VB-Trend 2019

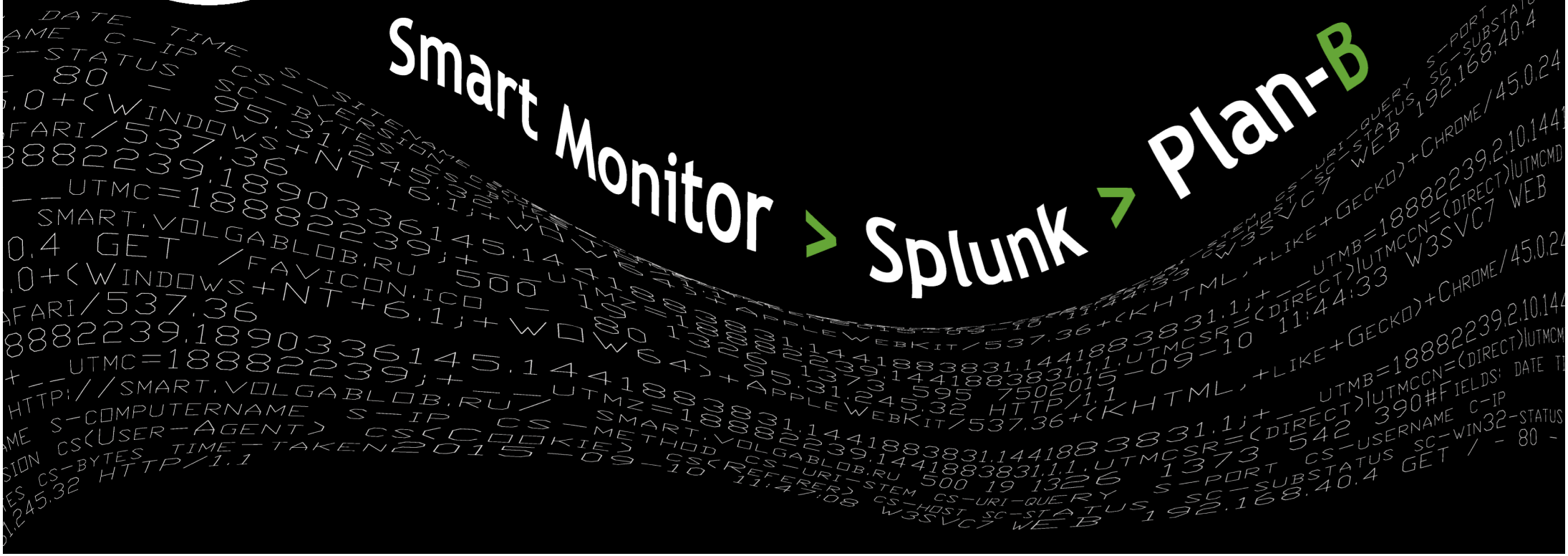
plan>b

VolgaBlob

Нагрузочное тестирование Smart Monitor

Splunk vs Elastic Stack

Smart Monitor > Splunk > Plan-B

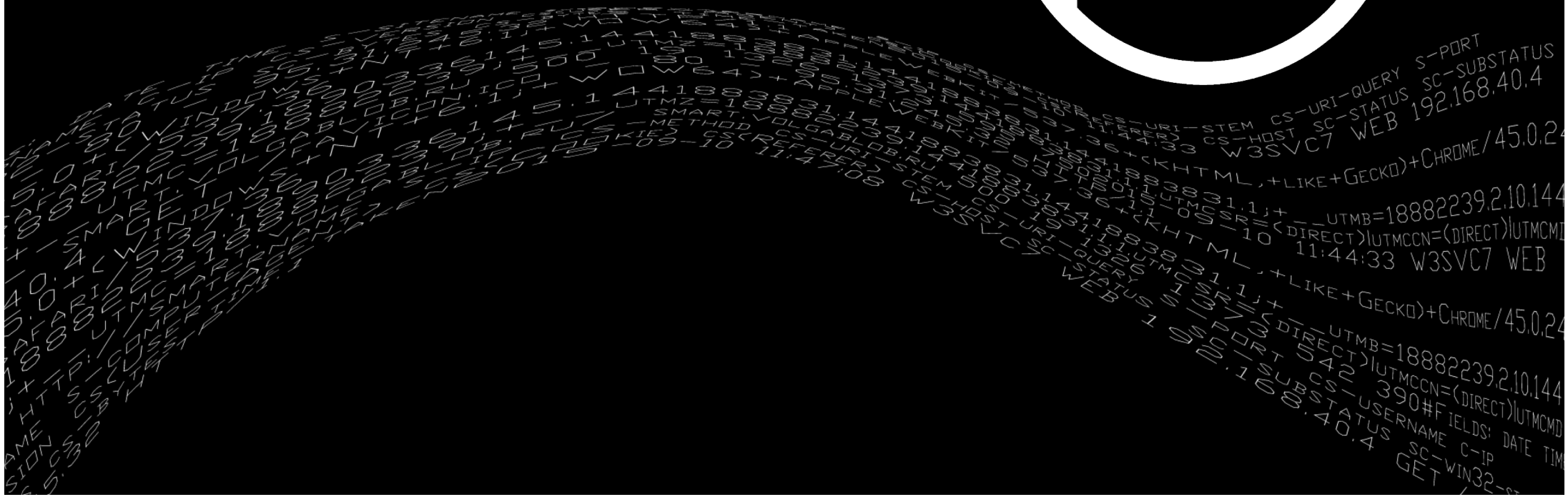


Что и где тестируем?

VB-Trend 2019

plan>b

VolgaBlob



Наборы данных для тестов

VB-Trend 2019 plan>b VolgaBlob

- Датасет №1: логи удаленного доступа - 119 ГБ
- Датасет №2: логи аутентификации - 524 МБ

- Разбор событий посредством регулярных выражений
- Формат данных:

```
{ '_time': '2019-09-28T04:02:16.000+03:00', 'user': 'khitrov.i',
'department_group.name': 'IT Security', 'service': 'Jira',
'src_ip': '10.10.64.13', 'action': 'OK' }
```

Тестовые стенды

- Splunk Enterprise
- Universal Forwarder
- 24 ядра CPU
- 64 Гбайт ОЗУ
- 600 Гбайт SAS 15k / SSD

- Elasticsearch OSS
- Filebeat OSS
- 24 ядра CPU
- 64 Гбайт ОЗУ
- 600 Гбайт SAS 15k / SSD

