

тодологии экстремального программирования, метода декомпозиции. Он позволяет существенно снизить затраты на обеспечение качества продуктов. Данный метод оправдал себя в долгосрочных проектах (длительностью более 1 года), где более заметна зависимость применяемой методологии от затрат. Метод помог снизить риск компании связать себя обязательствами по отношению к убыточным проектам на этапе принятия решения о вступлении в проект. Контроль качества на каждом этапе реализации проекта снижает себестоимость проекта, т. к. позволяет отслеживать ошибки на ранних этапах, когда их стоимость значительно ниже по сравнению с ошибками, выявленными на этапе тестирования и последующих этапах.

Суть предложенного метода можно кратко изложить как объединение работы аналитика и инженера QA (возможна даже ситуация для неболь-

ших компаний, чтобы это был один человек) на каждом этапе жизненного цикла разработки программного обеспечения. Аналитик и инженер QA обладают большим запасом знаний и опытом, позволяющим осуществлять контроль за качеством системы на всех уровнях: от логики системы до опечаток программистов. Поэтому, начиная с момента принятия решения о начале реализации проекта до сдачи системы в эксплуатацию, инженер QA и аналитик должны работать в тесном сотрудничестве с обменом ролями и дополнением обязанностей друг друга. Внедрение метода предполагает определенное время на перестроение работы аналитического отдела и отдела контроля качества, а также на адаптацию многих аналитических методов к процессам контроля качества, но результаты, выраженные в финансовой и временной экономии, оправдывают такие затраты.

СПИСОК ЛИТЕРАТУРЫ

1. Schulmeyer G.G. Handbook of Software Quality Assurance. – Norwood: Artech House, 2008. – 485 p.
2. Поллис Г., Огастин Л. Разработка программных проектов на основе Rational Unified Process (RUP). – М.: Бинوم, 2009. – 256 с.
3. Брагин Ю.В. Путь QFD: проектирование и производство продукции исходя из ожиданий потребителей. – М.: Центр качества, 2003. – 240 с.
4. Бек К. Экстремальное программирование: разработка через тестирование. – СПб.: Питер, 2003. – 224 с.

Поступила 13.04.2009 г.

УДК 004.623

ТЕХНОЛОГИЧЕСКИЕ АСПЕКТЫ ОЦЕНКИ ПРОИЗВОДИТЕЛЬНОСТИ СИСТЕМ ХРАНЕНИЯ И ОБРАБОТКИ БОЛЬШИХ ОБЪЕМОВ ДАННЫХ

А.В. Алубин, В.В. Грачев, С.А. Матвеев, О.Ф. Юдин, М.А. Сонькин*

ФГУП «НИИ «КВАНТ», г. Москва

*Томский политехнический университет

E-mail: sonkin@tpu.ru

Определены основные подходы к оценке производительности, проведен сравнительный анализ и протестированы системы хранения данных ведущих производителей.

Ключевые слова:

Системы хранения данных, системы обработки информации, вычислительные модули, локальная вычислительная сеть, производительность.

В настоящее время отсутствует общепринятая методика оценки производительности систем обработки информации (СОИ) специального назначения, что связано в первую очередь с отсутствием единиц для измерения количества вычислительной работы. Поэтому для оценки производительности используется широкая номенклатура величин — показателей производительности, которые и в отдельности, и в совокупности не удовлетворяют в полной мере потребностям теории и практики проектирования и эксплуатации СОИ.

Рассмотрим СОИ, состоящую из совокупности вычислительных модулей (ВМ) объединенных в локальную вычислительную сеть (ЛВС), причем каждый из ВМ занимается решением только одной выделенной задачи специального назначения. Структурная схема такой СОИ представлена на рис. 1.

Основными функциями СОИ могут быть:

- автоматическая обработка принятой информации;
- ведение архивов принятой и обработанной информации;

- работа с базами данных (БД) специального назначения.

В данной статье предлагается новый подход и обобщенный алгоритм оценки потенциальной и реализованной производительности СОИ специального назначения. Применение этого алгоритма требует разработки и использования специального программного обеспечения, предназначенного для получения част-

ных значений характеристик производительности отдельных ВМ, входящих в состав СОИ, и СОИ в целом.

