

## **Тема 9. СОВРЕМЕННЫЕ ТЕНДЕНЦИИ РАЗВИТИЯ КИС**

### **Контрольные вопросы**

1. Что такое OLAP?
2. Что такое технологии BI?
3. Когда базу данных можно отнести к Big Data?
4. Что такое «набор V» в Big Data?
5. Какие технологии обработки данных используются в Big Data?
6. Какие категории BI-продуктов Вам известны?
7. Какие технологии искусственного интеллекта применяются при построении BI-систем?
8. В чем заключается смысл виртуализации ресурсов?
9. Какие способы виртуализации Вам известны?
10. Как можно использовать виртуализацию при построении информационного пространства предприятия?
11. Какие имеются технологические основы для внедрения мобильных технологий в корпоративные информационные решения?
12. Какие элементы мобильных технологий могут использоваться в КИС?
13. Какие задачи приходится решать при внедрении мобильных технологий в КИС?
14. Дату определение корпоративного мобильного приложения.
15. Что такое облачные вычисления?
16. В чем достоинства и недостатки применения облачных технологий при реализации КИС?

### **9.1 Технологии анализа данных**

На рубеже 20 и 21 веков проводились активные работы по развитию технологий аналитической обработки данных с использованием корпоративных хранилищ данных. Многочисленные исследования в области аналитической обработки данных (OLAP – Online Analytics Processing) привели к созданию промышленных направлений обработки больших данных (Big Data) и интеллектуального анализа данных (BI – Business Intelligence).

Большие Данные, на сегодняшний момент, являются одним приоритетных направлений развития информационных технологий. За последние годы в ряде IT областей были накоплены огромные данные, которые стало невозможно обрабатывать традиционными способами. Принято считать, что база данных относится к виду базы Big Data, если:

- объем информации свыше нескольких петабайт;
- используется децентрализованный способ хранения информации;
- данные неструктурированы или частично структурированы, имеют слабую связь;
- используется горизонтальная модель масштабирования данных.

Рассматривая технологию Big Data, надо отметить ее характеристики, которые можно назвать «множество V»<sup>1</sup>:

- Volume – объем, накопленная база данных представляет собой большой объем информации, который трудоемко обрабатывать и хранить традиционными способами, для них требуются новый подход и усовершенствованные инструменты;
- Velocity – скорость, данный признак указывает как на увеличивающуюся скорость накопления данных, так и на скорость обработки данных (все более востребованными становятся технологии обработки данных в реальном времени);
- Variety – многообразие, т.е. возможность одновременной обработки структурированной и неструктурированной информации, имеющей разные форматы. Главное отличие структурированной информации – это то, что она может быть классифицирована. Примером такой информации может служить информация о клиентских транзакциях. Неструктурированная информация включает в себя видео, аудио файлы, свободный текст, информацию, поступающую из социальных сетей. Данная информация нуждается в комплексном анализе, чтобы сделать ее полезной для дальнейшей обработки;
- Veracity – достоверность данных, все большее значение пользователи стали придавать значимости имеющихся данных. Так, у интернет-компаний есть проблема по разделению действий, проводимых роботом и человеком на сайте компании, что приводит в конечном счете к затруднению анализа данных;
- Value – ценность накопленной информации. Большие Данные должны быть полезны компании и приносить определенную ценность для нее. К примеру, помогать в усовершенствовании бизнес-процессов, составлении отчетности или оптимизации расходов.

К наиболее распространенным подходам обработки Больших Данных можно отнести:

NoSQL (Not Only SQL – не только SQL). Включает в себя ряд подходов, направленных на реализацию базы данных, имеющих отличия от моделей, используемых в традиционных, реляционных СУБД. Их удобно использовать при постоянно меняющейся структуре данных. Например, для сбора и хранения информации в социальных сетях.

MapReduce – модель распределения вычислений. Используется для параллельных вычислений над очень большими наборами данных (петабайты и более). В программном интерфейсе не данные передаются на обработку программе, а программа – данным. Таким образом, запрос представляет собой отдельную программу. Принцип работы заключается в последовательной обработке данных двумя методами Map и Reduce. Map выбирает предварительные данные, Reduce агрегирует их.

Hadoop – используется для реализации поисковых и контекстных механизмов высоконагруженных систем. Отличительной особенностью

---

<sup>1</sup> В ранних работах можно встретить термин 3V: Volume, Velocity, Variety.

