

Онлайн-кинотеатр ivi всматривается в аналитику

Переход на Vertica позволил не только существенно повысить скорость получения аналитических отчетов, но и сделать их более сложными и глубокими

Вызовы

Низкая производительность и недостаточная гибкость аналитических систем

ivi — крупнейший онлайн-кинотеатр России, аудитория которого превышает 48 млн уникальных посетителей в месяц. Ежемесячно пользователи проводят на ivi более 70 млн часов и осуществляют более 250 млн видеопросмотров. На ivi приставлен обширный каталог фильмов, мультфильмов и сериалов. Всего фильмотека онлайн-кинотеатра насчитывает более 62 тыс. единиц контента — это картины всех шести студий-мейджоров, а также более 400 европейских и российских независимых кинопроизводителей. Кроме того, в 2017 году ivi подписал эксклюзивные партнерские соглашения с крупными российскими телеканалами — Первым каналом, НТВ и СТС, согласно которым, популярный контент каналов доступен на ivi, в том числе раньше, чем в эфире телеканалов и сразу после ТВ эфира. Зрителям также доступны оригинальные звуковые дорожки и субтитры. ivi можно смотреть во всем мире на огромном количестве устройств — Smart TV, web, Android, iOS, игровые приставки (в США доступны iOS и web).

Шесть лет назад производительность построенного на базе PostgreSQL прежнего хранилища данных, в которое стекается статистика о показах контента и рекламы (на их основе готовятся отчеты для рекламодателей и правообладателей),

«Благодаря Vertica мы получили полную гибкость работы с большими объемами данных и широкие возможности для выполнения произвольных запросов»

ДМИТРИЙ АНТОНОВ

руководитель отдела технической аналитики компании ivi

окончательно перестала устраивать и аналитиков компании ivi, и сотрудников ее бизнес-подразделений. Всем стало понятно, что пора переходить на другую платформу СУБД, рассчитанную на аналитическую работу с данными.

«Мы искали решение, которое позволило бы гибко работать с информацией, сохраняя “сырые” данные о каждом показе рекламы или контента, чтобы работа с данными стала максимально удобной и при этом обеспечивалась быстрая подготовка хотя бы стандартных отчетов», — вспоминает Дмитрий Антонов, руководитель отдела технической аналитики компании ivi.

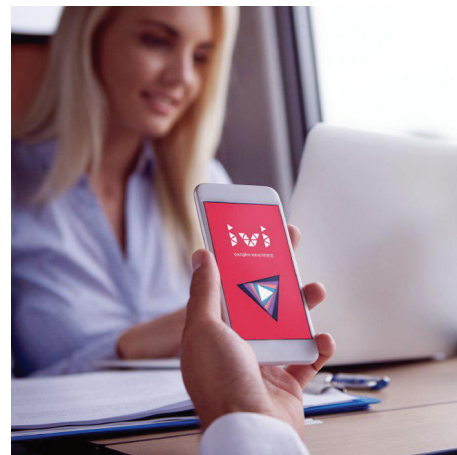
В числе основных требований, которые предъявлялись к новой аналитической платформе, — высокая производительность, ускоренное построение отчетов и быстрая загрузка данных в систему. Разумеется, учитывалась также финансовая сторона приобретения новой аналитической платформы и владения ею.

Решения

Быстрая и удобная основа для принятия бизнес-решений

По рекомендации одного из своих коллег ИТ-специалисты ivi развернули и протестировали бесплатно распространяемую версию Vertica Community Edition и убедились, что эта СУБД работает, по меньшей мере, на порядок быстрее той, что использовалась ранее. Кроме того, они сравнили возможности Vertica с конкурирующей аналитической платформой Sybase IQ, для чего был организован небольшой пилотный проект. После обсуждения его результатов выбор был сделан в пользу Vertica.

Внедрение Vertica активно поддерживал занимавший в то время пост технического директора Виктор Романчук. Новую платформу в ivi внедряли собственными силами, не обращаясь к внешним специалистам. Для развертывания системы были закуплены три сервера, сертифицированные на совместимость с Vertica.



О проекте

■ Отрасль:

интернет-торговля

■ Задача

Обеспечение выполнения точных расчетов с правообладателями и рекламодателями путем использования более информативных и гибких аналитических отчетов

■ Подход

Переход на высокопроизводительную аналитическую платформу и формирование эффективного окружения ETL, средств визуализации и выгрузки данных

■ Результаты

- + Существенное повышение скорости выполнения аналитических запросов и получение более детальных отчетов
- + Своевременное принятие различных управленческих решений на основе оперативной аналитической отчетности

■ Решение

СУБД Vertica

«С момента внедрения Vertica многое в компании изменилось, и теперь мы производим гораздо более серьезный анализ данных»

ДМИТРИЙ АНТОНОВ

руководитель отдела технической аналитики компании ivi

www.vertica.com
www.microfocus.com

Переход с Community Edition на коммерческую версию Vertica занял около полугода: специалисты ivi тщательно тестировали все отчеты, чтобы убедиться, что они работают исправно. «Мы получили удобное и надежное хранилище, в котором не возникает проблем, связанных с потерей данных или подключением к средствам визуализации. Подготовка отчетов заметно упростилась, а их разнообразие существенно увеличилось», — рассказывает Дмитрий Антонов.

