

УДК 004.942: 004.715

А.М. ХОШАБА

Винницький національний технічний університет

ИССЛЕДОВАНИЕ ПРОЦЕССОВ НАГРУЗОЧНЫХ ВОЗДЕЙСТВИЙ И ВОССТАНОВЛЕНИЙ ПРОИЗВОДИТЕЛЬНОСТИ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

В работе рассматриваются важные вопросы производительности вычислительных систем как: цели и задачи, общая модель производительности, фундаментальные проблемы определения оценок производительности. Названы понятия и наиболее часто встречаемые причины возникновения утечки производительности процессора и памяти. Сформулирована постановка задачи исследования производительности вычислительных систем в результате использования веб-сервисов, приведены примеры скриптов нагрузочного воздействия, наведена их характеристика и особенности работы. Приведен пример рекомендаций по созданию сценариев воздействия на веб-сервис. Показано расчет оценок производительности сценария воздействия на веб-сервис, наведен график выполнения работы сценария.

Ключевые слова: производительность вычислительных систем, нагрузочное воздействие, бенчмарки, веб-сервисы.

О.М. ХОШАБА

Вінницький національний технічний університет

ДОСЛІДЖЕННЯ ПРОЦЕСІВ НАВАНТАЖУВАЛЬНИХ ВПЛИВІВ І ВІДНОВЛЕННЯ ПРОДУКТИВНОСТІ ОБЧИСЛЮВАЛЬНИХ СИСТЕМ

У роботі розглядаються важливі питання продуктивності обчислювальних систем як: цілі та завдання, загальна модель продуктивності, фундаментальні проблеми визначення оцінок продуктивності. Названі поняття і причини виникнення витоків продуктивності процесора та пам'яті що найбільш часто зустрічаються на практиці. Сформульовано постановку задачі дослідження продуктивності обчислювальних систем в результаті використання веб-сервісів, наведені приклади скриптів навантажувального впливу, наведена їх характеристика та особливості роботи. Наведено приклад рекомендацій по створенню сценаріїв впливу на веб-сервіс. Показано розрахунок оцінок продуктивності сценарію впливу на веб-сервіс, наведений графік виконання роботи сценарію впливу.

Ключові слова: продуктивність обчислювальних систем, навантажувальний вплив, бенчмарки, веб-сервіси.

O.M. KHOSHABA

Vinnitsia National Technical University

INVESTIGATION OF PROCESSES OF LOADING EFFECTS AND RECOVERY OF COMPUTER SYSTEMS PERFORMANCE

In the article important issues of the performance of computing systems are considered such as: goals and tasks, a general performance model, fundamental problems in determining performance estimates. The concepts and the most common reasons for a processor and memory leak are identified. The statement of a research problem for processor and memory perform leak as a result of uses web services is formulated, examples of scripts of loading influence are

resulted, their characteristic and features of work are directed. An example of recommendations for creating scenarios of the impact on a web service is given. The calculation of the performance evaluations of the impact scenario on the web service is shown, the schedule for the performance of the script is plotted. An important part for study of the performance of computing systems was the development of benchmarks. The main lacks of existing benchmarks include the absence of: load and recovery models, their combined impact, distributed sources of impact. To study the performance of of computing systems in most cases create exposure scenarios. Such scenarios by the type of impact can have different dynamic (velocity) estimates and have different mathematical models. The study of the performance of computer systems using the example of a web service was performed in the operating system Linux Centos 7.0, kernel 3.10, the processor Intel Core i7-3667U, 2.0GHz, with 8 GB of RAM. The software was an apache web server. The load was effected using the httpperf utility, performance monitoring was sar, iostat, vmstat. Estimates of the performance of the exposure scenario were verified using mathematical modeling in the SciLab environment using special scripts. Also, the task of researching the performance of computing systems on the example of using web services is formulated, examples of stress loading scripts are given, their characteristics and work features are indicated. Particular importance is given to the model of impact formation. An example of recommendations for creating scenarios of the impact on a web service is given.