является то, что система защищена от выхода из строя любого из узлов кластера, так как каждый блок имеет, как минимум, одну копию данных на другом узле.

SAP HANA – высокопроизводительная NewSQL платформа для хранения и обработки данных. Обеспечивает высокую скорость обработки запросов. Еще одним отличительным признаком является то, что SAP HANA упрощает системный ландшафт, уменьшая затраты на поддержку аналитических систем.

Сегодня под термином Business Intelligence понимают инструменты для анализа данных, построения отчетов и запросов, которые могут помочь бизнес-пользователям преодолеть сложности с обработкой, интерпретацией и представлением данных для того, чтобы синтезировать из них значимую информацию. Эти инструменты в совокупности попадают в категорию, называемую «инструменты бизнес-интеллекта» (Business Intelligence Toolware).

Сегодня категории BI-продуктов включают: BI-инструменты и BI-приложения. BI-инструменты можно разделить на следующие виды:

- генераторы запросов и отчетов (Query/Report Generator — QRG);
- развитые BI-инструменты — прежде всего, инструменты оперативной аналитической обработки данных (Online Analytical Processing — OLAP);
- корпоративные BI-наборы (Enterprise BI Suites — EBIS) различной конфигурации, встраиваемые в ERP-системы;
- BI-платформы.

Многомерные OLAP-серверы, а также реляционные OLAP-механизмы являются BI-инструментами и инфраструктурой для BI-платформ, на базе которых разрабатываются разнообразные приложения с «заказными» пользовательскими интерфейсами. Указанные инструменты применяются для доступа к данным, их многомерного и многофакторного анализа и генерации отчетов по данным, которые чаще всего располагаются в различных витринах (оперативных складах), базах или хранилищах данных. В качестве примера BI-приложения можно указать информационную систему (подсистему) поддержки деятельности руководителя (Executive Support System — ESS).

BI-приложения обычно ориентированы на конкретные важные функции организации, такие как анализ тенденций рынка, анализ рисков, анализ и прогноз продаж, планирование бюджета и т.п. Они могут применяться и более широко — для построения систем сбалансированных показателей (Balanced Scorecard System) или для управления эффективностью предприятия в целом (Enterprise Performance Management).

BI-системы строятся на основе нечеткой логики, нейронных сетей, технологии активных агентов, генетических алгоритмов и их симбиозов.

## **9.2 Системы виртуализации**

В настоящее время все большую популярность набирают технологии виртуализации. В результате развития технологий производства компьютерной

техники увеличивается количество процессорных ядер, растет пропускная способность интерфейсов компьютеров, емкость и скорость систем хранения данных. В результате появляется возможность предоставления вычислительных ресурсов сервера (или логического объединения серверов) виртуальным вычислительным процессам. Наиболее распространенным примером является запуск нескольких операционных систем на одном сервере виртуализации (или кластере серверов виртуализации). Если рассматривать единое информационное пространство предприятия, то возникает возможность запуска различных серверных операционных систем на кластере виртуализации, что существенно снижает стоимость владения информационными ресурсами. Таким образом, технологии виртуализации в настоящее время становятся одним из ключевых компонентов современной ИТ-инфраструктуры крупных предприятий. Сейчас уже сложно представить построение нового серверного узла компании без использования технологии виртуализации. Определяющими факторами такой популярности, несмотря на некоторые недостатки, можно назвать экономию денег и времени, а также высокий уровень безопасности и обеспечение непрерывности бизнес-процессов.

Для осуществления виртуализации используется ряд способов, с помощью которых достигаются одинаковые результаты через разные уровни абстракции. У каждого способа есть свои достоинства и недостатки, но главное что каждый из них находит свое место в зависимости от области применения.

Можно считать, что самая сложная виртуализация обеспечивается эмуляцией аппаратных средств. В этом методе виртуальная машина (VM) аппаратных средств создается на хост-системе, чтобы эмулировать интересующее оборудование.

Другое использование эмуляции – это эмуляция оборудования, которая заключается в совместном развитии встроенного программного обеспечения и аппаратных средств. В этом методе VM аппаратных средств создается на хост-системе, чтобы эмулировать интересующее оборудование.

Существует большое множество аппаратных и программных решений в рассматриваемой области. Одним из перспективных направлений является создание распределённых систем виртуализации повышенной отказоустойчивости. Для построения таких систем можно использовать установку систем виртуализации поверх DRBD систем.