Следующие несколько лет в ivi использовалась платформа Vertica версии 7, развернутая на все тех же трех серверах, причем внешние системы хранения не использовались.

В качестве инструментария для подготовки отчетов и визуализации данных применяются продукты компании Tableau Software. Интеграция с Vertica реализована через ODBC-драйверы. Кроме того, к Vertica подключены созданные своими силами ETL-средства и модули выгрузки отчетности. В будущем, возможно, к ним добавится новая система подготовки отчетов на базе Microsoft Reporting Services.

В Vertica аккумулируются и анализируются данные о просмотре контента и рекламы, а также информация об оплате. «Это критически важные сведения, и их анализ должен выполняться максимально быстро и точно», — отмечает Дмитрий Антонов. Данные передаются из биллинговой системы и системных журналов (логов), в которых фиксируются показы рекламы и контента.

Онлайн-кинотеатр ivi смог существенно повысить качество управленческих решений, опираясь на отчеты, получаемые на основе аналитической платформы Vertica. Благодаря высокой ее производительности и удобству работы аналитики ivi используют для получения отчетов гораздо более сложные, чем прежде, запросы.

Непосредственными пользователями Vertica являются специалисты отдела аналитики компании ivi, продуктового отдела (они анализируют спрос на представленную в интернет-кинотеатре продукцию), а также менеджеры и сотрудники ряда других отделов — всего около полутора десятка человек. Отчетами, полученными на основе данных Vertica, пользуются около полутора сотен человек — и руководители, и рядовые сотрудники. Наиболее часто их изучают те, кто отвечает за размещение рекламы, подготовку каталога контента, выбор продукции для показа, а также за анализ продаж платного контента.

С помощью Vertica готовятся в первую очередь отчеты для акционеров и высшего руководства компании, а также сводятся воедино данные о взаиморасчетах с партнерами — рекламодателями и правообладателями. На их основе принимаются различные оперативные управленческие решения, в том числе об аудитории рассылок сообщений электронной почты, о выборе моделей контрактов с партнерами и многие другие.

Чтобы лучше понять своих клиентов, ivi собирает и анализирует данные, характеризующие поведение пользователей. Их хранение и анализ поначалу осуществлялись в Vertica Community Edition, однако после перехода на коммерческую версию платформы эту информацию стали собирать и обрабатывать средствами Hadoop, поскольку бюджет на хранение Больших данных в компании был ограничен и не позволял с необходимой частотой наращивать инфраструктуру.

Преимущества

Новый уровень разнообразия и глубины аналитики

Благодаря внедрению Vertica компании ivi удалось успешно решить проблему повышения производительности аналитической системы. «Произошел своего рода прорыв — разнообразие и глубина нашей аналитики существенно увеличились. Мы получили полную гибкость работы с данными, которые у нас накапливались, и широкие возможности для произвольных запросов: данные по-прежнему хранятся в “сыром” виде и изучаются под самыми разными углами зрения. С момента внедрения Vertica многое изменилось,

и теперь мы производим гораздо более серьезный анализ данных. Компания выросла, и круг сотрудников, которые пользуются аналитическими данными, увеличился примерно в четыре раза», — отмечает Дмитрий Антонов.

В данный момент инсталляция Vertica в ivi расширяется — происходит перенос прежней продуктивной среды на новый кластер из пяти серверов. Старый кластер из трех машин планируется использовать в качестве среды разработки и тестирования. Поскольку Vertica лицензируется по объему анализируемых данных, а не по аппаратным параметрам или количеству пользователей, при этом среды разработки предоставляются бесплатно, наращивание мощностей не требует приобретения дополнительных лицензий.

Одновременно с этим ivi производит апгрейд Vertica до новейшей версии 9.1, которая содержит множество дополнительных инструментов для облегчения «классических» задач хранилищ данных, в частности, упрощенное, ускоренное и экономичное (с точки зрения лицензирования) построение витрин с помощью функционала «плоских» таблиц (flattened tables). Кроме этого, пользователям открываются новые потенциально интересные возможности, например, поддержка различных сценариев использования машинного обучения в организации, от использования Vertica для хранения и подготовки данных и интеграции с существующими средами машинного обучения до распределенного расчета моделей в самой СУБД.

В дальнейшем планируется использовать Vertica не просто как хранилище «сырых» данных, которые применяются для подготовки отчетности и аналитики, а как систему накопления агрегированных данных и построения расширенных витрин данных.

«Переход на платформу Vertica способствует достижению нового уровня качества аналитики. Если заказчику нужен быстрый и при этом глубокий, сложный анализ данных, выбор платформы Vertica является оптимальным», — комментирует Евгений Степанов, руководитель направления Big Data Platform компании Micro Focus в России.