

Ко-маркетинг с Arenadata

- Направлен на совместное продвижение решений Arenadata и ваших услуг/решений
- Большой охват аудитории - выше узнаваемость бренда
- Экономия вашего маркетингового бюджета



```
var wanna_comarketing = true;  
while (wanna_comarketing) {
```

1. Выбор целевых отраслей, компаний и тем, про которые говорим;
2. Помощь в подготовке sales-kit для ваших продажников;
3. PR-активности - пресс-релизы, публикации в СМИ, Хабр, vc.ru, блоги и др.
4. Дистрибуция контента - email-рассылки, telegram, соц.сети, интернет-реклама;
5. Помощь в подготовке лендингов и страниц по нашим решениям;
6. Проведение вебинаров и митапов (пока только их), совместные выступления на конференциях;
7. Подготовка и публикация выполненных проектов (кейсы) - arenadata.tech, ваш сайт, “Проект года от GlobalCIO”, tadvisee и пр.;
8. Помощь в подготовке и размещение видео на нашем канале YouTube по вашему решению или кейсу;
9. Размещение карточки решения в нашем маркетплейсе решений (с января 21) - будем вести трафик с интернет-рекламы на этот раздел, готовьте описание решений и видео!

```
}
```

Публикации в блоге

Мы запустили в середине сентября новый сайт, на котором есть блог. В планах его активное развитие - привлечение целевого трафика (интернет-реклама + email + telegram).

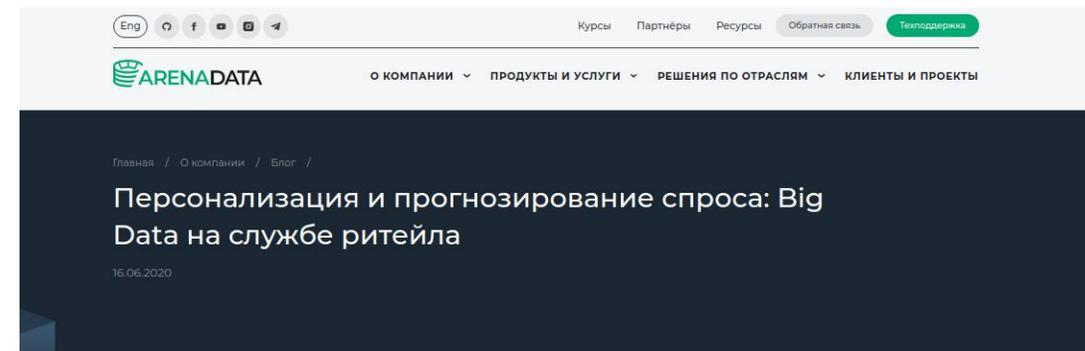
Публикуйте свои статьи в блоге бесплатно!

Требования к статье в блог:

- оригинальный текст, можно хороший рерайт + картинки
- статья, связанная с внедрением, настройкой или использованием стека Arenadata или общетематические по Big Data, AI, ML в отраслях
- продвижение статьи у себя в соц.сетях, в email-рассылке или telegram

А мы дополнительно направим СВОЙ трафик на неё.

С - синергия!



Любая сфера бизнеса генерирует огромные объёмы данных, анализируя которые можно добиться заметных успехов и конкурентных преимуществ. Михаил Гавриков, директор по продажам компании Arenadata, пообщавшись с известными российскими ритейлерами, рассказывает, как и для чего они используют технологии, построенные на Big Data.



Массивы данных в ритейле

Российские ритейлеры, особенно те из них, кто уже давно и успешно присутствуют на рынке, приходят к пониманию и принятию технологий продаж зарубежных коллег, самые успешные из которых личным опытом демонстрируют необходимость использования массивов данных о клиентах. Они поступают как из собственных систем лояльности, так и от партнёров. Полученные данные компании складывают в хранилища данных, где затем обогащают и анализируют их, строят аналитические модели и витрины (срезы, представляющие собой массивы тематической, законоправленной информации). В результате ритейлеры получают возможности для поиска дополнительных точек роста и автоматизации бизнеса.

Кому нужна СУБД Greenplum

О самом очевидном применении мы уже поговорили — такая система незаменима, когда данных слишком много. Если 2-4 терабайта можно как-то втиснуть на один-три сервера и даже обращаться к этим данным, то миллиард терабайт в обычную СУБД поместить весьма проблематично.

