



Установка Базы знаний InKnowledge



Оглавление

Введение	2
Подготовка к установке.....	2
Требования.....	2
Настройка межсетевых экранов	2
Установка ansible.....	5
Версии Python.....	5
Копирование инвентаря	5
Правка символических ссылок	6
Включение логирования	6
Настройка хостов и групп.....	6
Настройка подключения к хостам.....	8
Настройка групповых переменных	10
Запуск установки	17
Повторный запуск установки	17
Возможные проблемы и решения	17
Проверка корректности установки.....	18

Введение

Данный документ содержит информацию об установке Базы знаний InKnowledge.

Для установки InKnowledge будет использоваться Ansible - система управления конфигурациями, разрабатываемая компанией RedHat. Широко используется для приведения конфигураций систем к заданному состоянию, а также для развёртывания систем различной сложности.

При отсутствии возможности использования глобальных репозиториях производителей ПО (nginx, elastic, adopt и др.) необходимо использовать инструкции по скачиванию пакетов и установке в закрытый контур.

Подготовка к установке

Требования

- Рекомендуемые операционные системы на хостах - Linux семейства RedHat и Debian, протестировано на CentOS 7 и Ubuntu 20.
- Доступ в интернет на всех хостах
- Установку рекомендуется проводить на чистые системы без какого-либо предустановленного ПО
- Настроенный доступ к хостам по ssh с возможностью повышать привилегии через sudo
- Корректно настроенные межсетевые экраны на каждом из хостов
- Установленный ansible на управляющей станции версии не ниже 2.10
- Python3, установленный на управляющей станции
- Артефакты ansible для установки Базы знаний
- Заранее сгенерированные файлы для настройки HTTPS в nginx - приватные ключи и сертификаты в формате PEM

Настройка межсетевых экранов

Перед началом установки необходимо убедиться, что на всех хостах открыты нужные порты. В следующей таблице приведены необходимые открытые порты для входящего трафика на каждом компоненте системы InKnowledge:

Роль сервера	Порт/протокол	Источник	Описание
Liferay	8080/tcp	nginx	Доступ балансировщика к нодам портала
	9100/tcp	monitoring	Доступ prometheus к метрикам node-exporter
	9130/tcp	monitoring	Доступ prometheus к метрикам jmx-exporter

	multicast	liferay	Между нодами кластера необходимо обеспечить свободное хождение multicast-трафика для работы кластера
PostgreSQL	5432/tcp	все ноды liferay	Доступ портала к СУБД
	9100/tcp	monitoring	Доступ prometheus к метрикам node-exporter
	9187/tcp	monitoring	Доступ prometheus к метрикам postgres-exporter
ElasticSearch	9200/tcp	все ноды liferay	Доступ портала к сервису поиска
	9300/tcp	все ноды liferay	Доступ портала к сервису поиска
	9100/tcp	monitoring	Доступ prometheus к метрикам node-exporter
	9114/tcp	monitoring	Доступ prometheus к метрикам elasticsearch-exporter
NFS	2049/tcp	все ноды liferay	Доступ по протоколу nfs
	20048/tcp 20048/udp	все ноды liferay	Доступ по протоколу nfs, mountd
	111/tcp 111/udp	все ноды liferay	Доступ по протоколу nfs, rpc-bind
	9100/tcp	monitoring	Доступ prometheus к метрикам node-exporter
Nginx	80/tcp	все внешние пользователи	Доступ внешних пользователей к web-интерфейсу портала
	443/tcp	все внешние пользователи	Доступ внешних пользователей к web-интерфейсу портала
	9100/tcp	monitoring	Доступ prometheus к метрикам node-exporter
Monitoring	80/tcp	все внешние пользователи	Доступ внешних пользователей к web-интерфейсам мониторинга
	443/tcp	все внешние пользователи	Доступ внешних пользователей к web-интерфейсам мониторинга
	5020/tcp	все остальные компоненты InKnowledge	Доступ всех компонентов InKnowledge к интерфейсу приёма логов logstash

Кроме того, на каждом хосте должен быть доступен 22 порт для доступа с управляющей станции по ssh.

Установка ansible

Установить ansible можно через менеджер пакетов операционной системы (yum, apt), либо через pip:

```
[root@ansible ~]# python3 -m pip install ansible
```

Подробные инструкции по установке вы можете найти в официальной документации Ansible.

