

I. СЕРГІЄНКО, В. КОВАЛЬ

СКІТ — УКРАЇНСЬКИЙ СУПЕРКОМП'ЮТЕРНИЙ ПРОЕКТ

Сьогодні в Україні функціонує близько двох мільйонів персональних ЕОМ. Багато з них (майже 25 – 30%) — це сучасні Pentium'и з потужними офісними пакетами, зручними для різної діяльності, зокрема наукової. На жаль, є ще багато (35 – 40%) застарілих машин.

Проте персональні ЕОМ імпортного виробництва аж ніяк не можуть забезпечити розв'язання багатьох найважливіших задач економіки, техніки, національної безпеки нашої країни. Велика і надвелика їх обчислювальна розмірність, необхідність обробки з величезною швидкістю потужних інформаційних масивів роблять практично марними спроби розв'язання зазначених задач без використання сучасних високопродуктивних ЕОМ.

До останнього часу таких комп'ютерів в Україні практично не було. Це могло спричинити виникнення загрозливих ситуацій у різних галузях життєдіяльності країни, призвести до втрати провідних позицій у науці, загальмувати створення наукоємної продукції, технологій проектування та моделювання складних об'єктів і процесів тощо.

А придбати імпортні суперкомп'ютери завадила низка обставин: ембарго на постачання таких ЕОМ, їхня величезна вартість, практична неможливість їх модернізації, вимоги контролю використання комп'ютерів для тих чи інших цілей тощо.

Нині в Інституті кібернетики імені В.М. Глушкова НАН України створені та введені у дослідну експлуатацію дві високоефективні обчислювальні системи — суперкомп'ютери із кластерною архітектурою СКІТ-1 і СКІТ-2 на базі сучасних мікропроцесорів фірми Intel. За своїми характеристиками вони не поступаються світовим аналогам, а за можливістю ефективної інтелектуальної обробки великих обсягів знань і даних у ряді випадків істотно переважають сучасні зарубіжні зразки комп'ютерної техніки.

© СЕРГІЄНКО Іван Васильович. Академік НАН України. Директор Інституту кібернетики ім. В.М. Глушкова НАН України.

КОВАЛЬ Валерій Миколайович. Доктор технічних наук. Завідувач відділу цифрових математичних машин і систем того ж інституту (Київ). 2005.

Ці суперкомп'ютерні системи орієнтовані на розв'язання найважливіших для України задач економіки, техніки, екології, обороноздатності та безпеки, які мають велику й надвелику обчислювальну розмірність і потребують швидкісної обробки значних масивів даних [1].

Передбачено, що на базі зазначених суперкомп'ютерних систем буде створено потужний обчислювальний ресурс, надійно пов'язаний із користувачами різних організацій та регіонів країни, насамперед — з установами НАН України.

Зокрема, розроблені суперкомп'ютерні системи орієнтовані на розв'язання пріоритетних прикладних задач у таких галузях науки, як молекулярна біологія, генетика, матеріалознавство, фізика твердого тіла, ядерна фізика, фізика напівпровідників, астрономія, геологія тощо.

ІДЕОЛОГІЯ РОЗРОБОК

У процесі проектування суперкомп'ютерних систем виникає значний перелік питань, розв'язання яких потребує проведення низки експериментів, зокрема з вибору ефективної архітектури комплексу, апаратної платформи для побудови обчис-

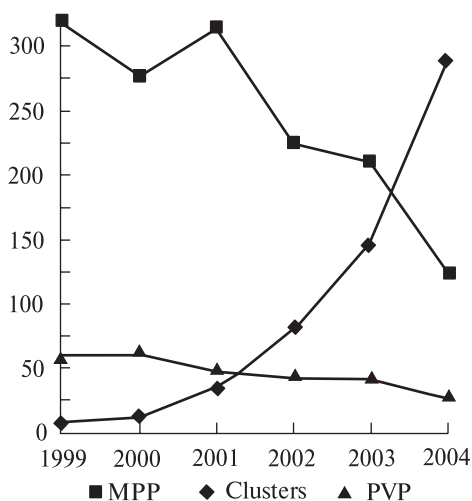


Рис. 1. Тенденції розвитку суперкомп'ютерів у світі

лювальних вузлів, комунікаційних інтерфейсів для вузлів зв'язку і т.п. [2].

При виборі архітектури комплексів ми спиралися на аналіз тенденцій розвитку суперкомп'ютерів у світі. Сьогодні можна виокремити п'ять найпоширеніших класів багатопроцесорних суперкомп'ютерних архітектур для паралельних обчислень:

- симетричні мультипроцесорні системи (SMP — Symmetric Multiprocessor Systems);
- системи з масовим паралелізмом (MPP — Massively Parallel Processing);
- системи з неоднорідним доступом до пам'яті і забезпеченням когерентності кеш-пам'яті (ccNUMA — cache coherent Non Uniform Memory Access);
- паралельні векторні системи (PVP — Parallel Vector Processing);
- кластерні системи (CS — Cluster Systems).

