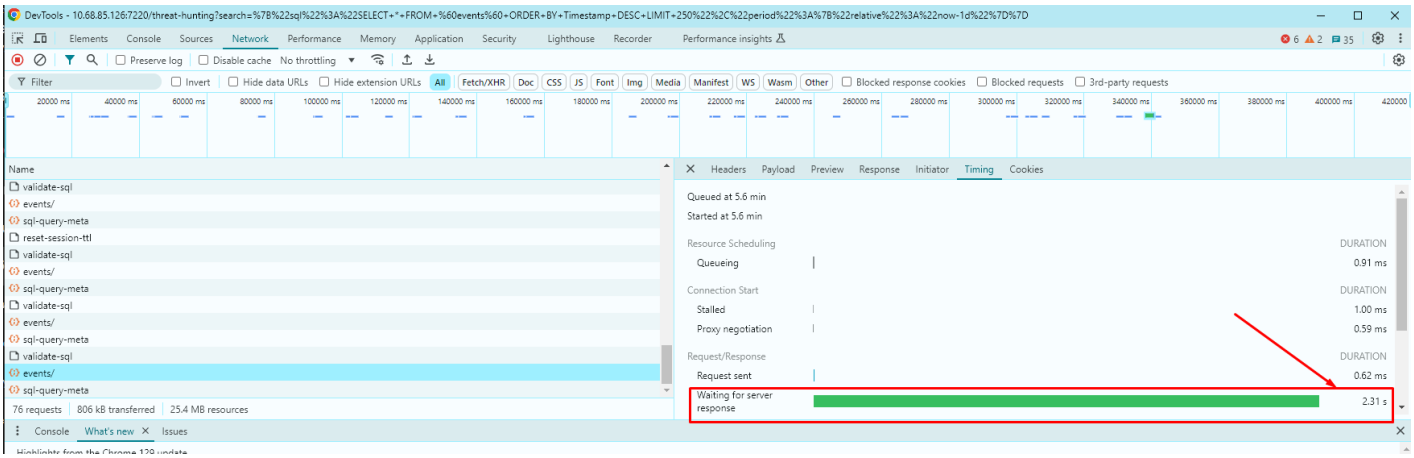
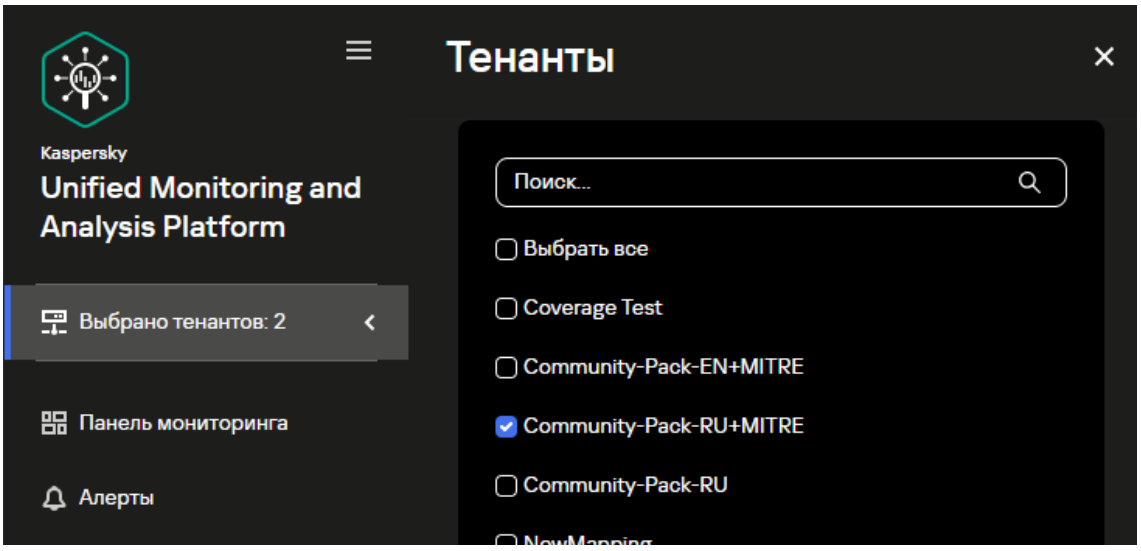
Создание оптимизированных запросов

Длительность выполнения запроса можно посмотреть при нажатии F12 (режим разработчика) в браузере Chrome после нажатия кнопки поиска:



При создании запросов всегда явно выбирайте необходимый тенант с нужными событиями:



Если известны конкретные поля для поиска, то можно использовать более глубокий поиск по событиям. Старайтесть **НЕ** использовать поиск по всем полям:

```
SELECT *
```

При поисках не увеличивайте LIMIT по умолчанию более чем 10000, все данные по поиску буферизируются в браузере.

