This page intentionally left blank.
Install from Amazon Marketplace

Contents:
- Product Limitations
- Install
  - Desktop Requirements
  - Sizing Guide
  - Prerequisites
  - Install Steps
- Documentation

This guide steps through the requirements and process for installing Trifacta® Wrangler Pro through the Amazon Marketplace.

Product Limitations

- Jobs must be executed on the Trifacta Server. No other running environment integrations are supported.
- Anomaly and stratified sampling are not supported in this deployment.
- When publishing single files to S3, you cannot apply an `append` publishing action.
- The EC2 instance, S3 buckets, and any connected Redshift databases must be located in the same Amazon region. Cross-region integrations are not supported at this time.

Install

Desktop Requirements

- All desktop users of the platform must have the latest version of Google Chrome installed on their desktops.
  - Google Chrome must have the PNaCl client installed and enabled.
  - PNaCl Version: 0.50.x.y or later
- All desktop users must be able to connect to the EC2 instance through the enterprise infrastructure.

Sizing Guide

**NOTE:** The following guidelines apply only to Trifacta Wrangler Pro.

Use the following guidelines to select your instance size:

**NOTE:** Trifacta Wrangler Pro enforces a maximum limit of 30 users.

<table>
<thead>
<tr>
<th>Size</th>
<th>max recommended users</th>
<th>Avg. size of jobs on Trifacta Server (GB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>5</td>
<td>5 GB</td>
</tr>
<tr>
<td>Medium</td>
<td>15</td>
<td>20 GB</td>
</tr>
<tr>
<td>Large</td>
<td>30</td>
<td>40 GB</td>
</tr>
</tbody>
</table>

Pre-requisites

Before you install the platform, please verify that the following steps have been completed.

1. **EULA.** Before you begin, please review the End-User License Agreement. See End-User License Agreement.
2. **S3 bucket.** Please create an S3 bucket to store Trifacta assets. In the bucket, the platform stores metadata in the following location:
3. **IAM policies.** Create IAM policies for access to the S3 bucket. Required permissions are the following:
   - The system account or individual user accounts must have full permissions for the S3 bucket:
     
     Delete*, Get*, List*, Put*, Replicate*, Restore*
   - These policies must apply to the bucket and its contents. Example:
     
     "arn:aws:s3:::my-trifacta-bucket-name"
     "arn:aws:s3:::my-trifacta-bucket-name/*"
   - See [https://console.aws.amazon.com/iam/home#/?policies](https://console.aws.amazon.com/iam/home#/?policies)

4. **EC2 instance role.** Create an EC2 instance role for this policy. See [https://console.aws.amazon.com/iam/home#/?roles](https://console.aws.amazon.com/iam/home#/?roles).
Install Steps

1. Launch Trifacta Wrangler Pro from AMI.
2. In the EC2 Console:
   a. Instance size: Select the instance size. See Sizing Guide.
   b. Network: Configure the VPC, subnet, firewall and other configuration settings necessary to communicate with the instance.
   c. Auto-assigned Public IP: You must create a public IP to access the Trifacta platform.
   d. EC2 role: Select the EC2 role that you created.
   e. Local storage: Select a local EBS volume. The default volume includes 100GB storage.

   NOTE: The local storage environment contains the Trifacta databases, the product installation, and its log files. No source data is ever stored within Trifacta Wrangler Pro.

   f. Security group: Use a security group that exposes access to port 3005, which is the default port for the platform.
   g. Create an AWS key-pair for access: This key is used to provide SSH access to the platform, which may be required for some admin tasks.
   h. Save your changes.
3. Launch the configured version of Trifacta Wrangler Pro.

   NOTE: From the EC2 Console, please acquire the instanceId, which is needed in a later step.

4. When the instance is spinning up for the first time, performance may be slow. When the instance is up, please navigate to the following:

   http://<public_hostname>:3005

5. When the login screen appears, enter the following:
   a. Username: admin@trifacta.local
   b. Password: (the instanceId value)

   NOTE: As soon as you login as an admin for the first time, you should immediately change the password. Select the User Profile menu item in the upper-right corner. Change the password and click Save to restart the platform.

6. From the application menu, select User menu > Admin Settings.
7. In the Admin Settings page, you can configure many aspects of the platform, including user management tasks, and perform restarts to apply the changes.
   a. In the Search bar, enter the following:

      aws.s3.bucket.name

   b. Set the value of this setting to be the bucket that you created for Trifacta Wrangler Pro.
8. The following setting must be specified.

   "aws.mode":"system",

   You can set the above value to either of the following:

<table>
<thead>
<tr>
<th>aws.mode</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>system</td>
<td>Set the mode to system to enable use of EC2 instance-based authentication for access.</td>
</tr>
<tr>
<td>user</td>
<td>Set the mode to user to utilize user-based credentials.</td>
</tr>
</tbody>
</table>
Details on the above configuration are described later.

9. Click **Save**.
10. When the platform restarts, you can begin using the product.

**SSH Access**

To access the Trifacta node using SSH:

```
ssh -i <path_to_key_file> centos@<tri_node_DNS_or_IP>
```

where the path to the key file is on your local system and the Trifacta node’s DNS or IP address is available through AWS.

**Documentation**

You can access complete product documentation online and in PDF format. From within the product, select **Help menu > Product Docs**.