

Literatura

1. *Aleksandrov A.P., Borovikov G.A., Eryomenko A.S.* Eksploatatsiya radiotekhnicheskikh kompleksov. M.: Sovetskoe radio, 1976. 280 s.
2. *Bazilevich L.A.* Modelirovanie organizatsionnykh struktur. L.: Izdatel'stvo LGU, 1978. 160 s.
3. *Pavlov A.N., Sokolov B.V.* Formirovanie i sovershenstvovanie organizatsionnykh struktur upravleniya: uchebno-metodicheskoe posobie. SPb.: SPbGUAP, 2005. 42 s.
4. *Petukhov G.B.* Osnovy teorii effektivnosti tselenapravlennykh protsessov. Ch. 1: Metodologiya, metody, modeli. M.: MO SSSR, 1989. 654 s.
5. *Putilin V.N.* Metodicheskie aspekty opredeleniya chislennosti lichnogo sostava voinskikh formirovanij // Strategicheskaya stabil'nost'. 2000. № 1.
6. *Chernykh Zh.V.* Metodicheskij podkhod k obosnovaniyu organizatsionno-shtatnoj struktury spatel'nykh tsentrov MChS Rossii // Nauchnye i obrazovatel'nye problemy grazhdanskoj zashchity. 2016. № 3. S. 20–24.

DOI: 10.25586/RNUV9187.20.03.P.083

УДК 061.68

А.И. Захаров, Г.А. Брякалов, Е.В. Чумакова

РАЗРАБОТКА МЕТОДИКИ РАСЧЕТА И ОЦЕНКИ
ЭКСПЛУАТАЦИОННЫХ ХАРАКТЕРИСТИК СРЕДСТВ СИСТЕМ
ХРАНЕНИЯ ИНФОРМАЦИИ ЦЕНТРОВ ОБРАБОТКИ ДАННЫХ

Разрабатываются основы методики расчета эксплуатационных характеристик и синтеза состава IT-оборудования системы хранения данных. Необходимость такой системы определяется возросшей потребностью в свободном доступе к данным и в возможности управления ими. Это достигается за счет использования комплекса специализированного аппаратного оборудования и программного обеспечения, предназначенных для организации хранения и передачи информации, что подчеркивает актуальность статьи и показывает важность задачи синтеза элементной базы системы хранения данных, методики ее расчета и предварительной оценки состава.

Ключевые слова: центры обработки данных, ЦОДы, системы хранения данных, СХД, IT-оборудование ЦОДа, расчет характеристик ЦОДа.

A.I. Zakharov, G.A. Bryakalov, E.V. Chumakova

ON THE QUESTION OF INFORMATION STORAGE SYSTEMS DATA
PROCESSING CENTERS ORGANIZATION

The article is concerned with laying the groundwork for methods of synthesis IT-equipment data storage system contents. The necessity of such system is defined by higher requirements of free data access and possibility to operate it. It is achieved by using the arrangement of specific hardware and software, which is meant for organization of storage and information transfer, that underlines the applicability of the article and shows the importance of the hardware components data storage system synthesis task, its calculation methodology and preliminary estimate of contents.

Keywords: data processing center, DPC, data storage system, DSS, IT-equipment of DPC, performance calculation of DPC.

Введение

В настоящее время в литературе, посвященной информационным технологиям, уделяется много внимания обработке и хранению информации в центрах обработки данных (ЦОДах). Это связано с большими объемами и высокой скоростью обработки информации, а также с интенсивным обменом ею и необходимостью надежного хранения. Повышенные требования к надежности хранения и времени доступа к данным предполагают создание специальной системы средств их хранения. Роль и важность такой системы определяются возросшей потребностью в свободном доступе к данным и возможностью управления ими.

Выделение средств хранения данных в отдельную систему позволяет решать проблему обеспечения упомянутого надежного хранения и доступа к данным. Кроме того, это создает условия для оформления системы хранения данных (СХД) в виде аппаратно-программной структуры, что предоставляет возможность использования средств хранения данных для решения поставленных задач. Иными словами, необходимость в СХД возникла, когда массивы хранимой и передаваемой информации достигли значительных размеров, что, в свою очередь, предопределило время и надежность доступа к ним.

