

azkaban_assistant 安装步骤

一、环境准备

1、在 linux 中安装 python2.7+

- 下载地址：<https://www.python.org/ftp/python/>
- 安装方法：<https://www.cnblogs.com/simuhunluo/p/7704765.html>

2、安装 python 包管理工具 pip

在 python2.7+ 安装 bin 目录下，执行如下命令

```
wget https://bootstrap.pypa.io/get-pip.py  
  
python get-pip.py
```

3、安装 python paramiko 模块

先安装依赖包 PyCrypto：`./pip install PyCrypto`

再安装 paramiko：`./pip install paramiko`

4、安装 python 的 web 框架 tornado-4.1

- 下载地址：<https://pypi.python.org/simple/tornado/>
- 安装方法：

https://blog.csdn.net/beatle_1983/article/details/50219879

5、安装 mysql5.6+

二、下载 azkaban_assistant 安装包，在 mysql 数据库新建 azkaban 配置信息存储表：

1、azkaban_assistant 下载地址：

https://github.com/cocofree/azkaban_assistant

2、在 linux 用户目录下新建一个“azkaban_assistant”目录，将下载的

azkaban_assistant 压缩包解压在该目录：

```
unzip azkaban_assistant-master.zip
```

3、在 mysql 中新建一个 azkaban 库，并在该库中执行建表脚本：

```
use azkaban;  
source /root/azkaban_assistant/azkaban_assistant-master/resources/azkaban_create_table.sql
```

azkaban_assistant 目录中的 resources 目录下的建表 sql 脚本执行完后，
会在 azkaban 库生成 17 张新表。

三、修改配置文件

1、修改 azkaban-web 的配置文件：

- `vim /root/azkaban_assistant/azkaban_assistant-master/azkaban-web-2.5.0/conf/azkaban.properties`
- 修改内容如下（加粗部分是需要根据自己环境或需要修改的）：

```
1 #Azkaban Personalization Settings  
2 azkaban.name=Test  
3 azkaban.label=My Local Azkaban  
4 azkaban.color=#EEC900  
5 azkaban.default.servlet.path=/index  
6 web.resource.dir=web/  
7 default.timezone.id=Asia/Shanghai  
8  
9 #Azkaban UserManager class  
10 user.manager.class=azkaban.user.XmlUserManager  
11  
12 user.manager.xml.file=/root/azkaban_assistant/azkaban_assistant-master/azkaban-web-  
2.5.0/conf/azkaban-users.xml  
13  
14 #有 ldap 使用 ldap，没有的话管理用户文件  
15 #user.manager.class=azkaban.user.LdapUserManager  
16  
17 #[ldapdemo]url=>ldap://ldap.nice.com/basedn=>CN=users,CN=accounts,DC=nice,D  
#C=com  
18  
19 #user.manager.ldap.url=  
20  
21 #user.manager.ldap.basedn=  
22  
23 #Loader for projects  
24  
25 executor.global.properties=/root/azkaban_assistant/azkaban_assistant-master/azkaban-
```

executor-2.5.0/conf/global.properties

```
21 azkaban.project.dir=projects
22
23 #数据库配置
24 database.type=mysql
25 mysql.port=3306
26 mysql.host=localhost
27 mysql.database=azkaban
28 mysql.user=root
29 mysql.password=root
30 mysql.numconnections=100
31
32 # Velocity dev mode
33 velocity.dev.mode=false
34
35 # Azkaban Jetty server properties.
36 jetty.hostname=localhost
37 jetty.maxThreads=25
38 jetty.ssl.port=8443
39 jetty.port=8081
40
jetty.keystore=/root/azkaban_assistant/azkaban_assistant-master/azkaban-web-2.5.0/ke
ystore
41 jetty.password=azkaban
42 jetty.keypassword=azkaban
43
jetty.truststore=/root/azkaban_assistant/azkaban_assistant-master/azkaban-web-2.5.0/k
eystore
44 jetty.trustpassword=azkaban
45
46 # Azkaban Executor settings
47 executor.maxThreads=70
48 executor.port=12321
49 executor.flow.threads=60
50
51 # mail settings 系统发件箱配置
52 mail.sender =
53 mail.host=
54 mail.user =
55 mail.password=
56 job.failure.email=
57 job.success.email=
58
59 lockdown.create.projects=false
```

```
60
61 cache.directory=cache
```

2、修改 azkaban-web server 的用户配置文件 azkaban-users.xml :

- `vim /root/azkaban_assistant/azkaban_assistant-master/azkaban-web-2.5.0/conf/azkaban-users.xml`

