
Using Puddle at H2O

Release v1.0.3

H2O.ai

May 28, 2019

CONTENTS

1	Starting Puddle at H2O	3
1.1	Logging Out	3
2	Create New Driverless AI System	5
2.1	Viewing Driverless AI Systems	7
2.2	Starting Driverless AI	8
2.3	SSH into the Driverless AI System	9
2.4	Editing the config.toml File	9
2.5	Stopping a System	10
2.6	Starting a System	10
2.7	Rebooting a System	11
2.8	Deleting a System	11
3	Create New H2O-3 System	13
3.1	Viewing H2O-3 System Information	14
3.2	Starting H2O-3	15
3.3	SSH into the H2O-3 System	17
3.4	Stopping a System	17
3.5	Starting a System	18
3.6	Rebooting a System	18
3.7	Deleting a System	18

Welcome to Puddle at H2O, H2O.ai's platform for running Driverless AI and H2O-3. This document describes how to log in and user Puddle. Installation is not covered here. Those instructions are provided in the [Installation and Administration Guide](#).

Additional Resources

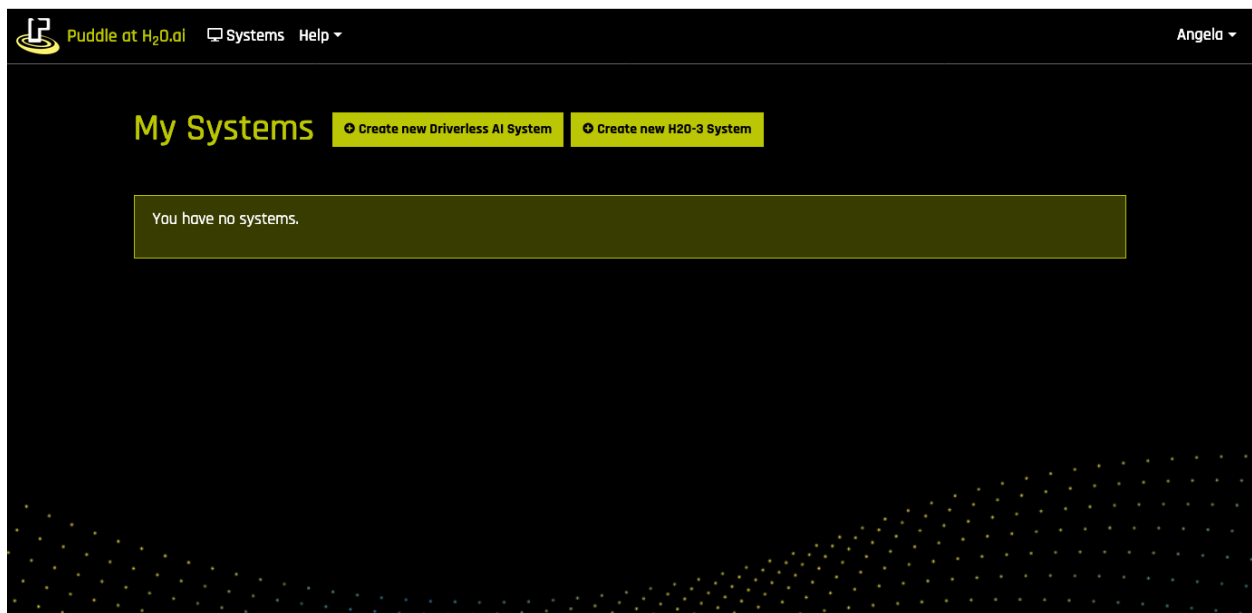
- [Driverless AI Documentation](#)
- [H2O-3 Documentation](#)

Have Questions?

If you have questions about using Puddle, post them on the [H2O.ai Community Slack workspace](#) in the **#cloud** channel. If you have not signed up for the H2O.ai Community Slack workspace, you can do so here: <https://www.h2o.ai/community/>.

STARTING PUDDLE AT H2O

Your admin will provide you with the URL for accessing Puddle. Sign in using your Microsoft Azure Active Directory account. Upon success, the home page displays and shows any current systems. From this page, you can create a new Driverless AI and/or H2O-3 system.



1.1 Logging Out

Click on your username in the upper-right corner to log out of Puddle. You will then have the option to log out of Azure as well.

