

# python操作impala

官网说明: <https://github.com/cloudera/impyla>

```
yum install gcc-c++ python-devel.x86_64 cyrus-sasl-devel.x86_64 cyrus-sasl-plain
```

(Ubuntu `sudo apt-get install libsasl2-dev` 就可以了)

## 安装pip

```
$ sudo yum install python-pip
```

```
pip3.5 install six
```

```
pip3.5 install bit_array
```

```
pip3.5 install thrift
```

```
pip3.5 install thrift_sasl==0.2.1
```

```
[root@bigdata-3 ~]# python3.5
Python 3.5.6 (default, Oct 9 2018, 19:30:31)
[GCC 4.8.5 20150623 (Red Hat 4.8.5-28)] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> from impala.dbapi import connect
>>> conn = connect(host='10.0.19.48', port=21050, database='bf_hpt', user='yarn', password=
>>> cur = conn.cursor()
>>> cur.execute('SHOW DATABASES')
>>> print(cur.fetchall())
[('impala_builtins', 'System database for Impala builtin functions'), ('account', ''), ('acc
, ''), ('baofoo_ma', ''), ('baofoo_rm', ''), ('baofoo_taskmaster', ''), ('bf_hpt', ''), ('cre
('tmp_baofoo_boas', ''), ('tmp_baofoo_cgw', ''), ('tmp_baofoo_cm', ''), ('tmp_baofoo_cm_v2',
), ('tmp_pay591', ''), ('tmp_tmp_baofoo_rm', ''), ('tmp_xy_ods', ''), ('xy_hpt', ''), ('xy_oo
>>> █
```

```
from impala.dbapi import connect
```

*# 需要注意的是这里的auth\_mechanism必须有, 但database不必须*

```
conn = connect(host='10.0.19.48', port=21050, database='bf_hpt', user='yarn',
password='baofoo@64', auth_mechanism='PLAIN')
cur = conn.cursor()

cur.execute('SHOW DATABASES')
print(cur.fetchall())

cur.execute('SHOW Tables')
print(cur.fetchall())

cur.close()
conn.close()
```

ubuntu版本参考:

[https://blog.csdn.net/gamer\\_gyt/article/details/52564335](https://blog.csdn.net/gamer_gyt/article/details/52564335)

方法二： 用odbc的方式操作

安装cloudera官网的odbc驱动 ClouderaImpalaODBC64.msi

注意用64位的python

```
import pyodbc

cfg = {'username': 'yarn', 'password': 'baofoo@64', 'database': 'default', 'dsn': 'Sample
Cloudera Impala DSN'}
connString = '''DSN={3};UID={0};PWD={1};Database=
{2}'''.format(cfg['username'], cfg['password'], cfg['database'], cfg['dsn'])
pyodbc.autocommit = True
conn = pyodbc.connect(connString, autocommit = True)
cursor = conn.cursor()

cursor.execute("SELECT code, description, total_emp, salary FROM sample_07 limit 10")
```

```
for row in cursor:  
    print(row.code, row.description)
```

window **python连接impala (安装impyla)** --测试没有成功

<https://www.cnblogs.com/dotafeiving/p/6956265.html>