

НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
УКРАИНЫ
"КИЕВСКИЙ ПОЛИТЕХНИЧЕСКИЙ ИНСТИТУТ"

На правах рукописи

АЛЬ - ДАУД АЛИ АСАД АХМАД

(Иордания)

УДК 681.324

ДИНАМИЧЕСКОЕ ВЫРАВНИВАНИЕ НАГРУЗКИ
В РАСПРЕДЕЛЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМАХ

Специальность 05.13.08 - "Вычислительные машины, системы и сети,
элементы и устройства вычислительной
техники и систем управления"

А В Т О Р Е Ф Е Р А Т
диссертации на соискание ученой степени
кандидата технических наук

Киев - 1996

004



00343955 (Т)

Диссертация является рукописью
Работа выполнена в Национальном
Украины "Киевский политехнический
вычислительной техники

Научный руководитель:

доктор технических наук,
профессор,
Луцкий Георгий Михайлович

Официальные оппоненты:

доктор технических наук,
профессор,
Зайченко Юрий Петрович

кандидат технических наук,
с.н.с.,
Антонюк Андрей Иванович

Ведущая организация:

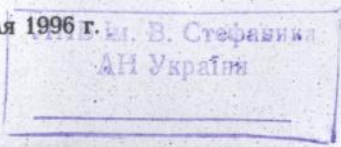
Институт проблем моделирования
в энергетике НАН Украины.

Защита состоится 20 мая 1996 г. в 14-30 часов на заседании
специализированного Совета Д 01.02.06 в Национальном техническом
университете Украины "Киевский политехнический институт"
(г.Киев, пр. Победы, 37, корп. 18, ауд. 306).

Отзывы на автореферат в двух экземплярах, заверенные печатью
учреждения, просим направлять по адресу:
252056, г. Киев, пр. Победы, 37, Ученому секретарю НТУУ "КПИ".

С диссертацией можно ознакомиться в библиотеке Национального
технического университета Украины "КПИ".

Автореферат разослан 20 апреля 1996 г.



Ученый секретарь
специализированного совета,
доктор технических наук,
профессор

Бузовский О.В. Бузовский

ДВ.34.726

АННОТАЦИЯ

Целью диссертационной работы является повышение эффективности распределенных вычислительных систем на основе развития методов и средств динамического выравнивания нагрузки между узлами вычислительных систем.

Для достижения данной цели в диссертации решаются следующие задачи:

1. Определение условий активизации процесса выравнивания нагрузки.
2. Определение очередности загрузки узлов вычислительной системы на основе сравнения их нагрузки.
3. Выбор и перемещение работ для их выполнения на других узлах системы.
4. Выделение критериев поиска оптимальных решений при решении задачи выравнивания нагрузки в распределенных вычислительных системах.
5. Разработка способов взаимодействия узлов системы при решении задачи выравнивания нагрузки.
6. Исследование централизованного и децентрализованного способов выравнивания нагрузки.
7. Выделение критериев выбора оптимальных алгоритмов выравнивания нагрузки.

Автор защищает следующие основные положения и результаты:

1. Способ синхронизации процессов выравнивания нагрузки на основе эстафетной передачи управления.
2. Алгоритмы динамического выравнивания нагрузки в распределенных вычислительных системах.
3. Алгоритмы контроля нагрузки в распределенных вычислительных системах и определения условий принятия решений о выравнивании нагрузки.
4. Имитационную модель процессов динамического выравнивания нагрузки в распределенных вычислительных системах.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Развитие вычислительной техники приводит к расширению сферы применения и увеличению размеров распределенных вычислительных систем. В условиях асинхронного и независимого поступления работ к узлам распределенных систем возникают ситуации, при которых одни узлы системы оказываются перегруженными работами, в то время как другие являются недогруженными. В указанных ситуациях использование механизмов перераспределения и выравнивания нагрузки позволяет существенно повысить эффективность использования системы и тем самым увеличить

ее пропускную способность, уменьшить время ожидания работ и время ответа.

Острая необходимость повышения эффективности распределенных систем и обеспечения своевременности выполнения работ делают разработку методов динамического выравнивания нагрузки особенно актуальной.

Методы исследований базируются на использовании основных положений теории имитационного моделирования, математической статистики, теории массового обслуживания, дисперсионного анализа и теории множеств.

Научная новизна работы заключается в разработке способа синхронизации процессов управления динамическим выравниванием нагрузки на основе эстафетной передачи управляющей информации, метода анализа и выбора алгоритмов динамического выравнивания нагрузки в распределенных вычислительных системах с помощью средств имитационного моделирования и эффективных алгоритмов динамического выравнивания нагрузки.

