

Redhat Upgrade Guide

erwin Data Quality V3.1.0

Prepared by



Introduction	1
Software Requirements (Auto Install)	2
PostgreSQL Installation (Only applicable for DB isolated Deployment)	2
Pre-Installation Setup	5
Configuration Setup	6
Details about the config file	6
Installation Process	7
Verification of Installation	8
Post-Installation Procedure	8

Introduction

This document provides a comprehensive, step-by-step process for upgrading DQLabs on Red Hat Server 9 from version 2.4.6 to version 3.1.0. It is designed for personnel with technical knowledge of Red Hat Server 9 and Linux Operating Systems.

The guide walks you through the installation steps, ensuring that you have a clear understanding of the process and any prerequisites specific to Redhat. By following this guide, you will be able to successfully set up the DQLabs 3.1.0 application and leverage its features in your Redhat Server 9 environment.

System Requirements

This section provides the minimum system and mandatory requirements that are needed to install the DQLabs application in the Linux environment successfully.

Category	Recommended
Operating system	Red Hat Enterprise Linux 9.x (Supported on Azure, AWS, and on-prem VMs) If AWS, use Amazon Machine Image (AMI) (HVM) with SSD Volume Type.
Processor	64-bit processor
Disk Space	100 GB, and it should be /directory or DQLabs installing folder, and the install user home directory should have at least 5GB

Package	Core and RAM Specifications
Bronze	4 Core 8 GB RAM
Silver	4 Core 8 GB RAM
Gold	8 Core 16 GB RAM
Platinum	16 Cores 32 GB RAM
Titanium	32 Cores 64 GB RAM

Postgres Server Prerequisites (Only in case of DB Isolated Deployments)

Operating System	Redhat 9.x
CPU Core	4 Cores or more
RAM	8GB or more

- Ensure to take a snapshot of the server before attempting to upgrade
- Dedicated Server:** DQLabs needs to be installed on a dedicated server.
- Internet Access:**

URL	Purpose	Required during
https://license-qa.dqlabs.cloud	This URL must be whitelisted to activate and manage the validity of the license key.	Required after Installation
https://s3.amazonaws.com	This URL must be whitelisted to allow binaries to be downloaded from the DQLabs repository.	Required only before installation. The file can be downloaded externally and moved to the server if needed
cdn.redhat.com subscription.rhsm.redhat.com	Official RHEL repositories are accessed to download and update the necessary packages Command	These repositories can be limited after updates and before installation
pypi.org pypi.python.org files.pythonhosted.org pythonhosted.org	These URL is required to download the Python packages.	Required after Installation

- License key:** A new DQ license key is required for activating DQLabs 3.1.0 upon upgrade

5. Ports to be opened: Ports Used for Internal Communication within the Application:

PostgreSQL	5432 (Mandatory)
Airflow	8080
HTTP	80
HTTPS	443

Software Requirements (Auto Install)

The requirements mentioned below are auto-installed with the script; the user should not manually install any of the software requirements in the DQServer

Services	Version
PostgreSQL	15.12
Python	3.11.9
Java	17.0.14
Airflow	2.8.1
Drivers	MSSQL, Oracle, PostgreSQL, MySQL ODBC/JDBC
Spark	3.5.1

Before proceeding with the upgrade, ensure the following:

1. PostgreSQL should be available on the same server from the fresh installation of erwin DQ 2.4.6. Upgradation will not accommodate Dual server deployments.

2. No third-party or external applications should be present on the server.

3. Database cleanup process
- Follow these steps to identify and remove excess rows from the core.drift_alert table in the dqlabs database.
- **Connect to the database**

None

```
sudo -u postgres psql -d dqlabs
```

When prompted, enter the database password (Password: Intel1234)

- **Delete groups exceeding 5,000 alerts**

None

```
#Run this query to remove all rows from core.drift_alert where the same (measure_id, run_id) appears more than 5,000 times
DELETE FROM core.drift_alert AS da
USING (
  SELECT
    drift_alert.measure_id,
    drift_alert.run_id
  FROM core.drift_alert
  JOIN core.metrics
    ON metrics.measure_id = drift_alert.measure_id
    AND metrics.run_id      = drift_alert.run_id
  GROUP BY
    drift_alert.measure_id,
    drift_alert.run_id
  HAVING COUNT(*) > 5000
) AS sub
WHERE
  da.measure_id = sub.measure_id
  AND da.run_id  = sub.run_id;
```

