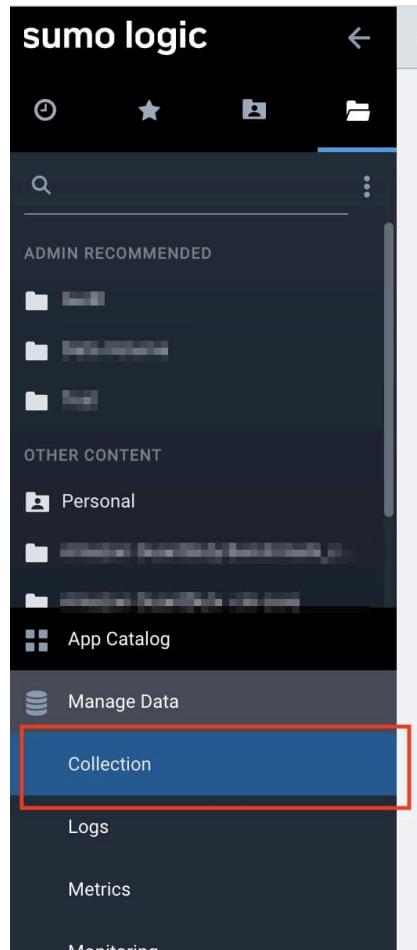


# How to ingest logs from S3 into Sumo Logic

 [dev.classmethod.jp/articles/sumo-logic-aws-s3](https://dev.classmethod.jp/articles/sumo-logic-aws-s3)

酒井剛

June 3, 2022



When trying to aggregate and store AWS logs within AWS, it is very common to store the data in S3. In this article, we will introduce the most basic method of importing logs from AWS, which is from S3.

Let's take a look.

## Introduction

There are two ways to import logs into Sumo Logic. In Sumo Logic, the function responsible for data collection is called **a collector**. There is an **agent-type (Installed Collector)** that is installed on a host or instance, and an **agentless type (Hosted Collector)** for services like S3 (Cloud Watch in an AWS environment, Microsoft 365 audit logs, etc.) that we will introduce here. Next, you define **a source** for each type of log you want to import, and link that source to a collector to begin log

collection. Depending on the type of log, the source is a setting that defines various settings such as the log location, credentials for accessing the logs, polling and log acquisition conditions, etc.

## procedure

---

The import settings are generally set up in the following steps:

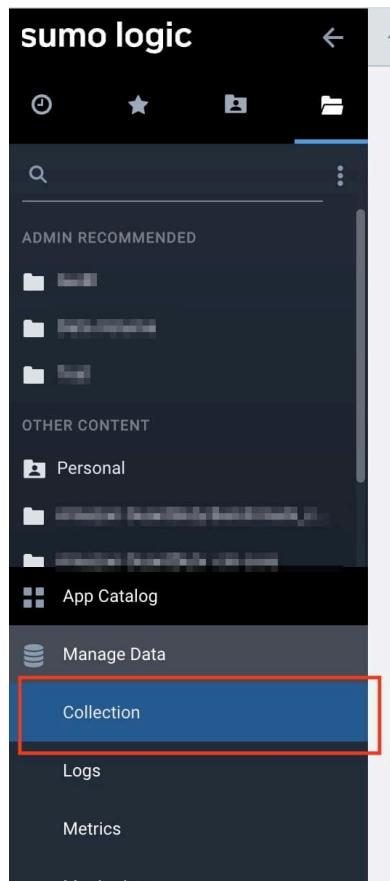
1. Create a Hosted Collector in Sumo Logic
2. In Sumo Logic, add a source to the collector you created.  
Set permissions (IAM role) in the AWS console during source configuration

Sumo Logic has designed its setup process to be simple, minimizing the burden on users. Log integration is possible in just a few steps.