Инструментарий

SAR и cURL

системные показатели и время поиска

Elasticsearch

информация индексах

Splunk

информация индексах

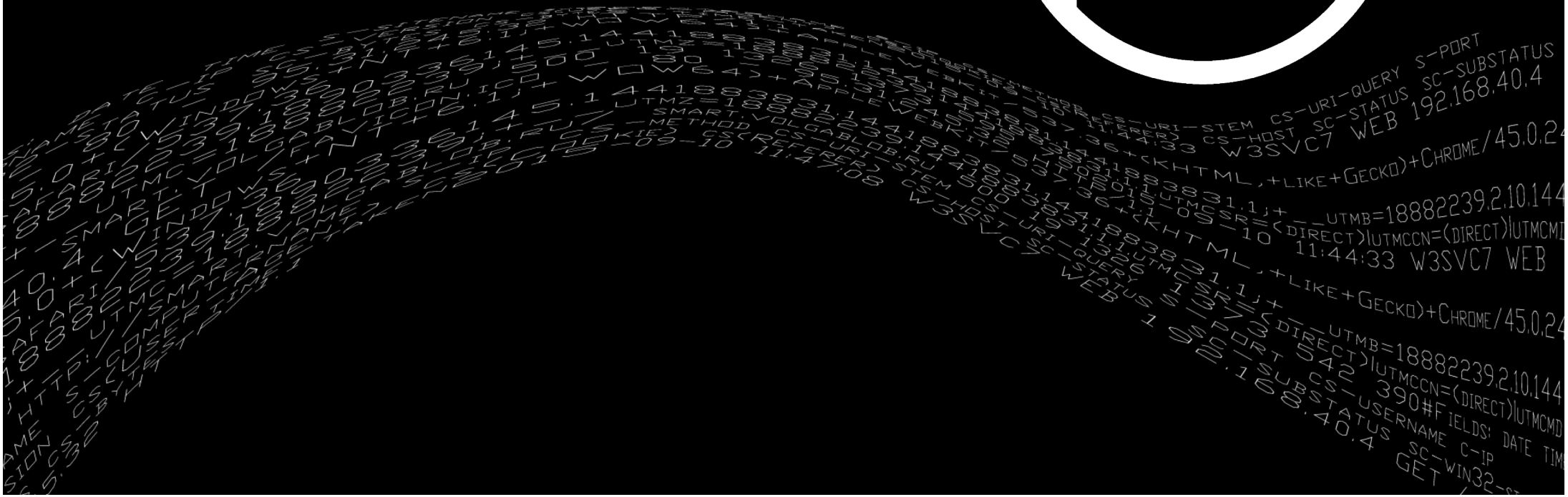
Загрузка данных

Тестируем процесс индексирования

VB-Trend 2019

plan>b

VolgaBlob



Условия тестирования

VB-Trend 2019 plan>b VolgaBlob

- Проверяемые варианты:
  - Splunk в режиме извлечения полей во время поиска
  - Splunk в режиме извлечения полей во время индексирования
  - Elastic Stack
- Тесты индексирования – затратная акция, один тест не повторялся несколько раз
- Измеряемые метрики:
  - общее время на загрузку данных
  - средний Indexing rate в процессе (Eps, KBps)
  - итоговый размер индекса

Тест на стандартных настройках

	Splunk> поля в search-time	Splunk> поля в index-time	Elastic Stack
Время индексирования	02:45:57	05:07:35	24:53:23
Indexing rate (kEps)	39.7	21.4	4.4
Indexing rate (MBps)	12.5	6.7	1.3
Размер индекса, ГБ	22.01	31.56	123.4

Комментарий: очень грустная ситуация для Elastic Stack ☹

Подкрутим настройки

VB-Trend 2019 plan>b VolgaBlob

- Что меняем в Elastic Stack?
  - маппинг и настройки индексов в Elasticsearch
  - параметры отправки и очереди в Filebeat
- Что меняем в Splunk?
  - алгоритмы компрессии и структуру индексов
  - количество конвейеров индексирования
- Что получаем?
  - индексы стали меньше
  - скорость загрузки вырастет
- Какой ценой?
  - значительно возрастает потребление ресурсов

Нужно подбирать под каждый кейс

Не рекомендуется больше 2!

Тест после перенастройки

VB-Trend 2019 plan>b VolgaBlob

	Splunk> поля в search-time	Splunk> поля в index-time	Elastic Stack
Время индексирования	00:39:04	01:14:19	04:35:22
Indexing rate (kEps)	168.9	88.8	23.9
Indexing rate (MBps)	53.2	27.9	7.5
Размер индекса, ГБ	15.87	27.76	60.8

Комментарий: общая тенденция сохраняется

Улучшение показателей Elastic

- Время индексирования: **на 80% меньше**
- Пропускная способность: **на 440% выше**
- Занимаемый объем: **на 50% меньше**