Для проведения сравнительного анализа систем хранения больших объемов данных были собраны и протестированы системы хранения данных (СХД) от трех производителей: IBM, HP, Adaptec. Структурная схема СХД на оборудовании IBM представлена на рис. 2.

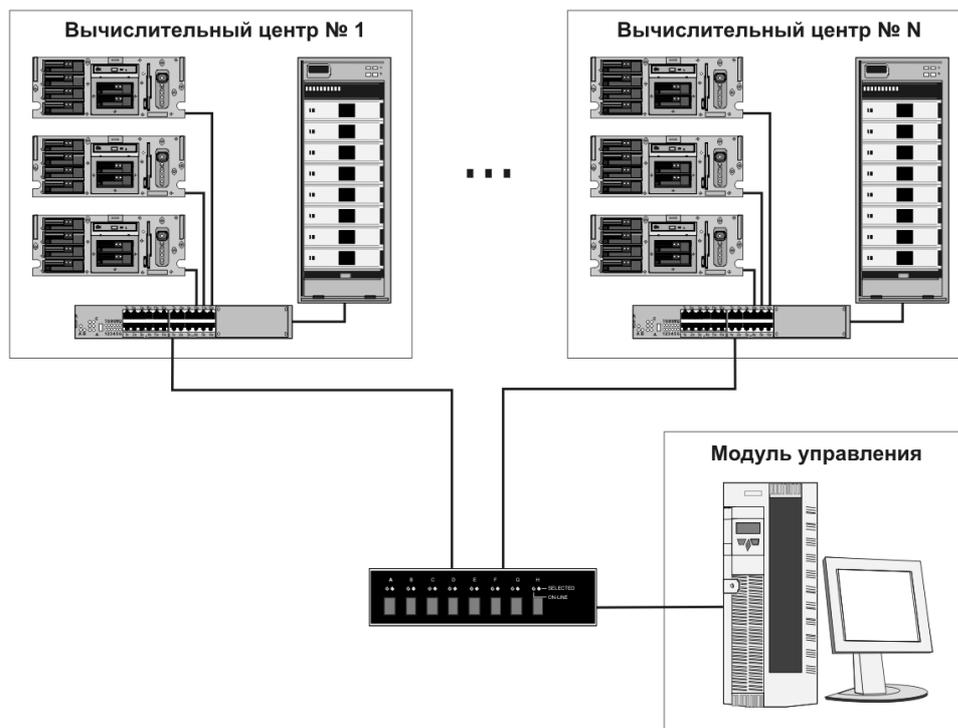


Рис. 1. Структурная схема СОИ

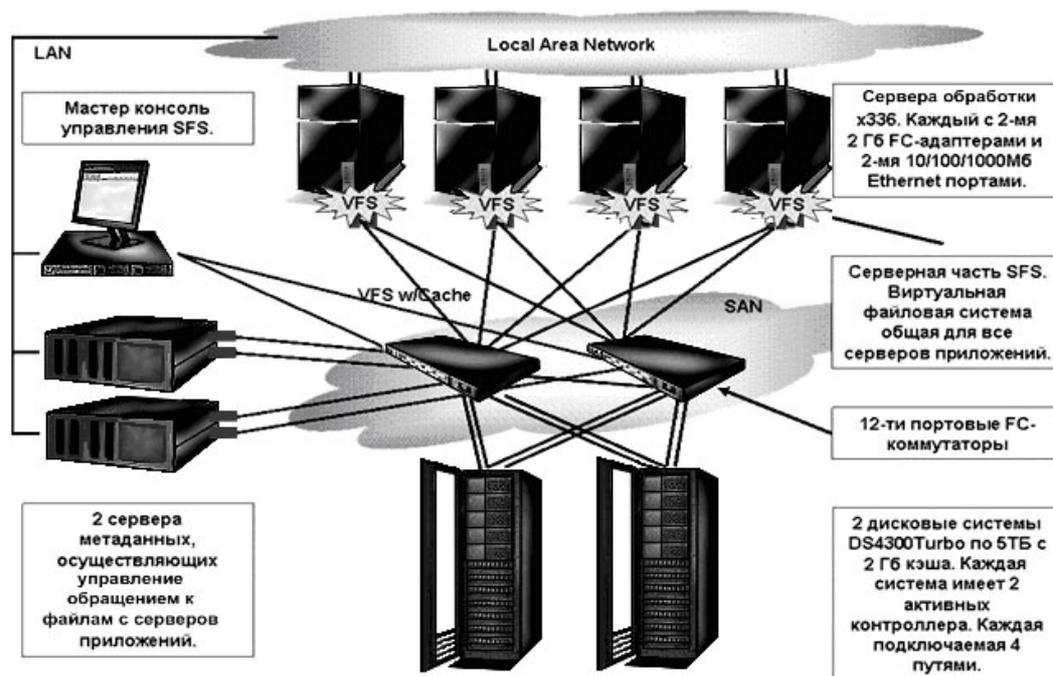


Рис. 2. Структурная схема СХД на оборудовании IBM

В состав системы хранения данных на базе оборудования IBM включены:

- IBM BladeCenter HS 20 (шасси блейд-серверов) с 7-ю блейд-серверами (спецификация: Intel Xeon 3.2 (2 шт.) / 4 Гб / IDE / 80 Гб (2 шт.), со встроенными FC-коммутаторами;
- SAN системы хранения IBM DS 4300 и DS 4500 емкостью по 2 Тб каждая.

На узлах обработки установлена операционная система MS Windows Server 2003 Enterprise Edition. Серверы метаданных IBM SAN FILE SYSTEM (SFS) функционируют под управлением ОС семейства Linux. На узлах обработки установлена клиентская часть системы IBM SFS, которая позволяет нескольким узлам