

# Анализ подходов к диагностике систем хранения данных

А. С. Макаров<sup>1</sup>, М. В. Болсуновская<sup>2</sup>, С. В. Широкова<sup>3</sup>, М. Б. Успенский<sup>4</sup>, А. А. Кузьмичев<sup>5</sup>

Санкт-Петербургский политехнический университет Петра Великого (СПбПУ)

<sup>1</sup>alexey.makarov@spbpu.com, <sup>2</sup>marina.bolsunovskaia@spbpu.com, <sup>3</sup>swchirokov@mail.ru,

<sup>4</sup>mikhail.uspenskiy@spbpu.com, <sup>5</sup>andrey.kuzmichev@spbpu.com

**Аннотация.** Для правильного выбора системы хранения данных требуется её тестирование, проверка того, подходит ли она по своим параметрам под требуемые задачи. Кроме того, диагностика необходима при эксплуатации. Данная статья представляет собой анализ и классификацию подходов к диагностике системы хранения данных, а также описание процесса тестирования и ожидаемые результаты.

**Ключевые слова:** система хранения данных; тестирование; диагностика; производительность

## I. ВВЕДЕНИЕ

Система хранения данных (СХД) – это комплексное программно-аппаратное решение по организации надёжного хранения информационных ресурсов и предоставления гарантированного доступа к ним. В настоящее время СХД повсеместно используются для хранения и обработки большого количества данных, а также в информационно-управляющих комплексах [1]. К хранению данных обычно предъявляются такие требования как: надёжность, скорость доступа к ним, возможность параллельного доступа к ним большого количества пользователей (серверов) и др. Для правильного подбора СХД может потребоваться тестирование нескольких моделей. Также проверка корректности работы СХД нужна уже непосредственно при эксплуатации. Целью данного исследования было классифицировать описанные в открытых источниках методы диагностики и тестирования СХД.

## II. МЕТРИКИ

Прежде, чем привести классификацию, следует выделить те метрики, которые измеряются при тестировании. Для оценки производительности СХД используют три основные метрики:

- Service Time, часто именуемый latency или response time, измеряется в миллисекундах и обозначает;
- IOPS – количество операций ввода-вывода в сек.;

---

Исследование выполнено при финансовой поддержке Министерства образования и науки Российской Федерации в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2014-2020 годы». Соглашение о предоставлении субсидии от 03.10.2017 г. № 14.581.21.0023. Уникальный идентификатор – RFMEFI58417X0025

- MB/s – количество переданных мегабайт в секунду [2].

Помимо основных метрик в тестах производительности SPC-1C от Storage Performance Council используются:

- общая емкость СХД, ASU (Application Storage Units, единицы хранения приложений);
- уровень защиты данных;
- оценочная конфигурация СХД (общая стоимость). [3]

Методика тестирования SPC-1C не описана в отчетах компании, поэтому её не представляется возможным включить в классификацию.

## III. МЕТОДЫ ДИАГНОСТИКИ

Проведенный нами анализ открытых источников позволил выделить два основных типа методов тестирования СХД: нагрузочное и функциональное. Кроме того, нами были рассмотрены другие типы тестов.

### A. Нагрузочные тесты

Основным источником для описания нагрузочных тестов стали статьи компании INLINE Technologies. На основе проведённого анализа можно сделать вывод, что в ходе нагрузочного тестирования обычно решаются следующие задачи:

- исследование процесса деградации производительности СХД при длительной нагрузке на запись (Write Cliff);
- анализ производительности СХД при различных профилях нагрузки;
- исследование влияния количества серверов, генерирующих нагрузку, на производительность СХД [4].

а) Тесты, реализующие длительную нагрузку типа random write.

Тесты выполняются посредством создания синтетической нагрузки на блочное устройство (Block Device), представляющее собой логический том типа stripe, созданный из N LUN (Logical Unit Number, виртуальный раздел в RAID-массиве) на тестируемой системе и предоставленных одному тестовому серверу.