Комментарий: система ОЧЕНЬ требовательна к настройкам

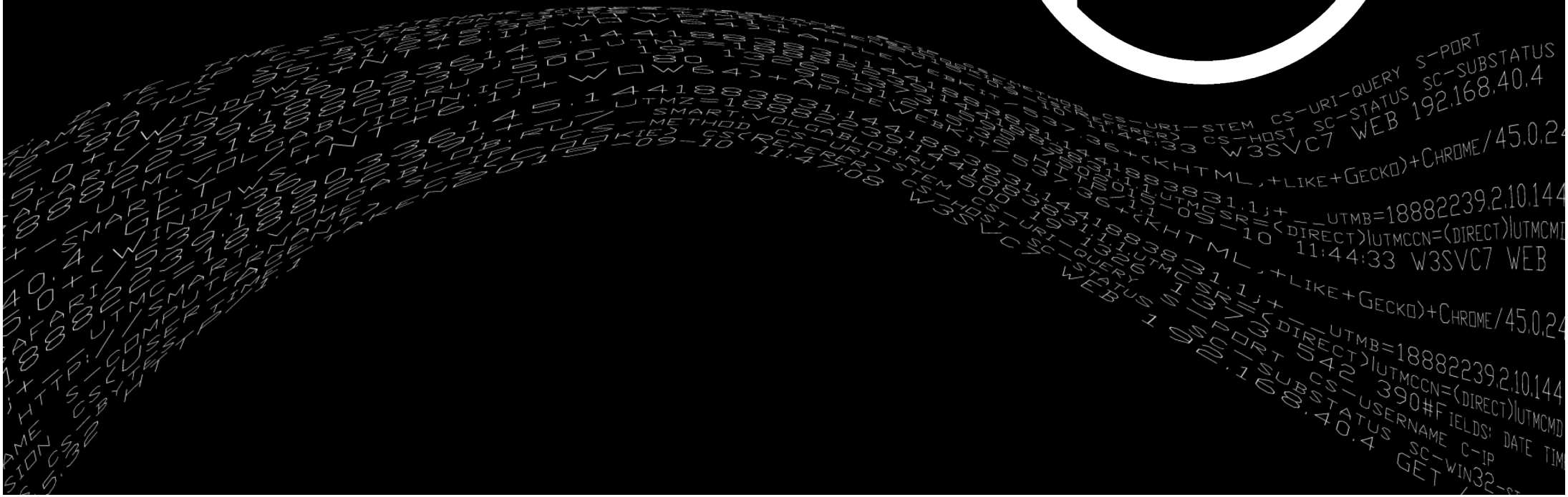
Поиск данных

Тестируем процессы поиска и выгрузки данных

VB-Trend 2019

plan>b

VolgaBlob



Условия тестирования

VB-Trend 2019 plan>b VolgaBlob

- Проверяемые варианты:
  - Splunk с извлечением полей во время поиска
  - Splunk с извлечением полей во время индексирования
  - Elastic Stack с query DSL
  - Elastic Stack с SME
- Тесты поиска – многократная проверка, по 5 запусков каждого запроса
- Измеряемые метрики
  - Среднее время выполнения запроса
  - Количество результатов

Что тестируем?

VB-Trend 2019 plan>b VolgaBlob

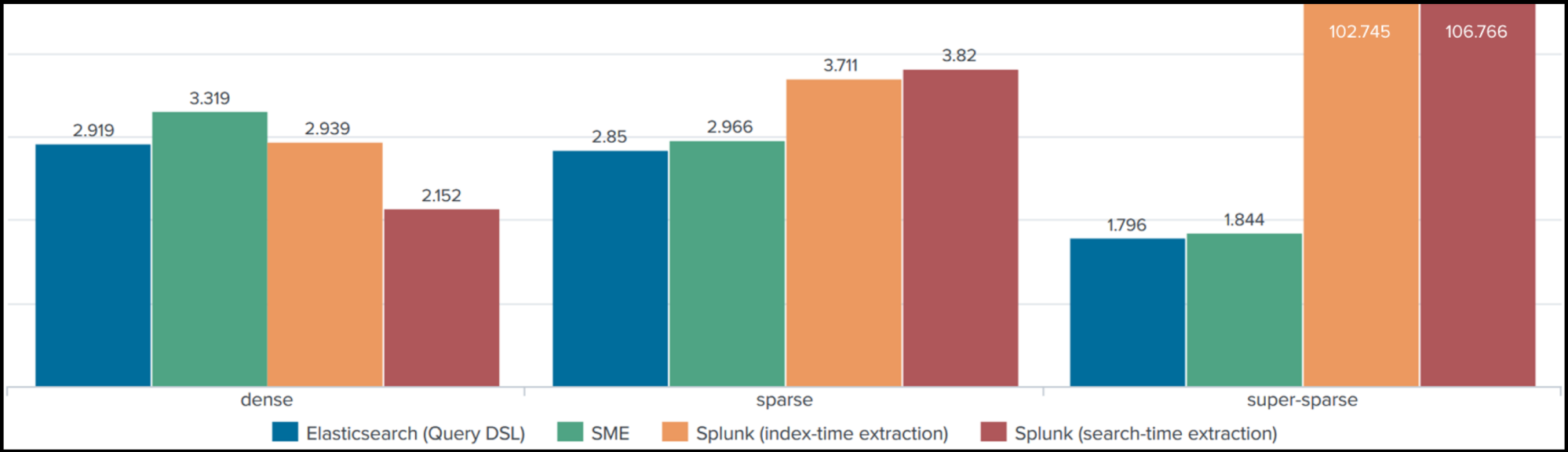
- Поиск данных для аналитики
  - до 10К записей
  - Dense / Sparse / Super-sparse поиск
- Поиск данных для экспорта
  - до 1М записей
  - Dense / Sparse / Super-sparse поиск

