

ОПЕРАТИВНИЙ РОЗПОДІЛ НАВАНТАЖЕННЯ НА МЕРЕЖІ ПЕРЕДАЧІ ДАНИХ

Розглянуто проблему розподілу навантаження в неоднорідних інформаційних системах для підвищення швидкодії територіально розподілених серверних систем з метою подальшої розробки алгоритму динамічного розподілу навантаження в гетерогенній мережі передачі даних.

The problem of loading division in the heterogeneous information systems is considered for the increase of fast-acting of the territorial distributed server systems with the purpose of subsequent development of dynamic loading division algorithm in the heterogeneous data communication network.

Ключові слова: передача даних, інформаційно-обчислювальні мережі.

Вступ

Одним із головних факторів, що забезпечує функціонування масштабної сучасної компанії є впровадження в ній сучасних досягнень у сфері інформаційно-обчислювальних технологій. Серед них одну з лідируючих позицій займають інформаційно-обчислювальні мережі, які на даний час є визначним фактором у багатьох компаніях, адже від них напряду залежить її нормальне функціонування.

Але створення великих та надійних інформаційних систем пов'язане з деякими проблемами:

- можливість аналізу мережевого трафіку;
- аналіз та передбачення навантаження на сервер;
- аналіз та передбачення навантаження на мережу та мережне обладнання;
- архітектура системи;
- маршрутизація пакетів всередині системи;
- функціональність розподілу навантаження.

Практична значимість дослідження полягає в створенні алгоритму, що ефективно долатиме проблеми, які виникають при збільшенні навантаження на інформаційно-обчислювальні мережі [1].

Постановка задачі

Мета, яку потрібно досягти при вирішенні даного питання, – розподіл навантаження в неоднорідних інформаційних системах для підвищення швидкодії територіально розподілених серверних систем.

Задача, яка стоїть в основі дослідження, – порівняльний аналіз існуючих алгоритмів розподілу навантаження на мережі; розробка математичної моделі інформаційної системи; розробка алгоритму динамічного розподілу навантаження в неоднорідному інформаційному середовищі; провести аналіз ефективності алгоритмів, розробка яких є ціллю дослідження та визначення найефективнішого серед них.

Загальні тенденції

Існує декілька способів підвищення швидкодії Інтернет-ресурсів, такі як збільшення пропускної здатності мережі, розробка більш ефективних додатків, модернізація апаратних та програмних засобів серверів та активного мережного устаткування. Ще один спосіб, що дає збільшення швидкодії вузлів мережі – збільшення кількості серверів та використання «дзеркального» контенту для розподілу навантаження в мережі. Розподіл навантаження на декілька серверів, дозволяє запобігти таким ситуаціям, коли дані, що передаються по мережі, перевантажують один сервер, в той час як інші простоюють. Для розподілу навантаження в мережі найчастіше використовуються DNS-сервери, що передбачають можливість кругової передачі IP-адреси першого вільного або найближчого серверу, що містить сайт, будь-якому клієнту, і як результат, розподіл навантаження між серверами в мережі [2, 3].

Ефективність функціонування розподільчої серверної мережі багато в чому залежить від того, яким саме чином всередині неї розміщена інформація. Проблема розміщення даних перед усім породжує цікаві задачі дискретної оптимізації, при якому розрізняють розміщення файлів та види служб. В першому випадку на полярних позиціях знаходиться повне дублювання усієї інформації на кожному сервері та повний розподіл, коли вся множина даних ділиться на N непересічних підмножин. Перше рішення неекономічне, так як дублювання витрачає значну частину ресурсів. Другий випадок – ненадійний в своєму роді, тому що не гарантує балансування навантаження (проблема виникає коли одні дані набувають більшої популярності, а як наслідок зростає кількість звернень до них, ніж до інших). Найкращим рішенням зазвичай являється дублювання однієї множини файлів і розподіл іншої. При такому зв'язку двох типів виникають задачі оптимального розподілу даних серверу на підмножини та вибір оптимальної кількості копій. Нарешті, існують алгоритми, що передбачають динамічну міграцію документів між хостингами.

