

О МОДЕЛИ ПОТРЕБЛЕНИЯ ЭЛЕКТРОЭНЕРГИИ КОМПАНИЕЙ, ПРЕДОСТАВЛЯЮЩЕЙ ОБЛАЧНЫЕ РЕСУРСЫ

Логачёв А. В.

Институт математики им. С. Л. Соболева СО РАН, Сибирский государственный университет геосистем и технологий
E-mail: avlogachov@mail.ru

Логачёва О. М.

Сибирский государственный университет геосистем и технологий
E-mail: omboldovskaya@mail.ru

Современные мобильные устройства широко используют облачные технологии. Облачная среда включает в себя один или несколько центров обработки данных. С ростом количества мобильных устройств и их программного обеспечения увеличивается количество серверов и другого оборудования в центрах обработки облачных данных. Это увеличение приводит к росту потребляемой электроэнергии и соответствующих затрат. Вопросы, связанные с энергетическим менеджментом облачных ресурсов, являются важными как с точки зрения защиты окружающей среды, так и с точки зрения максимизации прибыли от предоставления удаленных вычислительных ресурсов.

В работе предложена стохастическая модель, позволяющая оценить количество потребляемой электроэнергии компаниями, предоставляющими услуги, связанные с удаленной обработкой цифровых данных. Данная модель также может быть использована для моделирования поведения таких компаний.

Ключевые слова: облачные технологии, энергетический менеджмент, стохастическое моделирование, динамические модели, прогнозирование

DOI: 10.32324/2412-8945-2021-1-40-44

ABOUT THE MODEL OF ELECTRICITY CONSUMPTION BY A COMPANY THAT PROVIDES CLOUD RESOURCES

Logachov A. V.

Sobolev Institute of Mathematics SB RAS, Siberian State University of Geosystems and Technologies
E-mail: avlogachov@mail.ru

Logachova O. M.

Siberian State University of Geosystems and Technologies
E-mail: omboldovskaya@mail.ru

Modern mobile devices widely use cloud technologies. A cloud environment includes one or more data centers. As the number of mobile devices and their software increases, so does the number of servers and other hardware in cloud data centers. This increase leads to an increase in electricity consumption and related costs. Issues related to the energy management of cloud resources are important both from the point of view of protecting the environment and from the point of view of maximizing the profit from the provision of remote computing resources.

In this paper, we propose a stochastic model that allows us to estimate the amount of electricity consumed by companies providing services related to remote digital data processing. This model can also be used to simulate the behavior of such companies.

Keywords: cloud technologies, energy management, stochastic modeling, dynamic models, forecasting

DOI: 10.32324/2412-8945-2021-1-40-44

В последнее время в информационных технологиях (ИТ) большое распространение получили облачные технологии [2]. Облачная среда включает в себя один или несколько центров обработки данных, которые могут быть расположены в разных странах. Каждый центр обработки данных содержит множество физических серверов, на которых размещены виртуальные машины, отвечающие на запросы заказчиков. В частности, все больше и больше приложений исполь-

зуют комбинацию мобильных и облачных вычислительных технологий, чтобы полностью раскрыть свой потенциал. Облако обычно используется для предоставления дополнительных вычислительных возможностей, которые не могут быть эффективно реализованы в случае, когда используются только мобильные устройства. С ростом количества приложений, использующих облачные ресурсы, вычислительные центры, к которым эти приложения обращаются, становятся больше по масштабу и потребляют больше электроэнергии [10]. Потребление огромного количества электроэнергии приводит к большим выбросам

углекислого газа в атмосферу. Так, в работе [3] сообщалось, что в 2017 г. центры обработки данных потребляли около 3 % всей электроэнергии в мире и вырабатывали около 200 млн т углекислого газа в год. В то же время компании, предлагающие услуги, связанные с облачными вычислениями и обработкой данных, всегда стараются по максимуму использовать имеющееся аппаратное обеспечение и получить наибольшую прибыль от предоставления своих услуг. Поэтому вопросы, связанные с энергетическим менеджментом облачных ресурсов, являются важными как с точки зрения защиты окружающей среды, так и с точки зрения максимизации прибыли от предоставления удаленных вычислительных ресурсов.