Практическая ценность работы состоит в разработке имитационной модели, позволяющей анализировать условия применения различных алгоритмов динамического выравнивания нагрузки распределенной вычислительной системы, способов и средств повышения эффективности распределенных вычислительных систем: увеличении пропускной способности, уменьшении времени ожидания выполнения работ и времени ответа. Предложенные алгоритмы позволяют увеличить эффективность распределенных вычислительных системы на 30–89%.

А п п р о б а ц и я р а б о т ы. Основные положения диссертации докладывались и обсуждались на: семинаре "Выравнивание нагрузки в распределенных системах", Римский университет "La Sapienza", Италия, 1992 г.; семинаре "Управление транзакциями и восстановление мультибазы данных", Римский университет "La Sapienza", Италия, 1993 г.; семинаре "Использование стратегии эстафетной передачи при динамическом выравнивании нагрузки", Римский университет "La Sapienza", Италия, 1993 г.; конференции по распределенным алгоритмам, "Летняя школа по распределенным алгоритмам", Siena, Италия, 1994 г.; Международной конференции IASTED по Сетям. Orlando, Флорида США, 1996. ; Международной конференции IASTED по прикладной информатике, Инсбрукский центр конгрессов, Innsbruck, Австрия, 1996.

П у б л и к а ц и я. По теме диссертации опубликовано 4 печатных работы.

С т р у к т у р а и о б ъ е м р а б о т ы. Диссертация состоит из введения, четырех глав, заключения, приложения и списка литературы

из 90 наименований. Работа содержит 150 страниц машинописного текста, 11 таблиц, 43 рисунка.

Во введении обоснована актуальность темы диссертационной работы, сформулирована цель и задачи исследований, сформулированы научные результаты, выносимые на защиту, дана общая характеристика работы.

В первой главе рассматриваются основные элементы распределенной вычислительной системы, исследуются распределенные алгоритмы управления выравниванием нагрузки, способы и средства распределения ресурсов и нагрузки, задача выравнивания нагрузки формализуется как задача выбора алгоритма, выравнивающего нагрузку с наибольшей эффективностью при заданных внешних и внутренних условиях работы распределенной системы.

Во второй главе рассматриваются компоненты выравнивающей нагрузки системы и систематизация выравнивающих нагрузку алгоритмов. В этой главе также рассматриваются известные из литературы алгоритмы динамического выравнивания нагрузки, их достоинства и недостатки, обосновывается необходимость разработки новых алгоритмов выравнивания нагрузки.

В третьей главе разрабатывается имитационная модель процессов динамического выравнивания нагрузки распределенной вычислительной системы, описываются предлагаемые автором алгоритмы выравнивания нагрузки, рассматриваются особенности их работы, разрабатывается модель нагрузки распределенной системы, определяются критерии, необходимые для исследования предлагаемых автором алгоритмов

В четвертой главе анализируются результаты моделирования, полученных на основе предлагаемых автором алгоритмов и алгоритмов, широко используемых на практике; определяются условия применения предлагаемых автором алгоритмов и влияние различных параметров выравнивающих алгоритмов на показатели эффективности распределенной системы.

В приложении дается детальное описание имитационной модели.

СОДЕРЖАНИЕ РАБОТЫ

Развитие вычислительной техники приводит к расширению сферы применения и увеличению размеров распределенных вычислительных систем. В условиях асинхронного и независимого поступления работ к узлам распределенной системы могут возникать ситуации, когда одни узлы системы оказываются перегруженными работами, в то время как другие будут недогруженными. Неравномерная нагрузка узлов системы может привести к неэффективному использованию ресурсов системы. С другой стороны, перегрузка отдельных узлов приводит к образованию очередей работ, увеличению времени ожидания обслуживания работ и может привести к срыву сроков выполнения работ, когда результаты вычислений будут уже не нужны пользователю системы.

Выходом из ситуаций, когда одни узлы простаивают, а другие перегружены, является динамическое выравнивание нагрузки. Существует достаточно большое число алгоритмов, позволяющих выравнивать нагрузку распределенных вычислительных систем, однако каждый из них оказывается эффективным только при определенных условиях функционирования распределенных систем. В этой связи разработка новых, более эффективных, алгоритмов выравнивания нагрузки с более широкой областью применения является актуальной и важной.

С целью получения формальной постановки задачи динамического выравнивания нагрузки распределенной системы введем следующие обозначения:

WL — множество параметров, характеризующих рабочую нагрузку системы.

Job — множество параметров, характеризующих потребность отдельных работ в ресурсах системы.

Stream — множество параметров, характеризующих порядок поступления работ в систему.

T — множество условий, характеризующих сроки выполнения работ пользователей.

n — количество узлов распределенной системы.

CS — множество параметров, характеризующих коммуникационную сеть системы.

OS — множество параметров, характеризующих используемую операционную систему.

MEM — множество параметров, характеризующих используемые средства памяти системы.

CPU — множество параметров, характеризующих используемые процессоры системы.

