

АНАЛИЗ ЭКСПЛУАТАЦИОННОЙ НАДЕЖНОСТИ ОБОРУДОВАНИЯ СИСТЕМ ХРАНЕНИЯ ДАННЫХ

И. Д. Гуз, В. А. Острейковский

*Сургутский государственный университет,
idguzzz@gmail.com, ova@ivt.surgu.ru*

В статье даны определения сети хранения данных, системы хранения данных, приведена типовая структурная модель сети хранения данных, описаны проблемы, связанные с развитием инфраструктуры хранилищ данных, выполнен расчет показателей надежности дисковых подсистем. Целью работы является анализ надежности элементов дисковых подсистем на основе эксплуатационных данных. Основными свойствами, которыми должна обладать дисковая подсистема корпоративного уровня, являются надежность и отказоустойчивость. Доступность данных достигается за счет полного и частичного резервирования активных компонентов, а также продвинутыми функциями сохранения целостности данных. В качестве показателей надежности были выбраны интенсивность отказов и средняя наработка до отказа. После обработки входных данных были рассчитаны вероятности безотказной работы, вероятности отказов, интенсивности отказов, частоты отказов, а также построены графики. Были получены расчетные значения интенсивности отказов и средней наработки до отказа для элементов систем хранения данных.

Ключевые слова: сеть хранения данных, система хранения данных, дисковый модуль, надежность, отказоустойчивость, интенсивности отказов, средняя наработка до отказа, вероятности безотказной работы.

OPERATIONAL RELIABILITY ANALYSIS OF HARDWARE OF DATA STORAGE SYSTEMS

I. D. Guz, V. A. Ostreikovskiy

*Surgut State University,
idguzzz@gmail.com, ova@ivt.surgu.ru*

The article provides definitions of a storage area network, data storage systems, a typical structural model of a storage area network, describes the problems associated with the development of storage infrastructure. An overview of major storage manufacturers and reliability indicators calculations of storage equipment are given. The study aims to analyze the reliability of the elements of disk subsystems based on operational data. The main features that corporate-level data storage systems should have are reliability and fault tolerance. Data accessibility is achieved through the full and partial backup of active storage components, as well as through advanced data integrity functions. As for indicators of reliability, we chose the failure rate and the mean time to failure. After processing the input data, the probabilities of failure-free operation, the probability of failures, the failure rate, and the failure frequency are calculated, and the graphs are plotted. The calculated values of the failure rate and the mean time to failure for the data storage elements are obtained.

Keywords: storage area network, data storage system, disk module, reliability, fault tolerance, failure rate, mean time to failure, probability of failure-free operation.

Введение. Управление ресурсами хранения данных стало одной из самых важных стратегических проблем, стоящих перед сотрудниками отделов информационных технологий. Вследствие развития Интернета и коренных изменений в процессах бизнеса информация

накапливается с невиданной скоростью. Согласно данным компании IKS-Consulting, сегодня только на серверах открытых систем хранится не менее 200 петабайт информации, и этот объем удваивается каждые полтора года. Многие компании включились в своеобразное соревнование по преобразованию внутренних систем ведения бизнеса, чтобы использовать Интернет для его развития. Они перестраивают свои IT-системы для более полной поддержки приложений электронной коммерции, непрерывно работающих 24 часа в сутки, 7 дней в неделю, 365 дней в году [1].

Целью работы является анализ надежности элементов дисковых подсистем на основе эксплуатационных данных.

Задачи и функции сетевого хранения данных

Сетевое хранение данных позволяет решить многие текущие задачи в бизнесе, связанные с хранением информации, а именно:

- универсальный и совместный доступ к ресурсам;
- поддержание непрерывного роста объемов данных IT-систем;
- обеспечение непрерывной доступности при сохранении экономичности;
- обеспечение масштабируемости и высочайшей скорости работы хранилищ данных;
- упрощение управления ресурсами, связанного с их централизацией;
- повышение уровня защиты информации и отказоустойчивости.