Немного о запросах

- Dense search – больше 10% событий в индексе
  - Расчет статистики для панелей дашбордов
- Sparse search – от 1 до 5% событий в индексе
  - Корреляционные правила вида «выявление попыток подбора паролей»
- Super-sparse search – меньше 0.01% событий в индексе
  - Поиски вида «иголка в стоге сена», выбор редко встречающихся событий

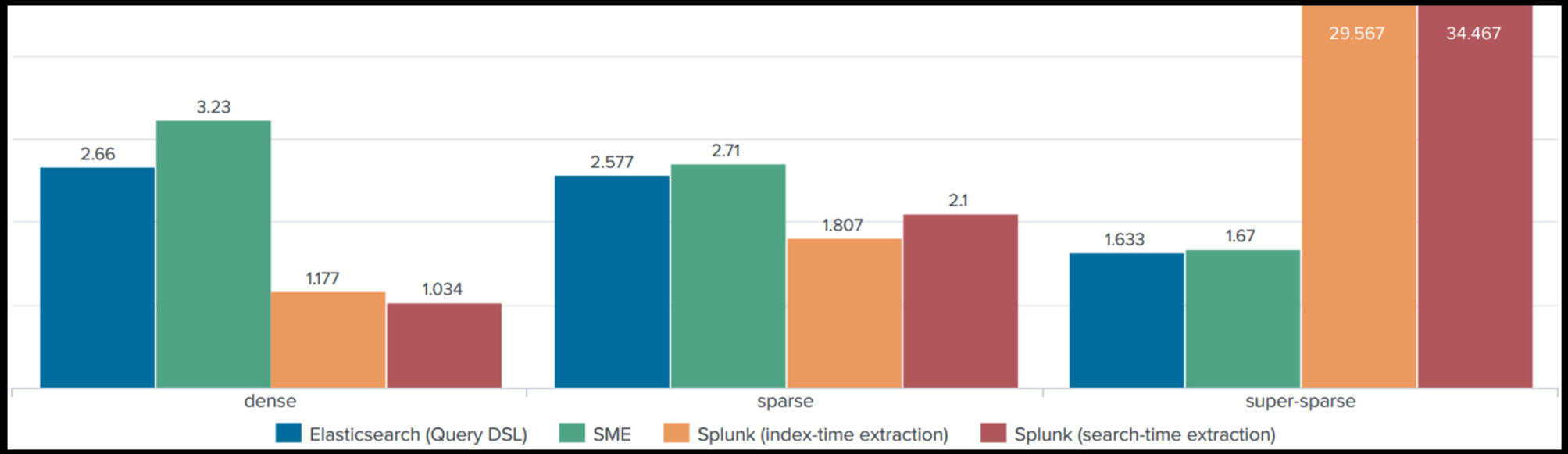
Поиск данных для аналитики - HDD

- Поиск последних 10К событий, связанных с нужным сервисом SSO
- Выбор сервиса определяет тип запроса