D — множество параметров, характеризующих используемые технические средства системы.

t_a — среднее время ответа на запрос пользователя при выравнивании нагрузки по a -му алгоритму.

t_{NoLB} — среднее время ответа на запрос пользователя при отсутствии механизмов выравнивания нагрузки.

q_a — средняя длина очереди узла при выравнивании нагрузки по a -му алгоритму.

q_{NoLB} — средняя длина очереди узла при отсутствии механизмов выравнивания нагрузки.

IN — множество внутренних условий функционирования системы.

OUT — множество внешних условий функционирования системы.

Alg — множество параметров возможных алгоритмов выравнивания нагрузки системы.

a — искомый вариант выравнивания нагрузки.

F — функциональное преобразование внутренних и внешних условий функционирования распределенной системы при конкретном варианте ее организации.

P — уменьшение времени ответа распределенной системы, получаемого от выравнивания нагрузки.

L_{overhead} — непроизводительные затраты времени, необходимые на организацию процессов выравнивания нагрузки.

Z — целевая функция задачи выравнивания нагрузки распределенной системы.

В принятых обозначениях задача выравнивания нагрузки может быть сформулирована следующим образом:

$$a = F(\text{IN}, \text{OUT}),$$

обеспечивающий

$$\max Z = P - L_{\text{overhead}},$$

при условии, что

$$a \in \text{Alg},$$

$$P = t_{\text{NoLB}} - t_a;$$

$$\text{OUT} = \{\text{WL}, \text{T}\},$$

$$\text{WL} = \{\text{Job}, \text{Stream}\},$$

$$\text{T} = \{ t_a < t_{\text{NoLB}}; q_a < q_{\text{NoLB}} \},$$

$$\text{IN} = \{n, D, \text{OS}, \text{Alg}\},$$

$$D = \{\text{CPU}, \text{MEM}, \text{CS}\}.$$

Главным объектом исследований данной работы является динамика функционирования распределенной системы, поэтому все перечисленные выше множества являются функциями времени, а целью решения задачи выравнивания нагрузки является определение составляющих этих функций.

Рассмотрим наиболее характерные алгоритмы выравнивания нагрузки — центральный и распределенный (DISTD), которые по своим характеристикам являются крайними представителями множества алгоритмов выравнивания нагрузки.

Работа центрального алгоритма основана на предположении, что существует единственный принимающий решения узел, которому известна нагрузка всех узлов, хотя бы с некоторой задержкой. Этот узел системы играет роль центрального планировщика нагрузки всех других узлов. Алгоритм работает следующим образом.

Центральный узел периодически получает сообщения об изменениях нагрузки от всех других узлов и принимает решения о выравнивании нагрузки. Вместе с тем центральный алгоритм — это алгоритм инициативного источника, поэтому для определения подходящего для перемещения процесса используется способ передачи сообщений о нагрузке. Если в некотором узле есть процесс, подходящий для перемещения, этот узел посылает центральному планировщику запрос и свое текущее значение нагрузки. После получения запроса центральный планировщик выбирает узел с наименьшей нагрузкой и определяет этот узел в качестве адресата и информирует узел-источник о новом назначении. Затем узел-источник перемещает выделенную работу в узел-адресат.

Многие авторы отмечают, что центральный алгоритм имеет наилучшие результаты по эффективности выравнивания нагрузки распределенной системы.

Недостатки этого алгоритма:

низкая надежность, так как отказ центрального планировщика приводит к отказу всей системы выравнивания нагрузки;

периодическая, каждые p единиц времени, перегрузка центрального узла при получении сообщений об изменениях нагрузки от всех узлов системы.

Другим примером алгоритма выравнивания нагрузки является алгоритм DISTD, который работает следующим образом.

Каждый узел периодически оповещает все другие узлы об изменениях своей нагрузки. Это означает, что каждый узел хранит информацию о нагрузке всей системы и периодически посылает сообщения о своей нагрузке другим узлам системы (если нагрузка узла не изменилась со времени последней посылки, то нет необходимости посылать сообщения другим узлам).

Если в некотором узле нагрузка превышает некоторый заданный заранее уровень и есть работа, подходящая для перемещения, этот узел в своем списке узлов ищет узел с наименьшей нагрузкой и, если эта нагрузка меньше нагрузки данного узла на заданную величину Δ , работа пересылается на этот узел. Если есть несколько узлов с наименьшей длиной очереди, произвольно выбирается один из них.

Достоинства алгоритма Distd:

полностью децентрализованное управление процессом динамического выравнивания нагрузки;

высокая надежность.

Недостатки алгоритма Distd:

высокие накладные расходы каждого узла (большое число сообщений, которыми периодически, каждые p единиц времени, обмениваются узлы), возрастающие линейно с ростом размера системы;

отсутствие синхронизации процессов выравнивания нагрузки, что приводит к дополнительным перемещениям работ к новым узлам-адресатам.

