

INFOWATCH

InfoWatch Data Discovery. Руководство по установке

03/09/2026

© АО “ИнфоВотч”

Тел./Факс +7 (495) 229-00-22

http://www.infowatch.ru

СОДЕРЖАНИЕ

1 Аудитория 5

2 Комплект документов 6

3 Техническая поддержка пользователей 7

4 Типы установки и обновления, совместимость с другими продуктами, аппаратные и программные требования 8

4.1 Типы установки и совместимость с другими продуктами InfoWatch 8

4.2 Типы обновления 10

4.3 Экспорт данных в Traffic Monitor 10

4.4 Аппаратные и программные требования 11

4.4.1 Аппаратные требования 11

4.4.2 Программные требования 12

5 Подготовка сервера 14

6 Настройка сетевых правил доступа 19

7 Установка Data Discovery 22

7.1 Варианты запуска установки Data Discovery 23

7.2 Модуль контентного анализа 23

7.3 Установка Data Discovery в интерактивном режиме 24

7.4 Установка Data Discovery с указанием параметров в командной строке 29

7.5 Установка Модуля контентного анализа в интерактивном режиме 34

7.6 Установка Модуля контентного анализа с указанием параметров в командной строке 38

8 Установка лицензии и плагина Traffic Monitor 44

8.1 Экспорт данных в Traffic Monitor 44

8.2 Импорт данных из Traffic Monitor 44

9 Устранение неполадок 46

10 Обновление Data Discovery 47

10.1 Настройки и данные Системы 47

10.2 Варианты обновления Data Discovery 47
10.3 Модуль контентного анализа 48
10.4 Особенности обновления кластера Центра расследований 48
10.5 Прямое обновление Data Discovery 49
10.6 Обновление Data Discovery с помощью сброса установки (reset) 53
10.7 Переустановка Data Discovery в другой кластер 57
10.8 Прямое обновление Модуля контентного анализа 59
11 Удаление Data Discovery 61

Настоящее руководство содержит сведения по установке, обновлению и удалению InfoWatch Data Discovery (далее Система или Data Discovery) – программного обеспечения, предназначенного для поиска конфиденциальной информации на общих сетевых ресурсах, рабочих станциях, серверах и в хранилищах документов.

1 Аудитория

Данное руководство предназначено для инженеров внедрения и офицеров безопасности, которые будут работать с Системой и заниматься ее администрированием. Руководство рассчитано на пользователей, знакомых с основами работы в среде операционных систем Linux и MS Windows.

2 Комплект документов

В документацию входят:

- «InfoWatch Data Discovery. Руководство по установке». Документ содержит описание процесса установки, обновления и удаления Системы.
- «InfoWatch Data Discovery. Руководство администратора». Документ содержит информацию по настройке Системы.
- «InfoWatch Data Discovery. Руководство пользователя». Содержит описание работы с задачами сканирования.

Сопутствующая документация по системе InfoWatch Traffic Monitor (далее Traffic Monitor) включает в себя:

- «InfoWatch Traffic Monitor. Руководство по установке». Содержит описание установки, настройки, обновления и удаления системы InfoWatch Traffic Monitor.
- «InfoWatch Traffic Monitor. Руководство администратора». Содержит информацию по администрированию системы InfoWatch Traffic Monitor (база данных, серверная часть).
- «InfoWatch Traffic Monitor. Руководство пользователя». Содержит описание работы с InfoWatch Traffic Monitor (настройка конфигурации, экспорт/импорт данных, подготовка политик для обработки объектов).
- «InfoWatch Traffic Monitor. Справочник по конфигурационным файлам». Содержит пояснения к часто используемым конфигурационным файлам.

Кроме основных руководств для пользователей и администраторов Системы предусмотрена База знаний lk.infowatch.ru, которая включает материалы по устранению неисправностей и ответы на вопросы. В случае отсутствия доступа к Базе знаний, просьба обращаться в службу технической поддержки компании InfoWatch по адресу support@infowatch.com.

3 Техническая поддержка пользователей

При возникновении проблем и вопросов, связанных с работой Системы, вы можете обратиться в службу технической поддержки:

- если вы приобрели продукт у партнера компании InfoWatch, то обратитесь в службу технической поддержки партнера;
- если продукт приобретен у компании InfoWatch напрямую, то обратитесь в службу технической поддержки компании InfoWatch по адресу support@infowatch.com.

Вы также можете посетить раздел технической поддержки на нашем сайте: https://www.infowatch.ru/services/support.

4 Типы установки и обновления, совместимость с другими продуктами, аппаратные и программные требования

4.1 Типы установки и совместимость с другими продуктами InfoWatch

InfoWatch Data Discovery является частью Центра расследований InfoWatch.

Центр расследований InfoWatch – это единый интерфейс для работы и эффективного взаимодействия следующих продуктов InfoWatch:

- Activity Monitor 3.4
- Data Access Tracker 2.4
- Data Discovery 2.2
- Prediction 3.4
- Vision 3.6

Каждый продукт может быть установлен и работать отдельно, однако использование нескольких продуктов одновременно существенно повышает возможности Офицера безопасности.

Основой для установки каждого продукта Центра расследований выступает Платформа. Инсталлятор каждого продукта содержит и Платформу, и продуктовые компоненты. Обе части обязательны к установке при развертывании продуктов.
И продуктовые компоненты, и Платформа поддерживают установку в двух режимах:

- `central`
- `office`

Режимы установки Платформы и продукта вы указываете при запуске инсталлятора. Если оба режима не указаны пользователем, инсталлятор по умолчанию использует значение `central` для обоих компонентов.

Data Discovery 2.2 реализован на базе Платформы версии 2.4.1.

Центр расследований поддерживает **распределенную установку**: центральная (`central`) и офисные (`office`) ноды Платформы объединены в кластер.
Вы можете установить компоненты продуктов и на центральную, и на офисные ноды Платформы. Это позволяет создать кластер Платформы с несколькими продуктами с учетом инфраструктуры, требований к безопасности и доступных мощностей компании.

> **Важно!**
>
> При распределенной установке примите во внимание следующие ключевые аспекты:
>
> - Продукты в кластере должны быть установлены на одной версии Платформы – в номерах версий должны совпадать первые два числа. Версия центральной ноды Платформы должна быть самой новой в кластере.
> - Если вы устанавливаете в кластер продукт на новой версии Платформы, то после этого обновите остальные продукты.

- В одном кластере можно развернуть только одну центральную ноду Платформы и каждый продукт в режиме `central` только в единственном экземпляре.
- Для работы Data Discovery в режиме `office` нужно, чтобы в кластере был экземпляр в режиме `central`.
- Если в кластере уже установлен продукт в режиме `central`, еще одна установка этого продукта в режиме `central` в том же кластере будет прервана ошибкой.
- Data Access Tracker поддерживает установку на любую ноду Платформы, но только в режиме `central`.

Продукты в одном кластере имеют:

- доступ ко всей функциональности в одной консоли управления;
- общие настройки;
- общую ролевую модель;
- общий доступ к данным независимо от того, на каком из них выполнена интеграция с источником, а также доступ к обработанным данным.

Важно!

Для подключения к веб-интерфейсу Центра расследований **всегда** используйте адрес или имя сервера, на котором установлена **Платформа в режиме `central`**.

Если в распределенной установке продукт в режиме `central` установлен на офисной ноде Платформы, для работы все равно используйте центральную ноду Платформы (`central`).

Продукты Центра расследований InfoWatch поддерживают **совместную установку** на одном сервере.

Data Discovery 2.2 реализован на базе Платформы 2.4.1, но поддерживает совместную установку и взаимодействие с продуктами и модулями на базе Платформы 2.5.
На Платформе **2.4.1** вы можете установить **Data Discovery 2.2 совместно** с:

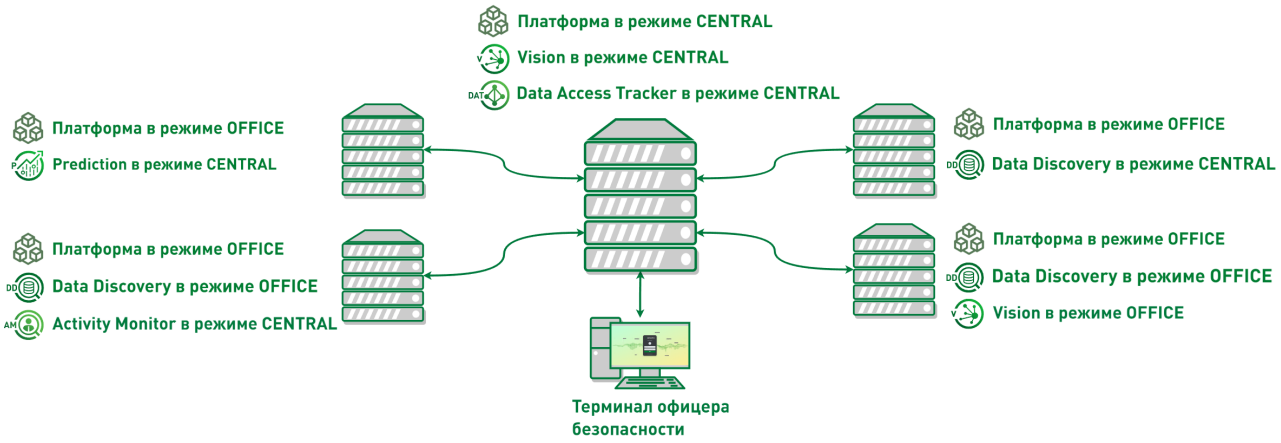
- **Activity Monitor 3.4**
- **Data Access Tracker 2.4**
- **Prediction 3.4**
- **Vision 3.6**

На Платформе **2.5** вы можете установить **Data Discovery 2.2 совместно** с:

- **Activity Monitor 3.5**
- **Data Access Tracker 2.4** (реализован на базе Платформы 2.4.1)
- **Prediction 3.5**
- **Vision 3.7**

Для расчета аппаратных требований при совместной установке учитывайте требования каждого продукта. Для точного расчета требований под конкретный проект обратитесь в службу технической поддержки InfoWatch.

Ниже представлен **пример** распределенного кластера Центра расследований с совместными установками нескольких продуктов.



Подробнее о процессе установки см. *Установка Data Discovery* .

4.2 Типы обновления

Data Discovery поддерживает:

- прямое обновление;
- обновление с помощью сброса и переустановки на прошлые данные.

Модуль контентного анализа поддерживает прямое обновление.

Подробнее о процессе обновления см. *Обновление Data Discovery* .

4.3 Экспорт данных в Traffic Monitor

Data Discovery может экспортировать данные в InfoWatch Traffic Monitor. Data Discovery взаимодействует с Traffic Monitor 7.1 и выше.

Для настройки см. статьи:

- *Установка лицензии и плагина Traffic Monitor*
- *InfoWatch Data Discovery. Руководство администратора. Зоны и серверы*

> **Примечание:**
>
> В Traffic Monitor 7.14 обновили предустановленные сертификаты и ключ. Учитывайте это, если Data Discovery будет взаимодействовать с Traffic Monitor 7.14 или более новыми версиями.

4.4 Аппаратные и программные требования

4.4.1 Аппаратные требования

Аппаратные требования Data Discovery зависят от сетевого окружения и предполагаемой нагрузки. Ниже приведена примерная конфигурация сервера Data Discovery для обработки 50 хостов в одной задаче сканирования:

Компонент	Рекомендуемые требования
ЦПУ	8 ядер, 16 потоков Intel Xeon или аналог, с поддержкой инструкций SSE 4.2
ОЗУ	16 ГБ
Диск	• Дисковый массив: RAID 5 • Диски: 10K HDD • Объем дискового массива: 400 ГБ **Примечание**: требуемый объем дискового массива зависит от количества и объема сканируемых файлов. **Важно!** Если использовать функциональность InfoWatch Data Analysis Service и/или Модуля контентного анализа, Data Discovery будет занимать дополнительный объем дискового пространства для хранения извлеченных текстовых контентов. Рекомендуем увеличить объем доступного дискового пространства на 40%, но при этом учитывать, что требуемое дополнительное пространство зависит от объема текстовых данных и особенностей документооборота компании.
Сеть	1 Гб/сек LAN

Примечание:

При расчете аппаратных требований для совместной установки рекомендуем суммировать требования Data Discovery и других продуктов для каждого компонента.

Например: если для продукта нужно 24 ГБ оперативной памяти, то для совместного функционирования Data Discovery и этого продукта нужно 40 ГБ оперативной памяти.

Вы можете установить модуль контентного анализа, используемый для обнаружения объектов защиты в текстовых данных, на одном сервере с Data Discovery или отдельно.
Если вы будете использовать **Модуль контентного анализа на одном сервере с Data Discovery** :

- Для сканирования хранилищ одной задачей с одним сканером добавьте к аппаратным требованиям Data Discovery **не менее**:
  - 2-х ядер ЦПУ;
  - 1 ГБ ОЗУ;
  - 50 ГБ дискового пространства.

