

RhodeCode CE/EE - Bug #5494

rccontrol's python package management causes slow VCS SSH

21.08.2018 16:27 - Peter Colledge

Status:	New	Start date:	21.08.2018
Priority:	Normal	Due date:	
Assignee:		% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:		Spent time:	0.00 hour
Sorting:		Affected Version:	v4.12
Commit Number:			

Description

Separating python modules into /opt/rhodecode/store/<hash>-python2.7-<name>/ causes python to search for dependencies that are not in their relevant directory, causing hundreds of thousands of missing file errors.

To debug:

Edit one troubleshooting user in ~rcuser/.ssh/authorized_keys_rhodecode to include command="strace -f -o /tmp/rc-ssh-wrapper.log ~/.rccontrol/..."

```
$ less /tmp/rc-ssh-wrapper.log
```

```
...
34782      0.000131  stat("/opt/rhodecode/store/y2xhi7pkj4cycnaxm2wkw0hhcmfsn846-python2.7-celery-4
.1.0/lib/python2.7/site-packages", {st_mode=S_IFDIR|0755, st_size=4096, ...}) = 0
34782      0.000140  stat("/opt/rhodecode/store/y2xhi7pkj4cycnaxm2wkw0hhcmfsn846-python2.7-celery-4
.1.0/lib/python2.7/site-packages", {st_mode=S_IFDIR|0755, st_size=4096, ...}) = 0
34782      0.000134  stat("/opt/rhodecode/store/y2xhi7pkj4cycnaxm2wkw0hhcmfsn846-python2.7-celery-4
.1.0/lib/python2.7/site-packages/os", 0x7ffd146ac630) = -1 ENOENT (No such file or directory)
34782      0.000133  open("/opt/rhodecode/store/y2xhi7pkj4cycnaxm2wkw0hhcmfsn846-python2.7-celery-4
.1.0/lib/python2.7/site-packages/os.so", O_RDONLY) = -1 ENOENT (No such file or directory)
34782      0.000179  open("/opt/rhodecode/store/y2xhi7pkj4cycnaxm2wkw0hhcmfsn846-python2.7-celery-4
.1.0/lib/python2.7/site-packages/osmodule.so", O_RDONLY) = -1 ENOENT (No such file or directory)
34782      0.000129  open("/opt/rhodecode/store/y2xhi7pkj4cycnaxm2wkw0hhcmfsn846-python2.7-celery-4
.1.0/lib/python2.7/site-packages/os.py", O_RDONLY) = -1 ENOENT (No such file or directory)
34782      0.000142  open("/opt/rhodecode/store/y2xhi7pkj4cycnaxm2wkw0hhcmfsn846-python2.7-celery-4
.1.0/lib/python2.7/site-packages/os.pyc", O_RDONLY) = -1 ENOENT (No such file or directory)
...
```

```
$ grep ENOENT /tmp/rc-ssh-wrapper.log | wc -l
```

```
256522
```

```
$ grep ENOENT /tmp/rc-ssh-wrapper.log | grep -v python | wc -l
```

```
965
```

Adapt the strace to -c:

% time	seconds	usecs/call	calls	errors	syscall
44.60	6.077716	675302	9	3	wait4
33.16	4.519439	22	208836	202905	open
15.13	2.061998	17	122457	53627	stat
3.76	0.511863	18	28472	62	lstat

Re-ordering the PYTHONPATH may be an idea.

History

#1 - 21.08.2018 17:28 - Marcin Kuzminski [staff]

Thanks for reporting this, interesting find. I however doubt it will speed up the SSH invocation script. There's much more overhead loading the pyramid environment than this. We started to investigate making it execute faster by making the env limited to required parts of SSH rather than whole "app"

#2 - 01.04.2019 12:33 - Alexey Petrov

Hello! Is there any news to fix this problem? In version 4.16.1, the problem persists.

#3 - 01.04.2019 12:40 - Marcin Kuzminski [staff]

Hi,

We haven't worked on those parts since it contributes to roughly 10% of the execution time of the ssh-wrapper.

The big slowdown part is the loading of our Application, which roughly now takes 5-8s. We want to prioritize on making that faster before jumping into this kind of optimizations.

#4 - 02.04.2019 11:31 - Alexey Petrov

Thanks for the fast reply! Optimization will take some time or you already have a solution that will be presented in the next updates?

#5 - 02.04.2019 11:56 - Marcin Kuzminski [staff]

Can you describe your use case, and why this is important for you?

There are already some changes in the code we did, but they aren't turned on.

<https://code.rhodecode.com/rhodecode-enterprise-ce/files/c75dd264c5f748c55e73391632ccd00a07f03f99/rhodecode/config/middleware.py#L249>

If you'd modify `should_load_all` function (<https://code.rhodecode.com/rhodecode-enterprise-ce/files/c75dd264c5f748c55e73391632ccd00a07f03f99/rhodecode/config/middleware.py#L67>) to e.g:

```
def should_load_all():
    """
    Returns if all application components should be loaded. In some cases it's
    desired to skip apps loading for faster shell script execution
    """
    is_ssh = os.environ.get('SSH_ORIGINAL_COMMAND', '')
    if is_ssh:
        return False
    return True
```

This should already be a big improvement for SSH wrapper speed.