



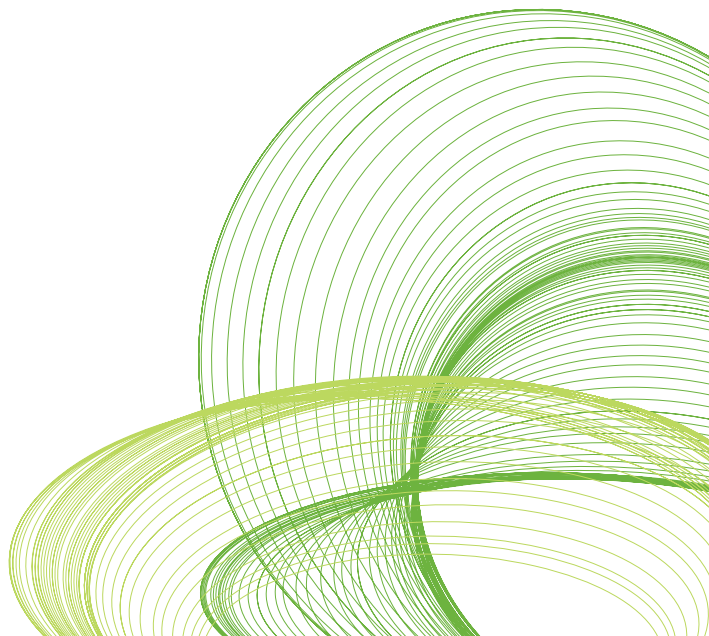
# Удобство ежедневной работы для тысяч пользователей по всему миру

---

Информационные материалы по технологиям QlikView

Опубликовано: октябрь, 2010 г.

[www.qlikview.ru](http://www.qlikview.ru)



## Содержание

---

<b>Наш критерий — практика</b>	3
<b>Всем веб-приложениям требуется поддержка масштабируемости</b>	3
<b>МЫ САМИ СТАЛКИВАЛИСЬ С ПРОБЛЕМАМИ МАСШТАБИРУЕМОСТИ</b>	4
<b>Наше решение: гибкость, распределенное размещение и высокая производительность</b>	5
МЫ ИСПОЛЬЗОВАЛИ ГИБКОСТЬ «ОБЛАКА»	5
МЫ РАЗМЕСТИЛИ СЕРВЕРЫ БЛИЖЕ К ПОЛЬЗОВАТЕЛЯМ	6
МЫ ИСПОЛЬЗОВАЛИ БОЛЕЕ МОЩНЫЕ СЕРВЕРЫ И УВЕЛИЧИЛИ ИХ КОЛИЧЕСТВО	7
<b>История успеха: «Kick It and Qlik It»</b>	8
<b>Инвестиции в масштабируемость окупались</b>	9
<b>Рекомендации</b>	10
<b>Приложение</b>	11
<b>Информационные материалы по технологиям QlikView</b>	11

## ЧТО ТАКОЕ МАСШТАБИРУЕМОСТЬ В КОНТЕКСТЕ БИЗНЕС-АНАЛИТИКИ?

Применительно к BI-платформам масштабируемость определяется следующими факторами:

- **Количество одновременно работающих пользователей.** Способна ли платформа поддерживать параллельную работу сотен или тысяч пользователей? Какое максимальное количество пользователей допустимо для BI-приложения?
- **География.** Можно ли масштабировать сетевые и вычислительные ресурсы, обеспечивая поддержку пользователей системы по всему миру?
- **Объем данных.** Способна ли система загружать большие объемы данных в память и агрегировать их на нескольких уровнях, чтобы удовлетворить все потребности пользователей?
- **Разработка.** Справляется ли архитектура с поддержкой параллельной работы десятков, а то и сотен разработчиков? Можно ли распространять BI-приложения дистанционно и с многократными итерациями без ущерба для их управляемости?

*В данном обзоре мы уделим особое внимание двум параметрам масштабируемости: количеству одновременно работающих пользователей и географии.*

## Наш критерий — практика

Наша компания стремится к тому, чтобы система QlikView использовалась повсеместно. Поначалу мы довольствовались получением благодарных отзывов от сотрудников отдельных предприятий. Количество таких отзывов росло, и распространение системы приняло лавинообразный характер. По мере того, как наша платформа для бизнес-аналитики (BI) обретала зрелость, QlikView превратилась из настольного приложения в гибкое универсальное решение, доступное практически везде, — на ПК, мобильном устройстве или через интернет-браузер. Система может интегрироваться в другие приложения. Наша задача — предоставить оптимальные инструменты бизнес-анализа всем, кому они необходимы.