Итак, системы хранения данных предназначены для организации надежного хранения данных, а также отказоустойчивого, высокопроизводительного доступа к устройствам хранения и представляют собой комплекс специализированного аппаратного оборудования и программного обеспечения, предназначенных для хранения и передачи информации [2; 9].

Все современные ЦОДы сталкиваются с проблемой постоянного увеличения объемов данных, которые обязательно нужно надежно хранить. Это могут быть данные оперативных систем корпоративных заказчиков ЦОДов, к которым предъявляются повышенные требования по доступности, или же архивные данные, требования доступности к которым ниже, зато их следует долго хранить (например, архивные или резервные копии).

У больших СХД для достижения высокого уровня сервиса становится актуальным повышение быстродействия системы. При этом большой СХД считается система, в которой обрабатываются значительные объемы данных (до петабайтов) или к ней подключены десятки серверов. Для небольших систем быстродействие можно повысить за счет более производительной аппаратной части системы.

Таким образом, на повестку дня остро встает вопрос правильной организации хранения данных с учетом требования по оперативности доступа к ним, а также оптимизации временных затрат при растущих объемах данных.

Анализ отечественных и зарубежных СХД [2; 3; 7] выявил ряд их положительных качеств, к которым можно отнести дешевизну и доступность ресурсов хранилища для любых компьютеров, простоту их коллективного использования, возможность применения блочных методов доступа и др.

Вместе с тем СХД свойственны и недостатки. Это, например, невозможность коллективного доступа к СХД с разных компьютеров, довольно высокая стоимость подключения к системам хранения и др.

Стремление снизить затраты на инфраструктуру СХД, более гибко и оперативно реагировать на новые требования к системам хранения данных, получить комплексную оценку их возможностей обуславливает актуальность исследований в этом направлении.

Структура СХД

Анализируя известные из литературы структуры СХД, можно выделить типовые фрагменты устройств, а также математическое обеспечение, характерные или обязательные для состава любой структуры системы хранения данных [3; 8]. К таковым можно отнести:

1. Устройства хранения (массивы дисков, ленточные библиотеки, оптические накопители).
2. Инфраструктуру доступа к устройствам хранения.
3. Подсистему резервного копирования и архивирования данных.
4. Программное обеспечение управления хранением.
5. Систему управления и мониторинга всей системы в целом.

Если коротко остановиться на физической основе структуры СХД, то следует отметить, что, с одной стороны, это специальная система для работы с носителями информации, с другой – система предоставления данных серверам для ее целевой обработки.

Таким образом, серверы предназначены для вычислений и обработки информации, а СХД принимает, хранит, отправляет данные и исключает возможность их утраты. Поскольку задачи у них разные, то раньше серверы размещались только в ЦОДах, а СХД – в хранилище данных.

