

УДК 004.7:004.89(045)

Іванкевич О.В., канд. техн. наук,
Лукашенко В.В.

ЗАСОБИ КЕРУВАННЯ ПОТОКАМИ ДАНИХ У РОЗПОДІЛЕНИХ ОБЧИСЛЮВАЛЬНИХ СИСТЕМАХ

Факультет комп'ютерних систем
Національного авіаційного університету

Робота присвячена створенню системи керування потоками даних у комп'ютерній мережі розподіленої обчислювальної системи на основі апарата експертних систем. Досліджено існуючі механізми одержання, зберігання й обробки великих обсягів даних у розподілених системах та особливості підвищення продуктивності комп'ютерних мереж. Запропоновано засоби інтелектуального керування потоками даних при використанні розподілених обчислювальних систем

Вступ

На сьогодні сфера застосування розподілених обчислювальних систем (РОС) безупинно розширюється, охоплюючи все нові застосування в різних галузях науки, бізнесу й виробництва. Можна виділити коло фундаментальних і прикладних проблем, ефективне рішення яких можливо тільки з використанням РОС. Це коло включає загальні задачі управління, прогнозування погоди, клімату й глобальних змін в атмосфері, науки про матеріали, побудову напівпровідникових приладів, астрономію, транспортні задачі, гідро- та газодинаміку, розпізнавання й синтез мови, розпізнавання зображень, тощо [1–5].

Якщо традиційно РОС застосовувалися в основному в науковій сфері для рішення обчислювальних завдань, що вимагають потужних обчислювальних ресурсів, то зараз через розвиток бізнесу різко зросла кількість компаній, що відводять використанню комп'ютерних технологій і електронному документообігу головну роль. У зв'язку із цим росте потреба в побудові обчислювальних систем, що базуються на мережній технології та обробляють інформацію для критично важливих додатків, пов'язаних з обробкою транзакцій, керуванням базами даних (БД) і обслуговуванням телекомунікацій.

Можна виділити дві основні сфери застосування таких систем: обробка транзакцій у режимі реального часу (*OLTP*, *on-line transaction processing*) і створення

сховищ даних для організації систем підтримки прийняття рішень (*Data Mining*, *Data Warehousing*, *Decision Support System*) [2]. Система для глобальних корпоративних обчислень насамперед є централізованою системою, з якою працюють практично всі користувачі в корпорації, і, відповідно, вона повинна увесь час перебувати в робочому стані. Як правило, рішення подібного рівня встановлюють у компаніях і корпораціях, де навіть короткочасні простої мережі можуть призвести до величезних збитків. Тому для організації такої системи не може бути використаний звичайний сервер зі стандартною архітектурою, цілком придатний там, де немає жорстких вимог до продуктивності й часу простою. Таким системам необхідно використовувати високопродуктивні РОС для глобальних корпоративних обчислень, які відрізняються такими характеристиками як підвищена продуктивність, масштабованість, мінімально припустимий час простою [2].

Не дивлячись на досягнення в області обробки даних у РОС, залишаються проблеми, що вимагають додаткових наукових досліджень – подальший розвиток методів розподілу навантаження при обробці великих масивів даних за допомогою розподілених серверних систем і розробка нових засобів створення швидких і оптимальних інформаційних систем з високою доступністю даних при збереженні або невеликому збільшенні вартості.

Мета

Метою роботи є дослідження існуючих механізмів одержання, зберігання й обробки даних у РОС та особливостей підвищення продуктивності комп'ютерних мереж при використанні РОС.

Аналіз існуючих систем керування РОС

Огляд літературних джерел [1–3] показав, що відомі методи і алгоритми оптимізації РОС, що обробляють БД великих розмірів і отримання результатів запитів, носять загальний або приватний характер, не враховують характерні для комп'ютерних мереж особливості, або складні для реалізації. У той же час складність комп'ютерних мереж увесь час збільшується, тому необхідне постійне вдосконалення систем керування РОС. Завдання надання якісних послуг по розподіленій обробці інформації є актуальними [4–5] і в значній мірі їх можуть вирішити інтелектуальні системи керування комп'ютерними мережами [3–6].

Інтелектуальна система керування здатна до «розуміння» і навчання відносно об'єкта керування, збурювань, зовнішнього середовища й умов роботи [7]. Основна відмінність інтелектуальних систем – наявність механізму системної обробки знань.

Інтелектуальні системи керування відрізняються від традиційних механізмом одержання, зберігання й обробки знань для реалізації своїх функцій. Розробка нових інтелектуальних алгоритмів керування комп'ютерними мережами дозволить створити РОС, продуктивність яких буде близька до пікової продуктивності мережі [4, 5, 7, 9].

Проведено порівняльний аналіз найпоширеніших систем керування РОС, який дозволив виявити їхні основні недоліки. Для аналізу були обрані системи: *HP OpenView Network Node Manager*; *IBM Tivoli NetView*; *Sun Solstice Domain Manager*.