CREATE NEW DRIVERLESS AI SYSTEM

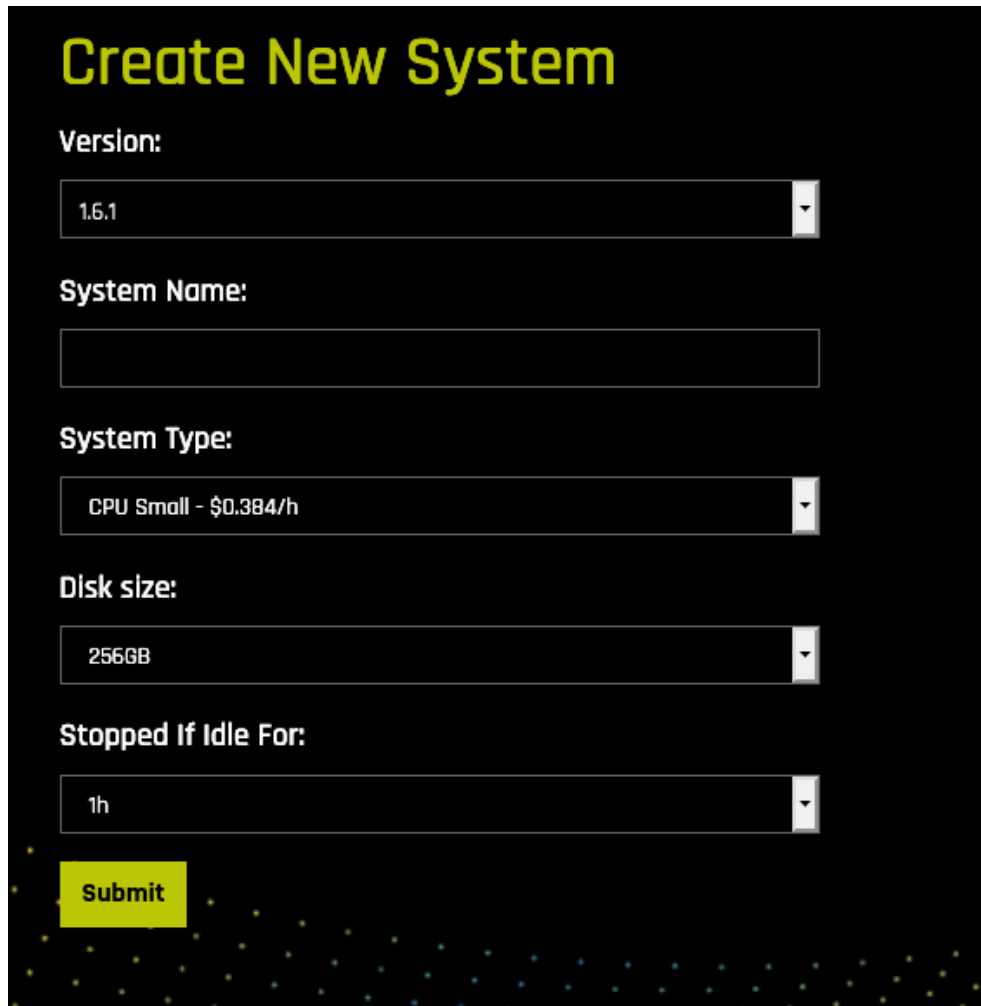
This section describes how to create a new Driverless AI system.

Note: You must have a Driverless AI license in order to run Driverless AI in Puddle. You can request a free 21-day trial license at <https://www.h2o.ai/try-driverless-ai/>.

1. Click **Create New Driverless AI System** on the Puddle Systems page.
2. Specify the following options to create the system:
 - Driverless AI Version: Available versions are added by Admins.
 - System Name: System names must be between 1 and 64 characters and contain only lowercase characters, numbers, and hyphens. It must start with a letter and end only with a number or letter. This is blank by default.
 - System Type: Please note the cost per hour for each system type.
 - CPU Small: \$0.384/h (default)
 - CPU Large: \$0.768/h
 - GPU Small: \$3.06/h
 - GPU Memory: \$4.336/h
 - GPU Compute: \$12.24/h

Note: Your account settings may include a limit as to the number of systems of a certain type that you can run. If you exceed that limit (for example, if you exceed the number of CPU Small systems that may configure for Driverless AI), then that option will not be available, and the least expensive option will then become the default.

- Disk Size: This can be 256GB (default), 512GB, or 1TB.
- Stopped If Idle For: This can be 30 min, 1 hour (default), 2 hours, 3 hours, or 4 hours.