Keywords: computer system performance, benchmarks, web services, sar, iostat, vmstat, SciLab.

Постановка проблеми

Основой производительности вычислительных систем (ВС) является моделирование нагрузочных воздействий (НВ) и восстановлений [1]. Производительность ВС может рассматриваться как самостоятельная дисциплина, так и в основе множества прикладных аспектов практической инженерии. Проектирование, разработка и сопровождение информационных систем (ИС) также невозможна без решения вопроса производительности ВС [2]. Эффективная разработка и сопровождение ИС в области изучения производительности требуют: определение модели предметной области и формирование спецификаций к проектированию данных систем; создание и генерацию заданных тестовых нагрузок; построение системы с использованием жизненного цикла и определения бизнес-процессов функционирования основных компонент ВС.

Разработка математических моделей и методов производительности ВС ставит своей целью не только выполнение моделирования процессов, происходящих в ВС, но и управление ими, устранение различных опасных последствий негативных воздействий [3].

Анализ последних исследований и публикаций

К основным задачам которые решаются с помощью моделей производительности ВС и программного обеспечения (ПО) в основном относятся следующие: оптимизация показателей выполнения ПО в среде ВС; поиск проблемных областей (“узкого места”) в работе (производительности) ВС или ПО; обнаружение аномалий в работе ВС или ПО, оценка работы прикладных программ, сервисов, систем управления базами данных, операционных систем, виртуальных и облачных информационных структур [3].

Общая модель производительности ВС может быть представлена в виде [1]:

$$k_{pn} \frac{d^n P}{dt^n} + \dots + k_{p1} \frac{dP}{dt} + k_0 P = k_{qm} \frac{d^m Q}{dt^m} + \dots + k_{q1} \frac{dQ}{dt} + k_0 Q \quad (1)$$

Фундаментальные проблемы определения оценок производительности ВС связаны с построением концептуальной математической модели, отбором и уменьшением пространства признаков при построении прогностических коэффициентов математических моделей, разработкой критериев определения состояний объектов и субъектов исследований для принятия оптимальных решений.

Проблеми производительности ВС, в частности утечка памяти и процессора, могут возникать в следствии самых разных причин. В общем случае, утечка производительности процессора представляет собой ситуацию (рис. 1), при которой в результате сбоев или целенаправленных действий происходит ненужная (лишняя, не контролируемая) работа одного из ядер процессора. Такая работа может сопровождаться чередованием загрузки ядер до 99-100% в результате которой не наблюдается зависания системы. Такая ситуация плохо контролируется пользователями по причине ее возникновения так как не наблюдается работа сторонних программ. Часто, при нахождении причины, вызвавшей такую ситуацию, невозможно удалить процесс, который поддерживает утечку, в результате чего приходится перезагружать систему.

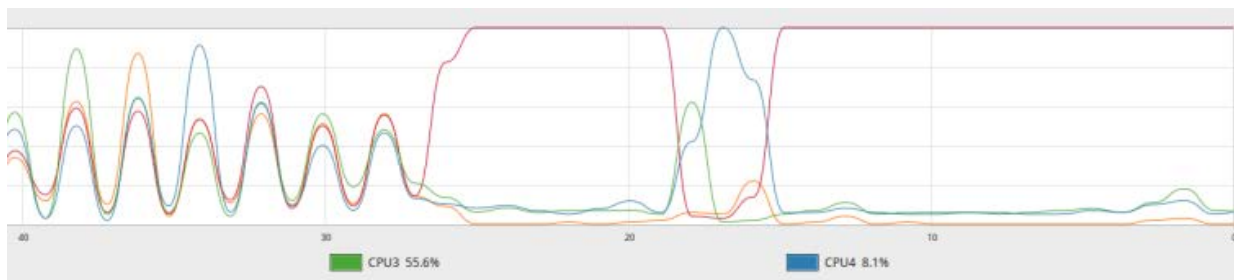


Рис. 1. Ситуация утечки производительности процессора.