DRBD (Distributed Replicated Block Device) – распределённое реплицируемое блочное устройство, предназначенное для построения отказоустойчивых кластерных систем на операционной системе Linux. DRBD выполняет зеркалирование (mirroring) по сети всех операций с блочным устройством. Условно можно считать DRBD сетевым RAID1.

Принцип работы DRBD заключается в том, что DRBD берёт данные, записывает их на локальный диск и пересылает на другой хост. На другом хосте они тоже записываются на диск.

Обычно DRBD-устройство работает на одном из узлов в режиме первичного (primary role), а на втором — в режиме вторичного или резервного (secondary role). Запись идёт на устройство, которое находится в режиме

главного, а на второе просто выполняется репликация. Такой режим применим для классических отказоустойчивых кластеров, его следует использовать, если на DRBD-устройстве непосредственно находятся традиционные, не кластерные файловые системы (ext3, XFS, JFS и т.д.).

Начиная с DRBD-8.0.08 можно заставить работать оба узла в режиме primary. Это даёт возможность монтировать кластерную файловую систему сразу на двух узлах одновременно.

Как говорилось выше, на DRBD устройства можно установить виртуальные операционные системы, управление которыми осуществляет гипервизор (например монитор виртуальных машин Xen).

### **9.3 Мобильные технологии**

Современное развитие мобильных технологий обусловило их широкое применение в корпоративных информационных системах. Карманные компьютеры, смартфоны и планшеты все больше проникают в корпоративную среду и превращаются для многих руководителей и сотрудников из личного мобильного устройства в рабочий инструмент. Сейчас многие компании и организации начинают централизованно внедрять и использовать мобильные технологии в интересах бизнеса.

Технологическими основами внедрения мобильных технологий в бизнесе являются:

- технологии мобильного интернета, позволяющего пользоваться услугами глобальных сетей и сервисов через сотовую связь;
- беспроводные технологии доступа не использующие сотовую связь (Wi-Fi, Bluetooth, ...);
- технологии IP-телефонии и телеконференций;
- технологии геолокации;
- технологии получения и обработки аудио, фото и видео информации;
- и многое другое.

В качестве элементов мобильных технологий могут применяться:

- электронные календари с возможностью обмена информацией между календарями учетных записей (Microsoft, Google, ...) и календарями корпоративных систем планирования и управления бизнес процессами;
- мобильные приложения IP-телефонии и видеоконференций (Skype, WhatsApp, Viber, ...);
- офисные мобильные приложения;
- специально разработанные корпоративные мобильные приложения.

При внедрении мобильных систем и технологий в корпоративную информационную систему приходится решать следующие задачи:

- интеграция мобильных устройств в ИТ- инфраструктуру предприятия;
- обеспечение безопасной работы мобильных устройств корпорации;

- разработка и адаптация работы бизнес и web приложений для мобильных устройств;
- обеспечение технической поддержки мобильных устройств корпорации.

Как становится ясно из перечисленных задач, внедрение мобильных технологий предусматривает не только разработку новых мобильных приложений корпорации, но и адаптация интерфейсов существующих информационных систем, прежде всего использующих web интерфейс.

Как известно, управление контентом сайта крупного предприятия является сложной задачей, которая классически автоматизируется посредством использования систем управления контентом CMS (Content Management System). С учетом появления мобильных версий сайтов в последнее время стали разрабатываться CMS, способные транслировать содержание стандартного web контента в контент сайта для мобильных устройств.

На сегодняшний день актуальной становится задача одновременного управления контентом как минимум трех приложений: стандартного web сайта, мобильного сайта и корпоративного мобильного приложения.

Таким образом, становится актуальным вопрос: «Что считать корпоративным мобильным приложением?»

Учитывая, что, как правило, корпоративные приложения подразумевают работу с большим числом данных в многостраничном режиме, то, на первый взгляд, логичным было бы разработка одного мобильного приложения с большим числом рабочих экранов. Однако, анализ тенденции развития мобильных приложений показывает, что желательна разработка ряда мобильных приложений, которые в совокупности «закрывают» функционал стандартного корпоративного приложения.

Таким образом, перспективные CMS должны уметь транслировать структуру корпоративного сайта не только на мобильную версию, но и на совокупность мобильных приложений. При этом эти системы должны предусматривать управление не только контентом, но единым оформлением всех видов приложений.

