# Configure Splunk with a Self-managed S3 Bucket

## Contents



## Introduction

This document describes how to configure Splunk with a self-managed S3 bucket.

## Overview

Splunk is a common tool for log analysis. It provides a powerful interface for analyzing large chunks of data, such as the logs provided by Cisco Umbrella for your organization's DNS traffic.

This article outlines the basics of getting Splunk set up and running so it is able to pull the logs from your S3 bucket and consume them. There are two main stages, one is to configure your AWS S3 Security Credentials to allow Splunk access to the logs, and the second is to configure Splunk itself to point at your bucket.

The documentation for the Splunk Add-on for AWS S3 is here, some of which has been copied verbatim into this document. For specific questions regarding Splunk setup, please refer to http://docs.splunk.com/Documentation/AddOns/latest/AWS/Description

This article has these sections:

- Prerequisites
- Stage 1: Configuring your security credentials in AWS (self-managed bucket only)
- Stage 2:  Setting up Splunk to pull DNS log data from your S3 bucket
    - Step 1: Setting up Splunk to pull DNS log data from self-managed S3 bucket
- Stage 3: Configuring Data Inputs for Splunk

## Prerequisites

The Splunk Add-on for Amazon Web Services supports these platforms.

- AWS Linux
- RedHat
- Windows 2008R2, 2012R2

### Splunk Enterprise system requirements

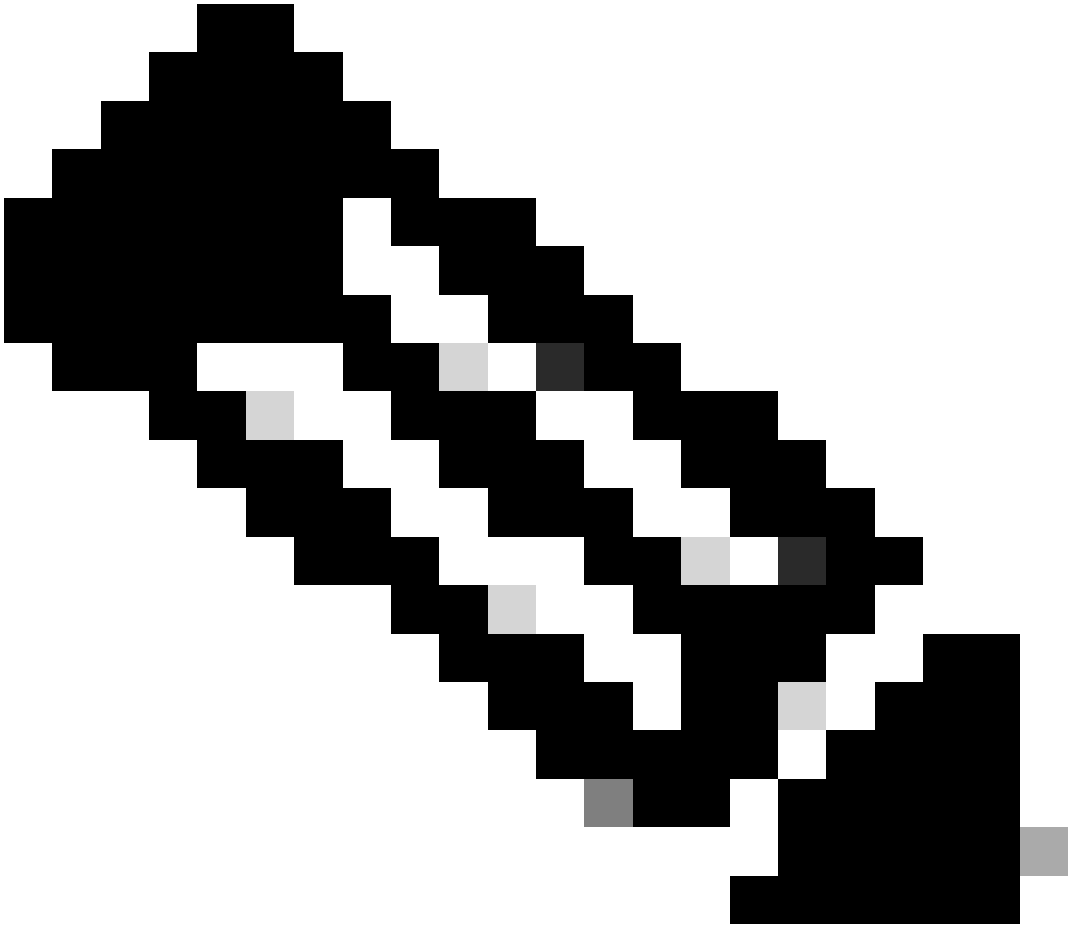
Because this add-on runs on Splunk Enterprise, all of the Splunk Enterprise system requirements apply. See the "System Requirements" Installation Manual in the Splunk Enterprise documentation. These instructions are for the Splunk Enterprise version 6.2.1.

