

АНАЛИЗ СТОХАСТИЧЕСКОЙ МОДЕЛИ РАСХОДОВ НА СОДЕРЖАНИЕ ГИБКОГО ВЫЧИСЛИТЕЛЬНОГО КЛАСТЕРА

А. Паньков

ГрГУ им. Я. Купалы

Гродно, Беларусь

a.pankov@gmail.com

В работе рассматривается стохастическая модель функционирования вычислительного кластера, построенного на базе технологий Amazon EC2 и Hadoop в виде замкнутой сети массового обслуживания (МО) с доходами (НМ-сети) в переходном режиме. Получена система разностно-дифференциальных уравнений (РДУ) для оценки затрат на содержание вычислительного кластера, описано ее решение.

Ключевые слова: НМ-сеть, вычислительный кластер.

1. ВВЕДЕНИЕ

Рассмотрим функционирование кластеров, созданных на основе системы EC2, предлагаемой компанией Amazon. Технология Amazon EC2 напоминает широко распространенную технологию по предоставлению VPS (Virtual Private Server) хостинга, главная концепция VPS позволяет запускать несколько компьютерных систем на одном физическом сервере. Amazon EC2, также как и VPS, позволяет запускать виртуальные сервера с различными операционными системами на физических серверах, однако Amazon берет деньги за использование серверов не ежемесячно, а за машино-часы и переданные данные между серверами (от \$0.10 до \$0.80 за машино-час, плюс \$0.10-\$0.20 за переданный гигабайт информации). При этом преимуществами Amazon EC2 являются такие, как легкость включения-выключения машин, создание образов новых машин, высокая скорость передачи данных между виртуальными машинами в локальной сети компании Amazon, относительная дешевизна использования виртуальных машин за чистое время — на время простоя машины можно выключить.

На основе технологии Amazon EC2 можно создать свою собственную вычислительную сеть, или кластер, состоящую из необходимого количества виртуальных машин (узлов кластера). Однако существует ограничение на максимальную конфигурацию кластера — на один аккаунт можно по умолчанию подключить до 20 виртуальных машин с максимальной конфигурацией: 8 ядер и 15 Гб памяти. На таком гибком, легко разворачиваемом и легко расширяемом кластере можно производить различные вычисления, требующие значительных затрат процессорного времени, а также машинной памяти.

На кластере, развернутом на мощностях Amazon EC2, можно производить вычисления с использованием системы Hadoop — открытой реализации технологии обработки данных MapReduce, разработанной компанией Google. Hadoop представляет собой платформу для построения распределенных приложений, способных обрабатывать огромные объемы данных. Система основывается на распределенном подходе к вычислениям и хранению информации.

Архитектуру приложения Hadoop можно представить следующим образом. Как правило, существует один выделенный главный сервер (ГС) или мастер (Master), который осуществляет координацию работы подчиненных ему серверов (ПС) — слэвов (Slave). На мастере запущено программное обеспечение для хранения и обработки информации о том, где и какие данные хранятся в распределенной файловой системе HDFS (Hadoop distributed file system), а также программное обеспечение для управления и распараллеливания работы вычислительных задач.

Работа приложения Hadoop осуществляется следующим образом. Оператор отправляет на вход ГС пакетное задание. Задание состоит из программы, осуществляющей обработку входных данных. Данные распределены между подчиненными серверами ПС. ГС, получив задание, отправляет копию программы на каждый второстепенный сервер. После этого ГС осуществляет планирование работы вычислительной задачи — разбивает входные данные на части, согласно установленному правилу (в зависимости от числа ПС, и т.д.), отправляет информацию о разбивке массива исходных данных на части каждому из ПС, а также информацию, с какой частью исходных данных работать каждому из ПС. После этого этапа происходит запуск программы обработки данных на каждом из ПС. Время обработки задания на ПС является случайным, экспоненциально распределенным с параметрами, зависящими от размера массива исходных данных, от числа ПС участвующих в обработке задания, от мощности оборудования, на котором запущены ПС. После окончания обработки задачи на ПС, ПС отправляет информацию об окончании вычислений ГС. Если на ГС еще есть части исходной задачи, ожидающие вычисления, ГС отправляет ПС следующий пакет заданий.

После окончания вычислений всех частей исходной задачи результаты вычислений могут отправляться ГС, где данные со всех ПС будут компоноваться в единый массив результата. Если на ГС еще имеются программы для произведения расчетов с только что полученными результатами, процедура повторяется. После окончания обработки всех программ, заданных на ГС, результат вычислений отправляется оператору.