Виявлено, що у відомих системах є примітивні алгоритми, які не дозволяють у повному обсязі вирішувати завдання

пов'язані з прогнозуванням поведінки РОС. Процес адміністрування таких систем залишається складним і трудомістким, потребує від адміністратора великого досвіду й знань. Тому необхідність у розробці систем керування з потужними засобами штучного інтелекту є актуальним завданням.

Завдання керування РОС містять у собі безліч різних підзадач – завдання пов'язані з оптимізацією параметрів мережного устаткування, завдання вибору оптимального режиму роботи устаткування, виявлення можливих неполадок при маршрутизації в мережі, несумісність мережного устаткування й т. і.

Система керування РОС

Під час роботи системи керування мережею РОС програмному модулю необхідно провести: аналіз топології керованої мережі; збір інформації з активного мережного устаткування; аналіз отриманої інформації й прийняття рішень; керування устаткуванням з урахуванням обробленої інформації.

Таким чином, система керування устаткуванням РОС являє собою програмне забезпечення, яке реагує на зміну станів пристроїв, прогнозує можливі неполадки в роботі мережі й дає можливість оперативно усувати неполадки.

Щоб якісно й ефективно адмініструвати РОС, система повинна вирішувати типові завдання інтелектуальних систем керування [9]: нормування показників; прогнозування, діагностика; оцінка внутрішніх параметрів системи; оцінка ефективності функціонування; оптимальне регулювання параметрів функціонування.

Для ухвалення рішення по керуванню комп'ютерною мережею немає необхідності в аналізі всіх параметрів бази *MIB*, тому що вона має більшу кількість несуттєвих для нашого завдання параметрів. Обрано 28 параметрів, які використані як базові, для побудови систем моніторингу.

Список з 28 параметрів можна значно розширити: додати параметри, що характеризують стан портів активних мере-

жних пристроїв або параметри, що відображають максимальну довжину пакетів, що пересилаються, але, на думку провідних виробників програм мережного моніторингу, такої необхідності немає.

У роботі вирішено завдання розробки експертної системи здатної управляти устаткуванням комп'ютерних мереж, що є

основою розподілених обчислювальних систем. Розроблено багаторівневу *NIMS* (*Network Intelligent Management System*), що представляє собою програмну реалізацію експертної системи (рис. 1). Систему *NIMS* створено за допомогою мови *Delphi* з використанням СУБД *MS SQL Server*.



Рис. 1. Структура запропонованої *NIMS*

Розроблено структуру модуля «База знань» (БЗ) (рис.2) для одержання інформації щодо топології керованої мережі, інформації з активного мережного устаткування та аналітики отриманої інформації, дозволяє обробляти більші обсяги інформації й одержувати нові знання, які використовуються системою для рішення завдань по керуванню РОС.

Реалізація алгоритмів оптимізації розміщення завдань на обробку по вузлах реальної комп'ютерної мережі передбачає наявність достовірних вихідних даних. Для визначення інтенсивності звернень до різних файлів БД створено резидентну програму "Аналізатор запитів". Ця резидентна програма була завантажена на всіх вузлах мережі. Час всіх комп'ютерів син-

хронізований. Протягом одного робочого дня накопичувалася інформація.

Програма аналізує звертання до файлів і протоколює в журнал дату, час, ім'я файлу, довжину запиту, час до відповіді, довжину відповіді. Протоколювання вноситься у внутрішній буфер і записується на диск тільки під час "холостого" ходу машини, тобто коли здійснюється введення з клавіатури або інші операції, що вимагають очікування відповіді.

За допомогою нерезидентної частини програми проведений аналіз інформації, що обробляється й виявлені тимчасові характеристики звернень до файлів. При нерівномірному завантаженні мережі спостерігалось як повна відсутність обігів, так і одночасний обіг із всіх робочих станцій.

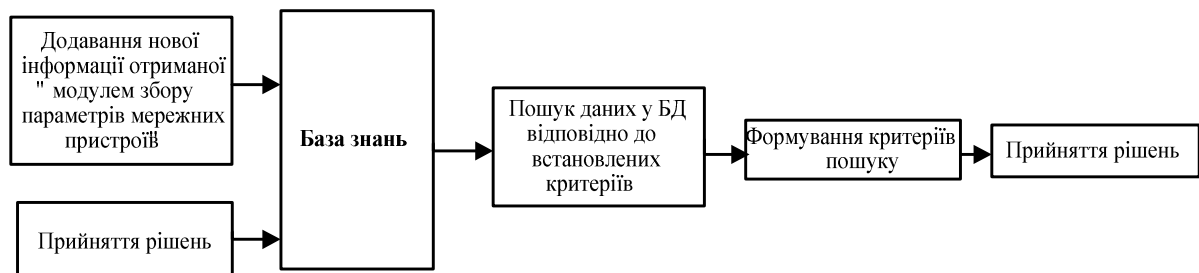


Рис. 2. Структура модуля «База знань»

Для перерозподілу пошукових запитів до інформаційно-довідкових файлів між вузлами комп'ютерної мережі призначена програма "Оптимізатор запитів у локальній мережі". Програма істотно зменшує ймовірність утворення черг, які при досягненні інтенсивністю запитів певного порога сильно "гальмують" роботу мережі.

