

РОЗПОДІЛЕНЕ ПАРАЛЕЛЬНЕ МОДЕЛЮЮЧЕ СЕРЕДОВИЩЕ

В.А. Святний, О.М. Солонін, Д.В. Надєєв, І. Степанов
кафедра ЕОМ ДонДТУ, sviatnyi@cs.dgtu.donetsk.ua
К. Ротермель, (IPVR) М. Цайтц (ISR) Штуттгартський університет
zeitz@uni-stuttgart.de

Abstract

Distributed parallel modeling environment (DPME). In this article is described hardware and software tools of DPME₁ and methods for access to its resources.

1. Вступ

Паралельне моделююче середовище для динамічних систем з зосередженими (ДСЗП) та розподіленими параметрами (ДСРП) - це нова форма системної організації сучасних засобів моделювання [1,2]. Реалізація середовища проводиться в рамках наукового співробітництва факультету ОТІ ДонДТУ та інститутів IPVR та ISR Штутгартського університету. Паралельні обчислювальні ресурси середовища, робочі місця та засоби доступу розташовані в Штуттгарті (HLRS - надпотужний комп'ютерний центр) та Донецьку на відстані коло 3 тис. км. Це обумовлює функціонування ресурсів за принципами розподілених обчислювальних систем [3]. Розглянемо побудову розподіленого паралельного моделюючого середовища (РПМС) та методичку доступу до його ресурсів.

2. Структура технічних засобів РПМС

Загальна структурна схема розподіленого паралельного моделюючого середовища (РПМС) представлена на рис. 1. Основними ресурсами є паралельні обчислювальні системи Cray та Hitachi. Лабораторії факультету ОТІ мають комп'ютери класу Пентіум (робочі станції), що об'єднані локально мережею з виходом в INTERNET. Вони також об'єднані в MPI-кластери на базі операційної системи Windows NT. Intel Paragon - це паралельна обчислювальна машина класу MMD, містить 11 вузлів, з'єднаних між собою двомірною решіткою. Intel Paragon використовується як тестова паралельна машина.

Система Cray T3E-900/512 має такі технічні характеристики: обчислювальна потужність - 461 GFLOP/S; кількість вузлів - 512; загальний об'єм оперативної пам'яті - 64 GB DRAM; об'єм оперативної пам'яті на вузлі - 128 MB DRAM; загальний об'єм дискової пам'яті 507 GB. Як особливості системи можна зазначити наявність розподіленої пам'яті, здатність до розширюємості, обмін повідомленнями, високу пропускну спроможність мережі комутації вузлів та підсистеми вводу/виводу, низьку латентність.

Система Hitachi SR8000 має обчислювальну потужність 128 GFLOP/S, 16 вузлів з 128 процесорами, з швидкістю передачі даних від вузла до вузла 1 GB/S та 8GB оперативної пам'яті у кожному вузлі. До позитивних властивостей сис-