Create New System

Version: 1.6.1

System Name:

System Type: CPU Small - \$0.384/h

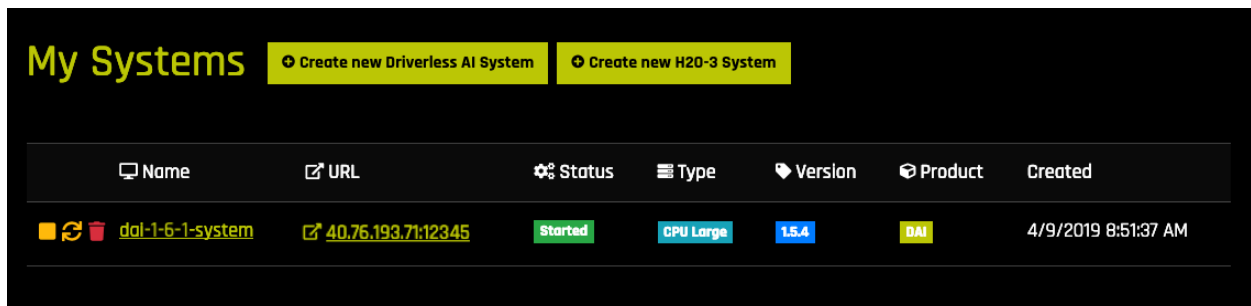
Disk size: 256GB

Stopped If Idle For: 1h

Submit




3. Click **Create System** when you are done. The system will begin provisioning. Note that this can take several minutes.

After the system has successfully started, it will appear on the My Systems page. At this point, you are ready to use Driverless AI.



My Systems

[Create new Driverless AI System](#) [Create new H2O-3 System](#)

Name	URL	Status	Type	Version	Product	Created
  dal-1-6-1-system	 40.76.193.71:12345	Started	CPU Large	1.5.4	DAI	4/9/2019 8:51:37 AM

2.1 Viewing Driverless AI Systems

Click on the Driverless AI System Name to view the configuration information and a list of current experiments.

The screenshot shows the configuration page for a Driverless AI system named 'dai-1-6-1-cpu-small'. The page is divided into two main sections: System Information and Experiments.

System Information:

- URL:** [40.76.193.71:12345](#)
- Status:** started
- Username:** `h2oa1`
- Password:** `h2oa1`
- Configuration:** [Edit config.toml](#)
- Current Session Cost:** \$0.05
- Total Cost:** \$0.29
- SSH Command:** `ssh -l angela_h2oa1.onmicrosoft.com#EXT#@msmarketplaceh2o.onmicrosoft.com 40.76.193.71`
- Stopped If Idle Till:** 4/29/2019 12:27:44 PM [Refresh timeout](#)
- Version:** 1.6.1
- Product:** DAI
- System Type:** CPU small
- Disk Size:** 256GB
- Updated:** 4/29/2019 11:36:32 AM
- Created:** 4/25/2019 2:41:53 PM

Experiments:

Description	Dataset	Target	Score	TestScore	Scorer	Progress	Status	Accuracy	Time	Interpretability	Training Duration	Created
sapavolo	CreditCard_...	default pay...	0.7784	0.7580	AUC	100%	completed	5	3	5	526.41s	4/26/2019 2:53:23 AM
pekofewe	CreditCard_...	default pay...	0.7846	NA	AUC	100%	completed	5	3	5	321.05s	4/26/2019 2:53:23 AM
ganuluhu	walmart_tr...	Weekly_Sal...	5094.8650	NA	RMSE	100%	completed	5	3	5	388.17s	4/25/2019 2:53:21 PM

This page provides general system information and Driverless AI model information (if any models exist)

2.1.1 System Information

- The URL for launching Driverless AI.
- The system status.
- The Username and Password for starting Driverless AI.
- An link to edit the config.toml file for that system.
- The current session cost.
- The total cost so far for this system.
- The SSH command to run in order to securely access the system that is running Driverless AI. (See *SSH into the Driverless AI System* below for more information.)
- The time when the system will stop if remaining idle. You can also refresh this timer.
- The product name and version currently running on the system.
- The system type and disk size.
- The updated and created dates.

