

МЕТОДИКА РАСПРЕДЕЛЕНИЯ ПРИКЛАДНЫХ ЗАДАЧ ПО ВИРТУАЛЬНЫМ МАШИНАМ С УЧЕТОМ НАДЕЖНОСТИ ИХ РЕШЕНИЯ

А. Г. БАСЫРОВ, Д. И. КАЗАНЦЕВ

*Военно-космическая академия им. А. Ф. Можайского, 197198, Санкт-Петербург, Россия
E-mail: zilan.ka@yandex.ru*

Рассмотрена задача оптимального использования ограничения ресурсов памяти и обеспечения надежности вычислительных процессов. Разработан алгоритм распределения прикладных задач по виртуальным машинам. Предложена методика определения рационального состава виртуальных машин.

Ключевые слова: виртуальные машины, распределение прикладных задач, вероятность безотказной работы

Введение. В связи с ростом вычислительных мощностей современных вычислительных систем в настоящее время все большую популярность набирают технологии виртуализации [1]. Обладая мощными физическими серверами, можно перенести в виртуальную среду все информационные и вычислительные ресурсы организации, создать виртуальные рабочие места [2, 3]. Технологии виртуализации становятся одним из ключевых компонентов современной ИТ-инфраструктуры организаций благодаря экономии средств и времени на эксплуатацию систем, высокому уровню безопасности и обеспечению непрерывности прикладных вычислительных процессов [4, 5].

В центрах обработки данных с использованием технологии виртуализации организуют работу публичных и частных вычислительных облаков [6] для удовлетворения потребностей организаций и отдельных пользователей в информационных и вычислительных ресурсах. Универсальность технологий виртуализации [7—11] вкуче с миниатюризацией аппаратных средств позволяют применять их не только на уровне организаций, но и на уровне отдельных технических систем, в которых реализуются многозадачные режимы обработки данных.

В работах [12—17], посвященных применению технологий виртуализации, исследуются вопросы оптимального распределения ограниченных вычислительных ресурсов физических серверов по виртуальным машинам. Однако в них при планировании загрузки серверов не учитывается надежность прикладных программ, функционирующих в виртуальной среде.

При развертывании виртуальных машин неизбежно решается ряд задач, связанных с оптимизацией использования ограниченных ресурсов памяти и обеспечением надежности вычислительных процессов. В настоящей работе представлен подход к подбору состава виртуальных машин для множества прикладных программ с учетом требований к надежности функционирования машин и ограничений на объем доступной памяти.

Постановка задачи рационального распределения прикладных задач по виртуальным машинам. Дано: множество прикладных программ $F = \{f_1, f_2, f_3, \dots, f_N\}$, каждая из которых характеризуется требованиями к размеру оперативной памяти $v(f_i)$ и вероятностью безотказной работы (ВБР) p_i на ЭВМ.

Требуется: определить рациональное число k виртуальных машин (ВМ) M_1, M_2, \dots, M_k , запускаемых на ЭВМ, и набор прикладных программ, одновременно выполняемых на каждой ВМ, при которых

- 1) обеспечивается одновременное выполнение каждой программы;

2) суммарные затраты оперативной памяти $V = \sum_{i=1}^k v(M_i)$, выделяемой ВМ, минимальны.

Здесь $v(M_i) = \sum_{j=1}^L v(f_j)$, L — число назначенных на ВМ M_i программ;

3) надежность $p(M_i)$ функционирования каждой ВМ не ниже требуемой — p_0 .

Пример состава ВМ, с распределением по ним прикладных программ, представлен на рис. 1. Отметим, что в состав каждой ВМ обязательно входит операционная система S .

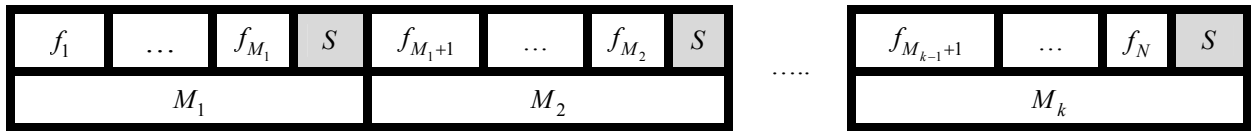


Рис. 1

Требование минимизации объема памяти, выделяемой ВМ, и неизбежность затрат памяти каждой ВМ как на операционную систему, так и другие программные компоненты, приводят к необходимости минимизировать количество ВМ.

С другой стороны, увеличение числа прикладных программ на одной ВМ неизбежно приводит к снижению надежности ее функционирования. Это связано с „зависанием“, сбоями в работе приложений ВМ в случае нарушения обработки одной из программ. Отсюда возникает противоречие между требованиями снижения затрат памяти, выделяемой ВМ, и требованиями к надежности их функционирования.