- Для сканирования файловых хранилищ с большим объемом данных одновременно двумя или более задачами сканирования, с одним сканером в каждой, добавьте **не менее**:
  - 4-х ядер ЦПУ;
  - 2 ГБ ОЗУ;
  - 50 ГБ дискового пространства.

При дополнительном увеличении предполагаемой нагрузки рекомендуем установить дополнительные Модули контентного анализа на отдельные серверы. Нагрузка будет распределена между ними равномерно.
Если вы будете использовать **Модуль контентного анализа на отдельном сервере**, обеспечьте **не менее**:

- 4-х ядер ЦПУ;
- 12 ГБ ОЗУ;
- 100 ГБ дискового пространства.

Если вы будете использовать **распознавание текста на изображениях (OCR), каждая запущенная задача** с этой технологией потребует добавить к аппаратным требованиям 8 ядер ЦПУ и 4 ГБ ОЗУ.

Важно!

Не рекомендуем устанавливать Data Discovery и Traffic Monitor на один сервер. Увеличение нагрузки от Traffic Monitor может привести к некорректной работе Data Discovery.

4.4.2 Программные требования

Для установки и работы Системы должно быть использовано следующее программное обеспечение:

Тип ПО	Варианты
Операционная система	• Astra Linux Special Edition, версии: • 1.7 в редакции "Воронеж" • 1.7.5 в редакциях "Смоленск" и "Орел" • 1.8.0 в редакции "Смоленск" • Red Hat Enterprise Linux, версии: • 7.9 • 8.9 • 8.10 • Oracle Linux, версии: • 7.9 • 8.4 • РЕД ОС, версии: • 7.3.3 • 8 • ОС Альт, версии: • Сервер 10.2 • СП 10

Тип ПО	Варианты
	Важно: модули **Netfilter** и **NAT** должны быть загружены в ядро ОС.
Браузер	• Google Chrome • Яндекс Браузер **Важно**: в браузере должна быть реализована аппаратная поддержка WebGL/WebGL2.

Data Discovery использует PostgreSQL 16.4. СУБД поставляется вместе с продуктом и не требует отдельной установки.

Модуль контентного анализа поддерживает установку на те же ОС, что и Data Discovery.

Для корректной работы версии Модуля контентного анализа и Traffic Monitor, конфигурация которого будет использована, должны совпадать. Data Discovery **2.2** использует Модуль контентного анализа **7.12.1**, **7.13** или **7.14**.

Примечание:

Модуль контентного анализа **7.14** реализован на базе Платформы 2.5.1, а Data Discovery **2.2** – на Платформе 2.4.1. Вы можете установить Модуль контентного анализа **7.14** и Data Discovery **2.2** на одну ноду, в этом случае оба компонента будут работать на Платформе 2.5.1. Если вы хотите использовать их в одном кластере на отдельных нодах, обратитесь к специалистам InfoWatch, чтобы обеспечить одну версию Платформы на всех нодах кластера.

Модуль контентного анализа **7.12.1** и **7.13** работает только на Платформе 2.4.1.

Для поиска похожих файлов Data Discovery **2.2** использует Data Analysis Service **1.5.0** или **1.6.0**.

5 Подготовка сервера

Перед установкой Data Discovery выполните шаги по подготовке сервера.

Шаги по подготовке сервера:

1. Программа установки написана на языке Python версии 2.7. Установите интерпретатор языка Python версии 2.7. Пакет с интерпретатором не входит в состав дистрибутива продукта.
2. Настройте правила POD сети или отключите межсетевой экран (подробнее см. статьи Базы знаний "Конфликты при взаимодействии firewalld и Kubernetes ", "Полное отключение межсетевого экрана").
3. Для корректной работы продукта настройте сетевые правила доступа (подробнее см. "Настройка сетевых правил доступа").
4. Data Discovery использует Kubernetes. Для корректной работы Kubernetes установите пакеты:
   - `socat`
   - `conntrack` или `conntrack-tools`. Пакет необходимо установить, даже если `libnetfilter_conntrack` уже установлен.
5. Если вы используете антивирус, настройте его для совместной работы с Системой (см. статью Базы знаний *Настройка антивируса для продуктов Платформы*).
6. **Если в операционной системе используется SELinux:**

   Для корректной работы используемого `containerd` обновите политики SELinux или отключите систему контроля доступа:

   - Чтобы **обновить** политики SELinux, скачайте и установите пакет `container-selinux`.
     Вы можете сделать это, выполнив команду:

     ```
     dnf install container-selinux
     ```

   - Чтобы **отключить** SELinux:
     a. Выполните команды от имени суперпользователя `root`:

     ```
     setenforce 0
     sed -i 's/SELINUX=enforcing/SELINUX=disabled/g' /etc/selinux/config
     ```

     b. Перезагрузите сервер с помощью команды:

     ```
     reboot
     ```

     c. Проверьте статус системы SELinux с помощью команды:

     ```
     sestatus
     ```

7. **Если вы устанавливаете продукт на сервер под управлением ОС Альт:**

   Установите пакеты:

- `python-modules-json`
- `python-modules-distutils`
- `python-modules-sqlite3`

8. **Если вы устанавливаете продукт в ОС Альт СП 10, выполните следующие действия:**

   a. В файле `/etc/apt/sources.list.d/altsp.list` раскомментируйте три последних строки, указывающие репозитории.

   Пример содержимого altsp.list с раскомментированными строками

   ```
   # update.altsp.su (IVK, Moscow)

   # ALT Certified 10
   #rpm [cert8] ftp://update.altsp.su/pub/distributions/ALTLinux c10f2/branch/x86_64
   classic gostcrypto
   #rpm [cert8] ftp://update.altsp.su/pub/distributions/ALTLinux c10f2/branch/x86_64-
   i586 classic
   #rpm [cert8] ftp://update.altsp.su/pub/distributions/ALTLinux c10f2/branch/noarch
   classic

   rpm [cert8] http://update.altsp.su/pub/distributions/ALTLinux c10f2/branch/x86_64
   classic gostcrypto
   rpm [cert8] http://update.altsp.su/pub/distributions/ALTLinux c10f2/branch/x86_64-
   i586 classic
   rpm [cert8] http://update.altsp.su/pub/distributions/ALTLinux c10f2/branch/noarch
   classic
   ```

   b. Обновите пакеты с помощью команды:

   ```
   apt-get update
   ```

   c. Установите требуемые пакеты с помощью команд:

   ```
   apt-get install python-modules-sqlite3
   apt-get install conntrack-tools
   ```

   d. Установите менеджер контейнеров Podman с помощью команды:

   ```
   apt-get install podman
   ```

   e. С помощью Podman загрузите образы:

   ```
   podman pull registry.altlinux.org/k8s-c10f2/kube-apiserver:v1.31.1
   podman pull registry.altlinux.org/k8s-c10f2/kube-controller-manager:v1.31.1
   podman pull registry.altlinux.org/k8s-c10f2/kube-scheduler:v1.31.1
   podman pull registry.altlinux.org/k8s-c10f2/kube-proxy:v1.31.1
   podman pull registry.altlinux.org/k8s-c10f2/pause:3.10
   podman pull registry.altlinux.org/k8s-c10f2/etcd:3.5.15-0
   podman pull registry.altlinux.org/k8s-c10f2/coredns:v1.11.3
   ```

f. Для присвоения имени образу выполните команду:

```
podman tag registry.altlinux.org/k8s-c10f2/pause:3.10 registry.k8s.io/pause:3.10
```

g. Установите Kubernetes с помощью команд:

```
apt-get install kubernetes-kubeadm kubernetes-kubelet kubernetes-crio cri-tools
systemctl enable --now crio kubelet
swapoff -a
```

h. Для корректного создания кластера отключите использование прокси-соединения:

```
unset http_proxy
unset https_proxy
```

i. Создайте кластер:

```
kubeadm init --pod-network-cidr=10.244.0.0/16 --kubernetes-version=1.31.1 --image-
repository=registry.altlinux.org/k8s-c10f2
```

j. После создания кластера на экран будет выведена инструкция. Выполните предложенные действия.

> **Примечание:**
>
> **Рекомендуется** выполнить действия из инструкции.
> Если вы не будете выполнять все действия из инструкции, чтобы продолжить установку продукта, выполните команду для настройки переменной окружения:
>
> ```
> export KUBECONFIG=/etc/kubernetes/admin.conf
> ```
>
> Команда позволит использовать `kubectl` в рамках текущей сессии подключения. При следующем подключении или перезагрузке выполните команду повторно.

k. Установите для Kubernetes сетевой провайдер Flannel, для этого загрузите образы:

```
podman pull registry.altlinux.org/k8s-c10f2/flannel:v0.25.7
podman pull registry.altlinux.org/k8s-c10f2/flannel-cni-plugin:v1.4.0-flannel1
```

l. Загрузите манифест Flannel:

```
wget https://github.com/flannel-io/flannel/releases/latest/download/kube-flannel.yml
```

m. Отредактируйте файл-манифест:

i. Строку `image: docker.io/flannel/flannel-cni-plugin:v1.1.2` замените на:

Строка для замены:

```
image: registry.altlinux.org/k8s-c10f2/flannel-cni-plugin:v1.4.0-flannel1
```

ii. Строки `image: docker.io/flannel/flannel:v0.22.0` замените на:

Строка для замены:

```
image: registry.altlinux.org/k8s-c10f2/flannel:v0.25.7
```

> **Примечание:**
>
> В конфигурационном файле две одинаковых строки вида `image: docker.io/flannel/flannel:v0.22.0`, замените обе.
>
> В связи с периодическими обновлениями Flannel версии в заменяемых строках могут отличаться.

n. Откройте на редактирование конфигурацию CoreDNS:

```
kubectl edit cm -n kube-system coredns
```

o. Добавьте следующее содержимое:

```
     rewrite stop {
        name substring iwplatform cluster answer auto
     }
```

Вид конфигурационного файла после редактирования:

```
Corefile: |
  .:53 {
      log
      errors
      health {
         lameduck 5s
      }
      rewrite stop {
         name substring iwplatform cluster answer auto
      }
      ready
      kubernetes cluster.local in-addr.arpa ip6.arpa {
         pods insecure
         fallthrough in-addr.arpa ip6.arpa
```

```
            ttl 30
        }
        prometheus :9153
        forward . /etc/resolv.conf {
            max_concurrent 1000
        }
        cache 30
        loop
        reload
        loadbalance
    }
```

p. Сохраните изменения и закройте редактор.
q. Примените изменения в манифесте и запустите Flannel с помощью команды:

```
kubectl apply -f kube-flannel.yml
```

r. Дождитесь запуска всех подов.
Вы можете проверить их статусы с помощью команды:

```
kubectl get pods -A
```

Все поды должны быть в статусе `Running`.
s. Смените директорию для временных файлов:

```
export TMPDIR=/var/tmp
```

9. **Если вы устанавливаете продукт на сервер под управлением ОС Astra Linux Special Edition:**

Отключите настройку `astra-nochmodx-lock`.
Для этого выполните команду:

```
astra-nochmodx-lock disable
```

10. Убедитесь, что на сервере установлены актуальные дата и время.

> **Примечание:**
>
> Пакеты не входят в состав дистрибутива продукта.
>
> Пакеты могут отсутствовать в репозиториях. В этом случае загрузите и установите требуемые пакеты вручную.

6 Настройка сетевых правил доступа

Для корректной работы должны быть разрешены соединения:

Доступ к веб-интерфейсу и SSH-консоли		
Рабочая станция Администратора → Сервер Data Discovery	TCP 22	Порт сервера Data Discovery. Используется для управления сервером Data Discovery с помощью протокола SSH
Соединение	Порт	Описание
Рабочая станция Офицера Безопасности → Центр расследований	TCP 443	Порт для подключения к веб-интерфейсу Центра расследований (центральная нода Платформы). Значение по умолчанию: **443**. Вы указываете значение во время установки Платформы. Чтобы изменить номер порта: 1. Произведите сброс установки Системы (см. "Устранение неполадок"). 2. Установите Систему повторно, указав новое значение.
Соединения с внешними системами		
Сервер Data Discovery → Сервер Traffic Monitor	TCP 9100	Порт сервера Traffic Monitor. Используется для передачи данных из Data Discovery в Traffic Monitor через XAPI
Сервер Data Discovery → LDAP-сервер	TCP 3268 TCP 389	Порты LDAP-сервера. Используются для интеграции Data Discovery с LDAP-каталогами
Сервер Data Discovery → LDAPS-сервер	TCP 3269 TCP 636	Порты LDAPS-сервера. Используются для интеграции Data Discovery с LDAPS-каталогами
Сервер Data Discovery → DNS-серверы	UDP 53	Порт DNS-сервера. Используется для разрешения DNS-запросов
Сервер Data Discovery → NTP-серверы	UDP 123	Порт NTP-сервера. Используется для синхронизации времени
Сервер Data Discovery → Почтовый сервер	Для протокола :	Порты почтового сервера. Указаны значения по умолчанию.

