exabeam

# Exabeam Site Collector Guide

Exabeam Management Platform - Version SMP 2021.1

Publication date April 30, 2021

**Exabeam**
1051 E. Hillsdale Blvd.
4th Floor
Foster City, CA 94404

1.844.392.2326

Have feedback on this guide? We'd love to hear from you!
Email us at docs@exabeam.com

## Table of Contents

# 1. Exabeam Site Collector

The Exabeam Site Collector is an application that securely and efficiently uploads event data to Exabeam SaaS services in the Exabeam Cloud Platform.

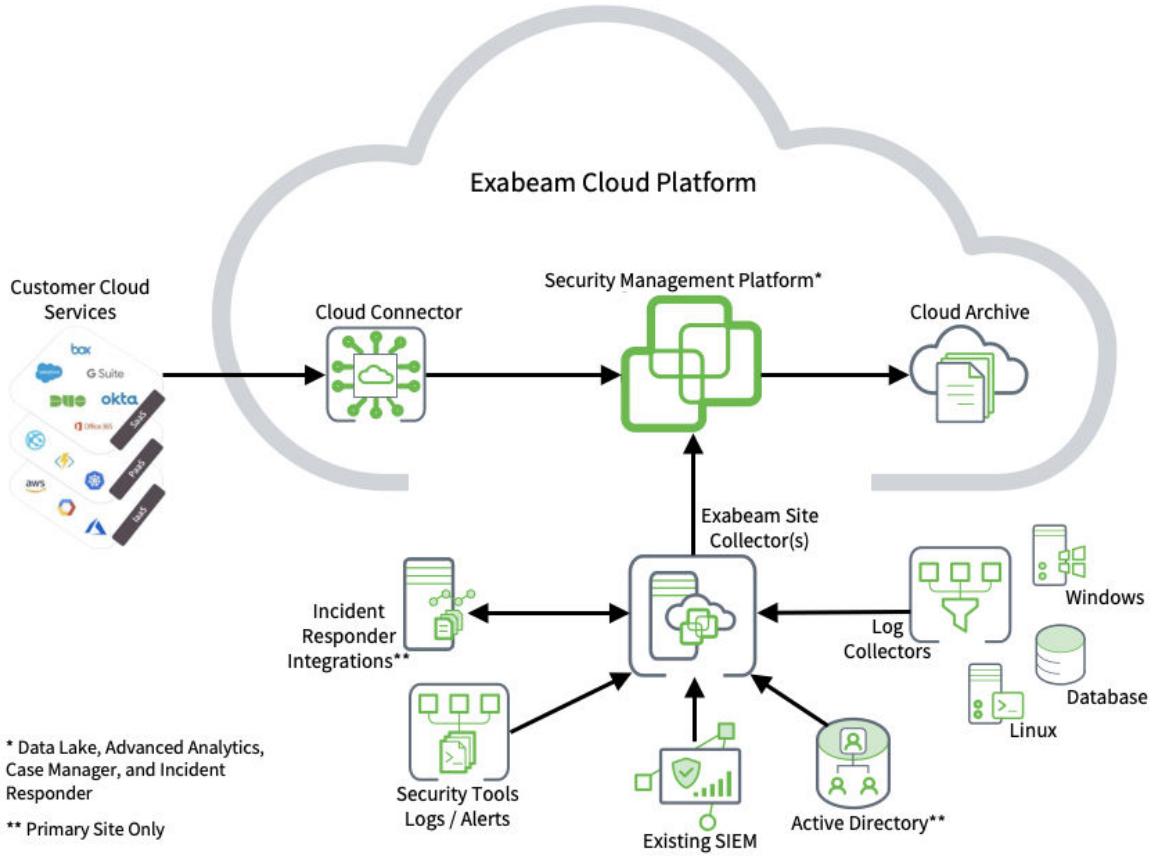



**Figure 1. Exabeam Site Collector High-level Architecture**

At a high level, the Exabeam Site Collector collects messages; transfers, persists, and uploads data; and connects to Exabeam Cloud Platform.

The Exabeam Site Collector is the managed entry point for logs to be routed to other processing tools, such as Exabeam Advanced Analytics, Exabeam Data Lake, Exabeam Incident Responder, and Exabeam Case Manager in the Exabeam Security Management Platform. Site collectors gather logs from external servers, systems, data centers, or Exabeam collectors (including Windows, File, and GZip).

The site collector routes collected logs to the Exabeam Security Management Platform. The site collector continuously queues and uploads logs as well as manages the forwarding rate and message backlog. Data is encrypted and compressed in transmission and while at rest in the Exabeam Cloud Platform. A persistent connection to the Exabeam Cloud Platform allows the site collector to connect to your assets such as Active Directory for context and authentication, access API for log repositories, and any Incident Responder actions.

You can deploy multiple site collectors as your log volume and sources grow. You can deploy them in any network, data center or virtual private cloud (VPC) as required. All SaaS service log data is collected via API using the Exabeam Cloud Platform. See Install Exabeam Site Collector for prerequisites and restrictions.

**IMPORTANT**
For information on configuring agent-based or server-side collectors, please refer to the *Exabeam Collector Guide*.

## 2. Exabeam Site Collector Network Ports

Apply the port configurations that match your deployment. These ports are required for the Exabeam Site Collector to operate correctly. In addition, communications for deployment-specific scenarios must also be allowed.

When you whitelist a syslog source, you may need to refer to the Exabeam Site Collector's certificate authority. You can whitelist Transport Layer Security (TLS) syslogs as a source that can be whitelisted. For more information on configuring , see Configure Transport Layer Security (TLS) Syslog Ingestion.

| Source | Destination | Port | Protocol | Description |
|---|---|---|---|---|
| **All Site Collectors** | DNS Server | 53 | DNS | DNS lookup |
| **All Site Collectors** | NTP Server | 123 | NTP | Time synchronization |
| **Administrator Network** | All Site Collectors | 22 | SSH | Administrator command line access to host via encrypted connection |
| **Log Sources** | All Site Collectors | 514<br>or<br>515 (TLS) | Syslog | Collector registration and monitoring and syslog ingestion port from log sources<br>Unidirectional traffic |
| **All Site Collectors** | <InstanceID>.beats.exabeam.com | 443 | HTTPS | Exabeam Site Collector registration and monitoring |
| **Primary Site Collector** | <InstanceID>.connect.exabeam.com | 443 | TCP | OpenVPN tunnel for on-premises deployments<br>**WARNING**<br>Do not configure more than one Open VPN connection per site collector and per SaaS tenant. Otherwise, network conflicts will occur. |
| **Primary Site Collector** | Domain Controller(s)<br>Global Catalog | 389 or 636<br>3268 or 3269 | LDAP<br>-or-<br>LDAPS | Active Directory context and administrator authentication |
| **All Site Collector** | accounts.google.com | 443 | HTTPS | Upload to Google Cloud Storage/ Pub-Sub |
| **All Site Collector** | *.googleapis.com<br>or<br>oauth2.googleapis.com<br>www.googleapis.com<br>storage.googleapis.com<br>pubsub.googleapis.com<br>accounts.googleapis.com | 443 | HTTPS | Upload to Google Cloud Storage/ Pub-Sub |

**IMPORTANT**
For Google Cloud Platform connections:

On-premises data centers must enable traffic on firewalls and proxies on TCP ports `account.google.com:443` and `*.googleapis.com:443` to access cloud services such as Cloud Storage .

Alternatively, you can use Private Google Access to connect to Google Cloud Platform services. For more information, see Private Google Access for on-premises hosts.

## 2.1. Additional Ports for Specific Configurations

If you are deploying additional services, review and configure appropriate ports if the following services match your environment:

| Source | Destination | Port | Protocol | Description |
|---|---|---|---|---|
| Exabeam Log Collector | Local Site Collector | 8484 | HTTPS | Exabeam Log Collector registration and monitoring |
| Exabeam Log Collector | Local Site Collector | 9092<br>9093 | KAFKA TCP | Windows and Linux event collection using Exabeam Log Collector |

**Table 1. Exabeam log collectors in Windows or Linux**

| Source | Destination | Port | Protocol | Description |
|---|---|---|---|---|
| Primary Site Connector | Splunk | 8089 | HTTPS | Log collection using Site Connector to poll systems directly |
| | QRADAR | 443 | | |
| | Other Log repositories/databases | Various | | |

**Table 2. Exabeam polling for log collection**

| Source | Destination | Port | Protocol | Description |
|---|---|---|---|---|
| Primary Site Collector | Orchestrated security products and servers | Various | HTTPS | Third-party integrations |

**Table 3. Third-party integration**

## 3. Exabeam Site Collector Specifications

The number of site collectors needed in your deployment depends on where the data is located, the data volumes and resilience requirements. Exabeam Site Collector hosts need to meet one of the following specifications to support the expected data volumes. Site collectors in parallel behind a load-balancer should be sized so there is no impact to data collection, if one of them is taken out of service. Please review the specifications in the table below that applies to your environment:

| Maximun Events Per Second (EPS) | Minimum CPU and Memory[1] | Maximum Agents/ Collectors | Operating System Volume | Storage Volume[2] |
|---|---|---|---|---|
| 1 - 1,000 | 4 CPU, 8 GiB RAM | 100 | 100 GB | 600 GB |
| 1,000 - 5,000 | 4 CPU, 8 GiB RAM | 100 | 100 GB | 3 TB |
| 5,000 - 20,000 | 8 CPU, 16 GiB RAM | 200 | 100 GB | 12 TB |
| 20,000 - 30,000 | 16 CPU, 32 GiB RAM | 500 | 100 GB | 18 TB |

**Table 4. Site Collector Ingestion Capacities**

[1] *GiB is Gibibyte, the unit on the Google Cloud Platform (GCP), and is equal to 1.07374 GB.*

[2] *Data Storage is based on an average message size of 1500 bytes and 24 hours data retention (default).*

Additionally, please ensure the following storage requirements and permissions are met:

- CentOS/RedHat 7.x
- `/` must have a minimum 100 GB is required for site collector operations
- `/tmp` must have full `root` permissions
- Ensure `/` and `/opt` are configured for disk usage
- `/data` is storage for Kafka data (sizing is based on the Site Collector Specifications above) with 300 GB or higher per EPS
- Default local retention is 24 hours or available free disk space in `/data` allocation

> **IMPORTANT**
> For capacity specifications that are not shown, please contact your Exabeam technical representative for assistance in calculating retention and EPS rates.

