Logging PageEvents to Splunk

A possible way to Log User- and Page-Access in Confluence is via the Event system - using Adaptavist's Scriptrunner for Confluence.

This ways has Pros and Cons - read Access Logging in Confluence. One Pro is that the POST to Splunk in in the backend; so we dont need to open for the receiving system in the Firewall for the world.

My site is mainly external as a website, with only one internal user, myself "bnp". In that situation, the PageViewEvent is not so interesting as if this was an internal system with multiple users.

Currently, I have found no way to correlate bot/spider/monitoring hits from the real PageViews.

Also, PageViewEvents only occur when a page is rendered and this gives back HTTP Code "200 OK" to the client. See Different Loggings for different logging compares.

In this working sample, we will post to a Splunk HTTP Event Collector - after setup and getting the Splunk Key, the collector need a POST like this:

```
{
    "time": 1426279439,
    "host": "localhost",
    "source": "datasource",
    "sourcetype": "txt",
    "index": "main",
    "event": { "hello": "world" }
}
```

We do eliminate the "time" field, as the POST is instantly from the Confluence server.

To achive this, we have setup a Script Event Handler:

the executes this script for every PageViewEvent:

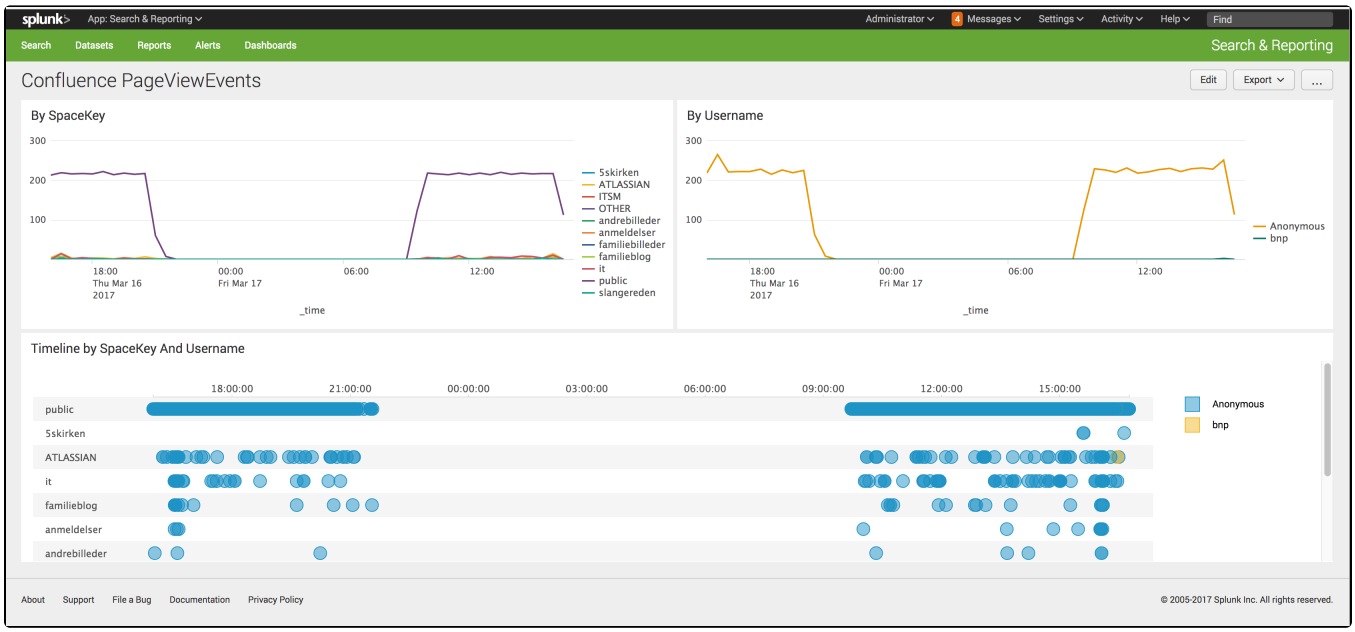
```
import com.atlassian.confluence.user.AuthenticatedUserThreadLocal
import com.atlassian.confluence.user.*;
import java.net.URL;
import java.net.URLEncoder;
import java.net.MalformedURLException;
import java.io.UnsupportedEncodingException;
import com.atlassian.confluence.pages.Page
import com.atlassian.confluence.pages.PageManager
import com.atlassian.confluence.spaces.Space
import com.atlassian.confluence.spaces.SpaceManager
import com.atlassian.sal.api.component.ComponentLocator
import com.atlassian.confluence.event.events.content.page.*

System.out.println("Start post2splunk.groovy")

def spaceManager = ComponentLocator.getComponent(SpaceManager)
def pageManager = ComponentLocator.getComponent(PageManager)
```

```
String userName="Anonymous"
def currentUser = AuthenticatedUserThreadLocal.get()
if (currentUser)
{
  userName=(String)currentUser.name
}

def event = event as PageEvent
String eventType=(String)event.toString()
eventType=eventType.replaceAll("com.atlassian.confluence.event.events.content.page.","")
eventType=eventType.substring(0, eventType.indexOf('@'))
eventType=eventType.replaceAll("Event","")

// keys to create unique nodes for counters
// https://docs.atlassian.com/confluence/5.9.7/com/atlassian/confluence/pages/Page.html

String spaceKey = event.page.getSpace().getKey()
String pageId = event.page.getIdAsString()
String pageName = event.page.getTitle()

def requestMethod = "GET";
def URLParam = []
def baseURL = "http://77.243.52.151:8088/services/collector"

def url = new java.net.URL(baseURL);
URLConnection connection = url.openConnection();
connection.setRequestMethod(requestMethod);
connection.doOutput = true
connection.setUseCaches(false);
connection.setRequestProperty("Content-Type", "application/json;charset=UTF-8");
connection.setRequestProperty('Authorization', 'Splunk XXXXXXX-XXXX-4D74-BB9E-64E3B3730D8E');

String jSon= "{\"host\": \"moserver\",\"source\": \"webaccess\",\"sourcetype\": \"webaccess\",\"index\": \"
webaccess\",\"event\":{"

jSon = jSon + "\"event-type\":\"" + eventType  + "\","
jSon = jSon + "\"space-key\":\"" + spaceKey + "\","
jSon = jSon + "\"confluence-page-title\":\"" + pageName + "\","
jSon = jSon + "\"confluence-page-id\":\"" + pageId + "\","
jSon = jSon + "\"username\":\"" + userName + "\""
jSon = jSon + "}}"

def writer = new OutputStreamWriter(connection.outputStream)
writer.write(jSon)
writer.flush()
writer.close()
connection.connect();
try
{
  connection.getContent()
}
catch (all)
{
}
String Status=connection.getResponseCode()
String Message=connection.getResponseMessage()
```

