

ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ ОДНОЙ ЭКСПОНЕНЦИАЛЬНОЙ ОКРЕСТНОСТИ ДЛЯ ЗАДАЧИ БАЛАНСИРОВКИ НАГРУЗКИ НА СЕРВЕРЫ

И.А. Давыдов

*Новосибирский государственный университет,
Институт математики им. С.Л. Соболева СО РАН, г. Новосибирск
e-mail: vann.davydov@gmail.com*

В докладе рассматривается задача балансировки нагрузки на серверы, возникающая при организации облачного хостинга веб приложений. Имеется набор серверов, на каждом из которых размещены диски (точнее, их образы). На дисках хранятся интернет сайты с разнородной информацией. Пользователи посещают сайты, создавая определенную нагрузку на сервер, на диске которого размещен сайт. Считается, что нагрузка на сайт меняется со временем и характеризуется несколькими параметрами, такими как загрузка центрального процессора сервера, выделение оперативной памяти, использование устройств ввода-вывода и др. Таким образом, по каждому диску известна активность пользователей в течение планового периода. Эта активность позволяет определить нагрузку на каждый сервер в каждый момент времени по каждому параметру. Если нагрузка по каждому параметру не превосходит заданного порога, то сервер находится в рабочем режиме. В противном случае сервер работает с перегрузкой. Чтобы избежать перегрузки, диски можно перемещать с одного сервера на другой. Такое перемещение требует определенных вычислительных затрат. Будем называть их накладными расходами. Предполагается, что для каждого диска известны накладные расходы по каждому параметру при его изъятии с сервера и подключении к любому другому серверу. Начальное распределение дисков по серверам считается известным. Задача состоит в том, чтобы перераспределить диски по серверам и достичь минимальной суммарной перегрузки на всем плановом периоде при ограничениях на накладные расходы по каждому серверу.

Для решения задачи балансировки нагрузки на серверы предложена новая окрестность экспоненциальной мощности. Элементами этой окрестности являются оптимальные решения задачи о назначениях, которая получается при изъятии одного диска с каждого сервера и перераспределения дисков между серверами оптимальным образом. Рассматриваются различные варианты выбора дисков для извлечения. Исследуются алгоритмы локального поиска с рандомизированными версиями такой окрестности. Обсуждаются результаты численных экспериментов. Приводится сравнение с результатами, полученными в работе [1].

ЛИТЕРАТУРА

1. Кочетов Ю. А., Кочетова Н. А. *Задача балансировки нагрузки на серверы.* — Вестник Новосибирского государственного университета. — 2013. т. 11, вып. 4, с. 71-76.