Where possible we recommend there is at least 2 site collectors deployed behind a load balancer for high availability performance. You can deploy as many site collectors as required for your logs processing. One site collector must have OpenVPN if your ingestion is to support LDAP polling, database logs, eStreamer logs and fetching by Advanced Analytics or Incident Responder accessing local endpoints.

# 4. Install Exabeam Site Collector

The Exabeam Site Collector to lets you upload log data from your data centers or virtual private clouds (VPCs) to Exabeam. Site collectors are designed to support most data centers with a single site collector, along with on-premises deployments. You may install more site collectors as log volumes grow.

If you are installing a site collector with Exabeam Advanced Analytics in your deployment, you will not be able to view the health of the site collector as it will not appear or monitored in the Advanced Analytics user interface. The host will be running in unmanaged mode. Please implement a custom monitoring solution to track the health of the node.

## 4.1. Prerequisites

Here are the prerequisites common to all environments before installing a site collector. Additional prerequisites may apply based on your deployment type.

- If there is a syslog source in your deployment, install a load balancer with two site collectors behind it to mitigate any potential data loss
- Ensure you have SSH login access to the site collector host
- If Security Enhanced (SE) Linux is enabled, use `permissive` mode to perform administrative tasks (such as installing, upgrading, and configuring) and then revert to `enforcing` mode after completing tasks
  (Run `getenforce` command to confirm status.)
- Site collector hostname's `A Record` must be resolvable using a domain name service (DNS)
  (Run `ping [hostname]` to confirm.)
- If UDP is being used, the source IP must still be routable from the site collector
- Determine which network zones will need to allow site collector traffic and the IP addresses and protocols that can be supported
- The `/tmp` partition on the site collector host is executable for root
- Ensure there is enough space and resources for site collector installation
- Services:
  - Network proxies are not supported where an on-premises endpoint is the log destination
  - Proxies with authentication are not supported
  - SSL authentication or interception is not supported
  - NTP client must be active and synchronized
    (Run `timedatectl` command to confirm.)
  - One site collector must have OpenVPN if your ingestion is to support LDAP polling, database logs, eStreamer logs and fetching by Advanced Analytics or Incident Responder accessing local endpoints
- Data collecting:
  - Use syslog, or secure syslog, over TCP where possible

- If you must use UDP for data transfer, still ensure the source host IP address is routable from Site Collector
- If syslog is not possible, determine the appropriate client, such as an Exabeam Log Collector, to support the collection method. See Exabeam Data Lake Collector Guide for more options

- Firewall:
  - Allow traffic between source and destination ports that match your deployment
  - Internet access is allowed at the site collector host
  - `firewalld` service is running on the site collector host
    (Run `systemctl status firewalld` to confirm `active (running)` status.)
  - Firewall rules should be built using FQDNs or domain wildcards

> **IMPORTANT**
> Do not use IP addresses for access to the Exabeam SaaS Environment. They are dynamic and can change based upon location and scaling of the service.

## 4.2. Install Site Collector Based on Deployment Environment Type

Follow the installation instructions that matches your deployment environment:

- Install a SaaS site collector with Data Lake in the environment
- Install a SaaS site collector for Advanced Analytics -only deployments (site collector in an unmanaged node)

For on-premises deployments, see Install and Upgrade Exabeam Site Collector for On-premises and Legacy Deployments.

## 4.3. Install Site Collector for Exabeam SaaS Data Lake

The following instructions are for a fresh Exabeam Site Collector installation so logs are sent to Exabeam's SaaS Data Lake .

### 4.3.1. PREREQUISITES

Ensure your environment meets all requirements before running a site collector installation. Please review prerequisites listed in Install Exabeam Site Collector in addition to the following:

- If you have a proxy,
  - Ensure that the proxy does not require site collector traffic to be authenticated
  - Configure both HTTPS and OpenVPN routes for:
    - On-premises to SaaS for data flow
    - SaaS to on-premises via OpenVPN for data such as LDAP polling

**NOTE**
OpenVPN must be used for:

1. Passing LDAP poll data

2. Using a DBlog collector in your deployment

3. Using eStreamer in your deployment

4. Fetching any on-premises SIEM / sources by Advanced Analytics

5. Connecting to on-premises endpoints by Incident Responder Actions

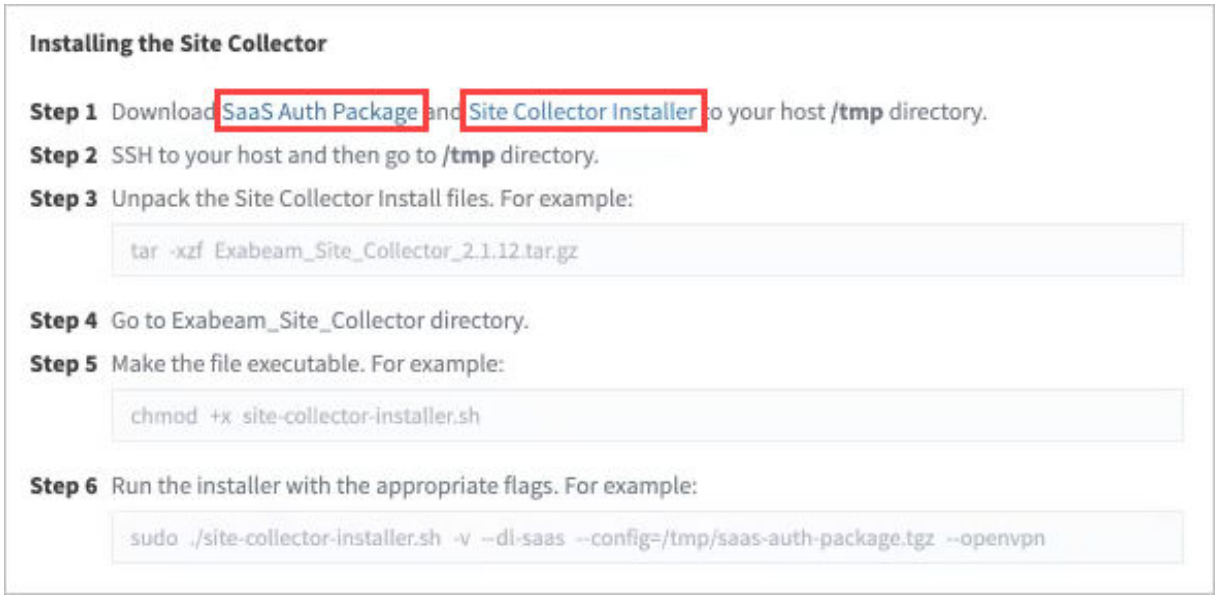
**Limitation:**

Only one OpenVPN connection can be active at a time. You need to have it installed onto more than one site collector (for active/standby option), disable the active service manually after installation.

Use commands:

```
sudo systemctl stop openvpn@<instanceID>
sudo systemctl disable openvpn@<instanceID>
```

1. In Data Lake, navigate to **Settings** > **SAAS Management** > **SaaS Site Collectors**.
2. Download the **SaaS Auth Package** and **Site Collector Installation Package**. These packages contain all required configurations and authentication data needed to access your SaaS tenant and installation package.




3. Use `scp` (secure copy) to place the files in the `/tmp` directory of the site collector host. (For help with this command, run `man scp`.)

```
scp <source_host>:<directory>/<package_file> <site_collector>:<directory>/
package_file>
```

4. Start a new terminal session using the an account with Administrator rights. Initiate a screen session. This is mandatory and will prevent accidental termination of your session.

   ```
   screen -LS [yourname]_[todaysdate]
   ```

5. Go to the `/tmp` directory and unpack the installation package only.

   ```
   cd /tmp
   tar -xzf <install_filename>.tar.gz
   ```

6. Go to the `Exabeam_Site_Collector` directory.

   ```
   cd Exabeam_Site_Collector
   ```

7. Make the files executable.

   ```
   chmod +x site-collector-installer.sh
   ```

8. Based on your deployment environment, execute one of the following installation commands:
   a. Installing site collector but with OpenVPN

      ```
      sudo ./site-collector-installer.sh -v --dl-saas --config=/tmp/saas-auth-
      package.tgz --openvpn
      ```

   b. Installing site collector and without OpenVPN

      ```
      sudo ./site-collector-installer.sh -v --dl-saas --config=/tmp/s-auth-
      package.tgz
      ```

   c. Installing site collector behind the proxy with OpenVPN

      ```
      sudo ./site-collector-installer.sh -v --dl-saas --config=/tmp/saas-auth-
      package.tgz --openvpn --proxy=<proxy_host_ip|proxy_hostname> --proxy-
      port=<proxy_port>
      ```

   d. Installing site collector behind the proxy without OpenVPN

      ```
      sudo ./site-collector-installer.sh -v --dl-saas --config=/tmp/saas-auth-
      package.tgz --proxy=<proxy_host_ip|proxy_hostname> --proxy-
      port=<proxy_port>
      ```

9. Once installation is complete, the prompt returns `Site collector installer complete`.
10. To verify that the site collector source has been installed, log into the Data Lake and navigate to **Settings** > **Collector Management** > **Collectors** to see the list of configured collectors.


> **NOTE**
> It is normal to find the **Site Collector Data Forwarder** service is shown as `Stopped` while another service is shown as `Running`. To verify if there is on-going ingestion, one of these services will show non-zero messages in the graph.

You can also send a test message via syslog and confirm it arrived at the destination via Data Lake after several minutes:

```
echo "test message [date_time] from [hostname|host_ip]" | nc localhost 514
```

If your site collector does not appear in the list and the test message did not reach its destination, review the known common scenarios in Troubleshoot for Exabeam Site Collector that can be resolved immediately.

## 4.4. Install Site Collector for Exabeam SaaS Advanced Analytics-only Deployment

The following instructions are for a fresh Exabeam Site Collector installation so your logs are sent to Exabeam's SaaS Advanced Analytics and where there is no Exabeam Data Lake deployed. The ingestion status for this configuration can be found on the Status Page in the Exabeam SaaS Cloud.

