LexPredict ContraxSuite Documentation

Quick Start Installation Guide for Linux

Release 1.0.4 - December 1, 2017

Getting Started

Licensing Information

ContraxSuite is available under a dual-license open-source model. Unless otherwise released by ContraxSuite, LLC, you can use and modify this software under the terms of the GNU AFFERO GENERAL PUBLIC LICENSE.

If you have question about licensing or would like to request a release from the AGPL license, please email license@contraxsuite.com.

ContraxSuite also relies on a software and data dependencies that are independently licensed. For more information about these dependencies and their respective licensing models, please refer to the Software and Data Dependency documentation in this release.

Support, Customization, Hosting, or Training

We can help! If you need assistance customizing, hosting, training, or supporting your ContraxSuite instance, please reach out to us to discuss options. You can always email us support@contraxsuite.com or create issue on Github.

Prerequisite Skills

This document assumes basic familiarity with the installation and configuration enterprise multi-tier applications. In particular, personnel with Linux, Java, and Python experience are best suited to successfully complete the installation process.

How to Get Support

For support or help in setting up the application, please contact support@contraxsuite.com.

Scope

This Quick Start Installation Guide is designed to meet the needs of two use cases:

- Local environments for **Developers** or **Researchers**
- **Evaluation** and **testing** by organizations

The end result of this process is an environment with:

- All services installed on a single server (Single Point of Failure)
- High-Availability (HA) and Resiliency not configured
- Backups and Disaster Recovery (DR) not configured
- Data Encryption not configured
- Performance and Load Testing not completed

CAUTION: If you are planning to use ContraxSuite in production or for real projects, you should carefully evaluate whether this architecture meets your requirements.

NOTE: As of Release 1.0.5, this installation procedure via Fabric will be replaced with Docker as the default installation process. If you are planning long-term support for your own ContraxSuite instance, you should consider which approach makes more sense for your environment.

Quick Start Installation Guide for Linux (64-bit)

Server Requirements

While ContraxSuite can be run on a variety of Linux distributions and versions, we recommend the following configuration for development and testing:

- Operating System: Ubuntu 16.04 LTS 64-bit
- Virtualization Supported: Yes
- Minimum Requirements:
    - CPU/vCPU: 4 cores / 4 vCPU
    - RAM: 8GB
    - Disk: 40GB SSD or more
    - IP Addresses: 1 static

Pre-Installation Tasks

There are only a few pre-installation tasks to complete when performing a Quick Start Installation.

Storage Architecture

ContraxSuite utilizes three data stores - a traditional relational database, a message broker, and an enterprise search index. While the persistence of data is important for all three of these systems, the primary storage burden is placed on the relational database and enterprise search index.

The quick start deployment does not allow for the configuration of storage architecture. All tasks will assume default package installation paths and install accordingly.

Quick Start Deployment

Purpose	Suggested Mountpoint	Suggested Type
OS, Application, RDBMS, etc.	/	SSD

From a sizing perspective, requirements will vary widely based on the type of document, number of documents, and ContraxSuite functionality applied. ContraxSuite and its software dependencies require approximately 2GB for complete installation However, the table below provides estimates for additional storage requirements per thousand documents:

Sizing Requirements per 1000 Documents (KiloDoc)

System	GB per KiloDoc
RDBMS	0.5
Enterprise Search	0.75

Network Architecture

ContraxSuite can be deployed either on a single server or flexibly distributed across two or more servers. The Quick Start Installation architecture relies on a single-server deployment, so no additional network configuration is required for ContraxSuite itself. However, in order for users to access the Quick Start instance over the network, HTTP or HTTPS traffic need to be allowed

Required System Software

The following software is required to execute the actual installation steps of the process below. Please refer to your underlying operating system support documentation or Google for instructions on installing these packages if not available.

- OpenSSH Client and Server
- git
- Python 3.x
- "sudo" or administrative privileges

Installing ContraxSuite

The Quick Start Installation process can be executed with an automated deployment script. This deployment script requires minimal configuration and manages the entire, end-to-end process of assessing system state, installing required dependencies, deploy ContraxSuite software, and performing basic network and user configuration.

The deployment automation software supports two models - Local Machine and Remote Machine installation. Local Machine installation executes the steps directly on the target server, whereas the Remote Machine installation executes installation and configuration over a remote SSH. Both methods rely on Fabric, a deployment and automation framework for Python.

Local Machine Installation

Edit Base Fabric Configuration

Local Machine installation first requires configuration of the base fabricrc configuration. This file can be found under the deployment repository at the following path:

- `base/fabricrc`

The following steps need to be performed in this file:

1. Change database credentials. The default database credentials are insecure defaults that should be updated.
2. Set the web application superuser credentials, including username, password, email.