2.1.2 Experiment Information

For each experiment run on Driverless AI through Puddle, the following information displays:

- A description that includes the experiment key.
- The training dataset used in the experiment.
- The target column.
- The validation score.
- The test scorer.
- The scorer used for the experiment.
- The experiment progress and status.
- The Accuracy, Time, and Interpretability options used for the experiment.
- The amount time it took to complete the experiment (in seconds).
- The time when the experiment was created.

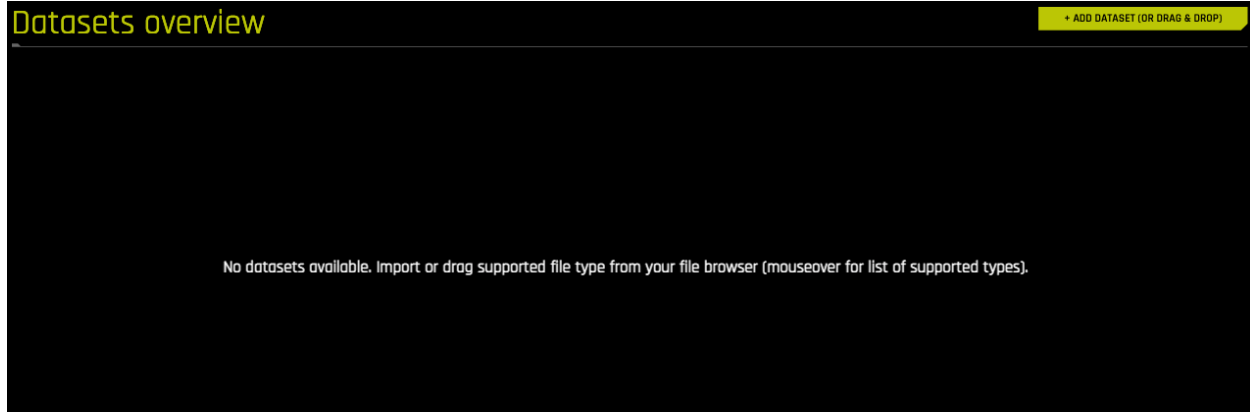
2.2 Starting Driverless AI

1. Click on the URL provided in the Driverless AI system page. This takes you to the DNS of the URL.
2. If this is your first time starting Driverless AI on this system, or if you have restarted the system, accept the license agreement.
3. Enter the Username and Password that are provided on the Driverless AI system page.



4. If this is your first time starting Driverless AI on this system, enter your license key. Note that if you do not have a license key, you can request a free 21-day trial license at <https://www.h2o.ai/try-driverless-ai/>.

Upon completion, Driverless AI will open on the Datasets Overview page. At this point, you can add or upload datasets and begin running experiments.



In Driverless AI, click on **Resources > Help** to view the Driverless AI User Guide. Additional documentation for Driverless AI is available at docs.h2o.ai.

2.3 SSH into the Driverless AI System

Puddle provides the ability to SSH into a system that is running Driverless AI.

1. Select the system that you want to SSH into.
2. On your local machine, run the provided SSH command.
3. You will be prompted to continue the connection. Type `yes`.
4. After the URL is added to your list of known hosts, you will be provided with a login URL and password. The message will be similar to the following:

```
To sign in, use a web browser to open the page https://microsoft.com/
↪devicelogin
and enter the code AZSVUQA9X to authenticate. Press ENTER when ready.
```

Open a browser and follow the instructions in the terminal message.

5. After you correctly enter your login and password in your browser, the following message will display.



Microsoft Azure Linux Virtual Machine Sign-In

You have signed in to the Microsoft Azure Linux Virtual Machine Sign-In application on your device. You may now close this window.

©2019 Microsoft [Terms of use](#) [Privacy & cookies](#) ...

Return to your terminal window and press Enter. Upon completion, you will receive a welcome message.

2.4 Editing the config.toml File

In Driverless AI, the `config.toml` file allows you to specify system-wide configuration options. These options are specified using environment variables. Perform the following steps to edit the `config.toml` file. Note that a system reboot (🔄) is required when changes are made to the `config.toml` file.

There are two ways to open the config.toml editor:

- Click the **Edit config.toml** link in the Driverless AI System information table.



- Click the **Edit config.toml** button (⚙️) beside your system name on the Driverless AI System page.

1. Specify the environment variables that you want to include.
2. Click **Submit**.
3. Reboot the system.

A list of available environment variables is included in the Driverless AI documentation for the [config.toml file](#). Note that this link points to the latest version of the config.toml file. The User Guide that's available in your system's Driverless AI under **Resources > Help** includes the config.toml file that matches your running Driverless AI version.