Доступ к веб-интерфейсу и SSH-консоли		
Рабочая станция Администратора → Сервер Data Discovery	TCP 22	Порт сервера Data Discovery. Используется для управления сервером Data Discovery с помощью протокола SSH
Соединение	Порт	Описание
	• TLS - 465 • START TLS - 587	Используются для отправки почтовых уведомлений Prediction и Activity Monitor. Data Discovery использует почтовый сервер только для регулярной отправки отчетов
Подключение к ресурсам для сканирования (указаны протоколы или хранилища)		
Сервер Data Discovery → SMB	TCP 139, 445	Порт хоста с SMB-подключением. Для сканирования общих сетевых ресурсов и рабочих станций под управлением ОС MS Windows
Сервер Data Discovery → SSH	TCP 22	Порт хоста с SFTP (SSH)-подключением. Для сканирования серверов и рабочих станций под управлением ОС семейства Linux
Сервер Data Discovery → SharePoint	TCP 443	Порт хоста с SharePoint-подключением. Для сканирования библиотек документов в SharePoint Server 2016 по протоколу HTTPS
Сервер Data Discovery → FTP	TCP 21	Порт хоста с FTP-подключением. Для сканирования файловых ресурсов по протоколу FTP
Сервер Data Discovery → Nextcloud	TCP 443	Порт хоста с Nextcloud. Для сканирования хранилищ данных Nextcloud по протоколу WebDAV
Сервер Data Discovery → Alfresco	TCP 443	Порт хоста с Alfresco. Для сканирования серверов системы управления контентом Alfresco Community Edition по протоколам HTTP/HTTPS
Сервер Data Discovery → Яндекс Диск	TCP 443	Порт хоста с Яндекс Диск. Для сканирования хранилищ Яндекс Диск с использованием REST API Диска. **Примечание**: для корректного сканирования Яндекс Диска должны быть доступны следующие домены:

Доступ к веб-интерфейсу и SSH-консоли		
Рабочая станция Администратора → Сервер Data Discovery	TCP 22	Порт сервера Data Discovery. Используется для управления сервером Data Discovery с помощью протокола SSH
Соединение	Порт	Описание
		• `https://downloader.disk.yandex.ru/` • `https://storage.yandex.net/` – включая все его поддомены первого уровня (`*.storage.yandex.net`)
Примечание: Для подключения к ресурсам сканирования указаны порты по умолчанию. Если в вашей сетевой инфраструктуре используются другие порты подключения, учитывайте это при создании задач сканирования и добавлении хостов		
Работа в кластере		
Сервер Платформы в режиме `office` → Сервер Платформы в режиме `central`	TCP 6443 TCP 2379, 2380 TCP 10251, 10252, 10255	Порты сервера Платформы в режиме `central` . Используются для взаимодействия Платформы в режиме `office` с Платформой в режиме `central`
Сервер Платформы в режиме `central` → Сервер Платформы в режиме `office`	TCP 10251, 10255 TCP 30000–32767	Порты сервера Платформы в режиме `office` . Используются для взаимодействия Платформы в режиме `central` с Платформой в режиме `office`
Сервер Платформы в режиме `central` ← → Сервер Платформы в режиме `office`	UDP 8472 TCP 10250	Используются для туннелирования трафика между узлами кластера
Сервер Платформы в режиме `central` ← → Сервер Платформы в режиме `office`	TCP 10256	Порт должен быть доступен на всех нодах кластера. Используется для проверки состояния компонента `kube-proxy`

7 Установка Data Discovery

Дистрибутив представляет собой архив `iw_discovery_setup_2.2.X.xxx.tar.xz`, где `xxx` – номер сборки. Архив содержит полный набор бинарных модулей образов контейнеров, необходимых для развертывания Платформы и компонентов Data Discovery без доступа к интернету (offline-установка), а также программу установки.

Для установки Data Discovery используйте командную строку.

Вы можете установить Data Discovery:

- на отдельном сервере;
- в кластере из серверов Data Discovery;
- на сервере совместно с другими продуктами;
- в кластере с другими продуктами.

Если вы устанавливаете Data Discovery в одном кластере с другими продуктами InfoWatch, **убедитесь, что версия Платформы совпадает у всех продуктов на всех нодах**. В противном случае продукты, установленные на старой версии Платформы, нужно обновить.

Чтобы узнать версию Платформы устанавливаемого продукта:

1. Распакуйте архив с дистрибутивом продукта. В нашем примере – в директорию `/ddiscovery`. Для этого перейдите в директорию и выполните команду:

   Пример для версии 2.2.0.191:

   ```
   tar -xvf iw_discovery_setup_2.2.0.191.tar.xz
   ```

2. Выполните команду:

   ```
   cat /ddiscovery/platform/setup.yaml | grep version
   ```

> **Примечание:**
>
> Если в кластере с Data Discovery 2.2 **Платформа будет обновлена до версии 2.5**, для повторного подключения сервиса `contentbroker` к базе данных **перезапустите** его.
>
> **Для этого:**
>
> 1. Выполните команду перезапуска:
>
>    ```
>    kubectl -n infowatch rollout restart deployment/contentbroker-central
>    ```
>
> 2. Проверьте статус сервиса с помощью команды:
>
>    ```
>    kubectl -n infowatch rollout status deployment/contentbroker-central
>    ```

Важно!

Все ноды кластера должны быть доступны друг для друга в сети. Для этого откройте необходимые порты (см. *Настройка сетевых правил доступа*).

Также перед установкой убедитесь, что:

- настроено разрешение имен нод средствами DNS;
  или
- в файле `/etc/hosts` каждой ноды определено соответствие между адресами и именами нод.

Для проверки доступности нод выполните команду между нодами кластера:

```
ping <имя_ноды>
```

7.1 Варианты запуска установки Data Discovery

Вы можете запустить установку Data Discovery:

- **в интерактивном режиме** – инсталлятор последовательно запросит параметры установки (см. *Установка Data Discovery в интерактивном режиме*);
- **с указанием параметров** установки в команде запуска инсталлятора (см. *Установка Data Discovery с указанием параметров в командной строке*).

В обоих вариантах запуска инсталлятор поддерживает и совместную установку с другими продуктами Центра расследований на одном сервере, и распределенную установку в кластере.

7.2 Модуль контентного анализа

Чтобы обнаруживать объекты защиты в просканированных текстовых файлах без отправки данных на анализ в Traffic Monitor, Data Discovery использует отдельный Модуль контентного анализа.

Модуль контентного анализа должен быть установлен в одном кластере с Data Discovery. Его можно установить на отдельный сервер или совместно с Data Discovery. Для корректной работы версии Модуля контентного анализа и Traffic Monitor, конфигурация которого будет использована, должны совпадать. Data Discovery **2.2** использует Модуль контентного анализа **7.12.1**, **7.13** или **7.14**.

Вы можете запустить установку Модуля контентного анализа:

- **в интерактивном режиме** – инсталлятор последовательно запросит параметры установки (см. *Установка Модуля контентного анализа в интерактивном режиме*);
- **с указанием параметров** установки в команде запуска инсталлятора (см. *Установка Модуля контентного анализа с указанием параметров в командной строке*).

В обоих вариантах запуска инсталлятор поддерживает и совместную установку с другими продуктами Центра расследований на одном сервере, и распределенную установку в кластере.

7.3 Установка Data Discovery в интерактивном режиме

> **Примечание:**
>
> Если вы устанавливаете Data Discovery на сервер, на котором уже установлены другие продукты Центра расследований, инсталлятор не запросит часть параметров, так как эти директории и настройки уже были указаны при первичной установке Платформы.

Чтобы установить Data Discovery:

1. Проверьте, соответствует ли сервер аппаратным и программным требованиям (см. *Типы установки и обновления, совместимость с другими продуктами, аппаратные и программные требования*).
2. Выполните шаги по подготовке сервера к установке Data Discovery (см. *Подготовка сервера*).
3. Создайте новую директорию на сервере (например, `distr_discovery`):

   ```
   mkdir distr_discovery
   ```

4. Скопируйте архив **`iw_discovery_setup _2.2.X.xxx.tar.xz`** в созданную директорию. В нашем примере будет использована версия `2.2.0.191`, то есть архив будет `iw_discovery_setup _2.2.0.191.tar.xz`.
5. Перейдите в директорию, в которую был скопирован архив.
6. Распакуйте архив с дистрибутивом в эту директорию:

   ```
   tar -xvf iw_discovery_setup_2.2.0.191.tar.xz
   ```

7. **Для установки продукта на ОС Альт СП 10:**

   Используйте Kubernetes из репозиториев ОС.
   Для этого:

   a. Выполните действия для ОС Альт СП 10 из статьи *Подготовка сервера*.
   Если вы уже выполнили подготовку, пропустите этот шаг.
   b. Выполните команду:

   ```
   kubectl taint nodes <имя_ноды> node-role.kubernetes.io/control-plane:NoSchedule-
   ```

   В нашем примере:

   ```
   kubectl taint nodes ddnode node-role.kubernetes.io/control-plane:NoSchedule-
   ```

   c. На следующем шаге при запуске инсталлятора дополнительно укажите опцию **`--externalK8S=True`**
8. Запустите программу установки с помощью команды вида:

```
./setup.py install --nodemode=<режим_установки_Платформы> --
productmode=<режим_установки_продукта>
```

Чтобы установить:

- **на центральную ноду Платформы продукт в режиме central:**

  ```
  ./setup.py install --nodemode=central --productmode=central
  ```

- **на центральную ноду Платформы продукт в режиме office:**

  ```
  ./setup.py install --nodemode=central --productmode=office
  ```

- **на офисную ноду Платформы продукт в режиме central:**

  ```
  ./setup.py install --nodemode=office --productmode=central
  ```

- **на офисную ноду Платформы продукт в режиме office:**

  ```
  ./setup.py install --nodemode=office --productmode=office
  ```

> **Примечание:**
>
> Если вы не укажете режимы установки Платформы и продукта, инсталлятор по умолчанию выберет режим `central`.
>
> Если вы устанавливаете Data Discovery на сервер первым и укажете только параметр `--nodemode`, инсталлятор использует такое же значение и для параметра `--productmode`.

> **Важно!**
>
> Во время установки установленные на сервере продукты Платформы могут быть недоступны.

9. Ознакомьтесь с условиями лицензионного соглашения. Лицензионное соглашение содержит несколько страниц. Для перехода на следующую страницу нажимайте **Enter**.
10. Чтобы принять лицензионное соглашение, введите " `y` " и нажмите **Enter**.

> **Примечание:**
>
> Если отклонить лицензионное соглашение, инсталлятор прервет процесс и завершит работу.

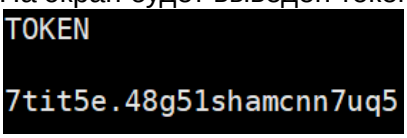
11. Если вы устанавливаете **центральную ноду Платформы** (`--nodemode=central`), введите IP-адрес сетевого интерфейса для взаимодействия с кластером в формате IPv4: `xxx.xxx.xxx.xxx` (по умолчанию: `0.0.0.0`) и нажмите **Enter**.
Если указать `0.0.0.0` , будут использованы все доступные сетевые интерфейсы.

> **Примечание:**
>
> Здесь и далее для использования значений, предложенных по умолчанию, нажимайте **Enter** без ввода значений.

12. Если вы устанавливаете **офисную ноду Платформы** (`--nodemode=office`):
    a. Укажите адрес центральной ноды Платформы в формате IP-адреса или доменного имени (например: `192.0.2.0` или `host.example.com`).
    b. Укажите токен для подключения к центральной ноде Платформы.
    Чтобы получить токен:
        i. Подключитесь к консоли управления центральной ноды Платформы по SSH.
        ii. Выполните в терминале команду для получения токена:

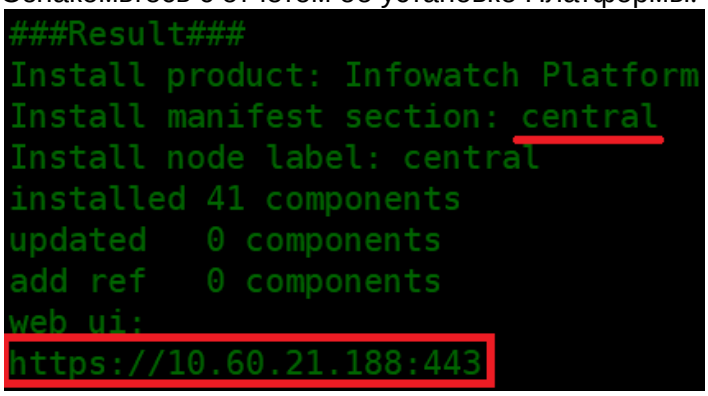
```
kubeadm token list
```

        На экран будет выведен токен:



        Токен в нашем примере: `7tit5e.48g51shamcnn7uq5`
        iii. Укажите лейбл ноды (по умолчанию: `office`) и нажмите **Enter**.
        Лейбл может состоять из букв латинского алфавита и/или цифр. Длина лейбла не должна превышать 32 символа.
13. Выделите объем оперативной памяти для размещения данных ClickHouse (по умолчанию: `80%`) и нажмите **Enter**.
14. Укажите путь для размещения данных ClickHouse (по умолчанию: `/opt/chdata`) и нажмите **Enter**.
15. Укажите путь для размещения данных PostgreSQL (по умолчанию: `/opt/pgdata`) и нажмите **Enter**.
16. Укажите путь для размещения данных NATS (по умолчанию: `/opt/natsdata`) и нажмите **Enter**.
17. Укажите путь для размещения бинарных данных (по умолчанию: `/opt/dsdata`) и нажмите **Enter**.
18. Укажите путь для временного хранения событий Traffic Monitor перед их загрузкой в Платформу (по умолчанию: `/opt/factsldrtmdata`) и нажмите **Enter**.
19. Если вы устанавливаете **центральную ноду Платформы**, укажите порт подключения к веб-интерфейсу (по умолчанию: `443`) и нажмите **Enter**.
20. Дождитесь завершения установки Платформы.