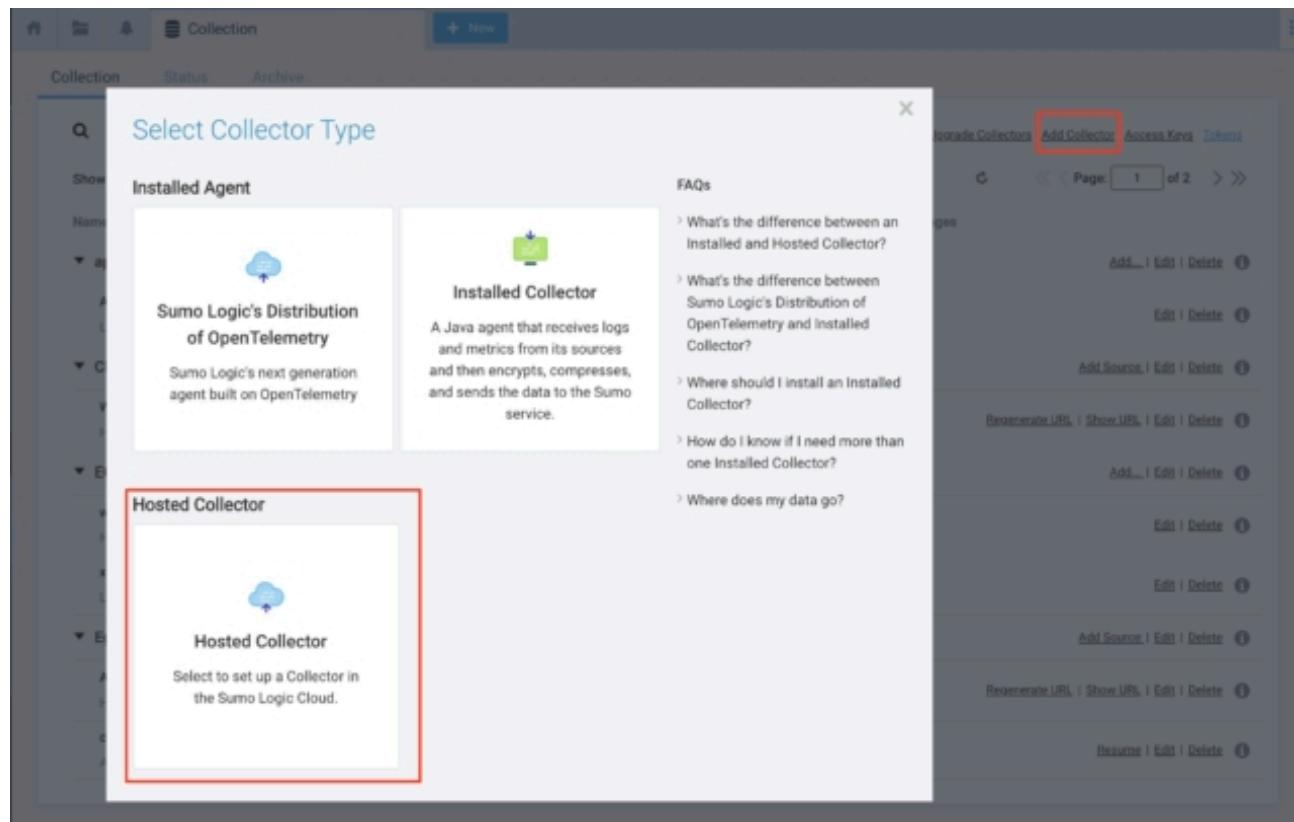
### 1. Create a Hosted Collector in Sumo Logic