По результатам тестов, формируются графики, совмещающие результаты тестов:

- IOPS, как функция времени.
- Bandwidth, как функция времени.
- Latency, как функция времени.

Проводится анализ полученной информации и делаются выводы о:

- наличие деградации производительности при длительной нагрузке на запись и на чтение;
- производительности сервисных процессов СХД (Garbage Collection) ограничивающих производительность дискового массива на запись при длительной пиковой нагрузке; [5]
- степени влияния размера блока операций ввода-вывода на производительность сервисных процессов СХД;
- объеме пространства, резервируемом на СХД, для нивелирования сервисных процессов СХД. [4]

Данная группа тестов состоит из четырех тестов, отличающихся суммарным объемом LUN, предоставленных с тестируемой СХД, размером блока операций ввода-вывода и направлением ввода-вывода (write или read):

- тест 1. Тест на запись, выполняемый на полностью размеченной СХД. Суммарный объем предоставленных LUN равен полезной емкости СХД;
- тест 2. Тест на запись, выполняемый на СХД, размеченной на 80%. Суммарный объем предоставленных LUN равен 80% от полезной емкости СХД;
- тест 3. Тест на чтение, выполняемый на полностью размеченной СХД;
- тест 4. Тесты на запись с изменяющимся размером блока: 1КБ, 8КБ, 16КБ, 32КБ, 64КБ, 1МБ, выполняемые на полностью размеченной СХД.

По результатам тестов формируются графики, совмещающие результаты двух тестов:

- IOPS, как функция времени;
- Latency, как функция времени. [6]

Также ряд тестов проверяет надежность, охватывающую 24-часовой период. Цель тестирования состоит в том, чтобы проверить, что хранилище может обрабатывать высокую нагрузку ввода-вывода в течение длительного периода времени. После этих стресс-тестов оба файла журнала и базы данных анализируются на предмет целостности:

- в файле журнала событий не должно быть обнаружено ошибок при проверке надежности хранения;
- не должно быть сообщений об ошибках для процесса определения контрольной суммы базы данных и журнала;
- должны отсутствовать сообщения об ошибках во время процесса резервного копирования на диск;

- не должно быть сообщений об ошибках для контрольной суммы базы данных в базе данных удаленного хранилища.

Первичное тестирование производительности СХД использует хранилище с максимальным устойчивым типом обмена ввода-вывода Exchange в течение двух часов. Тест показывает, сколько времени требуется, чтобы хранилище реагировало на ввод-вывод под нагрузкой. [7]

*b) Тесты производительности дискового массива при разных типах нагрузки.*

В ходе тестирования исследуются следующие типы нагрузок:

- профили нагрузки (randomrw, rwmixedread):
  - случайная запись 100%;
  - случайная запись 30%, случайное чтение 70%;
  - случайное чтение 100%.
- размеры блока: 1КБ, 8КБ, 16КБ, 32КБ, 64КБ, 1МБ;
- способы обработки операций ввода-вывода: асинхронный.

Тесты производятся в 3 этапа:

- для каждой комбинации перечисленных выше типов нагрузки путем вариации параметров numjobs и iodepth находится точка насыщения СХД, то есть такая комбинация jobs и iodepth, при которой достигается максимум iops, но задержка минимальна. Фиксируются показатели iops, latency, numjobs и qdepth;
- затем тесты проводятся аналогично предыдущему этапу, только определяется точка, при которой достигается приблизительно вдвое меньшая производительность;
- затем проводятся аналогичные тесты, ищется точка еще вдвое меньшей производительности.

