

# МЕТОДЫ ОЦЕНКИ ЗАГРУЗКИ ВЫЧИСЛИТЕЛЬНЫХ УЗЛОВ РАСПРЕДЕЛЕННОЙ ВЫЧИСЛИТЕЛЬНОЙ СИСТЕМЫ

С.К. Дорожкин

Балансировка нагрузки является на сегодняшний день одним из наиболее используемых приемов повышения производительности распределенных вычислительных систем за счет оптимального распределения заданий между узлами вычислительной системы. Под термином «система балансировки нагрузки» будем понимать инструментальное средство, предназначенное для перенаправления пользовательских запросов на наименее загруженный узел распределенной вычислительной системы. Для пользователя присутствие такой системы должно быть прозрачным. Задача системы балансировки нагрузки заключается в распределении запросов пользователей между наименее загруженными вычислительными узлами распределенной системы с целью наиболее полного использования всех имеющихся ресурсов. Для выполнения этой задачи могут использоваться различные методы оценки загрузки вычислительного узла. Система балансировки нагрузки должна оперировать данными, которые содержат информацию о состоянии вычислительных узлов распределенной системы, а также правилами, заданными администратором. На основании этих данных или правил и принимается решение о перенаправлении запроса наименее загруженному узлу.

Система балансировки нагрузки должна выполнять следующие задачи: контроль за нагрузкой и состоянием вычислительных узлов, выбор, в соответствии с заданным алгоритмом, узла для передачи на него запроса пользователя.

Все алгоритмы балансировки нагрузки можно разделить на две большие группы: пассивные и активные. К группе пассивных алгоритмов балансировки нагрузки относятся те, которые работают с таблицей, содержащей сведения об обслуживаемых узлах, такие как тактовая частота процессора, количество оперативной памяти, пропускная способность внутренней шины. Для распределения запросов пользователей пропорционально производительности вычислительного узла в таблицу могут быть добавлены весовые коэффициенты. Таблица формируется до начала работы системы балансировки и почти не меняется в течение ее работы. Эти алгоритмы, в свою очередь, делятся на две группы: статические и динамические. Статические алгоритмы основаны на распределении запросов пользователя между вычислительными узлами по определенному правилу, которое устанавливается заранее и не меняется в процессе работы выполнения алгоритма. Простейшим примером такого алгоритма является алгоритм с циклической выборкой (Round robin).

К группе динамических алгоритмов относятся те, которые изменяют правила обхода вычислительных узлов во время своей работы. В процессе своего функционирования система балансировки нагрузки, работающая по динамическому алгоритму распределения запросов, изменяет порядок обхода вычислительных узлов в зависимости от их состояния. Системы балансировки нагрузки, работающие по динамическим алгоритмам распределения нагрузки, самостоятельно формируют таблицу, описывающую состояние вычислительных узлов, не получая данных от самих вычислительных узлов. Примером может служить алгоритм, основанный на таком признаке загрузки узла, как количество установленных соединений. При установлении очередного соединения между пользователем и вычислительным узлом система балансировки находит в своей таблице узел с наименьшим на текущий момент количеством соединений и отправляет задание на исполнение в этот узел, пополняя при этом соответствующую запись в таблице. Когда задание выполнено и соединение закрывается, таблица снова корректируется.

Основным недостатком всех рассмотренных выше алгоритмов является отсутствие возможности автокорректировки их работы в процессе функционирования системы

балансировки, в случае возникновения неравномерного распределения нагрузки между вычислительными узлами. Правила распределения запросов устанавливаются заранее и либо не меняются вообще, либо изменяются незначительно. Такие алгоритмы целесообразно использовать только в комплексе с еще каким-либо алгоритмом, который реализует более гибкий подход к распределению запросов пользователей.

Активные алгоритмы распределения нагрузки характеризуются тем, что система балансировки нагрузки постоянно отслеживает уровень нагрузки и доступность вычислительных узлов, которые она контролирует, с тем, чтобы на основании имеющейся информации в любой момент передать задание пользователя наименее загруженному узлу [1, 2]. Здесь можно выделить два метода получения информации о текущем состоянии вычислительного узла. Первый метод – это так называемый внешний мониторинг состояния вычислительных узлов распределенной системы, второй метод – внутренний мониторинг. Системы балансировки нагрузки, основанные на внешнем мониторинге, как и все системы балансировки, основанные на алгоритмах, которые были рассмотрены выше, имеют централизованную структуру и располагаются на одном узле распределенной вычислительной системы. Система балансировки централизованно собирает данные о состоянии всех вычислительных узлов распределенной системы. При проведении внешнего мониторинга система балансировки нагрузки рассчитывает время отклика вычислительного узла, для чего направляет на него служебный запрос и замеряет время ответа. Самый простой вариант такого метода – это использование ping-тестов по протоколу управления сообщениями Internet Control Message Protocol (ICMP). Использование служебных запросов протокола ICMP позволяет системе убедиться в готовности вычислительного узла и определить, сколько времени необходимо для передачи данных на вычислительный узел от системы балансировки нагрузки и обратно. Если система балансировки не получает отклика от вычислительного узла после нескольких последовательных запросов, то данный узел считается недоступным и исключается из списка доступных. Если время ответа узла значительно увеличивается, то это означает, что соответствующий узел перегружен, и система балансировки не передает ему запросы пользователя. Однако задержка в получении ответа может не зависеть от загрузки вычислительного узла. Здесь задержки может вносить среда передачи данных. Если нагрузка на сеть в распределенной системе неравномерна, то данные о загрузке узлов, основанные на времени ответа узла, могут быть сильно искажены.