21. Ознакомьтесь с отчетом об установке Платформы.



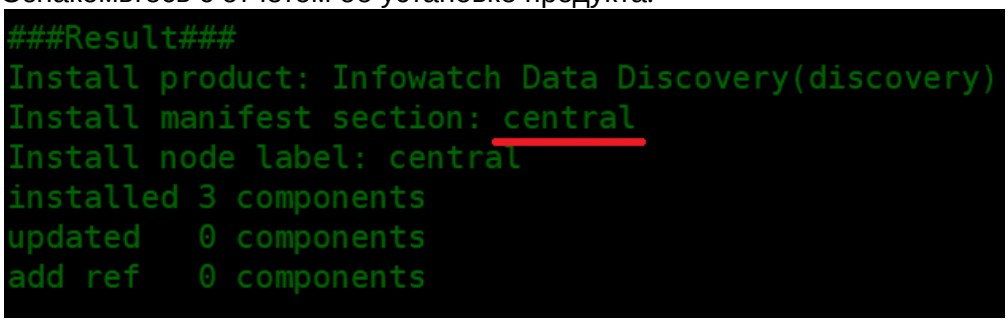
При установке центральной ноды Платформы в графе **`web ui`** указаны адрес и порт для подключения к веб-интерфейсу (в нашем примере – `https://10.60.21.188:443`).

22. Укажите путь для хранилища просканированных файлов (по умолчанию: `/opt/disctps`) и нажмите **Enter**. Этот путь будет использоваться на всех серверах Data Discovery кластера.

> **Важно!**
>
> Директория для хранилища должна быть пустой.
>
> После установки в хранилище должны находиться только файлы и директории, которые были добавлены Data Discovery. Если вы добавите другие файлы или директории в хранилище, Data Discovery может работать некорректно.

> **Примечание:**
>
> Рекомендуем, чтобы хранилище сервера находилось в отдельном разделе или диске. О настройке хранилища сервера см. *InfoWatch Data Discovery. Руководство администратора. Зоны и серверы*.

23. Дождитесь завершения установки Data Discovery.
24. Ознакомьтесь с отчетом об установке продукта:



25. Чтобы вывести на экран информацию о Платформе и продуктах в кластере, выполните команду:

```
./setup.py showproducts
```

26. Для работы Data Discovery в ОС должны быть дополнены настройки подсистемы inotify.

Для этого:

a. Проверьте наличие конфигурационного файла `/ etc/sysctl.d/discovery.conf` со следующим содержимым:

```
fs.inotify.max_user_instances = 32768
fs.inotify.max_user_watches = 32768
fs.inotify.max_queued_events = 262144
```

При наличии конфигурационного файла пропустите два следующих шага инструкции.

b. Если файл не был создан ранее скриптами инсталлятора, создайте его и добавьте содержимое из предыдущего шага.

c. Чтобы применить новые настройки подсистемы, выполните команду:

```
sysctl --load=/etc/sysctl.d/discovery.conf
```

27. Убедитесь, что все сервисы запущены, выполнив команду:

```
kubectl get pods -n infowatch
```

Команда выведет список сервисов. Каждый сервис должен иметь статус `Running`.

> **Примечание:**
>
> После установки Data Discovery скрипты инсталлятора **включают ротацию и удаление лог-файлов** (см. *InfoWatch Data Discovery. Руководство администратора.* *Ротация лог-файлов*).

28. Для использования продукта подключитесь к веб-интерфейсу Центра расследований. Для этого введите в браузере адрес и порт центральной ноды Платформы.

> **Примечание:**
>
> Если в кластере с установленным Data Discovery 2.2 **Платформа обновлена до версии 2.5**, для повторного подключения сервиса `contentbroker` к базе данных **перезапустите** его.
>
> **Для этого:**
>
> 1. На любой ноде Data Discovery выполните команду перезапуска:
>
> ```
> kubectl -n infowatch rollout restart deployment/contentbroker-central
> ```
>
> 2. Проверьте статус сервиса с помощью команды:
>
> ```
> kubectl -n infowatch rollout status deployment/contentbroker-central
> ```

7.4 Установка Data Discovery с указанием параметров в командной строке

Инсталлятор Data Discovery позволяет установить продукт с минимальным участием пользователя. Для этого запустите программу установки, указав в командной строке все требуемые параметры.

Перед запуском инсталлятора Data Discovery:

1. Проверьте, соответствует ли сервер аппаратным и программным требованиям (см. *Типы установки и обновления, совместимость с другими продуктами, аппаратные и программные требования*).
2. Выполните шаги по подготовке сервера к установке Data Discovery (см. *Подготовка сервера*).
3. Создайте новую директорию на диске (например, `distr_discovery`):

   ```
   mkdir distr_discovery
   ```

4. Скопируйте архив **`iw_discovery_setup_2.2.X.xxx.tar.xz`** в созданную директорию. В нашем примере будет использована версия `2.2.0.191`, то есть архив будет `iw_discovery_setup_2.2.0.191.tar.xz`.
5. Перейдите в директорию, в которую был скопирован архив.
6. Распакуйте архив с дистрибутивом в эту директорию:

   ```
   tar -xvf iw_discovery_setup_2.2.0.191.tar.xz
   ```

Ознакомьтесь со списком ключей в таблице ниже: с их помощью вы можете указать инсталлятору параметры установки продукта. Указанные значения по умолчанию – значения, предлагаемые инсталлятором в интерактивном режиме.

> **Примечание:**
>
> Если вы не укажете режимы установки Платформы и продукта, инсталлятор использует значение `central`.

Ключ	Значение по умолчанию	Доступные значения	Описание
`--nodemode`	`central`	`central, office`	Режим установки Платформы.
`--productmode`	`central`	`central, office`	Режим установки продукта.

Ключ	Значение по умолчанию	Доступные значения	Описание
`--externalK8S`		`True`	Использование Kubernetes из репозиториев ОС Альт СП.
`--loc`	`ru`	`ru`, `en`	Локализация продукта, язык для предустановленных сущностей. Параметр **необязателен** для запуска инсталлятора.
`--licagree`	`no`	`yes,no`	Принятие лицензионного соглашения.
`--disable-check-free-disk`	-	-	Отключение проверки свободного места на диске: • 20 ГБ для размещения образов; • 100 ГБ для размещения данных в указанных директориях. **Примечание**: если директории для хранения данных находятся на разных дисках, то проверится наличие 100 ГБ свободного места на каждом из них.
`--disable-check-total-ram`	-	-	Отключение проверки 8 ГБ ОЗУ на компьютере.
`--chmemsize`	`80%`	1–100% или числовое значение. Если не указать %, указанное значение будет принято за объем в ГБ.	Объем оперативной памяти, используемой для данных ClickHouse.
`--chdbpath`	`/opt/chdata`		Путь для размещения данных ClickHouse.
`--natspath`	`/opt/natsdata`		Путь для размещения данных NATS.
`--pgdbpath`	`/opt/pgdata`		Путь для размещения файлов базы данных PostgreSQL.

Ключ	Значение по умолчанию	Доступные значения	Описание
`--dsdatapath`	`/opt/dsdata`		Путь где будут располагаться бинарные данные.
`--factsldrtmdatapath`	`/opt/factsldrtmdata`		Путь для временного размещения событий Traffic Monitor перед их загрузкой в Платформу.
`--advertise`	`0.0.0.0`		IP-адрес сетевого интерфейса, на котором работает сервис `kubelet`. Указывайте при установке центральной ноды Платформы.
`--guiport`	`443`		Порт подключения к веб-интерфейсу. Указывайте при установке центральной ноды Платформы.
`--cluster`			Адрес центральной ноды Платформы в кластере. Указывайте при установке офисной ноды Платформы.
`--clustertoken`			Токен для подключения к центральной ноде Платформы. Чтобы получить токен, в терминале центральной ноды Платформы выполните команду: `kubeadm token list` Указывайте при установке офисной ноды Платформы.
`--nodelabel`	`office`		Лейбл ноды. Указывайте при установке офисной ноды Платформы.
`--tpspath`	`/opt/disctps`		Укажите путь для хранилища просканированных файлов. Этот путь будет использоваться на всех серверах Data Discovery кластера. Указывайте при установке центральной ноды Data Discovery.

Чтобы установить Data Discovery:

1. **Для установки продукта на ОС Альт СП 10:**

   Используйте Kubernetes из репозиториев ОС.
   Для этого:

   a. Выполните действия для ОС Альт СП 10 из статьи *Подготовка сервера*.
   Если вы уже выполнили подготовку, пропустите этот шаг.
   b. Выполните команду:

   ```
   kubectl taint nodes <имя_ноды> node-role.kubernetes.io/control-plane:NoSchedule-
   ```

   В нашем примере:

   ```
   kubectl taint nodes ddnode node-role.kubernetes.io/control-plane:NoSchedule-
   ```

   c. На следующем шаге при запуске инсталлятора дополнительно укажите опцию `--externalK8S=True`
2. Запустите программу установки Data Discovery с указанием параметров.

   Пример команды установки центральной ноды Платформы и продукта в режиме central:

   ```
   ./setup.py install --nodemode=central --productmode=central --licagree=yes --advertise=0.0.0.0 --chmemsize=80% --chdbpath=/opt/chdata --pgdbpath=/opt/pgdata --dsdatapath=/opt/dsdata --factsldrtmdatapath=/opt/factsldrtmdata --guiport=443 --natspath=/opt/natsdata --tpspath=/opt/disctps
   ```

   Пример команды установки офисной ноды Платформы и продукта в режиме office:

   ```
   ./setup.py install --nodemode=office --productmode=office --licagree=yes --nodelabel=lv-426 --cluster=10.60.21.188 --clustertoken=bvp2e1.e9zilgwg7v2urtsr --chdbpath=/opt/chdata --chmemsize=80% --natspath=/opt/natsdata --pgdbpath=/opt/pgdata --dsdatapath=/opt/dsdata --factsldrtmdatapath=/opt/factsldrtmdata
   ```

3. Если вы указали все обязательные параметры, инсталлятор запустит установку.
   Если вы не указали часть параметров, инсталлятор запустит интерактивный режим и попросит задать значения.
4. Дождитесь завершения установки Data Discovery.
   Чтобы вывести на экран информацию о Платформе и продуктах в кластере, выполните команду:

   ```
   ./setup.py showproducts
   ```

5. Для работы Data Discovery в ОС должны быть дополнены настройки подсистемы inotify.
   Для этого:

   a. Проверьте наличие конфигурационного файла `/ etc/sysctl.d/discovery.conf` со следующим содержимым:

```
fs.inotify.max_user_instances = 32768
fs.inotify.max_user_watches = 32768
fs.inotify.max_queued_events = 262144
```

При наличии конфигурационного файла пропустите два следующих шага инструкции.

b. Если файл не был создан ранее скриптами инсталлятора, создайте его и добавьте содержимое из предыдущего шага.

c. Чтобы применить новые настройки подсистемы, выполните команду:

```
sysctl --load=/etc/sysctl.d/discovery.conf
```

6. Убедитесь, что все сервисы запущены, выполнив команду:

```
kubectl get pods -n infowatch
```

Команда выведет список сервисов. Каждый сервис должен иметь статус `Running`.

> **Примечание:**
>
> После установки Data Discovery скрипты инсталлятора **включают ротацию и удаление лог-файлов** (о ротации и удалении лог-файлов см. *InfoWatch Data Discovery. Руководство администратора. Ротация лог-файлов*).

7. Для использования продукта подключитесь к веб-интерфейсу Центра расследований. Для этого введите в браузере адрес и порт центральной ноды Платформы.

