Web access and web error log correlation

You might to correlate web access logs with web server errors logs when doing the following:

- Managing web server performance

Prerequisites

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

- Splunk Enterprise or Splunk Cloud Platform
- Web server data

Example

While web access logs tell you when users experience errors and for which page requests, error logs indicate why the problem occurred. When these log sources are correlated, it may become evident that certain errors occur only for specific pages, browsers, tenants, or some other class of users. You want to correlate these two data sources for a clearer understanding of the impact specific errors have on users.

To optimize the search shown below, you should specify an index and a time range. In addition, this sample search uses Splunk Add-on for Apache Web Server. You can replace this source with any other web server data used in your organization.

1. Verify you deployed a web server add-on to the search heads, so that the needed tags and fields are defined. For more information, see About installing Splunk add-ons.
2. Run the following search:

   tag=web OR tag=error (sourcetype=apache:error OR (sourcetype=apache:access status>299))
   |eval status_group=case(status<300, "2xx", status<400, "3xx", status<500, "4xx", status<600, "5xx", true(), unknown)
   |eval log_event_type = if(sourcetype="apache:error", "apache_error", status_group)
   |timechart span=1h count BY log_event_type

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.

The information provided in Splunk Lantern is intended for informational and educational purposes only. All information is provided in good faith, however, Splunk disclaims any and all representations and warranties, express and implied, regarding the information provided, including without limitation any warranties and representations regarding the completeness, adequacy or accuracy of the information. You agree to take full responsibility for the results arising from the use of the information provided.
<table>
<thead>
<tr>
<th><strong>Splunk Search</strong></th>
<th><strong>Explanation</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><code>tag=web OR tag=error</code></td>
<td>Search for events that are tagged as web events.</td>
</tr>
<tr>
<td><code>(sourcetype=apache:error OR (sourcetype=apache:access status&gt;299))</code></td>
<td>Search only Apache errors and page requests with an HTTP status. If you are using a web server other than Apache, update the fields as necessary.</td>
</tr>
<tr>
<td>`</td>
<td>eval status_group=case(status&lt;300, &quot;2xx&quot;, status&lt;400, &quot;3xx&quot;, status&lt;500, &quot;4xx&quot;, status&lt;600, &quot;5xx&quot;, true(), unknown)`</td>
</tr>
<tr>
<td>`</td>
<td>eval log_event_type = if(sourcetype=&quot;apache:error&quot;, &quot;apache_error&quot;, status_group)`</td>
</tr>
<tr>
<td>`</td>
<td>timechart span=1h count BY log_event_type`</td>
</tr>
</tbody>
</table>

### Result

If not all the `log_event_type`es have a corresponding entry in the error log, you might need to look at the log level for errors. For instance, Apache does not log errors for pages if the logging level = warn.

A good next step is to look for 4xx errors that have corresponding `apache_error` log entries. For example, a 403 status found in a 4xx `log_event_type` should correspond to a permission error in the Apache error log. Similarly, a 404 status code should correspond to a file not found error. Corresponding errors can be seen by replacing `|timechart span=1h count BY log_event_type` with `|table _time log_event_type _raw`.

In the resulting table, you can look at related entries that are close in time for cause and effect. This information can be used for troubleshooting.