### 4.4.1. PREREQUISITES

Ensure your environment has met all requirements before running a site collector installation. Please review prerequisites listed in Install Exabeam Site Collector in addition to the following:

- If you have a proxy:
  - Ensure that the proxy does not require site collector traffic to be authenticated
  - Configure both HTTPS and OpenVPN routes for:
    - On-premises to the Exabeam SaaS Cloud for data flow
    - Exabeam SaaS Cloud to on-premises via OpenVPN for data such as LDAP polling

**NOTE**
OpenVPN must be used for:

1. Passing LDAP poll data

2. Using a DBlog collector in your deployment

3. Using eStreamer in your deployment

4. Fetching any on-premises SIEM / sources by Advanced Analytics

5. Connecting to on-premises endpoints by Incident Responder Actions

**Limitation:**

Only one OpenVPN connection can be active at a time. You need to have it installed onto more than one site collector (for active/standby option), disable the active service manually after installation.

Use commands:

```
sudo systemctl stop openvpn@<instanceID>
sudo systemctl disable openvpn@<instanceID>
```

1. Download SaaS Site Collector installation files from the Exabeam Community.
2. Download your authentication package file using the following URL template based on your `<instanceID>`.

   ```
   https://<instanceID>.aa.exabeam.com/api/setup/saas/authPackage
   ```

3. Place the files in the `/tmp` directory of the site collector host.
4. Start a new terminal session using the an account with administrator rights. Initiate a screen session. This is mandatory and prevent termination of your session.

   ```
   screen -LS [yourname]_[todaysdate]
   ```

5. Go to the `/tmp` directory and unpack the installation file only.

   ```
   cd /tmp
   tar -xzf <filename>.tar.gz
   ```

6. Go to the `Exabeam_Site_Collector` directory.

   ```
   cd Exabeam_Site_Collector
   ```

7. Make the files executable.

```
chmod +x site-collector-installer.sh
```

8. Based on your deployment environment, please execute one of the following installation commands:
   a. Installing site collector with OpenVPN

      ```
      sudo ./site-collector-installer.sh -v --aa-saas --config=/tmp/saas-auth-package.tgz --openvpn
      ```

   b. Installing site collector and without OpenVPN

      ```
      sudo ./site-collector-installer.sh -v --aa-saas --config=/tmp/saas-auth-package.tgz
      ```

   c. Installing site collector behind the proxy with OpenVPN

      ```
      sudo ./site-collector-installer.sh -v --aa-saas --config=/tmp/saas-auth-package.tgz --openvpn --proxy=<proxy_host_ip|proxy_hostname> --proxy-port=<proxy_port>
      ```

   d. Installing site collector behind the proxy without OpenVPN

      ```
      sudo ./site-collector-installer.sh -v --aa-saas --config=/tmp/saas-auth-package.tgz --proxy=<proxy_host_ip|proxy_hostname> --proxy-port=<proxy_port>
      ```

9. Once installation is complete, the prompt will return `Site collector installer complete`.
10. Verification checks must be made via the SaaS Status Page. The Status page is intended to show errors only and should not be used to verify throughput immediately after installation.
    You can send a test message via syslog and confirm it arrived at the log destination after several minutes:

    ```
    echo "test message [date_time] from [hostname|host_ip]" | nc localhost 514
    ```

    If no logs arrive at the destination after a few minutes or you cannot see a status for the site collector, review the known common scenarios in Troubleshoot for Exabeam Site Collector that can be resolved immediately.

# 5. Upgrade Exabeam Site Collector

Keep your site collectors up to date to take advantage of new features.

Follow the upgrade instructions that matches your deployment environment:

## 5.1. Prerequisites

Before upgrading your site collector, ensure prerequisites are met. Additional prerequisites may apply based on your deployment type.

- If you are adding a syslog source in your deployment, install a load balancer with two site collectors behind it to mitigate any potential data loss
- If Security Enhanced (SE) Linux is enabled, use `permissive` mode to perform administrative tasks (such as installing, upgrading, and configuring) and then revert the mode after completing tasks
- The `/tmp` partition on the site collector host is executable for root
- Ensure there is enough space for a site collector upgrade

## 5.2. Upgrade Site Collector Based on Deployment Environment Type

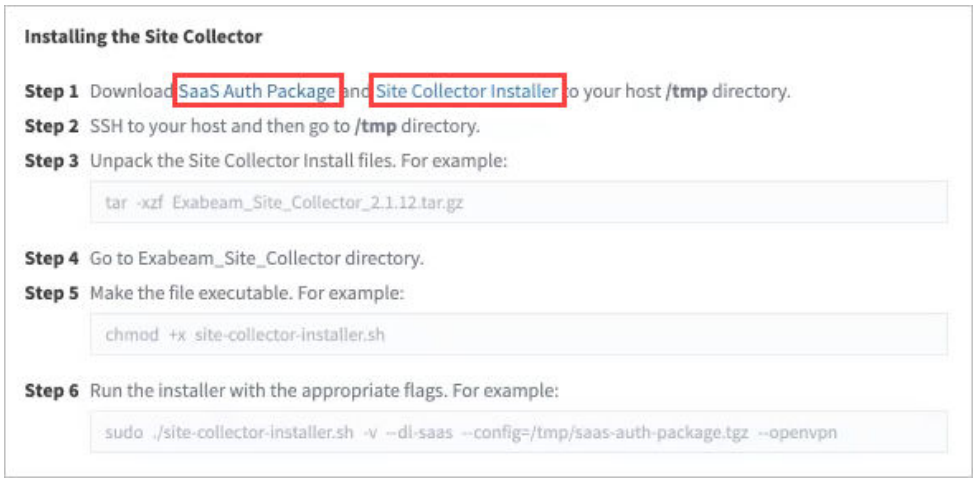
- Upgrade a SaaS site collector with Data Lake in the environment
- Upgrade a SaaS site collector for Advanced Analytics-only deployments (site collector in an unmanaged node)

For on-premises and legacy deployments, see Install and Upgrade Exabeam Site Collector for On-premises and Legacy Deployments.

## 5.3. Upgrade Exabeam Site Collector for SaaS with Exabeam Data Lake

The following instructions are for an Exabeam Site Collector upgrade if your logs are sent to Exabeam's SaaS Data Lake.

1. Ensure your environment has met all requirements before running a site collector upgrade.
2. In Data Lake, navigator to **Settings** > **Collector Management** > **Collectors**.
3. Click + ADD COLLECTOR to open the **Collector Artifacts** menu to get a list of **Site collectors**.



4. Download the **Site Collector Auth Package** and **Site Collector Installation Package**. These packages contain all required configurations and authentication data needed to access your SaaS tenant and installation packagge.
5. Place the files in the `/tmp` directory of the site collector host.
6. Start a terminal session to the site collector and initiate a screen session.

   ```
   screen -LS [yourname]_[todaysdate]
   ```

7. Go to the `/tmp` directory and unpack the downloaded files.

   ```
   cd /tmp
   tar -xzf <filename>.tar.gz
   ```

8. Go to the `Exabeam_Site_Collector` directory.

   ```
   cd Exabeam_Site_Collector
   ```

9. Make the files executable.

   ```
   chmod +x site-collector-installer.sh
   ```

10. Based on your deployment environment, please execute one of the following upgrade commands:
    a. Upgrade site collector behind the proxy with OpenVPN

       ```
       sudo ./site-collector-installer.sh -v --dl-saas --upgrade --config=/tmp/
       <instanceID>-auth-package.tgz --openvpn --proxy=<proxy_host_ip|
       proxy_hostname> --proxy-port=<proxy_port>
       ```

    b. Upgrade site collector behind the proxy without OpenVPN

       ```
       sudo ./site-collector-installer.sh -v --dl-saas --upgrade --config=/tmp/
       <instanceID>-auth-package.tgz --proxy=<proxy_host_ip|proxy_hostname> --
       proxy-port=<proxy_port>
       ```

    c. Upgrade site collector without proxy but with OpenVPN

       ```
       sudo ./site-collector-installer.sh -v --dl-saas --upgrade --config=/tmp/
       <instanceID>-auth-package.tgz --openvpn
       ```

    d. Upgrade site collector without proxy and without OpenVPN

       ```
       sudo ./site-collector-installer.sh -v --dl-saas --upgrade --config=/tmp/
       <instanceID>-auth-package.tgz
       ```

11. To verify that the site collector source has been upgraded, in Data Lake, navigate to **Settings** > **Collector Management** > **Collectors** to see the list of configured site collectors. The version should match the upgrade version.


> **NOTE**
> It is normal to find the **Site Collector Data Forwarder** service is shown as `Stopped` while another service is shown as `Running`. To verify if there is on-going ingestion, one of these services will show non-zero messages in the graph.

You can also send a test message via syslog and confirm it arrived at the destination via Data Lake after several minutes:

```
echo "test message [date_time] from [hostname|host_ip]" | nc localhost 514
```

If your site collector does not appear in the list and the test message did not reach its destination, review the known common scenarios in Troubleshoot for Exabeam Site Collector that can be resolved immediately.

### 5.4. Upgrade Site Collector for Exabeam SaaS Advanced Analytics-only Deployments

The following instructions are for an Exabeam Site Collector upgrade if your logs are sent to Exabeam's SaaS Advanced Analytics and where there is no Exabeam Data Lake deployed.

1. Ensure your environment has met all requirements before running a site collector upgrade.
2. Download SaaS Site Collector installation files from the Exabeam Community.
3. Download your authentication file package using the following URL template based on your `<instanceID>`.

   ```
   https://<instanceID>.aa.exabeam.com/api/setup/saas/authPackage
   ```

4. Place the files in the `/tmp` directory of the site collector host.
5. Start a terminal session to the site collector and initiate a screen session.