---

Click Collection under Manage Data on the sidebar

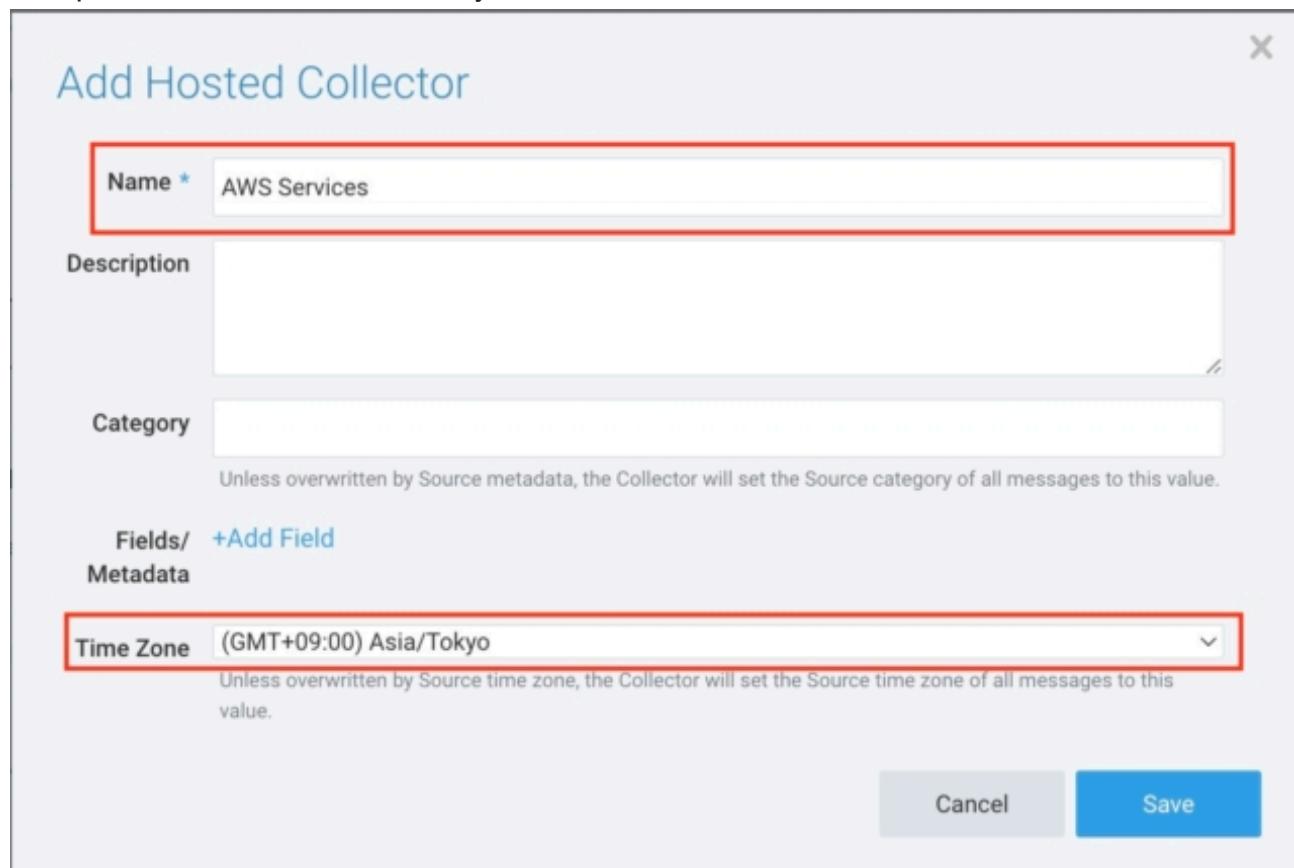


## Select Add Collector -> Hosted Collector



The screenshot shows the 'Collection' page in the Sumo Logic interface. A modal window titled 'Select Collector Type' is open. It contains three options: 'Installed Agent', 'Installed Collector', and 'Hosted Collector'. The 'Hosted Collector' option is highlighted with a red box. To the right of the modal, there is a sidebar with 'FAQs' and a list of collectors. The 'Add Collector' tab in the top navigation bar is also highlighted with a red box.

Set up a collector with a name of your choice and a time zone.

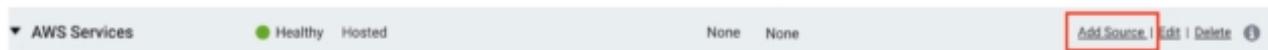


The screenshot shows the 'Add Hosted Collector' dialog box. It has fields for 'Name' (set to 'AWS Services'), 'Description', 'Category', 'Fields' (+Add Field), 'Metadata', and 'Time Zone' ((GMT+09:00) Asia/Tokyo). The 'Name' and 'Time Zone' fields are highlighted with red boxes. At the bottom, there are 'Cancel' and 'Save' buttons.