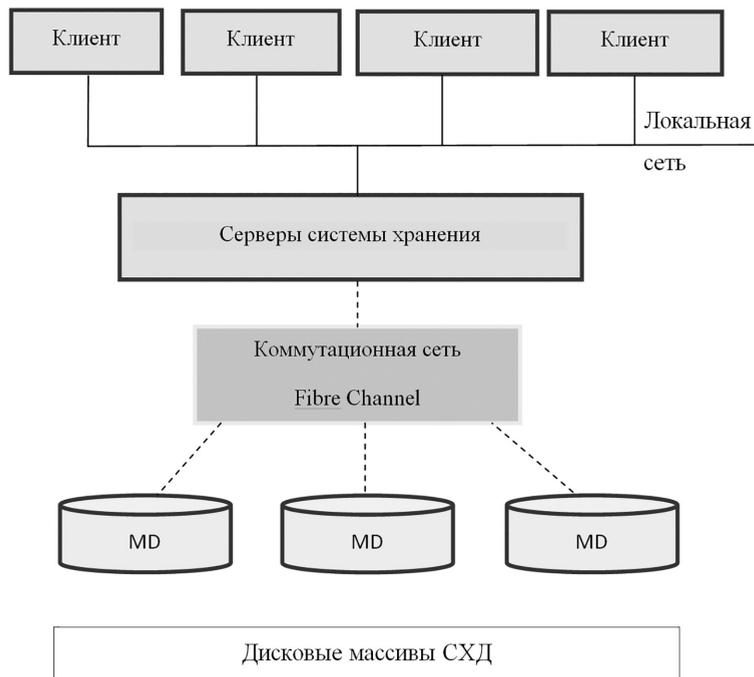
В настоящее время в СХД ставят процессоры с высокой производительностью, и СХД становится самостоятельным сервером, способным обрабатывать и хранить файлы, поэтому элементы структуры современных СХД (серверная часть и непосредственно хранилище данных) могут находиться как на разных площадках размещения оборудования, так и на одной.

По типу подключения СХД делятся на [2; 7; 9]:

- DAS (*Direct Attached Storage*) – устройство для хранения данных, подключенное непосредственно к серверу или к рабочей станции через интерфейс по протоколу SAS;
- NAS (*Network Attached Storage*) – отдельная интегрированная дисковая система со своей специализированной организационной структурой (ОС) и набором функций быстрого запуска системы и обеспечения доступа к файлам. Система подключается к обычной компьютерной сети (ЛВС);
- SAN (*Storage Area Network*) – это специальная выделенная сеть, объединяющая устройства хранения данных с серверами приложений. Она обычно строится на основе протокола *Fibre Channel* или протокола iSCSI [7].

При дальнейших рассуждениях остановимся именно на системе со специальной выделенной сетью SAN. Упрощенная структура такой СХД представлена на рисунке.

Если исходить из анализа недостатков СХД, отмеченных выше, то на содержательном уровне можно сформулировать задачу исследования системы хранения данных ЦОДа с учетом требований по их доступности и оптимизации временных затрат на доступ к информации как задачу синтеза СХД.



Вариант структуры СХД типа SAN

Методика расчета характеристик системы хранения информации центра обработки данных

Поставленную задачу на содержательном уровне можно сформулировать следующим образом.

На момент расчета хранилища данных центра обработки данных известно количество информации, которое необходимо обработать и сохранить в СХД. Например, оно составляет 5 Петафлоп (Пфлоп).

Требуется определить:

1. Площадь помещения и количество стоек для IT-оборудования машинного зала.
2. Количество серверов, необходимый дисковый массив, а также потребляемую ими электроэнергию.

Расчет суммарной производительности вычислительных средств СХД

Переведем исходное количество информации Q (5 Пфлоп), которое следует отправить в хранилище, из внесистемной в более удобную размерность – байты.

1 флоп = 8 байтов (при обработке данных в размере «двойное слово»).

1 Пфлоп = 10^{15} флоп.

5 Пфлоп = 5×10^{15} флоп = $8 \times 5 \times 10^{15}$ байт = 40 000 Терабайт (Тбайт).

Иными словами, в хранилище необходимо поместить 40 000 Тбайт информации.

С учетом того, что в сутках $T = 86\,400$ с, а системные затраты времени составляют примерно одну треть от T [4], т.е. $T_{\text{затр}} = 28\,800$ с, то при объеме пересылаемой информации

Захаров А.И. и др. Разработка методики расчета и оценки эксплуатационных...

$Q = 40\,000$ Тбайт получим следующую суммарную производительность Π вычислительных средств [6; 10]:

$$\Pi = Q / (T - T_{\text{затр}}) = 40\,000 / (86\,400 - 28\,800) \approx 0,7 \text{ Тфлопс,}$$

где Π – суммарная производительность вычислительных средств; Q – объем пересылаемой информации; T – время работы вычислительных средств; $T_{\text{затр}}$ – системные затраты времени.