Рекомендуемая версия к установке не ниже 2.10.

Версии Python

Внимание! Python версии 2 **не поддерживается!** Перед запуском установки Базы знаний необходимо убедиться, что Ansible установлен нужной версии и использует Python версии 3:

```
[root@ansible ~]# ansible --version
ansible 2.10.2
  config file = None
  configured module search path = ['/root/.ansible/plugins/modules',
'/usr/share/ansible/plugins/modules']
  ansible python module location = /usr/local/lib/python3.6/site-packages/ansible
  executable location = /usr/local/bin/ansible
  python version = 3.6.8 (default, Apr 2 2020, 13:34:55) [GCC 4.8.5 20150623 (Red
Hat 4.8.5-39)]
```

Копирование инвентаря

В каталоге `inknowledge-deployment/ansible/inventory` находятся примеры инвентарей. Возьмём за основу пример `sample`:

```
[root@ansible ~]# cd inknowledge-deployment/ansible
[root@ansible ansible]# cp -rp inventory/sample inventory/example
```

Далее будем работать с содержимым каталога `example`:

- `hosts.ini` - информация и хостах и их распределении по группам
- `group_vars/all.yml` - глобальные групповые переменные для всех групп хостов
- `group_vars/<groupname>.yml` - переменные для конкретных групп хостов

Правка символических ссылок

В проекте присутствуют символные ссылки, которые необходимы для корректной установки. При клонировании репозитория, либо при распаковке из архива они сохраняются, а при копировании через `scp` или `sftp` их с вероятностью 100% не окажется. Перед запуском установки следует проверить их наличие:

```
[root@ansible ~]# cd inknowledge-deployment/ansible/playbooks/
[root@ansible playbooks]# ls -al roles
lrwxrwxrwx. 1 root root 9 Jun4 09:20 roles -> ../roles/
[root@ansible ~]# cd inknowledge-deployment/ansible/playbooks/monitoring
[root@ansible monitoring]# ls -al roles
lrwxrwxrwx. 1 root root 9 Jun4 09:20 roles -> ../../roles/
[root@ansible ~]# cd inknowledge-deployment/ansible/roles/liferay/
[root@ansible liferay]# ls -al files
lrwxrwxrwx. 1 root root 17 Jun4 09:21 bash -> ../../../../bash/
```

Если их нет, то их надо создать:

```
[root@ansible ~]# cd inknowledge-deployment/ansible/playbooks/
[root@ansible playbooks]# ln -s ../roles/ roles
[root@ansible ~]# cd inknowledge-deployment/ansible/playbooks/monitoring/
[root@ansible monitoring]# ln -s ../../roles/ roles
[root@ansible ~]# cd inknowledge-deployment/ansible/playbooks/roles/liferay/files
[root@ansible files]# ln -s ../../../../bash/ bash
```

Включение логирования

При необходимости можно включить функцию логирования установки. Для этого нужно проверить, что в файле `ansible.cfg` присутствует нужная настройка `log_path`:

```
[root@ansible ~]# cd inknowledge-deployment/ansible/
[root@ansible ansible]# cat ansible.cfg
[defaults]
host_key_checking = False
allow_world_readable_tmpfiles = True
log_path = playbook.log
```

После включения этой настройки результаты работы `ansible` будут находиться в указанном файле в каталоге `inknowledge-deployment/ansible`.

Настройка хостов и групп

На данном этапе необходимо определиться с планируемой установкой InKnowledge - какие компоненты будут устанавливаться. Эти настройки указываются в файле инвентаря `hosts.ini`. Возможны следующие группы:

- **elasticsearch** - на хосты данной группы будет установлен Elasticsearch
- **postgresql** - на хосты данной группы будет установлен PostgreSQL
- **nfs** - на хосты данной группы будет установлен сервер NFS
- **liferay** - на хосты данной группы будет установлен кластер InKnowledge
- **nginx** - на хосты данной группы будет установлен балансировщик nginx
- **monitoring** - на хосты данной группы будут установлены средства мониторинга Prometheus, Grafana.

