

К.С. ТКАЧЕНКО

**Обработка данных на
компьютерных узлах облачной
распределенной среды с учетом
возмущающих В-событий трафика**

УДК 004.7

ФГАОУ ВО
«Севастопольский
государственный
университет»,
г.Севастополь

Предлагается формирование подхода по определению необходимых режимов работы компьютерных узлов на основе аналитико-модельных оценок их основных характеристик. В основе подхода лежит применение одноканальных систем массового обслуживания. Проводится вычислительный эксперимент, приводятся необходимые таблицы, графики.

Введение

Очень многие современные практические задачи решаются с использованием распределенных сред (РС). К их числу можно отнести и требующие повышенного уровня надежности и безопасности, если которые не будут соблюдены, то приведет к большим потерям и жертвам. Требуются меры для соблюдения этих уровней. Поскольку в основе РС используются компьютерные узлы обработки данных, то можно выработать стратегию для отдельных узлов. Затем, когда, появится такая необходимость, повторить действия для каждого из отдельных узлов в РС.

Компьютерный узел сложен для непосредственного управления, может иметь различные модели. Для практического исследования достаточно часто используются аналитические и имитационные модели. Широко распространены аналитические модели на основе систем массового обслуживания (СМО). Если исключить либо компенсировать погрешности от генератора псевдослучайных чисел, то благодаря имитационному

моделированию может построить больше моделей с нетипичными потоками заявок, чем аналитическому.

Целью настоящей публикации является подготовка аналитического подхода к параметрической настройке компьютерных узлов с нестационарным входным трафиком.

Обзор существующих исследований

Объединение вычислительных мощностей с целью поддержки программных платформ приводит к облачным сервисам [1]. Эти централизованно управляемые сервисы включают в себя не только аппаратно-программное обеспечение, но и виртуальные машины. Это вызывает зависимость от облачных провайдеров в плане сохранности потребительских данных, а также от доступа в интернет. Автоматическая балансировка нагрузки и высокая скорость обработки запросов возможны благодаря наличию специализированных брокеров. Для прогнозирования и принятия решений применяется анализ исторических данных.

В учебном процессе облачные вычисления позволяют доступ к вычислительным ресурсам организовать в режиме сотрудничества, доступности, экономности [2]. Для обмена и актуализации знаниями, их преобразовании в новые систематизированные комбинации, облачные платформы включают в себя инструментальные средства для модератора. Интегрированные среды программирования в облаке позволяют организовать различные формы учебной деятельности.

При открытом развертывании облачных вычислений, их средства организации могут подвергаться различным вредоносным угрозам [3]. Одной из основных угроз считается злоупотребление доверием. Эта угроза проистекает из того, что значительная совокупность функционала предоставляется провайдерами облачных услуг и в полной мере не контролируется потребителями. Кроме этого, сетевые и веб-службы подвержены агрессивному узловому трафику.

Для задач обработки данных в системах мониторинга окружающей среды существует подход по их динамической реструктуризации [4]. Этот подход в условиях нестационарного

входного трафика позволяет выполнять построение эффективных систем природного мониторинга.

Понятно, что существуют нерешенные ограничивающие проблемы, связанные с организацией функционирования процессов обработки информации в условиях, с одной стороны, подверженного внешним воздействиям трафика, с другой стороны, необходимостью выполнять в таких условиях управление компьютерным узлом. Эти ограничения невозможно снять исключительно аналитическим подходом, поскольку в результате для нахождения выходных характеристик систем придется прибегнуть к нахождению решений систем интегро-дифференциальных уравнений.

Можно с сильными допущениями такую задачу решать с применением мажорирующих систем [5]. При таком подходе общая K -канальная система с ограниченной N -очередью и произвольным законом потока заявок, произвольным законом обработки заявок $G/G1/K/N$ может быть представлена системой $M/M/K/N$. Имитационным моделированием доказываается и демонстрируется, что G/G -системы могут быть заменены M/M -системами с такой же конфигурацией и структурами, то есть K и N .

Для такого представления придется отказаться от предположения общности функций распределений времени поступления заявок и их обработки, а также наложить ограничение ординарности на входной поток.