Со временем QlikView все активнее применялась в работе крупных предприятий с тысячами и даже десятками тысяч сотрудников. Кроме того, система стала доступна и для посетителей публичных веб-сайтов. Возможности масштабирования и поддержки одновременной работы тысяч пользователей становились все более актуальным требованием, предъявляемым к платформам бизнес-аналитики. Компания QlikTech и сама столкнулась с этой проблемой после того как начала активно использовать свой демонстрационный сайт в маркетинговых целях.

В предлагаемом вашему вниманию информационном обзоре технологий QlikView описываются способы обеспечения масштабируемости, применявшиеся для организации доступа к демонстрационному сайту системы (demo.qlikview.com), которая, не снижая производительности и скорости реакции, ежедневно обслуживала тысячи посетителей. Мы познакомим вас с приложением «Kick It and Qlik It», позволяющим осуществлять бизнес-анализ в реальном времени. Оно было разработано специально для футбольных болельщиков, следивших за чемпионатом мира 2010 года. Данный материал адресован ИТ-профессионалам и технически подготовленным бизнес-пользователям, изучающим возможности платформ бизнес-аналитики и желающим получить более подробную информацию о масштабируемости системы QlikView.

## Всем веб-приложениям требуется поддержка масштабируемости

С момента появления Интернета вопрос масштабируемости никогда не терял своей актуальности. Тот, кто застал начальный этап развития коммерческой сети, помнит чудовищно медленные модемные соединения. В время шутили: глобальная сеть (world wide web) — это на самом деле глобальное ожидание (world wide wait).

Вот несколько фактов, касающихся масштабируемости:

- **Масштабируемость определяется доступными вычислительными мощностями и объемом памяти.** По мере улучшения качества интернет-соединений проблемы масштабируемости стали зависеть не столько от пропускной способности канала, сколько от вычислительных мощностей и памяти. Не все организации могут приобрести достаточно дорогое компьютерное оборудование, способное обеспечить одновременный доступ тысяч пользователей к их веб-приложениям. Устранить нехватку вычислительных мощностей и памяти позволили «облачные» архитектуры, впервые опробованные Amazon, eBay, Google и другими компаниями.

## ПРОБЛЕМА МАСШТАБИРУЕМОСТИ, С КОТОРОЙ СТОЛКНУЛАСЬ КОМПАНИЯ QLIKTECH

Поскольку спрогнозировать количество пользователей, которые будут одновременно посещать демонстрационный сайт QlickView из различных стран мира, не представлялось возможным, нам нужно было обеспечить необходимое масштабирование системы без излишних затрат на вычислительные ресурсы.

- **Проблемы масштабируемости проявляются в виде снижения скорости работы.** Когда один массив вычислительных ресурсов вынужден обрабатывать одновременно слишком большое число обращений, это сразу же снижает скорость реагирования системы, что мешает работе пользователей. Для преодоления узких мест «узких мест» веб-приложениям нужны более совершенные процессоры (CPU), высокая скорость ввода/вывода, оперативная память большей емкости. Увеличение памяти пойдет на пользу любому приложению, ведь производительность существенно падает, если оперативная память оказывается полностью исчерпана и сервер вынужден прибегать к виртуальной памяти.
- **Пики посещаемости веб-сайта создают дополнительные сложности.** Наибольшее число посетителей отмечается во время проведения международных рекламных акций или в связи с упоминаниями в престижных изданиях. Размещение ссылки на популярном сайте тоже может привести к резкому росту посещаемости. Нарастивание серверов в ожидании этого момента будет весьма дорогостоящим и неэффективным использованием ресурсов. Как же добиться гибкой масштабируемости, не тратясь на мощности, которые, возможно, никогда не будут использоваться?

### МЫ САМИ СТАЛКИВАЛИСЬ С ПРОБЛЕМАМИ МАСШТАБИРУЕМОСТИ

Когда QlikTech начала продвигать демонстрационный сайт, мы постарались обеспечить пользователям оптимальные условия работы. Если сайт функционирует плохо, посетители не выясняют, в чем причина. Им нет дела до того, где размещены серверы, неинтересны объяснения про недостаточную пропускную способность сети где-то на другом конце света. Они оценивают свой опыт обращения к системе QlickView в целом.

