СТОХАСТИЧЕСКАЯ МОДЕЛЬ ФУНКЦИОНИРОВАНИЯ ГИБКОГО ВЫЧИСЛИТЕЛЬНОГО КЛАСТЕРА НА БАЗЕ ТЕХНОЛОГИЙ AMAZON EC2 И HADOOP

А. В. Паньков

Гродненский государственный университет имени Янки Купалы
Гродно, Беларусь
E-mail: a.pankov@gmail.com

В работе рассматривается стохастическая модель функционирования вычислительного кластера, построенного на базе технологий Amazon EC2 и Hadoop в виде сети массового обслуживания (МО) в переходном режиме. Получена система разностно-дифференциальных уравнений для вероятностей состояний сети, и приведен способ нахождения вероятностей для частного случая.

Ключевые слова: сеть массового обслуживания, вычислительный кластер.

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 помесячно, 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), слэйвов (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 Главному серверу (ГС, Master) пакетное задание. Задание состоит из программы, осуществляющей обработку входных данных. Данные распределены между подчиненными серверами (ПС, Slave). 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.), отсылает 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 полученным результатами, процедура повторяется. После окончания обработки всех программ, заданных на ГС, результат вычислений отправляется оператору.

Моделью работы такого вычислительного кластера может служить открытая сеть МО, состоящая из n СМО, каждая из которых имеет m_i одинаковых линий обслуживания, рис. 1. ГС соответствует система S_n, ПС – СМО $S_1,...,S_{n-1}$, линии обслуживания соответствуют запущенным одновременно программам обработки данных на ПС. Обслуживание заявок в каждой линии системы S_i является экспоненциальным с интенсивностью $\mu_i(k_i)$, зависящей от числа заявок в этой системе k_i, $i=\overline{1,n}$. В сеть поступает пуассоновский поток заявок с интенсивностью λ, в систему S_n, и после обслуживания в ней с вероятностью p_{ni} переходит в СМО S_i, $i=\overline{1,n-1}$, либо с вероятностью p_{n0} покидает сеть, $\sum_{i=1}^{n-1} p_{ni}+p_{n0}=1$. Дисциплины обслуживания заявок в системах – FIFO, каждая из систем имеет неограниченное число мест для ожидания.

Под состоянием сети в момент времени t будем понимать вектор $k(t)=(k,t)=(k_1,k_2,...,k_n,t)$, где k_i – число заявок в системе S_i, $i=\overline{1,n}$. Обозначим через I_i – вектор размерности n, состоящий из нулей, за исключением i-ой компоненты, которая равна 1. В силу экспоненциальности времен обслуживания заявок случайный процесс $k(t)=(k,t)$ является марковским со счетным числом состояний. Возможны следующие переходы в состояние (k,t) за время Δt:

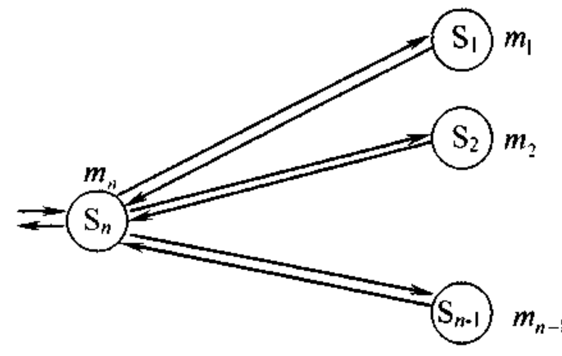

Рис. 1. Схема работы платформы Hadoop

а) из состояния $(k - I_n, t)$ с вероятностью $\lambda u(k_n)\Delta t + o(\Delta t)$;

б) из состояния $(k + I_n, t)$ с вероятностью $\mu_n(k_n+1)\min(m_n, k_n+1)p_{n0}\Delta t + o(\Delta t)$;

в) из состояния $(k + I_i - I_n, t)$ с вероятностью

$$\mu_i(k_i+1)u(k_n)\min(m_i, k_i+1)p_{in}\Delta t + o(\Delta t),\ i = \overline{1, n-1},\ \text{где}\ u(x) = \begin{cases} 1, & x > 0 \\ 0, & x \le 0 \end{cases};$$

г) из состояния $(k + I_n - I_i, t)$ с вероятностью

$$\mu_n(k_n+1)u(k_i)\min(m_n, k_n+1)p_{ni}\Delta t + o(\Delta t),\ i = \overline{1, n-1};$$

д) из состояния (k, t) с вероятностью

$$1 - \left[\sum_{i=1}^{n}\mu_i(k_i+1)\min(m_i, k_i+1) + \lambda\right]\Delta t + o(\Delta t);$$

е) из остальных состояний с вероятностью $o(\Delta t)$.