Примеры оптимизации поисковых запросов

Пример 1:

```
SELECT SourceAddress, DestinationUserName, DeviceEventClassID
FROM `events`
WHERE DeviceProduct = 'Windows'
AND NOT endsWith(SourceUserName,'$') AND SourceUserName != 'vasya'
AND (startsWith(SourceUserName,'adm-') OR startsWith(DestinationUserName, 'adm_'))
AND NOT inSubnet(SourceAddress, '10.10.10.150/24') AND NOT inSubnet(SourceAddress, '192.168.20.180/16')
GROUP BY SourceAddress, DestinationUserName, eviceEventClassID
```

Оптимизированный запрос:

```
SELECT SourceAddress, DestinationUserName, DeviceEventClassID
FROM `events`
WHERE DeviceProduct = 'Windows'
AND SourceUserName NOT LIKE '%$' AND SourceUserName != 'vasya'
AND (SourceUserName LIKE 'adm-%' OR DestinationUserName LIKE 'adm_%')
AND NOT inSubnet(SourceAddress, '10.10.10.150/24') AND NOT inSubnet(SourceAddress, '192.168.20.180/16')
GROUP BY SourceAddress, DestinationUserName, eviceEventClassID
```

Или можно таже попробовать отказаться от функции inSubnet:

```
SELECT SourceAddress, DestinationUserName, DeviceEventClassID
FROM `events`
WHERE DeviceProduct = 'Windows'
AND SourceUserName NOT LIKE '%$' AND SourceUserName != 'vasya'
AND (SourceUserName LIKE 'adm-%' OR DestinationUserName LIKE 'adm_%')
AND NOT (SourceAddress LIKE '10.10.10.%' OR SourceAddress LIKE '192.168.%.%')
GROUP BY SourceAddress, DestinationUserName, eviceEventClassID
```

Принцип:

- Уменьшение количества условий: Объединяйте условия, где это возможно, чтобы сократить количество операций.
- Использование подзапросов или CTE: Если набор данных большой, рассмотрите возможность использования подзапросов или общих табличных выражений (CTE), чтобы разбить запрос на более мелкие, более управляемые части.
- Избегание функций в предложении WHERE: Такие функции, как endsWith, startsWith и inSubnet, могут быть узкими местами производительности. Если возможно, попробуйте переписать условия исравнить скорость выполнения.

Запросы к системным таблицам

У ClickHouse есть системные таблицы (https://clickhouse.com/docs/en/operations/system-tables).

Одна из них - query_log (https://clickhouse.com/docs/en/operations/system-tables/query_log), содержит интересную информацию о запросах.

Чтобы обратиться к данным таблицами можно использовать клиент ClickHouse, который располагается по пути `/opt/kaspersky/kuma/clickhouse/bin/client.sh`

Можно заходить в консоль ClickHouse с помощью данного клиента, а также выполнять запросы через аргументы клиента, а также формировать запросы с помощью curl. В примерах ниже будет рассмотрен именно первый способ.

Вывод запросов с сортировкой по длительности

В примере ниже представлен запрос с сортировкой по длительности выполнения запросов за конкретный день

```
SELECT *
FROM system.query_log
WHERE query_kind = 'Select' AND current_database = 'kuma' AND type = 'QueryFinish' AND event_date =
'2024-01-22'
AND query_duration_ms != 0
ORDER BY query_duration_ms DESC LIMIT 10
```

Вывод запросов с сортировкой по использованной памяти

```
SELECT
query_start_time,
    query,
query_duration_ms,
    formatReadableSize(memory_usage)
FROM system.query_log
WHERE current_database = 'kuma' AND query_kind != 'Insert' AND type = 'QueryFinish'
ORDER BY memory_usage DESC
LIMIT 10 \G;
```

Вывод запросов, завершившихся ошибкой

```
SELECT *
FROM system.query_log
WHERE type IN ('ExceptionWhileProcessing', 'ExceptionBeforeStart') AND current_database = 'kuma'
LIMIT 10 \G;
```

Revision #4
Created 30 September 2024 12:49:00 by Boris Rzr
Updated 11 November 2024 11:42:11 by Boris Rzr