> **Примечание:**
>
> Если в кластере с установленным Data Discovery 2.2 **Платформа обновлена до версии 2.5**, для повторного подключения сервиса **`contentbroker`** к базе данных **перезапустите** его.
>
> **Для этого:**
>
> 1. На любой ноде Data Discovery выполните команду перезапуска:
>
> ```
> kubectl -n infowatch rollout restart deployment/contentbroker-central
> ```
>
> 2. Проверьте статус сервиса с помощью команды:
>
> ```
> kubectl -n infowatch rollout status deployment/contentbroker-central
> ```

7.5 Установка Модуля контентного анализа в интерактивном режиме

> **Примечание:**
>
> Модуль контентного анализа должен быть установлен в одном кластере с Data Discovery.
>
> Если вы устанавливаете Модуль контентного анализа на сервер, на котором уже установлены продукты Центра расследований, инсталлятор не запросит часть параметров, так как эти директории и настройки уже были указаны при первичной установке Платформы.
>
> Инструкция актуальна для всех поддерживаемых версий: 7.12.1, 7.13 и 7.14.

Чтобы установить Модуль контентного анализа:

1. Проверьте, соответствует ли сервер аппаратным и программным требованиям (см. *Типы установки и обновления, совместимость с другими продуктами, аппаратные и программные требования*).
2. Выполните шаги по подготовке сервера к установке Модуля контентного анализа (см. *Подготовка сервера*).
3. Создайте новую директорию на сервере (например, `distr_tech`):

   ```
   mkdir distr_tech
   ```

4. Скопируйте архив **`iw_tm_setup_7.13.0.xxx.tar.xz`** в созданную директорию.
   В нашем примере будет использована версия `7.13.0.167`, соответственно, архив – **`iw_tm_setup`**`_7.13.0.167.tar.xz`.
5. Перейдите в директорию, в которую был скопирован архив.
6. Распакуйте архив с дистрибутивом в эту директорию:

   ```
   tar -xvf iw_tm_setup_7.13.0.167.tar.xz
   ```

7. **Для установки модуля на ОС Альт СП 10:**

   Используйте Kubernetes из репозиториев ОС.
   Для этого:

   a. Выполните действия для ОС Альт СП 10 из статьи *Подготовка сервера*.
   Если вы уже выполнили подготовку, пропустите этот шаг.
   b. Выполните команду:

   ```
   kubectl taint nodes <имя_ноды> node-role.kubernetes.io/control-plane:NoSchedule-
   ```

   В нашем примере:

   ```
   kubectl taint nodes technode node-role.kubernetes.io/control-plane:NoSchedule-
   ```

c. На следующем шаге при запуске инсталлятора дополнительно укажите опцию `--externalK8S=True`

8. Запустите программу установки с помощью команды вида:

```
./setup.py install --nodemode=<режим_установки_Платформы> --
productmode=<режим_установки_модуля>
```

Чтобы установить:

- **на центральную ноду Платформы модуль в режиме central:**

```
./setup.py install --nodemode=central --productmode=central
```

- **на центральную ноду Платформы модуль в режиме office:**

```
./setup.py install --nodemode=central --productmode=office
```

- **на офисную ноду Платформы модуль в режиме central:**

```
./setup.py install --nodemode=office --productmode=central
```

- **на офисную ноду Платформы модуль в режиме office:**

```
./setup.py install --nodemode=office --productmode=office
```

Если вы не укажете режимы установки Платформы и модуля, инсталлятор по умолчанию выберет режим `central`.

> **Важно!**
>
> Во время установки могут быть недоступны продукты Платформы, установленные на сервере.

9. Ознакомьтесь с условиями лицензионного соглашения. Лицензионное соглашение содержит несколько страниц. Для перехода на следующую страницу нажмите **Enter**.
10. Чтобы принять лицензионное соглашение, введите " `y` " и нажмите **Enter**.

> **Примечание:**
>
> Если отклонить лицензионное соглашение, инсталлятор прервет процесс и завершит работу.

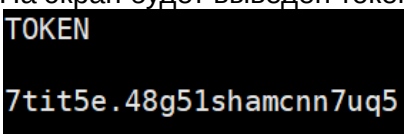
11. Если вы устанавливаете **центральную ноду Платформы** (`--nodemode=central`), введите IP-адрес сетевого интерфейса для взаимодействия с кластером в формате IPv4: " `xxx.xxx.xxx.xxx` " (по умолчанию: `0.0.0.0`) и нажмите **Enter**.
Если указать `0.0.0.0` , будут использованы все доступные сетевые интерфейсы.

> **Примечание:**
>
> Здесь и далее для использования значений, предложенных по умолчанию, нажмите **Enter** без ввода значений.

12. Если вы устанавливаете **офисную ноду Платформы** (`--nodemode=office`):
    a. Укажите адрес центральной ноды Платформы в формате IP-адреса или доменного имени (например: `192.0.2.0` или `host.example.com`).
    b. Укажите токен для подключения к центральной ноде Платформы.
    Чтобы получить токен для подключения к центральной ноде Платформы:
        i. Подключитесь к консоли центральной ноды Платформы по SSH.
        ii. Выполните в терминале команду для получения токена:

        ```
        kubeadm token list
        ```

        На экран будет выведен токен:

        Токен в нашем примере: `7tit5e.48g51shamcnn7uq5`
        iii. Укажите лейбл ноды (по умолчанию: `office`) и нажмите **Enter**.
        Лейбл может состоять из букв латинского алфавита и/или цифр. Длина лейбла не должна превышать 32 символа.
13. Выделите объем оперативной памяти для размещения данных ClickHouse (по умолчанию: `80%`) и нажмите **Enter**.
14. Укажите путь для размещения данных ClickHouse (по умолчанию: `/opt/chdata`) и нажмите **Enter**.
15. Укажите путь для размещения данных PostgreSQL (по умолчанию: `/opt/pgdata`) и нажмите **Enter**.
16. Укажите путь для размещения данных NATS (по умолчанию: `/opt/natsdata`) и нажмите **Enter**.
17. Укажите путь для размещения бинарных данных (по умолчанию: `/opt/dsdata`) и нажмите **Enter**.
18. Укажите путь для временного хранения событий Traffic Monitor перед их загрузкой в Платформу (по умолчанию: `/opt/factsldrtmdata`) и нажмите **Enter**.
19. Если вы устанавливаете **центральную ноду Платформы**, укажите порт подключения к веб-интерфейсу (по умолчанию: `443`) и нажмите **Enter**.
20. Дождитесь завершения установки Платформы.
21. Ознакомьтесь с отчетом об установке Платформы.
22. Модуль контентного анализа использует конфигурацию Traffic Monitor. Для загрузки конфигурации укажите адрес сервера с базой данных Traffic Monitor и нажмите **Enter**.

> **Важно!**
>
> Поддерживается взаимодействие **только с СУБД PostgreSQL**.

Если вы используете **Traffic Monitor 7.10 или более ранние** версии, объекты защиты не должны содержать элементы технологии "Автолингвист". Для работы модуля контентного анализа с конфигурациями этих версий Traffic Monitor удалите из объектов защиты элементы технологии "Автолингвист".

Примечание:

Если Traffic Monitor установлен на нескольких серверах или использует удаленную базу данных, укажите именно тот сервер, на котором развернута СУБД.

Если в кластере будет несколько серверов с Модулем контентного анализа, у них должна быть одинаковая конфигурация.

23. Укажите порт для подключения к базе данных Traffic Monitor (по умолчанию: `5433`) и нажмите **Enter**.
24. Для доступа к базе данных укажите пароль пользователя **`iwtm_linux`**.

Примечание:

Если вы не меняли имя пользователя и пароль при установке Traffic Monitor, укажите значение по умолчанию (см. InfoWatch Traffic Monitor. *Руководство по установке*. *Предустановленные серверные параметры*).

25. Дождитесь завершения установки Модуля контентного анализа.
26. Чтобы вывести на экран информацию о Платформе и продуктах в кластере, выполните команду:

```
./setup.py showproducts
```

27. Убедитесь, что все сервисы запущены, выполнив команду:

```
kubectl get pods -n infowatch
```

Команда выведет список сервисов. Каждый сервис должен иметь статус `Running`.

28. Для работы Data Discovery с Модулем контентного анализа нужны объекты защиты. Чтобы добавить их в Центр расследований, создайте и запустите интеграцию с Traffic Monitor (подробнее см. InfoWatch Data Discovery. *Руководство администратора*, *Интеграция с InfoWatch Traffic Monitor*). Рекомендуем добавлять объекты защиты из той же конфигурации Traffic Monitor, которая использована при установке Модуля контентного анализа.
Если в кластере не установлены продукты, которые используют события Traffic Monitor, в настройках интеграции вы можете выбрать "Протокол обмена данными 2.0", чтобы загрузить только конфигурацию, без событий. Это снизит нагрузку и сократит время добавления конфигурации в продукт.

После установки в **Настройках** Центра расследований в разделе **Зоны и серверы**:

- добавится сервер с ролью **Контентный анализ** – если Модуль контентного анализа установлен отдельно от Data Discovery;

- существующему серверу добавится роль **Контентный анализ** – если Модуль контентного анализа установлен на один сервер с Data Discovery.

Примечание:

Удаление компонентов Модуля контентного анализа выполняется аналогично продуктам на Платформе:

1. Перейдите в директорию на сервере, в которую был распакован архив с дистрибутивом.
2. Выполните команду:

```
./setup.py remove
```

7.6 Установка Модуля контентного анализа с указанием параметров в командной строке

Примечание:

Модуль контентного анализа должен быть установлен в одном кластере с Data Discovery.

Если вы устанавливаете Модуль контентного анализа на сервер, на котором уже установлены продукты Центра расследований, инсталлятор не запросит часть параметров, так как эти директории и настройки уже были указаны при первичной установке Платформы.

Инструкция актуальна для всех поддерживаемых версий: 7.12.1, 7.13 и 7.14.

Инсталлятор Модуля контентного анализа позволяет установить продукт с минимальным участием пользователя. Для этого запустите программу установки, указав в командной строке все требуемые параметры.

Перед запуском инсталлятора Модуля контентного анализа:

1. Проверьте, соответствует ли сервер аппаратным и программным требованиям (см. *Типы установки и обновления, совместимость с другими продуктами, аппаратные и программные требования*).
2. Выполните шаги по подготовке сервера к установке Модуля контентного анализа (см. *Подготовка сервера*).
3. Создайте новую директорию на сервере (например, `distr_cas`):

```
mkdir distr_cas
```

4. Скопируйте архив **`iw_tm_setup_7.13.0.xxx.tar.xz`** в созданную директорию.
   В нашем примере будет использована версия `7.13.0.167`, соответственно, архив – **`iw_tm_setup`**`_7.13.0.167.tar.xz`.
5. Перейдите в директорию, в которую был скопирован архив.

6. Распакуйте архив с дистрибутивом в эту директорию:

```
tar -xvf iw_tm_setup_7.13.0.167.tar.xz
```

Ознакомьтесь со списком ключей в таблице ниже: с их помощью вы можете указать инсталлятору параметры установки продукта. Указанные значения по умолчанию – значения, предлагаемые инсталлятором в интерактивном режиме.

> **Примечание:**
>
> Если вы не укажете режимы установки Платформы и продукта, инсталлятор использует значение `central`.

Ключ	Значение по умолчанию	Доступные значения	Описание
`--nodemode`	`central`	`central, office`	Режим установки Платформы.
`--productmode`	`central`	`central, office`	Режим установки модуля.
`--externalK8S`		`True`	Использование Kubernetes из репозиториев ОС Альт СП.
`--loc`	`ru`	`ru, en`	Локализация продукта, язык для предустановленных сущностей. Параметр **необязателен** для запуска инсталлятора.
`--licagree`	`no`	`yes, no`	Принятие лицензионного соглашения.
`--disable-check-free-disk`	-	-	Отключение проверки свободного места на диске: • 20 ГБ для размещения образов; • 100 ГБ для размещения данных в указанных директориях. **Примечание**: если директории для хранения данных находятся на разных дисках , то проверится наличие 100 ГБ свободного места на каждом из них.

Ключ	Значение по умолчанию	Доступные значения	Описание
`--disable-check-total-ram`	-	-	Отключение проверки 8 ГБ ОЗУ на компьютере.
`--chmemsize`	`80%`	1–100% или числовое значение. Если не указать %, указанное значение будет принято за объем в ГБ.	Объем оперативной памяти, используемой для данных ClickHouse.
`--chdbpath`	`/opt/chdata`		Путь для размещения данных ClickHouse.
`--natspath`	`/opt/natsdata`		Путь для размещения данных NATS.
`--pgdbpath`	`/opt/pgdata`		Путь для размещения файлов базы данных PostgreSQL.
`--dsdatapath`	`/opt/dsdata`		Путь где будут располагаться бинарные данные.
`--factsldrtmdatapath`	`/opt/factsldrtmdata`		Путь для временного размещения событий Traffic Monitor перед их загрузкой в Платформу.
`--advertise`	`0.0.0.0`		IP-адрес сетевого интерфейса, на котором работает сервис `kubelet`. Указывайте при установке центральной ноды Платформы.
`--guiport`	`443`		Порт подключения к веб-интерфейсу. Указывайте при установке центральной ноды Платформы.
`--cluster`			Адрес центральной ноды Платформы в кластере. Указывайте при установке офисной ноды Платформы.
`--clustertoken`			Токен для подключения к центральной ноде Платформы. Чтобы получить токен, в терминале центральной ноды Платформы выполните команду:

Ключ	Значение по умолчанию	Доступные значения	Описание
			`kubeadm token list` Указывайте при установке офисной ноды Платформы.
`--nodelabel`	`office`		Лейбл ноды. Указывайте при установке офисной ноды Платформы.
`--tm-db-host`			Адрес сервера с базой данных Traffic Monitor. Поддерживается взаимодействие только с СУБД PostgreSQL. Если в кластере будет несколько серверов с Модулем контентного анализа, у них должна быть одинаковая конфигурация. Если вы используете **Traffic Monitor 7.10 или более ранние** версии, объекты защиты не должны содержать элементы технологии Автолингвист. Для работы модуля контентного анализа с конфигурациями этих версий Traffic Monitor, удалите из объектов защиты элементы технологии Автолингвист.
`--tm-db-port`	`5443`		Порт для подключения к базе данных Traffic Monitor.
`--tm-db-pass-iwtm-linux`			Пароль пользователя **`iwtm_linux`** для доступа к базе данных.