Одним із підходів вирішення задачі розподілу або балансування навантаження в інформаційній мережі можна вважати Round-robin алгоритм. Round-robin – алгоритм розподілу навантаження розподіленої обчислювальної системи методом перебору її елементів по коловому циклу. Суть алгоритму полягає в наступному: нехай є N об'єктів, здатних виконати задану дію, і M завдань, які повинні бути виконані цими об'єктами. Припускається, що об'єкти n рівні по своїх властивостях між собою, завдання m мають рівний пріоритет. Тоді перше завдання ($m = 1$) призначається для виконання першому об'єкту ($n = 1$), друге –

другому і т.д., до досягнення останнього об'єкту ($m = M$). Тоді наступне завдання ($m = N+1$) буде призначено знову першому об'єкту і т.п. Простіше кажучи, відбувається перебір об'єктів, що виконують завдання, по циклу, або по колу (round), і після досягнення останнього об'єкту наступне завдання буде також призначено першому об'єкту. Вважається, що рішення задач може бути додатково розбите на кванти часу, при чому для продовження рішення в часі нумерація об'єктів (i , відповідно, призначені завдання) зсувається по колу на 1, тобто завдання першого об'єкту віддається другому, другого – третьому, і т.д., а перший об'єкт отримує завдання останнього, або звільняється для прийому нового завдання. Таким чином, алгоритм Round-robin стає алгоритмом розподілу часу або балансування навантаження.

З іншого боку, існують алгоритми динамічного перенесення навантаження, які називають RCL-стратегією. RCL – стратегія перенесення навантаження, в яку входять три алгоритми динамічного перенесення навантаження, запропоновані розробниками SPEEDES:

- випадковий алгоритм (random, R);
- алгоритм, заснований на комунікаціях (communication, C);
- алгоритм, заснований на обчисленні навантаження (load, L).

Алгоритм перенесення об'єкту або процесу з одного комп'ютера достатньо складний. Введення деяких розумних допущень в незначній мірі знижує складність цього алгоритму. Отже, зроблені наступні допущення:

- розподілена система є гетерогенною;
- користувач може інтерактивно взаємодіяти з машиною у будь-який момент часу;
- кількість обчислювальних вузлів в мережі обмежена;
- можлива зміна мережі;
- затримки на комунікацію в мережі залишаються незмінними;
- витрати на перенесення навантаження з одного на інший процесор порівняний з фізичними розмірами цього навантаження.

RCL стратегія використовує дворівневий процес ухвалення рішень, який об'єднує централізований і децентралізований підходи. Дворівневий процес ухвалення рішень припускає:

- центральний процес-координатор приймає рішення про перенесення навантаження;
- кожен процес, що посилає повідомлення, відповідальний за вибір навантаження, яке слід перенести.

На першому рівні вибирається центральний координатор для того, щоб серед всіх робочих станцій саме він міг ухвалювати рішення. Під час процесу перенесення навантаження всі робочі станції припиняють свою роботу на деякий проміжок часу, поки центральний координатор збирає інформацію про завантаження процесорів. Координатор аналізує інформацію, яка включає зведення про завантаження комп'ютерів і про розподіл навантаження в системі. Координатор, на підставі цієї інформації вирішує, чи слід робити дії по перенесенню навантаження з одного комп'ютера на інший.

Якщо стан системи такий, що перенесення навантаження необхідне (стан відповідає критерію перенесення навантаження), центральний координатор починає процедуру перерозподілу навантаження, керуючись швидкістю роботи комп'ютерів.