Giving us results to work on in Splunk (Where we already has created the index needed):

i	Time	Event
>	07/02/2017 17:59:59.000	{ [-] confluence-page-id: 4161541 confluence-page-title: Start event-type: PageView space-key: public username: Anonymous } Show as raw text host = moserver | source = webaccess | sourcetype = webaccess
>	07/02/2017 17:59:59.000	{ [-] confluence-page-id: 4161541 confluence-page-title: Start event-type: PageView space-key: public username: Anonymous } Show as raw text host = moserver | source = webaccess | sourcetype = webaccess
>	07/02/2017 17:59:15.000	{ [-] confluence-page-id: 75825174 confluence-page-title: Logging PageEvents to Splunk or Elasticsearch (or other) event-type: PageView space-key: ATLASSIAN username: bnp } Show as raw text host = moserver | source = webaccess | sourcetype = webaccess
>	07/02/2017 17:58:23.000	{ [-] confluence-page-id: 68648965 confluence-page-title: Alfresco Community 5.x Install Cookbook (Ubuntu 16.04) event-type: PageView space-key: it username: Anonymous } Show as raw text host = moserver | source = webaccess | sourcetype = webaccess
>	07/02/2017 17:58:01.000	{ [-] confluence-page-id: 4161541 confluence-page-title: Start event-type: PageView space-key: public username: Anonymous }

And we can build a dashboard:


splunk> App: Search & Reporting
Administrator
4 Messages
Settings
Activity
Help
Find
Search
Datasets
Reports
Alerts
Dashboards
Search & Reporting
Confluence PageViewEvents
Edit
Export
...
By SpaceKey
300
200
100
18:00
Thu Mar 16
2017
00:00
Fri Mar 17
06:00
12:00
_time
5skirken
ATLASSIAN
ITSM
OTHER
andrebilleder
anmeldelser
familiebilleder
familieblog
it
public
slangereden
By Username
300
200
100
18:00
Thu Mar 16
2017
00:00
Fri Mar 17
06:00
12:00
_time
Anonymous
bnp
Timeline by SpaceKey And Username
18:00:00
21:00:00
00:00:00
03:00:00
06:00:00
09:00:00
12:00:00
15:00:00
public
5skirken
ATLASSIAN
it
familieblog
anmeldelser
andrebilleder
Anonymous
bnp
About
Support
File a Bug
Documentation
Privacy Policy
© 2005-2017 Splunk Inc. All rights reserved.