В последнее время выходит большое количество работ, связанных с эффективностью потребления электроэнергии центрами обработки данных. Условно эти работы можно разделить на два класса: первый связан с вычислением потребляемой электроэнергии в зависимости от того, какой алгоритм распределения облачных ресурсов используется; второй связан с разработкой динамических моделей-алгоритмов, которые позволяют предсказать, какой объем облачных ресурсов понадобится в следующий момент времени. Мы ставим в этой работе следующую цель: на основе современных результатов по этим двум классам построить модель, которая позволит оценить количество электроэнергии, потребляемой центром обработки данных за фиксированный промежуток времени. По-видимому, в дальнейшем на основе этой модели могут быть разработаны алгоритмы, оптимизирующие работу компании, предоставляющей облачную среду.

Последующая часть работы структурирована в три раздела. В первом разделе мы рассмотрим результаты, связанные с детерминированным вычислением количества потребляемой электроэнергии. Во втором разделе будет рассмотрена модель, позволяющая прогнозировать загрузку серверов в зависимости от количества клиентов, обращающихся в центр обработки данных (здесь количество клиентов и объем задач, которые они собираются решать — случайные величины). В третьем разделе мы на основе моделей, рассмотренных в предыдущих двух разделах, получим модель, которая позволяет оценить количество потребляемой электроэнергии с учетом случайной составляющей.

Модель для вычисления количества потребляемой электроэнергии. В работах [5—7, 8] потребляемая мощность, связанная с вычислениями, определяется как функция от доли загрузки центрального процессора (ЦП) на k -м сервере, $1 \leq k \leq n$, где n — общее количество серверов в центре обработке данных, следующим образом:

$$P_{comp}^k(u) = 0,7P_{max}^k + 0,3P_{max}^k,$$

где P_{comp}^k — потребляемая k -м сервером мощность в случае, если процессор загружен на 100 %, $u \in [0, 1]$ —

доля загрузки ЦП (т. е. $u = 1$, если процессор загружен на 100 %).

Доля загрузки процессора зависит от времени $u = u(t)$, поэтому справедлива приведенная ниже формула, позволяющая вычислить объем потребляемой сервером электроэнергии, необходимой для вычислений на k -м сервере:

$$E_{comp}^k(t) = \int_0^t P_{comp}^k(u(s))ds, \quad (1)$$

где 0 — начальный момент времени, t — текущий момент времени.

Для вычисления количества потребляемой электроэнергии в процессе *динамической миграции* виртуальных машин (под динамической миграцией понимается перенос вычислений, связанных с решаемой в данный момент времени задачей, на другой сервер) в работе [6] предлагают использовать следующую формулу:

$$E_{migr}(t) = 4P_{migr} \sum_{j=1}^{m(t)} \frac{M_j}{B_j}, \quad (2)$$

где M_j — объем памяти, занимаемой j -й виртуальной машиной в момент миграции, B_j — пропускная способность в момент миграции j -й виртуальной машины, P_{migr} — мощность, потребляемая при перемещении единицы данных, $m(t)$ — количество миграций на промежутке времени $[0, t]$.

Отметим, что, согласно результатам работы [4], потребление электроэнергии в процессе динамической миграции в среднем составляет 10 % от общей потребляемой электроэнергии.

Для вычисления количества потребляемой электроэнергии, связанной с *переходом сервера из спящего режима в активный*, в работе [8] предлагается формула

$$E_{swit}(t) = \sum_{r=1}^{l(t)} \frac{P_{swit}^r T_{swit}^r}{2}, \quad (3)$$

где P_{swit}^r — средняя мощность, потребляемая r -м сервером в процессе перехода из спящего режима в активный, T_{swit}^r — время, необходимое для того, чтобы r -й сервер перешел из спящего режима в активный, $l(t)$ — количество серверов, которые переходили из спящего режима в активный на промежутке времени $[0, t]$.