К другим, наиболее распространенным из причин утечки производительности процессора и памяти могут являться заражение компьютерными вирусами и другими вредоносными программами, ошибки в работе операционных и файловых систем, неправильные пользовательские и системные настройки операционных систем, неправильная установка антивирусных и ресурсоемких программ (плагинов) на браузерах. В приведенных примерах в большинстве случаев отмечается ситуация утечки производительности процессора. Однако, на практике более известной и описанной проблемой среди пользователей является утечка памяти в процессе работы или завершения программ.

Цель исследования

Постановка задачи. Существует множество инструментов нагрузочного тестирования ВС и ПО. Однако, к их недостаткам следует отнести практическое отсутствие описания математической модели (основных параметров) НВ, к которым относятся: функциональная характеристика нагрузки (полиномы n -ой степени, степенные, экспоненциальные, показательные) и временные промежутки формирования отдельных серий запросов, общее время НВ (или количество серий запросов). Данные параметры нагрузочного воздействия существенно влияют на динамические характеристики работы ВС и ПО. Поэтому, к постановке задачи исследования процессов НВ и восстановлений относится:

- 1) создание и выполнение генерации сценария воздействия на веб-сервис с использованием функции полинома третьей степени;
- 2) определение модели работы веб-сервиса с ресурсами ВС;
- 3) выполнение исследования и определение оценок производительности ВС на примере веб-сервиса.

Изложение основного материала исследования

Теоретические сведения

Аппаратные средства ВС имеют определенную потенциальную составляющую в виде статических величин, которые позволяют различать программно-аппаратные комплексы (ПТК) друг от друга по возможности решения различных сложных. Обычно

пользователей интересуют потенциальные возможности в работе таких аппаратных компонентов ПТК как процессоры (такты частота), сетевые устройства (пропускная способность каналов передачи данных), видеоадаптеры (скорость прорисовки пикселей на экране), дисковые устройства (скорость чтения и записи данных).

При решении ряда задач ПО использует различные возможности этих и других аппаратных средств. Поэтому, такое динамическое использование (потребление) ресурсов аппаратных средств представляет собой кинетическую (динамическую) составляющую. Кинетическая (динамическая) составляющая использования аппаратных средств выражается функцией. Значения потенциальной и кинетической составляющей могут быть представлены в виде абсолютных и относительных величин.

Работа прикладной или системной задачи может быть представлена определенным воздействием на ресурсы ВС. По завершении задачи ресурсы, выделенные для ее решения, через некоторое время освобождаются. Поэтому воздействие задачи на ресурсы системы в виде потребления и расходования составляют два основных процесса производительности ВС: воздействия и восстановления. Часто, процессы воздействия и восстановления определяются как состояния или фазы функционирования объектов ПТК.

Для исследования производительности ВС выделяют субъектно-объектную модель которая состоит в следующем. В каждый фиксированный момент времени ВС представляет собой конечное множество компонентов, разделяемых по критерию активности на два подмножества: подмножество субъектов (S) и объектов (O) таких, что:

$$S \cap O = \emptyset; \quad (2)$$

В любой момент времени t_k подмножество субъектов и объектов не является пустым:

$$S \neq \emptyset; \quad (3)$$

$$O \neq \emptyset; \quad (4)$$

Объектом исследования процессов производительности ВС является аппаратная часть программно-технического комплекса (ПТК). К программной части ПТК принадлежат субъекты исследования процессов производительности ВС. В любой момент времени t при $\Delta t \rightarrow 0$ объект (O) может принадлежать только одному субъекту (S).