С целью преодоления недостатков существующих алгоритмов автором предлагаются метод синхронизации процессов выравнивания нагрузки с помощью стратегии эстафетной передачи и на его основе новые алгоритмы выравнивания нагрузки. Для описания предлагаемых алгоритмов используются несколько специальных терминов, которые объясняются ниже.

Под термином эстафета или маркер понимается специальное сообщение, передаваемое между узлами системы. Маршрут маркера определяется в каждом алгоритме по-своему. Для синхронизации процессов выравнивания нагрузки в системе вводится только один маркер, который определяется как привилегия: только узел, обладающий этой привилегией, имеет право активизировать процессы выравнивания нагрузки.

Под термином порог локальной нагрузки (T_1) понимается число ожидающих и выполняемых работ, которое выбирается таким образом, что если нагрузка узла равна или меньше (или больше, в зависимости от конкретного алгоритма) T_1 , то все работы выполняются на данном узле, иначе все прибывающие на этот узел работы ожидают выполнения на удаленных узлах. Оптимальное значение для T_1 зависит от рабочей нагрузки системы и от мощности используемой в узле ЭВМ (если рассматриваемая система состоит из ЭВМ с различной мощностью, для каждой ЭВМ назначаются свои значения T_1).

Первый из предлагаемых алгоритмов — это алгоритм инициативного ТР-сервера. В данном алгоритме владеющий маркером узел действует как инициативный сервер, то есть, если локальная нагрузка меньше, чем порог локальной нагрузки (T_1), этот узел ищет наиболее нагруженный узел и просит этот узел отдать одну или более работ с целью поднять свою нагрузку выше заданного порога. Поиск наиболее нагруженного узла выполняется путем послышки сообщений ко всем другим узлам и получения от них соответствующих ответов об их нагрузке. После завершения процесса выравнивания нагрузки маркер передается узлу с минимальной нагрузкой.

Второй из предлагаемых алгоритмов — это алгоритм инициативного ТР-источника. В данном алгоритме владеющий маркером узел действует как инициативный источник, то есть, если локальная нагрузка превышает порог локальной нагрузки (T_2), этот узел ищет наименее нагруженный узел и пересылает одну или более работ выбранному узлу. Поиск наименее нагруженного узла выполняется тем же самым способом, что и в алгоритме инициативного сервера, путем послышки сообщений ко всем другим узлам и получения от них соответствующих ответов об их нагрузке. После завершения процесса выравнивания нагрузки маркер передается узлу с максимальной нагрузкой.

Третий из предлагаемых алгоритмов — это адаптивный алгоритм инициативных ТР-сервера и ТР-источника. Одновременное использование обоих вышеописанных алгоритмов, инициативного ТР-сервера и инициативного ТР-источника, ставит адаптивный алгоритм впереди предыдущих двух алгоритмов по возможности влияния на выравнивание нагрузки системы. Адаптивный алгоритм объединяет в себе два алгоритма: алгоритм инициативного сервера и алгоритм инициативного источника. При этом, в зависимости от состояния нагрузки системы, активизируется тот или иной алгоритм. Адаптивный алгоритм работает следующим образом.

Узел, который владеет маркером, проверяет свою локальную нагрузку (LL). Если LL меньше чем заранее определенный порог (T_1), действует алгоритм инициативного сервера. Если LL больше чем заранее определенный порог (T_2), действует алгоритм инициативного источника. Если $T_1 < LL < T_2$, маркер передается следующему узлу в соответствии с последним действующим алгоритмом.

Четвертый из предлагаемых алгоритмов — это алгоритм Ramtha. В этом алгоритме владеющий маркером узел действует как администратор выравнивания нагрузки системы следующим образом.

Всем другим узлам передается запрос о состоянии их нагрузки. Ответы всех узлов собираются в один список и сортируются по возрастанию уровня текущей нагрузки узлов. Затем для каждого узла выбирается партнер таким образом, что узел с максимальной нагрузкой соединяется с узлом с минимальной нагрузкой, узел со вторым максимумом нагрузки — с узлом со вторым минимумом нагрузки и так далее. На следующем шаге алгоритма каждая пара самостоятельно решает, сколько работ необходимо переместить для выравнивания нагрузки между ними. Затем маркер передается следующему узлу сети. Если различие между максимальной и минимальной нагрузкой не больше чем заданная заранее величина s , выполнение алгоритма приостанавливается по истечении p секунд.

Во всех алгоритмах в качестве показателя нагрузки используются длины очередей в узлах. Формулы, используемые для вычисления длин очередей, будут приведены при описании модели процесса выравнивания нагрузки распределенной системы.

С целью решения поставленной задачи и оценки эффективности предлагаемых алгоритмов они сравнивались с наиболее характерными алгоритмами выравнивания нагрузки: центральным и DISTD, а также со случаем отсутствия в системе выравнивания нагрузки (NoLB).