## 2. In Sumo Logic, add a source to the collector you created.

### Sumo Logic Console

If you are creating a new collector, simply select Add Source or click "Add Source" for the collector you just created.



Select "Amazon S3" or "AWS CloudTrail" if you are storing your CloudTrail in S3.

On the source settings screen, first fill in the following fields:

The screenshot shows the 'Add Source' configuration page for AWS S3. The form includes fields for Name (VPC), Description, S3 Region (Others), and various API settings. It also includes fields for Bucket Name (vpc-sumologic) and Path Expression (AWSLogs/\*). The 'Collection should begin' dropdown is set to 24 hours ago. The Source Category field is set to aws/prod/vpc. The 'Fields' section has a '+Add Field' button.

Name\*: VPC  
Maximum name length is 128 characters.

Description

S3 Region: Others

Use AWS versioned APIs?  Yes  No  
Select 'Yes' to use the list-object-versions and get-object-version AWS S3 APIs.  
Selecting 'Yes' requires the credentials to have ListObjectVersions and GetObjectVersion permissions.

Bucket Name\*: vpc-sumologic

Path Expression\*: AWSLogs/\*  
Path expression to match one or more S3 objects.  
For example: ABC\*.log or ABC.log  
NOTE: Make sure the path does NOT start with a leading slash.

Collection should begin: 24 hours ago  
(starts approx. at 06/02/2022 12:00:00 AM)

Source Category: aws/prod/vpc  
Category metadata to use later for querying, e.g. prod/web/apache/access. This data is queried using the '\_sourceCategory' key name.

Fields: +Add Field

- Name: Any name you like is OK
- Bucket Name: Specify the bucket where the acquired logs are saved.
- Path Expression: Specify the path within the bucket to retrieve. (You can use an asterisk. The asterisk can only appear once in the path, so you can write it like this: path/\*.log.)
- Source Category: Add any value to the metadata to identify logs when searching logs. It is recommended to use a hierarchical definition such as "aws/prod/vpc". We recommend using this naming scheme before using Sumo in production. For more information, please refer to the blog below.

Under Access Method, click "Generate role-based access template." This will download a CloudFormation YAML template, which you can use to access the AWS management console and configure an IAM role that will allow Sumo to access the S3 logs in your AWS environment.

How should Sumo Logic access your AWS account? [Learn more](#)

Access Method\*  Role-based access (recommended)  Key access

Use an AWS CloudFormation template to create an IAM role:

[Generate role-based access template](#) [Learn more](#)

Or, manually create a role on your AWS IAM console using the following information:

Account ID: [REDACTED]  
External ID: [REDACTED]  
[Learn more](#)

