[Tip] Get AWS Account ID from CloudTrail logs with Sumo Logic #sumologic

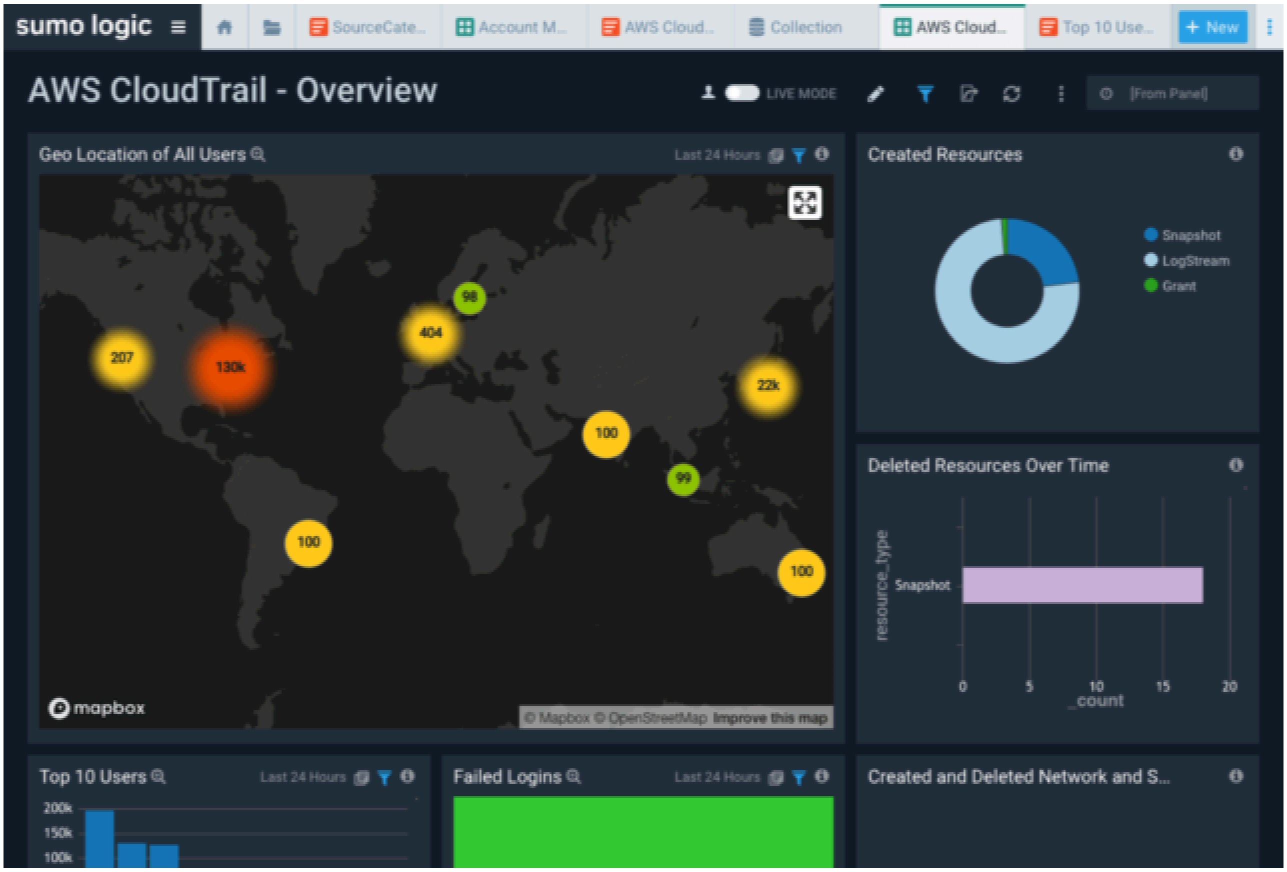
dev.classmethod.jp/articles/202005-sumologic-parse-aws-account-id-cloudtrail

渡辺聖剛

May 31, 2020

background

When monitoring CloudTrail with Sumo Logic, I sometimes find myself wondering, "Which AWS account did this log come from?"



This is because CloudTrail logs do not contain a field that indicates which AWS account ID the log occurred under.

`userIdentity.accountId`The information recorded as "User information" is merely "information about the user who performed the operation," so there may be many situations where this is not a problem in everyday life, but because the AWS system allows authority to be transferred to another AWS account, it is not perfect.

> How to Use External IDs When Granting Access to Your AWS Resources to Third Parties - AWS Identity and Access Management

In the first place, this information is not recorded in all logs, so it can be a bit difficult to use if you want to see everything.

So I thought about what to do and found a simple solution:
parsing the log file name (S3 object file name) would be good.

Commentary

As mentioned in the AWS documentation above, CloudTrail logs have object names in the following format:

```
AccountID_CloudTrail_RegionName_YYYYMMDDTHHmmZ_UniqueString.FileNameFormat
```

In reality, a default prefix will be added to this, so excluding the bucket name, it will look something like this:

```
prefix_name/AWSLogs/Account
ID/CloudTrail/region/YYYY/MM/DD/AccountID_CloudTrail_RegionName_YYYYMMDDTHHmm
Z_UniqueString.FileNameFormat.json.gz
```

> Searching CloudTrail Log Files - AWS CloudTrail

This file name can be obtained as in Sumo Logic `_sourceName`, so by parsing it you can get the AWS account name.

