

A faint, light gray wireframe globe is centered in the background, showing latitude and longitude lines.

May 23—27, 2022

AIRFLOW SUMMIT

Large, bright green abstract shapes are positioned on the left and right sides of the image. On the left, there is a thick, curved line and a solid rectangular block below it. On the right, there is a thick, curved line. Small green triangles point towards the center from the top right and bottom left.

Hey Maintainer (and user)

Exercise your empathy



Airflow Survey!

<https://bit.ly/AirflowSurvey22>

About me



Jarek Potiuk

Independent Open-Source Contributor and Advisor
Airflow Committer & PMC member
Twitter: @jarekpotiuk

What's empathy?

Empathy by Merriam-Webster

the action of understanding, being aware of, being sensitive to, and vicariously experiencing the feelings, thoughts, and experience of another of either the past or present without having the feelings, thoughts, and experience fully communicated in an objectively explicit manner

Empathy by Merriam-Webster

the action of **understanding**, being aware of, being sensitive to, and vicariously experiencing the feelings, thoughts, and experience of another of either the past or present without having the feelings, thoughts, and experience fully communicated in an objectively explicit manner

Empathy by Merriam-Webster

the action of understanding, **being aware of, being sensitive to**, and vicariously experiencing the feelings, thoughts, and experience of another of either the past or present without having the feelings, thoughts, and experience fully communicated in an objectively explicit manner

Empathy by Merriam-Webster

the action of understanding, being aware of, being sensitive to, and **vicariously experiencing** the feelings, thoughts, and experience of another of either the past or present without having the feelings, thoughts, and experience fully communicated in an objectively explicit manner

Empathy by Merriam-Webster

the action of understanding, being aware of, being sensitive to, and vicariously experiencing the **feelings, thoughts, and experience** of another of either the past or present without having the feelings, thoughts, and experience fully communicated in an objectively explicit manner

Empathy by Merriam-Webster

the action of understanding, being aware of, being sensitive to, and vicariously experiencing the feelings, thoughts, and experience of **another** of either the past or present without having the feelings, thoughts, and experience fully communicated in an objectively explicit manner

Empathy by Merriam-Webster

the action of understanding, being aware of, being sensitive to, and vicariously experiencing the feelings, thoughts, and experience of another of either the past or present **without having the feelings, thoughts, and experience** fully communicated in an objectively explicit manner

Empathy by Merriam-Webster

the action of understanding, being aware of, being sensitive to, and vicariously experiencing the feelings, thoughts, and experience of another of either the past or present without having the feelings, thoughts, and experience **fully communicated in an objectively explicit manner**

Empathy by Merriam-Webster

the action of understanding, being aware of, being sensitive to, and vicariously experiencing the feelings, thoughts, and experience of another of either the past or present without having the feelings, thoughts, and experience fully communicated in an objectively explicit manner

also : the capacity for this

Why we have to talk about it?

- Some people are more naturally empathetic than others
- Empathy can be taught
- You studied a lot about programming, right?
- When is the last time you took empathy course?

Apache Airflow community

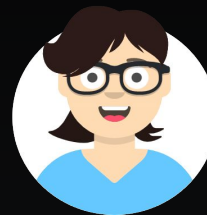
Who “we” are?



Committer
(Maintainer)



Contributor

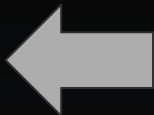


User

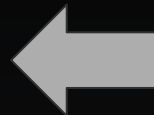
What's the journey ?



Committer
(Maintainer)



Contributor



User

This talk is not about



Contributor

Session

Adding an executor to Airflow: A contributor overflow exception



Vanessa Sochat

Research Software Engineer
@ Stanford University
Research Computing Center

Airflow 
Summit 2020



@AirflowSummit
/Airflow-Summit

Join from anywhere.
<https://airflowsummit.org>

<https://youtu.be/RKEmAshcreE>

Adding An Executor to Airflow: A Contributor Overflow Exception



Vanessa Sochat



Adding An Executor to Airflow: A Contributor Overflow Exception



Vanessa Sochat

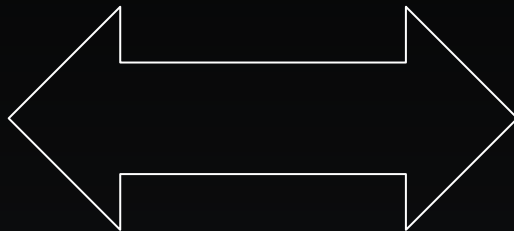
Be like Elmo!