- **Изначально все наши ресурсы размещались компактно.** В начале 2009 года мы установили демо-версию QlickView на серверах нашего ЦОД в Швеции, на родине QlikTech. В то время мы использовали три 64-битных (x64) сервера, каждый из которых был оснащен восьмиядерным процессором и 12 Гб памяти. Для балансировки нагрузки применялся маршрутизатор Cisco.
- **Низкая производительность серверов и «узкие места» в сети вызывали нарекания.** Серверы были мощными, но этого было недостаточно. Когда количество пользователей существенно возрастало, скорость реакции системы стремительно снижалась. Более того, для пользователей из Азии, Африки или Америки сетевые маршруты оказывались слишком продолжительными из-за локального размещения серверов. Результат? Время отклика системы возрастало и зависело от текущей нагрузки на сервер.

Необходимость изменений была очевидна: чтобы пользователи могли получить позитивные впечатления от использования QlickView, следовало позаботиться о масштабируемости системы.

## ПРЕДПРИНЯТЫЕ НАМИ ДЕЙСТВИЯ

- **Перемещение в «облако».** Мы перевели серверы QlikView из центра обработки данных в Швеции в американский ЦОД, размещенный в «облаке» Amazon EC2. Переход на гибкий «облачный» ресурс стал экономичным способом решения проблемы пиковых нагрузок.
- **Расширение географии.** К существующей конфигурации мы добавили еще две географически распределенные точки размещения серверов, приблизив их к пользователям. В результате сетевые маршруты сократились и улучшилось время отклика системы.
- **Увеличение мощности серверов.** Мы перешли с «мощных» серверов на «сверхмощные» и увеличили их минимальное количество в каждой точке с трех до пяти. Появилась возможность поддерживать большее количество пользователей одновременно.

## Наше решение: гибкость, распределенное размещение и высокая производительность

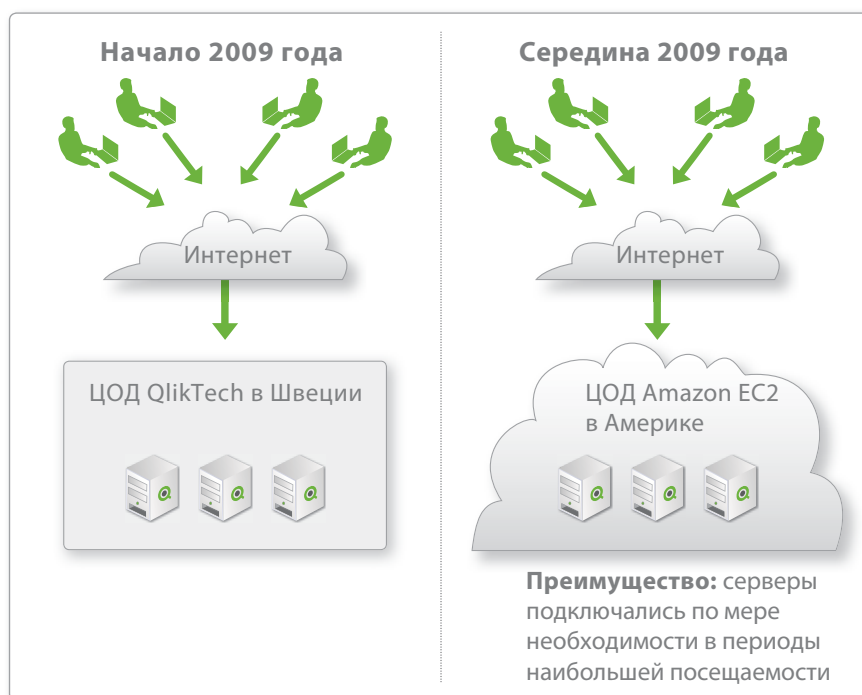
Для удовлетворения растущей потребности в масштабируемости мы предприняли несколько шагов. Во-первых, мы перенесли серверы QlikView из центра обработки данных QlikTech в «облако». Во-вторых, использовали дополнительные ЦОД, чтобы серверы располагались в трех точках земного шара. Наконец, мы заменили серверы на более мощные и увеличили их количество.

### МЫ ИСПОЛЬЗОВАЛИ ГИБКОСТЬ «ОБЛАКА»

В середине 2009 года мы перенесли серверы, поддерживающие сайт `demo.qlikview.com`, в «облако» (см. Рисунок 1), чтобы в сжатые сроки добиться именно той производительности, которая требовалась для маркетинговых инициатив. Мы использовали веб-сервис Amazon, известный как Amazon Elastic Compute Cloud (EC2), обеспечивающий масштабируемые вычислительные мощности и облегчающий перенос ИТ-инфраструктуры в Web.