Сеть хранения данных – это комплексное программно-аппаратное решение по организации надежного хранения информационных ресурсов и предоставления гарантированного доступа к ним. Сети хранения данных могут быть как частью, так и основой Центра Обработки Данных (далее – ЦОД).

Storage Area Network (далее – SAN) – это сеть хранения данных, которая обеспечивает подключение физических или виртуальных серверов к устройствам хранения данных (дискам, системам хранения, ленточным библиотекам и т. д.) и обмен данными между ними, чаще всего на уровне блоков.

Теоретически данному определению удовлетворяют различные технологии, включая такие стандартные сетевые протоколы, как IP/Ethernet, но на практике для использования SAN протоколы и технологии должны соответствовать ряду дополнительных требований. Стандарты, подобные IP, разрабатывались для приложений, допускающих потерю передаваемых данных, ошибки передачи и возникновение узких мест производительности. В то же время для SAN недопустимы потери или искажение данных и низкая производительность сети в течение долгого времени. Хотя потери данных в локальных и глобальных сетях допустимы для приложений и операционных систем, те же самые приложения разрабатывались исходя из допущения, что их система хранения данных работает быстро и надежно.

Таким образом, термину SAN можно дать следующее, более точное, определение: SAN – это высокопроизводительная и надежная сеть, предназначенная, прежде всего, для обеспечения связи и обмена данными на уровне блоков между физическими или виртуальными серверами и любыми устройствами хранения данных [2].

Важно осознавать, что SAN – это не конкретный протокол, а подход к организации хранения и обработки данных. Вопрос выбора тех или иных технологий, реализующих этот подход, остается за архитектором решения. Кроме того, необходимо четко осознавать момент, когда организация хранения данных является целесообразной или необходимой для организации. Возможен вариант подключения дискового массива к серверу напрямую, если дальнейшее развитие IT-инфраструктуры не планируется.

Типовая структурная модель сети хранения данных изображена на рис.1.