Аналитическое моделирование

СМО типа $M/M/K/N$ поглощают СМО типа $M/M/1/N$. Это удобно использовать в случаях, когда $K=1$. Например, если компьютерный узел рассматривать исключительно как единый сервис по предоставлению однотипных услуг, а именно, веб-сервер, файловый сервер и подобные. Аналитическое описание $M/M/1/N$ хорошо тем, что для расчета важнейших системных характеристик требуется меньшее число арифметических операций, чем для $M/M/K/N$. В результате, для $M/M/1/N$ программная реализация проще в вычислительном плане и гораздо более производительнее.

Расчет производится по известным аналитическим зависимостям:

$$\rho = \frac{\lambda}{\mu},$$

$$p_0 = \frac{1-\rho}{1-\rho^{N+2}}, \quad (1)$$

$$p_{отк} = p_{N+1} = p_0 \rho^{N+1}.$$

В этой формуле (1) ρ – загрузка, p_0 – вероятность простоя, $p_{отк}$ – вероятность отказа. В силу того, что $K=1$, должно быть $0 < \rho < 1$. На основе (1) для кортежей $\langle \lambda, \mu \rangle_i$ могут быть сформированы кортежи $\langle p_0, p_{отк} \rangle_i$ (для некоторых значений i).

В-события во входном трафике вызывают возмущения, приводящие к изменению интенсивности входного потока λ , что при неизменной производительности μ отражается как понижение производительности и повышение загрузки ρ . Каждое такое изменение может привести либо к потере заявок, либо к простоям. Как за потерю заявок, так и за простоями возникает критическая ситуация, затраты на ликвидацию которой можно оценить по штрафной функции, зависящей от p_0 и $p_{отк}$. Поэтому необходимо определить для заранее заданной величины $\lambda^* = const$ интервал значений производительности обработки $\mu \in [\mu^{min}; \mu^{max}]$, для которого существуют такие пары значений $\langle p_0^*, p_{отк}^* \rangle$:

$$\begin{cases} p_0^* \rightarrow min, \\ p_{отк}^* \rightarrow min. \end{cases} \quad (2)$$

Задача (2) является задачей нелинейного программирования, которую можно решить методами многокритериальной оптимизации. Чтобы минимизировать участие ЛПР (лица, принимающего решения) в решении (2), целесообразно применять метод множеств, эффективных по Парето. То есть, необходимо произвести для заранее известных значений $\lambda^* = const = \lambda^d$ и $[\mu^{min} = \mu^{d1}; \mu^{max} = \mu^{d2}]$:

1. Построение

$$Q^* = \{ \langle p_0, p_{отк} \rangle | p_0, p_{отк} \text{ по (1) для } \lambda^* = \lambda^d, \mu \in [\mu^{d1}; \mu^{d2}] \};$$

2. Выбор эффективной границы $Q^{**} \subseteq Q^*$ по (2).

Например, полагается, что $\lambda^d = 130$, $\mu^{d1} = 153$, $\mu^{d2} = 173$ и тип СМО М/М/1/22, то есть $\rho \in [0,75; 0,85]$. Выполняется расчет по (1), результаты заносятся в табл. 1.

Результаты вычислительного эксперимента

№	ρ	ρ_0	$\rho_{\text{отк}}$
1	0,750	0,250251	0,000335
2	0,753	0,247706	0,000358
3	0,755	0,245161	0,000384
4	0,758	0,242619	0,000410
5	0,760	0,240077	0,000439
6	0,763	0,237537	0,000469
7	0,765	0,234999	0,000502
8	0,768	0,232463	0,000536
9	0,771	0,229928	0,000572
10	0,773	0,227395	0,000611
11	0,776	0,224865	0,000652
12	0,778	0,222336	0,000695
13	0,781	0,219810	0,000741
14	0,783	0,217286	0,000790
15	0,786	0,214764	0,000842
16	0,788	0,212246	0,000897
17	0,791	0,209730	0,000955
18	0,794	0,207217	0,001017
19	0,796	0,204707	0,001082
20	0,799	0,202201	0,001150
21	0,801	0,199698	0,001223
22	0,804	0,197199	0,001300
23	0,806	0,194703	0,001381
24	0,809	0,192212	0,001466
25	0,812	0,189725	0,001557
26	0,814	0,187242	0,001652
27	0,817	0,184765	0,001752
28	0,819	0,182292	0,001858
29	0,822	0,179824	0,001970
30	0,824	0,177362	0,002087
31	0,827	0,174905	0,002211
32	0,829	0,172454	0,002340
33	0,832	0,170010	0,002477
34	0,835	0,167572	0,002621
35	0,837	0,165141	0,002771
36	0,840	0,162717	0,002930
37	0,842	0,160300	0,003096
38	0,845	0,157891	0,003270
39	0,847	0,155490	0,003453
40	0,850	0,153098	0,003644

На основании табл. 1 выполняется построение границы множества Парето-решений в соответствии с (2), результат изображается на рис. 1.

