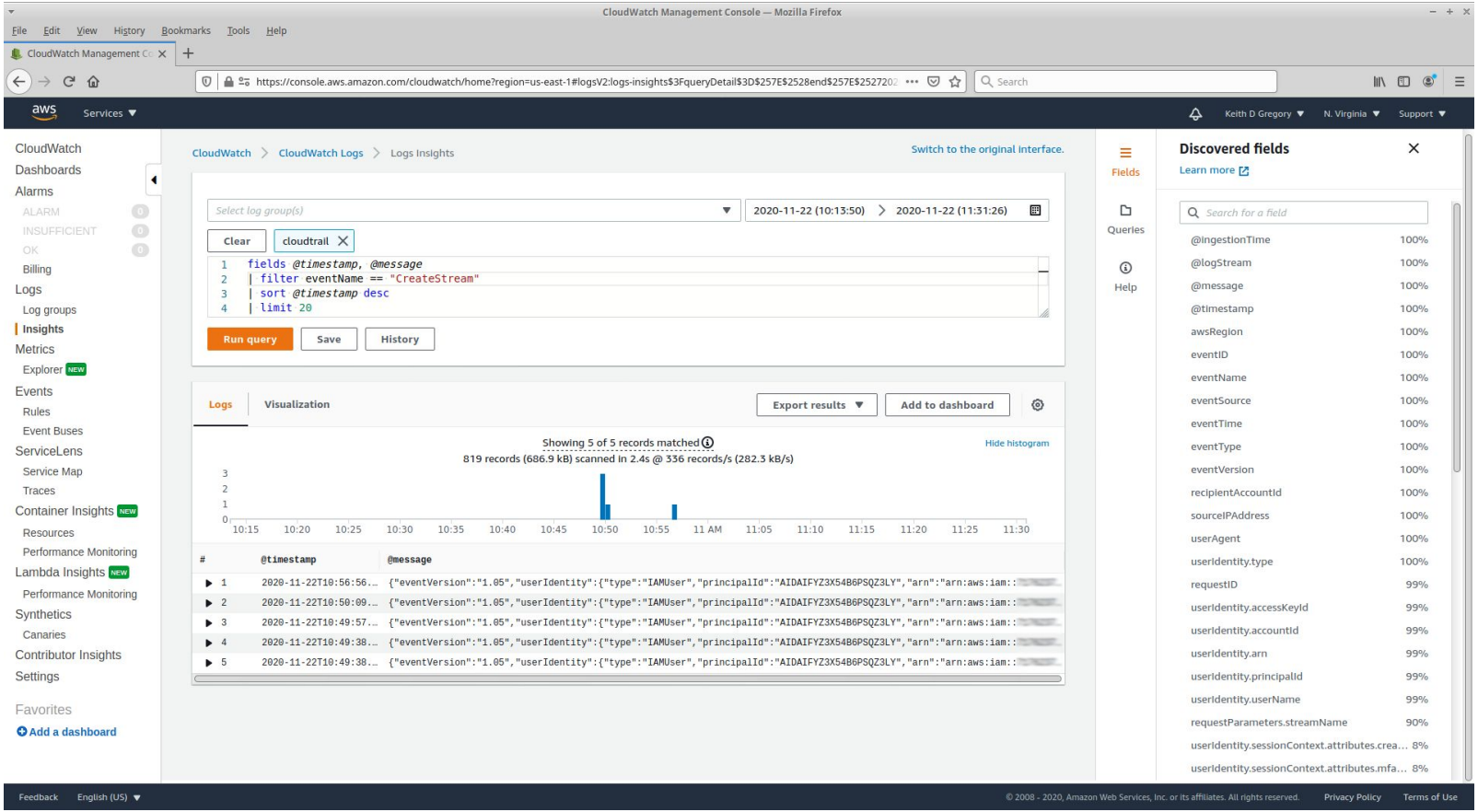
CHARIOT SOLUTIONS

Solutions

Search Engines to the rescue!