Выбор вычислительных средств для серверной части ИТ-оборудования и расчет площади серверной части машинного зала

Учитывая разнообразие типов ЭВМ и удобство компоновки их в машинном зале, оставим свой выбор на отечественном вычислительном комплексе ВК «Сивуч» [4; 5], характеристики которого приведены в таблице.

Технические характеристики ВК «Сивуч»

Наименование параметра	«С-1»-1	«С-4»-4	«С-4»-8	«С-4»-16
Количество ВК «Сивуч-1», шт.	1	4	8	16
Количество ИМС 1891ВМ7АЯ, шт.	36	144	288	576
Производительность, Тфлопс, не менее	0,4	1,6	3,2	6,4
Потребляемая мощность, кВт, не более	2,8	11,2	22,4	44,8

Исходя из общего необходимого объема хранилища $Q = 40\,000$ Тбайт и учитывая суммарную производительность вычислительных средств при суточной обработке информации $\Pi = 0,7$ Тфлопс, можно определить вариант состава ЭВМ, обладающих достаточной или более высокой производительностью. Например, для варианта использования ВК «Сивуч-1», производительность которого $\Pi_{\text{«С-1»}} = 0,4$ Тфлопс (см. табл.), она окажется ниже требуемых $0,7$ Тфлопс.

Следовательно, вариант с этим вычислительным комплексом не способен обработать поступивший объем информации Q за время $T_{\text{раб}} = T - T_{\text{затр}} = 57\,600$ с. С целью выполнения поставленной задачи по обработке поступившей информации следует для машинного зала выбрать два комплекса ВК «Сивуч-1» с суммарной производительностью $0,8$ Тфлопс.

Тогда появится возможность обработать и сохранить поступившую в СХД информацию.

$$Q = 2 \times \Pi_{\text{«С-1»}} \times T_{\text{раб}} = 2 \times 0,4 \times 57\,600 = 46\,080 \text{ Тфлоп.}$$

Вся площадь здания современного ЦОДа делится на серверную, хранилище данных и вспомогательную площадь.

Так как по статистическим данным на одну стойку ВК «Сивуч-1» отводится $2,5 \text{ м}^2$, то два ВК «Сивуч-1» займут серверную часть $S_{\text{с.ч}}$ площади в размере

$$S_{\text{с.ч}} = 2 \times 2,5 = 5 \text{ м}^2.$$

К вспомогательной площади ЦОДа отнесем комнату дежурной смены ($\approx 15 \text{ м}^2$), помещение для ЗИПа ($\approx 5 \text{ м}^2$), а также подходы и проходы к ВК «Сивуч-1» ($\approx 5 \text{ м}^2$).

Тогда вспомогательная площадь совместно составит $\approx 25 \text{ м}^2$.

Определение необходимого дискового массива

Согласно признаку классификации «тип накопителя», остановим выбор последнего на жестком диске как наиболее экономичном и управляемом накопителе.

Жесткие диски можно выбирать по следующим параметрам [1; 4]:

- конструктивное исполнение: HDD, SSD;
- по диаметру HDD в дюймах: 3,5; 2,5; 1,8 дюйма;
- по интерфейсу: ATA/IDE, SATA/NL, SAS, SCSI, SAS, FC;
- по классу использования: индивидуальные, корпоративные.

Требуемое число дисков для удобства эксплуатации объединим в массивы дисков по схеме RAID 50. Использование массива жестких дисков RAID 50 предполагает, что выполнение операций RAID 5, а затем RAID 0 увеличит производительность данного массива.

RAID 5 обеспечивает разделение данных с проверкой на паритет с возможностью хранения этих данных в разных хранилищах [3].

Для конструктивного размещения массива дисков RAID 50 на стойках СХД, установленных в машинном зале, выберем один из вариантов компоновки дисков на полках расширения фирмы IBM EXP3000, которая сочетает передовые технологии и высокую надежность. Полка расширения EXP3000 может содержать до 12 МД SATA и монтируется на стандартную стойку высотой 42U, которая вмещает до 20 таких полок.