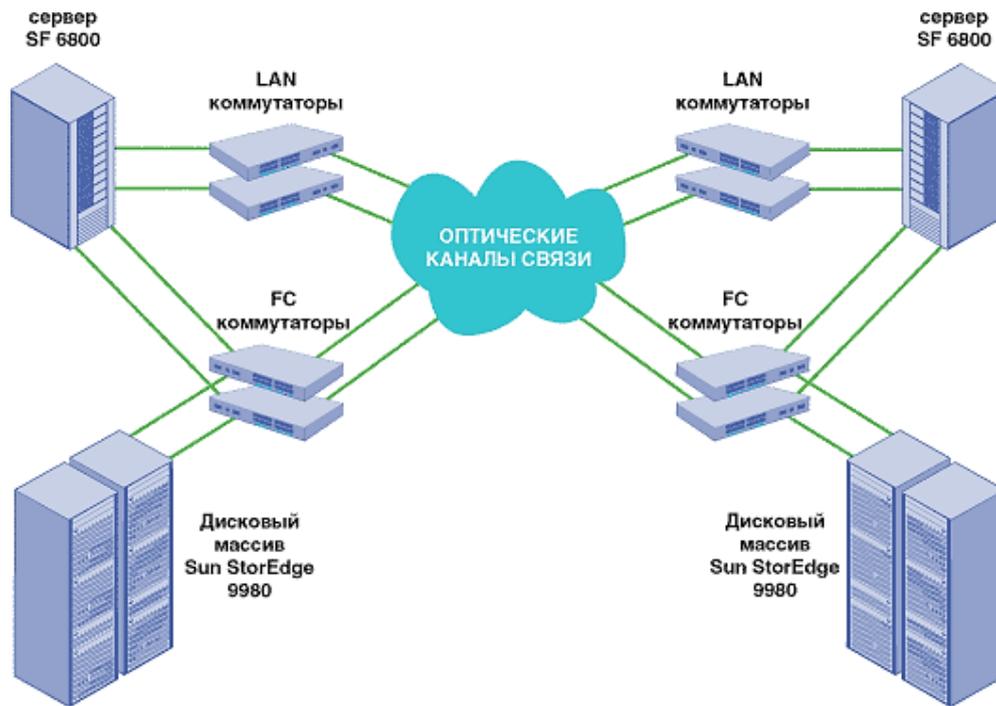


Рис. 1. Структурная модель сети хранения данных

Сеть хранения данных состоит из следующих функциональных модулей:

- серверного комплекса, включающего различные группы кластеров. Кластер рассматривается как совокупность однотипных серверов, объединенных коммутационной системой;
- системы хранения данных (далее – СХД), предназначенной для организации надежного хранения информационных ресурсов и предоставления к ним доступа со стороны серверов;
- сетевой инфраструктуры, включающей оборудование маршрутизации и коммутации, оборудование организации оптических каналов связи, оборудование подключения пользователей, средства обеспечения информационной безопасности.

Назначение систем хранения данных

Система хранения данных (или дисковая подсистема) – это совокупность специализированного оборудования и программного обеспечения, которая предназначена для хранения и передачи больших массивов информации. СХД позволяет организовать хранение информации на дисковых площадках с оптимальным распределением ресурсов [3].

На практике к СХД подключается не один сервер, а многие десятки и сотни. Это диктует ряд ключевых требований к системам такого рода:

1. Надежность и отказоустойчивость. В СХД предусмотрено полное или частичное резервирование всех компонент – блоков питания, путей доступа, процессорных модулей, дисков, кэша и т. д. Обязательно наличие системы мониторинга и оповещения о возможных и существующих проблемах.

2. Доступность данных. Обеспечивается продвинутыми функциями сохранения целостности данных (использование технологии RAID, создание полных и мгновенных копий данных внутри дисковой стойки, реплицирование данных на удаленную СХД и т. д.) и возможностью добавления (обновления) аппаратуры и программного обеспечения в горячем режиме без остановки комплекса.

3. Средства управления и контроля. Управление СХД осуществляется через web-интерфейс или командную строку, есть функции мониторинга и несколько вариантов оповещения администратора о неполадках. Доступны аппаратные технологии диагностики производительности.

4. Производительность. Определяется числом и типом накопителей, объемом кэш-памяти, вычислительной мощностью процессорной подсистемы, числом и типом внутренних и внешних интерфейсов, а также возможностями гибкой настройки и конфигурирования.

5. Масштабируемость. В СХД обычно присутствует возможность наращивания числа жестких дисков, объема кэш-памяти, аппаратной модернизации и расширения функционала с помощью специального программного обеспечения (далее – ПО). Все перечисленные операции производят без значительного переконфигурирования и потерь функциональности, что позволяет экономить и гибко подходить к проектированию ИТ-инфраструктуры [4].

Проблемы инфраструктуры систем хранения данных

Главной проблемой для большинства крупных корпоративных заказчиков стала на сегодняшний день разнородная инфраструктура СХД. Организациям нередко приходится поддерживать десятки СХД различных классов и поколений от разных производителей, поскольку разные приложения предъявляют разные требования к хранению данных. Так, критически важным транзакционным системам (биллинговым, процессинговым, ERP и т. п.) требуются высокая надежность и производительность, свойственные СХД верхнего ценового сегмента. Для аналитических систем нужны высокая производительность и низкая стоимость в расчете на единицу хранения, поэтому для них резервируются СХД с твердотельными дисками (SSD), а, например, для работы с файлами нужны функциональность и низкая стоимость, поэтому здесь применяются традиционные дисковые массивы. В разнородной инфраструктуре уровень утилизации СХД оказывается низким, общая стоимость владения (TCO) – непомерно высокой, управляемость – слабой, к тому же сложность такой инфраструктуры хранения, как правило, велика.