Теперь, используя формулы (1) — (3), мы можем найти общее количество электроэнергии, потребляемой центром обработки данных:

$$\begin{aligned} E_{total}(t) &= \sum_{k=1}^n E_{comp}^k(t) + E_{migr}(t) + E_{swit}(t) = \\ &= 0,7t \sum_{k=1}^n P_{max}^k + 0,3 \sum_{k=1}^n P_{max}^k \int_0^t u(s)ds + \\ &+ 4P_{migr} \sum_{j=1}^{m(t)} \frac{M_j}{B_j} + \sum_{r=1}^{l(t)} \frac{P_{swit}^r T_{swit}^r}{2}. \end{aligned} \quad (4)$$

Модель для оценки загрузки серверных ЦП и других параметров центра обработки данных. Будем предполагать, что клиенты поступают на центр обработки данных, согласно процессу рождения-гибели (см., например, [1]), переходная вероятность которого при $\Delta \rightarrow 0$ и $x > 0$ имеет следующий вид:

$$P(X(t+\Delta) = y | X(t) = x) = \begin{cases} \frac{x}{\mu} \Delta + 0(\Delta), & \text{если } y = x-1, \\ 1 - \frac{x}{\mu} \Delta - \gamma(\Delta) + 0(\Delta), & \text{если } y = x, \\ \gamma \Delta + 0(\Delta), & \text{если } y = x+1, \end{cases} \quad (5)$$

а при $\Delta \rightarrow 0$ и $x > 0$ и задается формулой

$$P(X(t+\Delta) = y | X(t) = 0) = \begin{cases} \frac{x}{\mu} \Delta + 0(\Delta), & \text{если } y = x-1, \\ 1 - \frac{x}{\mu} \Delta - \gamma(\Delta) + 0(\Delta), & \text{если } y = x, \\ \gamma \Delta + 0(\Delta), & \text{если } y = x+1, \end{cases} \quad (6)$$

где $X(t)$ — количество клиентов в момент времени t , $X(t+\Delta)$ — количество клиентов в момент времени $t+\Delta$, μ — среднее время обслуживания клиента, γ — интенсивность поступления клиентов. Отметим, что эта модель предложена в работе [9].

Обозначим a_z — «количество» ЦП, которые будут загружены задачами, поставленными z -м клиентом (здесь количество может быть не целым, также естественно, что на одном сервере может обслуживаться одновременно несколько клиентов). Очевидно, что мы заранее не знаем величину a_z , следовательно a_z — случайная величина. Будем предполагать, что задачи, которые собираются решать разные клиенты независимы в совокупности, а следовательно, независимыми являются и случайные величины $a_1, a_2, \dots, a_{X(t)}$.

В момент времени t все клиенты в совокупности загружат

$$A(t) = \sum_{z=1}^{X(t)} a_z \quad (7)$$

процессоров.

Для простоты будем предполагать, что все серверы центра обработки данных примерно одинаковы. В этом случае в момент времени t будут:

а) полностью загружены $[A(t)]$ — ЦП серверов (здесь $[A(t)]$ — целая часть числа $A(t)$);

б) и процессор одного сервера на $(A(t) - [A(t)])$ 100 %.

Наряду со случайным процессом $X(t)$ рассмотрим процесс $Y(t)$, который подсчитывает общее количество клиентов, обратившихся в центр обработки данных на промежутке времени $[0, t]$. Из

формул (5), (6) следует, что $Y(t)$ — процесс Пуассона с параметром $EY(t) = \gamma t$ (здесь $EY(t)$ — математическое ожидание случайной величины $Y(t)$).

Будем предполагать, что в момент обращения нового клиента его задачи будут размещены с вероятностями $p \in (0,1)$ и $q = 1 - p$ на серверах, находящихся в спящем режиме и активном режиме соответственно. Тогда количество серверов, которые переходили из спящего режима в активный на промежутке времени $[0, t]$, может быть вычислено по формуле

$$l(t) = \sum_{v=1}^{Y(t)} d_v. \quad (8)$$

Здесь случайные величины $d_1, \dots, d_{Y(t)}$ независимы и имеют распределение Бернулли с параметром p (т. е. если задача v -го клиента размещена на сервере, находящемся в спящем режиме, то $d_v = 1$, если на активном, то $d_v = 0$).