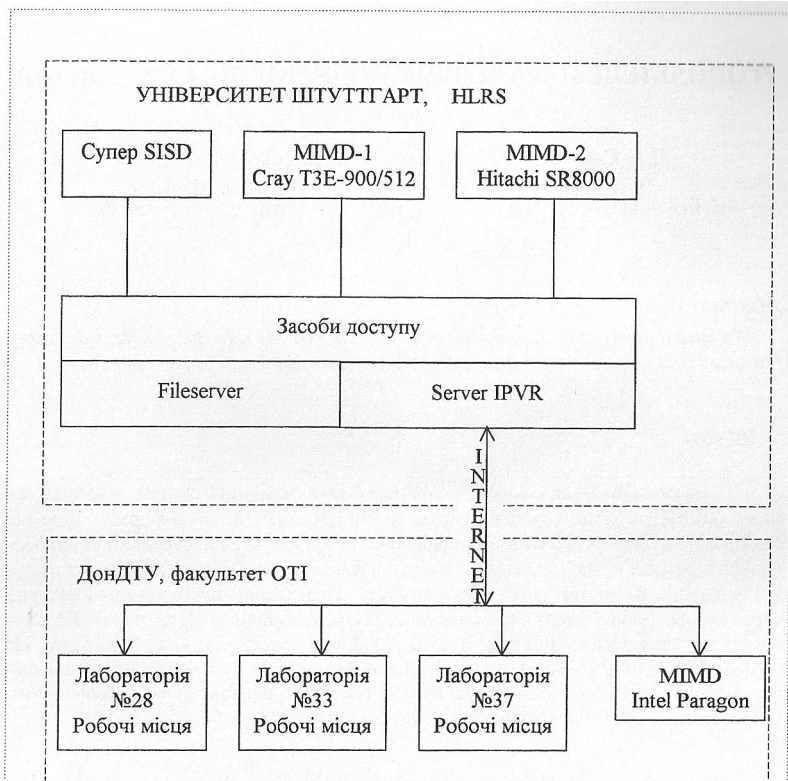


Рис. 1 Структура технічк засобів РПМС

теми треба віднести оперативну пам'ять с високою пропускнуо здібністю, векторне та скалярне програмування під SMP-кластер, Pseudo-Vector-Processing (PVP), наявність комерційних пакетів та систем COMPAS, MPI, OpenMP.

Файловий сервер SGI Power Challenge містить 2 центральних процесора Mips R4400 250 MHz, пам'ять 512 MB DRAM, дисковий накопичувач 1260 GB.

3. Структура системних програмних засобів

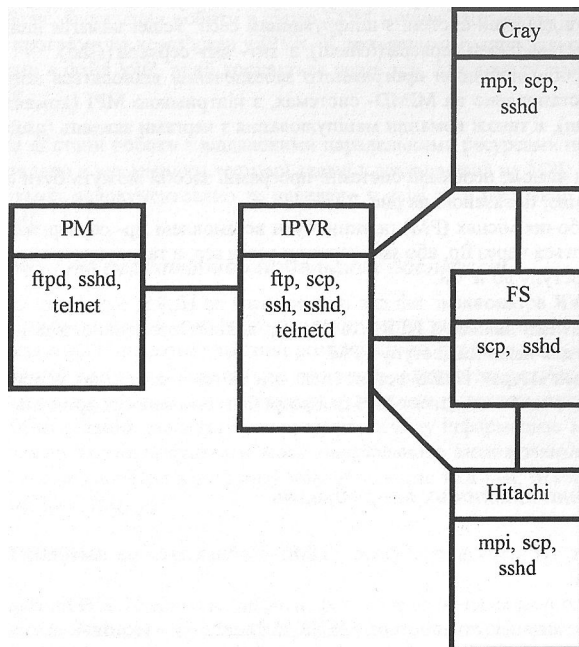
До системних програмних засобів відноситься усе програмне забезпечення, що підтримує функціонування компонентів структури РПМС.

Системні програмні засоби повинні:

дозволяти виконання операцій з файловими даними (копіювання PM->D?VR->HLRS, упорядкування файлів у домашніх каталогах та ін.)

- дозволяти створювати двоїчні виконувані модулі з програм, написаних на мові високого рівня;
виконувати процедуру авторизації при вході користувача в систему;
- забезпечувати взаємодію системних програм комп'ютерів з різною архітектурою.

Існує кілька засобів копіювання даних між віддаленими системами, але частіше використовуються ftp та scp/rscp [4]. При використанні ftp на комп'ютері встановлюються ftp-сервер (ftpd) та ftp-клієнт (ftp). Клієнтська програма ініціює зв'язок з сервером та отримує від нього дозвіл на роботу з файловими даними на сервері (читання/запис). Недолік підходу - необхідність знати команди ftp-сервера; щоб скопіювати файл на віддалений комп'ютер, часто потрібно ввести кілька довгих команд, особливо, якщо копіюється кілька файлів з багаторівневою



-Рис.2. Структура програмних засобів РПМС.

структурою каталогів. Крім того, на комп'ютер мало встановити клієнтську програму - треба ще займатися адмініструванням сервера. Позитивна сторона під-

ходу - повний контроль за виконанням операцій, можливість продовжити копіювання після аварійної зупинки.

Використання команд віддаленого копіювання `scp` або `sftp`, які базуються на віддаленому доступі до системи через `rsh` або `ssh`, відрізняється тим, що у `s-` формі виконується шифрування сесії, яке підвищує ступінь захисту при передачі даних. Для успішного використання команд необхідно, щоб на віддаленій системі було встановлено `ssh`- сервер (`sshd`). Простота та натуральність використання пов'язані тут з неможливістю продовжити операцію з місця аварійної зупинки.

Для операцій упорядкування даних використовуються класичні команди системи UNIX (`mkdir`, `cp`, `rm` і т.д.). Для виконання будь якої команди на віддаленій системі необхідно спочатку підключитись та ввійти до системи. Це можна здійснити використанням програмних засобів `telnet` або `ssh`. `Telnet` дозволяє виконати повний вхід у віддалену систему, як у локальну, а `ssh` запускає командну оболонку на віддаленій системі з шифруванням сесії. `Telnet` вимагає наявності на віддаленій системі `telnet`- сервера (`telnetd`), а `ssh` - `ssh`- сервера (`sshd`).