Для разрешения этого противоречия требуется определить минимальное количество виртуальных машин и распределить по ним прикладные программы с учетом надежности их функционирования.

Формализуем задачу, описывающую эту ситуацию. Найти минимальное значение k и множество булевых коэффициентов α_{ij} , $i = 1, \dots, N$, $j = 1, \dots, k$, таких, что

$$p(M_j) = \prod_{i=1}^N (1 - \alpha_{ij}(1 - p_i)) \geq p_d, \quad \forall j = 1, \dots, k, \quad (1)$$

$$\sum_{j=1}^k \alpha_{ij} = 1, \quad \forall i = 1, \dots, N, \quad (2)$$

где $\alpha_{ij} = 1$, если f_i выполняется на ВМ M_j ; $\alpha_{ij} = 0$ — в противном случае.

Первое условие требует, чтобы вероятность совместного выполнения всех прикладных программ на каждой ВМ была не ниже заданной, а второе — чтобы каждая прикладная программа была выполнена на одной из ВМ.

С математической точки зрения задачу можно сформулировать следующим образом: разбить группу положительных чисел, меньших единицы, на минимальное количество подгрупп так, чтобы их произведение в каждой подгруппе было не меньше заданного числа.

Алгоритм распределения прикладных задач по виртуальным машинам. Поставленная задача относится к классу NP -полных. Трудоемкость ее решения при полном переборе значений искомых коэффициентов составляет k^N .

Результат решения задачи может быть представлен в виде таблицы (табл. 1), причем в каждой строке только один из коэффициентов α_{ij} имеет значение 1, а значение $p(M_j)$, вычисляемое как произведение соответствующих значений p_i по каждому столбцу (ВМ) не менее чем p_d .

Таблица 1

Программа	ВБР	ВМ			
		1	2	...	k
f_1	p_1	α_{11}	α_{12}	...	α_{1k}
f_2	p_2	α_{21}	α_{22}	...	α_{2k}
f_3	p_3	α_{31}	α_{32}	...	α_{3k}
f_4	p_4	α_{41}	α_{42}	...	α_{4k}
f_5	p_5	α_{51}	α_{52}	...	α_{5k}
...
f_N	p_N	α_{N1}	α_{N2}	...	α_{Nk}

Для приближенного решения задачи, применяемого с целью сокращения времени расчета, можно предложить следующий эвристический алгоритм.

1. Нумерация программ в порядке невозрастания их ВБР:

$$p_1 \geq p_2 \geq p_3 \geq \dots \geq p_{N-1} \geq p_N.$$

2. Положить $k=0$.

3. $k:=k+1$. Если $k>N$, то переход к шагу 12.

4. $\beta:=1$, $p(M_j)=1$, $\alpha_{ij}=0$, $i=1, \dots, N$, $j=1, \dots, k$.

5. $i:=0$, $j:=0$.

6. $i:=i+1$. Если $i>N$, то переход к шагу 13.

7. $j:=j+\beta$.

8. $\alpha_{ij}:=1$.

9. $p(M_j):=p(M_j)p_i$. Если $p(M_j) < p_d$, то переход к шагу 3.

10. Если ($j=k$ и $\beta=1$) или ($j=1$ и $\beta=-1$), то $j:=j+\beta$, $\beta:= -\beta$.

11. Переход к шагу 6.

12. Выдача сообщения о невозможности решения задачи.

На рис. 2 показан порядок распределения шестнадцати программ по четырем ВМ в соответствии с представленным алгоритмом. Стрелками на рисунке показано направление размещения упорядоченных по неубыванию ВБР программ (увеличению их номеров).

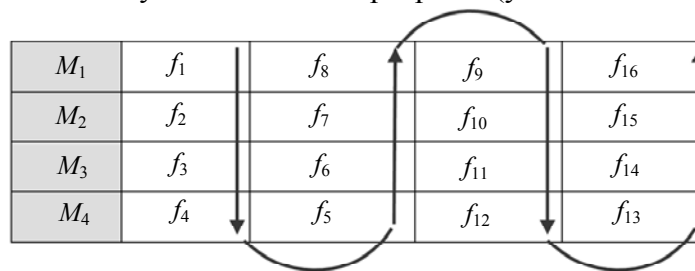


Рис. 2

Исследование эффективности предложенного алгоритма показало, что максимальное отклонение от оптимального значения ВБР комплекса программ, распределенных на ВМ по алгоритму, не превышает 0,025 с доверительной вероятностью 0,99. Моделирование проводилось методом Монте-Карло для 2—10 виртуальных машин и для 10—120 прикладных программ. Вычислительная сложность алгоритма $O(N^2)$, где N — число программ.