То есть **Greenplum нужна тем, у кого данных больше, чем очень много, то есть для работы с большими данными.**

Кроме того, хранить данные — часть дела. Если к записям нельзя обращаться за адекватное время и выполнять над ними нужные операции — толку от таких данных нет.

Поэтому **Greenplum нужна тем, кто не только хранит огромные объемы информации, но и активно с ними работает.**

Конечно, проблемы работы с большими объемами появились не вчера, инструменты под эти задачи на рынке есть: [ClickHouse](#), [Cassandra](#) и другие. Но после чтения документации можно заметить, что несмотря на общую сферу применения, у Greenplum есть особенности, которые четко определяют, когда точно нужна эта система, а когда стоит выбрать другую.

Сейчас мы поговорим о конкретных случаях и отличиях Greenplum от аналогов.



Облачная СУБД Arenadata DB

Для больших проектов на основе
Greenplum

Перейти

Mail.ru Cloud Solutions

**Завтра облачно**

1142 подписчика

Вы подписаны



1



Как работает Greenplum: аналитическая база данных для big data и больших проектов

5 августа  120 дочитываний  2,5 мин.

Greenplum — система управления данными из мира big data. Она нужна тем, кто анализирует и обрабатывает десятки терабайтов информации и кому тесно и некомфортно работать с обычными СУБД. Расскажем о том, что это за система, где и как ее использовать и в чем ее отличие от других систем, работающих с большими данными.



Самое главное: как устроена Greenplum

В основе Greenplum две вещи:

- знакомая многим база данных PostgreSQL;

Mail.ru Cloud Solutions

Мощная распределенная аналитическая база данных для больших проектов

Arenadata DB | Управляемая СУБД на основе Greenplum в
облаке

[Попробовать бесплатно](#)[Рассчитать стоимость](#)

Mail.ru Cloud Solutions

Быстро выполняйте сложные аналитические запросы с Arenadata DB на основе Greenplum

Arenadata DB — аналитическая база данных на основе Greenplum с открытым исходным кодом. Это массивно-параллельная СУБД, обладающая линейной масштабируемостью. Применяется в критически важных системах, работающих с большими объемами данных: объем данных в базе не ограничен, и она работает на 20% быстрее других СУБД.

Arenadata DB как сервис — быстрое кластерное решение, которое позволяет в несколько кликов развернуть базу для хранения и обработки больших данных, не вкладываясь в собственную инфраструктуру и ее поддержку.

[Попробовать бесплатно](#)

Результатов: примерно 95 200 (0,29 сек.)

Реклама · mcs.mail.ru/arenadata-db/облако ▾ 8 (499) 350-97-03

Arenadata DB в облаке - на 20% быстрее других СУБД

Снижение ТСО. Линейная масштабируемость. **Arenadata** DB в облаке. С поддержкой Mail.ru. 18+ Вся функциональность **Arenadata** DB с поддержкой от Mail.ru Group. Прозрачная оплата. Миграция «под ключ» Кастомные решения. MPP-архитектура. [Single Instance](#) - от 2,75 Р/ч. - Стоимость конфигурации · Ещё ▾

На основе Greenplum

Одновременная работа с данными из нескольких источников

Другие СУБД

PostgreSQL, MySQL, MongoDB, Redis, ClickHouse, Postgres Pro

arenadata.tech ▾

Arenadata — ведущий разработчик платформы сбора и ...

Ведущий разработчик платформ сбора и хранения данных для компаний, которые хотят изменить бизнес с помощью цифровизации процессов и ...

О компании

Arenadata — ведущий разработчик платформы ...

Arenadata DB

Arenadata DB (ADB) — аналитическая ...

Arenadata Hadoop

Продукты

Arenadata Enterprise Data Platform (EDP) — это ...

Arenadata QuickMarts

Arenadata QuickMarts (ADQM) — кластерная колоночная ...