   ```
   screen -LS [yourname]_[todaysdate]
   ```

6. Go to the `/tmp` directory and unpack the downloaded files .

   ```
   cd /tmp
   tar -xzf <filename>.tar.gz
   ```

7. Go to the `Exabeam_Site_Collector` directory.

   ```
   cd Exabeam_Site_Collector
   ```

8. Make the files executable.

```
chmod +x site-collector-installer.sh
```

9. Based on your deployment environment, please execute one of the following upgrade commands:
   a. Upgrade site collector behind the proxy with OpenVPN

```
sudo ./site-collector-installer.sh -v --aa-saas --upgrade --config=/tmp/
<instanceID>-auth-package.tgz --openvpn --proxy=<proxy_host_ip|
proxy_hostname> --proxy-port=<proxy_port>
```

   b. Upgrade site collector behind the proxy without OpenVPN

```
sudo ./site-collector-installer.sh -v --aa-saas --upgrade --config=/tmp/
<instanceID>-auth-package.tgz --proxy=<proxy_host_ip|proxy_hostname> --
proxy-port=<proxy_port>
```

   c. Upgrade site collector without proxy with OpenVPN

```
sudo ./site-collector-installer.sh -v --aa-saas --upgrade --config=/tmp/
<instanceID>-auth-package.tgz --openvpn
```

   d. Upgrade site collector without proxy and without OpenVPN

```
sudo ./site-collector-installer.sh -v --aa-saas --upgrade --config=/tmp/
<instanceID>-auth-package.tgz
```

10. You will not be able to view the status or health of the site collector in the Advanced Analytics console. The Status page is intended to show errors only and should not be used to verify throughput immediately after upgrading.
    Site collector operational checks must be run at the site collector host. You can send a test message via syslog and confirm it arrived at the destination, using:

```
echo "test message [date_time] from [hostname|host_ip]" | nc localhost 514
```

    If the test message did not reach its destination, review the known common scenarios in Troubleshoot for Exabeam Site Collector that can be resolved immediately.

# 6. Advanced Exabeam Site Collector Customizations

Exabeam Site Collector's standard features and configurations are usually enough to get your log ingestion and forwarding working. However, your deployment may require customizations.

Here are some customizations you can implement in your site collector:

- Log ingest filtering (generic)
- Log ingest from an external Kafka source
- Log forwarding filtering
- Customizable metadata tags in fields and logs
- Configure TLS Syslog ingestion
- Other supported configurations

## 6.1. Configure Transport Layer Security (TLS) Syslog Ingestion

For Hardware and Virtual Deployments Only

Exabeam Site Collector supports TLS Syslog ingestion. Using TLS certificates, you can implement a whitelist in your deployment.

1. Open port 515 at your firewall to log traffic.
2. Replace existing default authentication certificates with TLS certificates. Default certificates are located at:
   `/opt/logstash/inbound-certs/exa-ca.pem`
   `/opt/logstash/inbound-certs/syslog-inbound-key.pem`
   `/opt/logstash/inbound-certs/syslog-inbound-cert.pem`
3. Restart `logstash` services to apply the certificates.

   ```
   sudo systemctl restart logstash
   ```

## 6.2. Direct Kafka Input to Exabeam Site Collector

Kafka sources may reside in various locations in and out of your deployed environment that is protected behind a firewall. You can leverage Kafka for variety of use cases, including having it used as an interim logs storage/distribution point. A Kafka agent is an efficient means to send logs to a site collector to ingest and forward to SaaS or on-premises Data Lake deployments.

Before implementing an external Kafka source, confirm your environment meets the following criteria:

- No proxy services are in use between the site collector and Kafka source
- Your Kafka source has a network connection for data traffic to the site collector
- A site collector has been installed at the site collector host (For more information on site collector installation, see Install Exabeam Site Collector)
- Data is shipped in JSON or plain text format only

- Data is shipped without Kafka headers
- Messages are less than 1 Mb in size
- Kafka 1.1, and later, is in use

Supported deployment types:

- Kafka message ingestion without authentication
- TLS-configured Kafka using certificates, but without client login-password authentication

**IMPORTANT**
Not all TLS configurations are supported, please verify by consulting an Exabeam technical representative.

**NOTE**
Compression is used depending on the external Kafka configuration. Messages can be set with:

- GZip
- Snappy
- No compression

### 6.2.1. GET THE INSTALLATION PACKAGES

Ensure you have latest and matching Exabeam Site Collector installer version for you deployment. For more information, see Install Exabeam Site Collector.

Unpack the Exabeam Site Collector installer package at your Kafka host. See all available installation options with the `help` command:

```
sudo <Exabeam_Site_Collector_installer>/bin/external-kafka.sh --help
```

### 6.2.2. GENERATE AUTHENTICATION CERTIFICATES FOR SSL CONNECTION

If the connection between the Kafka source and site collector is to use SSL, generate authentication certificates before you start the Kafka installation. Authentication certificates need to be generated at the Kafka host. Copy the store files to the site collector host.

1. Run the `gen-kafka-certs.sh` at the external Kafka host (script found Exabeam Site Collector installer directory at the site collector).

   **WARNING**
   Generating new key and trust stores will affect existing authentication configurations. Therefore reconfigure existing SSL connections before running this script.

   ```
   #warning: reconfig >>
   sudo ./bin/gen-kafka-certs.sh
   ```

   A successful executed script will produce the following message:

```
Certificates generation process finished
Kafka CA certificate: .../kafka-ca.pem
Kafka client certificate: .../kafka-cert.pem
Kafka client key: .../kafka-key.pem
Kafka keystore file: .../kafka.keystore.jks
Kafka truststore file: .../kafka.truststore.jks
```

The keystore/truststore password used to generate the files is found in `gen-kafka-certs.sh`. Replace the default password before running this script.

```
cat gen-kafka-certs.sh | grep password=
```

2. Five files are generated. Copy the generated PEM files to the site collector host. (JKS files remain in place.)

```
kafka-ca.pem          # root certificate
kafka-cert.pem        # client cert
kafka-key.pem         # client key
kafka.keystore.jks    # kafka keystore
kafka.truststore.jks  # kafka truststore
```

### 6.2.3. INSTALL AN EXTERNAL KAFKA SERVICE

You should have the following:

- An unpacked copy of Exabeam Site Collector installation package at the Kafka host
- Names of Kafka topic(s) to subscribe to
- For SSL connections, copy authentication certificate files to the site collector host (see Generate Authentication Certificates for SSL Connection)
  - `kafka-ca.pem` (root certificate)
  - `kafka-cert.pem` (client certificate)
  - `kafka-key.pem` (client key)

Run the installation steps that best apply to your deployment environment and data flow:

- Install Kafka Message Ingestion Without Authentication
- Install SSL Connection with Server Verification on the Client Side
- Install SSL Connection with Client Authorization on the Server Side

#### *6.2.3.1. Install Kafka Message Ingestion Without Authentication*

Use this installation method if your environment does not need or support encrypted connections.

1. Configure a plaintext listener at the external Kafka host. Edit the following parameters in `kafka/config/server.properties`.

```
listeners=PLAINTEXT://0.0.0.0:9092
advertised.listeners=PLAINTEXT://<kafka_hostname|kafka_ip>:9092
```

2. Restart Zookeeper and Kafka services at the Kafka host to apply configurations. Verify that configurations are correct by verifying in the logs of both services at the Kafka host.
3. Run the following command at the site collector host in the Exabeam Site Collector installer directory:

```
sudo ./bin/external-kafka.sh --name=<name> --kafka-
hosts=<addr:port,addr:port> --kafka-topics=<topic1,topic2>

# Where the parameters are:
#  --name=<name>                          The unique name of Kafkabeat for
External Kafka (it can only contain upper and lowercase letters, and
numbers)
#  --kafka-hosts=<addr:port,addr:port>    Coma-separated list of External
Kafka brockers
#  --kafka-topics=<topic1,topic2>         Coma-separated list of External
Kafka topics
```

#### *6.2.3.2. Install SSL Connection with Server Verification on the Client Side*

This installation method requires `kafka-ca.pem` authentication file generated at the external Kafka host.

1. Configure the port at the external Kafka host. Edit the following parameters in `kafka/config/server.properties` for SSL with server verification on the client side.

```
listeners=PLAINTEXT://0.0.0.0:9092,SSL://0.0.0.0:9093
advertised.listeners=PLAINTEXT://<kafka_hostname|kafka_ip>:9092,SSL://
<kafka_hostname|kafka_ip>:9093
```

2. Configure the SSL options in `kafka/config/server.properties` at the external Kafka host.

```
security.protocol=SSL
ssl.client.auth=<none>
ssl.keystore.location=<full_path_to_kafka.keystore.jks>
ssl.keystore.password=<keystore_password>  # password used to generate file
ssl.truststore.location=<full_path_to_kafka.truststore.jks>
ssl.truststore.password=<truststore_password> # password used to generate
file
```

   Here is an example configuration with Kafka host paths, using server-based verification:

```
security.protocol=SSL
ssl.client.auth=none
ssl.keystore.location=/home/exabeam/certs/kafka.keystore.jks
ssl.keystore.password=test1234
ssl.truststore.location=/home/exabeam/certs/kafka.truststore.jks
ssl.truststore.password=test1234
```

3. Restart Zookeeper and Kafka services at the Kafka host to apply configurations. Verify that configurations are correct by verifying in the logs of both services at the Kafka host.

4. Run the following command at the site collector host in the Exabeam Site Collector installer directory:

```
sudo ./bin/external-kafka.sh --install --name=<connection_name> --kafka-
hosts=<kafka_hostname|kafka_ip>:9093 --kafka-topics=<kafka_topic> --
certificate-authority=/<full_path>/kafka-ca.pem
```

   Here is an example of an installation:

```
sudo ./bin/external-kafka.sh --install --name=test1 --kafka-
hosts=your.host.name:9093 --kafka-topics=your.topic --certificate-
authority=/path/to/kafka-ca.pem
```

#### *6.2.3.3. Install SSL Connection with Client Authorization on the Server Side*

This installation method requires `kafka-ca.pem`, `kafka-key.pem`, and `kafka-cert.pem` authentication file generated at the external Kafka host.