Обозначим m_v — количество динамических миграций, связанных с решением задач v -го клиента. Будем предполагать, что случайные величины $m_1, \dots, m_{Y(t)}$ независимы и имеют одинаковое распределение. Тогда количество миграций на промежутке времени $[0, t]$ может быть вычислено по формуле

$$m(t) = \sum_{v=1}^{Y(t)} dm_v. \quad (9)$$

Таким образом, нами введены все случайные элементы, связанные с оценкой параметров центра обработки данных.

Модель для прогнозирования количества потребляемой электроэнергии. Составим на основе моделей, рассмотренных в предыдущих двух разделах, модель, позволяющую оценить количество потребляемой электроэнергии с учетом случайного потока клиентов и случайного объема задач, которые они собираются решать.

Используя формулы (4)–(9), а) и б), а также предположение об однотипности серверов, входящих в центр обработки данных, будем иметь

$$\begin{aligned} E_{total}(t) &= \sum_{k=1}^n E_{comp}^k(t) + E_{migr}(t) + E_{swit}(t) = \\ &= P_{max} (0,3n + \int_0^t A(s) ds) + \\ &+ 4P_{migr} \sum_{j=1}^{m(t)} \frac{M_j}{B_j} + \frac{P_{swit} T_{swit}}{2} l(t). \end{aligned} \quad (10)$$

Здесь, в соответствии с формулами (8), (9), $l(t)$, $m(t)$ — случайные процессы:

$$l(t) = \sum_{v=1}^{Y(t)} d_v, \quad m(t) = \sum_{v=1}^{Y(t)} dm_v, \quad (11)$$

и в силу однотипности серверов

$$\begin{aligned} P_{max}^1 &= \dots = P_{max}^n = P_{max}, \\ P_{swit}^1 &= \dots = P_{swit}^l = P_{swit}, \\ T_{swit}^1 &= \dots = T_{swit}^l = T_{swit}, \end{aligned}$$

Легко увидеть, что если есть соответствующая выборка, то неизвестные параметры случайных элементов, входящих в модель (10), (11), могут быть оценены стандартными методами статистики случайных процессов. После оценки этих параметров может быть легко получена оценка среднего количества электроэнергии, потребляемой центром обработки данных. Отметим также, что модель (10), (11) может быть использована для симуляции работы компании, предоставляющей облачные ресурсы.

Таким образом, исходя из роста спроса на внедрение облачных технологий, актуальной яв-

ляется задача максимально эффективного энергопотребления центрами обработки данных. В настоящей работе предложена модель оценки количества электроэнергии, потребляемой центром обработки данных за фиксированный промежуток времени, на основе известных моделей вычисления количества потребляемой электроэнергии в зависимости от того, какой алгоритм распределения облачных ресурсов используется, а также динамических моделей-алгоритмов, которые позволяют предсказать, какой объем облачных ресурсов понадобится в следующий момент времени.

Список литературы

1. Гнеденко Б. В., Беляев Ю. К., Соловьев А. Д. Математические методы в теории надежности. М., 1965.
2. A view of cloud computing / M. Armbrust, A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica et al. // Communications of the ACM. 2010. Vol. 53, N. 4. P. 50—58.
3. Asad Z., Chaudhry M. A. R. A two-way street: Green big data processing for a greener smart grid // IEEE Systems Journal. 2017. N 11(2). P. 784—795.
4. Beloglazov A., Buyya R. Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers // Concurr Comput Pract Exp. 2012. N 24(13). P. 1397—1420.
5. Fan X., Weber W-D., Barroso L. A. Power Provisioning for a Warehouse-Sized computer // Proceedings of the 34th annual international symposium on computer architecture, ACM. 2007. P. 13—23.
6. Huang J., Wu K., Moh M. Dynamic Virtual Machine Migration Algorithms Using Enhanced Energy Consumption Model for Green Cloud Data Centers // International Conference on High Performance Computing & Simulation (HPCS), 2014. P. 902—910 (2020).
7. Power and performance management of virtualized computing environments via lookahead control / D. Kusic, J. O. Kephart, J. E. Hanson, N. Kandasamy, G. Jiang // Clust Comput. 2009. N 12(1). P. 1—15.
8. Saadi Y., El Kafhali S. Energy-efficient strategy for virtual machine consolidation in cloud environment // Soft Comput. 2020. N 24(2). P. 14845—14859.
9. Shi Y., Jiang X., Ye K. An Energy-Efficient Scheme for Cloud Resource Provisioning Based on CloudSim // IEEE International Conference on Cluster Computing, Austin, TX, USA, 2011. P. 595—599.
10. Virtual machine based energy-efficient data center architecture for cloud computing: A performance perspective / K. Ye, D. Huang, X. Jiang, H. Chen and S. Wu // Proceedings of the 2010 IEEE/ACM International Conference on Green Computing and Communications. 2010. P. 171—178.