Подобный алгоритм тестирования позволяет определить предельную производительность дискового массива при заданном профиле нагрузки, а также зависимость latency от нагрузки. [5]

По результатам тестов на основании данных формируется набор зависимостей для каждой комбинации следующих параметров нагрузки: профиля нагрузки, способа обработки операций ввода-вывода, глубины очереди, совмещающие в себе тесты с разными значениями блока ввода-вывода:

- IOPS, как функция от кол-ва процессов, генерирующих нагрузку;
- Bandwidth, как функция от количества процессов, генерирующих нагрузку;
- Latency (clat), как функция от количества процессов, генерирующих нагрузку. [4]

Так же можно выделить два вида нагрузки по другому принципу.

- LUN без дедупликации;
- LUN с дедупликацией. [8]

Возможен другой вариант нагрузочного теста, при котором генерируется нагрузка в 100 процессов случайного чтения, затем один процесс начинает выполнять операции 100% случайной записи блоками по 4 Кбайт, после чего к нему подключаются по одному дополнительные процессы. Зная максимальные значения и нужные нам соотношения читающих и пишущих процессов, мы можем варьировать нагрузку на СХД. В итоге получается весь диапазон производительности СХД данных при разной нагрузке. [9]

Кроме того, может генерироваться подражание активности ввода-вывода из четырех общих рабочих нагрузок бизнес-приложений:

- электронная почта. Утилита Microsoft Jetstress 2010 может быть использована для генерации трафика электронной почты;
- база данных: утилита Orion от Oracle может использоваться для генерации трафика базы данных;
- веб-сервер: стандартная утилита Iometer может использоваться для генерации трафика веб-сервера. Определение параметра ввода-вывода в данном случае состоит из случайных чтений различных размеров блоков.
- резервное копирование: утилита Iometer может использоваться для создания единого потока трафика с последовательным чтением большого блока.

Каждая из четырех рабочих нагрузок делается параллельно, при этом тест электронной почты Jetstress выполняется дольше всего (примерно три часа). [10]

Другие два типа тестов подобны первому, но с небольшими вариациями:

- тесты производительности дискового массива при разных типах нагрузки, генерируемой одним и двумя серверами на блочное устройство. [4]
- тесты производительности дискового массива при разных типах нагрузки, выполняемые на уровне блочного устройства. [11]

### В. Функциональное тестирование

В ходе функционального тестирования проверяется корректность работы некоторых функций СХД, таких как:

- SmartVirtualization (интеллектуальная гетерогенная виртуализация) при виртуализации LUN (не содержащих данных, содержащих данные);
- увеличение виртуализированных LUN;
- отказоустойчивость локальной дисковой системы;
- удаленная синхронная репликация СХД и др.

#### а) Тестирование функции SmartVirtualization при виртуализации LUN, не содержащих данных

Последовательность процесса тестирования:

- На СХД создаются 2 LUN, обычный и «тонкий», размером 50Г байт каждый;

- СХД конфигурируется для виртуализации этих LUN (инкапсулирования в eDevLUN);
- Средствами ПО управления дисковой системы производится проверка состояния внешних LUN (Health Status) и корректность отображения следующих параметров:
  - Capacity (емкость);
  - Device Manufacturer (производитель устройства);
  - Device Model (модель устройства);
  - Device WWN (World Wide Name, уникальный идентификатор, который определяет конкретное целевое устройство);
  - Thin LUN (Thin provisioning – это технология виртуализации СХД).
- Далее eDevLUN предоставляются серверу с установленным гипервизором;
- На eDevLUN создаются два Vmware Datastore (хранилища виртуальных машин) размером 50Гбайт. На Datastore создаются две виртуальные машины (по одной на каждом DataStore);
- Производится запуск и проверка корректности работы созданных виртуальных машин на операциях ввода-вывода (чтение и запись файлов).