1. Configure the port at the external Kafka host. Edit the following parameters in `kafka/config/server.properties`.

```
listeners=PLAINTEXT://0.0.0.0:9092,SSL://0.0.0.0:9093
advertised.listeners=PLAINTEXT://<kafka_hostname|kafka_ip>:9092,SSL://
<kafka_hostname|kafka_ip>:9093
```

2. Configure the SSL options in `kafka/config/server.properties` at the external Kafka host.

```
security.protocol=SSL
ssl.client.auth=<required>
ssl.keystore.location=<full_path_to_kafka.keystore.jks>
ssl.keystore.password=<keystore_password>  # password used to generate file
ssl.truststore.location=<full_path_to_kafka.truststore.jks>
ssl.truststore.password=<truststore_password> # password used to generate
file
```

   Here is an example configuration with Kafka host paths, using server-based verification:

```
security.protocol=SSL
ssl.client.auth=required
ssl.keystore.location=/home/exabeam/certs/kafka.keystore.jks
ssl.keystore.password=test1234
ssl.truststore.location=/home/exabeam/certs/kafka.truststore.jks
ssl.truststore.password=test1234
```

3. Restart Zookeeper and Kafka services at the Kafka host to apply configurations. Verify that configurations are correct by verifying in the logs of both services at the Kafka host.
4. Run the following command at the site collector host in the Exabeam Site Collector installer directory:

```
sudo ./bin/external-kafka.sh --install --name=<connection_name> --kafka-
hosts=<kafka_hostname|kafka_ip>:9093 --kafka-topics=<kafka_topic> --
```

```
certificate=/<full_path>/kafka-cert.pem --certificate-authority=/
<full_path>/kafka-ca.pem --key=/<full_path>/kafka-key.pem
```

Here is an example of an installation:

```
sudo ./bin/external-kafka.sh --install --name=test1 --kafka-
hosts=your.host.name:9093 --kafka-topics=your.topic --certificate-
authority=/path/to/kafka-ca.pem --certificate=/path/to/kafka-cert.pem --
key=/path/to/kafka-key.pem
```

### 6.2.4. GET A LIST OF EXTERNAL KAFKA SOURCES

Run the script with the `-list` flag. For example:

```
sudo <Exabeam_Site_Collector_installer>/bin/external-kafka.sh -list
```

### 6.2.5. UNINSTALL AN EXTERNAL KAFKA SERVICE

1. Use `external-kafka.sh --uninstall` to remove the Kafka service on the host.

   ```
   sudo ./bin/external-kafka.sh --uninstall --name=<kafka_broker_name>
   ```

   Here is an example of an uninstall instruction:

   ```
   sudo ./bin/external-kafka.sh --uninstall --name=test1
   ```

   A successful uninstall will produce messages like:

   ```
     Parsing current options

        - Action:  uninstall

        - Name:  test1

     Uninstalling...

        - Uninstalling External Kafka test1...

        - Uninstalling External Kafka manager for test1 ...

        - Deregister manager config:  /opt/exabeam/beats/test1/manager

        - Deregister manager agent:  abbdd3e5b92440899a44315c0bf9d56a

        - Uninstalling External Kafka worker for test1 ...

   [Removing the Kafkabeat for External Kafka test1 is done!]
   ```

2. Reset the listener port configuration at the external Kafka host.

### 6.2.6. TROUBLESHOOT KAFKA SERVICES

- If no data has been sent or received, verify that the site collector is running at the external Kafka host.

```
sudo systemctl status exabeam-kafka-<connecttion_alias>-collector
```

- If messages stop abruptly, inspect the logs at the Kafka host for error messages.

```
sudo cat /opt/exabeam/beats/<hostname>/worker/logs/kafkabeat
```

- If the Kafka log is larger than 1 MB, enable log truncation by editing the `processors` parameter in `/opt/exabeam/beats/<hostname>/worker/kafkabeat.yml`. Set the `max_bytes` to `1000000` bytes. Alternatively, but not at the same time, you can limit the event log size by the number of characters by editing `max_characters`.

```
processors:
- truncate_fields:
    fields:
    - message
    max_bytes: 1048576
    max_characters: 1000
```

- Verify that parameters used during installation are applicable for your environment. Review the full list of options using `<Exabeam_Site_Collector_installer>/bin/external-kafka.sh --help`:

```
-help                                  Print this help section
-uninstall                             Uninstall the Kafkabeat for External
Kafka
-list                                  List all the installed Kafkabeats for
External Kafka
-name=<name>                           The unique name of Kafkabeat for
External Kafka (it can only contain lowercase letters, and numbers)
-kafka-hosts=<addr:port,addr:port>     Coma-separated list of External Kafka
brockers
-kafka-topics=<topic1,topic2>          Coma-separated list of External Kafka
topics
-certificate-authority=<path>          The path to the certificate authority
file (*.pem) that is used in Kafka SSL configuration to verify the SSL
connection with Kafka server
-certificate=<path>                    The path to the certificate file
(*.pem) that is used in Kafka SSL configuration for client certificate
authorization (must be used with -key flag)
-key=<path>                            The path to the key file (*.pem) that
is used in Kafka SSL configuration for client certificate authorization
```

## 6.3. Filter Incoming Syslog Events in Exabeam Site Collector

For known threats in high-volume log scenarios, you can apply a filter on inbound syslog events to reduce the amount of data sent to your Exabeam SaaS deployment. You will edit the input configuration

file with the filter string. Filtering will capture syslog events with known threat patterns. It will not capture events using "not match" parameters (for example, filtering will not capture for "`X != event`").

1. SSH to the host of your site collector and create a back up of the configuration file to your home folder.

   ```
   cp   /opt/logstash/conf.d/syslog2kafka.conf  ~/01-syslog-input.conf.orig
   ```

2. Append the `filter` code to `/opt/logstash/conf.d/syslog2kafka.conf`. In this example, the filter uses `"drop-string-"` and `"drop-string-2"` as the filter.

   ```
   filter {
       if "drop-string-" in [message] or “drop-string-2" in [message] {
           drop { }
       }
   }
   ```

3. Restart the syslog service.

   ```
   sudo systemctl restart logstash

   # Verify service status
   sudo systemctl status logstash
   ```

   Use `sudo journalctl -e -u logstash` to display the syslog status log. A successful restart will resemble the log records below, ending with a `Successfully started` record.

   ```
   Mar 25 20:23:57 dl docker[21481]: [2019-03-25T20:23:57,726][INFO ]
   [logstash.inputs.tcp      ] Starting tcp input listener
   {:address=>"0.0.0.0:514"}
   Mar 25 20:23:57 dl docker[21481]: [2019-03-25T20:23:57,731][INFO ]
   [logstash.inputs.tcp      ] Starting tcp input listener
   {:address=>"0.0.0.0:515"}
   Mar 25 20:23:58 dl docker[21481]: [2019-03-25T20:23:58,229][INFO ]
   [logstash.pipeline        ] Pipeline main started
   Mar 25 20:23:58 dl docker[21481]: [2019-03-25T20:23:58,233][INFO ]
   [logstash.inputs.udp      ] Starting UDP listener {:address=>"0.0.0.0:514"}
   Mar 25 20:23:58 dl docker[21481]: [2019-03-25T20:23:58,253][INFO ]
   [logstash.inputs.udp      ] UDP listener started
   {:address=>"0.0.0.0:514", :receive_buffer_bytes=>"106496", :queue_size=>"200
   0"}
   Mar 25 20:23:58 dl docker[21481]: [2019-03-25T20:23:58,268][INFO ]
   [logstash.agent           ] Successfully started Logstash API endpoint
   {:port=>9600}
   ```

4. Verify the filter is working by sending the filter string to site collector's ingest port. In this example, the filter uses `drop-string-` and `drop-string-2` as the filter.

   ```
   logger -n localhost -T -P 514 test message from other system 1; logger -n
   localhost -T -P 514 test message from other system 2;logger -n localhost -T
   ```

```
-P 514 test message drop-string-1 1; logger -n localhost -T -P 514 test
message drop-string-2 1; logger -n localhost -T -P 514 test message drop-
string-1 2; logger -n localhost -T -P 514 test message drop-string-2 2;
logger -n localhost -T -P 514 test message from other system 3;
```

### 6.4. Filtering Outbound Logs in Exabeam Site Collector

Exabeam Site Collector supports log filtering before uploading the logs to configured destinations. The site collector can drop entire events if filtering conditions are matched. Site collectors use Kafkabeat for the outbound message processing. Configurations are made in `/opt/exabeam/beats/<TARGET>/worker/kafkabeat.yml`, where <TARGET> is the folder that contains the site collector destination setup.

The values in <TARGET> are the following options based on your deployment:

- For SaaS: `gcs1`
- For Data Lake on-premises: `lms1`
- For Advanced Analytics on-premises: `uba1`
- For custom destinations (with unique `destinationIDs` for each target): `<destinationID>`

The configuration resembles the following configuration block:

```
kafkabeat:
  inputs:
  - type: kafka
    ...
logging:
  ...
output:
  ...
path:
  ...
queue:
  ...
processors:
- drop_event:
    when:
      - <condition>
```

After updating configurations, restart Kafkabeat with the following command:

```
sudo systemctl restart exabeam-kafka-<TARGET>-collector
```

Kafkabeat does not parse the event. Rather, it sends the event as-is to the destination. Filtering is based solely on the `message` field of an event. Listed in the following table are the five conditions in the message field that are supported:

| Condition | Description | Condition Example |
|---|---|---|
| `contains` | Checks if a value is part of a field. The field can be a string or an array of strings. The condition accepts only a string value. | Check if an error is part of the event message<br>`contains:`<br>`  message: "Specific error"` |
| `regexp` | Checks the field against a regular expression. The condition accepts only strings. | Check if the event `message` has the src_ip that fits the IP range 192.168.0.1/16 (192.168.0.0 - 192.168.255.255)<br>`regexp:`<br>`  message: "src_ip=192\\.168\\.\\d{1,3}\\.\\d{1,3}"` |
| `or` | An operator that receives a list of criteria to match a single condition. | Determines match to either `message` criteria in `message ~= [DEBUG]` or `message ~= [TRACE]`<br>`or:`<br>`  - contains:`<br>`  message: "[DEBUG]"`<br>`  - contains:`<br>`  message: "[TRACE]"` |
| `and` | An operator that receives a list of criteria to match all of. | Determines match where both `http.response.code = 200` and `status = OK`<br>`and:`<br>`  - equals:`<br>`  http.response.code: 200`<br>`  - equals:`<br>`  status: OK` |
| `not` | An operator that receives the condition to negate. | Drops event if `message` does not match `message ~= [ERROR]`<br>`not`<br>`  contains:`<br>`  message: "[ERROR]"` |

Here is an example:

We need to filter out logs from a particular Filebeat that sends logs from a given IP address. The filter will match logs that contain the text `[OBSOLETE]` in content or that comes from the Filebeat with ID `573d5253-4e4e-4fff-92a5-8f2f227b3af1` and IP address `src_ip=195.164.*.*`.