Еще одна серьезная проблема – модернизация СХД. Зачастую СХД, приобретенная три-четыре года назад, уже не справляется с растущими объемами данных и требованиями к скорости доступа к ним, поэтому приобретается новая система, на которую переносятся данные с прежней. По сути, заказчики повторно платят за объемы хранения, требуемые для размещения данных и, кроме того, несут расходы на установку новой СХД и перенос данных на нее. При этом прежние СХД, как правило, еще не настолько устаревшие, чтобы отказываться от них полностью, поэтому заказчики пытаются приспособить их под другие задачи [5].

Статистические данные сбоев и отказов СХД

В процессе анализа надежности элементов дисковых подсистем было исследовано следующее оборудование СХД: 2 комплекта HP XP24000, 2 комплекта EMC VNX5700, 1 комплект EMC CX4-480, 2 комплекта HP EVA P6500, 1 комплект СХД HP EVA P6350, 11 комплектов СХД IBM Storwize V7000. Перечисленные дисковые подсистемы были разработаны крупнейшими производителями СХД, использовались для выполнения различных задач (виртуальная инфраструктура предприятия, ERP, документооборот), различаются как по уровню производительности, так и встроенным функциям отказоустойчивости. В табл. 1 представлена информация по исследуемым дисковым массивам [6].

Дисковые массивы IBM Storwize V7000, идентичные по комплектации, использовались под однотипные задачи, вследствие чего считаем, что 11 комплектов IBM Storwize V7000 представляют один дисковый массив.

По данным аналитического агентства IKS-Consulting [1] была собрана статистика отказов элементов дисковых подсистем компании «РТК-ЦОД», расположенной в г. Екатеринбурге, за 2014–2018 гг. Также были рассчитаны значения вероятности безотказной работы (далее – $P(t)$), вероятности отказа (далее – $Q(t)$), частоты отказов (далее – $f(t)$), интенсивности отказов (далее – $\lambda(t)$) элементов дисковых подсистем по следующим формулам:

$$P(t) = \frac{N_0 - n(t)}{N_0}, \quad (1)$$

где N_0 – исходное число работоспособных объектов;
 $n(t)$ – число отказавших объектов за время t .

$$Q(t) = 1 - P(t) \quad (2)$$

$$f(t) = \frac{n(t)}{N_0 \cdot \Delta(t)} \quad (3)$$

где $\Delta(t)$ – интервал времени.

$$\lambda(t) = \frac{n(t)}{(N_0 - n(t)) \cdot \Delta t} \quad (4)$$

$$T_{\text{cp}} = \frac{1}{\lambda}, \quad (5)$$

где T_{cp} – средняя наработка до отказа.

Данные статистики отказов элементов СХД представлены в табл. 2–6.

По полученным результатам были построены графики зависимости $P(t)$, $\lambda(t)$ дисковых модулей за 2014–2018 годы для исследуемых дисковых массивов (рис. 2–3) [7]. Отказы контроллеров и блоков питания не учитывались при построении графиков потому, что все исследуемые дисковые массивы имеют резервные компоненты, вследствие чего ни один отказ за 2014–2018 годы не привел к потере информации или останову дискового массива.

У дисковых групп, в отличие от контроллеров и блоков питания, есть ограниченный запас дисков, которые могут выйти из строя без угрозы потери информации. Кроме того, дисковые модули могут выходить из строя одновременно, целыми стеками, что, вероятнее всего, может привести к потере данных бизнес-критичных систем [8].

