

Орлов Александр Александрович

Магистрант НАЧОУ ВПО СГА

Направление: Информатика и вычислительная техника

Магистерская программа: Распределённые автоматизированные системы

Исследование распределенной системы хранения данных и оценка ее временных параметров

Аннотация. Статья посвящена исследованию распределенной системы хранения данных. В статье отражены основные технологии построения системы хранения данных, ее компоненты и инфраструктура. Особый интерес представляла оценка уровня производительности системы хранения данных. В качестве методики оценки временных параметров распределенной системы хранения данных было выбрано тестирование.

Ключевые слова: RAID-массивы, протокол iSCSI, технология Multipath I/O (MPIO), гипервизор, тестирование.

Глобальные объемы информации и рост ее ценности и значимости привели к необходимости создания и использования систем хранения данных, ставших залогом успешной работы компаний, осуществляющих деятельность в сфере информационных технологий и системной интеграции. Объектом исследования выбрана распределенная система хранения данных одной из динамично развивающихся компаний Ивановской области, в которой два узла системы хранения данных (СХД) и три сервера поддерживают Ethernet интерфейс и стек протокола iSCSI. В основу системы хранения данных заложена практика защиты информации на базе технологии RAID. В исследуемой СХД реализуются аппаратные RAID-массивы, отличающиеся от программных большей надежностью и минимальной нагрузкой на процессор и системную шину. Четыре дисковых массива СХД (по два в каждом узле)

работают в RAID 6. Узлы системы хранения данных реплицируются между собой.

Организация сетевых соединений принадлежит Ethernet коммутатору HP Switch with Premium Software с 8-портовым модулем. Для уменьшения нагрузки на центральный процессор и увеличения производительности на длинных расстояниях на коммутаторе используются Jumbo-кадры. В каждом из двух узлов системы хранения данных по одному серверному двухпортовому адаптеру Intel Ethernet и по два RAID контроллера с подключенными накопителями на жестких магнитных дисках.

Сетевые соединения реализуются посредством Storage area network (SAN). Сети хранения данных SAN – это сети, предназначенные для обеспечения подключения серверов к устройствам хранения (дискам, RAID-массивам и ленточным библиотекам) и обмена данными между ними, обычно на уровне блоков [1, с. 3]. Информационная сеть высокой производительности SAN в исследуемой системе организована на основе технологии iSCSI через существующую сетевую инфраструктуру 10Gb Ethernet. В сети Ethernet все компьютеры имеют доступ к общей шине, поэтому она может быть использована для передачи данных между любыми двумя узлами сети [2, с. 93]. Стандартизованный по RFC 3720 протокол iSCSI разрешает передавать SCSI-команды по сети.

Создавая логические пути «сервер-устройство хранения данных», технология многопутевого подключения Multipath I/O (MPIO) предоставляет возможность использования нескольких каналов ввода-вывода для повышения отказоустойчивости СХД. Один из компонентов MPIO в Windows Server, модуль DSM, предоставляет различные политики многопутевого ввода-вывода Microsoft (MPIO).

Технология виртуализации СХД позволяет серверу «видеть» дисковые массивы единым устройством с совокупной дисковой ёмкостью и функционалом. Совместно используемое хранилище для виртуальных машин, используемых в гипервизорах VMware и Hyper-V, создается с помощью

программного обеспечения StarWind iSCSI SAN & NAS, являющегося программной платформой виртуализации для построения iSCSI SAN сети и превращающего стандартный Windows Server в надежное хранилище данных.

Гипервизор предоставляет операционным системам, работающим под его управлением, сервис виртуальной машины. Гипервизор позволяет на одном сервере параллельно запускать несколько операционных систем и обеспечивать их изоляцию друг от друга.

Следующим этапом исследования СХД была оценка уровня ее производительности. В качестве методики оценки временных параметров распределенной системы хранения данных было выбрано тестирование быстродействия массивов и скоростных характеристик сети, для чего использовались утилиты – вспомогательные компьютерные программы:

ATTO Disk Benchmark для тестирования производительности жестких дисков и RAID-массивов, позволяющая выбирать размер блока для записи/чтения;

IOzone, тестирующая работу физических носителей и файловых систем;

SQLIO – простое средство тестирования операций Input/Output (I/O) дисковой подсистемы для определения производительности при разных типах и объемах доступа к данным, находящихся на обычных дисковых массивах или в системе Storage Area Network (SAN);

Iperf3 – для тестирования пропускной способности канала связи, рекомендованная для определения скорости прохождения трафика между двумя узлами сети;

PCATTCP – консольная программа для измерения скорости локальной сети

Платформа для тестирования – сеть хранения данных с тремя видами связи: VMware – switch; switch – узлы СХД; репликация между дисковыми массивами. Для определения производительности СХД в целом и на отдельных участках каналов связи тестировались все группы связей между компонентами сети хранения данных:

1. Определение конечной производительности с целью получения скоростных характеристик сети хранения данных в целом;
2. Оценка скорости подключения к СХД, без использования файловой системы VMware;
3. Замер собственной производительности дисковых массивов без влияния параметров других компонентов систем хранения данных;
4. Тестирование производительности сетевой составляющей СХД (ВМ на Windows Server→первый узел СХД, ВМ на Windows Server→второй узел СХД) для определения ее полной (10GbE) или частичной загруженности.