#### **9.4 Применение облачных технологий при построении информационного пространства предприятия**

Классический подход к автоматизации бизнес-процессов, сложивший в период 2000-2010 годов, означает, что организация должна обладать пулом серверного оборудования для предоставления сервисов для информационного обеспечения и надежной защиты её основных процессов. Очевидно, что для такой инфраструктуры необходимы дорогостоящие серверы и сетевое оборудование, включающее маршрутизаторы и файерволлы для обеспечения безопасности, а также клиентские станции для пользователей услуг и соответствующее программное обеспечение. Очевидно также, что многие

организации, особенно в сфере малого и среднего бизнеса, не могут позволить себе такой подход к организации информационной инфраструктуры.

Тенденция последних лет, ознаменовавшихся повсеместным распространением облачных сервисов и облачных способов хранения и обработки данных, открывает новые возможности для автоматизации и сопровождения бизнес-процессов.

Облачные вычисления (Cloud Computing) — это модель обеспечения сетевого доступа по требованию к пулу конфигурируемых вычислительных и информационных ресурсов, например, сетям передачи данных, серверам, устройствам хранения данных, приложениям и сервисам.

Такой подход предоставляет массу удобств пользователям, сокращает время ожидания и доступа к ресурсу, делает пользователя независимым в выборе ресурсов. Однако появляется и ряд новых проблем. Не все требуемые сервисы удастся в ближайшие годы перенести в облака. Есть ряд ограничений, которые пока не удаётся обойти. Требование к ширине канала передачи данных, к защите и шифрованию данных, отсутствие возможности работы в виртуальной среде и среде гипервизора — список можно продолжать.

Именно поэтому в настоящее время активно развивается новый подход к организации данных и работе с ними, на основе которого проектируются гибридные системы для информационного обеспечения процессов организации. В частности, предлагается постепенно, в ходе модернизации, перестраивать существующую информационную инфраструктуру, в инфраструктуру, построенную по принципу частного облака.

Это позволяет получить результат в виде быстрой миграции между публичным и частным облаком. Например, сегодня очень востребованы технологии совместной, распределённой работы. Одним из инструментов, позволяющим организовать такую работу, является корпоративный или образовательный портал. Такой портал позволяет организовать распределённую работу. Если инфраструктура будет построена по принципу частного облака, то для этого потребуется всего лишь копирование виртуальных машин в публичное облако. В случае классического подхода к формированию информационной инфраструктуры, проект миграции включал бы в себя развёртывание новой инфраструктуры и перенос старых данных в новую среду, что является достаточно трудоёмким проектом.

Такое решение может быть эффективно интегрировано с существующей сетью организации. Оно может быть реализовано в среде гипервизора, что позволяет эффективно управлять кластером виртуальных машин, при потребности переносить их в облако.

Таким образом, уже сегодня возможно построение гибридного решения — портала на корпоративном сервере и рабочих приложениях и массивов данных в частном облаке, с возможностью масштабирования в публичное облако, используя предложенную методику. Гибридный подход к формированию информационной инфраструктуры — подход, при котором часть ресурсов, для которых это целесообразно, выносится в публичное облако, а часть наиболее критических, бизнес-значимых сервисов остаются в пределах информационной инфраструктуры компании.

Такая информационная инфраструктура должна строиться по принципам частного облака, используя технологии виртуализации, для обеспечения возможности миграции сервисов, при необходимости, в оба направления. При использовании такого подхода, сервисы следует проектировать на основе одинаковых принципов, для организации единых подходов к обеспечению информационной безопасности.

Для обеспечения безопасной работы при формировании гибридной инфраструктуры следует придерживаться следующих основных правил:

- контроль строгого соблюдения политик информационной безопасности в распределенных частях организации;
- допуск к настройке сервисов в облаке или гибридном решении только профессионалов;
- согласование локальных политик и гарантий безопасности на уровне пользователей, сетевых провайдеров и владельцев сервисов;
- дополнительное сохранение данных локально, там, где это возможно;
- учёт значимости конкретного информационного ресурса предприятия для бизнеса в целом;
- учёт новых угроз при переносе любого контента во внешние сервисы.

Подход частного облака для организации инфраструктуры внутри компании, позволит обеспечить высокую доступность и отказоустойчивость таких сервисов, используя технологии виртуализации. С одной стороны, решение остаётся безопасным, с другой стороны, становится эффективным и масштабируемым. Проектируемая инфраструктура становится распределённой, что при грамотной её организации, снижает риски для предприятий в случае возникновения каких-либо чрезвычайных ситуаций.