Spike in downloaded documents per user on Salesforce cloud

You might need to look for a sudden, high-volume increase in downloaded documents when doing the following:

- Protecting a Salesforce cloud deployment

Prerequisites

In order to execute this procedure in your environment, the following data, services, or apps are required:

- Splunk Enterprise or Splunk Cloud Platform
- CRM, ERP, and other business application data
- Splunk Add-on for Salesforce

Example

A sudden, high-volume increase in downloaded documents can indicate unauthorized, non-compliant, and potentially malicious behavior. Because so many people in your organization have access to Salesforce, this is an activity you want to monitor for regularly.

To optimize the search shown below, you should specify an index and a time range.

1. Populate the lookup_sfdc_usernames lookup provided by the Salesforce Add-on with live values from your site.
2. Run the following search:

```
|EVENT_TYPE=DocumentAttachmentDownloads
|lookup lookup_sfdc_usernames USER_ID
|bucket _time span=1d
|stats count BY Username _time
|stats count AS num_data_samples max(eval(if(_time => relative_time(maxtime, "-1d@d"), 'count',null))) AS count avg(eval(if(_time<relative_time(maxtime,"-1d@d"),'count',null))) AS avg stdev(eval(if(_time<relative_time(maxtime,"-1d@d"),'count',null))) AS stdev BY Username
|eval lowerBound=(avg-stdev*2), upperBound=(avg+stdev*2)
|where 'count' > upperBound AND num_data_samples >=7
```

Search explanation

The table provides an explanation of what each part of this search achieves. You can adjust this query based on the
specifics of your environment.

<table>
<thead>
<tr>
<th>Splunk Search</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>EVENT_TYPE=DocumentAttachmentDownloads</td>
<td>Filter for the DocumentAttachmentDownloads EVENT_TYPE.</td>
</tr>
<tr>
<td></td>
<td>lookup lookup_sfdc_usernames USER_ID</td>
</tr>
<tr>
<td></td>
<td>bucket _time span=1d</td>
</tr>
<tr>
<td></td>
<td>stats count BY Username _time</td>
</tr>
<tr>
<td></td>
<td>stats count AS num_data_samples max(eval(if(_time &gt;= relative_time(maxtime, &quot;-1d@d&quot;), 'count',null))) AS count avg(eval(if(_time&lt;relative_time(maxtime,&quot;-1d@d&quot;),'count',null))) AS avg stdev(eval(if(_time&lt;relative_time(maxtime,&quot;-1d@d&quot;),'count',null))) AS stdev BY Username</td>
</tr>
<tr>
<td></td>
<td>eval lowerBound=(avg-stdev<em>2), upperBound=(avg+stdev</em>2)</td>
</tr>
<tr>
<td></td>
<td>where 'count' &gt; upperBound AND num_data_samples &gt;=7</td>
</tr>
</tbody>
</table>

**Result**

While there are no traditional false positives in this search, there will be a lot of noise. Every time this search runs, it will accurately measure a spike in the number of documents monitored.

How you handle these alerts depends on where you set the standard deviation. If you set a low standard deviation (2 or 3), you are likely to get a lot of events that are useful only for contextual information. If you set a high standard deviation...
(6 or 10), the amount of noise can be reduced enough to send an alert directly to analysts.

For most environments, these searches can be run once a day, often overnight, without a lag. If you want to run this search more frequently, or if this search is too slow for your environment, use a summary index that first aggregates the data.

When this search returns values, initiate your incident response process and identify the user demonstrating this behavior. Capture the time of the event, the user's role, and number of documents downloaded. If possible, determine the system used to download this data and its location. Contact the user and their manager to determine if the download is authorized, and then document that it was authorized and by whom. If you cannot find authorization, the user credentials may have been used by another party and additional investigation is warranted.