User and Maintainer



Committer
(Maintainer)



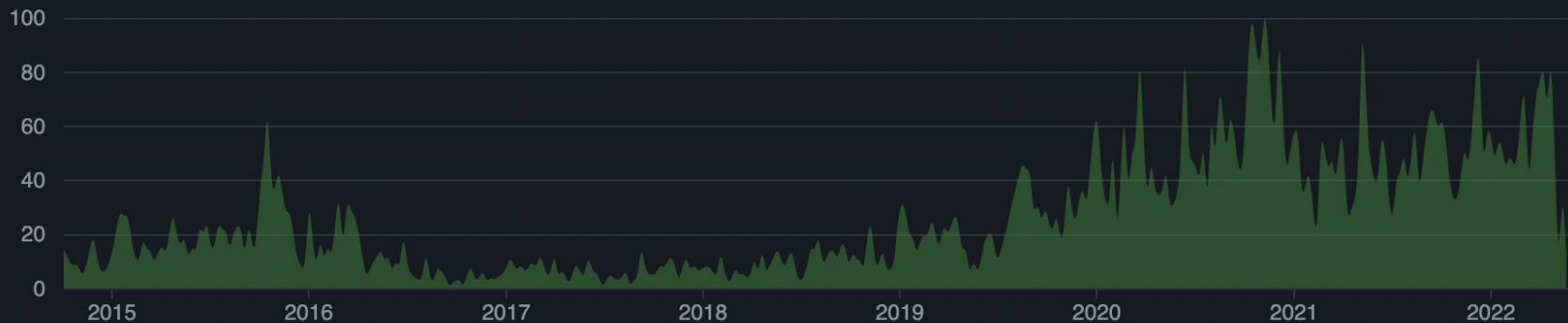
User

Hey maintainer

- Do you remember time when you were user?
- Did you know where to look in sources ?
- Did you know by heart all the configuration values?
- Did you know that someone already discussed it in Issue/devlist/discussion?
- Is it REALLY described in the documentation?
- Is this OK that the user might not know how airflow components are run?
- Is this OK that the user does not know Python?

Apache Airflow Pulse

Contributions to main, excluding merge commits and bot accounts



Hey user - do you know that:

- our maintainer to user ratio is less than 2 : 1000 ?
- we have more than 10 merged PRs a day ?
- there are PR with more than 100 comments ?
- there are about 80 - 100 messages generated a day from airflow repo activity?
- maintainers are not paid to help with troubleshooting ?
- there are SEARCHABLE(!) docs/issues/discussions/stack overflow?

What does it mean to be an empathetic maintainer?

You know nothing - John Snow

And now you can remove all the specific copy operators like `S3ToSFTP`, `AzureBlobbToGCS` ... Or, if they don't exist you don't need to implement them anymore!

What do you think ~~@turbaszk~~? I don't know how the Core development is organized (I know there's a Jira and a mailing list, but know nothing of the organization processes). This will be a core change. But if Airflow aims to thrive, it's necessary (IMHO)



Direct the users

Apache Airflow version

2.1.1

What happened

When I am API calling for users with limit 1000, it returns only 500 requests. How or where can I increase limit more than 500 requests? Looks like by default airflow API returns max 500 requests.

I am running following command but always return 500 requests: `https://airflow_url/api/v1/users?limit=1000`



mik-laj commented on 19 Apr

Member




Page size limit


To protect against requests that may lead to application instability, the stable API has a limit of items in response. The default is 100 items, but you can change it using `maximum_page_limit` option in `[api]` section in the `airflow.cfg` file.


<https://airflow.apache.org/docs/apache-airflow/stable/security/api.html>

Direct the users

**josh-fell** on 28 Jan Collaborator


Have you taken a look at `user_defined_macros` or `user_defined_filters` in the [Jinja templating docs](#)? They are pretty handy when you want to expose custom functions within Jinja templating.

 Unmark as answer ↑ 1 😊 1 reply

**a246530** on 28 Jan Author

Thank you for pointing this out. I was not fully aware of this. I just wrote a user defined filter to decode the xcom base64 value the way I initially suggested in my question and it worked just as expected.
Thanks again!

👍 1 🔥 1

 Write a reply

Answer selected by **a246530**

Check your assumptions



potiuik commented 13 days ago • edited ▾

Member



The `all` is a development set of extras and it only works welll in `main` branch.

Please use specific set of extras and providers you want to install rather than installing `all` - this should work fine.