Так как распределенные системы являются сложными и многомерными объектами исследования, в качестве основного метода исследования использовалось имитационное моделирование.

В основе имитационной модели распределенной системы лежат концепции модели отдельной работы, модели отдельного узла, модели коммуникационной сети и модели алгоритмов выравнивания нагрузки (Рис.1).

Модель отдельной работы связывается с распределением вероятностей двух случайных величин: интервала ($t_{arrival}$) времени между двумя поступлениями работ и времени ($t_{service}$) обслуживания отдельной работы. При этом время обслуживания включает затраты времени на выполнение операций ввода/вывода, работу операционной системы и оперативной памяти. Связь между основными временными показателями, используемыми в модели, представлена на рис.2.

Время ($t_{response}$) ответа системы на запрос (работу) пользователя определяется как разность между моментами времени поступления работы в систему и времени окончания обслуживания

$$t_{response} = T_{release} - T_{arrival}.$$

Время ожидания обслуживания работы (t_{wait}) пользователя определяется как разность между моментами времени поступления работы в систему и времени начала обслуживания

$$t_{wait} = T_{seize} - T_{arrival};$$

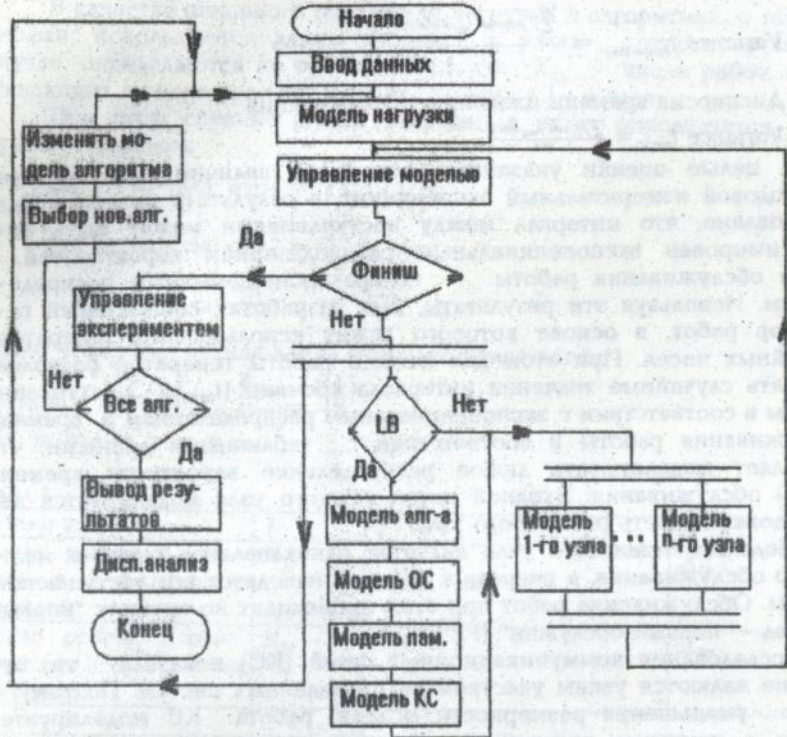


Рис.1.

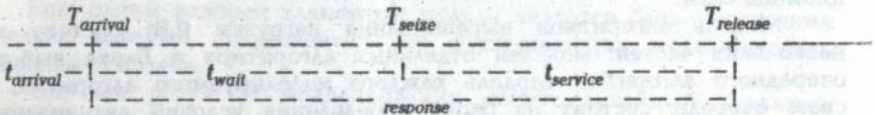


Рис.2.

Среднее время ответа системы

$$\text{Average } t_{\text{response}} = \frac{\sum_{j=1}^{n_{\text{job}}} t_{j, \text{response}}}{n_{\text{job}}};$$

Среднее время ожидания обслуживания

$$\text{Average } t_{\text{wait}} = \frac{\sum_{j=1}^{n_{\text{job}}} t_{j, \text{wait}}}{n_{\text{job}}};$$

Дисперсия времени ответа системы

$$\text{Variance } t_{\text{response}} = \frac{\sum_{i=1}^n (t_{\text{response } i} - t_{\text{response}})^2}{n-1};$$

Дисперсия времени ожидания обслуживания

$$\text{Variance } t_{\text{wait}} = \frac{\sum_{i=1}^n (t_{\text{wait } i} - t_{\text{wait}})^2}{n-1}.$$

С целью оценки указанных случайных величин был проведен полугодовой измерительный эксперимент, в результате которого было установлено, что интервал между поступлениями может быть аппроксимирован экспоненциальным распределением вероятностей, а время обслуживания работы — гиперэкспоненциальным распределением. Используя эти результаты, был разработан специальный генератор работ, в основе которого лежит использование генератора случайных чисел. При этом для каждой работы генератор позволяет получать случайные значения интервала времени (t_{arrival}) поступления работы в соответствии с экспоненциальным распределением и времени обслуживания работы в соответствии с табличными данными, что позволяет моделировать любое распределение вероятности времени (t_{service}) обслуживания. Входной поток каждого узла моделируется как последовательность работ этого узла.