Перенос демонстрационного сайта в «облако» позволил нам гибко наращивать и сокращать вычислительные мощности. Модель оплаты на основе фактического использования ресурсов, которую применял провайдер «облачного» сервиса, позволила производить оплату исходя из текущего трафика и расходовать средства на поддержку только нужных нам серверов — это стало эффективным способом решения проблемы пиков посещаемости без больших затрат на аппаратные или прочие ресурсы. Когда пиковые нагрузки заканчивались, от использования части серверов можно было отказаться: в течение нескольких часов мы ожидали подтверждения, что доступ к тем или иным серверам не востребован, и затем выполняли их отключение.

Рисунок 1: Мы использовали гибкость «облака»



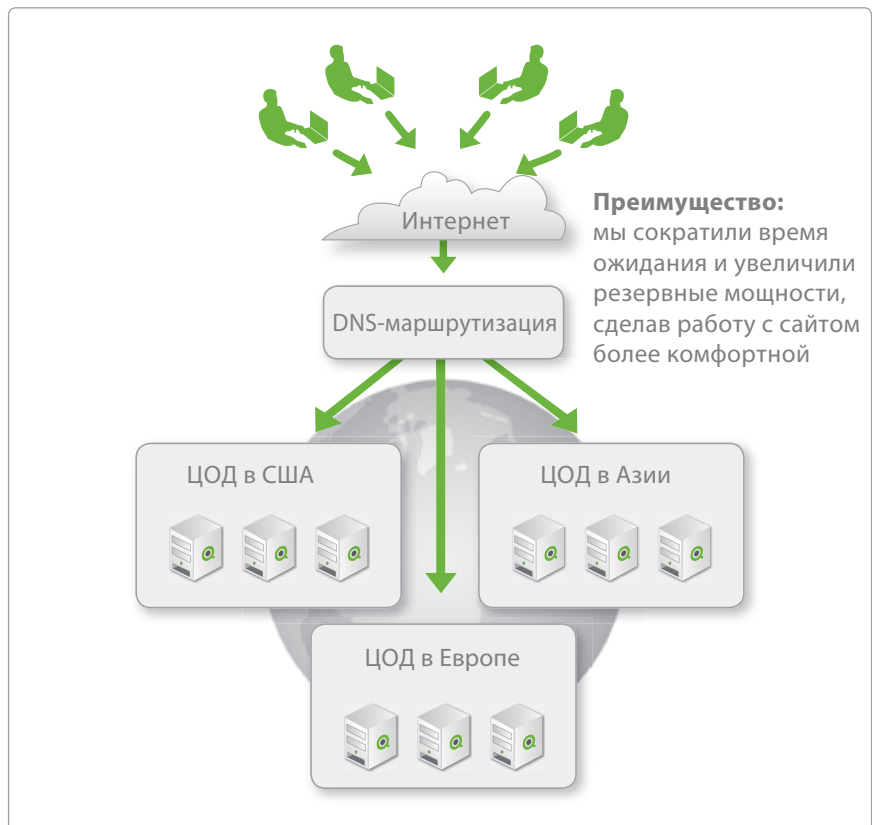
Источник: QlikTech

## МЫ РАЗМЕСТИЛИ СЕРВЕРЫ БЛИЖЕ К ПОЛЬЗОВАТЕЛЯМ

К весне 2010 года мы начали расширять охват наших маркетинговых кампаний. Нам пришлось учесть географический фактор, так как пользователям требовались комфортные условия для работы. Если бы все серверы QlikView были сосредоточены в одном ЦОД, американские пользователи, расположенные рядом с ЦОД Amazon EC2, получали бы наибольшие преимущества при обращении к демо-сайту, а тех, кто находился на другом конце земного шара, ожидало разочарование — функционирование системы оказалось бы слишком медленным. Подобное развитие событий могло оказать негативное влияние на развитие международного бизнеса QlikTech.

Решая эту проблему, мы распределили серверы географически (см. Рисунок 2) и создали три виртуальных кластера — в США, Европе и Азии. DNS-серверы перенаправляли пользователей к ближайшему серверу QlikView. При обращении пользователя к сайту demo.qlikview.com, главный DNS-сервер передавал его IP-адрес одному из двух других DNS-серверов, содержащих базу данных со всеми IP-адресами планеты с классификацией по странам. Затем DNS-сервер определял регион, соответствующий данной стране.

