

Data Aggregation Manual

Last updated: 06-24-2014

1. Installing Required Softwares

This manual is using **Ubuntu 12.04 LTS** Linux distribution with Python version 2.7

1.1 Python

Issue this following command to install Python related packages:

```
$ sudo apt-get install python2.7 python-psycpg2
```

By default Python should be has already installed on the Linux system.

2. Database Deployment

Create a table on a specific database, the table schema can be found in **database.sql** file in **dataaggregator.tar.gz** tarball. Follow this command to create a database named **report** and related tables with user **postgres**

```
$ cd aggregator
$ psql -h localhost -U postgres -c "CREATE DATABASE report"
$ psql -h localhost -U postgres report < database.sql
```

3. Configuration

Configuration can be changed by editing **aggregator.conf** file, the following table describes each of configuration key.

Key name	Default value	Description
date-range-use	yes	Whether to use date-range defined by date-range-start and date-range-end, if set to 'no' will use last x hours defined in date-last-hours
date-range-start	0000-00-00:00:00	Start of date to use as starting timestamp, format is YYYY-MM-DD,HH:MM
date-range-end	0000-00-00:00:00	End of date to use as ending timestamp, format is YYYY-MM-DD,HH:MM
date-last-hours	0	Last x hours to use when date-range-use is not used (set to 'no')
coordinate-boundary	yes	Whether to use defined boundary, if set to 'no' will scan all records to find all the possible boundaries
corner1-lat	0.000000	Latitude of boundary's corner #1

corner1-long	0.000000	Longitude of boundary's corner #1
corner2-lat	0.000000	Latitude of boundary's corner #2
corner2-long	0.000000	Longitude of boundary's corner #1
client-devices	client	Client devices to be reported, multiple clients is must separated by comma and no space or tab
grid-size	10	Grid box size in meter
motionless-max-second	10	Amount of seconds for motionless
pruning-inclusion	yes	Pruning inclusion, set to 'no' to disable
dbgps-host	localhost	Database server host
dbgps-port	5432	Database server port
dbgps-user	dbgps-user	Database username
dbgps-passwd	dbgps-passwd	Database pasword
dbgps-name	dbgpsname	GPS data database name
dbgps-table	gpsdata	GPS data table name
dbreport-host	localhost	Report database server host
dbreport-port	5432	Report database server port
dbreport-user	dbgps-user	Report database username
dbreport-passwd	dbgps-passwd	Report database pasword
dbreport-name	dbreportname	Report database name
dbreport-report-table	dbreprot-report-table	Report data table name
dbreport-param-table	dbreprot-param-table	Report parameter table name

4. Running

To run the aggregator use this following command template:

```
$ ./aggregator.py <configuration-file> [report-id]
```

If report-id is defined the program will read runtime parameters in **dbreport-param-table** based on given report ID number instead of using parameters defined in configuration file. Although a report-id has defined, aggregator still need to read configuration from configuration file for GPS and report database access.

Examples:

```
$ ./aggregator.py test.conf
```

Will run with all parameters that has read from **test.conf** configuration file.

```
$ ./aggregator.py test.conf 1004
```

Run with all parameters which has report ID **1004** that has been read from parameter (report ID) table.

Example output

```
$ ./aggregator.py test.conf
Boundary: north-west=40.627813,-89.476823 south-east=40.676072,-89.501648 length=5366.156m
width=2095.031m
Adjusted Boundary: north-west=40.627813,-89.476823 south-east=40.67641,-89.501698 length=5403.74m,
width=2099.251m
Grid Box: size=100m, x=54, y=21, total=1134
Time: mode=range start=1397800200 end=1403251140 (UTC)
Clients: macbookverizon pidrive
Max. Motionless: 10 second(s)
Pruning Inclusion: yes
```

Processing client 'macbookverizon'...

```
#1 nw=40.627813,-89.476823 local=25 ucast=19 bcast=0 mcast=0 ack=19
#2 nw=40.629813,-89.476823 local=54 ucast=44 bcast=0 mcast=0 ack=44
#3 nw=40.635813,-89.476823 local=39 ucast=30 bcast=0 mcast=0 ack=30
#4 nw=40.636813,-89.476823 local=146 ucast=112 bcast=0 mcast=0 ack=112
...
```

Processing client 'pidrive'...

```
#1 nw=40.635813,-89.476823 local=393 ucast=385 bcast=160 mcast=160 ack=381
#2 nw=40.636813,-89.476823 local=536 ucast=957 bcast=416 mcast=416 ack=957
#3 nw=40.637813,-89.476823 local=10 ucast=0 bcast=0 mcast=0 ack=0
...
```

Report ID: 1015

Inserting report data...

Inserting report parameters...

4. Problem and Solving

Some known error messages

Error message	Resolution
Report ID X was not found for reuse	Given report ID was not found in report ID (parameter) table
Invalid date: Invalid date-range-start	Date range start is invalid
Invalid date: Invalid date-range-end	Date range end is invalid
Invalid date: Invalid date-last-hours	Last hours value is invalid
Invalid date: date-range-start >= date-range-end	date-range-start is more than or equal to date-range-end
Invalid boundary: north-west-x >= south-east-x	north-west-x is more than or equal to south-east-x
Invalid boundary: north-west-y <= south-east-y	north-west-y is more than or equal to south-east-y

4.2. Examine Database Record

Each record should be self-explained in **database.sql** file.