SEAD Backend User Documentation

Table of Contents

Database

Installation

Installation of the Go Landing Zone and the Python API can be automated by using the Puppet modules included in the repository. The puppet modules are written for and assume to be executed on an Ubuntu 14.04 x64 Linux server. First you must install the prerequisites for running Puppet. From the terminal, execute:

sudo apt-get install puppet git

It is recommended that you also install fail2ban with the following command:

sudo apt-get install fail2ban

Copy the Puppet files onto the server (for example, by cloning the repository) and change to the DB directory.

cd DB/

If desired, configure the UNIX application user's password in puppet/modules/config. First add the user credentials to manifests/credentials.pp, then uncomment the password definitions in config/manifests/init.pp.

cd puppet/config
vi manifests/credentials.pp
vi manifests/init.pp
cd ../..

Copy the files to the /etc/puppet directory, and execute Puppet:

sudo rsync -avc puppet/ /etc/puppet/

sudo puppet apply puppet/manifests/site.pp

After Puppet has executed the modules correctly, the server should be listening on ports 8080 and 9000. Verify with netstat:

netstat -tln | egrep ':(8080|9000)'

Python API

The Python API is used to query the raw data contained in the database. It listens for HTTP GET requests on port 8080. When a request comes in to the URI "/" it returns a usage message, in Python style. Data can be requested of the API for devices by serial number by using the URI "/serial". For example, to qeury the API running locally for data from serial 7, you would execute a GET request to 127.0.0.1:8080/7. The official server is db.sead.systems.

Filters can be applied using the following GET parameters:

- **start_time:** The earliest data point to include in the query. Specified as a UTC Unix timestamp.
- **end_time:** The earliest data point to include in the query. Specified as a UTC Unix timestamp.
- **type:** The type of data to retrieve. Current options: I for current, T for temperature, W for power, V for voltage.
- **subset:** Select a subset of the data within the given filters such that the number of elements returned are evenly spaced within the range of data returned in the filter. This option may not return exactly specified number of data points. The number of data points returned should be close to the specified number. If the same query without a subset clause would return fewer than the specified number of rows, the subset option will have no effect.
- **limit:** Retrieve only the first x rows
- **json:** Format result as JSON
- **reverse:** Return the result in reverse order

Example API calls with parameters:

- http://db.sead.systems:8080/000002?type=T
  Returns all temperature data for the SEAD plug with serial number 000002.
- http://db.sead.systems:8080/000001?start_time=1416202050&end_time=1416202054
  Returns all types of data for the SEAD plug with serial number 000001 between the UTC Unix timestamps 1416202050 and 1416202054.
- http://db.sead.systems:8080/000001?type=W&subset=100
  Returns all power data for the SEAD plug with serial number 000001 downsampled to approximately 100 data points.

The service can be started using the init script installed via Puppet:

sudo service seadapi start

The Python API is the primary interface to be used by the front end designers to query the database for SEAD plug data, and facilitates the necessary action for our first user story: *As a GUI developer, I want to query the database for signals within a certain range*.

Our final goals for the API are currently being blocked by other teams. Querying based on device type is dependent on the analysis sub-team producing tested and working modules. Querying SEAD panel data is dependent on the SEAD panel team (not affiliated with CMPS 115) finishing the data transmission component of the SEAD panel.

Once we have the API prepared to accept information from the SEAD panel and the classification scripts, we will be able to generate a signature to classify the device based on the raw data received. This will allow us to proceed with our second user story: *As a GUI developer, I want to query the database for data by appliance type.*

Go Landing Zone

The Go Landing Zone is a high speed, concurrent TCP server which listens for connections from SEAD plugs on port 9000. The plug requests its configuration from the server before beginning to send the buffered data it collects. This stream of packets is decoded, and the data is bulk inserted into the Postgres database in the background to be later queried by the API.

The service can be started using the init script installed via Puppet:

sudo service landingzone start

The landing zone is not intended to directly interacted with by a user. The data collected by the landing zone is accessible via the API. The correct way to test if the landing zone is functioning properly is to connect a SEAD plug and check if new data becomes available through the API.

Once we have the groundwork necessary to organize the protocol coordination between the SEAD panel and the landing zone, we will be able receive data from the SEAD panel and perform a classification of the raw data to determine the appliance type. This will complete our final user story: *As a SEAD panel developer, I want to send raw appliance data to the server and a classification result to be rendered in the GUI.*

Classifier

Functionality

This classifier can tell a user what type of device is currently plugged into a SEAD plug. The system comes prepared to identify only the following: lamps, laptop computers, and microwaves, all of which must draw more than 40 watts.

Steps for Use:

Identifying a device is simple!

Follow these steps:

1. Make sure your computer is connected to the internet.

2. Have ready:
   - the serial number located on the back of your SEAD plug
   - the beginning and end times for the period in which the mystery device is plugged into the SEAD plug

3. In your browser, go to http://www.epochconverter.com/ and convert your beginning and end times into Unix epoch time. You will use these times in the next step.

4. Again in the browser, visit:

   http://db.sead.systems:8080/[serial_number]?start_time=[begining_time]&end_time=[end_time]&classify=1

   Where [serial_number], [beginning_time], [end_time] should be replaced with the appropriate values, without brackets.

5. The classification of your device will pop up on the screen.
6. NOTE: At the moment, the database interface with the classifier is not yet complete. While the classifier and database both work, the interface is as yet incomplete.

Classification Method:

The classification process begins with processing a set of raw power data, collected during the specified time period. The processed data takes the shape of a power spectrum unique to its type of device. This spectrum is evaluated by the Gaussian Naive Bayes classifier built into scikit-learn. The classifier knows the probability of certain spectrum attributes being associated with certain appliances' power signatures. The probability of the unidentified spectrum being created by each appliance is calculated, and the most probable origin is given as the appliance type.

If you would like to run the classifier on your own computer:

Note

On principle, the classifying should happen in the database server.
The following instructions are provided so that you can run the the classifier "by hand", for debugging purposes.

Before Using:

You need the following version of python installed:

- **python3.x**

Install python3.x from this website: https://www.python.org/downloads/.
Alternatively, install with your system's installer (such as apt-get for Debian Linux).

You also need the following python libraries:

- **scikit-learn:** Simple and efficient tools for data analysis.
- **scipy:** Utilized buy scikit-learn
- **numpy:** Utilized buy scikit-learn
- **matplotlib:** Utilized buy scikit-learn

Method 1: Using pip.

Recommended.
1. Follow the instructions on this website: https://pip.pypa.io/en/latest/installing.html#install-pip.
2. Run this command in command-line, with the appropriate privileges, for each package you want to i nstall:

sudo pip install <package name>

Method 2: Using apt-get. (*debian-based Linux systems only*)

This is not recommended. Apt-get's default sources for these libraries are not always up-to-date. May need to use manual install to get the latest versions in order to make the classifier work.
1. Run this command with appropriate privileges:

sudo apt-get install [package name]

Method 3: Manual install.

1. Google each python library needed; go to their homepages.
2. Follow the installation instructions on each package's homepage.

Lastly, you need the server. Install the server on your own computer, following instructions in the database portion of this manual (instructions begin on page 1).