Рисунок 2: Мы разместили серверы ближе к пользователям



Источник: QlikTech

## МЫ ИСПОЛЬЗОВАЛИ БОЛЕЕ МОЩНЫЕ СЕРВЕРЫ И УВЕЛИЧИЛИ ИХ КОЛИЧЕСТВО

Весной 2010 года, готовясь к крупной маркетинговой кампании, мы внесли новые изменения (см. Рисунок 3). Поскольку предполагалось, что новая инициатива привлечет на демонстрационный сайт тысячи посетителей, где они будут знакомиться с приложениями QlikView, мы сделали следующее:

- **Увеличили количество серверов.** Минимальное количество серверов в каждом ЦОД было увеличено с трех до пяти, что позволило системе обслуживать большое количество одновременно работающих пользователей. Это изменение было проделано быстро, без привлечения внутренних ИТ-ресурсов.
- **Воспользовались аппаратными возможностями Amazon.** Мы перешли с «мощных» серверов на «сверхмощные» виртуальные машины Amazon, оснащенные 15 Гбайт памяти, 8 виртуальными ядрами и более 1 Тбайт дискового пространства. Серверы работали на 64-битной платформе.

Рисунок 3: Мы увеличили серверные мощности



Источник: QlikTech

## История успеха: «Kick It and Qlik It»

---

Весной 2010 года мы создали специальное приложение QlikView, предназначенное для поклонников футбола (см. Рисунок 4). Оно получило название «Kick It and Qlik It». Наши маркетологи задействовали PR-инструменты, маркетинговые возможности поисковых машин и собственный сайт, чтобы донести эту информацию до миллионов пользователей. В специальной базе данных мы собрали информацию о сотнях тысяч потенциальных посетителей и отправили им письма с предложением опробовать работу «Kick It and Qlik It». Было разослано примерно 750 тыс. электронных сообщений на разных языках.

- **Мы загрузили приложение и данные в оперативной памяти.** Для повышения производительности платформа QlikView загружает в оперативную память данные и пользовательский интерфейс приложения (например, поля списков, диаграммы и графики). Приложение «The Kick It and Qlik It» содержало статистику по игрокам и командам. Мы настроили обновление данных на серверах QlikView так, чтобы болельщики получали мгновенный доступ к самой свежей информации.
- **Мы поддерживали ПО на виртуальных машинах в «облаке».** ПО для серверов QlikView Server выполнялось не на физических серверах, а на виртуальных машинах, размещенных в ЦОД Amazon EC2. Мы могли клонировать и группировать экземпляры приложения для обеспечения масштабирования. На одном из экземпляров размещался веб-сервер QlikView, обеспечивавший балансировку нагрузки на серверы приложений с помощью сервиса CloudWatch от Amazon.
- **Мы осуществляли автоматическое масштабирование.** Каждые 90 секунд служба CloudWatch оценивала производительность, текущее использование ресурсов и тенденции изменения спроса, включая загрузку процессоров, обращения к дискам и трафик, а затем определяла потребности в масштабировании. При выполнении определенных условий CloudWatch запускала новый экземпляр сервера QlikView. В течение нескольких минут веб-сервер QlikView перераспределял нагрузку между всеми виртуальными машинами в кластере.
- **Мы добились оптимальной производительности путем нагрузочного тестирования.** После проведения нагрузочного тестирования конфигурации с помощью таких инструментов, как BrowserMob, стало понятно, что необходимо настроить функцию автомасштабирования Amazon. Это позволит запускать новые экземпляры не за полчаса, как исходно планировалось, а за 15 минут. Кроме того, для обеспечения большей производительности мы решили отказаться от процедуры регистрации пользователей на серверах и веб-сервере QlikView. (При установке авторизация была включена по умолчанию.)



Рисунок 4: Приложение Kick It & Qlik It



Источник: QlikTech

## Инвестиции в масштабируемость окупилась

Географически распределенная облачная архитектура для пользовательского приложения QlikView оказалась успешной во всех отношениях.

- **Ранее ситуация была далека от оптимальной.** Периодические пики посещаемости приводили к отсутствию доступа или замедлению работы серверов, что вынуждало пользователей обращаться в службу поддержки.
- **Ключи к успеху: гибкость, распределенное размещение и высокая производительность серверов.** После того как мы перенесли в «облако» дополнительные серверы QlikView из трех регионов, посетители демонстрационного сайта перестали высылать сообщения о сбоях. Нам удалось справиться с пиковой нагрузкой, когда к системе «Kick It and Qlik It» получали доступ 8 тыс. пользователей в сутки.
- **Мы познакомили с QlikView миллионы людей.** Процент реагирования на нашу почтовую рассылку оказался на 15% выше, чем ожидалось. Почти 13 тыс. человек прошли регистрацию, чтобы получить доступ к приложению «Kick It and Qlik It». В среднем пользователи работали с системой более получаса — это гораздо больше обычного времени (нескольких минут), которое посетители проводят на сайтах других вендоров информационных систем.