Для аналізу ефективності подібних систем у світі використовується перелік найбільш продуктивних ЕОМ — список top500 [3]. Фахівці часто сперечаються з приводу того, чи відображає top500 усі тенденції в архітектурі суперкомп'ютерних систем, але це поки єдиний реальний опис, що поєднує практично всі надпотужні системи.

Аналіз останньої редакції переліку top500 (березень 2005 р.) переконливо свідчить, що кластерні архітектури стали домінуючими у світі.

Кластерна обчислювальна система — це набір стандартних програмно-апаратних компонентів, об'єднаних для розв'язання загальних задач. У кластері як процесорні елементи використовують стандартні однопроцесорні або симетричні мультипроцесорні системи SMP, а як міжпроцесорні швидкодіючі комунікації у вузлах кластерів — стандартні інтерфейси (SCI, QsNet, Infiniband, Myrinet, Gigabit Ethernet і деякі інші).

Створення суперкомп'ютерних систем із кластерною архітектурою сьогодні є од-

ним із найперспективніших світових напрямів розробки надпотужних обчислювальних комплексів — їхня кількість постійно зростає, обсяги фінансування на їх розробку збільшуються.

На рис.1 відображено світові тенденції розвитку суперкомп'ютерів для мультипроцесорних систем типу **MPP**, **PVP** і кластерних. Як показує аналіз переліку найпродуктивніших суперкомп'ютерів, кластери посідають у ньому домінуюче місце — 297 систем із 500, а кількість систем типу **MPP** і **PVP** істотно скорочується.

ВИБІР АПАРАТНОЇ ПЛАТФОРМИ ТА КОМУНІКАЦІЙНИХ ІНТЕРФЕЙСІВ

Ефективну апаратну платформу для побудови обчислювальних вузлів ми вибирали, ґрунтуючись на аналізі співвідношення коефіцієнта ціна/швидкодія. Відповідний аналіз провели на основі тестів із сімейства Specfp, що найчастіше використовуються для оцінки швидкодії обчислень плаваючою крапкою, і цін на ці системи на ринку України (рис. 2).

Ціновий аналіз здійснено з урахуванням вартості основних відзначених компонентів при побудові систем на обраних архітектурах, з огляду на вартість комунікаційних інтерфейсів і монтажних систем із розрахунку очікуваної обчислювальної потужності 300GFlops на тестах Linpack і, відповідно, 120 000 SpecFP. Проаналізувавши відображені на рис. 2 порівняльні величини досліджуваного коефіцієнта, ми вибрали для кластера СКІТ-1 платформу Intel Xeon 2,7, а для кластера СКІТ-2 — платформу Itanium-2. Окрім того, ця платформа для нас є ліпшою, адже вона має велику розрядність обчислень (64) із плаваючою крапкою.

Оскільки планується значне розширення розроблених кластерів СКІТ-1 і СКІТ-2, то перевага, яку ми віддали мікропроцесорам компанії Intel, цілком зрозуміла, тим паче,

Ціна змінної частини системи при швидкодії 120000 SPECfp2000 (дані за 2004 р.)

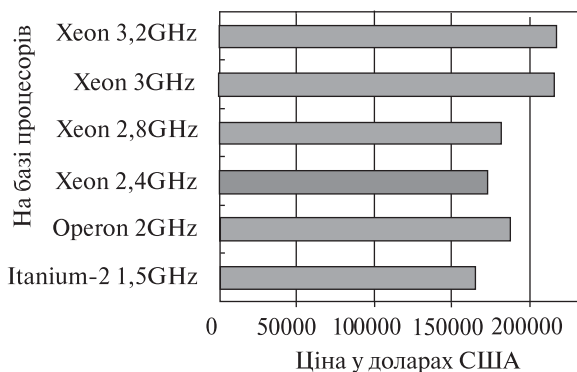


Рис. 2. Співвідношення ціна/швидкодія для платформ на різних процесорах

що у 2005 р. передбачається істотне підвищення їхніх характеристик — два ядра на одному кристалі, шина PCI Express, зниження енергоспоживання і т.п.

Проектування і вибір комунікаційних інтерфейсів для зв'язку вузлів здійснювали шляхом моделювання й аналізу різних видів цих інтерфейсів на задачах із пакета Gromacs (задачі молекулярної динаміки) [4]. Нас цікавили дві основні характеристики — латентність (яка затримка у процесі передачі пакета мінімальної величини) і пропускна здатність (швидкість передачі потоку даних). На рис. 3 і 4 наведено ці характеристики для різних типів комунікаційних інтерфейсів.