Тест считается успешным, если виртуальные машины корректно запустились, и операции ввода-вывода выполняются без ошибок. [12]

#### б) Тестирование функции SmartVirtualization при виртуализации LUN содержащих данные

Последовательность процесса тестирования такая же, как в предыдущем пункте, но добавляются следующие пункты после первого:

- Оба LUN презентуются серверу с установленным гипервизором;
- На LUN создаются два Vmware Datastore. На Datastore создаются две виртуальные машины (по одной на каждом DataStore);
- Производится запуск и проверка корректности работы созданных виртуальных машин на операциях ввода/вывода (чтение и запись файлов);
- После проверки корректности работы, виртуальные машины выключаются, и производится отключение сервера от дисковой системы. [12]

#### в) Тестирование возможности увеличения виртуализированных LUN

Тестирование производится на уже созданных и виртуализированных ранее (в предыдущем тесте) обычном и «тонком» LUN.

- размер LUN на дисковой системе увеличивается с 50 до 60-ти GB;
- проверяется корректность отображения его параметров на СХД;
- на СХД производится увеличение соответствующего eDevLUN;

- производится увеличение размеров DataStore на увеличенных eDevLUN и запуск имеющихся на них виртуальных машин.

Тест считается успешным, если виртуальные машины корректно запустились. [12]

#### *д) Тестирование локальной отказоустойчивости дисковой системы*

Тестирование производится на двух eDevLUN, созданных в предыдущих тестах, и двух внутренних LUN СХД:

- логические диски предоставляются серверу с установленным гипервизором;
- на каждом логическом диске создается свой Datastore. На каждом Datastore создается по две виртуальные машины под управлением ОС RHEL;
- на виртуальных машинах эмулируется нагрузка на файловую систему с использованием ПО filebench (<https://github.com/filebench/filebench/wiki>).

Проверяется работоспособность системы под нагрузкой при следующих отказах:

- выход из строя одного из контроллеров дисковой системы;
- потеря одного произвольного внешнего соединения дисковой системы с хостом VMware;
- потеря двух соединений между контроллерами дисковой системы с контроллерами виртуализируемой дисковой системы.

Тесты считаются успешными, если в результате отказов виртуальные машины, использующие ресурсы дисковой системы, остались в работоспособном состоянии, и доступ к ним не прерывался. [12]

#### *С. Другие типы тестов*

Кроме того, были выделены другие типы тестов, которые применяются не так часто:

- тестирование работы удаленной синхронной репликации СХД [12];
- работа Easy Tier [13];
- тестирование дедупликации и сжатия [14];
- резервное копирование и восстановление баз данных [7];
- тесты базы данных [15].

### IV. Выводы

В результате анализа научных статей и источников, описывающих различные варианты тестовых испытаний СХД, были определены основные подходы к тестированию и диагностике, выявлены особенности и виды нагрузочного и функционального тестирования. Для каждого из видов тестирования описаны этапы процесса тестирования, входные и выходные параметры и условия соответствия положительному результату испытаний. Проведенное исследование позволило определить набор и обобщенную структуру тестовых испытаний СХД, необходимую для

разработки стенда для нагрузочного и функционального тестирования различных конфигураций СХД.