Пример файла инвентаря **hosts.ini**:

```
[inknowledge:children]
nginx
liferay
elasticsearch
postgresql
nfs
monitoring

[nginx]
nginx-01 ansible_host=172.17.251.70

[liferay]
liferay-01 ansible_host=172.17.251.71
liferay-02 ansible_host=172.17.251.72

[elasticsearch]
elastic-01 ansible_host=172.17.251.73

[postgresql]
postgres-01 ansible_host=172.17.251.74

[nfs]
storage-01 ansible_host=172.17.251.75

[monitoring]
monitoring-01 ansible_host=172.17.251.76
```

В каждой группе, кроме Liferay, может находиться не более одного хоста. Это связано с тем, что в настоящий момент только Liferay может устанавливаться в кластерной конфигурации. Добавление более чем одного хоста в группу вызовет ошибку при запуске установки.

Хост, находящийся в группе monitoring, не может одновременно находиться в нескольких группах. Добавление хоста, входящего в группу monitoring, в любую другую группу вызовет ошибку при запуске установки.

Любой из хостов может находиться одновременно в нескольких группах, исключая группу monitoring. В таком случае, на него будет установлено несколько компонентов.

Вместе с тем, рекомендуется под каждый компонент выделять отдельную машину.

Обратите внимание на то, как указывается хост. Если вы указываете непосредственно ip-адрес, либо доменное имя, которое может разрешаться в ip-адрес, то можно указать хост вот так:

```
[liferay]
liferay-01
172.17.251.72
```

Если имя хоста не резолвится, то необходимо дополнительно указать его адрес:

```
[liferay]
liferay-01 ansible_host=172.17.251.71
liferay-02 ansible_host=172.17.251.72
```

Настройка подключения к хостам

Ansible подключается к хостам, как правило, с использованием протокола ssh. Настройки подключения следует указывать в файле `group_vars/all.yml`.

Необходимо указать тип подключения:

```
ansible_connection: ssh
```

Далее необходимо указать пользователя, под которым ansible будет подключаться к хостам:

```
ansible_user: user
```

Следует иметь в виду, что данный пользователь должен иметь права повышать привилегии при помощи sudo. Если для повышения привилегий необходим пароль, его следует указать в соответствующей опции:

```
ansible_become_password: password
```

Если повышение привилегий необходимо производить от имени другого пользователя, то это также нужно указать в соответствующей опции:

```
ansible_become_user: user
```

Рекомендуется настраивать подключение с использованием ssh-ключей. Для этого публичный ключ пользователя, под которым вы работаете на управляющей станции, должен быть предварительно добавлен в `authorized_keys` на всех хостах. Если для подключения вы используете закрытый ключ, отличный от `~/.ssh/id_rsa`, то в таком случае следует указать путь к закрытому ключу в соответствующем параметре:

```
ansible_ssh_private_key_file: /path/to/private/key
```

Если вы предпочитаете подключаться к хостам с использованием паролей, то предварительно необходимо установить утилиту `sshpass` на управляющую станцию:

```
[root@ansible ~]# yum install sshpass
```

После этого можно указать пароль в файле `group_vars/all.yml`:

```
ansible_ssh_password: password
```

Указывать пароль в конфигурационном файле не рекомендуется по соображениям безопасности. Вместо этого рекомендуется использовать утилиту `ansible-vault`. Для этого необходимо создать зашифрованный файл переменных:

```
[root@ansible ~]# cd inknowledge-deployment/ansible/inventory/example  
[root@snabile example]# ansible-vault create vars.yml
```

Далее необходимо дважды ввести новый пароль для файла и в открывшемся редакторе ввести и сохранить все нужные переменные с нужными значениями.

Кроме того, пароль можно вводить интерактивно при запуске утилиты `ansible-playbook`. Обратитесь к официальной документации Ansible для получения более подробной информации.