Аналіз зазначених інтерфейсів за їхніми характеристиками, а також за критерієм «ціна/швидкодія» спонукає до висновку про перевагу у використанні інтерфейсу Scalable Coherent Interface (SCI) для кластерних систем із кількістю вузлів менше 256. SCI дає змогу будувати тороїдальні топології, що дозволило за досить низьких витрат запропонувати кілька варіантів двовимірного тору з невеликим діаметром і з пропускною здатністю значно більшою, ніж за умови з'єднання через комутатор, оскільки при цьому забезпечувалася можливість передачі даних від одного вузла по двох каналах одночасно [5].

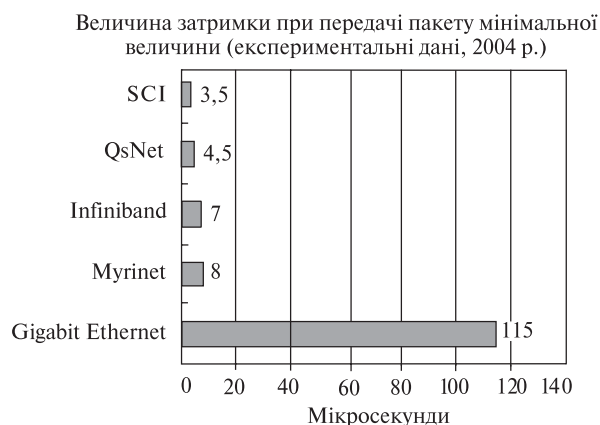


Рис. 3. Латентність комунікаційних інтерфейсів

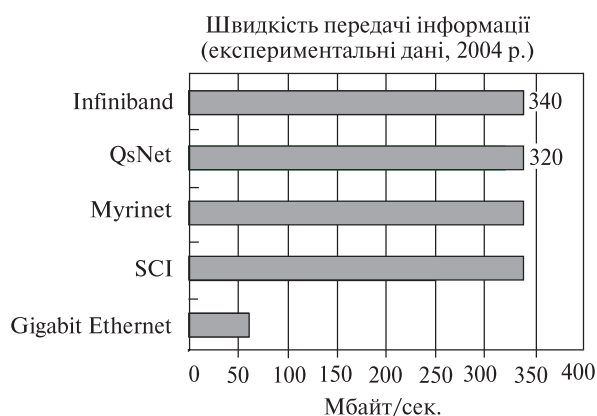


Рис. 4. Пропускна здатність комунікаційних інтерфейсів

Відомо, що **продуктивність** й **інтелектуальність** — найважливіші фактори, які визначають розвиток сучасних високопродуктивних комп'ютерів. Як уже вказувалося, **перший фактор** зумовив створення паралельних архітектур комп'ютерів, раціональним базисом яких є універсальні мікропроцесори, об'єднані у кластерні архітектури.

Другий фактор конкретизується за умови використання введеного В.М. Глушковым поняття «*машинний інтелект*», яким визначено і структуровано часто вживаний вислів «*внутрішній інтелект ЕОМ*», а інтелектуалізація ЕОМ розглядається як підвищення рівня машинного інтелекту.

У цьому напрямі в Інституті кібернетики імені В.М. Глушкова (ІК) НАН України впродовж останніх 5 — 6 років розробляються кластерні архітектури, орієнтовані на використання знань і названі **інтелектуальними розв'язувальними машинами (ІРМ)**. ІРМ реалізують паралельні мови високого та надвисокого рівнів і забезпечують **ефективну роботу зі складними структурами даних і знань великого обсягу** з використанням як **традиційно-обчислювальних задач** (математичної фізики, моделювання складних систем і об'єктів, обробки траєкторних вимірів тощо), **так і задач штучного інтелекту** (обробки і представлення знань, розпізнавання образів, діагностики, прогнозування і т.п.) [6].

Знання і складні структури даних у розроблених кластерах представлені у вигляді **орієнтованих графів** довільної складності — *дерев, семантичних мереж, що розростаються і змінюються у часі, тощо*. У комп'ютерах класу ІРМ можна будувати графи з багатьма мільйонами вершин і представляти знання з різних галузей.

Окрім того, орієнтовані на знання архітектури розроблених кластерів **добре інтегруються з архітектурами розподілених баз даних й істотно прискорюють процеси пошуку та обробки інформації**.

АПАРАТНО-ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ СКІТів

Сьогодні введені в дослідну експлуатацію суперкомп'ютери для інформаційних технологій (СКІТ). Саме інтелектуальна складова таких кластерів разом із розподіленими базами даних істотно відрізняє їх від того, що розробляється в цьому напрямі [7].

« **Кластер СКІТ-1** — 32 процесори, 16 вузлів на основі мікропроцесорів Intel Xeon 2,67 ГГц (1ГГц — 1×10^9 Гц) із розрядністю 32 біта і можливістю виконувати обчислення з 64- і 128-бітовою інформацією, піковою продуктивністю не менше 170 Гфлопс (1Гфлопс = 1 мільярд

операцій із плаваючою крапкою на секунду), яка може підвищуватися до 0,5–1 Тфлопс.