Таблица 1

Характеристики исследуемых СХД

Модель дискового массива	Класс оборудования	Объем памяти, ТБ	Кол-во дисков	Модель диска	Наработка до отказа дисковых модулей (данные производителя), час	Год ввода в эксплуатацию
HP XP24000	высокий	56,8	312	HP DKS2F-K300FC	1 600 000	2009
EMC VNX5700	высокий	181	284	Seagate ST3600057SS	1 600 000	2013
EMC CX4-480	средний	90	284	Seagate ST345085	1 600 000	2010
HP EVA P6500	средний	52	96	HP EG0600FBLSH	1 600 000	2014
HP EVA P6350	средний	17,45	68	HP EF0600FARNA	1 600 000	2012
IBM STORWIZE V7000	средний	10	55	IBM 85Y5862	1 600 000	2011

Таблица 2

Статистика отказов элементов дисковых подсистем за 2014 год

Модель дискового массива	Наименование единицы	Кол-во отказов дисков	Кол-во отказов контроллеров	Кол-во отказов блоков питания	Кол-во дисков в массиве	$P(t)$	$Q(t)$	$f(t), 10^{-6} \text{ час}^{-1}$	$\lambda(t), 10^{-6} \text{ час}^{-1}$
HP XP24000	XP24000-1	1	0	0	312	0,997	0,003	0,366	0,367
HP XP24000	XP24000-2	5	0	0	312	0,984	0,016	1,829	1,859
EMC VNX5700	VNX5700-1	2	0	0	284	0,993	0,007	0,804	0,810

Окончание табл. 2

Модель дискового массива	Наименование единицы	Кол-во отказов дисков	Кол-во отказов контроллеров	Кол-во отказов блоков питания	Кол-во дисков в массиве	$P(t)$	$Q(t)$	$f(t), 10^{-6}$ час ⁻¹	$\lambda(t), 10^{-6}$ час ⁻¹
EMC VNX5700	VNX5700-2	2	0	0	284	0,993	0,007	0,804	0,810
EMC CX4-480	CX4	16	0	0	212	0,925	0,075	8,615	9,319
HP EVA P6500	EVA-P6500-1	4	0	0	96	0,958	0,042	4,756	4,963
HP EVA P6500	EVA-P6500-2	1	0	1	96	0,990	0,010	1,189	1,202
HP EVA P6350	EVA-P6350-1	2	0	0	68	0,971	0,029	3,358	3,459
IBM STORWIZE V7000	IBM-V7000	8	3	0	605	0,987	0,013	1,509	1,530

Таблица 3

Статистика отказов элементов дисковых подсистем за 2015 год

Модель дискового массива	Наименование единицы	Кол-во отказов дисков	Кол-во отказов контроллеров	Кол-во отказов блоков питания	Кол-во дисков в массиве	$P(t)$	$Q(t)$	$f(t), 10^{-6}$ час ⁻¹	$\lambda(t), 10^{-6}$ час ⁻¹
HP XP24000	XP24000-1	5	0	0	312	0,984	0,016	1,829	1,859
HP XP24000	XP24000-2	8	1	0	312	0,974	0,026	2,927	3,004
EMC VNX5700	VNX5700-1	2	0	0	284	0,993	0,007	0,804	0,810
EMC VNX5700	VNX5700-2	2	0	0	284	0,993	0,007	0,804	0,810
EMC CX4-480	CX4	10	0	0	212	0,953	0,047	5,385	5,651
HP EVA P6500	EVA-P6500-1	7	0	0	96	0,927	0,073	8,324	8,979
HP EVA P6500	EVA-P6500-2	2	0	0	96	0,979	0,021	2,378	2,429
HP EVA P6350	EVA-P6350-1	1	0	0	68	0,985	0,015	1,679	1,704
IBM STORWIZE V7000	IBM-V7000	19	5	0	605	0,969	0,031	3,585	3,701