обработки обращаться и работать с одним дисковым пространством. Использование программного обеспечения в таком режиме позволяет сформировать общий логический диск, обеспечивая «прозрачную» работу удаленным пользователям

Таким образом, IBM SFS представляет собой некое средство кластеризации.

Второй образец собран на оборудовании HP по схеме «без единой точки отказа». Структурная схема СХД на оборудовании HP представлена на рис. 3. Все соединения по Faber Channel выполнены с дублированием.

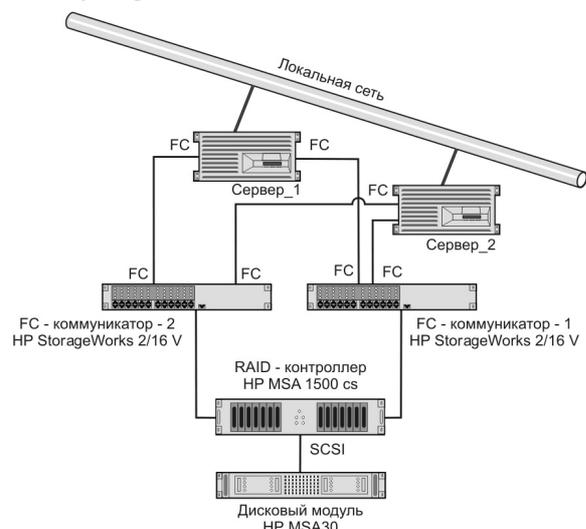


Рис. 3. Структурная схема СХД на оборудовании HP

Состав оборудования второго образца включает:

- Сервера – Intel Xeon 3.2 (2 шт.) / 4 Гб / SCSI 73 Гб (2 шт.);
- RAID-контроллер MSA 1500 sc с дисковым модулем MSA 30;
- FC-коммутаторы – HP StorageWorks Switch 2/16v.

В качестве третьего образца представлена высокопроизводительная система, собранная на оборудовании Adaptec, представленная на рис. 4.

Спецификация серверов: Intel Xeon 3.2 – 2 шт. / 4 Гб / SCSI 73 Гб – 2 шт.