Система разностно-дифференциальных уравнений для вероятностей состояний сети имеет вид:

$$\begin{gathered}\frac{dP(k,t)}{dt} = -\left\{\sum_{i=1}^{n}\mu_i(k_i+1)\min(m_i, k_i+1) + \lambda\right\}P(k,t) + \\ + \mu_n(k_n+1)\min(m_n, k_n+1)p_{n0}P(k + I_n, t) + \lambda u(k_n)P(k - I_n, t) + \qquad (1) \\ + \sum_{i=1}^{n-1}\Big(\mu_i(k_i+1)u(k_n)p_{in}\min(m_i, k_i+1)P(k + I_i - I_n, t) + \mu_n(k_n+1)u(k_i)p_{ni}\min(m_n, k_n+1)P(k - I_i + I_n, t)\Big).\end{gathered}$$

ВЕРОЯТНОСТИ СОСТОЯНИЙ СЕТИ ДЛЯ СЛУЧАЯ $n = 3$

Пусть $n = 3$ и все СМО сети являются однолинейными, т.е. $m_i = 1$, $i = \overline{1, n}$. Тогда система уравнений (1) примет вид

$$\frac{dP(k,t)}{dt} = -P(k,t)\left(\lambda + \sum_{i=1}^{n}\mu_i(k_i+1)\right) + \lambda u(k_n)P(k - I_n, t) + \mu_n(k_n+1)p_{n0}P(k + I_n, t) +$$

$$+\sum_{i=1}^{n-1}\big(\mu_i(k_i+1)p_{in}u(k_n)P(k+I_i-I_n,t)+\mu_n(k_n+1)p_{ni}u(k_i)P(k-I_i+I_n,t)\big). \tag{2}$$

Введем трехмерную производящую функцию, т.е. для $z=(z_1,z_2,z_3)$ положим

$$\psi(z,t)=\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\sum_{r=0}^{\infty}P(k_1=l,k_2=m,k_3=r,t)z_1^l z_2^m z_3^r=\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\sum_{r=0}^{\infty}P(l,m,r,t)z_1^l z_2^m z_3^r\,. \tag{3}$$

Если умножить каждое из уравнений системы (2) на $z_1^l z_2^m z_3^r$ и просуммировать по всем возможным значениям l, m, r, то получим

$$\frac{\partial\psi(z,t)}{\partial t}=-\lambda\psi(z,t)-\sum_{i=1}^{3}\sum_{l=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\mu_i(k_i+1)P(k,t)z_1^l z_2^q z_3^r+$$

$$+\sum_{l=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\big(\mu_n(k_n+1)p_{n0}P(k+I_n,t)+\lambda u(k_n)P(k-I_n,t)\big)z_1^l z_2^q z_3^r+$$

$$+\sum_{l=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\sum_{i=1}^{n-1}\big(\mu_i(k_i+1)u(k_n)p_{in}P(k+I_i-I_n,t)+\mu_n(k_n+1)u(k_i)p_{ni}P(k-I_i+I_n,t)\big)z_1^l z_2^q z_3^r\,. \tag{4}$$

В дальнейшем будем предполагать, что интенсивности обслуживания $\mu_i(k_i)$ линейно зависят от числа заявок, т.е. $\mu_i(k_i)=\mu_i k_i$, $i=\overline{1,n}$. Естественно, что если $k_i=0$, то $\mu_i(k_i)=0$. Этот факт позволяет снять ограничение функционирования сети в условиях высокой нагрузки, при которой в работе [1] найдены нестационарные вероятности состояний сети, также используя метод производящих функций. Выразив каждую сумму системы (4) через производящую функцию $\psi(z,t)$ и её производные, получим уравнение для производящей функции:

$$\frac{\partial\psi(z,t)}{\partial t}=-\left(\lambda(1-z_3)+\sum_{i=1}^{n}\mu_i\right)\psi(z,t)+\mu_1(z_3-z_1)\frac{\partial\psi(z,t)}{\partial z_1}+\mu_2(z_3-z_2)\frac{\partial\psi(z,t)}{\partial z_2}+$$

$$+\mu_3\big(-z_3+p_{30}+p_{31}z_1+p_{32}z_2\big)\frac{\partial\psi(z,t)}{\partial z_3}\,. \tag{5}$$

Уравнение (5) является линейным неоднородным уравнением с частными производными первого порядка. Его можно свести к однородному уравнению, если искать его решение в неявном виде [2]

$$V(z_1,z_2,z_3,t,\psi)=0\,. \tag{6}$$

Считая, что в равенстве (6) ψ есть функция от z_1, z_2, z_3, t и продифференцировав это равенство по независимой переменной, получим $\frac{\partial V}{\partial z_i}+\frac{\partial V}{\partial\psi}\cdot\frac{\partial\psi}{\partial z_i}=0$, $i=\overline{1,3}$, $\frac{\partial V}{\partial t}+\frac{\partial V}{\partial\psi}\cdot\frac{\partial\psi}{\partial t}=0$, откуда следует, что $\frac{\partial\psi}{\partial z_i}=-\frac{\partial V}{\partial z_i}\Big/\frac{\partial V}{\partial\psi}$, $i=\overline{1,3}$, $\frac{\partial\psi}{\partial t}=-\frac{\partial V}{\partial t}\Big/\frac{\partial V}{\partial\psi}$. Подставляя эти значения частных производных в уравнение (5) и перенося все члены в левую часть, получим