### СПИСОК ЛИТЕРАТУРЫ

- [1] Efremov, A.A., Loginova, A.V., Mikeladze, B.D., Shirokova, S.V. The models and technologies for supporting decision making in design of information-control complexes (Conference Paper) // Proceedings of 2017 20th IEEE International Conference on Soft Computing and Measurements, SCM 2017; Article # 7970742. PP. 846-848.
- [2] Анализ утилизации СХД. [Электронный ресурс]. URL: <https://habrahabr.ru/post/338438/>. – (Дата обращения: 24.12.2017)
- [3] Передовые практики тестирования производительности твердотельных дисков. Пояснение результатов тестирования производительности SPC-1C для твердотельных дисков. [Электронный ресурс]. URL: [https://www.seagate.com/files/www-content/product-content/pulsar-fam/\\_cross-product/ru/docs/bp-ssd-performance-tp623-2-1201ru.pdf](https://www.seagate.com/files/www-content/product-content/pulsar-fam/_cross-product/ru/docs/bp-ssd-performance-tp623-2-1201ru.pdf). – (Дата обращения: 24.12.2017)
- [4] Тестирование флеш СХД. Hitachi HUS VM с модулями FMD. [Электронный ресурс]. URL: [https://habrahabr.ru/company/inline\\_tech/blog/232309/](https://habrahabr.ru/company/inline_tech/blog/232309/). – (Дата обращения: 24.12.2017)
- [5] Агапкин, А. Тестирование флеш СХД. EMC XtremIO. [Электронный ресурс]. URL: <http://itg-td.blogspot.ru/2014/09/emc-xtremio.html>. – (Дата обращения: 24.12.2017)
- [6] Тестирование флеш СХД. IBM RamSan FlashSystem 820. [Электронный ресурс]. URL: [https://habrahabr.ru/company/inline\\_tech/blog/227887/](https://habrahabr.ru/company/inline_tech/blog/227887/). – (Дата обращения: 24.12.2017)
- [7] Hitachi Unified Storage 130 Dynamically Provisioned 8,000 Mailbox Exchange 2010 Mailbox Resiliency Storage Solution. [Электронный ресурс]. URL: <https://www.hitachivantara.com/en-us/pdf/white-paper/130-dynamically-provisioned-8000-mailbox-exchange-storage-solution.pdf>. – (Дата обращения: 24.12.2017)
- [8] Выгодная замена тиринговым массивам: обзор и тесты полностью флешевой железки HP ZPAR 7400. [Электронный ресурс]. URL: <https://habrahabr.ru/company/croc/blog/267395/>. – (Дата обращения: 24.12.2017)
- [9] Колганов, В. Флеш-СХД: все познается в сравнении. Storage News(4(60)). [Электронный ресурс]. URL: [http://www.storagenews.ru/60/CROC\\_Flash\\_test\\_60.pdf](http://www.storagenews.ru/60/CROC_Flash_test_60.pdf). – (Дата обращения: 24.12.2017)
- [10] Garrett, B. Lab Validation Report. Dell PowerVault MD36X0f Series. Mixed Workload Performance with Application Aware Data Management. [Электронный ресурс]. URL: <http://i.dell.com/sites/doccontent/shared-content/data-sheets/en/Documents/ESG-Lab-Validation-Dell-PowerVault-MD36X0f-Series-Jun-11-1.pdf>. – (Дата обращения: 24.12.2017)
- [11] Агапкин, А. Тестирование флеш СХД. Huawei Dorado 2100 G2. [Электронный ресурс]. URL: <http://itg-td.blogspot.ru/2015/08/huawei-dorado-2100-g2.html>. – (Дата обращения: 24.12.2017)
- [12] Агапкин, А. Решения для замещения. Тестирование возможностей СХД Huawei OceanStor V3. [Электронный ресурс]. URL: <http://itg-td.blogspot.ru/2015/04/huawei-oceanstor-v3.html>. – (Дата обращения: 24.12.2017)
- [13] Обзор и тестирование системы хранения данных Lenovo V3700 V2. [Электронный ресурс]. URL: [http://www.thg.ru/storage/lenovo\\_v3700\\_v2\\_obzor/onepage.html](http://www.thg.ru/storage/lenovo_v3700_v2_obzor/onepage.html). – (Дата обращения: 24.12.2017)
- [14] Garrett, B. ESG Lab Review. Platform-and Data-efficient Dell Storage SC Series. [Электронный ресурс]. URL: <http://i.dell.com/sites/doccontent/shared-content/data-sheets/en/Documents/ESG-Lab-Review-Dell-Storage-Center-7-Apr-2016.pdf>. – (Дата обращения: 24.12.2017)
- [15] Обзор и тестирование СХД IBM FlashSystem 820. [Электронный ресурс]. URL: <https://habrahabr.ru/company/safedata/blog/258701/>. – (Дата обращения: 24.12.2017)