Определение оценок производительности вычислительных систем

Оценки производительности ВС находятся в виде скалярных (количественных и качественных) и векторных величин, выражаются в виде функций [4]. К примеру, пусть в результате экспериментальных исследования по изучению сценария воздействия на заданный ресурс (устройства) получена с помощью инструментов мониторинга некоторая функция $v_T(t)$. Данная функция представляет собой скорость выполнения работы некоторой задачи (в данном случае сценария воздействия). Тогда, выполненную работу можно представить в виде функции:

$$w_T(t) = \int v_T(t) dt \quad (5)$$

Скорость изменения протекания процессов определяется из выражения:

$$v_T(t) = w'_T(t) \quad (6)$$

$$a_T(t) = v'_T(t) \quad (7)$$

$$a_{TD}(t) = a'_T(t) \quad (8)$$

Оценки скорости протекания процесса или скорости выполнения задачи (9), изменения скорости протекания процесса (10) и выполненной работы (11) получаем следующим образом:

$$V_T = \int_{t_1}^{t_2} a_T(t) dt \quad (9)$$

$$A_T = \int_{t_1}^{t_2} a_{TD}(t) dt \quad (10)$$

$$W_T = \int_{t_1}^{t_2} v_T(t) dt \quad (11)$$

К показателям производительности ВС, которые тесно связаны с параметрами функционирования субъектов (объектов) ВС относятся [2,4]: пропускная способность (throughput), время ответа (response time), утилизация (utilization) и доступность (availability). Данные оценки и показатели производительности могут быть применены для сценариев воздействия на объект (субъект) исследования.

Создание сценариев воздействия на веб-сервис

Важной составной частью исследования производительности ВС является разработка сценариев (бенчмарков). К основным недостаткам существующих бенчмарков относятся отсутствие: моделей нагрузки и восстановления, их комбинированного воздействия, распределенных источников воздействия.

Для исследования производительности ВС в большинстве случаев создают сценарии воздействия. Такие сценарии по типу воздействия могут иметь различные динамические (скоростные) оценки и характеризоваться следующей моделью:

$$D_{LOAD} = k_A \frac{d^2 Q}{dt^2} + k_V \frac{dQ}{dt} + k_W f_Q(t) \quad (12)$$

В этом случае, равнопеременные при $k_A \neq 0$ считаются равноускоренные (при $k_A > 0$) или равнозамедленные (при $k_A < 0$) нагрузочные воздействия. При $k_A = 0$ сценарии воздействия будут считаться равномерными по генерации запросов к субъекту (объекту) исследования.

Сформируем сценарии воздействия по первому случаю как совокупность известной функциональной зависимости групп запросов (Force of Impact, F_{Impact}) и временных интервалов (Pause, P_s) между ними:

$$F_{Common} = F_{Impact} + P_s \quad (13)$$

Для создания сценария воздействия можно принять следующую функциональную зависимость:

$$F_{Impact} = 500 * t \text{ [запросов / с]} \quad (14)$$

Тогда временные промежутки сценариев воздействия будут следующими:

для $S0 - 0$ с, $S1 - 1$ с, $S2 - 2$ с, $S3 - 3$ с.

Целью проведения равнопеременных воздействий является определения состояний объекта (субъекта) исследования. К примеру, состояния веб-сервиса при воздействии по сценарию $S0$ показаны в табл. 1. Различие в границах количества запросов определяется аппаратно-программным окружением и нагрузкой на ВС.

Целью проведения равномерных и равнопеременных воздействий на субъект исследования является определение основных процессов производительности ВС: НВ и

восстановления. В общем случае, производительность также характеризуется временным (или постоянным) прекращением или существенным понижением обработки данных объектом исследования в результате которого понижается его доступность по направляемым к нему запросам.

К примеру, скрипт, реализующий последовательность запросов по первому случаю (сценарий *S0*) показан на рис. 1а. Второй случай представлен также скриптом, показанным на рис. 1б. В случае тестирования, скрипты пишутся и выполняются в среде *bash* программирования. Для автоматизации исследований программы составляются на языках высокого уровня, к примеру, по настоящему исследованию – на *python*.