Arenadata Streaming

Mail.ru Cloud Solutions



Курсы Big Data, Hadoop, Arenadata, Kafka и Spark

Специализированный авторизованный Учебный центр для корпоративного обучения по Большим Данным

ШКОЛА БОЛЬШИХ ДАННЫХ

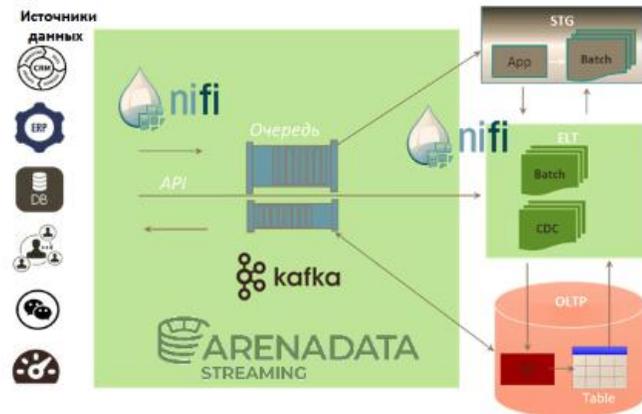
Поиск

- Главная
- Курсы
- Arenadata
- Блог
- Стоимость
- Регистрация
- О компании

Курсы Big Data, Hadoop, Arenadata, Kafka и Spark > News > Kafka > Что такое Arenadata Streaming: сочетаем Apache Kafka с NiFi в корпоративном масштабе

Что такое Arenadata Streaming: сочетаем Apache Kafka с NiFi в корпоративном масштабе

24 сентября, 2020 Анна Вичугова



Мы уже рассказывали про преимущества совместного использования Apache Kafka и NiFi. Сегодня рассмотрим, как эти две популярные технологии потоковой обработки больших данных (Big Data) сочетаются в рамках единого решения от отечественного разработчика — Arenadata Streaming. Читайте далее про основные сценарии использования и ключевые достоинства этого современного продукта класса Event Stream Processing в режиме реального времени.

Что такое Arenadata Streaming и зачем она нужна

Напомним, удобный веб-GUI и обилие готовых обработчиков событий (процессоров) с множеством настроек и поддержкой Python в NiFi отлично дополняют возможности сбора и агрегации потоковых данных Apache Kafka. Комбинация этих Big Data фреймворков дает максимальную гибкость разработчику Data Flow и инженеру Big Data, которые поддерживают этот конвейер обработки данных (data pipeline). Поэтому неудивительно, что именно эти две open-source технологии положены в основу Arenadata Streaming (ADS) - корпоративной платформы потоковых операций с большими данными в режиме онлайн. ADS также может использоваться в качестве корпоративной шины обмена данными, эффективно решая следующие задачи:

- получение данных в реальном времени из различных внешних систем, включая реляционные и NoSQL-СУБД;
- агрегация и преобразование потоков данных в распределенном режиме;
- надежное хранение данных в течение заданного периода времени;
- возвращение потребителям нужной информации, с автоматической балансировкой нагрузки и изоляцией критически важной информации от несанкционированного доступа.

АКЦИИ



Новое на сайте

- » Как очистить большие данные для Apache Spark SQL: краткий обзор Cleanframes
- » Как устроен конвейер аналитики больших данных на Apache Kafka и Druid в Netflix
- » Зачем вам UNION вместо JOIN в Apache Druid и семплирование больших данных в Spark Streaming: пример потоковой аналитики Big Data
- » 3 метода векторизации слов в PySpark
- » Почему вам нужно обучение разработке в Apache Kafka: 4 причины выбрать курсы от Школы Больших Данных