2.5 Stopping a System

Click the Stop button (⏹️) to halt a system that is in a “Started” state. No information will be lost when a system is stopped.


2.6 Starting a System

Click on the Start button (▶️) to start a system that is in a “Stopped” state. This will launch a new system with a new URL. All prior data will still be available from this URL.

2.7 Rebooting a System

Click the Reboot button () to reboot a system that is in a “Started” state. This will stop the system and launch a new system with a new URL. All prior data will still be available from the new URL. The entire process can take several minutes. **Note:** A reboot is required when you change the config.toml file.

2.8 Deleting a System

Click the Delete button () to completely remove a system. A confirmation page will display asking if you are certain about deleting the system. Click **Yes** complete the delete. This request deletes the system and destroys all data that is on the system.

CREATE NEW H2O-3 SYSTEM

This section describes how to create a new H2O-3 system. H2O-3 is open source and does not require a license.

1. Click **Create New H2O-3 System** on the Puddle Systems page.
2. Specify the following options to create the system:
 - H2O-3 Version: This defaults to the most current version of H2O-3.
 - System Name: This is blank by default.
 - System Type: Please note the cost per hour for each system type.
 - CPU Small: \$0.384/h (default)
 - CPU Large: \$0.768/h
 - GPU Small: \$3.06/h
 - GPU Memory: \$4.336/h
 - GPU Compute: \$12.24/h

Note: Your account settings may include a limit as to the number of systems of a certain type that you can run. If you exceed that limit (for example, if you exceed the number of CPU Small systems that may configure for H2O-3), then that option will not be available, and the least expensive option will then become the default.

- Disk Size: This can be 256GiB (default), 512GiB, or 1TiB.
- Stopped If Idle For: This can be 30 min, 1 hour (default), 2 hours, 3 hours, or 4 hours.

Create New System

Version:
3.24.0.2

System Name:

System Type:
CPU Large - \$0.768/h

Disk size:
256GB

Stopped If Idle For:
1h

Submit **Cancel**

3. Click **Create System** when you are done. The system will begin initializing. Note that this can take several minutes.

3.1 Viewing H2O-3 System Information

Click on the H2O-3 system to view its configuration information.

h2o-3-24-0-2-cpu-large	
URL	40.117.56.171:54321
Status	Started
Current Session Cost	\$0.55
Total Cost	\$0.55
SSH Command	<code>ssh -l angela_h2oai.onmicrosoft.com#EXT#@msmarketplaceh2o.onmicrosoft.com 40.117.56.171</code>
Stopped If Idle Till	5/2/2019 9:26:39 AM Refresh timeout
Version	3.24.0.2
Product	H2O-3
System Type	CPU Large
Disk Size	256GB
Updated	5/2/2019 8:26:39 AM
Created	5/2/2019 8:23:47 AM

3.1.1 System Information

- The URL for launching H2O-3. Click on the URL to open Flow. This takes you to the DNS of the URL. **Note:** You can also download the H2O-3 Python or R clients from <https://www.h2o.ai/download/>, and use this URL to launch H2O-3 in either Python or R. But be sure that the H2O-3 version in Puddle and the version downloaded locally match, or you will receive a version mismatch error.
- The system status.
- The current session cost.
- The total session cost.
- The SSH command to run in order to securely access the system that is running H2O-3. (See *SSH into the H2O-3 System* below for more information.)
- The time when the system will stop if remaining idle. You can also refresh this timer.
- The product name and version currently running on the system.
- The system type and disk size.
- The updated and created dates.

3.2 Starting H2O-3

After the system has completed initializing, you will be ready to use H2O-3. As indicated previously, H2O-3 can be started in Flow, Python, or R.

3.2.1 Starting H2O-3 in Flow

Click on the URL to launch H2O Flow. From Flow, you can begin adding datasets and building models. Refer to the Flow section in the *H2O-3 User Guide* for information on how to use H2O-3 in Flow.

3.2.2 Starting the H2O-3 Python Client

Perform the following steps to start the H2O-3 Python client. Run these commands in a Terminal window.

Note: These steps assume that the H2O-3 system created in Puddle is using the latest H2O-3 release. If your H2O-3 version in Puddle is not the latest, then refer to the [H2O-3 Prior Releases](#) page for information on downloading the H2O-3 Python client for your specific version.