### Umbrella requirements

This document assumes that your Amazon AWS S3 bucket has been configured in the Umbrella dashboard (Admin> Log Management) and is showing green with recent logs having been uploaded. For more information on Log Management,  see Cisco Umbrella Log Management in Amazon S3.

## Stage 1: Configuring your Security Credentials in AWS

---



**Note**: These steps are the same as those outlined in the article describing how to configure a tool to

---

download the logs from your bucket (How to: Downloading logs from Cisco Umbrella Log Management in AWS S3 ). If you have already performed those steps, you can simply skip to step 2, although you need the security credentials from your IAM user to authenticate the Splunk plugin to your bucket.

---

### Step 1

1. Add an access key to your Amazon Web Services account to allow remote access to your local tool and give the ability to upload, download and modify files in S3. Log in to AWS and click your account name in the upper-right hand corner. In the drop-down, choose **Security Credentials**.
2. You are prompted to use Amazon Best Practices and create an AWS Identity and Access Management (IAM) user. In essence, an IAM user ensures that the account that s3cmd uses to access your bucket is not the primary account (for example, your account) for your entire S3 configuration. By creating individual IAM users for people accessing your account, you can give each IAM user a unique set of security credentials. You can also grant different permissions to each IAM user. If necessary, you can change or revoke an IAM user's permissions at any time.
   For more information on IAM users and AWS best practice, read here: https://docs.aws.amazon.com/IAM/latest/UserGuide/best-practices.html

### Step 2

1. Create an IAM user to access your S3 bucket by clicking **Get Started with IAM Users**. You are taken to a screen where you can create an IAM User.
2. Click **Create New Users**, then go ahead and fill out the fields. Note that the user account cannot contain spaces.
3. After creating the user account, you are given only one opportunity to grab two critical pieces of information containing your Amazon User Security Credentials. **We highly recommend that you download these using the button in the lower right to back them up. They are not available after this stage in the setup.** Ensure you make a note of both your Access Key ID and Secret Access Key as we need them later when setting up Splunk.