Отзывы на Google

Школа Больших Данных

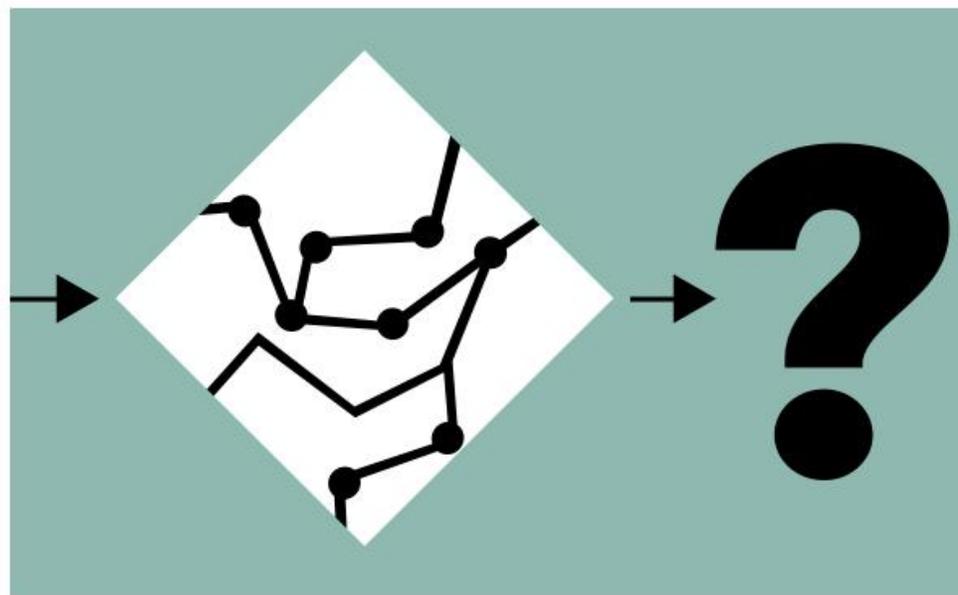


Поделиться ↗

Кейсы: как использовать big data на производстве

Производство Цифровизация Кейсы

Антон Балагаев, директор по консалтингу компании Arenadata, и Антон Денисов, директор по работе с ключевыми заказчиками IBS — о том, какие задачи и как производственные компании решают с помощью big data



Зачем big data на производстве

IBS



101,73

Рейтинг

Подписаться

КРОК Облачные сервисы

Облачная IaaS-платформа собственной разработки

SZinkevich 29 января 2020 в 10:06

Массивно-параллельная база данных Greenplum — короткий ликбез

Блог компании КРОК Облачные сервисы, IT-инфраструктура, Хранение данных, Облачные сервисы



Для Hadoop и Greenplum есть возможность получить готовый SaaS. И если Хадуп — известная штука, то Greenplum (он лежит в основе продукта ArenadataDB, про который далее пойдёт речь) — интересная, но уже менее «на слуху».

Arenadata DB — это распределённая СУБД на базе опенсорсного Greenplum. Как и у других решений MPP (параллельной обработки данных), для массивно-параллельных систем архитектура облака далека от оптимальной. Это может снижать производительность аж до 30 % (обычно меньше). Но, тем не менее, эту проблему можно нивелировать (о чём речь пойдёт ниже). Кроме того, стоит покупать такую услугу из облака, часто это удобно и выгодно в сравнении с развёртыванием собственного кластера.

В гайдах явно указывается on-premise, но сейчас многие осознают масштаб удобства облака. Все понимают, что некая деградация производительности будет, но это настолько всё равно супер по удобству и скорости, что уже есть проекты, где этим жертвуют на каких-то этапах вроде проверки гипотез.

Если у вас есть хранилище данных больше 1 ТБ и транзакционные системы — не ваш профиль по нагрузке, то ниже — рассказ, что можно сделать как вариант. Почему 1 ТБ? Начиная с этого объёма использование MPP эффективнее по соотношению производительность/стоимость, если сравнивать с классическими СУБД.

Когда использовать?

Когда классическая single-node СУБД по архитектуре не подходит для ваших объёмов. Частый случай — новое хранилище данных объёмом больше 1 ТБ. MPP СУБД сейчас в тренде, и Greenplum — один из лучших на рынке по соотношению производительность/стоимость, если сравнивать с классическими СУБД.

ИНФОРМАЦИЯ

Table with 2 columns: Field (Дата основания, Локация, Сайт, Численность, Дата регистрации) and Value (2 марта 2018 г., Москва Россия, cloud.croc.ru, Неизвестно, 27 февраля 2018 г.)

БЛОГ НА ХАБРЕ

Как упростить жизнь IT-разработчику и ускорить time-to-market с месяца до двух дней

1,5k 2

Опыт смены SAP-хостинга: как мигрировать системы, чтобы не было мучительно больно

1,7k 7

Как мы превратили «статистическую аномалию» в сервис: новый уровень хранения в облаке

3,7k 5

Почему при Covid-19 увеличилась переподписка, и как это проверить

5,4k 16

Как мы проверяли производительность новых процессоров в облаке для ТС по тесту Гилёва

10,3k 30

Packer, Terraform и Ansible: деплой кластера Kubernetes за час

9,9k 3

Массивно-параллельная база данных Greenplum — короткий ликбез

5,2k 7

Дизайн сервиса: как Amazon на...