$$\mu_1(z_1-z_3)\frac{\partial V}{\partial z_1}+\mu_2(z_2-z_3)\frac{\partial V}{\partial z_2}+\mu_3\big(-z_3+p_{30}+z_1p_{31}+z_2p_{32}\big)\frac{\partial V}{\partial z_3}+\frac{\partial V}{\partial t}-$$

$$-\left((1-z_3)\lambda+\sum_{i=1}^{3}\mu_i\right)\psi\frac{\partial V}{\partial\psi}=0\,. \tag{7}$$

Это уравнение является однородным линейным уравнением относительно V.

Уравнения характеристик для него имеют вид:

$$\frac{dz_1}{dx}=\mu_1(z_3-z_1),\ \frac{dz_2}{dx}=\mu_2(z_3-z_2),\ \frac{dz_3}{dx}=\mu_3\left(p_{30}-z_3+p_{31}z_1+p_{32}z_2\right),$$

$$\frac{dt}{dx}=-1,\ \frac{d\psi}{dx}=\left(\lambda(1-z_3)+\sum_{i=1}^{3}\mu_i\right)\psi. \tag{8}$$

Известно [3], что если найдены четыре первые независимые интегралы уравнений характеристик (8)

$$\varphi_i(z_1,z_2,z_3,t,\psi)=C_i,\ i=\overline{1,4}, \tag{9}$$

то все решения уравнения (9) определяются из равенства $\Phi(\varphi_1,\varphi_2,\varphi_3,\varphi_4)=0$, где Φ – произвольная дифференцируемая функция. Из уравнений (10) имеем:

$$t(x)=-x+C_4. \tag{10}$$

$$\begin{aligned}z_1(x)=&\left(C_1-\mu_1\mu_2p_{30}M_1(x)\right)R_1(x)+\left(C_2-\mu_2\mu_3p_{30}M_2(x)\right)p_{32}\mu_1\mu_3A_2(x)+\\&+\left(C_3-\mu_3p_{30}M_3(x)\right)\mu_1A_1(x)\end{aligned}, \tag{11}$$

$$\begin{aligned}z_2(x)=&\left(C_1-\mu_1\mu_2p_{30}M_1(x)\right)p_{31}\mu_2\mu_3A_2(x)+\left(C_2-\mu_2\mu_3p_{30}M_2(x)\right)R_2(x)+\\&+\left(C_3-\mu_3p_{30}M_3(x)\right)\mu_2A_3(x)\end{aligned}, \tag{12}$$

$$\begin{aligned}z_3(x)=&\left(C_1-\mu_1\mu_2p_{30}M_1(x)\right)p_{31}\mu_3A_1(x)+\left(C_2-\mu_2\mu_3p_{30}M_2(x)\right)p_{32}\mu_3A_3(x)+\\&+\left(C_3-\mu_3p_{30}M_3(x)\right)R_3(x)\end{aligned}, \tag{13}$$

$$\psi(x)=C_4e^{\int_1^x(\mu_1+\mu_2+\mu_3+\lambda(1-z_3(\tau)))d\tau}, \tag{14}$$

где $A_1(x)$, $A_2(x)$, $A_3(x)$, $M_1(x)$, $M_2(x)$, $M_3(x)$, $R_1(x)$, $R_2(x)$, $R_3(x)$ – некоторые функции вида $\sum_{i=1}^{3}\alpha_ie^{\beta_ix}$, $\alpha_i,\beta_i\in R$.

Согласно определению производящей функции (3), вероятность $P(k,t)$, где $k=(k_1,k_2,k_3)$, равна коэффициенту при $z_1^{k_1}z_2^{k_2}z_3^{k_3}$ в разложении функции $\psi(z_1,z_2,z_3,t)$ в ряд. Выразив правую часть выражения для $\psi(x)$ через z_1, z_2, z_3, t, используя соотношения (11)-(14), а затем разложив в ряд по степеням z_1, z_2 и z_3, получим вид для функции вероятностей состояний сети.

ЛИТЕРАТУРА

1. *Matalytski, M.* Analisys of the open queueing networks using method of generating functions / M. Matalytski, S. Gomanczuk, A. Pankov // Zbior artykulow Symposjuma Naukowego Instytutu Matematyki Pol. Czest. – Czestochowa: PCz, 2002.

2. *Матвеев Н. М.* Методы интегрирования обыкновенных дифференциальных уравнений / Н.М. Матвеев. М.: Высшая школа, 1967. 564 с.

3. *Самойленко А. М.* Дифференциальные уравнения. Примеры и задачи / А.М. Самойленко, С.А. Кривошея, Н.А. Перестюк. М.: Высшая школа, 1989. 383 с.