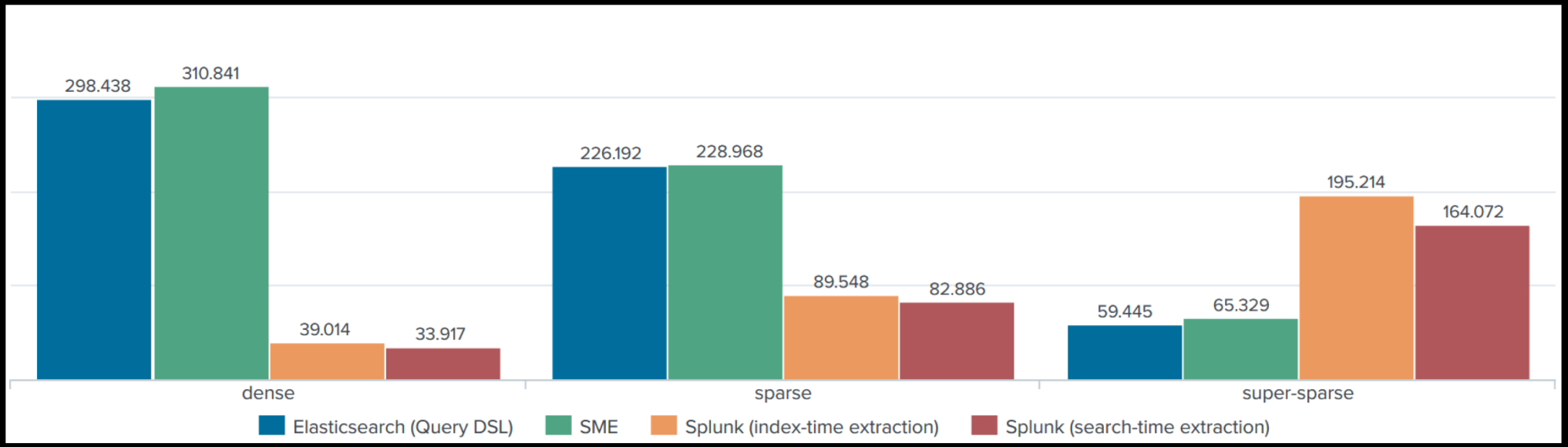
Поиск данных для аналитики - SSD

- Поиск последних 10К событий, связанных с нужным сервисом SSO
- Выбор сервиса определяет тип запроса