You should not use it to install all dependencies of airflow in released airflow version. It is impossible to make it really without heavy complication of setup `oy`. I will make it clear in the docs because I think it is not really described well.



potiuik mentioned this issue 13 days ago

Clarify that bundle extras should not be used for PyPi installs #23697

Merged



potiuik added a commit to potiuik/airflow that referenced this issue 13 days ago



Clarify that bundle extras should not be used for PyPi installs ...

affefa0



potiuik commented 13 days ago

Member



Documentation clarification in #23697

Communicate via code

```
252 + >&2 echo "WARNING! You should run the image with GID (Group ID) set to 0"
253 + >&2 echo "          even if you use 'airflow' user (UID=50000)"
254 + >&2 echo
255 + >&2 echo " You started the image with UID=$(id -u) and GID=$(id -g)"
256 + >&2 echo
257 + >&2 echo " This is to make sure you can run the image with an arbitrary UID in the future."
258 + >&2 echo
259 + >&2 echo " See more about it in the Airflow's docker image documentation"
260 + >&2 echo "      http://airflow.apache.org/docs/docker-stack/entrypoint"
261 + >&2 echo
262 + # We still allow the image to run with `airflow` user.
263 + return
```

Invite the users to contribute

Apache Airflow version

2.3.0 (latest released)

What happened

We used to configure the Database using the `AIRFLOW__CORE__SQL_ALCHEMY_CONN_CMD` Environment variable. Now the config option moved from CORE to DATABASE. However, we intended to keep backward compatibility as stated in the [Release Notes](#).

Upon 2.3.0 update however, the `_CMD` suffix variables are no longer recognized for database configuration in Core - I think due to a missing entry here:

```
airflow/airflow/configuration.py  
Line 135 in 8622808
```

```
135      ('database', 'sqlalchemy_conn'),
```

What you think should happen instead

We should only get a deprecation warning but the

Are you willing to submit PR?

☐ Yes I am willing to submit a PR!



potiuik commented 21 days ago • edited

Member

Indeed seems like a bug. Would you like to submit a PR fixing it @c-thiel ? Seems like an easy fix - and you could become one of the > 2000 contributors.



c-thiel mentioned this issue 22 days ago

Add backward compatibility for core__sql_alchemy_conn__cmd #23441

Merged



mik-laj closed this as completed 23 days ago

Invite the users to contribute

I think one could add that to the docs.

Surely. Feel free. <https://airflow.apache.org/docs/apache-airflow/stable/start/local.html> - at the bottom right you will find "suggest a change on this page" and you can open a PR directly clicking on it. I think as a user who struggled with finding it, you are the best to describe it in the way that other users like you will find useful. Airflow is created by more than 1800 contributors - most of them voluntarily - and small contributions like that are the best way to start "giving back" to the community for the free software you get :)



Unmark as answer



1



3 replies



boredland on 20 Dec 2021 Author



would you accept a PR removing that hard-coded port? "8080" is the default port. I removed that bit and it still started on 8080, as expected. So removing it would not break anything and ~~between "fix(standalone): remove hardcoded webserver-port #20429"~~ another port.



potiuk on 20 Dec 2021 Collaborator

sure. But some "docs" around that would be nice too (i.e. ~~between "fix(standalone): remove hardcoded webserver-port #20429"~~



1



boredland on 20 Dec 2021 Author

ok, lets further discuss that [here](#).



Write a reply

fix(standalone): remove hardcoded webserver-port #20429



Merged

kaxil merged 3 commits into [apache:main](#) from [boredland:patch-1](#) on 21 Dec 2021