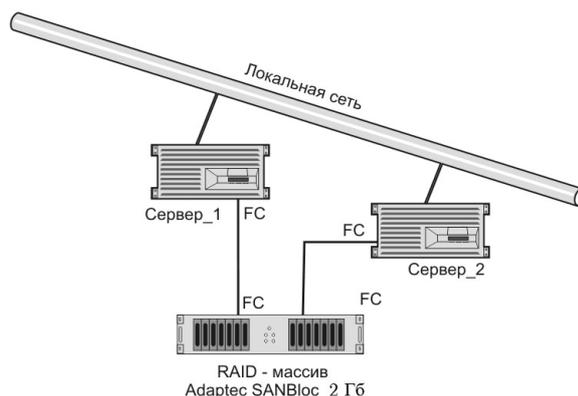


Рис. 4. Структурная схема СХД на оборудовании Adaptec

SAN система хранения имеет два независимых RAID-контроллера и дисковую емкость 2 Тб.

На серверах установлена ОС MS Windows 2003 Server Enterprise Edition. Узлы кластера (серверы) объединены в Windows-кластер на основе единой дисковой подсистемы. Дисковая подсистема подключена к узлам кластера через Faber Channel.

Ниже приведены таблицы и графики, отображающие основные результаты тестирования представленных систем. Тестирование проводилось на RAID-массивах с изменением уровней и количества используемых физических дисков.

Тестирование проводилось Intel IOMetr v.2004.07.30. В табл. 1 представлены результаты, полученные на оборудовании IBM, а в табл. 2 – на оборудовании HP и Adaptec.

Сравнив полученные результаты тестирования представленных СХД, можно сделать вывод, что система на оборудовании IBM показала максимальную производительность, а система на оборудовании HP – аутсайдер с минимальными показателями. Объединенные результаты измерений, проведенных с использованием блоков большого размера (443 Кбайт), представлены в табл. 3 и на рис. 5.

Структурные элементы VM СОИ обладают следующими характеристиками:

- **устройства ввода-вывода** – пропускной способностью ($r_{ввод}$, $r_{вывод}$);
- **устройства памяти** – объемом памяти ($C_{опер}$, $C_{диск}$);
- **процессоры** – быстродействием ($r_{проц}$).

Принимая во внимание особенности построения и специфику функционирования СОИ, в качестве частных показателей, характеризующих производительность VM, предлагается рассмотреть следующие:

- **для процессора** – количество выполняемых инструкций за единицу времени при решении;
- **для оперативной памяти** – количество выполняемых операций ввода вывода и чтения записи за единицу времени;
- **для дисковой подсистемы** – скорость чтения с диска и записи на диск данных за единицу времени;

Таблица 1. Результаты измерений производительности файловой системы на оборудовании IBM

No SFS					
RAID 0			RAID 1		
IO	IO ps	Mb ps	IO	IO ps	Mb ps
1	117,609589	51,498302	1	50,395337	21,684676
2	166,362153	72,196988	2	55,32962	23,744530
4	211,914938	70,317448	4	56,607866	24,471043
8	245,529549	107,757028	8	58,731007	24,514347
16	277,120047	117,027247	16	61,637836	26,453736
32	274,931667	119,216676	32	62,906162	27,487261
64	292,270001	126,362333	64	63,653648	27,733638
128	290,718859	126,252329	128	62,123263	26,793882
SFS					
RAID 0			RAID 1		
IO	IO ps	Mb ps	IO	IO ps	Mb ps
1	114,863608	49,922411	1	56,266927	24,769101
2	146,207977	62,882427	2	63,970722	27,161911
4	171,996315	74,718975	4	67,984951	28,111954
8	186,456324	81,691451	8	65,481424	28,218195
16	191,476866	82,777355	16	68,644108	29,028060
32	188,183353	80,812845	32	81,534497	36,893524
64	186,578630	79,441652	64	86,729082	36,303370
128	188,293087	79,957955	128	77,195697	33,218619
No SFS					
RAID 10			RAID 5		
IO	IO ps	Mb ps	IO	IO ps	Mb ps
1	90,4725150	40,521999	1	130,942946	55,786955
2	113,783247	48,755791	2	135,545668	59,956351
SFS					
RAID 10			RAID 5		
IO	IO ps	Mb ps	IO	IO ps	Mb ps
1	84,3017750	37,893433	1	116,42890	50,728821
2	106,372728	45,159202	2	146,20542	62,276511

Примечание:
 1. SFS – SAN file system.
 2. No SFS / SFS – тесты проводились с SFS и без нее.
 3. RAID 0 / 5 / 10 – использовались 4 диска, RAID 1–2 диска.
 4. IO – Outstanding I/O (кол-во одновременных потоков ввода-вывода в тесте).
 5. IO ps – Input/Output per second (количество операций ввода-вывода в с.)