- **Кластер СКІТ-2** — 64 процесори, 32 вузли на основі мікропроцесорів Intel Itanium 2, 1,4 ГГц, із розрядністю 64 біта і можливістю виконувати обчислення з 128- і 256-бітовою інформацією, піковою продуктивністю не менше 270–300 Гфлопс, яка може зростати до 2–2,5 Тфлопс (1Тфлопс = 1 трильйон операцій із плаваючою крапкою на секунду), системою збереження даних обсягом 1 Тбайт (1 Тбайта = 1 трильйон байт) і можливістю нарощувати обсяг до 10–15 Тбайт.

У структурі кластерів використовуються три раніше згадані комунікаційні мережі — SCI, Gigabit Ethernet, Fast Ethernet.

Кожний із кластерів СКІТ-1 і СКІТ-2 (вузли і керуюча ЕОМ) є масивом обчислювальних модулів, пов'язаних між собою трьома локальними обчислювальними мережами (ЛОМ) — високошвидкісною мережею SCI (пропускна здатність каналу — 2000 Мбіт/с), мережею Gigabit Ethernet (пропускна здатність каналу — до 1000 Мбіт/с) і мережею Fast Ethernet (пропускна здатність каналу — до 100 Мбіт/с).

Мережа SCI забезпечує високошвидкісний обмін між вузлами у процесі обчислень. Топологія зв'язків вузлів кластера — це двовимірний тор (16-вузловий кластер — 2×8 чи 4×4 вузлів, 32-вузловий кластер — 2×16 чи 4×8). При обміні даними між двома вузлами у мережі SCI із використанням протоколів MPI може досягти пропускної здатності на рівні 300–350 Мбайт/с. Мережа Gigabit Ethernet (топологія Ethernet) виконує функцію з'єднання всіх вузлів кластера з керуючою ЕОМ і файлом-сервером. Мережа Fast Ethernet призначена для початкового завантаження програм і даних у вузли, а також для передачі службової інформації про перебіг обчислювального процесу.

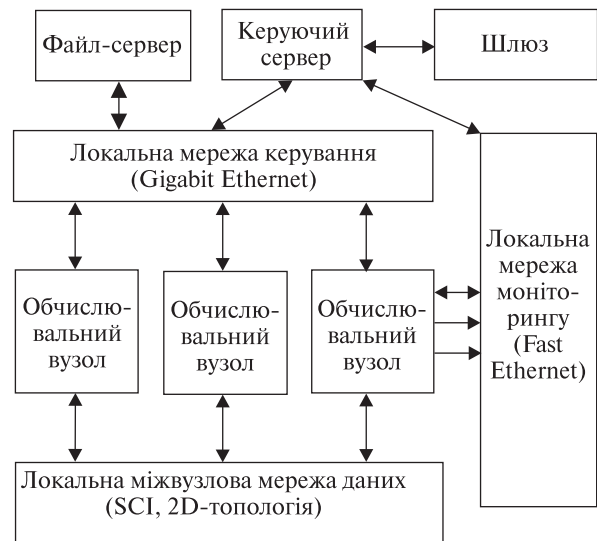


Рис. 5. Загальна структура кластерів СКІТ-1 і СКІТ-2

Продуктивність систем СКІТ-1 і СКІТ-2 (64-розрядні обчислення)

Найменування	Характеристика	
	СКІТ-1	СКІТ-2
Пікова продуктивність на 1 процесор (Gigahertz = 1×10^9 herz)	2,67	1,4
Цілочисельних операцій на сек. (IPS)	$1,34 \cdot 10^9$	$5,6 \cdot 10^9$
Операції з плаваючою крапкою на сек. (FLOPS)	$5,34 \cdot 10^9$	$5,6 \cdot 10^9$
Загальна пікова продуктивність		
Цілочисельних операцій на сек. (IPS)	$4,3 \cdot 10^{10}$	$3,58 \cdot 10^{11}$
Операції з плаваючою крапкою на сек. (FLOPS)	$1,7 \cdot 10^{11}$	$3,58 \cdot 10^{11}$
Швидкість системної шини (Gigabate/Sec)	4,2	6,4
Реальна продуктивність кластера на тесті Linpack (FLOPS)	$1,125 \cdot 10^{11}$	$2,8 \cdot 10^{11}$

Загальну структуру і продуктивність кластерів СКІТ-1 і СКІТ-2 відображають рис. 5 і таблиця.

Високий рівень машинного інтелекту цих систем дає змогу використовувати кластери для широкого кола традиційних обчислювальних задач і задач слабоформалізованих галузей (прийняття рішень за умов невизначеності, класифікація і формалізація понять, планування дій і пошук закономірностей у великих масивах даних, правдоподібних міркувань на основі автоматизації дедуктивних та індуктивних висновків тощо).