A sample log resembles:

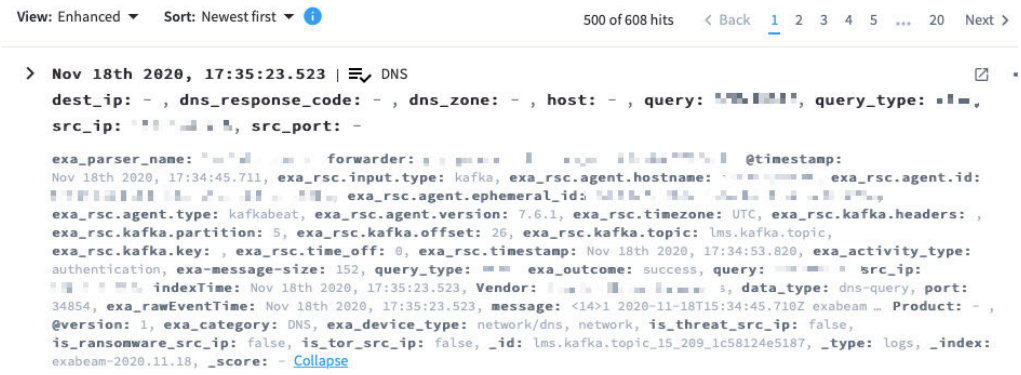
```
[OBSOLETE] - Mar 24 15:00:34 2020 rt=1585062034 device=110.90.230.153
name=Stephanie Kim
```

The filter is written to apply all three possible matching criteria:

```
processors:
- drop_event:
    when:
      or:
      - and:
        - contains:
            message: "\"id\":\"573d5253-4e4e-4fff-92a5-8f2f227b3af1\""
        - regexp:
            message: "src_ip=195\\.163\\.\\d{1,3}\\.\\d{1,3}"
      - contains:
          message: "[OBSOLETE]"
```

### 6.5. Metadata Tags in Exabeam Site Collector Fields and Logs

By default, events processed through the Exabeam Site Collector contain metadata. They have the prefix `exa_rsc`, such as in `exa_rsc.input.type`, `exa_rsc.agent.hostname`, and `exa_rsc.timezone`. Here is an example event with default tags in the payload.




### 6.6. Add OpenVPN After Exabeam Site Collector Installation

If you have an existing and running Exabeam Site Collector on your host and you need to add OpenVPN, you can rerun the installation script with express instruction to apply the OpenVPN only. This avoids reinstalling the entire site collector package.

You will use the same installation script for new installations, upgrades, and feature add-ons. (Go to the instructions for your deployment in Install Exabeam Site Collector or Upgrade Exabeam Site Collector to see package download steps.)

In the directory where the installation file has been placed, use the following to add OpenVPN to the existing and running site collector hosts:

```
./site-collector-installer.sh -v --dl-saas --feature=openvpn
```

If a reinstallation of OpenVPN is needed, use:

```
./site-collector-installer.sh -v --dl-saas --feature=openvpn --reinstall
```

### 6.7. Supported Exabeam Site Collector Changes

For a list of all options, use `site-collector-installer.sh --help`.

Below are all supported changes that you can make. For other changes, please contact Exabeam Customer Success for further guidance.

- Operating system updates
- Site Collector server IP changes
  - Once the IP of the server has been changed, edit the below line in the file: `/opt/kafka/config/server.properties`
  - `advertised.listeners=EXTERNAL_PLAINTEXT://MYIP:9092, EXTERNAL_SSL://MYIP:9093, INTERNAL_SSL://localhost:9094`

    - Then, restart Kafka: `sudo systemctl restart kafka`
- Log retention change (default is 24 hours)
    - Edit the below line in the file: `/opt/kafka/config/server.properties`
    - `log.retention.hours=24`
    - Then, restart Kafka: `sudo systemctl restart kafka`
- RAM allocation to logstash
    - Edit the below lines appropriately in this file: `/opt/logstash/config/jvm.options`
        - `-Xms16g`
        - `-Xmx16g`
    - Then, restart Kafka: `sudo systemctl restart logstash`
- RAM allocation to Kafka
    - Edit the numbers before the 'G' in this file: `/opt/kafka/bin/kafka-server-start.sh`
    - `export KAFKA_HEAP_OPTS="-Xmx5G -Xms5G"`
    - Then, restart Kafka: `sudo systemctl restart kafka`

# 7. Troubleshoot for Exabeam Site Collector

Below are troubleshooting steps for common scenarios found in the Exabeam Site Collector.

In order to run commands given in this chapter, you must be able to log into site collector host and start a terminal session. You will initiate a screen session to prevent termination of your session.

```
screen -LS [yourname]_[todaysdate]
```

If the scenarios do not apply to your issue, please capture diagnostics data and contact Exabeam Customer for further assistance.

## 7.1. Scenario 1: Collector or its status does not appear in the console and no logs reach destination

After installation, update, or during loss of data throughput, if the site collector does not appear in the Exabeam SaaS Status page or Exabeam Data Lake Collectors list (navigate to **Settings** > **Collector Management** > **Collectors**), verification must be run at the site collector host to ensure necessary services are running and there is throughput.

Run the following command to check all Exabeam Site Collector Services:

```
sudo /opt/exabeam/tools/sc-services-check.sh
```

Here is an excerpt from the response of a working site collector where Datadog and OpenVPN is not deployed:

```
Check all Site Collector services

Check Zookeeper service...
●zookeeper.service
Loaded: loaded (/etc/systemd/system/zookeeper.service; enabled; vendor preset:
disabled)
Active: active (running) since ...
Main PID: 5887 (java)
CGroup: /system.slice/zookeeper.service●5887 java -Xmx512M -Xms512M -server -
XX:+UseG1GC -XX:MaxGCPauseMillis=20 -XX:InitiatingHeapOccupancyPercent=35 -
XX:+ExplicitGCInvokesConcurrent -Djava.awt.headless=true -Xloggc:/opt/kafka/
bin/../logs/zookeeper-gc.log -verb...Mar 09 18:28:17 centos7 zookeeper-server-
start.sh[5887]: [2021-03-09 18:28:17,472] INFO Processed session termination
for sessionid: 0x178183f22620001
(org.apache.zookeeper.server.PrepRequestProcessor)...

Check Kafka service...
●kafka.service
Loaded: loaded (/etc/systemd/system/kafka.service; enabled; vendor preset:
disabled)
Active: active (running) since ...
Main PID: 6028 (java)
CGroup: /system.slice/kafka.service●6028 java -Xmx4G -Xms4G -server -XX:+UseG1GC
```

```
-XX:MaxGCPauseMillis=20 -XX:InitiatingHeapOccupancyPercent=35 -
XX:+ExplicitGCInvokesConcurrent -Djava.awt.headless=true -Xloggc:/opt/kafka/
bin/../logs/kafkaServer-gc.log -verbos...Mar 09 18:28:39 centos7 sh[6028]:
[2021-03-09 18:28:39,206] INFO [Partition lms.kafka.topic-2 broker=1]
lms.kafka.topic-2 starts at Leader Epoch 0 from offset 0. Previous Leader Epoch
was: -1 (kafka.cluster.Partition)...

Check Logstash service...
●logstash.service
Loaded: loaded (/etc/systemd/system/logstash.service; enabled; vendor preset:
disabled)
Active: active (running) since …
Main PID: 7244 (java)
CGroup: /system.slice/logstash.service●7244 /bin/java -Xms4g -Xmx4g -cp /opt/
logstash/logstash-core/lib/jars/commons-compiler-3.0.8.jar:/opt/logstash/
logstash-core/lib/jars/google-java-format-1.1.jar:/opt/logstash/logstash-
core/lib/jars/guava-19.0.jar:/opt/logs...
Mar 09 18:28:58 centos7 sh[7244]: transactional.id = null
Mar 09 18:28:58 centos7 sh[7244]: value.serializer = class
org.apache.kafka.common.serialization.StringSerializer...

DataDog is not installed...
OpenVPN is not installed...

Check SC Forwarder service...
●exabeam-rsc-forwarder.service - exabeam-rsc-forwarder
Loaded: loaded (/etc/systemd/system/exab
```

Review the output from your site collector and to see if any required services have failed to load. Ensure that the parameters for each service shown in the response is correct. If you have services that failed to load, run diagnostics and send the output to Exabeam for further assistance. If no services failed to load, then determine whether there are network issues impeding data throughput.

### 7.2. Scenario 2: Collector is healthy but no logs are transmitted or received

If using OpenVPN, check the OpenVPN client logs for error messages, then check if the OpenVPN client is running on site collector server.

```
# check status of openvpn client
sudo systemctl status openvpn@<instanceID>

# check logs of openvpn client
sudo journalctl -fu openvpn@<instanceID>
```

Resolve the issues cited in the error messages. Logs should start appearing without an OpenVPN service restart.

If you have confirmed that the network is allowing traffic yet no logs are being sent or received at the configured destinations, run diagnostics and send the output to Exabeam for further assistance.

### 7.3. Scenario 3: Exabeam Advanced Analyticsunable to pull LDAP data

Firewall rules may not be applied properly (for example, those in SE Linux) for your environment. Run the following command:

```
firewall-cmd --zone=public --add-forward-
port=port=389:proto=tcp:toport=389:toaddr=$<dns_server> --permanent
firewall-cmd –reload
```

Generate LDAP context again and the data should become available without restarting additional services.

If the above command still fails to resolve the issue, run the following alternative command:

```
firewall-cmd --add-forward-port=port=389:proto=tcp:toport=389:toaddr=
$<dns_server> --permanent
firewall-cmd –reload
```

**NOTE**
Port 389, given in this example, is used for LDAP. It may be different for your organization. Please confirm with your organization's network configuration.