- для файловой системы – количество операций чтения и записи файлов в единицу времени.

Таким образом, требуется найти значения системной производительности:

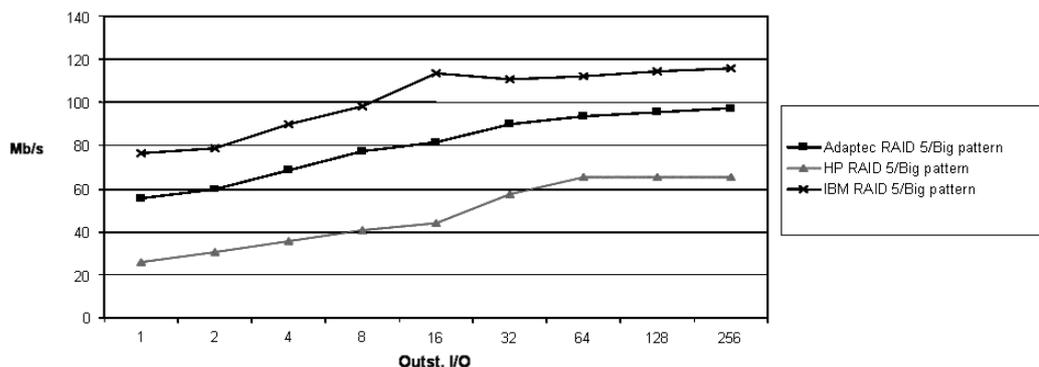


Рис. 5. График зависимости объема информации при операциях ввода-вывода в секунду от количества потоков

Таблица 2. Результаты измерений производительности файловой системы на оборудовании HP и Adaptec

HP MSA 1500CSi / 10 HDD					
RAID 10 / Database			RAID 5 / Database		
Outst. I/O	IO ps	Mb	Outst. I/O	IO ps	Mb
16	432,46	3,3786	16	935,67	7,3099
64	599,883	4,6866	64	943,97	7,3747
256	635,364	4,9638	256	978,46	7,6442
RAID 10 / FileServer			RAID 5 / FileServer		
Outst. I/O	IO ps	Mb	Outst. I/O	IO ps	Mb
16	400,788	4,2929	16	953,44	10,317
64	555,934	5,9592	64	1129,6	12,229
256	603,47	6,4804	256	1191	12,915
Adaptek SANBlock 2 Гб / 10 HDD					
RAID 10 / Database			RAID 5 / Database		
Outst. I/O	IO ps	Mb	Outst. I/O	IO ps	Mb
16	1707,16	13,337	16	1381,7	10,795
64	1709,39	13,355	64	1703,3	13,307
256	2024,59	15,817	256	1876,3	14,658
RAID 10 / FileServer			RAID 5 / FileServer		
Outst. I/O	IOPs	Mb	Outst. I/O	IOPs	Mb
16	1602,11	17,067	16	1350,4	14,667
64	1828,94	19,578	64	1827,3	19,774
256	1977,82	21,189	256	2041,7	22,057

Таблица 3. Сравнительные результаты тестирования оборудования различных производителей. IO означает количество операций ввода-вывода