Методика определения рационального состава виртуальных машин

1. На основании данных о состоянии аппаратного и программного обеспечения, полученных системой мониторинга сети [18], определить ВБР каждой прикладной программы для выбранного периода t ее функционирования:

$$p_i = e^{-\lambda_i t}, \quad (3)$$

где p_i — ВБР i -й программы в течение времени t (в часах), λ_i — интенсивность сбоев (отказов) i -й программы — число сбоев задачи за единицу времени (час) — определяется на основе статистики функционирования задачи на продолжительном интервале времени (день, месяц, год).

2. Задать директивную ВБР p_d виртуальной машины.

3. Если существуют программы, для которых $p_i \leq p_d$, то каждая такая программа устанавливается на отдельную ВМ и в последующем распределении задач по ВМ не участвует.

4. При незначительном числе программ (≤ 5) можно применить метод полного перебора для распределения задач по ВМ, в противном случае — алгоритм, описанный выше.

5. В связи с изменением во времени (3) ВБР задач необходимо периодически переназначать задачи по ВМ. Для этого следует спланировать распределение задач на ближайший период. Период планирования выбирается исходя из интенсивности сбоев задач.

Пример. В табл. 2 представлены сведения по отказам прикладных программ (ПП), на основании которых оценена ВБР программ на три месяца.

Таблица 2

Сведения по отказам прикладных программ и значения ВБР

№	Прикладные программы	Код ПП	Среднее число отказов в год	$\lambda, 10^5/\text{ч}$	Вероятность безотказной работы ВМ		
					январь	февраль	март
1	Web-сервер	f_1	0,3	3,42	0,97484	0,95266	0,92870
2	Сервер контроллера домена, DNS, DHCP	f_2	0,2	2,28	0,98316	0,96819	0,95188
3	Почтовый сервер	f_3	0,4	4,57	0,96660	0,93739	0,90608
4	Файловый сервер	f_4	0,05	0,57	0,99576	0,99195	0,98775
5	Система резервного копирования	f_5	0,15	1,71	0,98734	0,97605	0,96369
6	Система мониторинга ЛВС	f_6	0,1	1,14	0,99154	0,98397	0,97564

Применим рассмотренный выше алгоритм распределения ПП по ВМ для исходных данных (табл. 2) на одну (рис. 3, в), две (б) и три (а) машины.

а)

ВМ	Код ПП		ВБР ПП на ВМ
M_1	f_4	f_3	0,89498
M_2	f_6	f_1	0,90608
M_3	f_5	f_2	0,91732

б)

ВМ	Код ПП			ВБР ПП на ВМ
M_1	f_4	f_2	f_1	0,87318
M_2	f_6	f_5	f_3	0,85191

в)

ВМ	Код ПП						ВБР ПП на ВМ
M_1	f_4	f_6	f_5	f_2	f_1	f_3	0,76243

Рис. 3

Заключение. Предложенный подход позволяет оптимизировать распределение прикладных программ по виртуальным машинам с учетом надежности их функционирования. Методика может быть с успехом применена и при планировании назначения программ по физическим вычислителям (микроконтроллерам, компьютерам, серверам).

Использование методики позволит обоснованно выбирать состав виртуальных машин для многозадачных вычислительных систем.