Таблица 1

Характеристика состояний субъекта исследования (веб-сервиса)

Название состояния	Количество запросов (10^3 c^{-1})	Критерий определения состояния
Стабильное (нормальное)	от 1 до 2,5	Отсутствие очереди запросов, отсутствие ошибок
Нагрузочное	от 2,5 до 8	Появление очереди запросов, отсутствие ошибок
Стрессовое	свыше 8	Появление ошибок, потеря запросов

<pre>#!/bin/bash for i in {100..9000..500} do httpperf --hog --server localhost --uri "/" -- num-conn \$i --num-call 1 --timeout 5 --rate \$i --port 80 done</pre> <p style="text-align: center;">а)</p>	<pre>#!/bin/bash for i in {1..180} do httpperf --hog --server localhost --uri "/" -- num-conn 1000 --num-call 1 --timeout 5 -- rate 1000 --port 80 done</pre> <p style="text-align: center;">б)</p>
--	---

Рис. 2. Bash скрипты, реализующие различные последовательности запросов к веб-сервису: а) равнопеременные (равноускоренные) запросы; б) равномерные запросы

Результаты исследования производительности ВС на примере веб-сервиса

Исследование производительности ВС на примере веб-сервиса выполнялось в операционной системе Linux Centos 7.0, ядре 3.10, процессоре Intel Core i7-3667U, 2.0GHz, оперативная память 8ГБ. Программное обеспечение: веб-сервер *apache*. Нагрузочное воздействие осуществлялось с помощью утилиты *httperf*, мониторинг - с помощью *sar*, *iostat*, *vmstat*.

Для скрипта (рис. 2а) согласно оценкам (6-8) принималась следующая модель:

$$v_T(t) = 0.1428t^3 - 5.2727t^2 + 553.2215t - 492.2981; \quad (15)$$

$$a_T(t) = 0.4284t^2 - 10.5454t + 553.2215; \quad (16)$$

$$w_T(t) = 0.0357t^4 - 1.7576t^3 + 276.6108t^2 - 492.2981t. \quad (17)$$

Данная модель (15) для $v_T(t)$ является полиномом третьей степени и имеет следующие характеристики качества аппроксимации:

Индекс корреляции $R = 0.9994$;

Индекс детерминации $R^2 \approx 0.9987$;

Средняя ошибка аппроксимации $A \approx 3.7013\%$

Оценки производительности сценария воздействия проверялись с помощью

математического моделирования в среде SciLab с помощью скрипта в сокращенном виде (рис. 3а). Результаты работы скрипта определения динамических оценок производительности сценария воздействия $S0$ показаны на рис. 3б: количество выполненной работы задачей ($w_T(t)$) – верхняя линия, скорость выполнения процесса ($v_T(t)$) – нижняя.