Следующий метод реализации внешнего мониторинга основан на отправке служебных пакетов подконтрольным узлам с целью получения более детальной информации о загрузке вычислительного узла. Метод, основанный на ping-тестах по протоколу ICMP, диагностирует только стек протоколов IP. Протоколы верхних уровней не затрагиваются простым ping-тестом, и для получения данных о функционировании протоколов верхних уровней необходимо использовать более сложные запросы. Примером реализации этого метода внешнего мониторинга является метод, при котором система балансировки нагрузки устанавливает соединение с вычислительным узлом по протоколу TCP, для чего требуется осуществить обмен служебными сообщениями, состоящий из трех этапов. После установления TCP соединения система балансировки немедленно разрывает его, чтобы не занимать дополнительные ресурсы вычислительного узла. Нагрузка на стек TCP вычислительного узла оценивается по общему времени, необходимому на установление соединения с узлом. При этом серьезную нагрузку испытывает среда передачи данных, так как количество служебной информации, передаваемой по сети, существенно возрастает.

Следующим методом оценки загрузки вычислительного узла с использованием внешнего мониторинга является подход, позволяющий обеспечивать мониторинг времени отклика и готовности как самого вычислительного узла, так и приложений, работающих на нем [2]. В данном случае время отклика приложения определяется как ин-

тервал времени между отправкой запроса на предоставление данных и до момента объявления о готовности к передаче.

Во всех этих методах можно учитывать производительность вычислительного узла и вводить весовые коэффициенты. Кроме того, в отличие от пассивных алгоритмов балансировки, весовые коэффициенты можно динамически менять в процессе работы системы балансировки, что позволяет менять нагрузку на вычислительный узел более гибко.

Вывод, который можно сделать после анализа методов оценки загрузки вычислительных узлов с использованием внешнего мониторинга, следующий. Эти методы позволяют получить данные о загрузке узла, выраженные только во времени ответа на запрос системы балансировки. О таких существенных характеристиках вычислительного узла, как состояние процессора или процессоров, состояние памяти или подсистемы ввод/вывода, с помощью внешнего мониторинга нельзя получить каких-либо сведений. Кроме того, серьезным недостатком всех методов внешнего мониторинга является то, что загрузка вычислительного узла определяется по времени, которое необходимо для передачи и получения служебного пакета. Большое время ответа вычислительного узла может и не зависеть от загрузки самого узла, а быть таким из-за перегрузок в сети.

Для получения более детальной информации о состоянии вычислительного узла необходимо использовать методы второй группы.

Ко второй группе методов, с использованием активных алгоритмов для оценки загрузки вычислительного узла, относится внутренний мониторинг состояния узла. При таком подходе на каждый узел распределенной системы помещается программа-агент, которая собирает данные об узле, на котором она находится. Данные о состоянии отдельных компонентов через определенный интервал времени передаются на центральную машину системы балансировки. Главным отличием от метода внешнего мониторинга является то, что все узлы распределенной вычислительной системы активно участвуют в процессе сбора информации о загрузке. Вычислительные узлы самостоятельно инициируют сбор и обработку информации о своей загрузке. На основную машину системы балансировки данные от программ-агентов могут передаваться либо после запроса центральной (диспетчерской) части системы балансировки нагрузки, либо программы-агенты могут самостоятельно передавать сведения о загрузке подконтрольного узла на центральный узел через определенный интервал времени.

Преимуществом метода является существенное уменьшение нагрузки на коммуникационную среду, так как служебные данные передаются меньшими порциями по сравнению со служебной информацией, передаваемой при использовании методов балансировки нагрузки с использованием внешнего мониторинга состояния вычислительных узлов.

Другим преимуществом этой группы методов является то, что использование внутреннего мониторинга позволяет получить более детальную и объективную информацию о состоянии вычислительных узлов и отдельных их компонентов.

Основным недостатком методов балансировки нагрузки, основанных на внутреннем мониторинге состояния вычислительного узла, является присутствие резидентной программы на каждом узле вычислительной системы. Служебные программы-агенты поглощают определенное количество ресурсов подконтрольного вычислительного узла. Объем требуемых ресурсов зависит от того, по скольким параметрам идет оценка загрузки вычислительного узла [3], идет ли обработка этих данных на самом узле или они передаются на центральную машину системы балансировки нагрузки.

Для уменьшения количества требуемых программой-агентом ресурсов необходимо выявить, какие функции она должна выполнять. Основная функция программы-агента – предоставление данных о загрузке подконтрольного ей вычислительного узла. Поэтому при разработке системы балансировки нагрузки важным моментом является

определение критериев загрузки вычислительного узла. От выбора критерия загрузки зависит то, какие ресурсы будут затребованы программой-агентом, какие накладные расходы лягут на сеть, связывающую отдельные узлы распределенной вычислительной системы.