Edit Local Installation Fabric Configuration

Local Machine installation next requires configuration of the local installation fabricrc configuration. This file can be found under the deployment repository at the following path:

- `local/fabricrc`

The following steps need to be performed in this file:

1. Set user and sudo password.
2. Change `public_ip`, `dns_name` to match desired network adapter and address.
3. Confirm `hosts = localhost` is set.
4. If HTTPS is desired:
    a. Uncomment `https_redirect`.
    b. Uncomment and set `cert_email` to your email address.
5. Set `jqwidgets_zip_archive_path` to the full path for the jqWidgets distribution ZIP file.
6. If using the default UI, set `theme_zip_archive_path` to be the full path to the Canvas theme distribution ZIP file.

Edit Local Installation Application Configuration

Local Machine installation next requires configuration of the local application settings configuration. This file can be found under the deployment repository at the following path:

- `local/local_setting.py`

The following steps need to be performed in this file:

- Set django's secret key based on SECRET_KEY documentation.
- Setup your email backend, e.g., sendgrid.
- Setup project and ADMINS.

NOTE: Passwordless local SSH can simplify the local installation process. If you are not comfortable configuring SSH keys or using ssh-agent to run the installation process, please follow the steps below.

Organizations should carefully evaluate the security of passwordless local SSH and key management before relying on these steps for production usage.

Passwordless SSH Configuration:

1. Generate new SSH key for local usage:
    a. `ssh-keygen -f ~/.ssh/id_rsa`
2. Add server keys to known hosts:
    a. `ssh-keyscan -H localhost >> ~/.ssh/known_hosts.`
3. Allow user to ssh to itself:
    a. `cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys`
4. Confirm local ssh without prompt: ssh localhost

Execute Local Installation Process

Once all configuration steps are complete, execute the local installation process with the following command:

- `bash setup_local.sh | tee -a log.txt`

Installation logs are stored in log.txt and can be reviewed or shared with Support.

NOTE: This command is not meant to be run as root. Do not sudo prior to executing the command.

Remote Machine Installation

Edit Base Fabric Configuration

Remote Machine installation first requires configuration of the base fabricrc configuration. This file can be found under the deployment repository at the following path:

- `base/fabricrc`

The following steps need to be performed in this file:

1. Change database credentials. The default database credentials are insecure defaults that should be updated.
2. Set the web application superuser credentials, including username, password, email.

Edit Remote Installation Fabric Configuration

Remote Machine installation next requires configuration of the remote installation fabricrc configuration. This file can be found under the deployment repository at the following path:

- `remote/fabricrc`

The following steps need to be performed in this file:

1. Set user and sudo password.

2. Change `public_ip`, `dns_name` to match desired network adapter and address.
3. If HTTPS is desired:
    a. Uncomment `https_redirect`.
    b. Uncomment and set `cert_email` to your email address.
4. Set `jqwidgets_zip_archive_path` to the full path for the jqWidgets distribution ZIP file.
5. If using the default UI, set `theme_zip_archive_path` to be the full path to the Canvas theme distribution ZIP file.

Edit Remote Installation Application Configuration

Remote Machine installation next requires configuration of the remote application settings configuration. This file can be found under the deployment repository at the following path:

- `remote/local_setting.py`

The following steps need to be performed in this file:

- Set django's secret key based on SECRET_KEY documentation.
- Setup your email backend, e.g., sendgrid.
- Setup project and ADMINS.
- Add the remote machine PEM key in the `remote/` directory.

Execute Remote Installation Process

Once all configuration steps are complete, execute the remote installation process with the following command:

- `bash setup_remote.sh | tee -a log.txt`

Installation logs are stored in log.txt and can be reviewed or shared with Support.

NOTE: This command is not meant to be run as root. Do not sudo prior to executing the command.

Default ContraxSuite UI Setup

Most ContraxSuite implementations involve customized user interface and user experience development. However, to facilitate the testing and evaluation of ContraxSuite, a default UI can be configured for ContraxSuite out-of-the-box. This default UI uses two packages detailed in the list in the Software and Data Dependency documentation. Once the Third Party Theme and Software is licensed, the default theme can be installed using the automation described above in Local Machine Installation or Remote Machine Installation.

More Documentation

This Quick Start guide focuses on simple single-server deployments for development, research, or testing. As a result, it does not cover many of the architecture, installation, configuration, or testing topics that most organizations will encounter.

For more information, organizations should refer to these documents below:

- Installation and Configuration Guide
- Software and Data Dependencies
- System Administration and Monitoring Guide
- Architecture Diagram
- Data Model Diagram