Для определения временных и скоростных параметров посредством iSCSI target медленный дисковый массив подключался к VMware двумя способами: сначала через файловую систему VMware, затем для исключения влияния файловой системы VMware использовали прямой доступ с виртуальной машины VMware на один из узлов СХД.

Для тестирования быстродействия массивов был сформировали набор из одиннадцати тестов, в том числе три варианта подключения VMware к дисковым массивам, используя iSCSI инициатор, через сетевой коммутатор и три варианта подключения ВМ с Windows Server к узлам СХД.

Группа 1:

Тест 1: VMware к распределённому iSCSI target на первом дисковом массиве.

Тест 2: VMware к распределённому iSCSI target на втором дисковом массиве.

Тест 3: VMware к iSCSI target на медленном дисковом массиве второго узла СХД.

Тест 4: ВМ с Microsoft Windows к распределённому iSCSI target на первом дисковом массиве.

Тест 5: ВМ с Microsoft Windows к распределённому iSCSI target на втором дисковом массиве.

Тест 6: VM с Microsoft Windows к медленному дисковому массиву второго узла СХД.

Группа 2:

Тест 7: Прямое подключение iSCSI target в VM, через VMware, медленного массива из первого узла СХД (в обход использования файловой системы VMware).

Группа 3:

Тест 8: Тестирование первого дискового массива на первом узле СХД.

Тест 9: Тестирование второго дискового массива первого узла СХД.

Группа 4:

Тест 10: Сеть между VM с Windows Server и первым узлом СХД.

Тест 11: Сеть между VM с Windows Server и вторым узлом СХД.

Пиковая скорость передачи данных – это наибольшая скорость, которую разрешается достигать пользовательскому потоку в течение оговоренного небольшого промежутка времени T [3, с. 178]. Пришлось отказаться от глубины очереди и случайного чтения/записи, поскольку интерес представляли значения пиковой производительности. Каждый замер повторялся 4 раза для получения более точных результатов. При замере скорости в 5 тесте не представилось возможным получить скорость чтения и записи, так как ПО выдавало ошибку, в результате – отказались от ПО IOzone.

Набор переменных параметров в ПО, используемом в тестах:

размер блока передаваемых данных – 1024Б (приближённое к MTU 1500), 4096Б (стандартный размер блока данных NTFS), 8192Б (приближённое к MTU 9198), 512КБ (используемый размер блока RAID на дисках SAS), 1024КБ (используемый размер блока RAID на дисках SATA и один из используемых размеров блока VMFS), 4096КБ (один из используемых размеров блока VMFS), 8192КБ;

максимальный размер пакета TCP – 9198;

продолжительность теста – ограничена 1 минутой;

общий размер передаваемых данных в 2ГБ выбран для обеспечения нагрузки, близкой к пределу системы, исходя из-за предположения высокой производительности системы.

В рамках данной статьи не представляется возможным показать полученные результаты тестирования, которые были сведены в таблицы с последующей графической интерпретацией, поэтому позволю ограничиться кратким выводом.

Как правило, «узким местом» выступает дисковая подсистема, но исходя из результатов тестирования, стало понятно, что дисковый массив – вполне полноценное устройство, а слабым местом является сетевая составляющая. Все сети имеют свои предельные возможности, очевидно, что существует и некий предел сети исследуемой системы, не позволяющий дисковому массиву работать на полную мощность. Это подтверждают диаграммы, составленные на основе полученных результатов. На диаграмме тестирования первого дискового массива на первом узле СХД (Тест 8) максимальная запись в IOzone составляет 875,55 МБ/сек, а на диаграмме тестирования сети между ВМ с Windows Server и первым узлом СХД (Тест 10) было хорошо видно, что ширина полосы пропускания полосы 308,37 МБ/сек (ПО РСАТТСР). Тестирование исследуемой СХД за все время ее эксплуатации проводилось впервые. Практическая ценность полученных результатов в возможности использования их для дальнейшего исследования сети. Ограничение сети SAN по производительности вызвано, скорее всего, не самой сетью, а подключенными к ней устройствами. Чтобы подключенные к SAN устройства не стали «узким местом» сети и не ограничивали возможности дисковых массивов, необходимо проанализировать внешние и внутренние характеристики устройств, подключенных к SAN – это основание для нового исследования.

Литература

1. Джад Д. Основы проектирования SAN. Сан-Хосе: Infinity Publishing, 2008.

2. Кузин А.В. Компьютерные сети. М.: Инфра-М, 2011.
3. Олифер В., Олифер Н. Компьютерные сети. Принципы, технологии, протоколы. СПб.: Питер, 2010.

© Бюллетень магистранта 2015 год № 1