Применениепередовых методов в защитеинформации в создаваемой Государственной информационнойсистеме“Электронное управление госорганов Кыргызской Республики”

АсаналиевБакытОмушевич
эксперт по IT
т: 0771802086, E-mail: omushtegin@gmail.com

В мире с развитием ИКТ многократно возрастают возможности по добыванию, сбору, обработке, хранению, поиску, отображению и передаче информации. Параллельно и столь же интенсивно разрабатываются способы и средства уничтожения, искажения, несанкционированного доступа к информации, ее блокирования, а также нарушения функционирования информационных систем, которые относятся к наиболее уязвимым элементам инфраструктуры. В связи с этим многократно повышается роль и значение информационной безопасности в информационных системах государственного и ведомственного управления.

В нашей республике Постановлением правительства Кыргызской Республики №651 от 17 ноября 2014 года утверждена Программа правительства Кыргызской Республики по внедрению электронного управления («электронное правительство») в государственных органах исполнительной власти и органах местного самоуправления Кыргызской Республики на 2014-2017 годы. Для достижения целей впрограммеприоритетные задачи сформулированыиз 7 пунктов,но безопасность информации не рассмотрена должным образом, а стоит на второстепенном плане (1).

Если сравнитьГИС состроительным объектом, товстроительствена первом планеуделяется внимание набезопасностьзданий, учитывая все возможные угрозы.Безопасность зданий это безопасность людей живущих в нем.Аналогично в ГИС безопасность информации это безопасность государства, безопасность людей живущих в этом государстве.

Информационная безопасность не сводится исключительно к защите от несанкционированного доступа информации и шифрования данных, это принципиально более широкое понятие.

Из этого определения можно сделать вывод, что защита информации - довольно емкая и многогранная задача, охватывающая не только определение необходимости защиты информации, но и способы защиты, от чего защищать, когда защищать, чем защищать и каким должна быть эта защита. В этом контексте разработка комплексной системы защиты информации для определённого типа информационных систем является на сегодняшний день актуальной проблемой.

В современном мире длясоздания технологии защиты информации для конкретнойИССначала определяютполитику безопасности

информации.Политика безопасности ИС строится из требований Госстандартабезопасности информации.

Например, британский стандарт безопасности информации BS 7799 первая часть. BS 7799 Part 1 — CodeofPracticeforInformationSecurityManagement(Практические правила управления информационной безопасностью) описывает 127 механизмов контроля, необходимых для построения системы управления информационной безопасностью.Этот документ служит практическим руководством по созданию СУИБ(2). Не буду останавливаться на других стандартах, принятых в разных странах.Разработка таких документовна государственном уровне в нашей республикена сегодня актуальна.

В большинствеслучаяхв разрабатываемых информационных системах не учитывается безопасность информации или остается на задний план.Нельзя начинать техническую разработку, не имея тщательно проработанную систему информационной защиты. Если начинать с решения наиболее очевидных задач, не обращая внимания на потенциально существующие риски, то такая система будет непрерывно находиться в стадии разработки и переделки или не будет долго существовать.

В следующей части, не углубляясь в технических терминах и деталях,я хочу кратко изложить свою точку зрения как построить безопасную среду в создаваемойГИС “Электронное управлениегосорганов Кыргызской Республики”.

Задача, которую должны выполнять большинство ГИС - это хранение и обработка данных, обладающих разной структурой. Поэтому в основе большинства информационных систем лежит среда хранения и доступа к данным. Среда должна обеспечить уровень надежности хранения, отказоустойчивости и эффективности доступа.

Реляционные системы управления базами данных (РСУБД) стали сегодня основным инструментом создания информационных систем.РСУБД хранит данные в таблицах с жестко определенной схемой, таблицы состоят из строк и столбцов, между таблицами существуют отношения. Большинство ИС, созданных на основе РСУБД, работающих с данными малого и среднего объема, остаютсявне конкуренции по простоте использования, гибкости и широте возможностей.Однако, когда обработка информации будет иметь дело с большими объемами данных удобства РСУБД будет сопряжены огромными потерями производительности, а механизмыраспределения данных будут чрезвычайно затруднены.

