# Kaspersky MLAD: ?????? ?????????? ??????

## 1. ???????? ??????

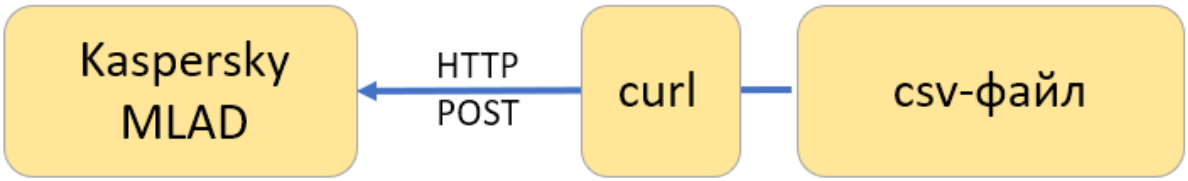
После создания дерева активов в Kaspersky MLAD, для обучения моделей и проведения исторического инференса потребуется достаточная выборка исторических данных. Данные можно накопить в процессе работы программы при наличии подключения к источнику данных, например, по протоколу OPC UA. Но если у вас уже есть накопленная выборка, то наиболее быстрым и предпочтительным способом является импорт данных при помощи встроенного HTTP-коннектора.

Данный документ описывает процесс настройки HTTP-коннектора и импорта данных из csv-файла при помощи POST-запросов.

Для передачи csv-файла в MLAD можно использовать стандартную утилиту, например, curl. В данном документа также описан процесс импорта при помощи инструмента csv2http (предоставляется по запросу).

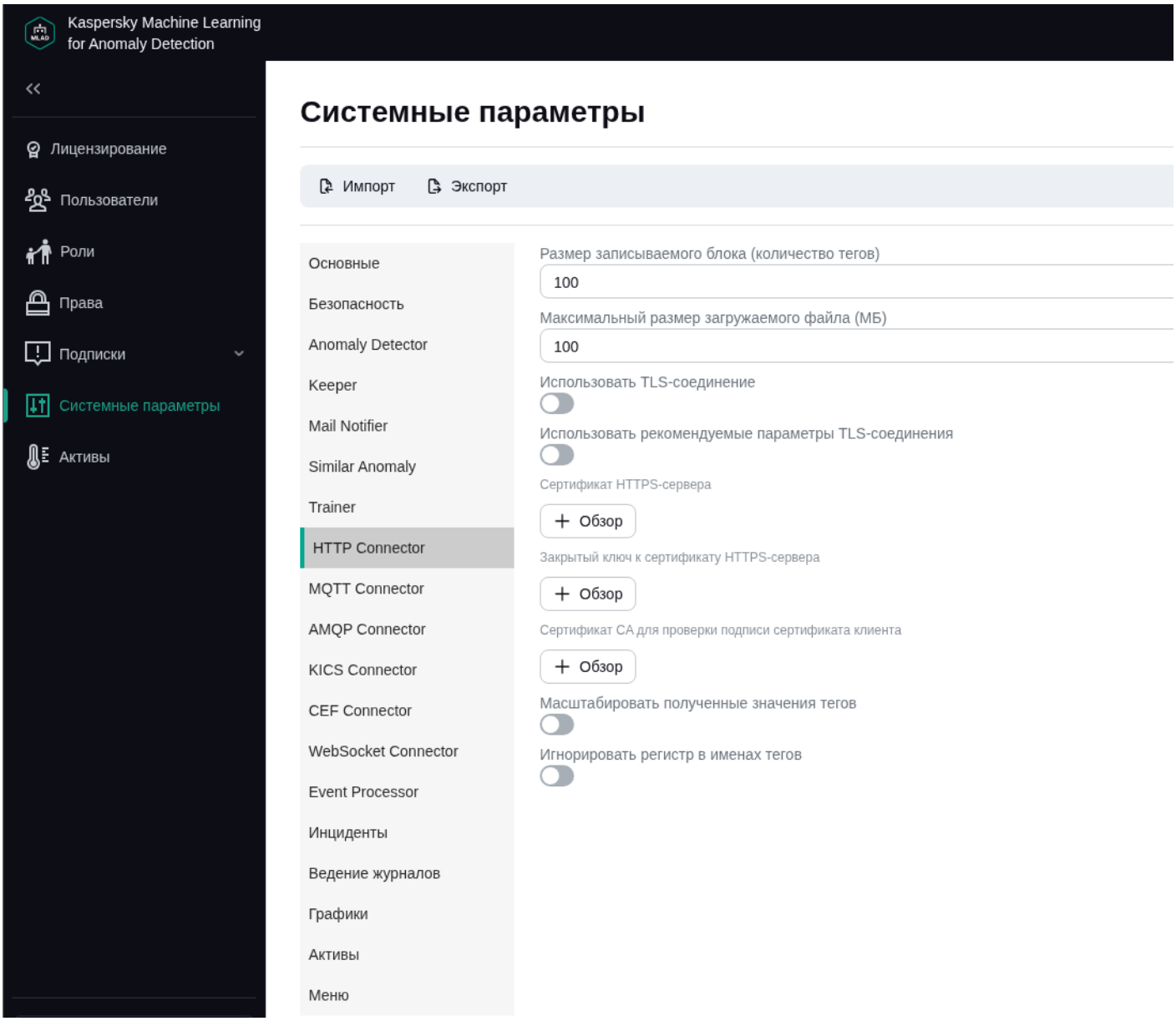
Для работы потребуется операционная система Linux (в данном документе используется Debian 13).