Програма "Оптимізатор запитів" складається із трьох частин: інсталяція, перехоплення запитів і обробка запитів.

Алгоритм роботи інсталяційної частини включає:

1) перевірка наявності себе в оперативній пам'яті (ОП). Якщо в ОП не була раніше завантажена, то триває подальша установка. У протилежному випадку програма завершує свою роботу;

2) відкриття робочих сокетів. Якщо сокети вже були відкриті іншою програмою, або ж відкриття не відбулося з технічних причин, то про це видається повідомлення й програма припиняє подальшу установку;

3) заповнення внутрішньої таблиці місцезнаходження інформаційних файлів з файлу журналу. Якщо інформаційно-довідковий файл журналу не знайдений або записи в ньому не відповідають не-

обхідному формату, то видається повідомлення й установка програми припиняється;

4) одержання мережної адреси даного вузла;

5) одержання адрес векторів використовуваних переривань;

6) одержання інформації з векторів переривань на свій оброблювач;

7) вихід із програми зі збереженням її резидентної частини в оперативній пам'яті.

Розроблена програма повинна бути встановлена на всіх вузлах мережі, які будуть використовувати переваги перерозподілу запитів. Тому, для щоденного використання може виявитися доцільним включити завантаження програми у файл авто запуску операційної системи або у файл завантаження мережі.

Показники, що характеризують роботу центрального сегмента РОС стандарту *Ethernet* з кількістю робочих станцій рівним 300 у період після впровадження *NIMS* наведені у табл. Підвищення середньої швидкості передачі даних клієнтам у моменти пікового завантаження й зменшення обсягів трафіку, що буферизується комутатором, підтверджує ефективність розробленої системи керування.

Таблиця 1. Показники центрального сегмента РОС

Середня швидкість передачі даних від файлового сервера клієнтам, Мбіт/с	5,77
Середня швидкість передачі даних від файлового сервера клієнтам у момент пікового завантаження, Мбіт/с	18,03
Середня кількість перевантажень, обумовлених перевищенням вхідного трафіка в комутатор над вихідної, у годину	189
Середнє завантаження буфера комутатора під час перевантаження, Мб	2,64
Коефіцієнт кореляції між інтенсивністю вихідного трафіка файлового сервера й завантаження буфера комутатора	0,59

Запропонована *NIMS* крім допомоги в роботі адміністратору мережі, діагностує роботу мережних пристроїв, прогнозує збої, дозволяє самостійно управляти роботою мережі, налаштовувати параметри мережних пристроїв для забезпечення ма-

ксимальної продуктивності мережі в цілому.

Висновки

1) Проведено аналіз існуючих інтелектуальних систем керування РОС;

2) запропоновано систему керування навантаженням мережі в РОС для обробки великих обсягів даних;

3) результати роботи системи свідчать про можливість значного збільшення швидкості обробки даних великого обсягу у РОС за рахунок використання механізмів оптимізації завантаження вузлів мережі;

4) запропонована система може бути використана у інформаційних системах для яких ефективність роботи з даними залежить від можливостей системи швидко обробляти великі масиви інформації;

5) у подальшому планується отримати кількісні оцінки ефективності запропонованої системи.

Список літератури

1. Жуков І.А., Дровозов В.І., Масловський Б.Г. Експлуатація комп'ютерних систем та мереж: Навч. посібник. – К.: НАУ, 2007. – 368 с.

2. Жуков И. А. Новые компьютерные технологии проектирования телекоммуникаций гражданской авиации Украины // Проблемы информатизації та управління. – К.: НАУ, 2009. – Вип. 3 (27). – С. 62–72.

3. Иванкевич О.В., Кременецкий Г.М., Мазур В.І. Інформаційні системи та структури даних: Навч. посібник. – К.: НАУ, 2006. – 232 с.

4. Олифер В. Г., Олифер Н. А. Компьютерные сети. Принципы, технологии, протоколы: Учебник для вузов. – 3-е изд. – С.Пб.: Питер, 2006. – 958 с.

5. Иванкевич А.В. Использование специализированных программных комплексов на базе распределенных хранилищ данных авиапредприятий Украины // Проблемы информатизації та управління. – Вип. 1(23). – К.: НАУ, 2008. – С. 138–142.

6. Иванкевич А.В., Салим Аль Шибани. Организация системы распределенной обработки запросов к серверам баз данных в компьютерных сетях // Збірник наукових праць за результатами міжнародної науково-практичної конференції "Мікропроцесорні пристрої та системи в автоматизації виробничих процесів". –

№3. – Хмельницький: Технологічний університет Поділля, 2007. – Т.1. – С. 82–85

7. Кузнецов И.П. Семантические представления. – М.: Наука, 1985.– 186 с.

8. Глущенко В.В. Диагностико-прогнозирующие системы управления информационными процессами в сетевых объектах. – СПб.: СПГУВК, 1999. – 285 с.

9. Жуков И. А., Зайченко Ю. П., Печурин Н. К. Модель распределения информационных ресурсов в компьютерных сетях // Проблемы информатизації та управління. – К.: НАУ, 2005. – Вип. 3 (14). – С. 9–14.

Подано до редакції 22.10.2010