Conversation 4



Commits 3



Checks 36



Files changed 2



boredland commented on 20 Dec 2021 • edited

Contributor

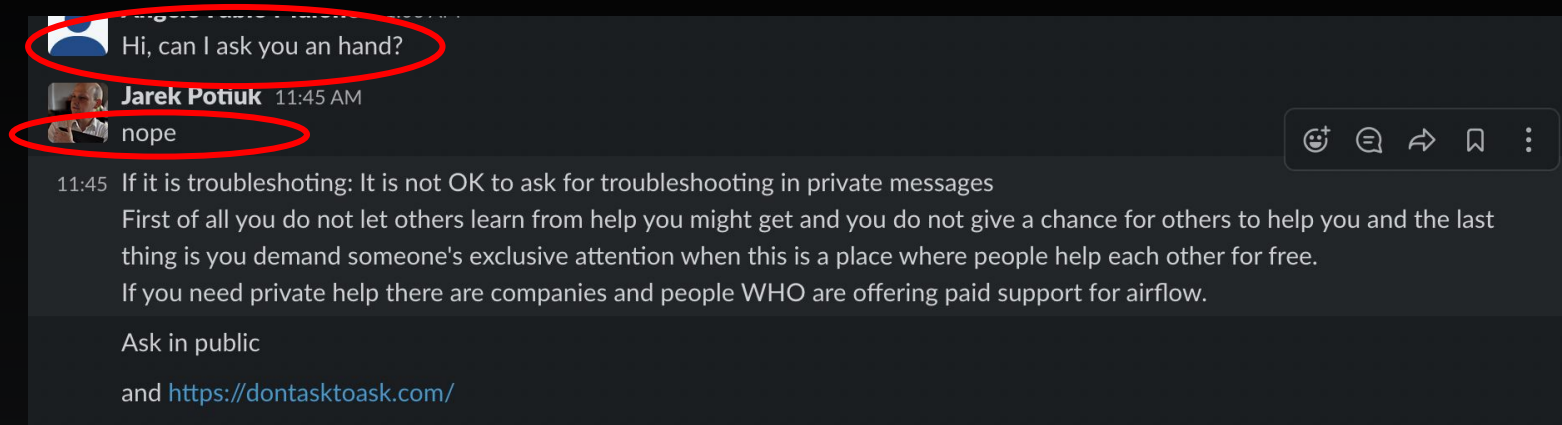
Answer selected by **boredland**

Empathy and assertiveness



In order to be able to help others,
you must help yourself first.

Set the boundaries



- You cut yourself from help of others
- You demand exclusive attention for free
- You don't let others learn from help you might get

What does it mean to be an empathetic user?

Read The Friendly Manual

How to achieve dependencies between tasks #23032

Unanswered

Lenchok asked this question in Q&A



Lenchok on 15 Apr

...

Apache Airflow version

2.2.2

What happened

My requirement is that the post-task will run only if the pre-dependency runs successfully. If the pre-dependency does not run or fails, the post-task will not run even when the time is up. How to implement this function?



potiuk on 15 Apr

Collaborator

edited ...

Look at our docs and search for "task dependencies". You will find answer on your own without unnecessarily engaging other people. The documentation we have is pretty comprehensive: <https://airflow.apache.org/docs/>



Mark as answer



1



0 replies



Write a reply

Don't ask to ask

<https://dontasktoask.com/>

Don't ask to ask, just ask

Every now and then, in online chat rooms I hang around in, someone pops in and says something in the lines of,



Foobar123: Any Java experts around?

This is bad form, for several reasons. What the person is *actually* asking here is,

- You ask for too much responsibility
- Laziness
- So, to summarize, don't ask *"Any Java experts around?"*, but rather ask *"How do I do [problem] with Java and [other relevant info]?"*

Use Github Discussions more

How to reproduce

What should we do to reproduce the problem? If you are not able to provide a reproducible case, please open a [discussion](#) instead.

Write

Preview

H B I ≡ <> 🔗 ≡ ≡ ☑ @ ↗

Please make sure you provide a reproducible step-by-step case of how to reproduce the problem as minimally and precisely as possible. Keep in mind we do not have access to your cluster or DAGs. Remember that non-reproducible issues will be

Ask a question or get support

Ask a question or request support for using Apache Airflow

↗ Open

Using issues for questions

This issue was moved to a discussion.

You can continue the conversation there. [Go to discussion](#) →

How do I install airflow using helm on arm architecture #23777

New issue

🔒 Closed

👤 1 person done 📅 11 days ago 🗓️ closed this issue 9 days ago 💬 1 comment



commented 9 days ago

Official Helm Chart version

1.5.0 (latest released)

Apache Airflow version

2.3.0 (latest released)

Kubernetes Version

1.18.2

Helm Chart configuration

No response

Docker Image customisations

No response

What happened

No response

What you think should happen instead

No response

How to reproduce

No response

Assignees

No one—assign yourself

Labels

area:helm-chart kind:bug

Projects

None yet

Milestone

No milestone

Development

Create a branch for this issue or link a pull request.

Notifications

Customize

🔔 Unsubscribe

You're receiving notifications because you're watching this repository.

1 participant



🔓 Unlock conversation

📌 Pin issue ⓘ

Help to help yourself

I am deploying airflow in Kubernetes with this chart: <https://github.com/apache/airflow/tree/master/chart> and I am using:

- AKS Kubernetes version: 1.16.13
- PostgreSQL hosted in Azure Database for PostgreSQL version: 10

If I use the PostgreSQL database sub chart it works perfectly.