Створенню кластерів СКІТ-1 і СКІТ-2 та введенню їх у дослідну експлуатацію сприяла тісна співпраця Інституту кібернетики ім. В.М. Глушкова НАН України з фірмою «Юстар» (м. Київ) і компанією Intel, які забезпечили технічну і консультативну підтримку у використанні обладнання з процесорами Intel Xeon та Intel Itanium2.

ЗАГАЛЬНОСИСТЕМНЕ ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ

Компоненти загальносистемного програмного забезпечення підтримують усі етапи розробки паралельних програм користувачів, а також забезпечують безпосереднє виконання процесів змістовної обробки на розв'язувальному полі. Вони функціонують на всіх вузлах і на керуючій ЕОМ. Операційні системи кластерів – ALT Linux для СКІТ-1 та Red Hat Enterprise Linux AS – для СКІТ-2, а загальне середовище паралельного програмування – модель передачі повідомлень MPI/SCI (Message Passing Interface). Крім того, загальносистемне програмне забезпечення включає оптимізовані компілятори мов паралельного програмування C, C++, Фортран та інші.

ПРИКЛАДНЕ ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ

Розроблені суперкомп'ютери СКІТ-1 і СКІТ-2 дають змогу розв'язувати принципово нові надскладні задачі великої роз-

мірності у сферах науки, економіки, екології, сільського господарства, техніки, безпеки, в космічній та інших галузях. Уже сьогодні на сімействі кластерних суперЕОМ реалізовано низку прикладних пакетів для створення інформаційних технологій розв'язання важливих класів задач практичного застосування. Скорочення часу постановки, дослідження та розв'язання задач досягається не тільки за рахунок організації паралельних обчислень, а й завдяки інтелектуальному програмуванню, що враховує адаптацію алгоритму і синтезовані програми щодо властивостей задачі, беручи до уваги технічні і математичні можливості суперкомп'ютера.

Зокрема, в Інституті кібернетики імені В.М. Глушкова, Інституті космічних досліджень, Інституті програмних систем НАН України розроблено програмні пакети, які реалізовані на кластерах СКІТ-1 і СКІТ-2. Охарактеризуємо створені інформаційні технології, клас задач і сфери реалізації таких технологій.

1. Розв'язання проблем екології ґрунтів за великих обсягів повнозв'язаних даних. Добування корисних копалин, створення підземних комунікацій та споруд різноманітного призначення, забудівля схилів і рівнинних територій висотними будинками, замулення малих річок, порушення дренажних систем тощо спричиняють зникнення водоносних горизонтів, підвищення рівнів ґрунтових вод, зсуви, просідання територій, забруднення водоносних горизонтів артезіанських свердловин. Ці негативні явища особливо характерні для Донецького регіону, курортної зони Трускавця, Дніпропетровська, Києва, Одеської, Херсонської, Запорізької, Закарпатської областей.

Створена технологія дозволяє оцінювати наслідки різноманітної діяльності, прогнозувати розвиток процесів, притаманних ґрунтовим середовищам (підтоплення, зсуви, просідання ґрунтів тощо); визначати ефек-

тивність різноманітних інженерних засобів запобігання негативним явищам у ґрунтових середовищах.

2. Обробка даних сейсмічної розвідки земної кори. Технологію пошуку та видобування корисних копалин створено у співпраці з Українським державним геологорозвідувальним інститутом та Інститутом геофізики ім. С.І. Субботіна НАН України. Впроваджена вона в усіх геофізичних експедиціях, у багатьох організаціях нафтогазового, вугільного та рудного секторів економіки України.

Обробка й інтерпретація геофізичних даних потребують значного обсягу розрахунків (від кількох днів до кількох тижнів безперервно), що безпосередньо впливає на якість прийняття рішень за обмеженого часу. Часовий ліміт також перешкоджає поширенню складних, але точніших методів розрахунків. Враховуючи вартість буріння однієї свердловини (близько 1 млн доларів США), переведення таких розрахунків із ПК на потужні кластери дасть значний економічний ефект. Реалізовано паралельні розрахунки важливого класу задач; час обчислень скорочено майже пропорційно кількості вузлів кластера, наприклад у 32 рази. До того ж зменшився час очікування результатів, підвищилися якість і точність геологічних досліджень, оскільки тепер можна будувати деталізованіші моделі протягом обмеженого часу, відведеного на обробку сейсмічного матеріалу.

3. Аналіз та прогнозування природних і техногенних процесів в атмосфері. Ця інформаційна технологія призначена для чисельного розв'язання прогностичних рівнянь конкретних моделей циркуляції атмосфери, моделювання і прогнозування погодних умов в Україні з різною завчасністю та розв'язання еволюційних задач переносу і дифузії в атмосфері пилогазової суміші, викинутої промисловими підприємствами за проектного режиму та під час аварій. Роз-

роблено методи параметризації процесів підмержевих масштабів і моделювання мезо- та макромасштабних ефектів розсіювання в атмосфері газів, аерозолів і твердих домішок від постійно діючих чи миттєвих промислових джерел за метеорологічних умов з урахуванням рельєфу підстилаючої поверхні і стратифікації атмосфери. У нинішній версії пакета реалізовано інтерполяцію метеорологічних величин з макромасштабної мережі у мезомасштабну.