- **Exit psql**

None

```
\q
```

Pre-Installation Setup

Step 1: Log in to the Application server with appropriate privileges and install wget & the p7zip Plugin using the below command:

```
None
sudo yum install wget -y
```

Step 2: Download the Prerequisite File from S3 into the directory where the application should be installed

```
None
wget
https://s3.us-east-1.amazonaws.com/erwin-2.0/code/linux/application-code/3.1.0/On\_premise/Packages\_3.1.0/Erwin-Redhat-Upgrade-Packages/Erwin-prerequisites.tar
```

Step 3: Extract the Downloaded Tar File

```
None
tar -xvf Erwin-prerequisites.tar
```

Step 4: Execute the following commands:

```
None
#Navigate to the p7zip_file folder and Install the p7zip Plugin
cd prerequisites/p7zip_file

#Install the p7zip plugin:
sudo rpm -ivh *.rpm

#Remove the redundant tar file
sudo rm -rf ~/Erwin-prerequisites.tar

#Return to Prerequisites Folder
cd ~
```

Step 5: Download the Installation File by using the commands below into the directory where the application should be installed

```
None
wget
https://s3.us-east-1.amazonaws.com/erwin-2.0/code/linux/application-code/3.1.0/On\_premise/Packages\_3.1.0/Erwin-Redhat-Upgrade-Packages/Erwin-installer.7za
```

Step 6: Extract the Installation File by using the commands below:

```
None
sudo 7za x Erwin-installer.7za
```

Step 7: Remove the redundant zip file after extraction

```
None
sudo rm -rf Erwin-installer.7za
```

Configuration Setup

Step 1: Navigate to the Pre-requisites directory

```
None

#If the application is installed on home directory
cd prerequisites/

#If the application is installed on a custom directory
cd <custom_directory>/prerequisites/
```

Ensure that the config.txt and Erwin-installer.sh are in the same directory

```
[ec2-user@ip-172-31-78-245 ~]$ cd prerequisites/
[ec2-user@ip-172-31-78-245 prerequisites]$ ls
config.txt  Erwin-installer-main.sh  Erwin-installer-sub.sh  p7zip_file
[ec2-user@ip-172-31-78-245 prerequisites]$
```

Step 2: Edit the config.txt file using the command below:

```
None

sudo vi config.txt
```

Step 3: Update the config.txt file by clicking “I” to get into insert mode

Details about the config file

- 1. **Source and Destination Locations:** Please verify that the SOURCE_LOCATION path specified in the config.txt is accurate and ensure the installation file is downloaded correctly to that location.
 - a. **SOURCE_LOCATION:** Define the source location path.
Example - (/home/ec2-user/Erwin-installer)
 - b. **DESTINATION_LOCATION:** Define the new destination location path (Provide a new directory)
Example - (/home/ec2-user/erwindq)
- 2. **Rollback Options:**
 - a. **ROLLBACK_ENABLED:** yes or no
 - i. **yes** - To Rollback to the previous version (Used only after upgrade when you prefer to go back to the older version)
 - ii. **no** - To proceed with the upgrade

```
##### Erwin #####

# shellcheck disable=SC2034
SOURCE_LOCATION=/home/ec2-user/Erwin-installer
DESTINATION_LOCATION=/home/ec2-user/DQ310

# Rollback Inputs
ROLLBACK_ENABLED=no

~
```

Step 4: Go to command mode by pressing the escape key. Save and exit the editor using :wq! and press Enter

Installation Process

Step 1: Set Permissions and execute the script using the commands below:

```
None

#Navigate to the prerequsites directory
cd prerequisites

#Grant permissions to the script
sudo chmod 777 Erwin-installer-main.sh
sudo chmod 777 Erwin-installer-sub.sh
```

```
#Execute the setup script
./Erwin-installer-main.sh

#After installing all the required services, proceed to execute the sub-script for migration
./Erwin-installer-sub.sh
```