Таблица 4

Статистика отказов элементов дисковых подсистем за 2016 год

Модель дискового массива	Наименование единицы	Кол-во отказов дисков	Кол-во отказов контроллеров	Кол-во отказов блоков питания	Кол-во дисков в массиве	$P(t)$	$Q(t)$	$f(t), 10^{-6}$ час ⁻¹	$\lambda(t), 10^{-6}$ час ⁻¹
HP XP24000	XP24000-1	8	0	0	312	0,974	0,026	2,927	3,004
HP XP24000	XP24000-2	8	0	0	312	0,974	0,026	2,927	3,004
EMC VNX5700	VNX5700-1	13	2	0	284	0,954	0,046	5,225	5,476
EMC VNX5700	VNX5700-2	12	0	0	284	0,958	0,042	4,823	5,036
EMC CX4-480	CX4	12	0	0	212	0,943	0,057	6,462	6,849
HP EVA P6500	EVA-P6500-1	2	1	0	96	0,979	0,021	2,378	2,429
HP EVA P6500	EVA-P6500-2	5	0	0	96	0,948	0,052	5,946	6,272
HP EVA P6350	EVA-P6350-1	3	1	0	68	0,956	0,044	5,036	5,269
IBM STORWIZE V7000	IBM-V7000	18	1	5	605	0,970	0,030	3,396	3,501

Таблица 5

Статистика отказов элементов дисковых подсистем за 2017 год

Модель дискового массива	Наименование единицы	Кол-во отказов дисков	Кол-во отказов контроллеров	Кол-во отказов блоков питания	Кол-во дисков в массиве	$P(t)$	$Q(t)$	$f(t), 10^{-6}$ час ⁻¹	$\lambda(t), 10^{-6}$ час ⁻¹
HP XP24000	XP24000-1	7	0	0	312	0,978	0,022	2,561	2,620
HP XP24000	XP24000-2	8	0	0	312	0,974	0,026	2,927	3,004
EMC VNX5700	VNX5700-1	19	0	1	284	0,933	0,067	7,637	8,185
EMC VNX5700	VNX5700-2	21	0	0	284	0,926	0,074	8,441	9,115
EMC CX4-480	CX4	5	0	0	212	0,976	0,024	2,692	2,757
HP EVA P6500	EVA-P6500-1	2	0	0	96	0,979	0,021	2,378	2,429
HP EVA P6500	EVA-P6500-2	8	0	0	96	0,917	0,083	9,513	10,378
HP EVA P6350	EVA-P6350-1	2	1	0	68	0,971	0,029	3,358	3,459
IBM STORWIZE V7000	IBM-V7000	22	5	1	605	0,964	0,036	4,151	4,308

Таблица 6

Статистика отказов элементов дисковых подсистем за 2018 год

Модель дискового массива	Наименование единицы	Кол-во отказов дисков	Кол-во отказов контроллеров	Кол-во отказов блоков питания	Кол-во дисков в массиве	$P(t)$	$Q(t)$	$f(t), 10^{-6}$ час ⁻¹	$\lambda(t), 10^{-6}$ час ⁻¹
HP XP24000	XP24000-1	4	0	0	312	0,987	0,013	1,464	1,483
HP XP24000	XP24000-2	12	0	0	312	0,962	0,038	4,391	4,566
EMC VNX5700	VNX5700-1	25	1	0	284	0,912	0,088	10,049	11,019
EMC VNX5700	VNX5700-2	29	1	0	284	0,898	0,102	11,657	12,982
EMC CX4-480	CX4	9	0	1	212	0,958	0,042	4,846	5,061
HP EVA P6500	EVA-P6500-1	1	0	0	96	0,990	0,010	1,189	1,202
HP EVA P6500	EVA-P6500-2	1	0	0	96	0,990	0,010	1,189	1,202
HP EVA P6350	EVA-P6350-1	1	0	0	68	0,985	0,015	1,679	1,704
IBM STORWIZE V7000	IBM-V7000	20	3	4	605	0,967	0,033	3,774	3,903