Настройка групповых переменных

В целом подготовка к установке сводится к распределению хостов по группам в файле инвентаря и дальнейшей настройке групповых переменных. Далее будут расписаны настройки каждой из групп.

group_vars/all.yml

Параметр	Описание	Значение по умолчанию
Группа настроек подключения к хостам	См. предыдущий раздел	
openjdk_version	Версия AdoptOpenJDK, которая будет устанавливаться на хосты. Должна быть 11.0.5+10 в текущей версии InKnowledge.	11.0.5+10
liferay_admin_instance_domain_name	Доменное имя админского инстанса. Это доменное имя также будет использоваться при создании vhost в случае установки балансировщика nginx	localhost
max_user_count	Максимальное количество пользователей, которые будут работать с системой	
closed_env	Установите этот параметр в true, если сервера, на которые производится установка, отсутствует выход в интернет. При установке данного параметра вам понадобится в файлах переменных указать пути к скачиванию необходимых для установки артефактов	false

group_vars/elasticsearch.yml

Параметр	Описание	Значение по умолчанию
es_version	Версия ElasticSearch, которая будет устанавливаться. Должна быть 7.9.3	7.9.3
es_plugins	Список официальных плагинов, необходимых для установки. Должен включать в себя следующие плагины: <ul style="list-style-type: none"> analysis-icu analysis-kuromoji 	

	<ul style="list-style-type: none"> • analysis-smartcn • analysis-stempel 	
es_custom_plugin_url	Ссылка на архив плагина analysis-morphology версии 7.9.3 . Следует использовать этот параметр, если вы хотите скачать архив из интернета. При использовании данного параметра вам могут понадобиться следующие два параметра	
es_custom_plugin_username	Логин для basic-аутентификации при скачивании плагина analysis-morphology версии 7.9.3 из интернета	
es_custom_plugin_password	Пароль для basic-аутентификации при скачивании плагина analysis-morphology версии 7.9.3 из интернета	
es_custom_plugin_archive	Путь к архиву плагина analysis-morphology версии 7.9.3 на управляющей станции. Используйте данный параметр, если у вас есть уже скачанный плагин. Обратите внимание, что данный параметр имеет приоритет над параметром es_custom_plugin_url - если указаны оба, то использоваться будет es_custom_plugin_archive .	
es_cluster_name	Имя устанавливаемого кластера ElasticSearch. Не рекомендуется изменять данный параметр	LiferayElasticsearchCluster
es_http_port	HTTP порт ElasticSearch	9200
es_transport_port	Транспортный порт ElasticSearch	9300
jvm_heap_factor	Коэффициент, определяющий процент от общего количества оперативной памяти, который будет выделен ElasticSearch в качестве heap size в случае, если ElasticSearch устанавливается на выделенный сервер. Не	0.7

	рекомендуется устанавливать выше 0.7	
--	--------------------------------------	--

group_vars/liferay.yml

Параметр	Описание	Значение по умолчанию
<code>libreoffice_setup</code>	Устанавливать LibreOffice или нет	true
<code>libreoffice_version</code>	Версия LibreOffice	6.3.5.2
<code>inknowledge_url</code>	Ссылка на архив InKnowledge. Следует использовать этот параметр, если вы хотите скачать архив из интернета. При использовании данного параметра вам могут понадобиться следующие два параметра	
<code>inknowledge_username</code>	Логин для basic-аутентификации при скачивании InKnowledge из интернета	
<code>inknowledge_password</code>	Пароль для basic-аутентификации при скачивании InKnowledge из интернета	
<code>inknowledge_archive</code>	Путь к архиву InKnowledge на управляющей станции. Данный параметр следует использовать, если у вас есть уже скачанный архив. Обратите внимание, что данный параметр имеет приоритет над параметром <code>inknowledge_url</code> - если указаны оба, то использоваться будет <code>inknowledge_archive</code> .	
<code>es_remote</code>	Использовать внешний Elasticsearch или нет. Если в инвентаре вы указали группу <code>elasticsearch</code> , то данный параметр автоматически будет установлен в true. Если группу <code>elasticsearch</code> вы не указали, но у вас есть установленный сервер Elasticsearch, который вы хотите использовать, то данный параметр необходимо выставить в true, а также указать следующие три параметра. Если вы хотите использовать встроенный Elasticsearch, то данный параметр необходимо выставить в false. Не	