С учетом возможности масштабирования дисков путем подключения дополнительных полок расширения фирмы IBM EXP3000 общее количество дисков на одной стойке может быть доведено до 240. Тогда при использовании дисков SATA емкостью 750 Гбайт с подключением к системе SAN их максимальная емкость на стойке Q_c составит 180 Тбайт.

Отсюда можно определить общее количество стоек K_c с набором магнитных дисков

$$K_c = Q / Q_c = 40\,000 / 180 \approx 222 \text{ стойки.}$$

Определение общей площади и необходимого количества электроэнергии для ИТ-оборудования СХД

Часть общей площади, отводимой под хранилище данных, включающей 222 стойки по $0,5 \text{ м}^2$ на стойку (т.е. 111 м^2), проходы и коридор между стойками дисков, занимающие порядка 14 кв.м , в сумме составит 125 м^2 .

Общую площадь СХД ($S_{\text{общ}}$) можно определить как сумму площадей серверной ($S_{\text{с.ч}}$), хранилища данных (S_A) и вспомогательной ($S_{\text{всп}}$).

$$S_{\text{общ}} = S_{\text{с.ч}} + S_A + S_{\text{всп}} = 5 + 125 + 25 = 155 \text{ м}^2.$$

Необходимое количество электроэнергии $P_{\text{общ}}$ для ИТ-оборудования СХД можно определить как сумму потребляемой серверами электроэнергии P_c (см. табл.) и массивом дисков хранилища данных $P_{\text{х.д}}$.

Захаров А.И. и др. Разработка методики расчета и оценки эксплуатационных...

Принимая во внимание, что одна стойка вмещает 240 дисков, а каждый диск потребляет ≈ 10 Вт электроэнергии, то каждая стойка потребляет 2400 Вт или 2,4 кВт. Отсюда следует, что все 222 стойки СХД могут потреблять 532,8 кВт электроэнергии.

Тогда необходимое количество электроэнергии $P_{\text{общ}}$ для IT-оборудования СХД с учетом небольшого запаса составит

$$P_{\text{общ}} = P_c + P_{\text{х.д}} = 5,6 + 532,8 = 538,4 \text{ кВт} \approx 540 \text{ кВт.}$$

Итак, по ходу изложения статьи были затронуты вопросы, связанные с проектированием и разработкой систем хранения данных, размещаемых в рамках единого здания ЦОДа, предпринята попытка комплексной оценки структуры и состава средств IT-оборудования как центра обработки данных, так и системы их хранения.

Заключение

В статье представлены формулировка и решение задачи синтеза состава IT-оборудования, необходимого для нормального функционирования современного хранилища данных ЦОДа. В ходе исследования выполнены ориентировочные расчеты размещения средств обработки данных и их хранения на единой совмещенной территории машинного зала, что по своей сути подтверждает решение задачи синтеза для проектируемой СХД.

Статья носит прикладной характер и предназначена для разработчиков СХД и инженерного состава, связанного с их эксплуатацией.

Литература

1. Аладьев В.З. Компьютерная хрестоматия. Справочное руководство. Работа жесткого диска. М.: Российская энциклопедия, 2012. 337с.
2. Архипенков С.Я., Голубев Д.В., Максименко О.Б. Хранилища данных. М.: Диалог-МИФИ, 2002. 528 с.
3. Барсегян А.А. и др. Методы и модели анализа данных OLAP и Data Mining. Гл. 4–5, 7. СПб.: БХВ-Петербург, 2004.
4. Вычислительный комплекс «Сивуч-1». Руководство по эксплуатации. ТВГИ. 466535. 130-01 РЭ. М.: МЦСТ, 2014. 124 с.
5. Голубев Д., Лобанов А. Сети хранения данных (SAN) // Jet Info. 2002. № 9. С. 2–16.
6. Захаров А.И., Брякалов Г.А., Михайлова П.И., Чумакова Е.В. Методика расчета и оценки состава IT-оборудования центра обработки данных // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2019. Вып. 2. С. 31–40.
7. Захаров А.И., Брякалов Г.А., Чмыхова Я.В. Методика оценки возможностей вычислительных средств центра обработки данных // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2019. Вып. 1. С. 124–129.
8. Ким А.К. и др. Микропроцессоры и вычислительные комплексы семейства «Эльбрус»: учебное пособие. СПб.: Питер, 2013. 272 с.
9. Ушаков Н.Н. Технология элементов вычислительных машин. Структуры хранения информации. М.: Высшая школа, 2001. 413 с.
10. SATA и другие интерфейсы жестких дисков // ИТС.ua. URL: http://its.ua/articles/sata_i_dругие_interfejsy_zhestkih_diskov_22528 (дата обращения: 08.07.2020).