- 修改内容如下（加粗部分是需要根据自己环境或需要修改的）：

```
1 <azkaban-users>
2 <user username="azkaban" password="azkaban" roles="admin" groups="azkaban" />
3 <user username="metrics" password="metrics" roles="metrics"/>
4 <user username="admin" password="admin" roles="admin,metrics" />
5 <role name="admin" permissions="ADMIN" />
6 <role name="metrics" permissions="METRICS"/>
7 </azkaban-users>
```

3、修改 azkaban-executor 的配置文件：

- `vim /root/azkaban_assistant/azkaban_assistant-master/azkaban-executor-2.5.0/conf/azkaban.properties`

- 修改内容如下（加粗部分是需要根据自己环境或需要修改的）：

```
1 #Azkaban
2 default.timezone.id=Asia/Shanghai
3
4 # Azkaban JobTypes Plugins
5 azkaban.jobtype.plugin.dir=plugins/jobtypes
6
7 #Loader for projects
8
9 executor.global.properties=/root/azkaban_assistant/azkaban_assistant-master/azkaban-executor-
10 2.5.0/conf/global.properties
11 azkaban.project.dir=projects
12
13 #数据库配置
14 database.type=mysql
15 mysql.port=3306
16 mysql.host=localhost
17 mysql.database=azkaban
18 mysql.user=root
19 mysql.password=root
20 mysql.numconnections=100
21
22 # Azkaban Executor settings
23 executor.maxThreads=70
24 executor.port=12321
```

```
23 executor.flow.threads=60
```

4、修改 schedule/conf/nice.cfg

- vim /root/azkaban_assistant/azkaban_assistant-master/schedule/conf/nice.cfg

- 修改内容如下：

```
[master_azkaban]
host=localhost
port=3306
db=azkaban
username=root
password=root
charset=utf8
[job_replace_param]
default_dir=/tmp
```

四、启动服务

1、启动 azkaban web 服务器：

- 启动命令：

```
cd /root/azkaban_assistant/azkaban_assistant-master/azkaban-web-2.5.0
bin/azkaban-web-start.sh
```

- 启动 azkaban web 服务器时报错：

```
Invalid maximum heap size: -Xmx4G

The specified size exceeds the maximum representable size.

Could not create the Java virtual machine
```

解决方法：

修改文件

azkaban_assistant/azkaban-web-2.5.0/bin/azkaban-web-start.sh，将文件中的 AZKABAN_OPTS="-Xmx4G" 修改为 AZKABAN_OPTS="-Xmx512M"。设置的大小按照机器的存储而定，如果设置太大可能无法启动，设置太小会内存溢出。

- 启动成功标志：

```
2018/11/05 01:21:59.470 +0900 INFO [AzkabanWebServer] [Azkaban] Server running on ssl port 8443.
```

2、启动 azkaban executor：

- 启动命令：

```
cd /root/azkaban_assistant/azkaban_assistant-master/azkaban-executor-2.5.0  
bin/azkaban-executor-start.sh
```

- 启动报错：

```
[root@hadoopmaster azkaban-executor-2.5.0]# Error occurred during initialization of VM  
Could not reserve enough space for 3145728KB object heap
```

解决方法：

修改文件

/root/azkaban/azkaban-executor-2.5.0/bin/azkaban-executor-start.sh

, 将 AZKABAN_OPTS="-Xmx3G"修改为 AZKABAN_OPTS="-Xmx512M"

- 启动成功标志：

```
2018/11/05 01:37:12.983 +0900 INFO [AzkabanExecutorServer] [Azkaban] Azkaban Executor Server  
started on port 12321
```

3、启动 job 管理服务：

- schedule/webapp 下执行：

startup.sh 或 restart.sh,查看 schedule_web.log 日志是否有报错。

- 如果 schedule_web.log 日志报错：No module named MySQLdb

安装 MySQL-python 模块：

```
./pip install mysql-python
```

4、登陆 https://ip 地址:8443，用户名和密码均为 admin。

先创建项目，再点击“任务配置”进行配置。

五、注意事项

1、因为所有任务都是远程服务调用，所以需要提前打通 azkaban 启动用户到各服务器的 ssh 权限！

2、邮件、短信接口已预留好，在 schdule/util/alarm.py 自定义，默认只在日志中打印出信息

3、远程脚本是否正常结束，是通过 shell 执行 script execute status ["\$?"]（最后一个命令的状态返回值）来判断的，所以下面的脚本：

```
sh test.sh

echooo 'hello world'
echo 'hello world'
```

虽然第一行出错，但最后一行继续执行并返回正确（相当于第一行的异常被捕获了），所以脚本仍认为是成功，使用的时候需要注意下。

4、在 azkaban 页面上 kill 任务时，仅会 kill 本地监听的脚本，远程脚本会继续执行下去（该特性暂时不做调整）。