Liet. matem. rink., T. 43, spec. nr., 277–281

Investigation of PC cluster productivity in quantum mechanical molecular computations

Egidijus PANKEVIČIUS, Rimantas-Jonas RAKAUSKAS (KA),
Juozas ŠULSKUS (KA, VU), Svajonė VOŠTERIENĖ (KA)
e-mail: egidijus.pankevicius@lrs.lt, juozas.sulskus@ff.vu.lt, svajone.vosteriene@tmk.lka.lt

1. Introduction

Clusters based on PCs running Linux have become the cheapest "supercomputers"in the academic and commercial field. Such cluster TAURAS was created in the Applied Sciences Department in Lithuanian Military Academy. The SCore cluster system software is built on top of Linux and Cluster TAURAS parameters are comparable with other ones located in academic areas. Cluster is registered in the top500 cluster list [1]. Created computational facilities were applied to investigate electronic structure and vibration spectra of trinitrotoluene molecule by means of nonempirical quantum mechanical computation methods.

The accurate nonempirical quantum chemical computations (molecular electronic structure and vibration spectra investigations) are essential in spectroscopy. The basis of these investigations is quantum mechanics, but in order to achieve quantitative results comparable with experimental we need supercomputer's power or parallel computers cluster. On the other hand the highly sophisticated quantum chemical computations have been become the routine tests of supercomputers performance. The tasks solved were used to compare the possibilities of SCore cluster system software with Beowulf [2]. During investigations the GAMESS quantum chemistry computer code was used [3].

2. Score cluster system software

The high performance-computing environment on the PC Cluster TAURAS cannot be realized without the SCore cluster system software on top of Linux. It consists of a communication facility called PM, MPI-1 implemented on PM called MPICHPM/ CLUMP, a global operating system called SCore-D, and a multi-threaded programming language called MPC++. To realize the high performance system using commodity hardware and software, the following key technologies have been employed:

- a user-level zero-copy message transfer mechanism between nodes and one copy message transfer mechanism within a node by a high performance communication facility called PM.

- a high performance MPI-1 implementation called MPICH-PM/CLUMP that integrates both zero-copy message transfer and message passing facilities in order to maximize performance.
- multi-user environment realized by an operating system daemon called SCore-D.

SCore cluster system is introduced to show that a compact and well maintainable PC cluster using commodity hardware can be built on top of Linux. Since the PM kernel-level driver's functionality is significant to realize high performance on top of Linux. PM provides a virtual network mechanism called PM channel [4]. The channel provides reliable data gram communication, instead of connection-oriented communication such as that of TCP/IP.

3. Quantum chemical computations

As was stated quantum chemical computations have been performed by means of ab initio quantum mechanical calculations using GAMESS computer code [5]. The basis of all such calculations is solving of multidimensional time independent Shrödinger equation

$$\hat{H}\Psi_i = E_i\Psi_i, \tag{1}$$

here $\hat{H}$ – Hamilton operator of investigated quantum particle system; Ψ_i – wave function of the system; E_i – allowable energy value. Hamiltonian operator and wave function of the system depend from the coordinates of all particles (electrons and nucleus). We use non-relativistic Hamilton operator for electronic part of the problem [4, 9]

$$\hat{H}_e = \sum_{p=1}^{N} h_p + \sum_{p>1}^{N} g_{pg}, \tag{2}$$

here

$$h_p = \frac{1}{2}\nabla_p^2 - \sum_{\mu=1}^{M} \frac{Z_\mu}{r_{p\mu}}, \tag{3}$$

is one electronic part of Hamiltonian operator; M – is the number of nucleus; Z_μ – the charge of μth nuclei; $r_{p\mu}$ – the distance between μ-th nuclei and pth electron;

$$g_{pq} = \frac{1}{r_{pq}} \tag{4}$$

is the two electronic part of Hamiltonian. We use the atomic units system in all formulae.