266,12

Рейтинг

Подписаться

КРОК

IT-компания

PavelEgorov 5 ноября 2020 в 14:24

Как мы искали неработающие датчики на «УРАЛХИМЕ» (первый проект Data Lake)

Блог компании КРОК, IT-инфраструктура, Хранилища данных



«УРАЛХИМ» делает удобрения. № 1 в России — по производству аммиачной селитры, например, входит в топ-3 отечественных производителей аммиака, карбамида, азотных удобрений. Выпускаются серные кислоты, двух-трёхкомпонентные удобрения, фосфаты и многое другое. Это всё создаёт агрессивные среды, в которых выходят из строя датчики.

Мы строили Data Lake и заодно охотились на те датчики, которые замерзают, выходят из строя, начинают давать ложные данные и вообще ведут себя не так, как должны себя вести источники информации. А «фишка» в том, что невозможно строить матмодели и цифровые двойники на базе «плохих» данных: они просто не будут правильно решать задачу и давать бизнес-эффект.

Но современным производствам нужны Data Lake'и для дата-сайентистов. В 95 % случаев «сырые» данные никак не собираются, а учитываются только агрегаты в АСУТП, которые хранятся два месяца и сохраняются точки «изменения динамики» показателя, которые вычисляются специально заложенным алгоритмом, что для дата-сайентистов снижает качество данных, т. к., возможно, может пропустить «всплески» показателя... Собственно, примерно так и было на «УРАЛХИМЕ». Нужно было создать хранилище производственных данных, подключить к источникам в цехах и в MES/ERP-системах. В первую очередь это нужно для того, чтобы начать собирать историю для дата-сайенса. Во вторую очередь — чтобы дата-сайентисты имели площадку для своих расчётов и песочницу для проверки гипотез, а не нагружали ту же самую, где крутятся АСУ ТП. Дата-сайентисты пробовали сделать анализ имеющихся данных, но этого не хватило. Данные хранились прореженные, с потерями, часто...

ИНФОРМАЦИЯ

Table with 2 columns: Field (Дата основания, Локация, Сайт, Численность, Дата регистрации) and Value (1 января 1992 г., Москва Россия, croc.ru, 1 001-5 000 человек, 24 августа 2010 г.)

ВИДЖЕТ



ССЫЛКИ

Облако КРОК cloud.croc.ru

КРОК на Youtube www.youtube.com

FACEBOOK



Миграция системы подготовки отчетности банка ВТБ с Oracle SuperCluster на российское решение Arenadata DB

ЗАКАЗЧИК

ПАО ВТБ

РУКОВОДИТЕЛЬ ПРОЕКТА СО СТОРОНЫ ЗАКАЗЧИКА



Владислав Отбоев

руководитель службы управления ИТ-поставщиками, департамент организации и управления ИТ

ИТ-ПОСТАВЩИК

Arenadata

ГОД ЗАВЕРШЕНИЯ ПРОЕКТА

2020

СРОКИ ВЫПОЛНЕНИЯ ПРОЕКТА

Сентябрь, 2019 - Сентябрь, 2020

МАСШТАБ ПРОЕКТА

3000 человеко-часов

ЦЕЛИ

ВТБ столкнулся с ситуацией вывода из поддержки продуктов компании Oracle и дефицитом физического пространства в связи с расширением. Понадобилось определиться с целевой платформой для системы подготовки отчетности исходя из строгих требований регулятора в части времени подготовки отчетности и прицела на импортозамещение.



ПРЕДМЕТНАЯ ОБЛАСТЬ

- ERP-системы
- ITSM
- CRM-системы
- ECM-системы
- Бизнес-приложения (BPM,PLM,ToIP)
- Аналитические решения и большие данные
- Сервисы
- Создание и модернизация инфраструктуры
- Облачные среды
- Отечественная разработка
- Импортозамещение
- Мобильные решения
- AI/RPA/IoT
- Аутсорсинг
- IoT и RPA

СКАЛА

ФОРСАЙТ

Расписание

Спикеры

Контакты



33.825

Вебинар

«Форсайт» + Arenadata: особенности применения различных СУБД в аналитике больших данных

17 ноября в 11:00 (мск)

[зарегистрироваться](#)

50.826

58.062

2.305

ФОРСАЙТ

Быть или не быть?



ИЛИ



Обсудим ко-маркетинг?



Александр Осипов
oav@arenadata.io