Простой аналог: РСУБД можно сравнит с почтальоном нанимающего обоздля доставки грузов клиентам.Но вскоре выясняется, что с увеличением грузауменьшаетсяскорость передвижения обоза.Почтальону придется увеличить мощность, то есть добавлятьбыков. Управление обозом с несколькими быками сопровождается рисками итрудностями, так как загруженный обоз в любое время может поломаться, некоторые быки могут выйти из строя, апочтальон должен своевременнонаходитьпосылку каждого клиента, принимать посылку для нового адресата и хранить в удобном месте.

С увеличением заказа почтальону придется нанимать еще других обозов и.т. Все эти факторы отрицательно воздействуют на эффективность работы почтальона,а когда заказы идут большими потоками, управлять и содержать будет сложна.

Мы живем в эпоху информационного взрыва, когда рост информации, доступный каждому человеку, превосходитвсякое воображение, когда каждый человек способен создавать мегабайты, гигабайты и даже терабайты информации, хранить её и передавать другим.

По данным ежегодного исследования "Цифровая вселенная" проводимого IDC темпы роста информации удваиваются каждые два года, к 2020 году она достигнет 44 триллион гигабайтов.

Наша республика не останется на стороне от мировой тенденции. 4 марта 2016 года Государственная регистрационная служба при Правительстве Кыргызской Республики сообщила, что завершается оцифровка адресных листов по всей республике. Будут переведены в цифровой формат 5 млн. адресных листов. Следующим шагом оцифровки будет свидетельства рождений школьников от 1-го до 11-го классов. После будут оцифровываться дети до 7 лет, а также свидетельства о браке. Это только вершина айсберга, при полном переходе в электронное управление цифровые документы достигнут несколько десятка миллионов и будут расти каждый день, объем электронных данных на языке компьютера увеличится досотни, сотни миллионов мегабайтов.

Для того,чтобы обрабатывать и хранить возрастающий объем данных применяется вертикальное и горизонтальное масштабирование. Вертикальное масштабирование подразумевает увеличение мощности компьютерных машин с увеличением объема информации. Но более крупные машины становятся все более и более дорогими, не говоря уже о том, что существует физический предел их укрупнения (в нашем примере почтовой обоз).

Создание все более и более мощных серверных машин не обязательно дает оптимальное решение крупномасштабных задач.Сегодня быстро набирает оборот альтернативный подход –объединение множества недорогих стандартных компьютеров в единую распределенную систему. Чтобы понять,почему распределенные системы (масштабирование по горизонтали) оказались популярнее монолитных серверов (масштабирование по вертикали)высокопроизводительному компьютеру с четырьмя каналами ввода/вывода со скоростью 100Мб/с каждый потребуется три часа, чтобы только прочитать набор данных объемом 4 ТБ. Если же воспользоваться четырьмя стандартными компьютерами объединенных в единую распределенную систему, то эти машины могут читать этот набор данных параллельно, достигая более высокой пропускной способности. При этом распределенная система, собранная из этихкомпьютеров, оказывается дешевле одной высокопроизводительной серверной машины.

Однако, концепция параллельного чтения и записи данных на нескольких дисках не так проста. Необходимо учесть сбои оборудования,

безопасности хранимой информации. Каким-то образом надо создать отношения между данными на разных компьютерах.

Рост информации с геометрической прогрессией первые почувствовалитакие гиганты как Google, Yahoo, Microsoft. Им приходилось обрабатывать терабайты, петабайты данных. Имеющие инструменты оказались не приспособлены к обработке столь больших объемов данных

Googlestала первой компанией, начавшей разработку программно-платформенную систему работающих с распределенными данными. Систему называли MapReduce, версию с открытым кодом -Hadoop.

Ныне Hadoop составляет основную часть вычислительной инфраструктуры многих работающих веб гигантовне толькоGoogle, Yahoo, Facebook,Twitter, но и других более традиционных информационных систем, имеющих дел с большими данными.