4. Кластерні обчислення задач обробки супутникової інформації. Попередня обробка космічних знімків «Метеосата» і побудови маски хмарності важлива для короткочасного прогнозування хмарності в інтересах авіакомпаній, передбачення розвитку надзвичайних ситуацій метеорологічного характеру, планування експериментів на борту «Січ-1М». Як геостаціонарний метеорологічний космічний апарат, супутник «Метеосат-7» постачає знімки розміром 10 Мбайт щопівгодини («Метеосат» 2-го покоління — кожні 15 хвилин, обсяг знімка — 100 Мбайт). Розв'язання задачі побудови маски хмарності з використанням марковських полів полягає у визначенні апостеріорної ймовірності сегментації знімка. Алгоритм надзвичайно працездатний, потрібна оперативна обробка, тому доцільна кластерна система. Проведені експерименти свідчать про восьмиразове підвищення ефективності розв'язання задачі на кластері з 32-ма процесорами. Цього досить для задач короткострокового прогнозування у реальному часі.

5. Моделювання структурно-технологічних змін у перехідній економіці. Розроблені інформаційні технології і методи розв'язання широкого кола задач економічного, фінансового аналізу на макrorівні та визначення головних економічних показників ВВП:

♦ інформаційна технологія середньострокового прогнозування основних макроеконо-

мічних показників застосовується у стратегічному плануванні й управлінні виробничою та комерційною діяльністю, розподілі обмежених ресурсів і капіталів, проектуванні та розробці можливих сценаріїв дій під час розв'язання господарських та соціально-політичних проблем, для аналізу тенденцій розвитку економіки України в середньостроковій перспективі, а також для оцінки ефективності та можливих наслідків вибору тих чи інших стратегій розвитку на основі моделей і методів середньострокового прогнозування головних макроекономічних показників;

- технологія моделювання розвитку економіки України на макrorівні за різними можливими сценаріями, що дає змогу визначити наслідки тих чи інших економічних або соціальних рішень. Технологія базується на використанні комплексу взаємопов'язаних макроекономічних моделей. Це дозволяє враховувати вплив на економіку України таких чинників, як зміна величини імпорту енергоносіїв, частки інвестицій у реальному ВВП, у структурі витрат зведеного бюджету, в рівні основних відсоткових ставок, інфляції, курсу гривні до долара США. Технологія може використовуватися у роботі різних міністерств України;
- економіко-математичні методи для середньо- та довгострокового моделювання економіки України;
- інформаційна технологія для комплексного прогнозування головних макроекономічних показників перехідної економіки країни, що дає змогу отримувати оцінки економічних або політичних рішень. Система має багаторівневий захист від несанкціонованого доступу до різних груп даних й уможливорює різноманітні розрахунки прогнозів основних макроекономічних показників для фіксованої моделі;
- нова методика аналізу стану економічної безпеки держави та конструювання

інтегрального індексу безпеки на базі оригінальних процедур нормування економічних показників з урахуванням їхніх порогових та граничних значень, порушення яких істотно впливає на сталий розвиток соціально-економічних процесів у державі. Методику детально опрацьовано на прикладі безпеки інвестиційної сфери України;

- оригінальна інформаційна технологія індуктивного моделювання, заснована на методі групового врахування аргументів для моніторингу складних економічних та екологічних процесів. Технологія базується на методах пошуку прихованих закономірностей у статистичних даних для виявлення потенційно несприятливих тенденцій розвитку соціально-економічних й екологічних процесів. Розробка призначена для підвищення ефективності процесів прийняття поточних управлінських рішень завдяки оперативному виявленню можливих негативних наслідків;
- нова інформаційна технологія визначення аномальних тенденцій у фінансово-економічній діяльності підприємств. Технологія може використовуватися для з'ясування тенденцій до банкрутства банків і підприємств, а також з метою виявлення ознак тіньової діяльності підприємств на основі обробки даних статистичної звітності за економічними показниками. Перевірка на реальних даних показала високу ефективність реалізованого підходу (90 – 92%).

6. Задачі керування рухомими об'єктами, зокрема задача динамічного комівояжера в режимі реального часу. Технологія призначена для протиракетної оборони, охорони об'єктів, запобігання терористичним актам, зіткненню (скажімо, літаків на злітній смузі) тощо. Класичний варіант задачі статичного комівояжера розглянуто як тестовий приклад, а динамічний варіант з рухо-

ними об'єктами обчислюється у режимі реального часу. Реалізовані алгоритми для статичного комівояжера – наближені (знайдений розв'язок не завжди оптимальний), а для динамічного – точні (розв'язок оптимальний).