Ниже приведена упрощенная схема подключения, используемая в описываемой в данном документе интеграции.

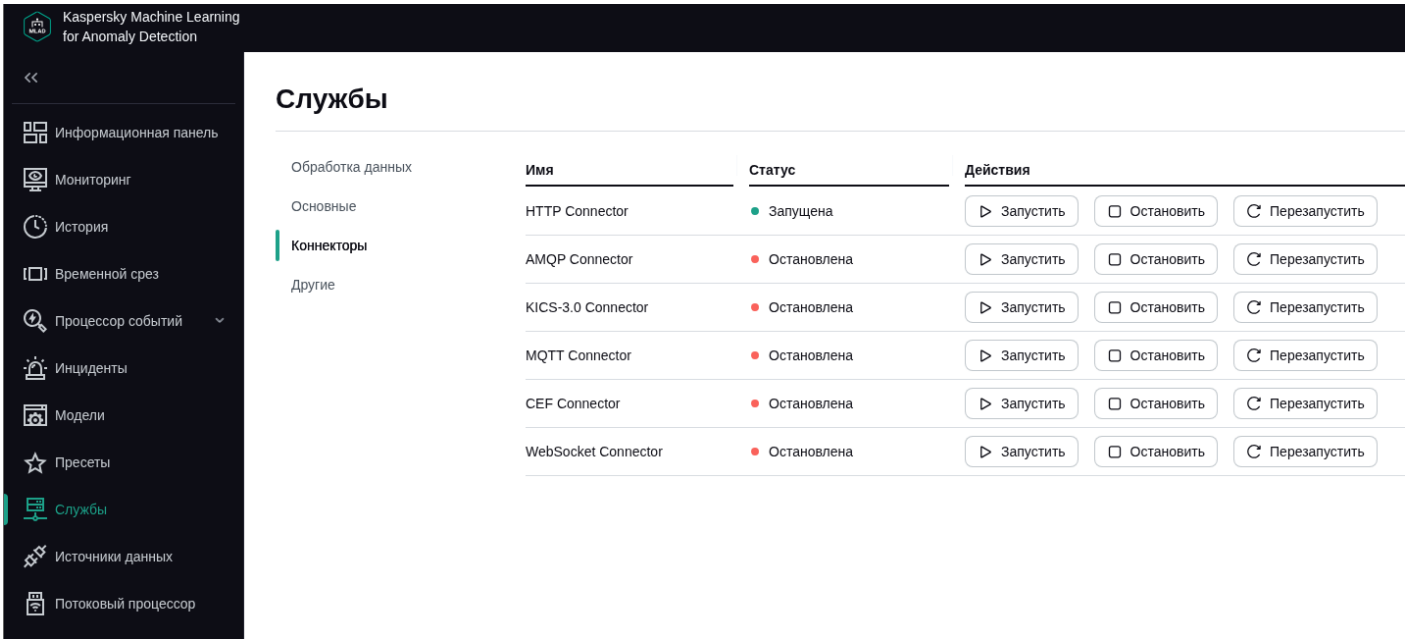



## 2. ????????? ??????

Первым шагом необходимо в MLAD перейти в **Меню администратора** -> **Системные параметры -> HTTP Connector.** Здесь рекомендуется отключить использование TLS-соединения, чтобы упростить процесс настройки Kaspersky MLAD и ввод его в эксплуатацию. Если в дальнейшем по соображениям безопасности потребуется перейти на HTTPS-соединение, здесь же можно включить TLS-обратно и подгрузить необходимые сертификаты.




После этого необходимо запустить HTTP-коннектор в **Службах**.



Теперь мы можем отправлять данные в MLAD в виде csv файлов через POST-запросы. Более подробно настройка HTTP-коннектора и загрузка данных описаны в руководстве пользователя.

Csv-файл должен иметь следующую структуру:

> **метка времени; имя тега; значение тега**

Метка времени должна быть в формате **%Y-%m-%dT%H:%M:%S** (или **%Y-%m-%d %H:%M:%S**).

Имя тега должно соответствовать имени тега в активах MLAD.

Если значение тега содержит дробную часть, используйте точку при отделении целой и дробной частей.

Файл не должен содержать заголовочную строку (строку с названием столбцов).

Пример содержимого файла:

```
2026-01-12 08:00:00;Tag1;11.10
2026-01-12 08:00:00;Tag2;37
2026-01-12 11:00:00;Tag1;12.10
2026-01-12 11:00:00;Tag2;38
```

Пример отправки CSV-файла в MLAD через curl по протоколу методом `POST` на порт 4999 сервера Kaspersky MLAD:

```
curl -F "file=@<имя файла>.csv" -X POST "http://<IP-адрес или доменное имя сервера Kaspersky MLAD>:4999/upload"
```

Не рекомендуется загружать файлы размером более 10-20 МБ. Это может привести к потере отдельных точек данных при загрузке. Если вы хотите загрузить большой csv-файл (что весьма вероятно), необходимо разбить его на более мелкие файлы и загрузить их по очереди с небольшой задержкой. Можно использовать, к примеру, такой скрипт (запускается из директории, в которой находятся csv-файлы:

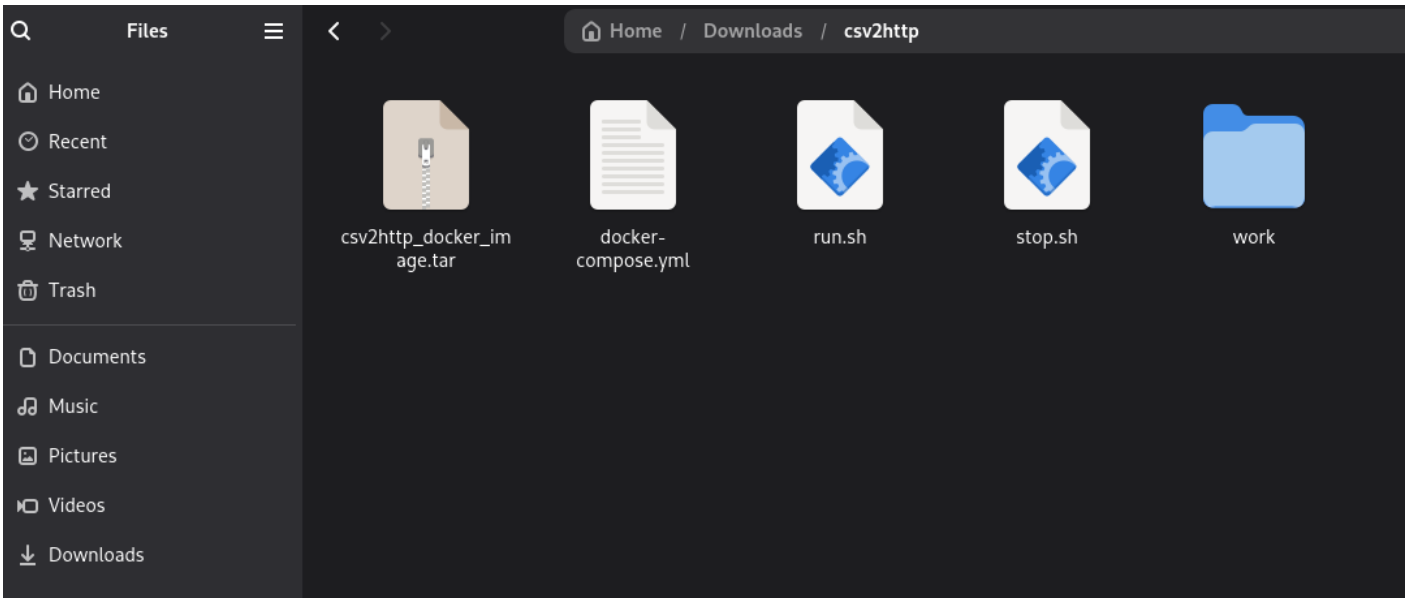
```
#!/bin/bash
# Укажите адрес вашего сервера
SERVER_URL="localhost:4999/upload"
# Перебор всех csv-файлов в директории
for file in *.csv; do
```

```
    # Завершение скрипта, если csv-файлы не найдены
    [ -e "$file" ] || { echo "No CSV files found."; exit 0; }
    echo "Uploading $file to $SERVER_URL ..."
    # Отправка файла в MLAD через POST-запрос
    curl -X POST -F "file=@$file" "$SERVER_URL"
    sleep 0.5
    echo
done
```

**Kaspersky MLAD будет добавлять часовой пояс, указанный в настройках программы к меткам времени, указанным в csv-файле. Имейте это в виду при подготовке файла для импорта.**

Данный метод загрузки данных с одной стороны является достаточно простым и не требует использования дополнительных инструментов, но с другой стороны накладывает жесткие ограничения на формат и размер загружаемого csv-файла, и требует редактирования метки времени.

Альтернативно можно использовать утилиту csv2http (предоставляется по запросу), которая автоматически подготавливает файл для загрузки в соответствии с настройками и загружает его. Для ее запуска на машине потребуется установленный docker и docker-compose (можно запускать на машине с MLAD).




Перед началом использования утилиты необходимо загрузить локальный docker-образ (укажите корректный путь до tar-файла):

```
docker load -i /home/mlad-user/Downloads/csv2http/csv2http_docker_image.tar
```

Утилита запускается при помощи скрипта в корневой директории **./run.sh**
Для остановки утилиты используется скрипт **./stop.sh**

Вы можете изменить настройки программы в файле **csv2http.yml**. После изменения параметров необходимо остановить и заново запустить утилиту. Ключевые настройки, на которые нужно обратить внимание:

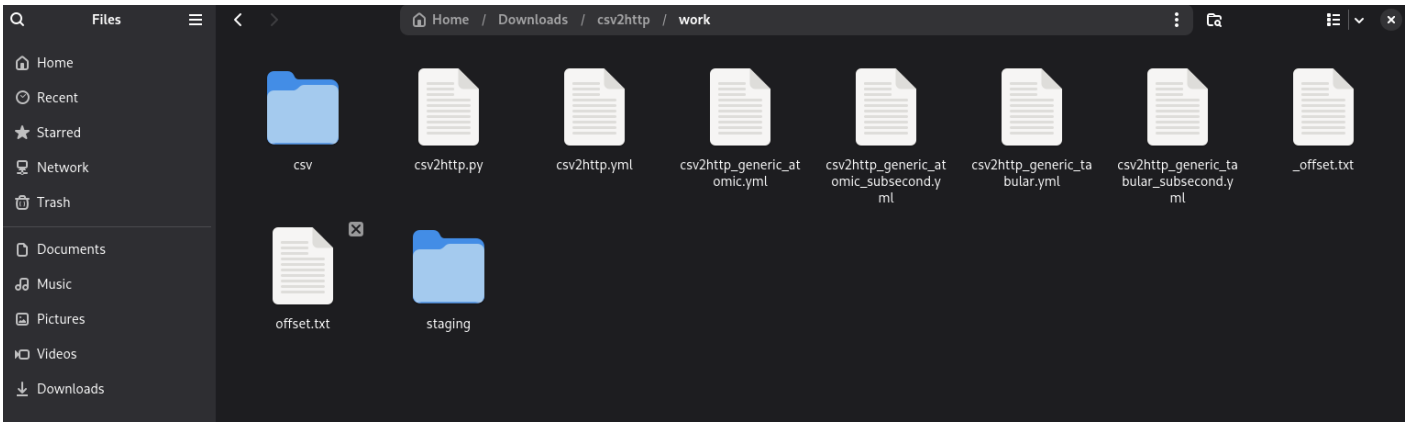
**hours_from_GMT** - количество часов, указанных в данном параметре должно соответствовать часовому поясу, заданному в настройках MLAD.
**tabular** - если значение данного параметра равно **false**, то данные формат csv должен соответствовать формату, описанному выше (**метка времени; имя тега; значение тега)**, если значение **true**, то файл должен иметь табличную форму, как показано в следующем примере:

| timestamp | param1 | param2 | param3 | param4 |
|---|---|---|---|---|
| 2025-12-28T00:00:00 | 70 | 53 | 49 | 66 |
| 2025-12-28T00:01:00 | 71 | 52 | 49 | 66 |
| 2025-12-28T00:02:00 | 70 | 53 | 49 | 66 |
| 2025-12-28T00:03:00 | 71 | 53 | 49 | 66 |
| 2025-12-28T00:04:00 | 70 | 52 | 49 | 66 |
| 2025-12-28T00:05:00 | 70 | 52 | 49 | 66 |
| 2025-12-28T00:06:00 | 70 | 53 | 49 | 66 |

**timestamp_position** - номер столбца, содержащего метки времени (нумерация начинается с 0). Релевантно только в табличном режиме, при переключении в стандартный режим закомментируйте эту строку.
**tag_position** - номер столбца, содержащего имена тегов (нумерация начинается с 0). Релевантно только в стандартном режиме, при переключении в табличный режим закомментируйте эту строку.
**value_position** - номер столбца, содержащего значения тегов (нумерация начинается с 0). Релевантно только в стандартном режиме, при переключении в табличный режим закомментируйте эту строку.
**skip_lines** - количество строк в начале файла, которое вы хотите пропустить при обработке. Полезно, если файл содержит дополнительную информацию в начале.
**has_header** - значение **true**, если используется табличный режим. При переключении в стандартный режим укажите **false**.
**timestamp_format** - укажите формат меток времени, используемый в вашем файле, например, **'%Y-%m-%dT%H:%M:%S'** или **'%d-%b-%y %H:%M:%S'**

**field_separator** - укажите разделитель, который используется в вашем csv-файле.
**url** - укажите ip или доменное имя вашего сервера MLAD. Если MLAD расположен на этой же машине, используйте значение по умолчанию '**http://host.docker.internal:4999/upload**'.

Поместите ваш csv-файл в директорию ./work/csv:




Утилита периодически сканирует данную директорию на предмет новых файлов и обрабатывает их. Обработанные фалы помещаются в директорию **./work/staging**, после чего загружаются в MLAD стандартными средствами и удаляются. Если директория staging больше не содержит файлов, значит загрузка завершена.

Вы можете подключиться к журналу работы утилиты для траблшутинга загрузки. Для этого используйте следующую команду:

> **docker logs csv2http-smb2mlad1**
> используйте флаг **-f**, если вы хотите отслеживать новые записи в журнале

> Если загрузка файлов завершилась досрочно (например, скриптом **./stop.sh**), и вы хотите загрузить файлы заново, то вам необходимо удалить все файлы из директорий **./work/csv** и **./work/staging**, а также удалить файл **./work/offset.txt**.

---

Revision #7
Created 4 June 2026 09:35:05 by Эльдар Юсуфов
Updated 4 June 2026 15:14:51 by Эльдар Юсуфов