Built-in Metadata - Sumo Logic

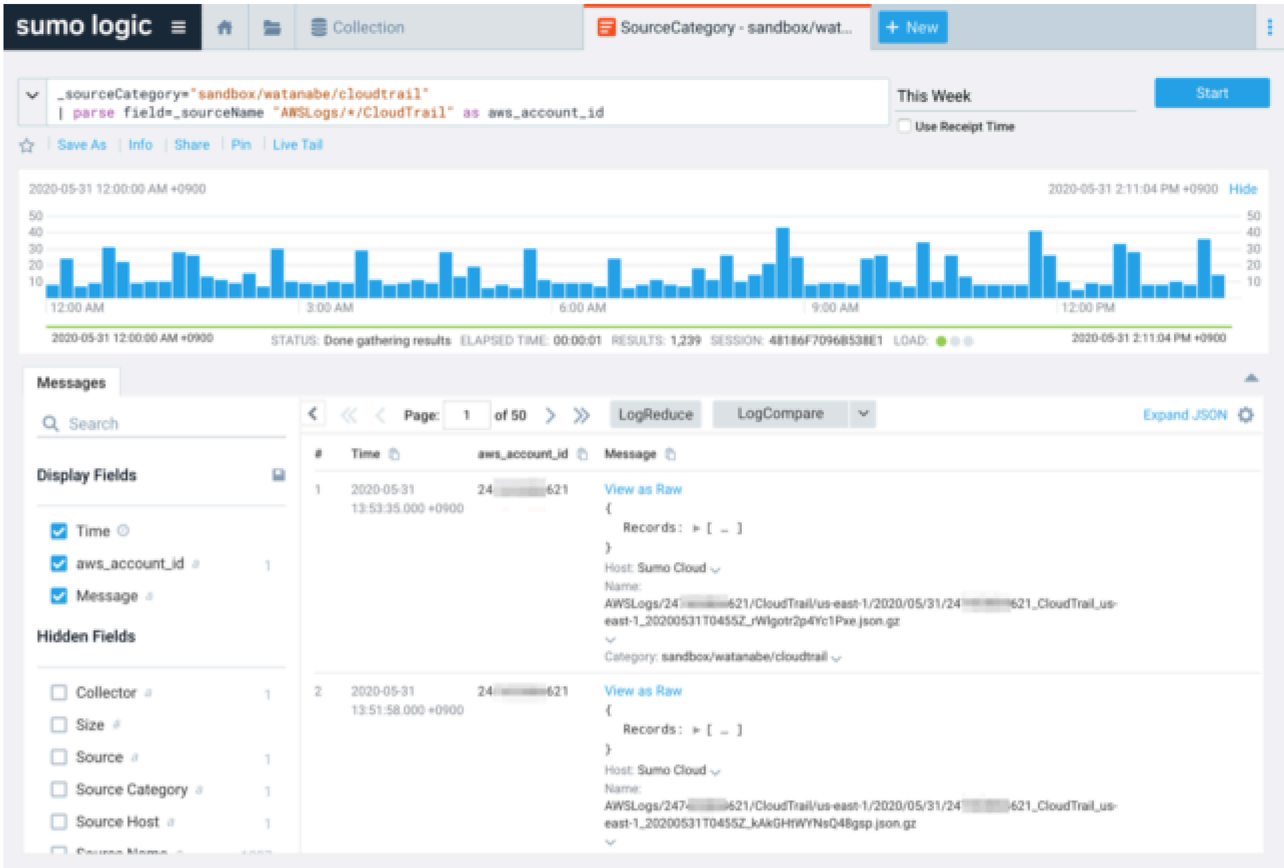
> `_sourceName`
>
> *The name of the log file, determined by the path you entered when you configured the Source.*

I think the easiest way to retrieve it `parse`is to use a phrase `_sourceName`and then use pattern matching.

```
| parse field=_sourceName "AWSLogs/*/CloudTrail" as aws_account_id
```

- Parse field option - Sumo Logic
- Parse field options - Sumo Logic

When I actually searched, I got something like this:



```
_sourceCategory="sandbox/watanabe/cloudtrail"
| parse field=_sourceName "AWSLogs/*/CloudTrail" as aws_account_id
```

This Week
□ Use Receipt Time

2020-05-31 12:00:00 AM +0900 — 2020-05-31 2:11:04 PM +0900

STATUS: Done gathering results ELAPSED TIME: 00:00:01 RESULTS: 1,239 SESSION: 48186F7096B538E1 LOAD: ●

Page: 1 of 50

Display Fields
☑ Time
☑ aws_account_id 1
☑ Message

Hidden Fields
□ Collector 1
□ Size
□ Source 1
□ Source Category 1
□ Source Host 1
□ Source Name

#	Time	aws_account_id	Message

1	2020-05-31
	13:53:35.000 +0900
		24[redacted]621	View as Raw
			{
			Records: ▸ […]
			}
			Host: Sumo Cloud
			Name: AWSLogs/24[redacted]621/CloudTrail/us-east-1/2020/05/31/24[redacted]621_CloudTrail_us-east-1_20200531T0455Z_rWlgotr2p4Yc1Pxe.json.gz
			Category: sandbox/watanabe/cloudtrail

2	2020-05-31
	13:51:58.000 +0900
		24[redacted]621	View as Raw
			{
			Records: ▸ […]
			}
			Host: Sumo Cloud
			Name: AWSLogs/247[redacted]621/CloudTrail/us-east-1/2020/05/31/24[redacted]621_CloudTrail_us-east-1_20200531T0455Z_kAkGHtWYNsQ48gsp.json.gz

summary

I tried to get parameters that are not in the log from the file name.
Sumo Logic allows you to drill down based on information that is not necessarily in the log text or source category, so please try it out.