

Status report on Crawling

JUNHO LEE

2018.07.22

outline

2

- ▶ 1. bash shell for Ubuntu server
- ▶ 2. Crawling status

1. bash shell for Ubuntu server

3

- ▶ ls
- ▶ cd
- ▶ pwd
- ▶ rm
- ▶ Mv
- ▶ cp / scp
- ▶ mkdir
- ▶ ssh

1. bash shell for Ubuntu server

- ▶ Access PYKU server by shell (git bash , docker quick terminal, terminal)
 - ▶ ssh [USERID@154.8.160.186](#)
 - ▶ scp localfile [USERID@154.8.160.186:/home/USERID](#)
 - ▶ Try 'ls', 'cd', 'pwd',.....
 - ▶ mysql -uroot -p
 - ▶ show databases;
 -
 -
 -

2. Crawling Status

5

- ▶ On the PYKU server do : `mysql -u root -p`
insert the password

`show database;`

`use BEIJING;`

`show tables;`

`use SEOUL;`

`show tables;`

- ▶ Serious problem on BAIDU index DATA

```
mysql> show databases;
+-----+
| Database |
+-----+
| information_schema |
| BEIJING |
| SEOUL |
| TEST |
| mysql |
| performance_schema |
| sys |
+-----+
```

```
mysql> show tables;
+-----+
| Tables_in_SEOUL |
+-----+
| PM2p5_predicted_V1 |
| Since08_18Y_Hourly |
| Since14_18Y_Daily |
| Since96_18Y_Daily |
+-----+
```

```
mysql> show tables;
+-----+
| Tables_in_BEIJING |
+-----+
| BAIDU_FeiYan |
| BAIDU_KeSou |
| BAIDU_PianTaoTiYan |
| BAIDU_SangZiTeng |
| BAIDU_YuMen |
| Since14_18Y_Daily |
+-----+
```