Чтобы установить Модуль контентного анализа:

1. **Для установки модуля на ОС Альт СП 10:**

   Вы должны использовать Kubernetes из репозиториев ОС.
   Для этого:

   a. Выполните действия для ОС Альт СП 10 из статьи *Подготовка сервера*.
   Если вы уже выполнили подготовку, пропустите этот шаг.
   b. Выполните команду:

```
kubectl taint nodes <имя_ноды> node-role.kubernetes.io/control-plane:NoSchedule-
```

В нашем примере:

```
kubectl taint nodes technode node-role.kubernetes.io/control-plane:NoSchedule-
```

c. На следующем шаге при запуске инсталлятора дополнительно укажите опцию `--externalK8S=True`

2. Запустите программу установки с указанием параметров.

Пример команды установки центральной Платформы и модуля в режиме central:

```
./setup.py install --nodemode=central --productmode=central --licagree=yes --advertise=0.0.0.0 --chmemsize=80% --chdbpath=/opt/chdata --pgdbpath=/opt/pgdata --dsdatapath=/opt/dsdata --factsldrtmdatapath=/opt/factsldrtmdata --guiport=443 --natspath=/opt/natsdata --tm-db-host=10.60.22.33 --tm-db-port=5443 --tm-db-pass-iwtm-linux=xxXX1234
```

Пример команды установки офисной Платформы и модуля в режиме office:

```
./setup.py install --nodemode=office --productmode=office --licagree=yes --nodelabel=lv-426 --cluster=10.60.21.188 --clustertoken=bvp2e1.e9zilgwg7v2urtsr --chdbpath=/opt/chdata --chmemsize=80% --natspath=/opt/natsdata --pgdbpath=/opt/pgdata --dsdatapath=/opt/dsdata --factsldrtmdatapath=/opt/factsldrtmdata --tm-db-host=10.60.22.33 --tm-db-port=5443 --tm-db-pass-iwtm-linux=xxXX1234
```

3. Если вы указали все обязательные параметры, инсталлятор запустит установку.
   Если вы не указали часть параметров, инсталлятор запустит интерактивный режим и попросит задать значения.
4. Дождитесь завершения установки Модуля контентного анализа.
5. Чтобы вывести на экран информацию о Платформе и продуктах в кластере, выполните команду:

```
./setup.py showproducts
```

6. Убедитесь, что все сервисы запущены, выполнив команду:

```
kubectl get pods -n infowatch
```

   Команда выведет список сервисов. Каждый сервис должен иметь статус `Running`.
7. Для работы Data Discovery с Модулем контентного анализа нужны объекты защиты. Чтобы добавить их в Центр расследований, создайте и запустите интеграцию с Traffic Monitor (подробнее см. InfoWatch Data Discovery. *Руководство администратора. Интеграция с InfoWatch Traffic Monitor*). Рекомендуется добавлять объекты защиты из той же конфигурации Traffic Monitor, которая использована при установке Модуля контентного анализа.
   Если в кластере не установлены продукты, которые используют события Traffic Monitor, в настройках интеграции вы можете выбрать "Протокол обмена данными 2.0", чтобы

загрузить только конфигурацию, без событий. Это снизит нагрузку и сократит время добавления конфигурации в продукт.

После установки в **Настройках** Центра расследований в разделе **Зоны и серверы**:

- добавится сервер с ролью **Контентный анализ** – если Модуль контентного анализа установлен отдельно от Data Discovery;
- существующему серверу добавится роль **Контентный анализ** – если Модуль контентного анализа установлен на один сервер с Data Discovery.

Примечание:

Удаление компонентов Модуля контентного анализа выполняется аналогично продуктам на Платформе:

1. Перейдите в директорию на сервере, в которую был распакован архив с дистрибутивом.
2. Выполните команду:

```
./setup.py remove
```

8 Установка лицензии и плагина Traffic Monitor

8.1 Экспорт данных в Traffic Monitor

Data Discovery позволяет экспортировать копии файлов для анализа в Traffic Monitor. Для этого в Traffic Monitor необходимы **плагин и лицензия на модуль перехвата**.

Чтобы добавить лицензию:

1. Войдите под учетной записью Офицера безопасности в Консоль управления сервера Traffic Monitor.
2. Перейдите в раздел **Управление** -> **Лицензии**.
3. В области **Лицензии** нажмите .
4. Добавьте полученный от сотрудников Технической поддержки InfoWatch файл лицензии **tm_license_yyyy_mm_dd.license**, используя Проводник.

После успешного добавления лицензии информация о ней будет доступна в разделе **Управление** -> **Лицензии**.

Для экспорта данных используйте **предустановленный плагин** в Консоли управления Traffic Monitor:

- InfoWatch Crawler - для Traffic Monitor версии 7.3 и ниже;
- InfoWatch Data Discovery - для Traffic Monitor выше 7.3.

Плагин отображается, только если в Traffic Monitor установлена действующая лицензия.

Информация о плагине будет доступна в разделе **Управление → Плагины** в Консоли управления Traffic Monitor. Для последующей отправки данных в Traffic Monitor необходим токен, который создается автоматически и доступен в разделе **Управление → Плагины → Токены**.

Подробнее о настройке экспорта данных см. InfoWatch Data Discovery. Руководство администратора, Зоны и серверы.

8.2 Импорт данных из Traffic Monitor

Для получения данных из Traffic Monitor в веб-интерфейсе Центра расследований необходимо настроить интеграцию (подробнее см. *InfoWatch Data Discovery*. Руководство администратора, Интеграция с InfoWatch Traffic Monitor). Для этого в Traffic Monitor должны быть добавлены **лицензия на экспорт данных и плагин Центра расследований**. Токен для интеграции создается автоматически после добавления плагина. Токен доступен в разделе **Управление** > **Плагины** Консоли Управления Traffic Monitor, на вкладке **Токены** плагина **InfoWatch Investigation Centre**.

> **Примечание:**
>
> Data Discovery не использует данные по событиям Traffic Monitor. Функциональность могут использовать другие продукты, установленные в кластере Центра расследований.

В **Traffic Monitor 7.10 и более новых** версиях бессрочная лицензия на экспорт данных в Центр расследований и плагин **InfoWatch Investigation Centre** предустановлены.

В **Traffic Monitor 7.9 и более ранних** версиях добавьте лицензию на экспорт данных и плагин вручную, как описано ниже.

Чтобы добавить лицензию вручную:

1. Войдите под учетной записью Офицера безопасности в Консоль управления сервера Traffic Monitor.
2. Перейдите в раздел **Управление** > **Лицензии**.
3. В области **Лицензии** нажмите .
4. Добавьте полученный от сотрудников Технической поддержки InfoWatch файл лицензии **tm_license_yyyy_mm_dd.license**, используя Проводник.

После успешного добавления лицензии информация о ней будет доступна в разделе **Управление** > **Лицензии**.

Важно!

Для экспорта данных в Центр расследований в лицензии для Traffic Monitor должны быть указаны:

- компания, выпускающая лицензию - **IW**;
- система-получатель данных - **Platform**.

Чтобы добавить плагин вручную:

1. Войдите под учетной записью Офицера безопасности в Консоль управления сервера Traffic Monitor.
2. Перейдите в раздел **Управление** > **Плагины**.
3. В области **Плагины** нажмите .
4. В открывшемся окне нажмите **Выбрать файл**.
5. Добавьте архив с файлом плагина, используя Проводник.
   Плагин представляет собой zip-архив `platform/tmplugin/platform.zip`. Папка с плагином расположена в директории, куда был распакован архив дистрибутива.

После успешного добавления плагина информация о нем будет доступна в разделе **Управление** > **Плагины**.

Как загрузить плагин и лицензию в Traffic Monitor одним архивом, см. статью базы знаний "Как привязать плагин к лицензии".

9 Устранение неполадок

Во время установки Системы возможны сбои: прерывания работы программы установки, отключение электричества и т.д. В этом случае Система будет установлена некорректно. Для решения этой проблемы следует сбросить установку. После сброса все шаги по установке необходимо повторить. После сброса все данные сохранятся в Системе. Поэтому при повторной установке на вопрос об удалении существующих данных выберите ответ **Нет**.

Чтобы сбросить установку Системы:

1. Если Data Discovery и Prediction установлены совместно, остановите расчет данных Prediction перед сбросом установки (подробнее см. *Обязательные действия для обновления продуктов на Платформе, которые установлены совместно с Prediction*).
2. Запомните или сохраните пути к текущим директориям хранения данных. Получить список директорий можно, выполнив команду:

   ```
   cat /var/lib/iwplatform/data/volumes.txt
   ```

3. Запомните или сохраните лейблы нод. Чтобы посмотреть список лейблов в кластере, выполните команду:

   ```
   kubectl get node -n infowatch -o jsonpath="{range .items[*]}
   {.metadata.annotations.flannel\.alpha\.coreos\.com/public\-ip}
   {.metadata.labels.kubernetes\.io/hostname} {.metadata.labels.iwnode}{'\n'}{end}"
   ```

4. Перейдите в директорию, где распакован дистрибутив продукта.
5. Выполните команду сброса установки Системы:

   ```
   /setup.py reset
   ```

6. После сброса при установке присвойте нодам кластера те же лейблы, которые были до начала процедуры сброса.

Важно!

При сбросе установки Системы автоматически производится сброс других продуктов Платформы, установленных на сервере.

Если вы сбросили установку на центральной ноде Платформы, станут недоступны компоненты всех продуктов в кластере. В этом случае выполните сброс на каждой офисной ноде. После установки Data Discovery повторно установите и другие продукты в кластере.

10 Обновление Data Discovery

Data Discovery поддерживает обновление с версий:

- **2.0**
- **2.1**

> **Примечание:**
>
> Для обновления с более ранних версий, сначала обновите Data Discovery до версий из списка выше.

Дистрибутив представляет собой архив `iw_discovery_setup_2.2.X.xxx.tar.xz`, где `xxx` – номер сборки. Этот архив содержит полный набор бинарных модулей образов контейнеров, необходимых для развертывания Платформы и компонентов Data Discovery без доступа к интернету (offline-установка). Кроме этого, данный архив содержит программу обновления.

Для обновления Data Discovery используйте командную строку.

10.1 Настройки и данные Системы

- **Задачи сканирования**:
  - Обработка ресурсов в рамках задач сканирования продолжится с того места, где она закончилась до обновления Системы.

    > **Важно!**
    >
    > Перед обновлением остановите все задачи сканирования в интерфейсе Data Discovery. Запустите задачи вручную, когда обновление будет завершено на всех серверах кластера.

- **Конфигурационные файлы**:
  - Если обновление проводится с помощью сброса текущей установки, в некоторых конфигурационных файлах могут быть установлены значения по умолчанию. Если вы редактировали конфигурационные файлы Системы, проверьте значения в файлах после обновления.

10.2 Варианты обновления Data Discovery

> **Важно!**
>
> В случае совместной установки Data Discovery с Prediction нужно остановить расчет данных Prediction на время обновления Data Discovery. Подробнее об этом см. *InfoWatch Prediction. Руководство по установке. Прямое обновление Prediction* или *Обновление Prediction с помощью сброса установки*.

Data Discovery 2.2 поддерживает:

- прямое обновление

- обновление с помощью сброса текущей установки и переустановки на старые данные

Если Data Discovery и другие продукты установлены в разных кластерах, вы можете обновить продукты в другом кластере и установить в него новую версию Data Discovery (подробнее см. *Переустановка Data Discovery в другой кластер*).

Важно!

Если вы обновляете Data Discovery в одном кластере с другими продуктами InfoWatch, убедитесь, что версия Платформы совпадает у всех продуктов на всех нодах. В противном случае продукты, установленные на старой версии Платформы, нужно обновить.