Bibliography

1. Gnedenko B.V., Beljaev Ju.K., Solov'ev A.D. Matematicheskie metody v teorii nadezhnostij. Moscow, 1965.
2. A view of cloud computing / M. Armbrust, A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica et al. Communications of the ACM. 2010. Vol. 53, N 4, pp. 50—58.
3. Asad Z., Chaudhry M.A.R. A two-way street: Green big data processing for a greener smart grid. IEEE Systems Journal. 2017. N 11(2), pp. 784—795.
4. Beloglazov A., Buyya R. Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers. Concurr Comput Pract Exp. 2012. N 24(13), pp. 1397—1420.
5. Fan X., Weber W-D., Barroso L.A. Power Provisioning for a Warehouse-Sized computer. Proceedings of the 34th annual international symposium on computer architecture, ACM. 2007, pp. 13—23.
6. Huang J., Wu K., Moh M. Dynamic Virtual Machine Migration Algorithms Using Enhanced Energy Consumption Model for Green Cloud Data Centers. International Conference on High Performance Computing & Simulation (HPCS), 2014, pp. 902—910 (2020).
7. Power and performance management of virtualized computing environments via lookahead control / D. Kusic, J.O. Kephart, J.E. Hanson, N. Kandasamy, G. Jiang. Clust Comput. 2009. N 12(1), pp. 1—15.
8. Virtual machine based energy-efficient data center architecture for cloud computing: A performance perspective / K. Ye, D. Huang, X. Jiang, H. Chen, S. Wu. Proceedings of the 2010 IEEE/ACM International Conference on Green Computing and Communications. 2010, pp. 171—178.
9. Saadi Y., El Kafhali S. Energy-efficient strategy for virtual machine consolidation in cloud environment. Soft Comput. 2020. N 24(2), pp. 14845—14859.
10. Shi Y., Jiang X., Ye K. An Energy-Efficient Scheme for Cloud Resource Provisioning Based on CloudSim. IEEE International Conference on Cluster Computing, Austin, TX, USA, 2011, pp. 595—599.

ЛОГАЧЁВ Артём Васильевич — кандидат физико-математических наук, доцент, старший научный сотрудник, Институт математики им. С. Л. Соболева СО РАН, доцент, Сибирский государственный университет геосистем и технологий, Новосибирск, Российская Федерация. E-mail: avlogachov@mail.ru

Artem V. LOGACHOV — Candidate of Physical and Mathematical Sciences, Associate Professor, Senior Researcher, Sobolev Institute of Mathematics SB RAS Associate Professor, Siberian State University of Geosystems and Technologies, Novosibirsk, Russian Federation. E-mail: avlogachov@mail.ru

ЛОГАЧЁВА Ольга Михайловна — кандидат физико-математических наук, доцент, доцент, Сибирский государственный университет геосистем и технологий, Новосибирск, Российская Федерация. E-mail: omboldovskaya@mail.ru

Olga M. LOGACHOVA — Candidate of Physical and Mathematical Sciences, Associate Professor, Siberian State University of Geosystems and Technologies, Novosibirsk, Russian Federation. E-mail: omboldovskaya@mail.ru