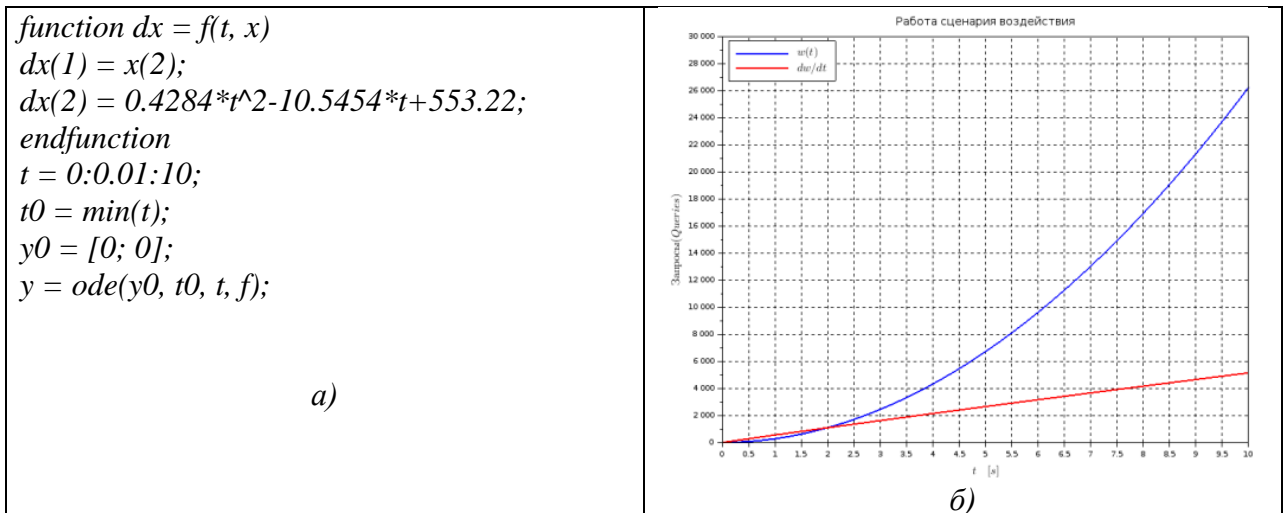


Рис. 3. Скрипт среды окружения SciLab, позволяющий получить оценки производительности сценария воздействия $S0$ а) и результат его работы б)

Выводы

В работе рассмотрены важные вопросы производительности вычислительных систем к которым относятся цели и задачи, общая модель производительности, фундаментальные проблемы определения оценок производительности. Названы наиболее часто встречаемые причины возникновения утечки производительности процессора и памяти. Сформулирована постановка задачи исследования производительности вычислительных систем на примере использования веб-сервисов, приведены примеры скриптов нагрузочного воздействия, наведена их характеристика и особенности работы. Особое значение уделено модели формирования воздействия. Приведен пример рекомендаций по созданию сценариев воздействия на веб-сервис.

Показана оценка производительности сценариев воздействия на веб-сервис, наведены графики выполнения работы на основе скриптов, написанных в среде SciLab.

Список использованной литературы

1. Хошаба А.М. Моделирование состояний при нагрузочных воздействиях на сервисы вычислительной системы / А.М. Хошаба // XIII Міжнародна науково-практична конференція "Математичне та імітаційне моделювання систем": Тези допов. – Київ, 2018. – С. 269-272.
2. Хошаба А.М. Концептуальная модель производительности вычислительных систем / А.М. Хошаба // XI Міжнародна науково-практична конференція "Математичне та імітаційне моделювання систем": Тези допов. – Київ, 2016. – С. 417-420.
3. Хошаба О.М. Дослідження надійності функціонування систем з керування потоками даних у комп'ютерних мережах / А.М. Хошаба // II Міжнародна науково-практична конференція "Інформаційні технології в освіті, науці і техніці": Тези допов. – Черкаси, 2014. – С. 107-108.

4. Хошаба А.М. Построение информационной системы для компьютерного моделирования производительности программно-технического комплекса / А.М. Хошаба // Вестник Херсонского национального технического университета. – 2015. – №3(56). – С. 303-307.

References

1. Khoshaba, A. M. Modelirovanie sostoyaniy pri nagruzochnykh vozdeystviyah na servisii vyichislitelnoy sistemyi. XIII Mizhnarodna naukovo-praktychna konferentsiia "Matematychnye ta imitatsiine modeliuvannia system": Tezy dopov. Kyiv. 269-272. (2018)
2. Khoshaba, A. M. Kontseptualnaya model proizvoditelnosti vyichislitelnykh system. XI Mizhnarodna naukovo-praktychna konferentsiia "Matematychnye ta imitatsiine modeliuvannia system": Tezy dopov. Kyiv. 417-420. (2016)
3. Khoshaba, A. M. Doslidzhennia nadiinosti funktsionuvannia system z keruvannia potokamy danykh u kompiuternykh merezhakh. II Mizhnarodna naukovo-praktychna konferentsiia "Informatsiini tekhnolohii v osviti, nautsi i tekhnitsi": Tezy dopov. Cherkasy. 107-108. (2014)
4. Khoshaba, A. M. Postroenie informatsionnoy sistemyi dlya kompyuternogo modelirovaniya proizvoditelnosti programmno-tehnicheskogo kompleksa. Vestnik Hersonskogo natsionalnogo tehnicheskogo universiteta. 3(56), 303-307. (2015)