CHARIOT SOLUTIONS

CloudWatch Logs Insights



CloudWatch Logs Pros/Cons

Pro: set up with a few clicks

Pro: easy to enable log monitors, alerts

Con: every Insights search is a unique filter

- You need to know what the event looks like

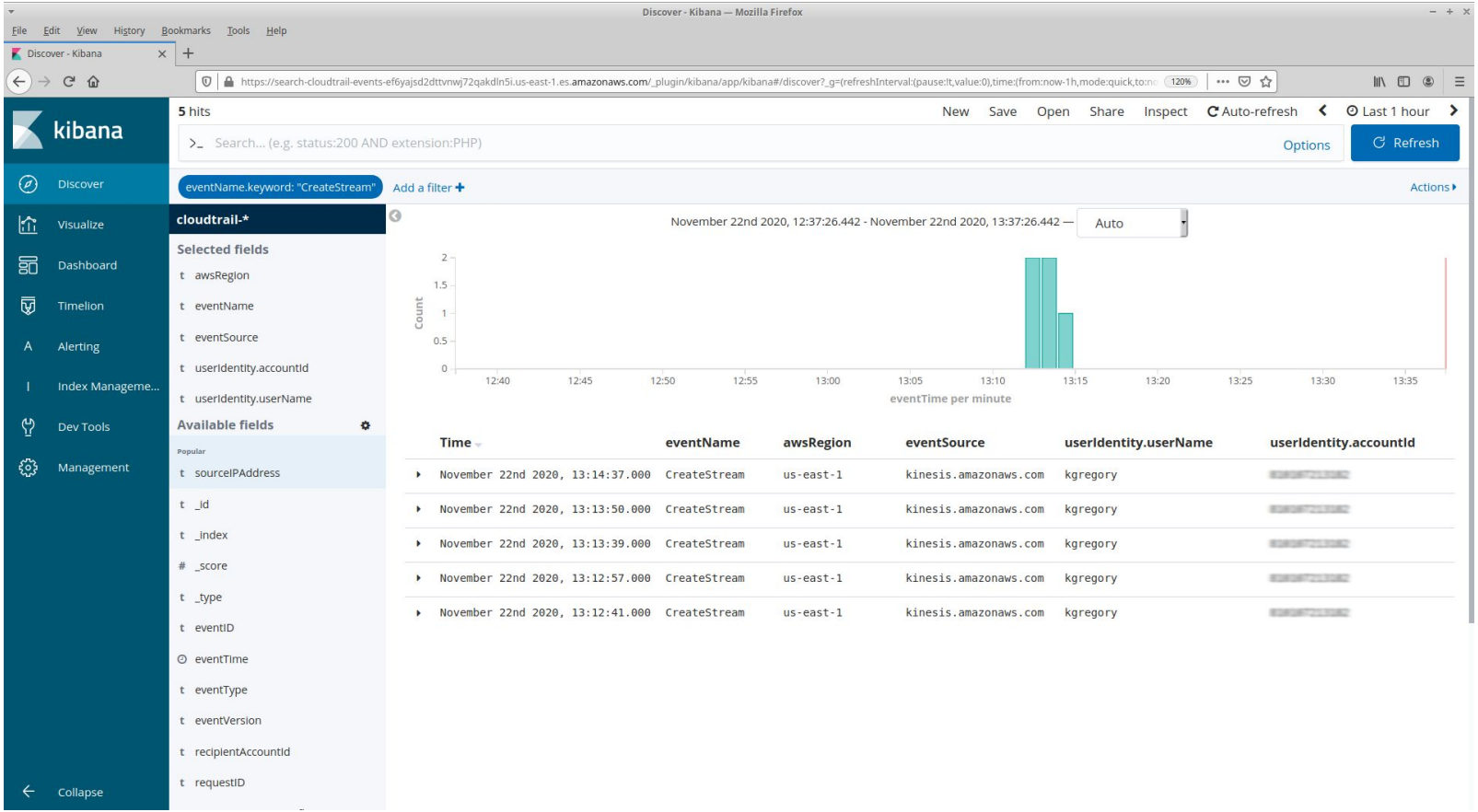
Con: you pay per search

- It's only $0.005 per GB of data scanned, but it adds up

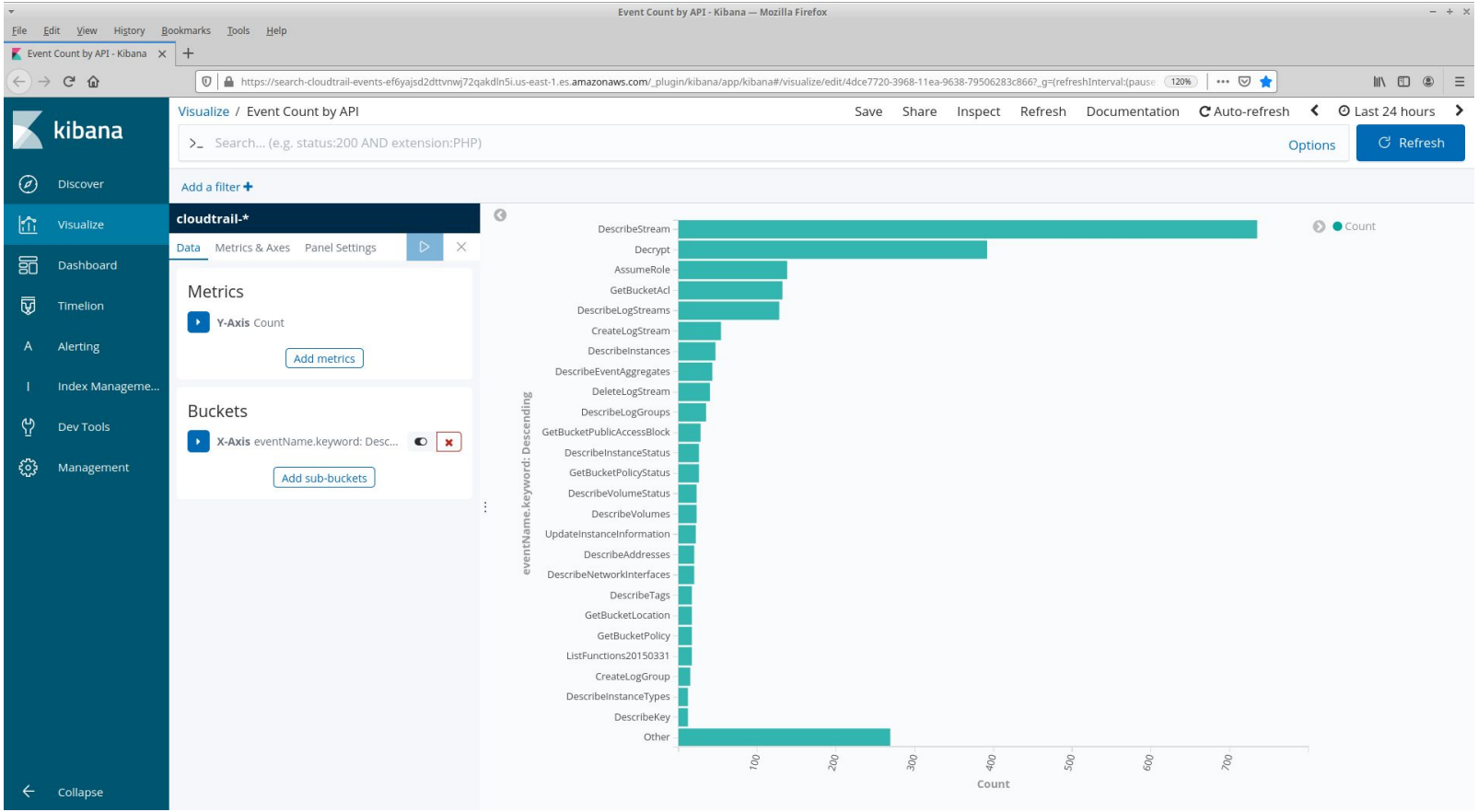
Con: it's slow

CHARIOT SOLUTIONS

Elasticsearch



Elasticsearch Visualizations



Elasticsearch Pros/Cons

Pro: extremely flexible

- Search by any value in the event (as long as it's indexed)
- Easily narrow searches by time range; filter undesired values

Pro: fast!

- Most searches can be based on keywords

Con: you have to configure event pipeline

- Elasticsearch doesn't like that every event is different!
- Logstash has a plugin to parse events

CHARIOT SOLUTIONS

Third-party Providers

The usual suspects: DataDog, Logz.io, Loggly, …

Easy to get started; flexible and fast searches

Cost-effectiveness depends on log volume

Use Cases

Including a few that aren’t supported

Who Started That EC2 Instance?

Event type: `RunInstances`

Filter by: `responseElements.instanceId`

Look at `userIdentity`

CHARIOT SOLUTIONS

Who Accessed That Bucket?

Must have data events enabled

Filter by `requestParameters.bucketName`

- Optional: filter by `eventName`

Look at `userIdentity`, `sourceIPAddress`