Якщо в перенесенні навантаження немає необхідності, комп'ютер повинен розіслати повідомлення про це всім іншим комп'ютерам. В цьому випадку дії другого рівня ігноруються, і кожен комп'ютер продовжує обробляти свої локальні події. Якщо, знаходячись на другому рівні, робочі станції отримають від центрального координатора сигнал про необхідність міграції навантаження, вони починають процес вибору частини навантаження для перенесення його на інший хост. При виборі використовують RCL-стратегію. Після того, як частини навантаження вибрані на всіх тих робочих станціях, що пересилають дані, починається процес міграції. Послідовність кроків, що робляться при міграції навантаження між двома комп'ютерами, включає упаковку даних і їх відправку з комп'ютера-відправника, і прийом і розпаковування на комп'ютері-одержувачі. В кінці кожного кроку міграції комп'ютері-одержувачі розсилає всім іншим повідомлення про завершення процедури міграції і про нове розміщення даних. Після цього кожен комп'ютер продовжує обробку.

Розглянемо різні підходи до балансування навантаження, засновані на технології клієнт-сервер. Виділяють наступні категорії:

- клієнтські;
- засновані на DNS;
- диспетчерські;
- серверні.

Клієнтський підхід реалізує вибір сервера на стороні клієнта. Клієнти можуть вибрати один з веб-серверів випадковим чином або вибрати найбільш відповідний сервер, використовуючи механізми інтелектуального вибору. Браузер Netscape Navigator використовує клієнтський підхід при доступі до своїх сайтів. Коли користувач проглядає домашню сторінку Netscape, Navigator випадковим чином вибирає один з серверів і направляє йому запит користувача. Випадковий вибір не може гарантувати балансування завантаження і доступність сервера. Інтелектуальний вибір сервера може бути реалізований з використанням Java аплетів, запущених на стороні клієнта для визначення стану серверів і затримок в мережі. В цьому випадку може бути вибраний найбільш відповідний сервер і запит відправлений саме йому. Недолік даного підходу полягає у великій тимчасовій затримці, викликаний визначенням станів серверів.

Підхід, заснований на DNS, – це рішення на стороні системи. Сервер служби імен доменів – це

механізм маршрутизації для розподілених веб-серверів. Він може вибрати один з веб-серверів для обробки запиту шляхом відображення уніфікованого покажчика інформаційного ресурсу (URL) на IP адресу веб-сервера. Проте DNS-сервер може стати вузьким місцем в процесі маршрутизації. Існують програмні продукти для розподілу навантаження між багатьма географічно розподіленими серверами, в яких DNS-сервер визначає і доступність серверів, і тимчасову затримку в мережі для вибору відповідного сервера, що використовує технологію клієнт-сервер.

Щоб отримати повний контроль над маршрутизацією клієнтських запитів диспетчерський підхід виконує відображення адреси на рівні IP. Він визначає групу веб-серверів з одиночною віртуальною IP-адресою, якою є IP-адреса диспетчера. Диспетчер у свою чергу діє як центральний планувальник і має повний контроль над маршрутизацією запитів. Диспетчер ідентифікує кожен сервер з приватною адресою і, якщо потрібно, перенаправляє клієнт-серверний пакет, переписавши IP-адресу. У основі лежить механізм на основі TCP-маршрутизатора, TCP-маршрутизатор грає роль диспетчера. Основним недоліком підходу з переписуванням пакету є накладні витрати на зміну адреси.

Інша схема: перенаправлення пакетів – перенаправлення на рівні HTTP. Тут диспетчер розподіляє вхідні запити по серверах через механізми перенаправлення, HTTP, що надаються. Диспетчер може перенаправити запит посилкою відповіді клієнта, при цьому в заголовку відповіді вказана адреса сервера. В цьому випадку клієнт ще раз відправляє запит, але вже безпосередньо на сервер.

Серверний підхід використовує механізм дворівневої диспетчеризації. DNS-сервер спочатку визначає сервер, на який йде клієнтський запит. Після цього кожен сервер може перевизначити запит до іншого сервера. Це децентралізована стратегія перерозподілу навантаження, в якому всім серверам дозволена участь в процесі балансування навантаження. Існує реалізація цього підходу на основі веб-сервера Apache. Ця реалізація дозволяє провести однорідний перерозподіл HTTP запитів від переобтяжених серверів до тих, які мало використовуються. Запит передається по ланцюжку функцій ухвалення рішень, званих функціями-кандидатами. Кожна функція-кандидат переупорядковує набір доступних серверів і вибирає метод перерозподілу запиту (HTTP перенаправлення або через HTTP-проксі), ґрунтуючись на інформації про ресурси.