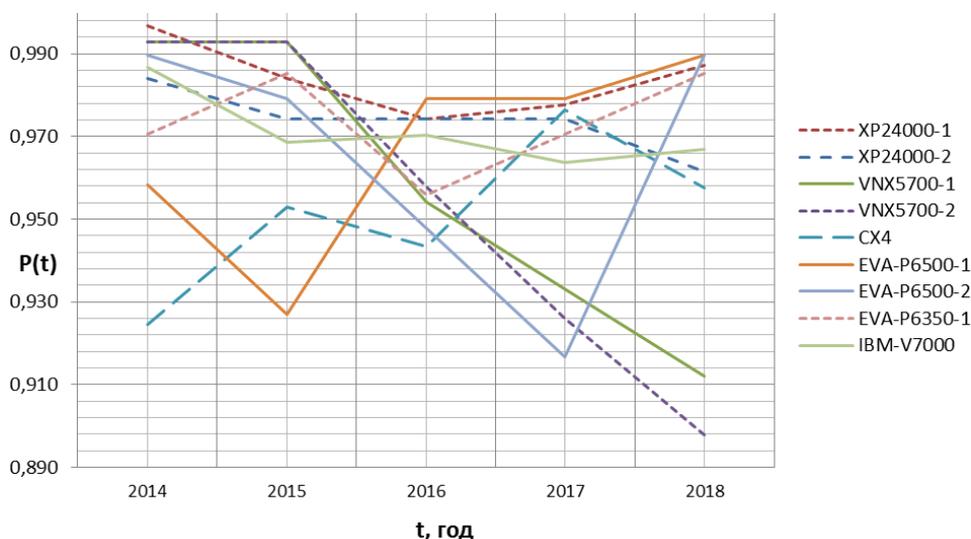


Рис. 2. Вероятность безотказной работы дисковых носителей СХД за 2014–2018 годы

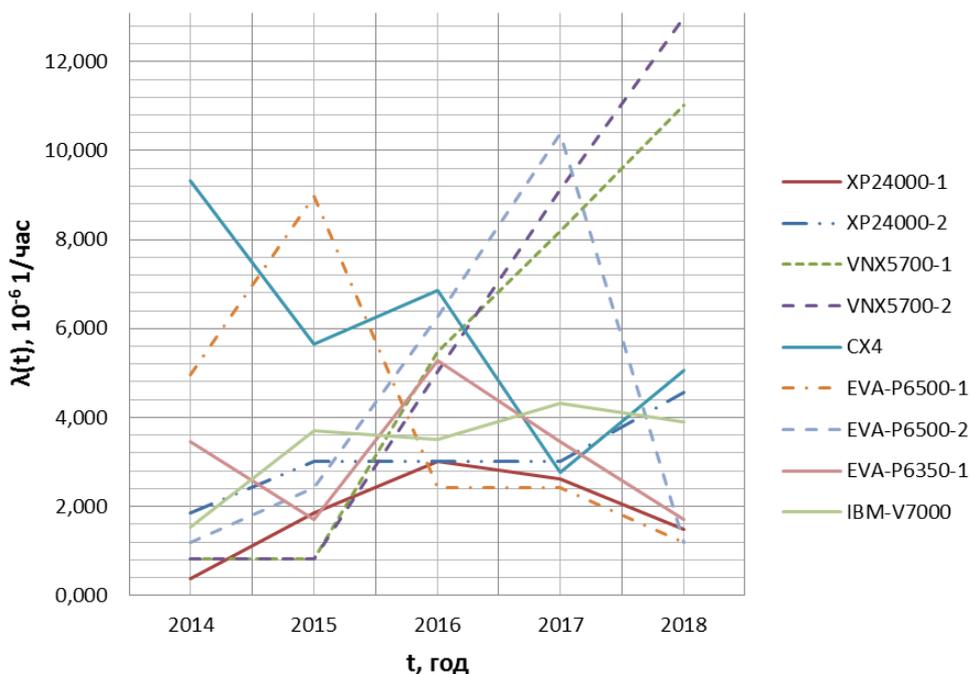


Рис. 3. Интенсивность отказов дисковых носителей СХД за 2014–2018 годы

По данным, приведенным в табл. 1–5, по формулам 4 и 5 были рассчитаны средние значения интенсивности отказов (далее – λ_{cp}) и средняя наработка до отказа (далее – T_{cp}) дисковых модулей исследуемых дисковых подсистем, также приведены значения средней наработки на отказ (далее – T_{vendor}), предоставленные компаниями-производителями дисковых носителей (табл. 7) [9].