Чтобы узнать версию Платформы обновляемого продукта:

1. Распакуйте архив с дистрибутивом продукта. В нашем примере, в директорию `/ddiscovery`. Для этого перейдите в директорию и выполните команду:

   Пример для версии 2.2.0.191:

   ```
   tar -xvf iw_discovery_setup_2.2.0.191.tar.xz
   ```

2. Выполните команду:

   ```
   cat /ddiscovery/platform/setup.yaml | grep version
   ```

10.3 Модуль контентного анализа

Data Discovery взаимодействует с Модулем контентного анализа 7.12.1, 7.13 и 7.14. Модуль контентного анализа поддерживает прямое обновление с версий 7.11 и 7.12.

10.4 Особенности обновления кластера Центра расследований

Если в кластере установлено несколько продуктов, перед обновлением учтите особенности обновления каждого из них.

1. Если в кластере установлен **Prediction**, на время любого обновления **остановите** его **расчет** данных (подробнее см. *InfoWatch Prediction. Руководство по установке. Прямое обновление Prediction* или *Обновление Prediction с помощью сброса установки*).
2. Если на обновляемой ноде установлен **Activity Monitor**, **обновите** его **первым**.
3. Если вы обновляете продукты с помощью сброса установки (`reset`):
   a. Чтобы сохранить настройки зон и серверов в новой версии Data Discovery, перед сбросом установки подключитесь к ноде, на которой Data Discovery установлен в режиме `central` и **остановите сервис Discoperator** с помощью команды:

   ```
   kubectl scale --replicas=0 deployment discoperator-<лейбл_ноды> -n infowatch
   ```

Пример для Data Discovery в режиме central на офисной ноде с лейблом lv426:

```
kubectl scale --replicas=0 deployment discoperator-lv426 -n infowatchkubectl scale
--replicas=0 deployment discoperator-lv426 -n infowatch
```

b. **Первым** должен быть **переустановлен Activity Monitor**.
c. **Prediction 3.4** должен быть установлен **последним**.

Также перед обновлением кластера рекомендуем ознакомиться с инструкциями всех установленных продуктов.

10.5 Прямое обновление Data Discovery

Важно!

После обновления на версию 2.2 останутся только данные последних завершенных сканирований, **данные прошлых сканирований будут удалены**.

Чтобы обновить распределенный кластер Центра расследований:

1. **Обновите центральную ноду** кластера и все установленные на ней продукты.
2. **Обновите офисные ноды** и все установленные на них продукты.

Если вы используете удаленное подключение по протоколу SSH, чтобы избежать проблем в случае разрыва соединения с сервером, рекомендуем использовать утилиту Screen. При отключении от утилиты запущенные в ней процессы не прервутся, что позволит безопасно продолжить обновление.

Чтобы обновить Data Discovery:

1. **Остановите все задачи** сканирования в веб-интерфейсе Data Discovery.
2. Подготовьте сервер к обновлению (см. *Подготовка сервера*).
3. Создайте новую директорию на диске (например, `distr_discovery`):

```
mkdir distr_discovery
```

4. Скопируйте архив **`iw_discovery_setup_2.2.x.xxx.tar.xz`** в созданную директорию. В нашем примере будет использована версия `2.2.0.191`, то есть архив будет `iw_discovery_setup_2.2.0.191.tar.xz`.
5. Перейдите в директорию, в которую был скопирован архив.
6. Распакуйте архив с дистрибутивом в эту директорию:

```
tar -xvf iw_discovery_setup_2.2.0.191.tar.xz
```

7. Убедитесь, что в разделе с директорией `/var/lib/containerd` доступно не менее 50 ГБ дискового пространства.
   Чтобы проверить доступное дисковое пространство, используйте команду:

```
df -h /var/lib/containerd
```

Примечание:

Чтобы освободить дисковое пространство, вы можете удалить неиспользуемые образы контейнеров прошлых версий.

Не удаляйте контейнеры текущей установленной версии. Учитывайте, что с помощью команд ниже вы удалите все версии контейнеров, кроме указанной.

Выполните следующие шаги:

a. Создайте файл `/etc/crictl.yaml` или дополните его следующим содержимым:

```
runtime-endpoint: unix:///run/containerd/containerd.sock
image-endpoint: unix:///run/containerd/containerd.sock
timeout: 2
pull-image-on-create: false
disable-pull-on-run: false
```

b. Выполните команды:

i. Для удаления контейнеров Платформы (где `x.x.x.x` – **текущая установленная** версия Платформы):

```
crictl rmi $(crictl image | awk '($2 !~ /x.x.x.x|latest/ && $1 ~ /
docker.infowatch.ru\/platform-core-release/ && $1 !~ /flannel/) {print
$3}')
```

ii. Для удаления контейнеров продукта Data Discovery (где `x.x.x.x` – **текущая установленная** версия Data Discovery):

```
crictl rmi $(crictl image | awk '($2 !~ /x.x.x.x|latest/ && $1 ~ /
docker.infowatch.ru\/discovery/) {print $3}')
```

iii. Для удаления контейнеров Docker:

```
rm -rf /var/lib/docker
```

8. **На время обновления** для миграции данных ClickHouse понадобится **дополнительное свободное пространство** на диске помимо того, что было выделено для работы инсталлятора и перехода на новую версию. После завершения обновления дополнительное пространство будет освобождено.
Чтобы определить объем дополнительного пространства:

a. Для подключения к ClickHouse на ноде с Data Discovery в режиме `central` выполните команду:

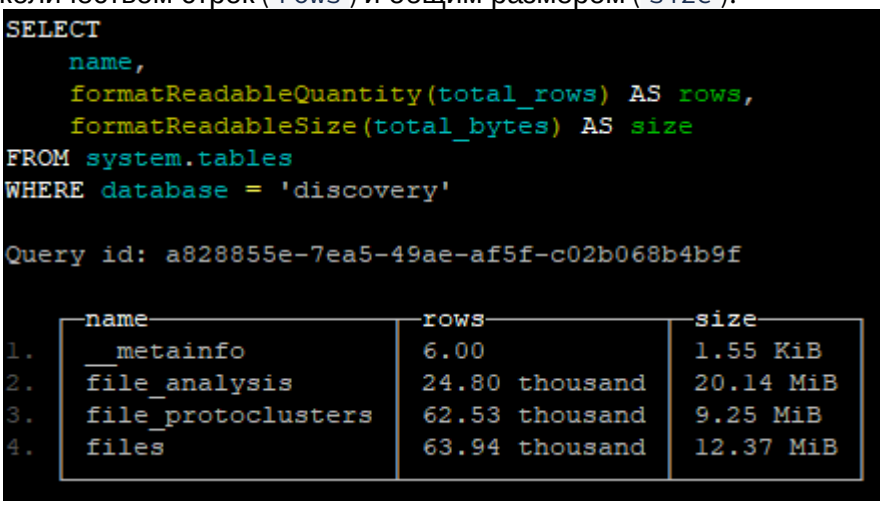
```
kubectl exec --stdin --tty -n infowatch deploy/clickhouse-central -- clickhouse-
client -d discovery --user discovery --password $(kubectl get secret discoperator-
```

```
discovery-central-db-secret -n infowatch -o jsonpath='{.data.service_ch_password}' |
base64 --decode)
```

b. Для получения размеров таблиц с данными выполните запрос:

```
SELECT
    name,
    formatReadableQuantity(total_rows) AS rows,
    formatReadableSize(total_bytes) AS size
FROM system.tables
WHERE database = 'discovery'
```

В результате выполнения запроса будут выведены названия таблиц (`name`) с количеством строк (`rows`) и общим размером (`size`).



c. Для определения нужного дополнительного пространства **определите наибольшее значение** из трех:
   - размер таблицы **`files` x 3**
   - размер таблицы **`file_analysis` x 2**
   - размер таблицы **`file_protoclusters` x 2**

   Пример:

   Размеры таблиц: `files` = 29 ГБ, `file_analysis` = 42 ГБ, `file_protoclusters` = 30 ГБ.

   В этом примере нужно выделить дополнительно 87 ГБ, так как 29 * 3 = 87 – это наибольшее значение.

> **Примечание:**
>
> Учитывайте также, что миграция данных потребует дополнительного времени при обновлении. Чем больше данных в таблицах, тем больше времени займет миграция. Например, миграция 100 млн. записей занимает ~12 минут дополнительного времени.

9. Запустите программу обновления:

```
./setup.py update
```

Важно!

При обновлении с версии 2.0 метаданные файлов, которые остались в хранилище после удаления задач сканирования, будут удалены.

Во время обновления уже установленные на сервере продукты Платформы могут быть недоступны.

10. Укажите путь для временного хранения событий Traffic Monitor перед их загрузкой в Платформу (по умолчанию: `/opt/factsldrtmdata`) и нажмите **Enter**.
11. Дождитесь завершения обновления Data Discovery. После обновления продукта инсталлятор автоматически удалит образы контейнеров прошлых версий.
12. Убедитесь, что все сервисы запущены. Для этого выполните команду:

    ```
    kubectl get pods -n infowatch
    ```

    Команда выведет список сервисов. Каждый сервис должен иметь статус `Running`.
13. Для работы Data Discovery в ОС должны быть дополнены настройки подсистемы inotify. Для этого:
    a. Проверьте наличие конфигурационного файла `/etc/sysctl.d/discovery.conf` со следующим содержимым:

    ```
    fs.inotify.max_user_instances = 32768
    fs.inotify.max_user_watches = 32768
    fs.inotify.max_queued_events = 262144
    ```

    При наличии конфигурационного файла пропустите два следующих шага инструкции.
    b. Если файл не был создан ранее скриптами инсталлятора, создайте его и добавьте содержимое из предыдущего шага.
    c. Чтобы применить новые настройки подсистемы, выполните команду:

    ```
    sysctl --load=/etc/sysctl.d/discovery.conf
    ```

14. Чтобы вывести на экран информацию о Платформе и продуктах в кластере, выполните команду:

    ```
    ./setup.py showproducts
    ```

    В выводе команды будет отражена информация по всем серверам кластера.
15. Если на сервере установлены и другие продукты, обновите их до версий, совместимых с Data Discovery 2.2 (см. *Типы установки и обновления, совместимость с другими продуктами, аппаратные и программные требования*).
16. Если Data Discovery установлен в распределенном кластере, обновите Платформу и продукты на других серверах кластера до версий, совместимых с Data Discovery 2.2.

17. Очистите кеш браузера, в котором работали до обновления продукта. Веб-интерфейс Центра расследований доступен по адресу центральной ноды Платформы.

Примечание:

Если в кластере с установленным Data Discovery 2.2 **Платформа обновлена до версии 2.5**, для повторного подключения сервиса `contentbroker` к базе данных **перезапустите** его.

Для этого:

1. На любой ноде Data Discovery выполните команду перезапуска:

```
kubectl -n infowatch rollout restart deployment/contentbroker-central
```

2. Проверьте статус сервиса с помощью команды:

```
kubectl -n infowatch rollout status deployment/contentbroker-central
```

10.6 Обновление Data Discovery с помощью сброса установки (reset)

Важно!

Чтобы обновить весь распределенный кластер Центра расследований с помощью сброса (reset) установки, действуйте в следующей последовательности:

1. Выполните **сброс на офисных** нодах с продуктами в режиме `office`.
2. Выполните **сброс на офисных** нодах с продуктами в режиме `central`.
3. Выполните **сброс центральной** ноды.
4. **Установите центральную** ноду кластера.
5. **Установите офисные** ноды с продуктами в режиме `central`.
6. **Установите офисные** ноды с продуктами в режиме `office`.

В рамках обновления кластера выполняйте сброс установки только один раз для каждой ноды.

Обновление с помощью сброса установки состоит из следующих шагов:

1. Сброс установки **с помощью нового инсталлятора**. При сбросе данные Системы не будут удалены.
2. Установка новых версий Платформы и продукта с указанием уже имеющихся папок с данными.

Примечание:

Если вы используете удаленное подключение по протоколу SSH, чтобы избежать проблем в случае разрыва соединения с сервером, рекомендуем использовать утилиту Screen. При отключении от утилиты запущенные в ней процессы не прервутся, что позволит безопасно продолжить обновление.