Verification of Installation

Once the installation process is complete, it is essential to verify that the DQLabs application and related services are running successfully on your Red Hat machine. Perform the following verifications:

Step 1: Ensure the following output is shown at the end of the script execution

```
airflow-webserver is running.
airflow-scheduler is running.
apache2 is running.
dqlabs is running.
true
*****
All services are running. DQ installation completed successfully!
*****
```

Note: The execution log will be stored on the server and can be accessed using the following commands:

```
None

#View the logs
vi script.log
```

Patch Process

Step 1: Download the Patch script into the DQ Server under the same user with which DQ was installed

```
None
wget
https://s3.us-east-1.amazonaws.com/erwin-2.0/code/linux/application-code/3.1.0/On_premise/erwin_v310_patchscript.sh
```

Step 2: Open the script file

```
None
sudo vi erwin_v310_patchscript.sh
```

Step 3: Update the variables

```
None
#Provide the directory path where the script is downloaded
PATCH_DIR="/home/dqlabs/"

#Provide the path to Hitach-server folder
TARGET_DIR="/home/dqlabs/dq310/DQLabs-Server"

#Provide the destination directory of the application (Same as TARGET_DIR, exclude DQLabs-Server)
PARENT_DIR="/home/dqlabs/dq310"
```

Step 4: Save the file and exit

None
:wq!