IO	Adaptec RAID 5 / Big pattern	HP RAID 5 / Big pattern	IBM RAID 5 / Big pattern
1	55,786955	25,786340	76,592400
2	59,956351	30,694560	78,605400
4	68,783475	35,893240	89,745900
8	77,252749	40,785670	98,263500
16	81,803282	43,912625	113,459205
32	90,045720	57,542720	110,673200
64	93,692008	65,478850	112,012472
128	95,307823	65,337680	114,378200
256	97,402740	65,191516	116,076920

$$P=[p_1, V_1, p_2, V_2, p_3, V_3, p_4, V_4],$$

где p_1, p_2, p_3, p_4 – параметры загрузки устройств (процессора, оперативной памяти, дисковой системы, файловой системы соответственно), которые определяются при совместной работе технических средств под управлением операционной системы VM, а $V_1, V_2,$

V_3, V_4 – параметры, которые характеризуют значения быстродействия технических средств (процессора, оперативной памяти, дисковой системы, файловой системы соответственно). Под быстродействием предлагается понимать число операций, выполняемых ЭВМ и устройствами за единицу времени.

Показатель использования устройств (загрузка) определяется по следующей формуле:

$$p_i = T_i / T,$$

где T_i – время работы устройства, а T – общее время работы системы.

На основании исследований [1-3] и анализа реальных систем хранения и обработки больших объемов данных предлагается обобщенный алгоритм оценки производительности СОИ специального назначения, который состоит из следующих этапов:

1 этап. Определение состава, назначения и основных функций СОИ. Определение сценариев СПО и показателей производительности СОИ.

2 этап. Выбор подхода к оцениванию производительности:

1-й подход – использование стандартных средств;

2-й подход – использование специальных программных средств.

3 этап. Выполнение операций в зависимости от выбранного подхода – проведение испытаний, измерений и сравнение результатов.

4 этап. Определение критерия эффективности работы СОИ и ее производительности. На данном этапе осуществляется выбор конфигурации аппаратных средств для сравнительного тестирования.

5 этап. Обобщение полученных результатов.

Таким образом, созданный стенд, специальное алгоритмическое и программное обеспечение позволили получить значения характеристик для оценивания производительности отдельных вычислительных модулей, входящих в состав систем обработки информации и оценить потенциальную и реализованную производительности, а также имитацию процессов автоматической обработки информации.

СПИСОК ЛИТЕРАТУРЫ

1. Лацис А.О. Как построить и использовать суперкомпьютер. – М.: Бестселлер, 2003. – 274 с.
2. Воеводин В.В., Воеводин Вл.В. Параллельные вычисления. – СПб.: БХВ-Петербург, 2002. – 608 с.

3. Антонов А.С. Параллельное программирование с использованием технологии MPI. – М.: Изд-во МГУ, 2004. – 71 с.

Поступила 15.05.2009 г.

УДК 369:519.2

НЕПАРАМЕТРИЧЕСКОЕ ОЦЕНИВАНИЕ НЕТТО-ПРЕМИЙ ДЛЯ СМЕШАННОГО СТРАХОВАНИЯ ЖИЗНИ

Г.М. Кошкин, Н.В. Ланкина

Томский государственный университет
Отдел проблем информатизации ТНЦ СО РАН, г. Томск
E-mail: Lankina_Nata@mail.ru

Рассматривается задача оценивания нетто-премии в условиях смешанного страхования жизни. Синтезируется непараметрическая оценка нетто-премии, находится главная часть асимптотической среднеквадратической ошибки оценки и ее предельное распределение. Приводятся результаты статистического моделирования.

Ключевые слова:

Нетто-премия, смешанное страхование жизни, асимптотические свойства, непараметрические оценки.

Введение и постановка задачи

Эффективность финансовой деятельности страховой компании зависит от правильного расчета нетто-премии для различных видов страхования необходимых категорий и возрастных групп населения [1]. В долгосрочном страховании жизни при расчетах премий за риск учитывается динамика ценности денег, основанная на процентной ставке δ с непрерывно начисляемым процентом по

вкладу [2–6]. В этом случае для выработки управляющих решений страховой фирме следует предварительно оценить нетто-премию, которая гарантирует фирме средний нулевой доход.

Ранее в работах [3, 4] в условиях непараметрической неопределенности изучались оценки нетто-премий для различных видов индивидуального страхования, а в [5, 6] – в случае коллективного страхования. В данной работе рассматривается за-