7. Оптимальний вибір орбіт і структури космічних систем зондування Землі та їх раціональна експлуатація.

Розв'язуються дві задачі:

- оптимізація орбіт і структури супутникових угруповань космічних апаратів;
- раціональне оперативне планування завантаження апаратних комплексів космічних апаратів, тобто складання розкладів спостережень вказаних районів території.

8. Пошук оптимального розміщення сервісних центрів. Сервісні центри (краєниці, склади, аварійні служби, оператори мобільного зв'язку, станції «швидкої допомоги», заправні станції та інші постачальники послуг) обслуговують споживачів на певній території. Споживачі послуг нерівномірно розміщені на площині або вздовж певної лінії. Потреба клієнта у послугі з географічної точки може бути випадковою; з кількох сервісних центрів клієнт вибирає найближчий, можливо, із заданою ймовірністю. Задача полягає у розташуванні сервісних центрів так, щоб сукупні очікувані витрати на обслуговування всіх споживачів послуг були мінімальними. Пакет на кластері у десятки разів скоротив час виконання цього класу оптимізаційних багатоекстремальних задач стохастичного програмування.

9. Побудова перешкодозахищеного коду максимального об'єму методом гілок і границь. *NP*-складну задачу знаходження максимальної незалежної множини розв'язано для побудови кодів, що забезпечують надійну передачу інформації. Потребу у розпаралелюванні зумовлено великою розмірністю задачі. З літератури відомо про розв'язання

задачі на графі 1tc1024 за 200 годин, а час розв'язання за нашою технологією різко скорочений – він триває близько 36 хвилин.

10. Оцінка і класифікація альтернативних рішень щодо поточного стану складної системи на базі нечітких даних та знань недетермінованої природи. Сфера застосування технології – економічне оцінювання інвестиційних проектів на транспорті; виявлення проривів нафтопроводів низького тиску на основі нечітких даних щодо забруднення навколишнього середовища; прогнозування ушкодження посівів зернових; з'ясування слабких місць проектів і їхніх причин; діагностика та лікування захворювань; оцінювання надійності функціонування систем тощо.

Реалізовані алгоритми байєсівської класифікації для високопродуктивних паралельних систем на базі даних та знань у точковому й інтервальному поданні дали змогу:

- значно скоротити час аналізу інформації (як апіорної, так і апостеріорної) у процесі розв'язання задачі в інтервальному інформаційному просторі з досягненням прийнятної точності результатів;
- розв'язувати задачі класифікації станів досліджуваних систем у нечіткому інформаційному просторі з прийнятною точністю результатів, що неможливо на ПК.

11. Моделювання прямих та обернених задач динаміки систем із розподіленими параметрами. Розглянуто задачі проектування природоохоронних заходів й управління якістю природного середовища, які розв'язують у випереджальному режимі під час планування великих проектів, пов'язаних з впливом на довкілля. При цьому розв'язано обернену задачу: визначення межі припустимих антропогенних навантажень за критеріями екологічної безпеки і стабільності. Алгоритмічна та програмна

реалізація на кластерній системі розв'язання таких задач застосовує розроблений та обґрунтований математичний апарат аналітичного моделювання динаміки систем з розподіленими параметрами.

12. Кластеризація даних за допомогою генетичних алгоритмів. Це конче потрібно для розв'язання задачі пошуку закономірностей щодо закупівель продукції у маркетингових дослідженнях споживчого ринку і планування виваженої торговельної політики. Розробку цього пакета інтелектуального аналізу даних за допомогою Data Mining зумовлено майбутнім великим попитом на такі системи в Україні. Застосування генетичних алгоритмів дає змогу використовувати максимально повну сферу пошуку і максимальну універсальність формулювання задачі, а проблему повільної швидкості збіжності генетичних алгоритмів розв'язано додаванням до класичної схеми генетичних алгоритмів оператора розвитку.

Обробка даних реального комерційного підприємства (понад 5000 фактів продажу продукції) на ПК (2 GHz) зайняла 4 години. Оскільки для великих підприємств обсяг даних перевищує мільйони одиниць, розв'язання цієї задачі на ПК неможливе, а кластер пришвидшив обчислення у десятки разів.

ПЕРСПЕКТИВИ

Виконані дослідження зі створення кластерів СКІТ-1 і СКІТ-2 ми розглядаємо лише як перший етап розвитку вітчизняних суперкомп'ютерів. Вони істотно розширили фронт робіт з інтелектуалізації інформаційних технологій, але широка інформатизація держави потребує значно більших зусиль. Тому наступний етап дослідження найдоцільніше присвятити підвищенню характеристик розроблених суперкомп'ютерів та їх програмного забезпечення, що відповідає сучасним тенденціям у світовому обчислювальному машинобу-

дуванні. Реалізація цього завдання дасть змогу розв'язувати важливі державні проблеми забезпечення безпеки країни, керування процесами економічного розвитку, а також складні науково-технічні задачі, які мають велику і надвелику обчислювальну розмірність і потребують швидкої обробки потужних масивів даних та знань.