2. СЕТЕВАЯ МОДЕЛЬ КЛАСТЕРА

Моделью функционирования такого вычислительного кластера может служить замкнутая сеть МО с центральной СМО, состоящая из n СМО, СМО S_i имеет m_i одинаковых линий обслуживания. ГС соответствует система S_n , i -му ПС — СМО S_i , $i = \overline{1, n-1}$, линии обслуживания соответствуют запущенным одновременно программам обработки данных на ПС, заявками являются пакетные задания для вычисления, а также результаты этих вычислений. Обслуживание заявок в каждой линии систе-

мы S_i является экспоненциальным с интенсивностью μ_i , $i = \overline{1, n}$. Заявки после обслуживания в системе S_n с вероятностью p_{ni} переходят на обслуживание в СМО S_i , $i = \overline{1, n-1}$, откуда, по завершению обслуживания, переходят в СМО S_n , $\sum_{i=1}^{n-1} p_{ni} = 1$.

Дисциплины обслуживания заявок в системах – FIFO, каждая из систем имеет неограниченное число мест для ожидания.

Пусть K – общее число заявок в сети. Тогда состояние сети в момент времени t можно описать вектором

$$k(t) = (k, t) = (k_1, k_2, \dots, k_n, t) = (k_1, k_2, \dots, k_{n-1}, K - k_1 - k_2 - \dots - k_{n-1}, t),$$

где k_i – число заявок в системе S_i , $i = \overline{1, n}$.

Будем считать, что при переходе заявки из периферийной СМО S_i в центральную заявка приносит СМО S_i расход в размере $R_{in}(k)$, $i = \overline{1, n-1}$, а СМО S_n – тот же доход. Если же происходит переход заявки из центральной СМО S_n в периферийную, то расход СМО S_n увеличивается на величину $R_{ni}(k)$, а доход последней увеличивается на эту же самую величину, $i = \overline{1, n-1}$. Кроме того, за каждый малый промежуток времени Δt СМО S_i увеличивает свой доход на величину $r_i(k)\Delta t + o(\Delta t)$, $i = \overline{1, n}$. Тогда полученная НМ-сеть может служить моделью изменения расходов на содержание вычислительного кластера, построенного на мощностях, предоставляемых компанией Amazon. Величина $r_i(k) > 0$ характеризует стоимость использования одной виртуальной машины в единицу времени. Величины $R_{in}(k)$ и $R_{ni}(k)$ равны расходам, выраженным в денежных величинах, которые несет владелец кластера за пересылку данных между виртуальными машинами, $R_{in}(k) > 0$, $R_{ni}(k) > 0$, $i = \overline{1, n-1}$. Примерами такого обмена данными являются пересылка пакетных заданий от одной виртуальной машины к другой и пересылка результатов вычислений. Принимая во внимание то, что обработка данных происходит по технологии MapReduce, данные в результате такой обработки, как правило, уменьшаются в объеме. Поэтому можно предположить, что объем пакетного задания, пересылаемого от ГС к ПС в несколько раз меньше, чем результат вычислений, пересылаемый от ПС к ГС. Так как стоимость пересылки данных пропорциональна объему пересылаемых данных, то можем сделать предположение о том, что $R_{ni}(k) > R_{in}(k)$, $i = \overline{1, n-1}$. Используя данную модель можно оценить денежные расходы на содержание отдельной виртуальной машины, а также расходы на содержание целого кластера, которые несет его владелец.

Следуя результатам, полученным в [1, 2], можно получить систему РДУ для ожидаемых доходов рассматриваемой НМ-сети:

$$\begin{aligned} \frac{dv(k, t)}{dt} = & \sum_{i=1}^{n-1} r_i(k) - \sum_{j=1}^n \mu_j u(k_j) v(k, t) + \\ & + \sum_{i=1}^{n-1} [\mu_n u(k_n) p_{ni} R_{ni}(k + I_i - I_n) - \mu_i u(k_i) R_{in}(k - I_i + I_n)] + \end{aligned}$$

$$+ \sum_{i=1}^{n-1} [\mu_i u(k_i) v(k - I_i + I_n, t) + \mu_n u(k_n) p_{ni} v(k + I_i - I_n, t)], \quad (1)$$

где $v(k, t)$ – полный ожидаемый доход, который получает сеть МО за время t , если в начальный момент времени сеть находится в состоянии k , μ_i – интенсивность обслуживания заявок в СМО S_i , $i = \overline{1, n}$, $u(x)$ – функция Хевисайда, I_i – вектор размерности n , состоящий из нулей, за исключением i -ой компоненты, которая равна 1. Для решения данной системы уравнений можно применить методы, разработанные в [3].