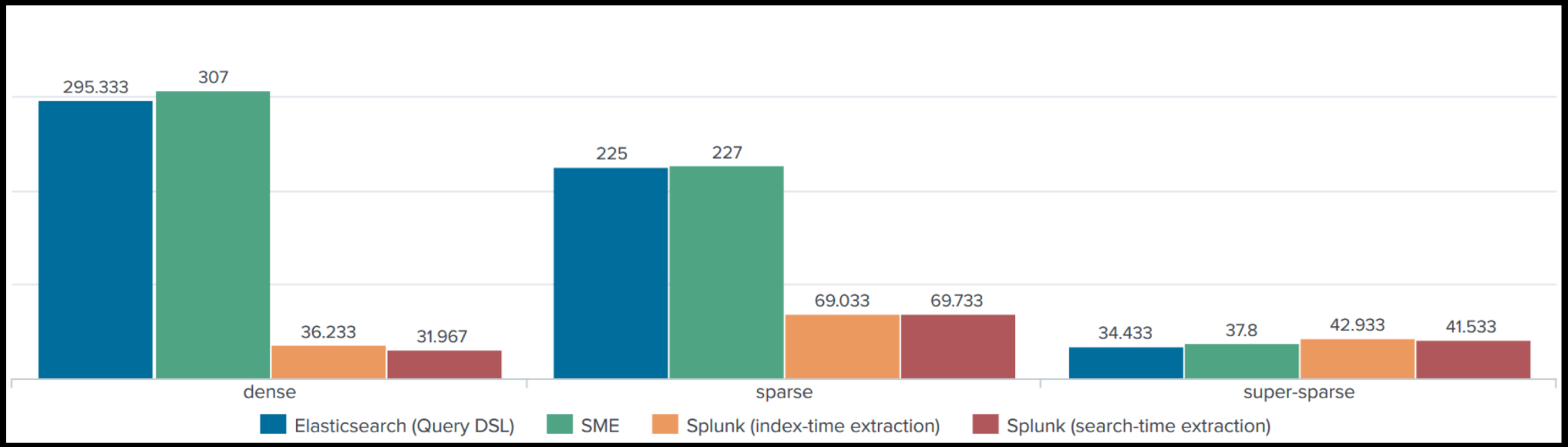
Поиск данных для экспорта - HDD

- Выгрузка 1M событий, связанных с нужным сервисом SSO
- Выбор сервиса определяет тип запроса



Поиск данных для экспорта - SSD

- Выгрузка 1M событий, связанных с нужным сервисом SSO
- Выбор сервиса определяет тип запроса

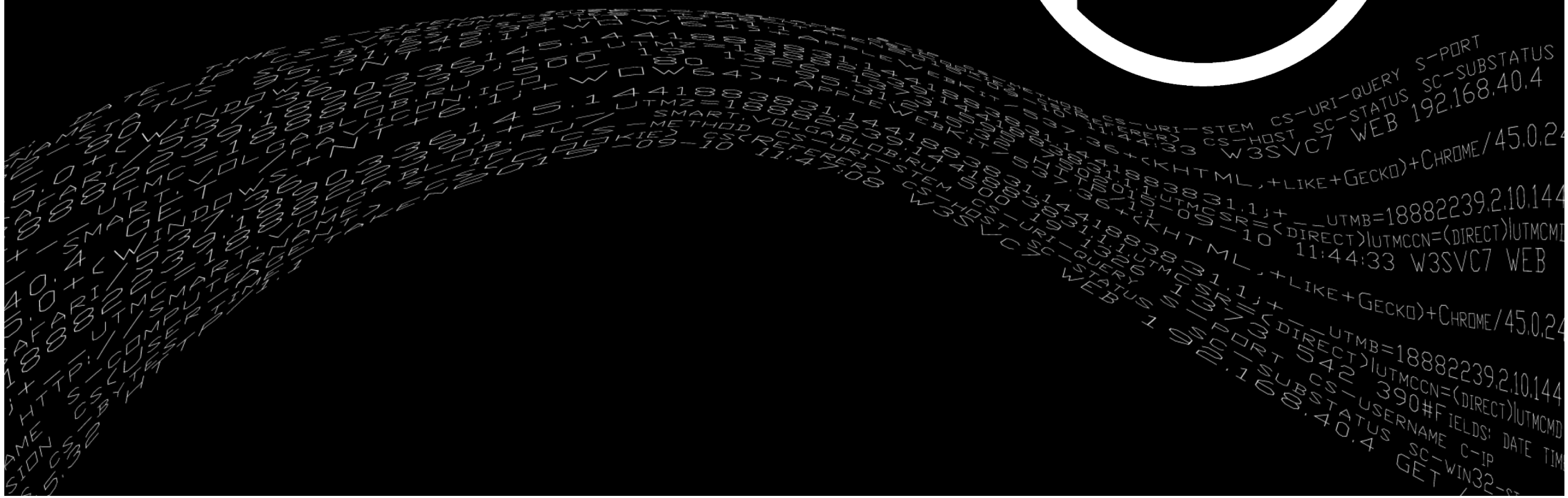


Что в итоге?