Literatura

1. *Alad'ev V.Z.* Komp'yuternaya khrestomatiya. Spravochnoe rukovodstvo. Rabota zhestkogo diska. M.: Rossijskaya entsiklopediya, 2012. 337s.
2. *Arkhipenkov S.Ya., Golubev D.V., Maksimenko O.B.* Khranilishcha dannykh. M.: Dialog-MIFI, 2002. 528 с.
3. *Barsegyan A.A. i dr.* Metody i modeli analiza dannykh OLAP i Data Mining. Gl. 4–5, 7. SPb.: BKhV-Peterburg, 2004.
4. Vychislitel'nyj kompleks "Sivuch-1". Rukovodstvo po ekspluatatsii. TVGI. 466535. 130-01 RE. M.: MTsST, 2014. 124 s.
5. *Golubev D., Lobanov A.* Seti khraneniya dannykh (SAN) // Jet Info. 2002. № 9. С. 2–16.
6. *Zakharov A.I., Bryakalov G.A., Mikhajlova P.I., Chumakova E.V.* Metodika rascheta i otsenki sostava IT-oborudovaniya tsentra obrabotki dannykh // Vestnik Rossijskogo novogo universiteta. Seriya "Slozhnye sistemy: modeli, analiz i upravlenie". 2019. Vyp. 2. S. 31–40.
7. *Zakharov A.I., Bryakalov G.A., Chmykhova Ya.V.* Metodika otsenki vozmozhnostej vychislitel'nykh sredstv tsentra obrabotki dannykh // Vestnik Rossijskogo novogo universiteta. Seriya "Slozhnye sistemy: modeli, analiz i upravlenie". 2019. Vyp. 1. S. 124–129.
8. *Kim A.K. i dr.* Mikroprotsessory i vychislitel'nye komplekсы semejstva "El'brus": uchebnoe posobie. SPb.: Piter, 2013. 272 s.
9. *Ushakov N.N.* Tekhnologiya elementov vychislitel'nykh mashin. Struktury khraneniya informatsii. M.: Vysshaya shkola, 2001. 413 с.
10. SATA i drugie interfejsy zhestkikh diskov // ITC.ua. URL: http://itc.ua/articles/sata_i_drugie_interfejsy_zhestkih_diskov_22528 (data obrashcheniya: 08.07.2020).

DOI: 10.25586/RNU.V9187.20.03.P.090

УДК 123.2

В.А. Максимов, А.В. Калюжный, В.Е. Салимоненко

АЛГОРИТМ СИММЕТРИЧНОГО ШИФРОВАНИЯ
ДЛЯ ПРОГРАММНОГО КОМПЛЕКСА ЗАЩИЩЕННОГО ОБМЕНА
ИНФОРМАЦИИ В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ

Представлен оригинальный алгоритм симметричного шифрования, а также описан программный комплекс информационного обмена, использующий данный алгоритм. Отличительными особенностями алгоритма являются невысокая вычислительная сложность и малая информационная избыточность, вносимая в исходные открытые данные. Программный комплекс может быть использован в автоматизированных системах специального назначения либо в других системах, где требуется защищенный обмен информацией.

Ключевые слова: шифрование, алгоритм, защита информации.