Що стосується генерації та розподілу динамічного вмісту, то ефективність планування робіт багато в чому визначається бізнес-логікою та технологіями розробки Web-додатків.

При дослідженні варто звернути увагу та використовувати методи теорії масового обслуговування, імітаційного моделювання, об'єктно-орієнтованого програмування.

Основна реалізація, як теоретичної так і практичної частини дослідження, має бути реалізована у вигляді програмного комплексу, що виконує моделювання роботи неоднорідної інформаційної системи.

На даний момент можливе вирішення даної проблеми за допомогою апаратних та програмних засобів, а також кластерних систем.

Апаратні продукти

Серед апаратних засобів найбільший досвід на ринку засобів розподілу навантаження має компанія Cisco Systems. Компанія контролює, мабуть, найбільшу частку ринку, а сьогодні його довгожителем є продукт LocalDirector. Cisco по праву вважається провідним постачальником маршрутизаторів, завдяки чому, зокрема, і отримала перевагу над іншими виробниками таких продуктів. Проте, не можна стверджувати, що продукти від Cisco, а саме маршрутизатори найбільш придатні для розподілу навантаження на інформаційно-обчислювальні мережі. Звичайно, адміністраторам, знайомим з інтерфейсом командного рядку і Internetwork Operating System (IOS), простіше впоратися із завданнями та конфігурувати дані пристрої, ніж їх не інформованим колегам, проте операції по розподілу навантаження від LocalDirector прозорі для маршрутизаторів.

LocalDirector – це програмно-апаратний комплекс на базі мікропроцесорів Intel Pentium, що працює під управлінням різновиду IOS, від компанії Cisco. Він забезпечує розподіл навантаження в неоднорідній інформаційній мережі таким чином, щоб сервер завжди був доступним. Пристрій має два інтерфейси: порти Ethernet 10/100 Мбіт/с, які поставляють в стандартній комплектації, і, крім того, за бажанням, є можливість встановлення інтерфейсу FDDI. Одному з інтерфейсів присвоюється одна, або декілька віртуальних IP-адрес [4].

Програмні продукти

Серед програмних засобів розподілу навантаження можна виділити компанію Resonate Software, що випускає три продукти, кожен з яких втілює собою певний рівень розподілу навантаження.

Система Central Dispatch – це програмний продукт, який працює на усіх платформах NT або Solaris. З точки зору архітектури – це не міст, комутатор або маршрутизатор, як розглянуті вище продукти. Central Dispatch визначає віртуальну IP-адресу, але тільки після того, як Central Dispatch Primary Scheduler перенаправив вхідний трафік на дійсну IP-адресу, після чого сервер взаємодіє безпосередньо з користувачем в процесі з'єднання. Таким чином, на відміну від апаратних засобів, Primary Scheduler не входить до маршруту даних після того як він перенаправить трафік на конкретний сервер. Primary Scheduler може бути як сервером даних, так і диспетчером. Secondary Scheduler забезпечує захист від збоїв, беручи на себе обов'язки Primary Scheduler.

Кластерні системи

Кластерні з'єднання також використовують для розподілу навантаження в інформаційно-обчислювальних мережах. Під кластерами, як правило в таких випадках, розглядають системи загального

призначення, що розподіляють потоки між клієнтами мережі у цілях підвищення її швидкодії. Таким системам потрібні для роботи спеціально орієнтовані для даних цілей ОС та спеціально розроблені програмні продукти. Більшість проблем з мережами, а саме та, що розглядається, проблема перевантаження мережі, можна вирішити за допомогою використання кластерних систем. Одним із вдалив прикладів кластерних систем загального призначення є продукт компанії «Novell – Cluster Services». Вона представляє собою систему для кластеризації серверів, що забезпечує доступність даних та змогу керування важливими даними на досить високому рівні, включаючи як дані на серверах так і їх сервіси, служби, додатки. Novell – Cluster Services – це багатовузловий кластерний продукт для NetWare 6, підтримуючим eDirectory, режими «перенос при збої» (failover), та «повернення після відновлення» (failback), а також міграцію (балансування навантаження) окремих ресурсів.