Пример

Рассмотрим вычислительный кластер, состоящий из 3 узлов, $n = 3$. Моделью изменения расходов на его содержание может служить замкнутая НМ-сеть с доходами. Пусть $K = 4$, $\mu_3 = 25$, $\mu_i = 5$, $i = \overline{1, 2}$. Состояниями такой сети будут $(0,0,4)$, $(0,1,3)$, $(0,2,2)$, $(0,3,1)$, $(0,4,0)$, $(1,0,3)$, $(1,1,2)$, $(1,2,1)$, $(1,3,0)$, $(2,0,2)$, $(2,1,1)$, $(2,2,0)$, $(3,0,1)$, $(3,1,0)$, $(4,0,0)$, переобозначим их соответственно $1, \dots, 15$.

В общем случае, вероятности переходов заявок между системами сети могут контролироваться пакетными заданиями по обработке данных. Однако мы рассмотрим случай, когда каждый процесс по обработке данных на ПС равновероятно получает одинаковую долю входных данных для обработки. В этом случае $p_{ni} = 1/2$, $p_{in} = 1$, $i = \overline{1, 2}$.

Известно, что стоимость использования самого простого типа виртуальных машин у компании Amazon составляет 0.1 у.е. в сутки. Предположим, что все 3 виртуальные машины в кластере являются одного простого типа, тогда $r_i = 0.1$, $i = \overline{1, 3}$.

Доходы от переходов между состояниями НМ-сети приведены в таблице 1.

Таблица 1 Доходы от переходов между состояниями сети

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$R_{ni}(k)$, у.е., $i = 1, n - 1$	4	3	3	4	4	3	4	5	3	4	3	2	4	2	4
$R_{in}(k)$, у.е., $i = 1, n - 1$	1.5	2	2	2	1.5	2	2	2	2	1	2	2	1	2	2

Для нахождения стоимости содержания рассматриваемого вычислительного кластера можно использовать систему уравнений (1), сведя ее к системе линейных неоднородных ОДУ и решив ее методом преобразований Лапласа [3]. На рис. 1–2 приведены графики изменения расходов на содержание кластера в зависимости от начального состояния НМ-сети.

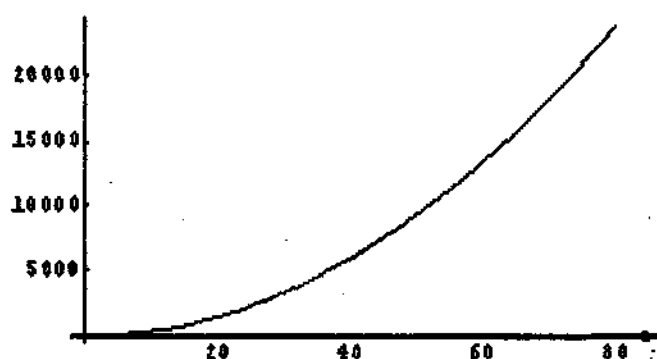


Рис. 1. Рост расходов на содержание кластера, если начальным состоянием НМ-сети является $k = 2$

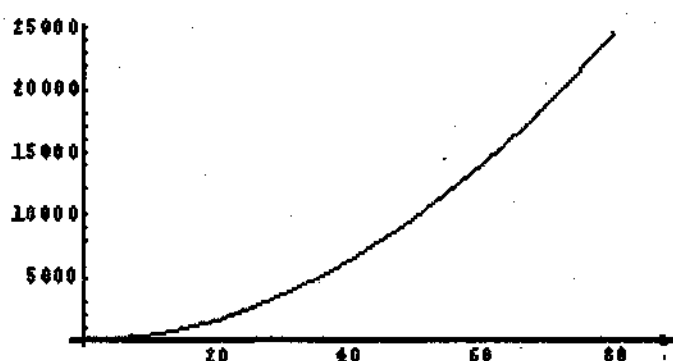


Рис. 2. Рост расходов на содержание кластера, если НМ-сеть начала функционирование в состоянии $k = 15$

ЛИТЕРАТУРА

1. Паньков А. Анализ сетей массового обслуживания с центральной системой в переходном режиме // Массовое обслуживание: потоки, системы, сети: материалы междунар. науч. конф. BWWQT-2007. г. Гродно. 29 янв. – 1 февр., 2007. Вып. 19. С. 201–206.
2. Паньков А. В. Анализ доходов в замкнутой марковской сети с центральной системой // Сборник научных статей международной конференции "Теория вероятностей, случайные процессы, математическая статистика и приложения". г. Минск. 21–25 февр., 2005 г. С. 215–222.
3. Matalytsky M., Pankov A. Analysis of one Markov queueing network with incomes // Queues: flows, systems, networks: Proceedings of the Intern. Conf. BWWQT-2005, Minsk, February 22–24, 2005. P. 124–128.