I modify the values.yml to connect to the external PostgreSQL database:

```
# Airflow database config
data:
  # If secret names are provided, use those secrets
  metadataSecretName: ~
  resultBackendSecretName: ~

  # Otherwise pass connection values in
  metadataConnection:
    user: <Myuser>
    pass: <Mypass>
    host: <server>
    port: 5432
    db: airflow
    sslmode: disable
  resultBackendConnection:
    user: <Myuser>
    pass: <Mypass>
    host: <server>
    port: 5432
    db: airflow
    sslmode: disable
```

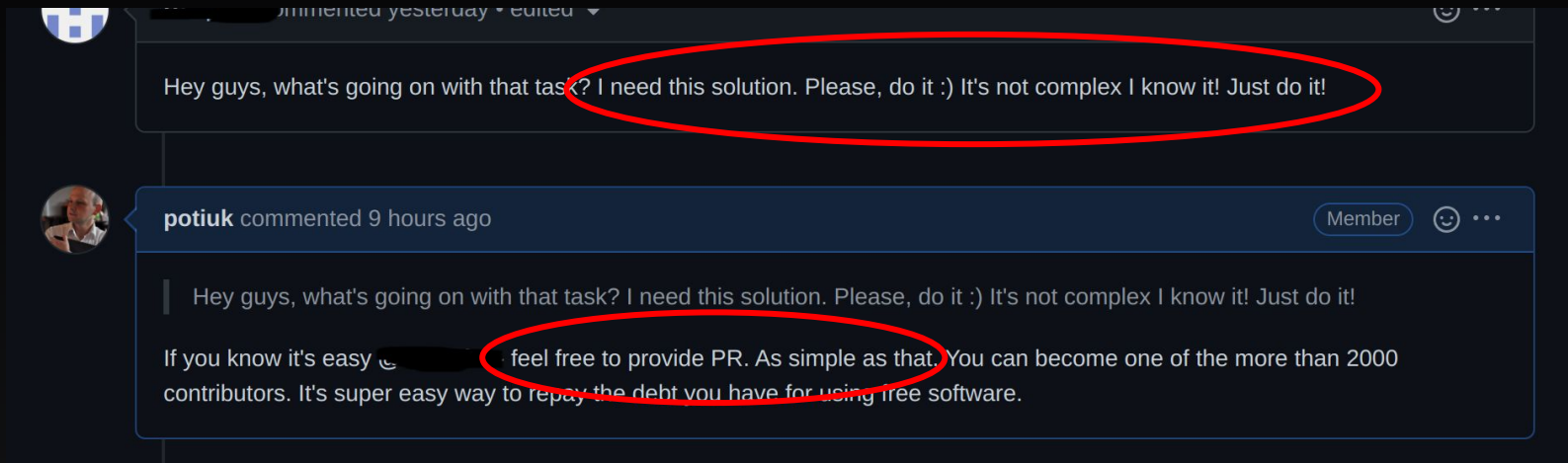
When I deploy it the scheduler and webserver pods fail. I have checked the connection string and using it manually works fine

Add value

The screenshot shows a GitHub issue thread with four comments. Red circles highlight specific parts of the comments to illustrate 'adding value'.

- Comment 1 (ashb):** The entire comment is circled. It states that the issue is fixed in version 2.2.2 and will be fixed when 2.2.2 is out in a few days. The text is: "Just confirmed this works on Mysql with #19425 when going from 2.1.2 to v2-2-test branch, so this will be fixed when 2.2.2 is out in a few days".
- Comment 2 (ashb):** The entire comment is circled. It states that the issue is closed as completed. The text is: "ashb closed this as completed on 9 Nov 2021".
- Comment 3 (User):** The entire comment is circled. It asks if the issue has been resolved. The text is: "hi, i'm facing this issue, anyone resolved it?".
- Comment 4 (potiuk):** The entire comment is circled. It explains that the issue is closed and fixed in 2.2.2, and advises the user to open a new issue if they face a similar problem. The text is: "This issue is closed and fixed in 2.2.2. If you face simiilar issue @ashb for 2.2.2 and 2.2.3 - please open a new one where you specify all details - this is likely different issue similar to this one. It's really not helpful to say 'I have the same issue' without providing no details when issue is closed already."