До засобів побудови програмного забезпечення відносяться компілятори C/C++, що встановлено на MBVTO- системах, з підтримкою MPI (команди `shpic`, `mpicc`, `mpirun`), а також команди маніпулювання з чергами завдань (`qsub`, `qstat` та ін.).

Таким чином, необхідні системні програмні засоби можуть бути представлені структурою, показаною на рис. 2.

На робочих місцях (PM) повинні бути встановлені `ftp`- сервер, якщо дані з `n>VR` копіюються через `ftp`, або `ssh`- сервер у разі з `scp`, а також програма `telnet` для отримання доступу до IPVR.

На `rPVR` встановлені: `ssh` для підключення до HLRS, `scp` та `ssh`- сервер для обміну файловими даними з HLRS та PM, `ftp` для альтернативного зв'язку з PM, `telnet`- сервер для надання доступу з PM.

На комп'ютерах HLRS встановлені `scp` та `ssh`- сервер для обміну файловими даними, а також комплект MPI для розробки паралельних програм.

4. Проведення тестових випробувань

Процес тестових випробувань РПМС розбивається на наступну послідовність операцій:

- 1) Підключення до комп'ютера у IPVR та копіювання тестових даних з PM.

Підключення до `rPVR` відбувається через `telnet` (рис.2.). Після входу до системи визначається похідна EP- адреса (команда `who`) та робиться перехід у каталог, де зберігаються тестові програми (`cd MPI_Tests`).

Для копіювання даних з PM використовується `ftp` (рис.2). Після входу на `ftp`- сервер PM робиться перехід у каталог, де знаходяться похідні файли (`cd incoming`), потім викликається командна оболонка для створення потрібної системи каталогів (команда `!`), далі виконується процес завантаження даних (`mget`).

Після закінчення роботи з ftp- сервером (виконання quit) тестові дані знаходяться у каталозі на комп'ютері EPVR.

2) Підключення до MMD-системи Cray у HLRS та копіювання даних з ffVR.

Підключення до Cray (рис.2) виконується через ssh, що запускається на комп'ютері у TPVR

Копіювання даних виконується через команду scp.

3) Компіляція та виконання програм на Cray.

Для компіляції програми необхідно зробити перехід у відповідний каталог (команда cd), відкомпілювати програму (команда mpicc) та запустити, вказавши потрібне число процесорів для виконання (mpirun з ключем pr).

4) Копіювання результатів з HLRS на IPVR та вихід з HLRS.

Після завершення роботи з супер-EOM можна скопіювати результати виконання програми на комп'ютер у IPVR. У попередньому прикладі результати були записані в файл log2, який копіюється через scp, для виходу треба ввести команду exit.

Всі ці стани роботи з віддаленими паралельними ресурсами перевірені експериментально з виконанням тестової задачі з документації на MPI. Протокол дій на всіх станах задокументовано та наведено у звіті та в інструкції для користувачів РПМС.

5. Висновки та перспективи подальших досліджень

Виконано перший етап робіт з компонування та експериментальної перевірки працездатності апаратної частини розподіленого паралельного моделюючого середовища. В результаті користувачі отримали можливість реалізовувати та досліджувати як паралельні моделі складних динамічних систем, так і підсистеми діалогу, балансування завантаження ресурсів, обміну інформацією моделюючого середовища, а також побудувати мову паралельного моделювання. Вирішення цих проблем заплановане в угоді про співробітництво між ISR E³VR та факультетом ОТІ на 2000-2003 рр.

6. Література

1. Святний В.А. Проблеми паралельного моделювання складних динамічних систем,- Наукові праці ДонДТУ, серія "Інформатика, кібернетика та обчислювальна техніка", випуск 6, Донецьк, 1999, с. 6-14.
2. Parallele Simulationstechnik ^eldman L.P., Lapko V.V., Svjatnyj V.A., Grilles E.D., Reuter A., Rothermel K., Zeitz M./ - Наукові праці ДонДТУ, серія "Проблеми моделювання...", вип. 10, Донецьк, 1999, с. 9-19
3. Tanenbaum A.S. Distributed Operating Systems. Practice hall, 1995.
4. <http://www.hlrs.de/flw-access> - HW&Access