Моделью отдельного узла является одноканальная система массового обслуживания, в очереди к которой находятся все поступающие работы. Обслуживание работ при этом происходит по правилу "первый пришел — первый обслужен" (FIFO).

Исследования коммуникационных сетей (КС) показали, что эти сети не являются узким участком распределенных систем. Поэтому, с целью уменьшения размерности модели, работа КС моделируется только с помощью временных задержек, добавляемых ко времени ответа системы при имитации процессов выравнивания нагрузки. Таким образом, передача сообщений и перемещение работ моделируются без имитации движения сообщений по каналам коммуникационной сети.

Модель алгоритмов выравнивания нагрузки (LB) состоит из нескольких частей: моделей отдельных алгоритмов и блока выбора очередного алгоритма. Модель каждого моделируемого алгоритма, в свою очередь, состоит из блока определения условий активизации процесса LB и моделей, учитывающих связанные с этим алгоритмом непроизводительные временные затраты на организацию LB: затраты операционной системы (ОС), затраты устройств памяти и коммуникационной сети. При этом в состав модели коммуникационной сети входят модели двух типов КС: широковещательных и коммутационных.

Блок выбора очередного алгоритма для моделирования предназначен для замены моделируемого алгоритма после очередного прогона модели. Это позволяет управлять имитационным экспериментом с помощью изменения входных данных и частично автоматизировать эксперимент.

В качестве основного показателя нагрузки в алгоритмах, и также в модели, используются длины очередей в узлах, которые, в общем случае, определяются на основе табл.1, где $X_{i,j}$ - число работ, пребывающих в очереди в i -й момент времени в j -м узле.

При этом средние длины очередей в узлах определяется следующим образом:

$$\text{для 1-го узла: } avg1 = (\sum_{i=1}^n X_{i,1})/m_1;$$

$$\text{для 2-го узла: } avg2 = (\sum_{i=1}^n X_{i,2})/m_2;$$

$$\text{для } n\text{-го узла: } avg_n = (\sum_{i=1}^n X_{i,n})/m_n.$$

Тогда в общем средняя длина очередей будет равна

$$\text{average } L = \frac{\sum_{i=1}^n X_{i,1} + \sum_{i=1}^n X_{i,2} + \dots + \sum_{i=1}^n X_{i,n}}{m_1 + m_2 + \dots + m_n} = \frac{\sum_{i=1}^n \sum_{j=1}^n X_{i,j}}{m},$$

где $m_1 = m_2 = \dots = m_n = m$.

Табл.1

Номер узла	1	2	...	n
Число работ в очереди в 1-й момент времени	X_{11}	X_{12}	...	X_{1n}
Число работ в очереди в 2-й момент времени	X_{21}	X_{22}	...	X_{2n}
..
Число работ в очереди в m-й момент времени	$X_{m,1}$	$X_{m,2}$...	$X_{m,n}$

Еще одним важным элементом модели является блок управления моделью, который предназначен для синхронизации моделирующих процессов в модели.

Так как результат моделирования носит случайный характер, завершающим этапом моделирования является обработка результатов моделирования с помощью дисперсионного анализа.

С помощью описанной выше модели имитировалась работа распределенной вычислительной системы, состоящей из ряда обрабатывающих узлов (до 40), соединенных обеспечивающей передачу данных сетью и использующих распределенную файловую систему. Целью данного эксперимента было исследование характеристик различных подходов к обмену информацией о нагрузке и перераспределению работ, а также влияние на выравнивание нагрузки следующих факторов:

размера системы, то есть числа обрабатывающих узлов в системе; уровня нагрузки системы;

значения настраиваемых параметров алгоритмов выравнивания нагрузки (табл.2);
 непроизводительных затрат на обмен сообщениями и работами;
 наличия перемещаемых работ;
 согласования действия узлов в процессе выравнивания нагрузки.

Табл.2

Обозначение	Назначение
L1	Порог локальной нагрузки для алгоритма источника
L2	Порог локальной нагрузки для алгоритма сервера
TW	Алгоритм Ramtha прекратит работу на (TW) секунд (нагрузка в системе выровнена).
P6	Интервал обмена нагрузкой (для центрального алгоритма).
L6	Порог локальной нагрузки (для центрального алгоритма).
L7	Интервал обмена нагрузкой (для алгоритма DISTD).
P7	Интервал обмена нагрузкой (для алгоритма DISTD).
Δ	Минимальная дельта нагрузки для передачи работ (для алгоритма DISTD).
N	Количество хост-машин в системе.
α	Входной коэффициент использования (для генератора).
Msend	Стоимость отправки сообщения (в единицах времени ЦП).
Mreceiv	Стоимость приема сообщения (в единицах времени ЦП).
JobTraп	Стоимость передачи работы (в единицах времени ЦП).
Su	Параметр коэффициента использования системы (используется для подпитки генератора работ).