Step 5: Grant permission to execute the script

None
sudo chmod 777 erwin_v310_patchscript.sh

Step 6: Execute the script

None
./erwin_v310_patchscript.sh

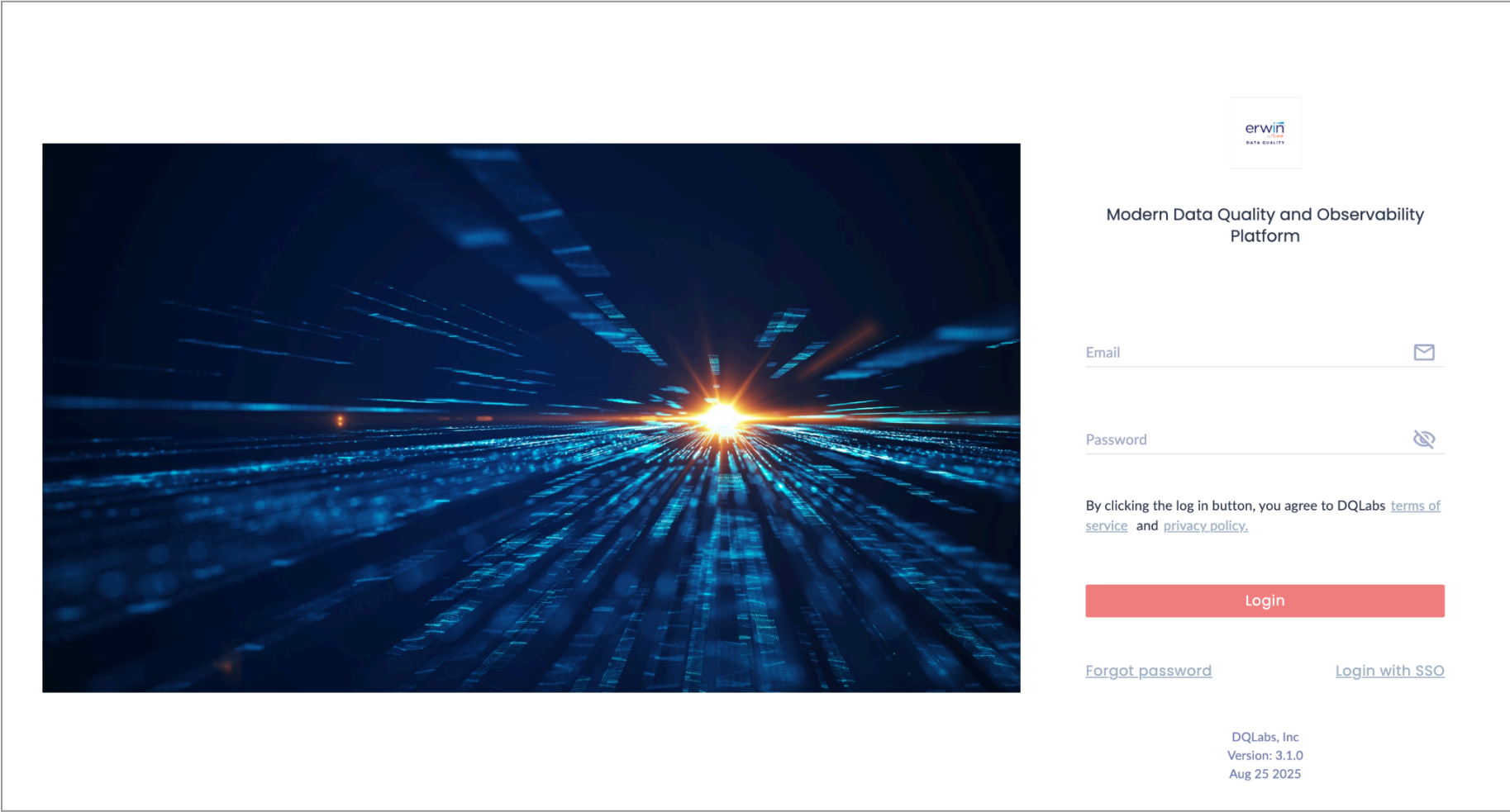
Post-Installation Procedure

After completing the installation and verifying the successful setup of DQLabs on your RedHat machine, you can now log in to the application and proceed with the final steps. Follow the instructions below:

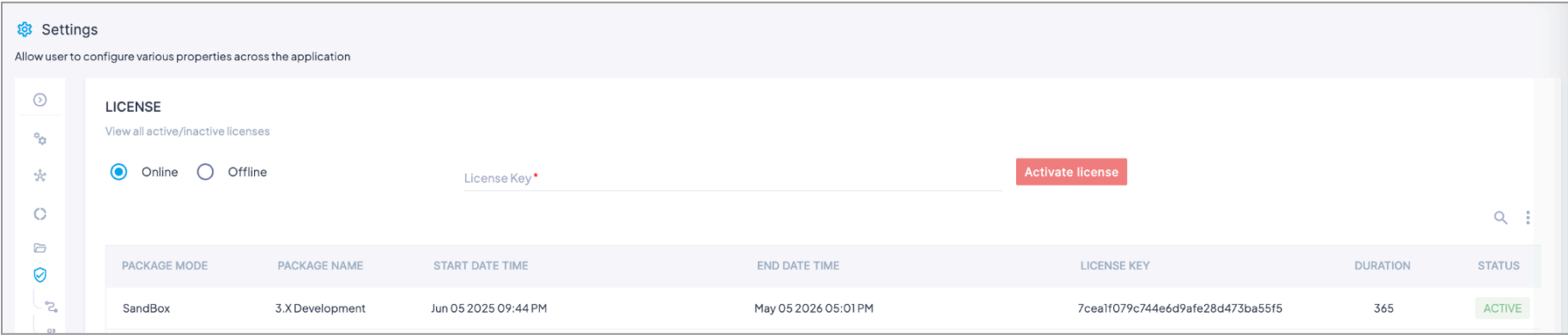
Step 1: Launch any supported web browser on your machine

Step 2: In the address bar of the browser, enter the IP address or DNS name used during installation

Step 3: The browser will load the DQLabs application, and you will be presented with the login page



Step 4: After logging in to the application, navigate to Settings > Security > License. Activate your new 3x license key.



Migration Notes:

The following modules will not be carried over after the upgrade to version 3.1.0:

Items that are not migrated	Description
Custom User Roles	User-based access control in version 3.1.0 differs from version 2.4.6 and cannot be directly migrated. While users themselves will be migrated, they will be assigned a default user role. Any custom roles that need to be recreated will require manual reassignment to the appropriate users.
Dashboards and Reports	Due to structural changes in the UI and visual reporting framework, custom dashboards and reports will not be migrated and will need to be recreated in the new version.
User Activity log	Changes in the way the logs are captured between the versions