1. Install dependencies (prepending with `sudo` if needed):

```
pip install requests
pip install tabulate
pip install "colorama>=0.3.8"
pip install future
```

Note: These are the dependencies required to run H2O. A complete list of dependencies is maintained in the following file: <https://github.com/h2oai/h2o-3/blob/master/h2o-py/conda/h2o/meta.yaml>.

2. Run the following command to remove any existing H2O module for Python.

```
pip uninstall h2o
```

3. Use `pip` to install this version of the H2O Python module.

```
pip install -f http://h2o-release.s3.amazonaws.com/h2o/latest_stable_Py.html ↵
↪h2o
```

Note: When installing H2O from `pip` in OS X El Capitan, users must include the `--user` flag. For example:

```
pip install -f http://h2o-release.s3.amazonaws.com/h2o/latest_stable_Py.html ↵
↪h2o --user
```

4. Initialize H2O in Python and optionally run a demo to see H2O at work. In the `h2o.init()` command, be sure to change `<h2o_3_system_address>` to the URL provided in Puddle for this H2O-3 system.

```
python
import h2o
h2o.connect(ip="<h2o_3_system_address>", port=54321)
h2o.demo("glm")
```

3.2.3 Starting the H2O-3 R Client

Perform the following steps to start the H2O-3 Python client. Run the commands in R, copying and pasting these commands one line at a time.

Note: These steps assume that the H2O-3 system created in Puddle is using the latest H2O-3 release. If your H2O-3 version in Puddle is not the latest, then refer to the [H2O-3 Prior Releases](#) page for information on downloading the H2O-3 R client for your specific version.

1. The following two commands remove any previously installed H2O packages for R.

```
if ("package:h2o" %in% search()) { detach("package:h2o", unload=TRUE) }
if ("h2o" %in% rownames(installed.packages())) { remove.packages("h2o") }
```

2. Next, download packages that H2O depends on.

```
pkgs <- c("RCurl","jsonlite")
for (pkg in pkgs) {
  if (! (pkg %in% rownames(installed.packages()))) { install.packages(pkg) }
}
```

3. Download and install the H2O package for R.

```
install.packages("h2o", type="source", repos=(c("http://h2o-release.s3.
↪amazonaws.com/h2o/latest_stable_R")))
```

4. Initialize H2O and optionally run a demo to see H2O at work. In the `h2o.init()` command, be sure to change `<h2o_3_system_address>` to the URL provided in Puddle for this H2O-3 system.

```
library(h2o)
h2o.connect(ip="<h2o_3_system_address>", port=54321)
demo(h2o.kmeans)
```

3.3 SSH into the H2O-3 System

Puddle provides the ability to SSH into a system that is running H2O-3.

1. Select the system that you want to SSH into.
2. On your local machine, run the provided SSH command.
3. You will be prompted to continue the connection. Type `yes`.
4. After the URL is added to your list of known hosts, you will be provided with a login URL and password. The message will be similar to the following:

```
To sign in, use a web browser to open the page https://microsoft.com/
↪devicelogin
and enter the code AZSVUQA9X to authenticate. Press ENTER when ready.
```

Open a browser and follow the instructions in the terminal message.

5. After you correctly enter your login and password in your browser, the following message will display.



Microsoft Azure Linux Virtual Machine Sign-In

You have signed in to the Microsoft Azure Linux Virtual Machine Sign-In application on your device. You may now close this window.


©2019 Microsoft [Terms of use](#) [Privacy & cookies](#) ...

Return to your terminal window and press Enter. Upon completion, you will receive a welcome message.


3.4 Stopping a System

Click the Stop button (■) to halt a system that is in a “Started” state. No information will be lost when a system is stopped.


3.5 Starting a System

Click on the Start button () to start a system that is in a “Stopped” state. This will launch a new system with a new URL. All prior data will still be available from this URL.

3.6 Rebooting a System

Click the Reboot button () to reboot a system that is in a “Started” state. This will stop the system and launch a new system with a new URL. All prior data will still be available from the new URL. The entire process can take several minutes. **Note:** A reboot is required when you change the config.toml file.

3.7 Deleting a System

Click the Delete button () to completely remove a system. A confirmation page will display asking if you are certain about deleting the system. Click **Yes** complete the delete. This request deletes the system and destroys all data that is on the system.