СПИСОК ЛИТЕРАТУРЫ

1. Виртуализация приходит в массы // Журнал сетевых решений/LAN [Электронный ресурс]: <<http://www.osp.ru/lan/2008/04/4964875/>>.
2. Васильев А. С., Казанцев Д. И., Мещеряков С. Е. Обеспечение отказоустойчивости сервисов инфо-телекоммуникационной среды вуза на основе построения единого центра обработки данных с применением технологии виртуализации // Научный обозреватель. 2015. № 4. С. 59—63.
3. Ковригин С. Н., Янюшкин В. В. Основы анализа IT-инфраструктуры открытого контура автоматизированной системы информационного обеспечения подготовки космонавтов для создания интегрированного тренажерного комплекса // Программные продукты и системы. 2015. № 4. С. 16—21.
4. Казанцев Д. И., Мещеряков С. Е., Немиров А. В. Подходы к обеспечению функционирования инфо-телекоммуникационной системы „электронный вуз“ // Universum: технические науки. 2015. № 4—5(17). С. 3.
5. Эсаулов К. А., Забузов В. С., Казанцев Д. И. Повышение доступности сервисов путем создания реконфигурируемой системы с использованием средств виртуализации // Вестн. Российского нового университета. 2015. № 10. С. 83—86.
6. Басыров А. Г., Казанцев Д. И. К вопросу создания инфотелекоммуникационной инфраструктуры вуза // Вестник компьютерных и информационных технологий. 2013. № 2(104). С. 37—39.
7. Романова А. О. Виртуализация в высокопроизводительных вычислительных системах // Наука и образование: науч. издание МГТУ им Н.Э. Баумана. 2011. № 3. С. 12.
8. Коуров А. В. Современные гипервизоры как основа инфраструктуры образовательной организации // Вестн. Шадринского гос. педагогического ин-та. 2015. № 1(25). С. 142—148.
9. Снытко А. С., Терников А. А. Особенности модернизации центра обработки данных и космоцентра // Программные продукты и системы. 2015. № 4. С. 22—27.
10. Назиров Р. Р., Коноплев В. В. „Вычислительное облако“ как средство эффективной организации вычислительных ресурсов в центре обработки научных данных // Информационное общество. 2013. № 1—2. С. 17—25.
11. Технология облачных вычислений. Мир телекома [Электронный ресурс]: <<http://mirtelecoma.ru/magazine/elektronnaya-versiya/31>>.
12. Козловский А. Л. Модели, методы и алгоритмы распределения ресурсов виртуализованных вычислительных кластеров. Дис. ... канд. техн. наук. М.: Научно-исследовательский и экспериментальный ин-т автомобильной электроники и автотракторного электрооборудования, 2012.
13. Ролика А. И., Боданюк М. Е. Алгоритмы последовательного анализа вариантов в задаче распределения виртуальных машин по физическим серверам в центрах обработки данных // Адаптивні системи автоматичного управління. 2012. № 21(41). С. 61—69.
14. Вдовин П. М., Зотов И. А. и др. Сравнение различных подходов к распределению ресурсов в центрах обработки данных // Изв. РАН. Теория и системы управления. 2014. № 5. С. 71—83.
15. Зотов И. А., Костенко В. А. Алгоритм распределения ресурсов в центрах обработки данных с единым планировщиком для различных типов ресурсов // Изв. Российской академии наук. Теория и системы управления. 2015. № 1. С. 61—71.
16. Парфёнов Д. И. Исследование распределения ресурсов в интерактивных сервисах инфокоммуникационных сетей. Дис. канд. техн. наук. Самара: Поволжская гос. акад. телекоммуникаций и информатики, 2014.
17. Игнатова Н. А. Методы и алгоритмы оптимизации ресурсного обеспечения сложных информационно-вычислительных систем на железнодорожном транспорте. Дис. ... канд. техн. наук. М.: Московский гос. ун-т путей сообщения МГУПС (МИИТ), 2014.
18. Забузов В. С., Казанцев Д. И., Белая Т. И., Швецов А. С. Способ организации контроля качества обслуживания в инфо-телекоммуникационной сети на примере ВКА им. А. Ф. Можайского // Научный обозреватель. 2014. № 12. С. 56—57.

Сведения об авторах

- Александр Геннадьевич Басыров** — д-р техн. наук, профессор; ВКА им. А. Ф. Можайского, кафедра информационно-вычислительных систем и сетей; начальник кафедры; E-mail: alxanderbas@mail.ru
- Денис Иванович Казанцев** — ВКА им. А. Ф. Можайского, проектно-конструкторское бюро отдела организации научных работ и подготовки научно-педагогических кадров; начальник бюро; E-mail: zilan.ka@yandex.ru

Рекомендована кафедрой
информационно-вычислительных систем
и сетей

Поступила в редакцию
18.04.16 г.

Ссылка для цитирования: Басыров А. Г., Казанцев Д. И. Методика распределения прикладных задач по виртуальным машинам с учетом надежности их решения // Изв. вузов. Приборостроение. 2016. Т. 59, № 10. С. 822—827.

**METHOD OF APPLIED PROBLEMS ALLOCATION TO VIRTUAL MACHINES
CONSIDERING THE RELIABILITY OF THEIR SOLUTION****A. G. Basyrov, D. I. Kazantsev**

A. F. Mozhaysky Military Space Academy, 197198, St. Petersburg, Russia
E-mail: zilan.ka@yandex.ru

The task of optimizing the use of limited resources of memory and the reliability of the computational processes is considered. An algorithm for applied problems allocation to virtual machines is developed. A technique of definition of rational structure of virtual machines is proposed.

Keywords: virtual machines, applied problems allocation, probability of failure-free performance

Data on authors

- Alexander G. Basyrov** — Dr. Sci., Professor; A. F. Mozhaysky Military Space Academy, Department of Information Systems and Networks; Chief of the Department; E-mail: alxanderbas@mail.ru
- Denis I. Kazantsev** — A. F. Mozhaysky Military Space Academy, Department of Organization of Science and Academic Staff Training; Chief of the Department; E-mail: zilan.ka@yandex.ru

For citation: Basyrov A. G., Kazantsev D. I. Method of applied problems allocation to virtual machines considering the reliability of their solution // Izv. vuzov. Priborostroenie. 2016. Vol. 59, N 10. P. 822—827 (in Russian).

DOI: 10.17586/0021-3454-2016-59-10-822-827