Кластерні служби Novell Cluster Services (NCS), що входять до поставки NetWare 6, мають ліцензію на використання двох комп'ютерів у кластері, але існує можливість розширення цієї кількості до 32-х вузлів. NCS забезпечує захист від збоїв та балансування навантаження критичних мережеских ресурсів, включаючи дані, що містяться на сервері, додатків, серверних ліцензій, служб і т.д.

Один кластер NetWare може містити в собі сервери з різноманітним обладнанням. Адміністратори мережі можуть точно вказати, який із серверів кластера візьме на себе функції серверу, що вийшов з ладу, тим самим розподілити визначені додатки та дані по найбільш відповідним серверам в кластері.

Висновок

Актуальність даної теми в наш час очевидна, так як постійне зростання робочих станцій збільшує навантаження на мережу та програмно-апаратні засоби, потрібно створити ефективний механізм для забезпечення стабільної роботи інформаційно-обчислювальних систем та мереж.

Поряд з вже розглянутими методами існує ще декілька «гібридних» методів та маловідомих рішень. При порівняльному аналізі вищеописаних методів, найбільшу увагу заслуговує програмно-апаратний комплекс LocalDirector від компанії Cisco, так як він є своєрідним «гібридом», що забезпечує виконання своїх функцій найефективніше, маючи можливість виконувати роботу як на апаратному так і на програмному рівні.

Література

1. Аллан Леинванд Конфигурирование маршрутизаторов Cisco / Аллан Леинванд, Брюс Пински; [пер. с англ.] – [2-е изд.]: М.: Вильямс, 2001. – 368 с.
2. Моисеев Т.Н., Распределение информационных потоков данных в распределенных многосерверных системах / Моисеев Т.Н. – Воронеж: Научная книга, 2005. – 145 с.
3. Рогачко Е.С. Динамическое распределение нагрузки в сетях массового обслуживания / Рогачко Е.С. – Саратов, 2007. – 102 с.
4. Французов Д. Оценка производительности вычислительных систем. Открытые системы / Французов Д. – М., 1996. – С. 58-66.

Надійшла 23.5.2010 р.

УДК 389: 638.011.54

В.Т. КОНДРАТОВ

Институт кибернетики им. В.М. Глушкова НАН Украины

МЕТОДЫ ИЗБЫТОЧНЫХ ИЗМЕРЕНИЙ: ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ И КЛАССИФИКАЦИЯ

У статті викладена сутність методів надлишкових вимірювань, основні визначення і класифікація. Наведено конкретний приклад створення методу надлишкових вимірювань при лінійній функції перетворення вимірювального каналу. Показано, що дійсно вони забезпечують досягнення системи цілей та отримання системи знань про вимірювальну фізичну величину і про стан засобу вимірювань.

In present paper the essence of redundant measurements methods, the basic definitions and classification is stated. The concrete example of creation of redundant measurements methods is resulted at linear function of transformation of the measuring channel. It is shown, that is valid they provide achievement of system of the purposes and reception of system of knowledge of required physical size and about a measuring apparatus condition.

Ключові слова: избыточные измерения, классификация.

Введение

Третье тысячелетие – это тысячелетие новой эры высокоточных измерений, поскольку на основе общенаучной методологии системного подхода и информативной избыточности в 2001 году была создана новая стратегия измерений, инвариантных к воздействиям внешних дестабилизирующих факторов. Данная стратегия гарантирует получение точности измерений, соизмеримой с точностью физических величин (ФВ), воспроизводимых мерой или стандартными образцами, а также высокую метрологическую надежность