VB-Trend 2019

plan>b

VolgaBlob



Сравнение: индексирование

VB-Trend 2019 plan>b VolgaBlob

Время индексирования

10 — И ”Scheme on the fly”, и index-time extraction быстрее

7 — Низкая производительность Elastic, исправляется настройкой компонентов

Использование ресурсов сервера

10 — И индексирование, и отправка данных достаточно экономичны

5 — Высокое потребление процессорных ресурсов

Размер индексов

10 — Более компактная структура хранения при любых настройках

7 — Высокая зависимость от настроек index mappings

Сравнение: поиск

Запросы данных малого объема (~10K событий)

8	Высокая зависимость от скорости СХД, медленные super-sparse поиски	10	Практически не зависит от СХД, одинаково быстр на любых типах запросов	10	Практически не зависит от СХД, одинаково быстр на любых типах запросов

Запросы данных большого объема (~1M событий)

10	Более слабая зависимость скорости поиска от объема данных	5	Медленная выдача больших объемов данных	5	Медленная выдача больших объемов данных

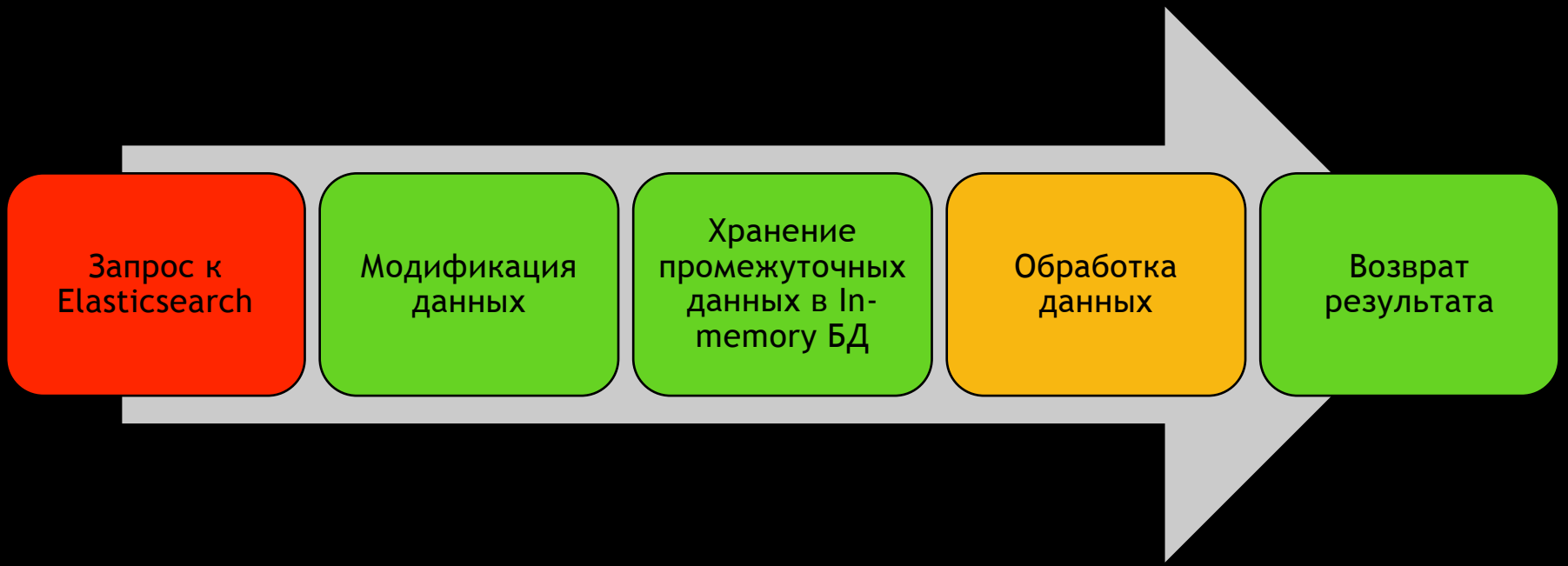
Запросы с постобработкой событий

10	Широкий набор команд обработки данных, удобный синтаксис	4	Громоздкий синтаксис, не везде можно использовать	6	Растущий набор команд обработки данных, удобный синтаксис

Запросы расчета статистики

10	Широкий набор команд и функций, высокая скорость работы	4	Неточность результатов, множество ограничений при расчете	6	Растущий набор команд и функций, точность расчетов

Проблемные места SME



Что с этим делать?

режим Smart Search

- Выдача результатов как можно раньше
- Анализ запроса и оптимизация процесса выполнения
- Конвейерное выполнение команд

Спасибо за внимание!

Вопросы? :)