Role ARN\*

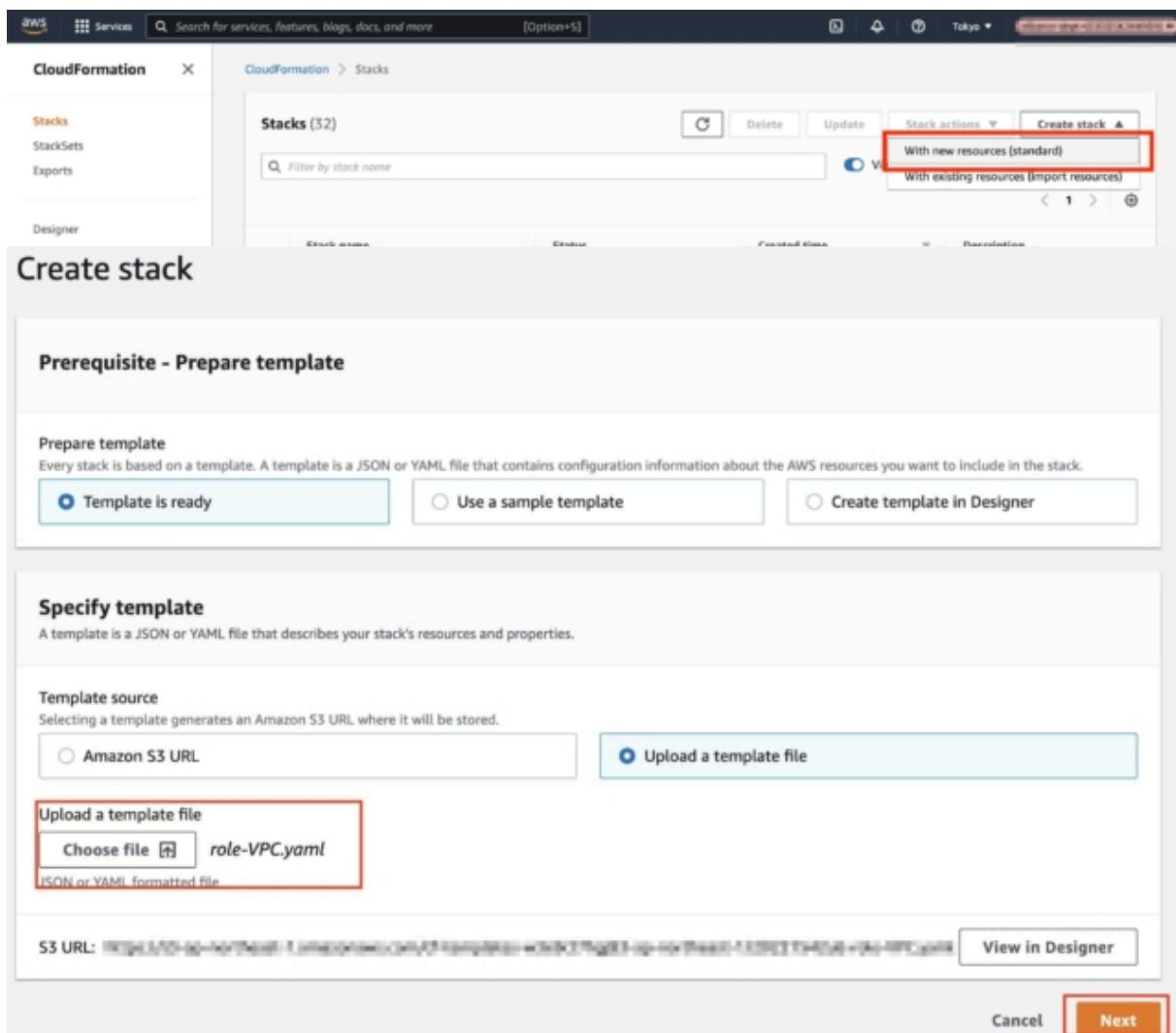
**CloudFormation Template - IAM Role-based Access**

```
AWSTemplateFormatVersion: '2010-09-09'
Description: A CloudFormation template that creates a role for authenticating with Sumo's AWS integrations.
Resources:
  SumoRole:
    Type: AWS::IAM::Role
    Properties:
      AssumeRolePolicyDocument:
        Version: '2012-10-17'
        Statement:
          - Effect: Allow
            Principal:
              AWS: arn:aws:iam:[REDACTED]
            Action: sts:AssumeRole
            Condition:
              StringEquals:
                sts:ExternalId:
                  [REDACTED]
        Path: "/"
      Policies:
        - PolicyName: SumoPolicy
          PolicyDocument:
            Version: '2012-10-17'
            Statement:
              - Effect: Allow
                Action:
```

[Close](#) [AWS CloudFormation Console](#) [Download](#)

## AWS Console

- Using the YAML file downloaded from CloudFormation in the AWS console, create an IAM role that will allow Sumo Logic to call the Get API for the bucket you specified earlier.



- Once CloudFormation is complete, check the ARN of the created IAM role from the output of the created stack.

sumo-vpc					
Stack info	Events	Resources	Outputs	Parameters	Template
<b>Outputs (1)</b>					
Key	Value		Description		Export name
SumoRoleARN	arn:aws:iam::123456789012:role/sumo-vpc-role		ARN of the created role. Copy this ARN back to Sumo to complete the source creation process.	-	

## Sumo Logic Console

- Once you have completed the above steps, return to the Sumo Logic console and enter the ARN of the role you just created.