	рекомендуется использовать встроенный ElasticSearch в продуктивных инсталляциях.	
es_host	Если вы установили es_remote в true и не указали в инвентаре группу elasticsearch, то в данном параметре необходимо указать адрес внешнего сервера ElasticSearch. Если вы указали группу elasticsearch, то данный параметр будет установлен автоматически.	
es_port	Если вы установили es_remote в true и не указали в инвентаре группу elasticsearch, то в данном параметре необходимо указать транспортный порт внешнего сервера ElasticSearch. Если вы указали группу elasticsearch, то данный параметр будет установлен автоматически.	
es_major_version	Если вы установили es_remote в true и не указали в инвентаре группу elasticsearch, то в данном параметре необходимо указать мажорную версию внешнего сервера ElasticSearch. Важно указать этот параметр, потому что от него зависит путь к конфигу коннектора в Liferay. Если вы указали группу elasticsearch, то данный параметр будет установлен автоматически.	
liferay_kerberos_use	Настраивать аутентификацию Kerberos или нет	false
liferay_krb_principal	Принципал Kerberos	
liferay_krb5_path	Путь на управляющей станции к файлу krb5.conf	
liferay_keytab_path	Путь на управляющей станции к файлу keytab	
liferay_tomcat_port	Базовый порт, на котором будет работать tomcat на всех экземплярах	8080
liferay_wizard	Группа параметров для файла portal-setup-wizard.properties	

mail	Группа параметров для подключения к почтовому серверу, сохраняются в файле portal-ext.properties	
liferay_db	Группа параметров для подключения к внешнему серверу PostgreSQL. Если вы указали группу postgresql в инвентаре, то эти параметры будут установлены автоматически. Если у вас есть установленный сервер PostgreSQL, который вы хотите использовать, то вам следует указать эти параметры.	
nfs_server	Адрес внешнего сервера NFS. Использование NFS обязательно в кластерной конфигурации. Если вы не указали группу nfs в инвентаре, вы должны указать адрес внешнего сервера здесь. Если вы устанавливаете кластер Liferay, не включили в инвентаре группу nfs и не указали этот параметр, то при старте установки возникнет ошибка.	
nfs_server_share_dir	Путь к сетевой папке на сервере NFS. Данный параметр используется только в случае использования внешнего сервера	
nfs_version	Версия протокола NFS. Данный параметр используется только в случае использования внешнего сервера	
jvm_heap_factor	Коэффициент, определяющий процент от общего количества оперативной памяти, который будет выделен Liferay в качестве heap size в случае, если Liferay устанавливается на выделенный сервер. Не рекомендуется устанавливать выше 0.7	0.7

group_vars/monitoring.yml

Параметр	Описание	Значение по умолчанию
nginx_setup	Устанавливать реверсивный прокси nginx перед интерфейсами мониторинга или нет.	true

nginx_domain_name	Доменное имя, по которому будут доступны интерфейсы мониторинга в случае установки nginx. Данное имя должно разрешаться в ip-адрес хоста, на который устанавливаются элементы мониторинга	
https_use	Использовать https или нет в случае установки nginx	true
https_cert_path	Путь к файлу сертификата в формате PEM на управляющей станции	
https_key_path	Путь к файлу приватного ключа в формате PEM на управляющей станции	
nginx_basic_auth	Использовать базовую авторизацию или нет для доступа в интерфейсы Prometheus и Alertmanager	true
nginx_basic_auth_user	Имя пользователя для базовой авторизации	
nginx_basic_auth_pass	Пароль пользователя для базовой авторизации	
alertmanager_global	Секция global конфигурационного файла alertmanager.yml	
alertmanager_route	Секция route конфигурационного файла alertmanager.yml	
alertmanager_receivers	Секция receivers конфигурационного файла alertmanager.yml	

group_vars/nfs.yml

Параметр	Описание	Значение по умолчанию
<code>nfs_share_user_uid</code>	UID пользователя, который будет владельцем сетевой папки. Пользователь с таким же UID будет создан на хостах liferay	
<code>nfs_share_dir</code>	Каталог на хосте, который будет расшарен	

group_vars/nginx.yml

Параметр	Описание	Значение по умолчанию
<code>https_use</code>	Использовать https или нет	true
<code>https_cert_path</code>	Путь к файлу сертификата в формате PEM на управляющей станции	
<code>https_key_path</code>	Путь к файлу приватного ключа в формате PEM на управляющей станции	

group_vars/postgresql.yml

Параметр	Описание	Значение по умолчанию
<code>postgre_version</code>	Версия сервера postgresql	11.10
<code>postgresql_db_name</code>	Имя базы данных. Рекомендуется не изменять.	inknowledge
<code>postgresql_user_name</code>	Имя пользователя базы данных. Рекомендуется не изменять.	liferay
<code>liferay_connection_static</code>	Количество дополнительных разрешённых подключений к базе данных	50

Запуск установки

После настройки инвентаря необходимо перейти в каталог ansible и запустить установку. Если вы используете ansible-vault, следует добавить использование дополнительных переменных и запрос пароля:

```
[root@ansible ~]# cd inknowledge-deployment/ansible
[root@ansible ~]# ansible-playbook -i inventory/example/hosts.ini -e
"@inventory/example/vars.yml" --ask-vault-pass playbooks/deploy_liferay.yml
```

В процессе установки будет появляться информация о шагах установки. После успешного завершения необходимо проверить корректность установки.