Чтобы обновить Data Discovery:

1. **Остановите все задачи** сканирования в веб-интерфейсе Data Discovery.
2. Подготовьте сервер к обновлению (см. *Подготовка сервера*).
3. Создайте новую директорию на диске (например, `distr_discovery`):

   ```
   mkdir distr_discovery
   ```

4. Скопируйте архив **`iw_discovery_setup_2.2.X.xxx.tar.xz`** в созданную директорию. В нашем примере будет использована версия `2.2.0.191`, соответственно, архив – `iw_discovery_setup_2.2.0.191.tar.xz`.
5. Перейдите в директорию, в которую был скопирован архив.
6. Распакуйте архив с дистрибутивом в эту директорию:

   ```
   tar -xvf iw_discovery_setup_2.2.0.191.tar.xz
   ```

7. Если Data Discovery и Prediction установлены совместно или планируется совместно их использовать, остановите расчет данных Prediction на время обновления Data Discovery (подробнее см. Обязательные действия для обновления продуктов на Платформе, которые установлены совместно с Prediction).
8. После сброса при установке новой версии присвойте нодам кластера те же лейблы, которые были до обновления.
   Чтобы посмотреть список лейблов в кластере, выполните команду:

   ```
   kubectl get node -n infowatch -o jsonpath="{range .items[*]}
   {.metadata.annotations.flannel\.alpha\.coreos\.com/public\-ip}
   {.metadata.labels.kubernetes\.io/hostname} {.metadata.labels.iwnode}{'\n'}{end}"
   ```

9. Чтобы в новой версии сохранить настройки зон и серверов, перед сбросом установки подключитесь к ноде, на которой установлен Data Discovery в режиме **`central`** и **остановите сервис Discoperator** с помощью команды:

   ```
   kubectl scale --replicas=0 deployment discoperator-<лейбл_ноды> -n infowatch
   ```

   Пример для Data Discovery в режиме central на офисной ноде с лейблом lv426:

   ```
   kubectl scale --replicas=0 deployment discoperator-lv426 -n infowatch
   ```

10. До сброса установки запомните или сохраните пути к текущим директориям хранения данных. Получить список директорий можно с помощью команды:

```
cat /var/lib/iwplatform/data/volumes.txt
```

11. Выполните **сброс** текущей установки **с помощью инсталлятора новой версии**:

```
./setup.py reset
```

> **Важно!**
>
> Если на сервере установлены другие продукты на базе Платформы, установки этих продуктов будут сброшены. При сбросе установки на центральной ноде продукты на офисных нодах перестанут функционировать.
>
> После установки новой версии Data Discovery повторно установите продукты Платформы, указав директории с имеющимися данными.
>
> Если на этом сервере вы **уже обновили другой продукт** до совместимой версии, **не выполняйте повторный сброс** при обновлении Data Discovery, так как требуемая версия Платформы уже установлена.
>
> Если продукт установлен **на офисной ноде в режиме** `central`, после сброса установки у добавленных ролей пользователя будет **сброшен доступ к продуктам**.

12. Дождитесь окончания процесса сброса установки Платформы и продукта.
13. Если вы обновляетесь с переходом на другой сервер или на новую ОС, скопируйте на другой раздел или на новый сервер директории с данными, которые были указаны при установке.

> **Примечание:**
>
> В **Data Discovery 2.0 и более новых** – по умолчанию по пути `/opt`.
> Обязательно копируйте директории с сохранением прав и владельцев. Для копирования используйте стандартные средства, например утилиту `rsync`.

14. Убедитесь, что в разделе с директорией `/var/lib/containerd` доступно не менее 40 ГБ дискового пространства.
Чтобы проверить доступное дисковое пространство, используйте команду:

```
df -h /var/lib/containerd
```

15. Чтобы освободить дисковое пространство, вы можете удалить директорию с образами контейнеров прошлых версий `/var/lib/containerd`.
Для этого выполните команду:

```
rm -rf /var/lib/containerd
```

16. Перезагрузите сервер с помощью команды:

```
reboot
```

17. Установите новую версию Data Discovery (см. *Установка Data Discovery*).

> **Важно!**
>
> Когда инсталлятор потребует указать директории для размещения данных, указывайте директории с данными прошлой версии и выбирайте "**n**" при ответе на вопрос "Очищать ли папки перед установкой?".
>
> Если вы ранее обновлялись **с версии 1.10 и более ранних**, учитывайте, что с версии 2.0 пути установки по умолчанию изменились. Если ранее вы использовали пути по умолчанию, при установке Data Discovery 2.2 и более новых по умолчанию инсталлятор не обнаружит данные прошлой версии и создаст директории по пути `/opt`. Поэтому переместите данные в `/opt` или укажите старые директории при установке новой версии.
>
> На всех нодах кластера с Data Discovery должен быть одинаковый путь к хранилищу просканированных файлов (по умолчанию: `/opt/disctps`).

18. Если до сброса на сервере были установлены и другие продукты, установите версии продуктов, совместимые с Data Discovery 2.2 (см. Типы установки и обновления, совместимость с другими продуктами, аппаратные и программные требования).
19. Если Data Discovery установлен в распределенном кластере, обновите Платформу и продукты на других серверах кластера до версий, совместимых с Data Discovery.
20. Если после обновления вам нужно освободить место на диске, вы можете удалить неиспользуемые контейнеры docker или containerd. Подробнее, см. в статье Как удалить неиспользуемые контейнеры docker и containerd?.
21. Очистите кеш браузера, в котором работали до обновления продукта. Веб-интерфейс Центра расследований доступен по адресу центральной ноды Платформы.

> **Важно!**
>
> Если Data Discovery установлен **на офисной ноде в режиме** `central`, в результате сброса прошлой установки у добавленных ролей пользователя **сброшен доступ к продуктам**. После обновления перейдите в **Настройки** и заново настройте доступ к продуктам в этих ролях (о ролях см. *InfoWatch Data Discovery. Руководство администратора. Роли*).

> **Примечание:**
>
> Если в кластере с установленным Data Discovery 2.2 **обновлена Платформа до версии 2.5**, для повторного подключения сервиса `contentbroker` к базе данных **перезапустите** его.
>
> **Для этого:**
>
> 1. На любой ноде Data Discovery выполните команду перезапуска:

```
kubectl -n infowatch rollout restart deployment/contentbroker-central
```

2. Проверьте статус сервиса с помощью команды:

```
kubectl -n infowatch rollout status deployment/contentbroker-central
```

10.7 Переустановка Data Discovery в другой кластер

Если Data Discovery и другие продукты установлены в разных кластерах, вы можете установить новую версию Data Discovery в кластер к другому продукту. В результате вы сможете работать с Data Discovery и другими продуктами в одном веб-интерфейсе Центра расследований, а продукты будут иметь общие интеграции и возможность обмена данными.

Важно!

Учитывайте, что при переустановке Data Discovery в другой кластер **не будут перенесены**:

- пользователи;
- роли;
- интеграции;
- расписания запуска задач сканирования.

В результате в новом кластере Центра расследований останутся только пользователи, роли и интеграции, которые были до добавления Data Discovery.

После переустановки **настройте заново расписания** запуска задач сканирования.

Чтобы переустановить Data Discovery в другой кластер:

1. В кластере, в который хотите установить новую версию Data Discovery, обновите продукты до версий, совместимых с Data Discovery 2.2 (см. "Типы установки и обновления, совместимость с другими продуктами, аппаратные и программные требования").
2. На сервере Data Discovery выполните шаги по подготовке сервера к обновлению (см. "Подготовка сервера").
3. Создайте новую директорию на диске (например, `distr_discovery`):

```
mkdir distr_discovery
```

4. Скопируйте архив **`iw_discovery_setup_2.2.X.xxx.tar.xz`** в созданную директорию. В нашем примере будет использована версия `2.2.0.191`, соответственно, архив - `iw_discovery_setup_2.2.0.191.tar.xz`.
5. Перейдите в директорию, в которую был скопирован архив.
6. Распакуйте архив с дистрибутивом в эту директорию:

```
tar -xvf iw_discovery_setup_2.2.0.191.tar.xz
```

7. Выполните **сброс** текущей установки **с помощью инсталлятора новой версии**:

```
./setup.py reset
```

8. Дождитесь завершения процесса сброса установки Платформы и продукта.
9. Убедитесь, что в разделе с директорией `/var/lib/containerd` доступно не менее 40 ГБ дискового пространства.
   Чтобы проверить доступное дисковое пространство, используйте команду:

```
df -h /var/lib/containerd
```

10. Чтобы освободить дисковое пространство, вы можете удалить директорию с образами контейнеров прошлых версий `/var/lib/containerd`.
    Для этого выполните команду:

```
rm -rf /var/lib/containerd
```

11. Перезагрузите сервер с помощью команды:

```
reboot
```

12. Установите **офисную ноду Платформы с новой версией Data Discovery** (см. "Установка Data Discovery").

> **Важно!**
>
> Используйте команду для установки на офисную ноду Платформы
>
> ```
> ./setup.py install --nodemode=office --productmode=central
> ```
>
> Когда инсталлятор запросит, укажите адрес и токен центральной Платформы нового кластера.
>
> Когда инсталлятор потребует указать директории для размещения данных, указывайте директории с данными прошлой версии и выбирайте «**n**» при ответе на вопрос «Очищать ли папки перед установкой?».
>
> Если вы обновлялись ранее **с версии 1.10 и более ранних**, учитывайте, что с версии 2.0 пути установки по умолчанию изменились. Если ранее вы использовали пути по умолчанию, при установке Data Discovery 2.2 и более новых по умолчанию инсталлятор не обнаружит данные прошлой версии и создаст директории по пути `/opt`. Поэтому переместите данные в `/opt` или укажите старые директории при установке новой версии.

На всех нодах кластера с Data Discovery должен быть одинаковый путь к хранилищу просканированных файлов (по умолчанию: `/opt/disctps`).

13. Если после обновления вам необходимо освободить место на диске, вы можете удалить неиспользуемые контейнеры docker или containerd. Подробнее, см. в статье "Как удалить неиспользуемые контейнеры docker и containerd?".
14. Очистите кеш браузера, в котором работали до обновления продукта.
Веб-интерфейс Центра расследований доступен по адресу центральной ноды Платформы.

10.8 Прямое обновление Модуля контентного анализа

Чтобы обновить Модуль контентного анализа:

1. **Остановите все задачи** сканирования в веб-интерфейсе Data Discovery.
2. Выполните шаги по подготовке сервера к установке Data Discovery (см. "Подготовка сервера").
3. Создайте новую директорию на сервере (например, `distr_tech`):

```
mkdir distr_tech
```

4. Скопируйте архив **`iw_tm_setup_7.13.0.xxx.tar.xz`** в созданную директорию.
В нашем примере будет использована версия `7.13.0.167`, соответственно, архив - **`iw_tm_setup`** `_7.13.0.167.tar.xz`.
5. Перейдите в директорию, в которую был скопирован архив.
6. Распакуйте архив с дистрибутивом в эту директорию:

```
tar -xvf iw_tm_setup_7.13.0.167.tar.xz
```

7. Запустите программу обновления:

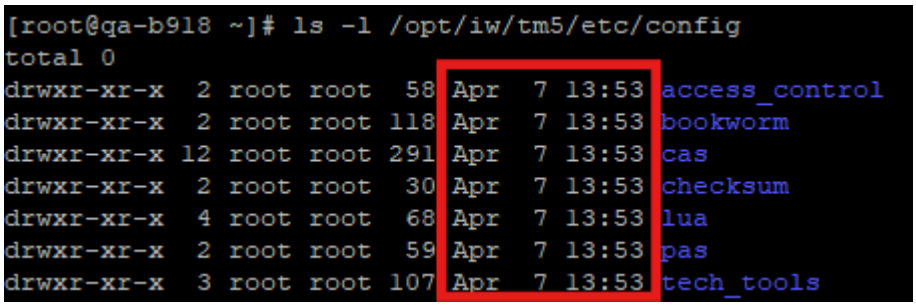
```
./setup.py update
```

8. Укажите путь для временного хранения событий Traffic Monitor перед их загрузкой в Платформу (по умолчанию: `/opt/factsldrtmdata`) и нажмите **Enter**.
9. Дождитесь завершения обновления.
10. Чтобы сервис `updater` гарантированно загрузил последнюю версию конфигурации Traffic Monitor, рекомендуем удалить прошлую конфигурацию с помощью команды:

```
rm -rf /opt/iw/tm5/tmp
rm -rf /opt/iw/tm5/etc/config
```

11. Через некоторое время после удаления сервис `updater` загрузит последнюю конфигурацию Traffic Monitor.
Чтобы проверить даты создания или изменения с конфигурацией, выполните команду:

```
ls -l /opt/iw/tm5/etc/config
```



12. После обновления модуля вы можете запустить ранее остановленные задачи сканирования в веб-интерфейсе Data Discovery.

11 Удаление Data Discovery

Чтобы удалить Data Discovery:

1. **Остановите все задачи** сканирования в веб-интерфейсе Data Discovery.
2. Перейдите в директорию на сервере, в которую был распакован архив с дистрибутивом при установке Data Discovery.
3. Выполните команду:

   ```
   ./setup.py remove
   ```

4. Дождитесь завершения удаления. Программа удалит компоненты Системы, включая внутренние пароли продукта и связанные данные. Данные Data Discovery удалены не будут. При необходимости вы можете удалить данные вручную из директорий, указанных при установке.

Примечание:

Если вы удалили Data Discovery, установленный в режиме central, компоненты Data Discovery, установленные в режиме office на других нодах, не будут функционировать и не будут удалены автоматически. Удалите их вручную.

Примечание:

В случае распределенной установки, если продукт установлен на офисной ноде и на данной ноде отсутствуют другие продукты, то после удаления продукта на ноде также рекомендуется удалить элементы платформы. Для этого после выполнения команды `./setup.py remove`, также выполните команду `./setup.py reset`.

Если продукт установлен на центральной ноде или если на офисной ноде также установлены другие продукты, выполнять `./setup.py reset` категорически не рекомендуется.