How should Sumo Logic access your AWS account? [Learn more](#)

Access Method\*  Role-based access (recommended)  Key access

Use an AWS CloudFormation template to create an IAM role:

[Generate role-based access template](#) [Learn more](#)

Or, manually create a role on your AWS IAM console using the following information:

Account ID: [REDACTED]  
External ID: [REDACTED]  
[Learn more](#)

Role ARN\*

▼ Advanced Options for Logs

Enable Timestamp Parsing  Extract timestamp information from log file entries

Time Zone  Use time zone from log file. If none is detected use:

Ignore time zone from log file and instead use:

Timestamp Format  Automatically detect the format  Specify a format

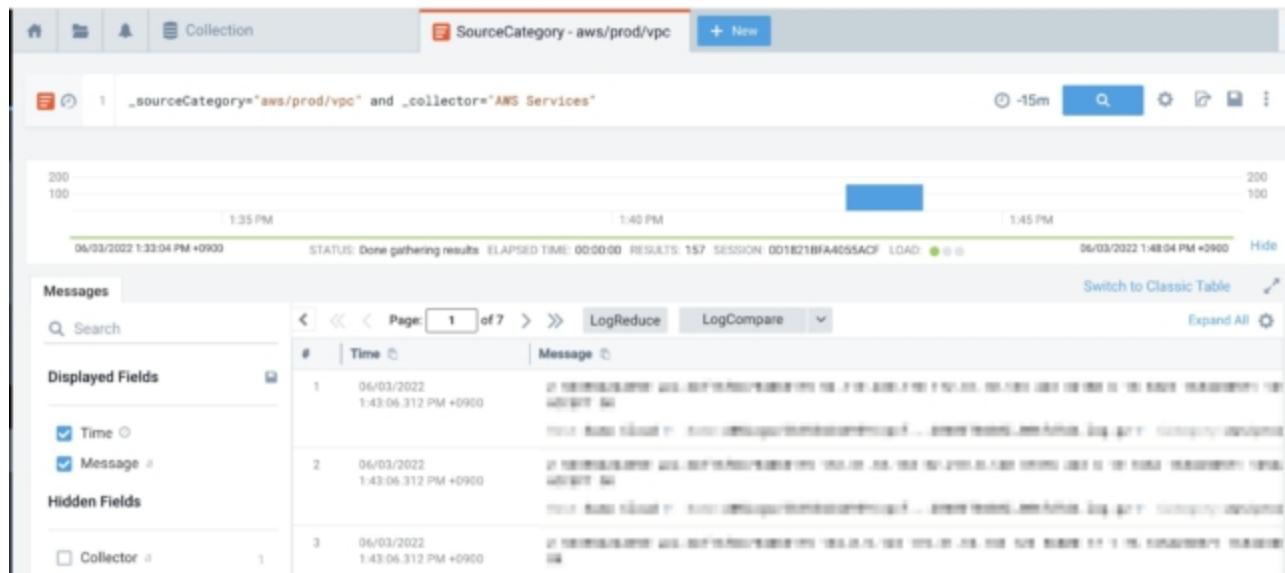
Enable Multiline Processing  Detect messages spanning multiple lines  
 Infer Boundaries - Detect message boundaries automatically  
Please note, Infer Boundaries may not be accurate for all log types.  
 Boundary Regex - Expression to match message boundary e.g. (?<!\r\n\r+)

► Processing Rules for Logs [What are Processing Rules?](#)

[Cancel](#) [Save](#)

## Verify that logs are being captured

Check that the logs are being imported by going to the Collection source or by opening a new Log Search screen. If they are not being imported, wait a while and try again.



## summary

This time, we introduced the most basic method of ingesting logs from AWS: ingesting logs from S3. We hope this article will be useful for those who are considering trying out Sumo Logic.