В случае возникновения проблем установка прервётся.

В среднем при хорошей производительности сети установка длится от 10 до 20 минут.

Повторный запуск установки

В самом начале установки запускаются плейбуки проверки корректности инвентаря и групповых переменных. В случае явных несоответствий требованиям установка прервётся с указанием текста ошибки и рекомендациями по устранению.

В большинстве случаев прервавшуюся с ошибкой установку можно запустить заново, но стоит иметь в виду следующие моменты:

- из-за особенностей дистрибуции LibreOffice и InKnowledge повторный запуск вызовет их повторное скачивание и установку
- если в архиве InKnowledge есть бэкап, то он будет развёрнут заново с заменой существующих данных

Перед повторным запуском необходимо проверить инвентарь и файлы групповых переменных, чтобы убедиться, что все нужные переменные указаны верно.

В случае если повторная установка также завершается с ошибкой, необходимо провести установку ещё раз на чистые системы.

Возможные проблемы и решения

Какие могут быть проблемы во время установки и как их решать:

- Ошибки, связанные с некорректной настройкой инвентаря. Ansible первым делом запускает плейбуки проверки инвентаря. Если будут обнаружены логические ошибки, либо ошибки в названии указанных файлов, то установка прервётся с указанием текста ошибки и рекомендациями по устранению. После устранения замечаний установку необходимо запустить заново.
- Ошибки, связанные с некорректными версиями софта. Иногда разработчики ПО удаляют некоторые версии софта из своих репозиториях. Необходимо убедиться, что все указанные версии ПО доступны в соответствующих репозиториях.

- Ошибки, связанные с некорректными ссылками/логинами/паролями. Нужно проверить, что данные указаны верно, после чего запустить установку заново.
- Ошибки, связанные с недоступностью ресурсов. Например, репозиторий не ответил вовремя, либо что-то пошло не так во время ssl handshake. В таком случае установку нужно запустить заново.

Проверка корректности установки

Проверьте, что при попытке зайти браузером на адрес, указанный в переменной `liferay_admin_instance_domain_name` в файле `group_vars/all.yml`, вы видите открывшийся интерфейс портала InKnowledge. Обратите внимание, что portalу требуется некоторое время (до 5 минут) для полного запуска.

Следующие сервисы должны быть в статусе running:

- группа хостов liferay
 - liferay.service
 - libreoffice.service
 - node-exporter.service
- группа хостов postgresql
 - postgresql-11.service (postgresql.service в случае установки на Ubuntu)
 - node-exporter.service
 - postgres-exporter.service
- группа хостов elasticsearch
 - elasticsearch.service
 - node-exporter.service
 - elasticsearch-exporter.service
- группа хостов nfs
 - node-exporter.service
- группа хостов nginx
 - nginx.service
 - node-exporter.service
- группа хостов monitoring
 - node-exporter.service
 - blackbox-exporter.service
 - prometheus.service
 - alertmanager.service
 - grafana-server.service
 - nginx.service

Если установка включала в себя мониторинг, то по адресу, указанному в переменной `nginx_domain_name` файла `group_vars/monitoring.yml`, должны открываться соответствующие интерфейсы:

- `http://<nginx_domain_name>/prometheus`
- `http://<nginx_domain_name>/alertmanager`
- `http://<nginx_domain_name>/grafana`

Для доступа в интерфейс Grafana по умолчанию используются учётные данные

admin:admin. При первом входе будет необходимо их изменить.

Для доступа в интерфейсы Prometheus и Alertmanager используются учётные данные для базовой авторизации, указанные в файле `group_vars/monitoring.yml`.

В Grafana на дашборде Overview должно быть видно, что все сервисы запущены и готовы принимать запросы.