Hadoop предназначен для создания и запускараспределенных программных приложений, для обработки большого объема данных и состоит изпервичногоузла имен, вторичного узла имен, резервного узла имен и узлов данных. Каждый узел работает на отдельной компьютерной машине.

Рассмотрим в контексте безопасность информации:

Узел имен – это в своем роде дирижер оркестра, который управляет всеми действиями, он знает всех данных где, что лежит.Узел имен также знает, на каких узлах данных хранятся все блоки заданного файла.

Узлы данных — это основная «рабочая сила» файловой системы. Они читают и записывают блоки (по требованию клиентов или узла имен), а также периодически передают узлу имен список сохраняемых ими блоков.

Без узла имен файловая система становится неработоспособной, как без дирижера оркестр. Более того при уничтожении машины, на которой работает узел имен, все файлы в файловой системе будут потеряны, потому что восстановить их по блокам узлов данных будет невозможно. Для этогорассмотрены 2 механизма решения этой задачи.

Hadoop можно настроить таким образом:

1. Чтобы узел имен записывал свое устойчивое состояние в несколькихузлах данных.Вторичный узел имен, работающий на отдельном компьютере,хранит все образы пространства имен узлов данных, который может использоваться в случае сбоя основного узла.
2. Что резервный узел, записывает все действия первичного и вторичного узлов имен, который в случаях сбоя первичного и вторичного узла запускается моментально.

Узлы данных другдруга дублируют, в случаях отказа некоторых узлов данных другие выполняют эти задачи. Например:принцип такойцифра 56789- напервом узле данных сохраняются 5,6,7, на втором узле данных - 7,8,9, на третьем-9,5,6и т.д. Как видно, каждая цифра дублируется несколько раз. Местонахождение конкретных данных никто не сможет знать, данные постоянно перемещаются и шифруются.Кроме этого каждый узел защищён собственной защитой. Например: в традиционных серверах злоумышленник взломает сервера данных и все. А в Hadoop, чтобы пробраться до

определенного файла придется взломать несколько десятков, а то и сотни компьютеров, это невозможно (3), (4).

Использование современных технологийкак Hadoop в создании государственной информационной системы “Электронное управление госорганов в КР” решает несколько задач:

1. Безопасность хранения и эффективность доступа данных.
2. Отказоустойчивость. АрхитектураHadoop разработана с учетом возможности частых отказов.
3. Скорость. Операции выполняются параллельно нанескольких компьютерах.
4. Работа с большими объемами данных.Данные масштабируются горизонтально, можно с легкостью добавлять, удалять узловые машины.
5. Дешевизна. Поскольку Hadoop работает на стандартных компьютерах,отпадает вопрос приобретения и обслуживания дорогих серверных оборудований.
6. Надежность и простота.Оборудования со временем морально устаревают, поэтому их нужно заменять на более новые и совершенные. Замена больших серверов сопровождается определенными сложностями. Замена узловых машин не составляет особого труда, их легкостью можно удалять, добавлять, заменять.

В конце хочу сказать, что на начальном этапе для внедрения электронного управления госорганов достаточно 15-20 стандартных компьютеров с жестким диском объемом 1 Тб. Со временем их можно увеличить. С точки зрения безопасности,чтобы уберечь от непредвиденных угрозцелесообразно их размещать на разных местах.

Излагать все возможностиHadoopна одном или на двух страницах невозможно, в специализированной литературе полностью описаны мощь и прелесть этого инструмента.

Использованная литература

1. Программа Правительства Кыргызской Республики по внедрению электронного управления («электронное правительство») в государственных органах исполнительной власти и органах местного самоуправления Кыргызской Республики на 2014-2017 годы - Бишкек 2014 г.
2. Википедия
3. Том Уайт. Подробное руководства “Hadoop” --Питер 2013
4. ЧакЛем. Hadoop в действии – издательство ДМК Москва 2012
5. www.kabar.kg