CHARIOT SOLUTIONS

Is There Anything To Worry About?

Visualize events by region, source IP address

Filter out AWS-generated events

Switch to search mode to explore unexpected results

CHARIOT SOLUTIONS

~~Is someone attacking me~~?

Events appear in the trail 10-20 minutes after they happen

Not all failed requests are captured

CHARIOT SOLUTIONS

~~What permissions does my app need~~?

Applications often use API calls that aren't captured

CHARIOT SOLUTIONS

~~What's doesn't my ABAC policy work~~?

User identity doesn't contain `aws:Principal` tags

CHARIOT SOLUTIONS

Wrapping Up

CHARIOT SOLUTIONS

You Can Trigger Lambdas

Via EventBridge (formerly CloudWatch Events)

Trigger format is … challenging

It's Not The Only Tool

Bucket access logs are cheaper alternative to data events

- Makes more sense if you care about public bucket accesses

Cost Explorer often has better view of "what's happening"

CHARIOT SOLUTIONS

Technology in the Service of Business.

Chariot Solutions is the Greater Philadelphia region's top IT consulting firm specializing in software development, systems integration, mobile application development and training.

Our team includes many of the top software architects in the area, with deep technical expertise, industry knowledge and a genuine passion for software development.

Visit us online at chariotsolutions.com.

CHARIOT SOLUTIONS