Knowing better



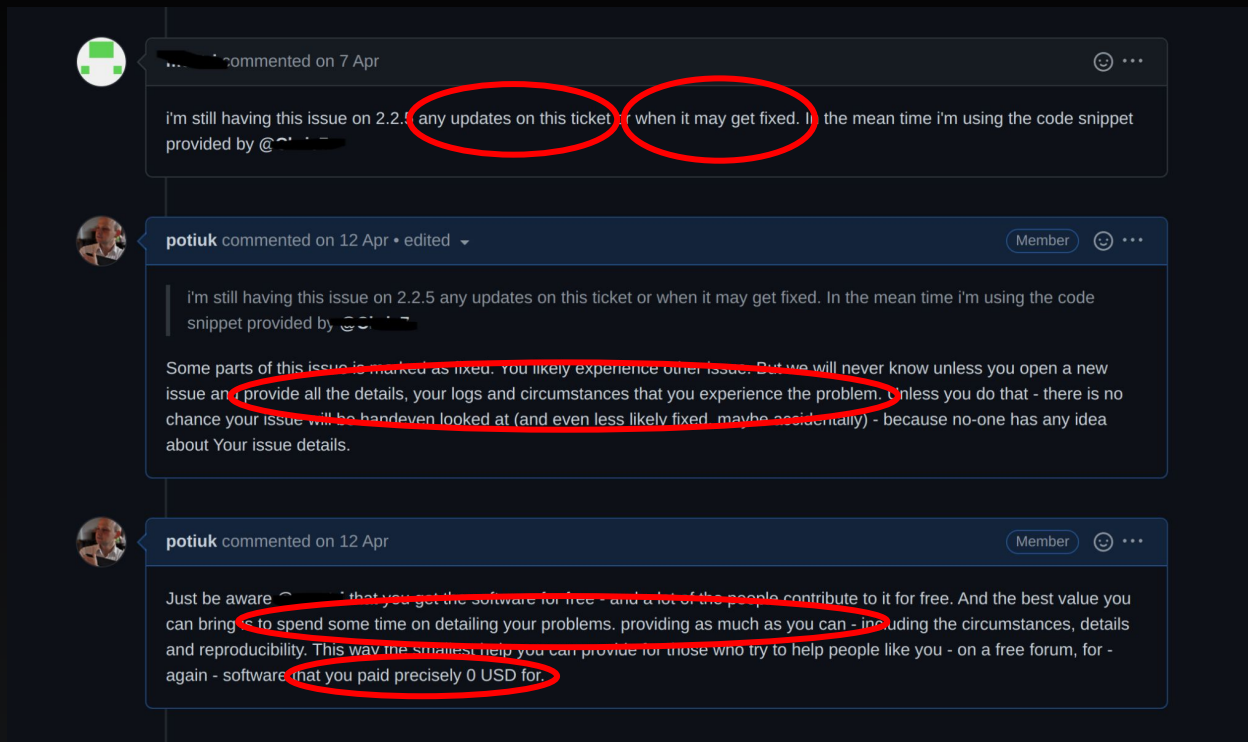
Hey guys, what's going on with that task? I need this solution. Please, do it :) It's not complex I know it! Just do it!

potiuk commented 9 hours ago Member

Hey guys, what's going on with that task? I need this solution. Please, do it :) It's not complex I know it! Just do it!

If you know it's easy  feel free to provide PR. As simple as that. You can become one of the more than 2000 contributors. It's super easy way to repay the debt you have for using free software.

Community is not a helpdesk



The screenshot shows a GitHub issue thread with three comments. Red circles highlight specific phrases in each comment:

- Comment 1:** "i'm still having this issue on 2.2.5 any updates on this ticket or when it may get fixed. In the mean time i'm using the code snippet provided by @..."
- Comment 2:** "Some parts of this issue is marked as fixed. you likely experience other issue. But we will never know unless you open a new issue and provide all the details, your logs and circumstances that you experience the problem. Unless you do that - there is no chance your issue will be handeven looked at (and even less likely fixed. maybe accidentally) - because no-one has any idea about Your issue details."
- Comment 3:** "Just be aware @... that you get the software for free - and a lot of the people contribute to it for free. And the best value you can bring is to spend some time on detailing your problems. providing as much as you can - including the circumstances, details and reproducibility. This way the smallest help you can provide for those who try to help people like you - on a free forum, for - again - software that you paid precisely 0 USD for."

Thank you !

Q&A

Survey! <https://bit.ly/AirflowSurvey22>