Если программа-агент будет собирать данные о загрузке различных компонентов вычислительного узла, то для обработки этих данных потребуется время и ресурсы самого узла. С другой стороны, оценка загрузки только одного компонента может не дать объективных данных о состоянии всего узла в целом. На наш взгляд, наиболее удачным показателем загрузки вычислительного узла является количество заданий, находящихся в очереди на исполнение. Число уже ожидающих заданий напрямую влияет на время ожидания вновь поступивших. Так как распределенная система обычно создается под конкретную задачу или область задач, то и запросы пользователей будут однотипны. Отсюда можно заключить, что критерий загрузки, основанный на определении количества заданий, находящихся в очереди на исполнение, является достаточно объективным.

Программа-агент, которая отслеживает только состояние очереди заданий на исполнение, не требует большого количества вычислительных ресурсов, что позволит уменьшить и размер самой программы, находящейся в памяти вычислительного узла.

Другой вопрос, который также влияет на ресурсоемкость программы-агента – это передача обработанной информации на центральную машину системы балансировки. Существует два подхода к решению этой задачи: программа-агент сама инициирует передачу данных через определенный интервал времени или передает данные по запросу центрального узла системы балансировки.

Первый метод позволяет постоянно иметь свежие данные о загрузке вычислительных узлов, но, с другой стороны, повышается нагрузка на сеть, так как постоянно идет передача служебных данных. Второй метод позволяет получать данные от всех узлов по запросу центрального узла в тот момент, когда необходимо поставить новый запрос клиента в очередь на исполнение.

Выбор одного из этих методов зависит от нескольких условий. Как отмечалось выше, распределенная система создается для решения определенного круга задач, поэтому заранее известно приблизительное время на обработку одного задания. Если задание ресурсоемкое, требует длительной обработки и интенсивного обмена между вычислительными узлами, то имеет смысл использовать метод, где данные о загрузке вычислительных узлов поступают по запросу центральной машины системы балансировки нагрузки. Если же задания пользователей выполняются за короткий промежуток времени, например, запросы пользователей к web-узлу, то тогда предпочтительнее использовать метод внутреннего мониторинга с «активными» программами-агентами, которые сами периодически передают данные о загрузке своих вычислительных узлов на центральную машину-диспетчер.

Системы балансировки нагрузки, построенные на активных алгоритмах оценки загрузки вычислительного узла, рассмотренных выше, получают одностороннюю информацию о текущем состоянии вычислительного узла. Внутренний мониторинг позволяет получить достаточно точную информацию о загрузке самого вычислительного узла и его отдельных компонентов, внешний мониторинг предоставляет данные о загрузке вычислительного узла по времени его отклика, таким образом, тестируется сегмент сети, в котором находится данный вычислительный узел. Для получения наиболее объективной картины загрузки как вычислительного узла, так и сетевых соединений наиболее целесообразно, на наш взгляд, использовать комбинированный подход. На каждом узле вычислительной системы должна находиться программа-агент, которая собирает данные о загрузке подконтрольного узла. Помимо этого, центральная машина системы балансировки нагрузки должна периодически тестировать сеть для определе-

ния наиболее загруженных сегментов. Это особенно важно для систем, в которых идет активное взаимодействие узлов друг с другом при обработке задания. Такой подход позволит регулировать нагрузку не только на вычислительных узлах, но и в сети.

Из всего вышесказанного можно сделать вывод, что на сегодняшний момент для эффективной балансировки нагрузки в распределенной вычислительной системе необходимо использовать комплексный подход. Классические алгоритмы балансировки нагрузки не позволяют достигнуть наилучших результатов в увеличении производительности распределенной вычислительной системы. Предлагаемый метод балансировки нагрузки основан на использовании нескольких из перечисленных выше алгоритмов в комплексе. В основу системы балансировки нагрузки должен быть положен метод внутреннего мониторинга состояния вычислительного узла. Только он позволяет получить объективную картину загрузки узла и данные о загрузке отдельных компонентов узла. Критерием загрузки вычислительного узла предлагается считать число заданий, находящихся в очереди на исполнение. Кроме того, необходим мониторинг состояния сети, связывающей отдельные узлы распределенной вычислительной системы. Третий компонент должен учитывать то обстоятельство, что современные распределенные вычислительные системы состоят из разнообразных, с точки зрения их вычислительной мощности, компонентов. Для учета этого факта необходимо использовать весовые коэффициенты, которые будут учитывать потенциальную вычислительную мощность каждого узла.

Именно использование комплексного подхода к балансировке нагрузке позволит добиться наилучших показателей в увеличении производительности распределенной вычислительной системы.

### Литература

1. Либман Л. Философия распределения нагрузки // Журнал сетевых решений LAN. 2000. №5.
2. Тао Чжоу Системы балансировки нагрузки Web серверов // Windows 2000 Magazine. 2000. № 3.
3. Воеводин В.В., Воеводин Вл.В. Параллельные вычисления. СПб: БХВ-Петербург, 2003.