В результате проведенных автором имитационных экспериментов было установлено, в частности, следующее:

предлагаемые алгоритмы улучшают эффективность распределенной системы на 30–84 %;

большее улучшение эффективности может быть получено за счет увеличения размера системы;

на основе существующих алгоритмов улучшение показателей выравнивания нагрузки достигается для небольшого количества узлов,

порядка нескольких десятков, а для предлагаемых алгоритмов улучшение, связанное с увеличением размера системы, имеет устойчивый характер; однако, из-за ограниченного роста пропускной способности при выравнивании нагрузки, целесообразно использовать алгоритмы выравнивания в системах из нескольких десятков узлов и нет большого стимула вводить выравнивание нагрузки в системах большего размера,

выравнивание нагрузки уменьшает среднее время ответа системы для широкого диапазона сообщений и предположений о непроизводительных затратах на перераспределение работ;

при использовании стратегии эстафетной передачи, алгоритм инициативного сервера превосходит по быстродействию алгоритм инициативного ресурса и при этом исключается проблема перегрузки узлов работами;

существуют два условия для выбора момента перемещения процессов: когда число процессов в локальной очереди превышает локальный порог нагрузки и когда перемещаемые процессы готовы продолжать удаленное выполнение;

сравнивать рабочую нагрузку различных узлов при выборе последовательности загрузки узлов распределенной системы можно используя передачу сообщений и стратегию эстафетной передачи; соответствующий компьютер должен иметь приоритет для активизации процесса выравнивания нагрузки;

для перемещения должна выбираться первая работа, которая превысит локальный порог нагрузки, если эта работа допускает удаленное выполнение; если работа относится к категории перемещаемых процессов, тогда проверяется вторая работа и так далее; при этом предполагается дисциплина обслуживания в порядке поступления;

самым лучшим адресатом для перемещаемой вычислительной работы является узел с минимальной нагрузкой;

поиском недогруженных узлов заниматься только узел, владеющий маркером.

процесс выравнивания нагрузки должен активизироваться, когда узел обнаруживает свое переполнение;

наиболее значимыми параметрами, которые необходимо учитывать при решении вышеупомянутых задач являются уровень нагрузки всей системы, локальный порог, число узлов в системе;

если данные не доступны или устарели, то обычно в этом случае возникает одна основная проблема: пересылка работы для удаленного выполнения с перегруженного узла на другой перегруженный узел, которая в свою очередь будет передавать этот процесс на другой узел, то есть вызывается бессмысленная пересылка работы;

достичь компромисса между эффективностью системы и непроизводительными затратами выбранного решения относительно выравнивания нагрузки можно уменьшая число сообщений, которые необходимы для принятия решения о выравнивании нагрузки;

избежать перегрузки недогруженного узла можно, если использовать стратегию эстафетной передачи, при этом возможность передачи

излишка работ в одно и тоже время более чем одному узлу исключается;

сотрудничество между узлами необходимо только в плане информационного обмена, а не для принятия решений относительно выравнивания нагрузки;

информация относительно выравнивания нагрузки должна собираться там, где принимаются решения.

предложенные автором алгоритмы обеспечивают достаточное улучшение эффективности системы, они были проверены для двух типов коммуникационных систем и доказали свою эффективность;

Пример зависимостей, полученных в результате моделирования, представлен на рис.3.

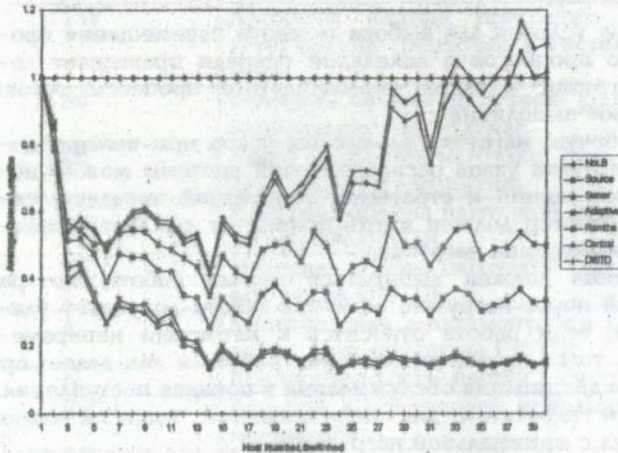


Рис.3

Рис.3. (1 — средняя длина очереди; 2 — число узлов системы; 3 — без выравнивания нагрузки; 4 — алгоритм инициативного TP-источника; 5 — алгоритм инициативного TP-сервера; 6 — адаптивный алгоритм; 7 — алгоритм Ramtha; 8 — центральный алгоритм; 9 — алгоритм DISTD)

Эти зависимости демонстрируют преимущество предлагаемых автором алгоритмов перед центральным и DISTD — алгоритмами и случае отсутствия выравнивания нагрузки (NoLB).