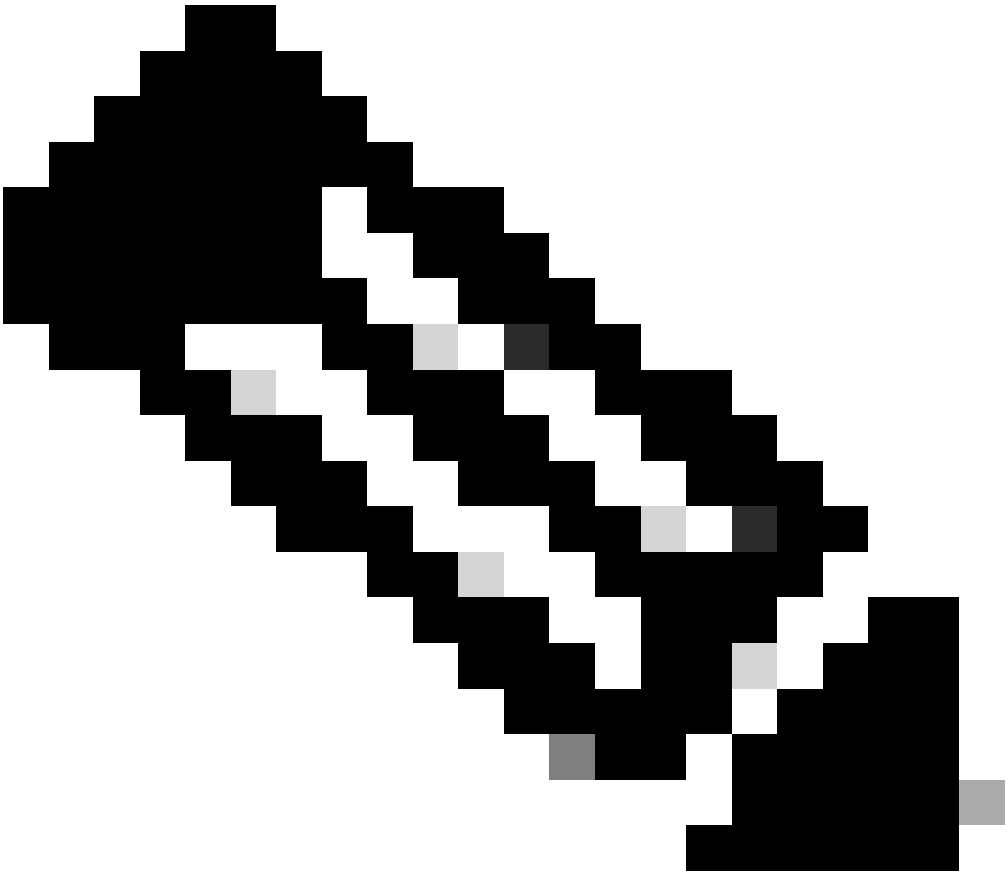
### Step 3

1. Next, you want to add a policy for your IAM user so they have access to your S3 bucket. Click the user you have just created and then scroll down through the users' properties until you see the Attach Policy button.
2. Click **Attach Policy**, then enter 's3' in the policy type filter. This shows two results: "AmazonS3FullAccess" and "AmazonS3ReadOnlyAccess".
3. Select **AmazonS3FullAccess** and then click **Attach Policy**.

## Stage 2: Setting up Splunk to pull DNS log data from your S3 bucket

### Step 1 : Setting up Splunk to pull DNS log data from self-managed S3 bucket

1. Start by installing the "Splunk Add-on for Amazon Web Services" to your Splunk instance. Open your Splunk dashboard and click **Apps**, or click **Splunk Apps** if it appears on your dashboard. Once in the Apps section, type "s3" in the search window to find "Splunk Add-on for Amazon Web Services", and install the app.



**Note**: You likely need to restart Splunk during the installation.
Once it is installed, you see Splunk Add-on for AWS with the folder name 'Splunk_TA_aws' now listed under Apps.

2. Click **Set up** to configure the app. This is the point where you need the Security Credentials from Stage 1 in this documentation.
   The setup requires these fields be entered:

   - A friendly name—the name you use to refer to this integration
   - Your AWS account Key ID (from stage 1)
   - Your password (your AWS account Secret Key, also from stage 1)

You can also set any local proxy information if it is required for Splunk to reach AWS, as well as adjusting logging. The setup screen looks like this:
3. Once you have added relevant information, click **Save** and the Splunk Add-on for Amazon Web Services are fully configured.

# Stage 3: Configuring Data Inputs for Splunk

1. Next, you want to configure the data input for Amazon Web Services S3. Navigate to **Settings > Data > Data Inputs** and under Local Inputs, you now see a list of various Amazon inputs including S3 at the bottom of the list.

2. Click **AWS S3** to configure the input.
3. Click **New**.
4. You are required to provide these pieces of information:

    - Enter a friendly name for your S3 integration.
    - Select your AWS account from the dropdown. This is the friendly name you provided in step 1.
    - Select your S3 bucket from the dropdown. This is the bucket name as specified in your Umbrella dashboard (Settings > Log Management).
    - Select the S3 key name from the dropdown. Every item in your bucket is listed, we recommend picking the top level directory \dns-logs\ , which includes all of the files and directories under it.

    - There are several options under "Message system configuration", we recommend leaving these as is—default settings.
    - There are additional options under "More settings." Of note is the "Source type", which is aws:s3 by default. We recommend leaving this as is, but if you do change it, the filter for your logs in the Search changes from what is described in Step 3 of these instructions.

Fill in the details, and your data input looks similar to this:

4. Click **Next** to finalize your details.
   You are taken to a screen that shows that the input was created successfully

## Step 3

Perform a quick search to see if your data is being imported properly. Just paste *sourcetype="aws:s3"* into the Search window in the upper right and then select "Open sourcetype="aws:s3" in search

This takes you to a screen similar to the one where you see the events from your organizations' DNS logs. Here, the Cisco Umbrella mobile service is blocking social media on an iPhone. You can also use the source of the filename to filter against a particular batch of logs.

After this point, the cron job in the background continues to run and pull down the latest sets from log information from your bucket.

There is a lot more you can do with Splunk beyond what has been outlined in this article, and if you have had a chance to experiment with using this data in your security response procedure, we would love to hear from you. Send any feedback, questions or concerns to umbrella-support@cisco.com and reference this article.