### 7.4. Capture Site Collector Diagnostics Using Exabeam Support Package

The Exabeam Customer Success team may require additional data to further assist in troubleshooting. Exabeam has provided a script to support capturing diagnostic information about your site collector. Please run the command based on your situation:

- To collect data just after a site collector installation:

  ```
  sudo /tmp/Exabeam_Site_Collector/bin/support-package.sh
  ```

- To collect data from a running site collector:
  - For site collector versions before 2.1.11

    ```
    sudo /opt/exabeam/bin/support-package.sh
    ```

  - For site collector version 2.1.11 and later

    ```
    sudo /opt/exabeam/tools/support-package.sh
    ```

# Install and Upgrade Exabeam Site Collector for On-premises and Legacy Deployments

Configure and gather information as outlined in this section before attempting to install or upgrade your site collector.

## 8.1. Prerequisites

Ensure your environment met all requirements before running a site collector installation. Please review prerequisites listed in Install Exabeam Site Collector in addition to the following:

- Have the following information for all log sources that will send data to the site collector:
  - Product and vendor
  - Hostname and IP address
  - Network zone of the log source
  - Ingest method and access port
  - Log throughput capacity in events per second (EPS)
  - Log storage capacity in GB
  - Associated site collector
- Routes through firewalls and proxies are not supported in on-premises deployments

### 8.1.1. ON-PREMISES INSTRUCTIONS BY DEPLOYMENT

Select the instructions that best matches your deployment environment:

Install Site Collector for Exabeam Data Lake On-premises Deployments

Installing Site Collector for Exabeam Advanced Analytics On-premises Deployments

Upgrade Site Collector for Exabeam Data Lake On-premises Deployments

Upgrade Site Collector for Exabeam Advanced Analytics On-premises Deployments

### 8.1.2. LEGACY DEPLOYMENTS

For site collectors in versions 1.0.0/1.0.3 and those before version 202004.1, use the following instructions to migrate services.

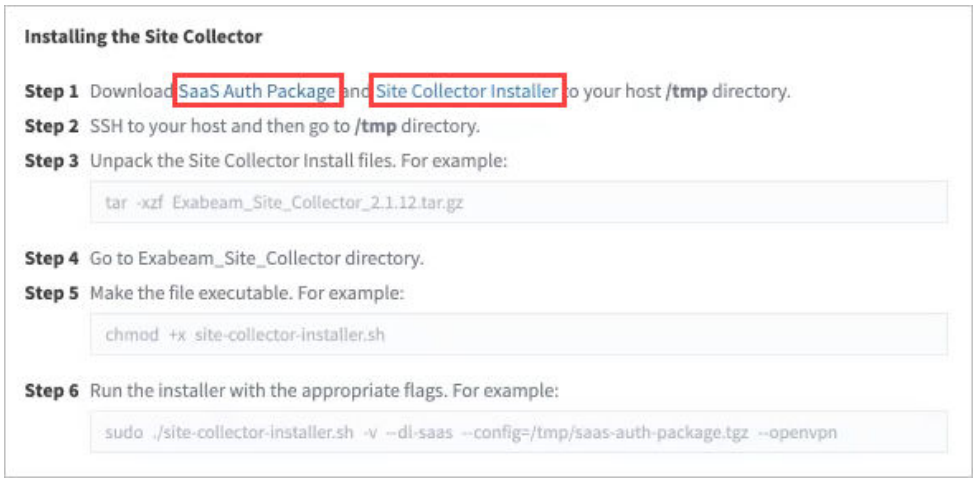
> ⚠ **WARNING**
> **For CentOS deployments** -- As CentOS 8.x will be reaching its end-of-life (December 31, 2021), we strongly recommend deploying site collectors on CentOS 7.x.

## 8.2. Install Site Collector for Exabeam Data Lake On-premises Deployments

For Data Lake in Appliance or Virtual Deployments Only

Follow these instructions for a fresh Exabeam Site Collector installation if your logs are to be sent to Exabeam Data Lake destination deployed on an appliance or virtual platform (excluding Exabeam SaaS Cloud).

1. Ensure your environment has met all requirements before running a site collector installation.
2. In Data Lake, navigate to **Settings** > **Collector Management** > **Collectors**.
3. Click + ADD COLLECTOR to open the **Collector Artifacts** menu to get a list of **Site collectors**.




4. Download the **Site Collector Auth Package** and **Site Collector Installation Package**. These packages contain all required configurations and authentication data needed to access your SaaS tenant and installation package.
5. Use `scp` (secure copy) to place the files in the `/tmp` directory of the site collector host. (For help with this command, run `man scp`.

```
scp <source_host>:<directory>/<package_file> <site_collector>:<directory>/
package_file>
```

6. Start a new terminal session using the an account with administrator rights. Initiate a screen session. This is mandatory and will prevent accidental termination of your session.

```
screen -LS [yourname]_[todaysdate]
```

7. Go to the `/tmp` directory and unpack the installation package only.

```
cd /tmp
tar -xzf <install_filename>.tar.gz
```

8. Go to the `Exabeam_Site_Collector` directory.

```
cd Exabeam_Site_Collector
```

9. Make the files executable.

```
chmod +x site-collector-installer.sh
```

10. Run following installation commands:

```
sudo ./site-collector-installer.sh -v --dl-on-prem --config=/tmp/sc-auth-
package.tgz
```

11. Once installation is complete, the prompt will return `Site collector installer complete`.
12. To verify that the site collector source has been installed, log into the Data Lake and navigate to **Settings** > **Collector Management** > **Collectors** to see the list of configured collectors.


> **NOTE**
> It is normal to find the **Site Collector Data Forwarder** service is shown as `Stopped` while another service is shown as `Running`. To verify if there is on-going ingestion, one of these services will show non-zero messages in the graph.

You can also send a test message via syslog and confirm it arrived at the destination via Data Lake after several minutes:

```
echo "test message [date_time] from [hostname|host_ip]" | nc localhost 514
```

If your site collector does not appear in the list and the test message did not reach its destination, review the known common scenarios in Troubleshoot for Exabeam Site Collector that can be resolved immediately.

## 8.3. Installing Site Collector for Exabeam Advanced Analytics On-premises Deployments

Follow these instructions for a fresh Exabeam Site Collector installation if your logs are to be sent to Exabeam Advanced Analytics destination deployed on an appliance or virtual platform (excluding Exabeam SaaS). In this configuration, you will not be able to view the status or health of the site collector in the Advanced Analytics console.

1. Ensure your environment has met all requirements before running a site collector installation.
2. Download SaaS Site Collector installation files from the Exabeam Community.
3. Place the files in the `/tmp` directory of the site collector host.
4. Start a new terminal session using the an account with administrator rights. Initiate a screen session. This is mandatory and will prevent accidental termination of your session.

```
screen -LS [yourname]_[todaysdate]
```

5. Go to the `/tmp` directory and unpack the downloaded file.

```
cd /tmp
tar -xzf <filename>.tar.gz
```

6. Go to the `Exabeam_Site_Collector` directory.

```
cd Exabeam_Site_Collector
```

7. Make the files executable.

```
chmod +x site-collector-installer.sh
```

8. Based on your expected load, execute one of the following installation commands:
   a. Installing site collector without EPS limit

```
sudo ./site-collector-installer.sh -v --aa-on-prem --aa-
listener=<listener_ip>:514
```

   b. Installing site collector with EPS limit

```
sudo ./site-collector-installer.sh -v --aa-on-prem --aa-
listener=<listener_ip>:514 --eps-limit=2048
```

9. Once installation is complete, the prompt will return `Site collector installer complete`.
10. Site collector operational checks must be run at the site collector host. You can send a test message via syslog and confirm it arrived at the destination, using:

```
echo "test message [date_time] from [hostname|host_ip]" | nc localhost 514
```

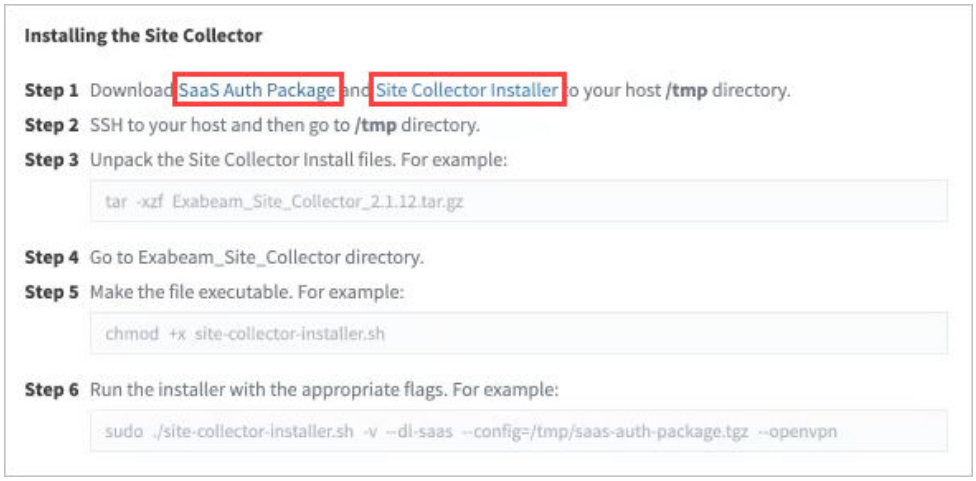
    If no logs arrive at the destination after a few minutes, review the known common scenarios in Troubleshoot for Exabeam Site Collector that can be resolved immediately.

## 8.4. Upgrade Site Collector for Exabeam Data Lake On-premises Deployments

For Data Lake in Appliance or Virtual Deployments Only

The following instructions are for an Exabeam Site Collector upgrade if your logs are sent to Exabeam Data Lake deployed on Exabeam hardware or virtual platform (excluding Exabeam SaaS).

1. Ensure your environment has met all requirements before running a site collector upgrade.
2. In Data Lake, navigate to **Settings** > **Collector Management** > **Collectors**.
3. Click + ADD COLLECTOR to open the **Collector Artifacts** menu to get a list of **Site collectors**.




4. Download the **Site Collector Auth Package** and **Site Collector Installation Package**. These packages contain all required configurations and authentication data needed to access your SaaS tenant and installation package.
5. Place the files in the `/tmp` directory of the site collector host.
6. Start a terminal session to the site collector and initiate a screen session.