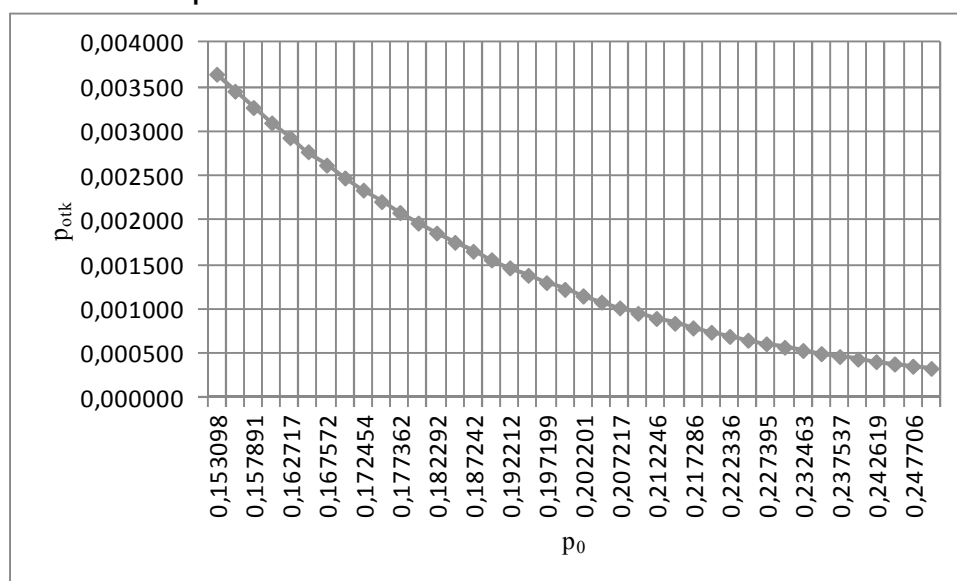


Рис. 1. Граница облака Парето-решений

Из анализа табл. 1 и рис. 1 видно, что для случая M/M/1/22 с $\lambda^* = 130$ и $[\mu^{min} = 153; \mu^{max} = 173]$, должно быть $\mu^{eff} \in [157; 163]$.

Заключение

В результате работы описан подход, который позволяет определять производительность обработки заявок компьютерными узлами на основе моделей одноканальных систем. С помощью проведенного вычислительного эксперимента получен численный пример нахождения эффективного решения.

Литература

1. Коваленко О.С. Обзор проблем и состояний облачных вычислений и сервисов / О.С. Коваленко, В.М. Курейчик // Известия Южного федерального университета. Технические науки, т. 132, №. 7, 2012. С. 146–153.
2. Голицына И.Н. Использование облачных вычислений в образовательном процессе / И.Н. Голицына, А.Н. Афзалова // Образовательные технологии и общество, т. 17, № 2, 2014. С. 450–459.
3. Ковалев Д. Информационная безопасность облачных вычислений / Д. Ковалев // Т-Сотт – Телекоммуникации и Транспорт, №. S1, 2011. С. 14–16.
4. Ткаченко К.С. Поточно-структурный подход к построению распределенных сред систем мониторинга / К.С. Ткаченко, И.А. Скاتков // Системы контроля окружающей среды. Севастополь: ИПТС. 2017. Вып. 9 (29). С.41–44.

5. *Скатков А.В.* Обеспечение гарантоспособности систем мониторинга окружающей среды на основе параметрического мажорирования / А.В. Скатков, И.А. Балакирева // Экологическая, промышленная и энергетическая безопасность – 2017. Сборник статей по материалам научно-практической конференции с международным участием. Под редакцией Ю.А. Омельчук, Н.В. Ляминой, Г.В. Кучерик. 2017. С. 1233–1236.

ТКАЧЕНКО К.С.

KSTKACHENKO@SEVSU.RU