Таблица 7

Значения показателей надежности элементов дисковых подсистем

Модель дискового массива	Модель дискового модуля	λ_{cp} (час ⁻¹)	T_{cp} (час)	T_{vendor} (час)
XP24000-1	HP DKS2F-K300FC	0,00000187	535742,45	1600000
XP24000-2	HP DKS2F-K300FC	0,00000309	323883,02	1600000
VNX5700-1	Seagate ST3600057SS	0,00000526	190122,25	1600000
VNX5700-2	Seagate ST3600057SS	0,00000575	173895,48	1600000
CX4	Seagate ST345085	0,00000593	168703,45	1600000
EVA-P6500-1	HP EG0600FBLSH	0,00000400	249986,51	1600000
EVA-P6500-2	HP EG0600FBLSH	0,00000430	232751,70	1600000
EVA-P6350-1	HP EF0600FARNA	0,00000312	320619,21	1600000
IBM-V7000	IBM 85Y5862	0,00000339	295124,74	1600000

Выводы. После обработки входных данных было сделано следующее:

1. Рассчитаны значения вероятности безотказной работы $P(t)$, вероятности отказов $Q(t)$, интенсивности отказов $\lambda(t)$, частоты отказов $f(t)$ элементов дисковых подсистем за 2014–2018 гг. [9].

2. Построены графики зависимости вероятности безотказной работы $P(t)$, интенсивности отказов $\lambda(t)$ за 2014–2018 г. для исследуемых дисковых массивов.

3. Рассчитаны значения показателей надежности элементов дисковых подсистем.

В некоторых ситуациях расчетные значения средней наработки на отказ T_{cp} и значения средней наработки на отказ T_{vendor} , предоставленные производителями СХД, отличаются в 10 раз, что говорит о высокой загруженности дисковых массивов. Решением могло бы стать проведение модернизации используемого оборудования, закупка нового более современного оборудования СХД или перераспределение нагрузки между действующими СХД [10].

Литература

1. Мирин С., Башилов Г., Патрикеев Д. Облачный провайдинг 2018–2022: экономика, стратегии, бизнес-модели: информационные и облачные технологии. Декабрь 2018 г. URL: <http://www.iksconsulting.ru/reports-91.html> (дата обращения: 03.07.2019).
2. Орлов С. А. Оптимальный ЦОД // Журнал сетевых решений Lan. 2011. С. 50–61.
3. Марк Фарли Сети хранения данных. 2-е изд. М. : Лори, 2004. 576 с.
4. Вильчинский Н. В. От хранения данных к управлению информацией. СПб. : Питер, 2016. 544 с.
5. Гоменюк А. Р., Сопенко С. И. Строим центр обработки данных // Корпоративные системы. 2017. № 5. С. 6–11.
6. Таненбаум Э., Бос Х. Современные операционные системы. 4-е изд. СПб. : Питер, 2015. 1120 с.
7. Таненбаум Э., Уэзеролл Д. Современные компьютерные сети. 5-е изд. СПб. : Питер, 2015. 960 с.
8. Джош Джад. Основы проектирования SAN. М. : Brocade Россия и СНГ, 2008. 589 с.
9. Острейковский В. А. Теория надежности. М. : Высш. шк., 2003. 463 с.
10. Гмурман В. Е. Теория вероятностей и математическая статистика. М. : Высш. образование, 2008. 479 с.