   ```
   screen -LS [yourname]_[todaysdate]
   ```

7. Go to the `/tmp` directory and unpack the downloaded files.

   ```
   cd /tmp
   tar -xzf <filename>.tar.gz
   ```

8. Go to the `Exabeam_Site_Collector` directory.

   ```
   cd Exabeam_Site_Collector
   ```

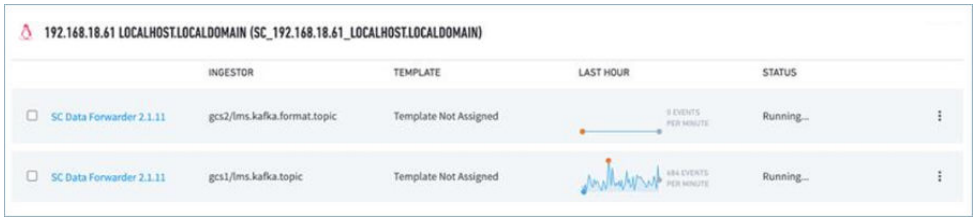
9. Make the files executable.

   ```
   chmod +x site-collector-installer.sh
   ```

10. Run following upgrade commands:

    ```
    sudo ./site-collector-installer.sh -v --dl-on-prem --upgrade --
    config=/tmp/sc-auth-package.tgz.tgz
    ```

11. To verify that the site collector source has been upgraded, log into the Data Lake and navigate to **Settings** > **Collector Management** > **Collectors** to see the list of configured collectors.



> **NOTE**
> It is normal to find the **Site Collector Data Forwarder** service is shown as `Stopped` while another service is shown as `Running`. One of these services will show non-zero messages in the graph if there is ongoing ingestion, which would be the indicator to verify with.

You can also send a test message via syslog and confirm it arrived at the destination via Data Lake after several minutes:

```
echo "test message [date_time] from [hostname|host_ip]" | nc localhost 514
```

If your site collector does not appear in the list and the test message did not reach its destination, review the known common scenarios in Troubleshoot for Exabeam Site Collector that can be resolved immediately.

## 8.5. Upgrade Site Collector for Exabeam Advanced Analytics On-premises Deployments

For Advanced Analytics in Appliance or Virtual Deployments, in Unmanaged Mode Only

The following instructions are for an Exabeam Site Collector upgrade if your logs are sent to Exabeam Advanced Analytics deployed on Exabeam hardware or virtual platform (excluding Exabeam SaaS Cloud).

1. Ensure your environment has met all requirements before running a site collector upgrade.
2. Download SaaS Site Collector installation files from the Exabeam Community.
3. Place the files in the `/tmp` directory of the site collector host.
4. Start a terminal session to the site collector and initiate a screen session.

   ```
   screen -LS [yourname]_[todaysdate]
   ```

5. Go to the `/tmp` directory and unpack the downloaded files.

   ```
   cd /tmp
   tar -xzf <filename>.tar.gz
   ```

6. Go to the `Exabeam_Site_Collector` directory.

   ```
   cd Exabeam_Site_Collector
   ```

7. Make the files executable.

   ```
   chmod +x site-collector-installer.sh
   ```

8. Based on your expected load, execute one of the following upgrade commands:
   a. Upgrade site collector without EPS limit

      ```
      sudo ./site-collector-installer.sh -v --aa-on-prem --upgrade --aa-
      listener=<listener_ip>:514
      ```

b. Upgrade site collector with EPS limit

```
sudo ./site-collector-installer.sh -v --aa-on-prem --upgrade --aa-
listener=<listener_ip>:514 --eps-limit=2048
```

9. Site collector operational checks must be run at the site collector host. You can send a test message via syslog and confirm it arrived at the destination, using:

```
echo "test message [date_time] from [hostname|host_ip]" | nc localhost 514
```

If the test message did not reach its destination, review the known common scenarios in Troubleshoot for Exabeam Site Collector that can be resolved immediately.

## 8.6. Migrate Legacy Site Collector to New Exabeam SaaS Site Collector

For site collectors in versions 1.0.0/1.0.3 and those before version 202004.1, the following instructions will guide you through redirecting feeds and agents in a step-wise manner to mitigate data loss. Please have a operating site collector host before attempting a migration.

> ⚠ **WARNING**
> Exabeam recommends that upgrading your site collector happen at a different host than the legacy site collector. This ensures there is no data loss in the transition. Otherwise, expect loss of data during the time between the shutdown of the legacy site collector and when the new site collector starts receiving data.

Exabeam recommends installing the new site collector on a host that is not running the existing site collector, using the following procedure:

1. Deploy a new site collector onto the dedicated host.
2. If there is a syslog source, we highly recommend that you install a load balancer to work with the new site collector to ensure high availability.
3. Switch syslog traffic to the new site collector or load balancer. Verify that messages pending in Kafka were successfully sent. Please ensure you have removed port forwarding rules for ports 5514 and 5515. Use ports 514 and 515 for Syslog ingestion.
4. Turn off Logstash services on the old site.
5. Direct feeds from agent collectors, if any, to the new site collector and restart them. Verify that messages pending in Kafka were successfully sent.
6. Let the old site collector continue to run so all remaining messages process to completion. Verify that messages pending in Kafka were successfully sent. The lag queue in Kafka must be 0.
7. Turn off the old site collector.

If you do not have an available host to install that is not currently running the existing site collector, you must first uninstall the legacy site collector and then install the new site collector.

# 9. Uninstall Exabeam Site Collector

It is assumed that logs are being held in queue, log collection or directed to another site collector (on a different host) while the uninstall process is happening. Otherwise, data loss will occur.

To uninstall the site collector, SSH to the host and apply the following command:

```
sudo ./site-collector-installer.sh -v --uninstall

# Or, uninstall in silent mode
sudo ./site-collector-installer.sh -v --uninstall -a
```

## 9.1. Uninstall a Legacy Site Collector

Once your Exabeam Site Collector has been established, you can uninstall legacy site collectors. During the uninstall process, there is a risk of data loss if ingest queues are not empty. An automatic uninstall script is available to minimize your risk:

1. Download the original Exabeam Site Collector installation package, either from
   - Exabeam Community or
   - In Exabeam Data Lake, go to **Settings** > **SaaS Management** > **SaaS Site Collectors** > click the link in **Step 1** of **Installing the Site Collector**.
2. Place the downloaded in the `/tmp` directory of the legacy site collector host.
3. Start a new terminal session using the an account with Administrator rights. Initiate a screen session. This is mandatory and will prevent accidental termination of your session.

   ```
   screen -LS [yourname]_[todaysdate]
   ```

4. Go to the `/binary` directory and make `uninstall-legacy-sc.sh` executable.

   ```
   cd /binary
   chmod +x uninstall-legacy-sc.sh
   ```

5. Stop ingestion service for the legacy collector.

   ```
   sudo systemctl stop exabeam-lms-server
   ```

6. Check for pending messages that need to be uploaded.

   ```
   sudo /opt/kafka/bin/kafka-consumer-groups.sh --group gcs --describe –
   bootstrap-server localhost1:9092 | grep lms.kafka.topic | awk "{ sum+=\$5}
   END { print sum}"
   ```

7. Run uninstall script from the `/binary` directory.

   ```
   sudo ./uninstall-legacy-sc.sh
   ```

## 10. Supported Exabeam Site Collector Changes

For a list of all options, use `site-collector-installer.sh --help`.

Below are all supported changes you can make. For other changes, please contact your Exabeam Customer Success at Exabeam Community for further guidance.

- Operating system updates
- Site collector server IP changes

  1. Once the IP of the server has been changed, edit the listener and SSL parameters in `/opt/kafka/config/server.properties` with the new IP.

     ```
     advertised.listeners=EXTERNAL_PLAINTEXT://<new_ip>:9092, EXTERNAL_SSL://
     <new_IP>:9093, INTERNAL_SSL://localhost:9094
     ```

  2. Restart Kafka.

     ```
     sudo systemctl restart kafka
     ```

- Log retention change (default is 24 hours)

  1. Edit `log.retention.hours` in `/opt/kafka/config/server.properties`.

     ```
     log.retention.hours=24
     ```

  2. Restart the `kafka` service.

     ```
     sudo systemctl restart kafka
     ```

- RAM re-allocation to `logstash`

  1. Edit the `-Xm` parameters in `/opt/logstash/config/jvm.options`, like the ones shown here:
     - `-Xms<ram>g`
     - `-Xmx<ram>g`
  2. Restart the `kafka`.

     ```
     sudo systemctl restart logstash
     ```

- RAM re-allocation to `kafka`

  1. Edit the `-Xm` parameters in the `export` variable in `/opt/kafka/bin/kafka-server-start.sh`.

```
export KAFKA_HEAP_OPTS="-Xmx<ram>G -Xms<ram>G"
```

2. Restart the `kafka` service.

```
sudo systemctl restart kafka
```

# Appendix A. Glossary of Terms

| Term | Definition |
|---|---|
| Active Directory | Microsoft directory services for Windows networks. |
| Advanced Analytics | Exabeam's Advanced Analytics provides user and entity behavior intelligence on top of existing SIEM and log management data repositories to detect threats by analyzing activities in the attack chain. |
| Cloud Connector | An ingestion mechanism that collects and uploads logs to Exabeam services from over cloud services such as AWS, Salesforce, and other cloud security, identity and access management, infrastructure and business applications. |
| Data Lake | An Exabeam log management system that orchestrates data collection, indexing, and visualization. |
| Exabeam Cloud Platform | A multi-tenant platform-as-a-service (PaaS) product that extends Exabeam’s security information and event management (SIEM) solution with capabilities unique to Exabeam along with cloud storage, data graphing and integrations |
| Kafka Log Collector | The ingestion point for syslogs where data is compressed for optimal transmission to site collectors. |
| Logstash | A collection engine that can data gather and normalize data from disparate sources for uniformed processing. |
| OpenVPN | An open-source virtual private network system that creates secure point-to-point or site-to-site network connections. |
| Primary Site Collector | Where there is a series of site collectors in a deployment, the Primary Site Collector is the master site collector with OpenVPN authentication to the log destination. |
| Watchdog | A monitoring daemon that ensures all critical services are operating. |