## Рекомендации

---

Мы гордимся тем, что можем предоставлять нашу систему в виде «облачного» сервиса. Добиться этого было непросто. Мы извлекли много полезных уроков, получили необходимый опыт и теперь можем сформулировать несколько рекомендаций:

- **Определите, какие серверы нужно масштабировать.** В нашем случае требовалось масштабирование серверов приложений, а не веб-сервера QlikView. Основная нагрузка должна приходиться на серверные ресурсы.
- **Испытайте конфигурацию с помощью нагрузочного тестирования.** Своевременное тестирование исходной архитектуры показало, что необходимо менять конфигурацию. Повторение тестов в конце процесса подтвердило, что мы на верном пути.
- **Размещайте серверы ближе к пользователям.** Если ваши пользователи географически распределены, серверы и веб-серверы QlikView должны находиться максимально близко к ним. Не теряйте производительность из-за сетевых ограничений, которые существуют на другом конце земного шара.
- **Отключите регистрацию пользователей.** Процесс регистрации сокращает число пользователей, которые могут одновременно подключаться к серверу, поэтому мы рекомендуем отключать авторизацию в приложениях и обслуживать пользователей анонимно.

## Приложение

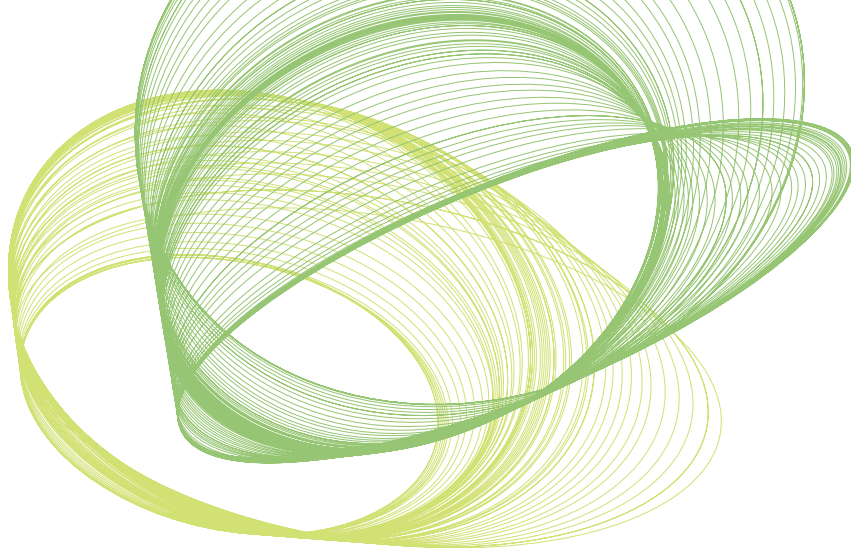
---

### **Информационные материалы по технологиям QlikView**

Обзор архитектуры QlikView, октябрь 2010 г.

Ассоциативный анализ: подавляющее превосходство QlikView, октябрь 2010 г.

# QlikView



[www.qlikview.ru](http://www.qlikview.ru)

---

© 2010 QlikTech International AB. Все права защищены. QlikTech, QlikView, Qlik, Q, Simplifying Analysis for Everyone, Power of Simplicity, New Rules, The Uncontrollable Smile и другие продукты и услуги компании QlikTech, а также визуальное отображение являются товарными знаками и зарегистрированными торговыми марками QlikTech International AB. Все другие названия компаний, продуктов и услуг, использованные в этом документе, являются товарными знаками или зарегистрированными торговыми марками соответствующих владельцев.

Сведения, представленные в документе, могут быть изменены без предупреждения. Данная публикация носит сугубо информационный характер и не содержит никаких гарантий. QlikTech не будет нести ответственности за возможные ошибки и неточности, содержащиеся в данной публикации. Исключительный перечень гарантийных обязательств на продукты и услуги QlikTech, если такие гарантии предусмотрены, приводится в соответствующих гарантийных условиях, сопровождающих продукты и услуги. Ничего из вышесказанного не может являться или быть трактовано как дополнительные гарантии.