

Website Input - Task #2498

Task # 2484 (Closed): Python 3 support

Fix unable to find geckodriver in path

11/06/2019 09:43 PM - Luke Murphey

Status:	Closed	Start date:	11/06/2019
Priority:	Normal	Due date:	
Assignee:	Luke Murphey	% Done:	100%
Category:	Input: Browser Renderer	Estimated time:	0.00 hour
Target version:	4.5.8		
Description			

Associated revisions

Revision 727 - 11/07/2019 10:47 PM - Imurphey

Fixing loading of the browser driver on Python 3

Reference #2498

Revision 728 - 11/07/2019 11:35 PM - Imurphey

Updating to geckodriver 0.26.0

Reference #2498

Revision 729 - 11/08/2019 02:25 AM - Imurphey

Updating the darwin geckodriver

Reference #2498

Revision 730 - 11/08/2019 02:25 AM - Imurphey

Changing Firefox to use headless mode

Reference #2498

Revision 731 - 11/08/2019 03:54 AM - Imurphey

Converting the content from the Firefox driver to binary

Reference #2498

History

#1 - 11/06/2019 09:59 PM - Luke Murphey

```
[exec] =====  
[exec] ERROR: test_get_result_basic_auth (__main__.TestBrowserRenderingFirefox)  
[exec] -----  
[exec] Traceback (most recent call last):  
[exec]   File "../src/bin/selenium/webdriver/common/service.py", line 74, in start
```

```
[exec]     stdout=self.log_file, stderr=self.log_file)
[exec] File "/var/opt/ci/splunk/py3/lib/python3.7/subprocess.py", line 775, in __init__
[exec]     restore_signals, start_new_session)
[exec] File "/var/opt/ci/splunk/py3/lib/python3.7/subprocess.py", line 1522, in _execute_child
[exec]     raise child_exception_type(errno_num, err_msg, err_filename)
[exec] FileNotFoundError: [Errno 2] No such file or directory: 'geckodriver': 'geckodriver'
[exec]
[exec] During handling of the above exception, another exception occurred:
[exec]
[exec] Traceback (most recent call last):
[exec] File "/var/lib/go-agent/pipelines/web-input/tests/unit_test_web_server.py", line 89, in _decorat
or
[exec]     return func(self, *args, **kwargs)
[exec] File "unit.py", line 932, in test_get_result_basic_auth
[exec]     content = client.get_url("http://admin:changeme@127.0.0.1:" + str(self.web_server_port) + "/")
[exec] File "../src/bin/website_input_app/web_driver_client.py", line 254, in get_url
[exec]     self.driver = self.get_driver()
[exec] File "../src/bin/website_input_app/web_driver_client.py", line 334, in get_driver
[exec]     driver = webdriver.Firefox(profile, log_path=make_splunkhome_path(['var', 'log', 'splunk', 'ge
ckodriver.log']))
[exec] File "../src/bin/selenium/webdriver/firefox/webdriver.py", line 142, in __init__
[exec]     self.service.start()
[exec] File "../src/bin/selenium/webdriver/common/service.py", line 81, in start
[exec]     os.path.basename(self.path), self.start_error_message)
[exec] selenium.common.exceptions.WebDriverException: Message: 'geckodriver' executable needs to be in PA
TH.
```

#2 - 11/07/2019 06:23 PM - Luke Murphey

Questions:

- Is the binary executable?
 - Yes
- Is the binary listed in the path?
- Are there other reports of this issue?

Observations:

- This works on the Mac

#3 - 11/07/2019 11:35 PM - Luke Murphey

Now getting: selenium.common.exceptions.WebDriverException: Message: invalid argument: can't kill an exited process

Questions:

- What version of the geckodriver, Firefox and Selenium is being used?
 - <https://stackoverflow.com/questions/52534658/webdriverexception-message-invalid-argument-cant-kill-an-exited-process-with>
 - geckodriver: 0.16.1
 - selenium: 3.4.1
 - firefox: Mozilla Firefox 67.0
- Is there anything in the geckodriver.log file?
 - Yes
 - `/var/opt/ci/splunk/py3$ sudo tail var/log/splunk/geckodriver.log`

```
1573171275575  mozrunner::runner      INFO    Running command: "/usr/bin/firefox" "--marionette" "--for
eground" "--no-remote" "--profile" "/tmp/rust_mozprofilemm6SAA"
Error: no DISPLAY environment variable specified
1573171276749  mozrunner::runner      INFO    Running command: "/usr/bin/firefox" "--marionette" "--for
eground" "--no-remote" "--profile" "/tmp/rust_mozprofilepv3Mya"
Error: no DISPLAY environment variable specified
1573171277917  mozrunner::runner      INFO    Running command: "/usr/bin/firefox" "--marionette" "--for
eground" "--no-remote" "--profile" "/tmp/rust_mozprofile6HrQf8"
Error: no DISPLAY environment variable specified
1573171279116  mozrunner::runner      INFO    Running command: "/usr/bin/firefox" "--marionette" "--for
eground" "--no-remote" "--profile" "/tmp/rust_mozprofileZIIhQb"
Error: no DISPLAY environment variable specified
1573171280286  mozrunner::runner      INFO    Running command: "/usr/bin/firefox" "--marionette" "--for
eground" "--no-remote" "--profile" "/tmp/rust_mozprofilehxrcNK"
Error: no DISPLAY environment variable specified
```

- Does setting headless mode fix it?
 - https://developer.mozilla.org/en-US/docs/Mozilla/Firefox/Headless_mode

Observations:

- Seems like the old geckodriver doesn't support Firefox 67

#4 - 11/07/2019 11:43 PM - Luke Murphey

It appears that this won't work with Mac with the Developer Edition of Firefox. See <https://github.com/mozilla/geckodriver/issues/1488>

#5 - 11/08/2019 03:45 AM - Luke Murphey

Getting authentication failures now:

I'm getting this on Linux:

```
[exec] =====
[exec] ERROR: test_form_auth_auto_discover_form_fields (__main__.TestFormAuthenticationFirefox)
[exec] -----
[exec] Traceback (most recent call last):
[exec]   File "unit.py", line 1293, in test_form_auth_auto_discover_form_fields
[exec]     self.assertEqual(result['match'][0], "Auth success")
[exec] IndexError: list index out of range
[exec]
[exec] =====
[exec] FAIL: test_form_auth_spider (__main__.TestFormAuthenticationFirefox)
[exec] -----
[exec] Traceback (most recent call last):
[exec]   File "unit.py", line 1235, in test_form_auth_spider
[exec]     self.assertEqual(len(results), 4)
[exec] AssertionError: 1 != 4
```

I'm getting these failures on Mac:

```
[exec] =====
[exec] ERROR: test_scrape_page (__main__.TestBrowserRenderingFirefox)
[exec] -----
[exec] Traceback (most recent call last):
[exec]   File "/Users/lmurphey/git/splunk-web-input/tests/unit_test_web_server.py", line 89, in _decorato
r
[exec]     return func(self, *args, **kwargs)
[exec]   File "unit.py", line 905, in test_scrape_page
[exec]     results = web_scraper.scrape_page(url_field.to_python("http://127.0.0.1:" + str(self.web_serve
r_port) + "/html"), selector_field.to_python("h1"), output_matches_as_mv=True, browser=self.BROWSER)
[exec]   File "../src/bin/website_input_app/web_scraper.py", line 758, in scrape_page
[exec]     additional_fields=additional_fields, **kw)
[exec]   File "../src/bin/website_input_app/web_scraper.py", line 452, in get_result_single
[exec]     encoding = self.detect_encoding(content, web_client.get_response_headers())
[exec]   File "../src/bin/website_input_app/web_scraper.py", line 257, in detect_encoding
[exec]     matched_encoding = find_meta_charset.search(content)
[exec] TypeError: cannot use a bytes pattern on a string-like object
[exec]
[exec] =====
[exec] ERROR: test_form_auth (__main__.TestFormAuthenticationFirefox)
[exec] -----
[exec] Traceback (most recent call last):
[exec]   File "unit.py", line 1217, in test_form_auth
[exec]     results = web_scraper.scrape_page(data_url, selector_field.to_python("h1"), browser=self.BROWS
ER)
[exec]   File "../src/bin/website_input_app/web_scraper.py", line 758, in scrape_page
[exec]     additional_fields=additional_fields, **kw)
[exec]   File "../src/bin/website_input_app/web_scraper.py", line 452, in get_result_single
[exec]     encoding = self.detect_encoding(content, web_client.get_response_headers())
[exec]   File "../src/bin/website_input_app/web_scraper.py", line 257, in detect_encoding
[exec]     matched_encoding = find_meta_charset.search(content)
[exec] TypeError: cannot use a bytes pattern on a string-like object
[exec]
[exec] =====
[exec] ERROR: test_form_auth_auto_discover_form_fields (__main__.TestFormAuthenticationFirefox)
[exec] -----
[exec] Traceback (most recent call last):
[exec]   File "unit.py", line 1289, in test_form_auth_auto_discover_form_fields
[exec]     results = web_scraper.scrape_page(data_url, selector_field.to_python("h1"), browser=self.BROWS
ER)
[exec]   File "../src/bin/website_input_app/web_scraper.py", line 758, in scrape_page
[exec]     additional_fields=additional_fields, **kw)
[exec]   File "../src/bin/website_input_app/web_scraper.py", line 452, in get_result_single
```

```
[exec]     encoding = self.detect_encoding(content, web_client.get_response_headers())
[exec] File "../src/bin/website_input_app/web_scraper.py", line 257, in detect_encoding
[exec]     matched_encoding = find_meta_charset.search(content)
[exec] TypeError: cannot use a bytes pattern on a string-like object
[exec]
[exec] =====
[exec] ERROR: test_form_auth_spider (__main__.TestFormAuthenticationFirefox)
[exec] -----
[exec] Traceback (most recent call last):
[exec]   File "unit.py", line 1233, in test_form_auth_spider
[exec]     results = web_scraper.scrape_page(data_url, selector_field.to_python("h1"), page_limit=5, browser=self.BROWSER)
[exec]   File "../src/bin/website_input_app/web_scraper.py", line 758, in scrape_page
[exec]     additional_fields=additional_fields, **kw)
[exec]   File "../src/bin/website_input_app/web_scraper.py", line 452, in get_result_single
[exec]     encoding = self.detect_encoding(content, web_client.get_response_headers())
[exec]   File "../src/bin/website_input_app/web_scraper.py", line 257, in detect_encoding
[exec]     matched_encoding = find_meta_charset.search(content)
[exec] TypeError: cannot use a bytes pattern on a string-like object
[exec]
[exec] -----
```

#6 - 11/08/2019 03:45 AM - Luke Murphey

It appears that `web_client.get_url()` returns Unicode from the Firefox client.

#7 - 11/08/2019 04:15 AM - Luke Murphey

- Status changed from New to Closed

- % Done changed from 0 to 100