Зокрема, суперкомп'ютерні системи слід орієнтувати передусім на розв'язання актуальних прикладних задач у таких галузях науки, як молекулярна біологія, генетика, матеріалознавство, фізика твердого тіла, ядерна фізика, фізика напівпровідників, астрономія, геологія тощо.

Передбачається, що на базі нових суперкомп'ютерних систем **буде створено потужний об'єднаний обчислювальний ресурс**, надійно пов'язаний із користувачами різних організацій та регіонів України, у тому числі з установами НАН України. Цього ресурсу буде цілком достатньо для забезпечення обчислювальними потужностями суперкомп'ютерів практично всіх можливих користувачів нашої країни.

Найперше завдання на даному шляху — подальший розвиток інтелектуальних суперкомп'ютерів класу СКІТ-1, СКІТ-2 для масового використання у різних організаціях та регіонах України, нових інформаційних технологій розв'язання складних задач у галузі науки і техніки, ряду типових проблемно орієнтованих комплексів на основі архітектур із засобами інтелектуального програмування, а також мережевих конфігурацій, які забезпечують високу і надвисоку продуктивність систем. Науковцями нашого інституту виконано розробки таких архітектур у широкому діапазоні продуктивності, з оптимізацією завантаження, розподілом функцій з підготовки задач та їх розв'язання, організацією розподіленого ієрархічного керування. Ноу-хау в галузі новітніх фізичних постановок задач і

методів їх розв'язання забезпечать високу ефективність обробки інформації та використання ЕОМ.

На першому етапі (2005 — 2006 роки) необхідно створити нижні та середні моделі ряду СКІТ (2 — 5 Тфлопс, 3 — 5 штук) на основі сучасних високонадійних комплектуючих; на другому етапі (2007 — 2008 роки) — потужні суперкомп'ютери (15 — 20 Тфлопс, 2 — 3 штуки) та суперкомп'ютерні центри й інтелектуальні мережі масового інформаційного сервісу.

1. *Сергієнко І.В.* Информатика та комп'ютерна техніка. — К.: Наук. думка, 2004. — 432 с.
2. *Коваль В.Н., Сав'як В.В., Сергієнко І.В.* Тенденции развития современных высокопроизводительных систем // УсиМ. — 2004. — № 6. — С. 31—43.
3. www.top500.org
4. www.gromacs.org
5. *Коваль В.Н., Сав'як В.В.* Мультипроцессорные кластерные системы: планирование и реализация // Искусственный интеллект. — 2004. — № 3. — С.117—126.
6. *Koval V., Bulavenko O., Rabinovich Z.* Parallel Architectures and Their Development on the Basis of Intelligent Solving Machines // Proc. Int. Conf. on Parallel Computing in Electrical Engineering. — Warsaw (Poland). — 2002. — P. 21—26.
7. SCIT — Ukrainian supercomputer project / *Koval V.N., Ryabchun S.G., Savyak V.V., Serghienko I.V., Yakuba A.A.* // Proc. International Conference KDS. — 2005. — Bulgaria, Varna, 2005.

І. Сергієнко, В. Коваль

СКІТ — УКРАЇНСЬКИЙ СУПЕРКОМП'ЮТЕРНИЙ ПРОЕКТ

Резюме

Детально охарактеризовано перші вітчизняні суперкомп'ютери із кластерною архітектурою — СКІТ-1 і СКІТ-2, які створені в Інституті кібернетики імені В.М. Глушкова НАН України і відповідають кращим світовим аналогам. Показано можливості цих систем у розв'язанні надскладних задач великої розмірності у сферах науки, економіки, екології, техніки, безпеки тощо. Реалізовані на кластерах СКІТ-1 і СКІТ-2 програмні пакети дали змогу створити ряд інформаційних технологій, що знайшли застосування у різних галузях економіки, у прогнозуванні природних і техногенних процесів. Окреслено шляхи подальшого розвитку інтелектуальних суперкомп'ютерів класу СКІТ-1 і СКІТ-2 для широкого використання у багатьох організаціях і в різних регіонах України.

І. Serhiyenko, V. Koval.

SCIT — UKRAINIAN SUPERCOMPUTER PROJECT

Summary

The detail characteristic of the first local supercomputers SCIT -1 and SCIT -2 with cluster architecture is presented. These computers were created in V.M. Glushkov Institute of Cybernetics of Ukrainian NAS and correspond to the best world counterparts. The system capacity to solve supercomplex tasks in such fields as science, economy, ecology, engineering, security is demonstrated. The software implemented on clusters SCIT-1 and SCIT-2 enabled development of a range of information technologies applied in various economics branches, in forecasting of natural and man-caused processes. The ways of further development of SCIT-1 and SCIT-2 class intellectual supercomputers for common use in many organizations and different regions of Ukraine are described.