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

1. Разработан способ синхронизации процессов выравнивания нагрузки на основе эстафетной передачи управления, позволяющий избежать перегрузки узлов распределенной системы и уменьшить количество пересылок работ.

2. Разработаны алгоритмы динамического выравнивания нагрузки в распределенных вычислительных системах, позволяющие существенно повысить эффективность системы.

3. Разработаны алгоритмы контроля нагрузки в распределенных вычислительных системах и определены условия принятия решений, позволяющие своевременно идентифицировать перегрузку узлов системы и активизировать процессы выравнивания нагрузки.

4. Установлено, что наиболее значимыми параметрами, которые необходимо учитывать при решении задач выравнивания нагрузки, являются уровень нагрузки всей системы, порог нагрузки узла и число узлов в системе.

5. Разработана имитационная модель процессов динамического выравнивания нагрузки в распределенных вычислительных системах, позволяющая исследовать процессы выравнивания нагрузки относительно основных показателей эффективности распределенных систем.

6. Установлено, что при выравнивании нагрузки компромисс между эффективностью системы и непроизводительными затратами времени достигается на основе уменьшения числа сообщений, необходимых для принятия решений.

РАБОТЫ, ОПУБЛИКОВАННЫЕ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Аль-Дауд Али, Луцкий Г.М., Валид А. Саями. An Adaptive Source, Server-Initiative Algorithm for Dynamic Load Balancing In Distributed System, #: 240-067, IASTED International Conference on Networks. Orlando, Florida USA, 8-10, January 1996.

2. Аль-Дауд Али, Луцкий Г.М., Валид А. Саями. An Adaptive Source, Server-Initiative Algorithm for Dynamic Load Balancing In Distributed System, - IASTED International Conference on APPLIED INFORMATICS, which is to be hold at the Innsbruck Congress Center, Innsbruck, Austria, February 20-22, 1996, under the number #: 243-012.

3. Аль-Дауд Али, Луцкий Г.М., Чьоффи Д. Ramtha: Управляющий алгоритм динамического выравнивания нагрузки распределенной системы // Деп. в ГНТБ Украины 08.04.96. № 910-Ук 96.

4. Аль-Дауд Али, Луцкий Г.М., Чьоффи Д. Алгоритмы эстафетной передачи для динамического выравнивания нагрузки распределенной системы // Деп. в ГНТБ Украины 08.04.96. № 911-Ук 96.

Аль- Дауд Али Асад Ахмад, Иордания.

Динамическое выравнивание нагрузки в распределенных вычислительных системах.

Диссертация в виде рукописи на соискание ученой степени кандидата технических наук по специальности 05.13.08 - "Вычислительные машины, системы и сети, элементы и устройства вычислительной техники и систем управления" НТУУ "КПИ" Киев, 1996.

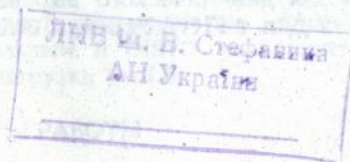
В диссертационной работе рассмотрены проблемы повышения эффективности распределенных вычислительных систем на основе развития методов и средств динамического выравнивания нагрузки между узлами вычислительных систем. Исследованы и определены условия активизации процесса выравнивания нагрузки, очередность загрузки узлов вычислительной системы на основе сравнения их нагрузки, критерии поиска оптимальных решений; предложены алгоритмы динамического выравнивания нагрузки, алгоритмы контроля нагрузки и определения условий принятия решений о выравнивании нагрузки. Предложена имитационная модель процессов динамического выравнивания нагрузки в распределенных вычислительных системах.

Al-Dahoud Ali Asa'd Ahmad

Dynamic load balancing in distributed computing systems

This scientific work is a manuscript to submit one's thesis for candidate's science in speciality 05.13.08 - Computers, systems and networks, elements and units of computer technique and control systems.

The aim of this thesis is to work and research methods and means of dynamic load balancing in distributed computing systems. A new method and four new algorithms were proposed, analyzed and compared with two standard algorithms. A model were constructed and the simulation was our analytical method.



Ключеві слова : розподілені обчислювальні системи, вирівнювання навантаження, маркерна передача управління, балансировка.

Al. Al-Dahoud

Підл. до друку 12.04.96. Формат 60×84^{1/8}. Папір
друк. № 1. Друк офсетний. Умовн. друк. арк. 0,93.
Умовн. фарбо-відб. 1,00. Облік-вид. арк. 1,0
Тираж 100 Зам. № 6-1587.

Фірма «ВІПОЛ».
252151, Київ, вул. Волинська, 60.

278636

AB 34.726