In one particle approximation the molecular wave function for closed shell electronic system (with odd number of electrons) is presented by Slatter determinant

$$\Psi = \frac{1}{\sqrt{N!}} \det\left|\Psi_1\bar{\Psi}_1\Psi_2\bar{\Psi}_2\ldots\Psi_{\frac{N}{2}}\bar{\Psi}_{\frac{N}{2}}\right|, \tag{5}$$

where N is the total number of electrons;

$$\Psi_i = \varphi_i \alpha, \bar{\Psi}_i = \varphi_i \beta, \tag{6}$$

ϕ_i and $\bar{\Psi}_i$ are one electron wave functions depending from ith electron coordinates: ϕ_i – spatial part of electron wave function; and β – spin orbitals.

Molecular orbitals (MO) ϕ_i are often expanded in the basis of Cartessian Gaussian atomic orbitales (AO):

$$\varphi_i = \sum_k, X_k T_{ki}, \tag{7}$$

$$X_{GO}(n,l,m) = N(\alpha,l,m,n) x^l y^m z^n e^{-\alpha r^2}, \tag{8}$$

or in the matrix form

$$\varphi = XT, \tag{9}$$

here $N(\alpha,l,m,n)$ – normalization coefficient; T – matrix with expansion coefficients for MO in the columns.

The first order variation of the full energy with additional requirement of orthogonality between MOs leads to matrix equation for MO coefficients matrix and one electronic energies matrix.

$$\hat{H^F} T = S T_\varepsilon, \tag{10}$$

where S is the matrix of overlap integrals between AO ($S_{ij} = \int X_i X_j \, \mathrm{d}_\tau$).

$$\hat{H^F} = \hat{h} + \sum_n (2\hat{J}_n - \hat{K}_n), \tag{11}$$

Coulomb and exchange operators are defined accordingly

$$\hat{J}_n \varphi_m(\mu) = \int \frac{|\varphi_n(\nu)|^2}{r_{\mu\nu}} \mathrm{d}\vec{r_\nu} \varphi_m(\mu), \tag{12}$$

$$\hat{K}_n \varphi_m(\mu) = \int \frac{\varphi_n^*(\nu)\varphi_m(\nu)}{r_{\mu\nu}} \mathrm{d}\vec{r_\nu} \varphi_m(\mu). \tag{13}$$

As operators $\hat{J}$ and $\hat{K}$ depend themselves from MO φ_i , (10) is usually solved by iterational self consistent field method (SCF). So all the problem is concerned with evaluation of great number of integrals in AO basis with (11) operator and the use of linear algebra matrix computational methods. One of the methods to go beyond the one electron approximation during investigation of molecular electronic structure is use of Moller–Plesset perturbation theory (zero order Hamiltonian $\hat{H}_0$ is the sum of the one-electron operators). Then for closed shell electronic systems the first non zero correction

to Hartfee-Fock energy equals:

$$E_0{}^2 = \sum_{a,b,r,s}^{\frac{N}{2}} \frac{\langle \varphi_a^* \varphi_b^* \mid \varphi_r \varphi_s \rangle \langle \varphi_r^* \varphi_s^* \mid \varphi_a \varphi_b \rangle}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s} - \sum_{a,b,r,s}^{\frac{N}{2}} \frac{\langle \varphi_a^* \varphi_b^* \mid \varphi_r \varphi_s \rangle \langle \varphi_r^* \varphi_s^* \mid \varphi_b \varphi_a \rangle}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s}, \tag{14}$$

where a, b, r and s each signify spatial MO. During MP2 computations transformation of many center integrals from AO to MO basis is necessary, what leads to extensive use of matrix algebra in computations and much greater use of intercomputer communications in the case of parallel computations.

4. Solution of the optimization problem and results of calculations

GAMESS should automatically build for parallel execution to run using the Distributed Data Interface over SHMEM, MPI-1, or TCP/IP sockets. We've recompiled GAMESS using SCore implementation of MP-1. Computations were performed in one electron Hartree–Fock (HF) approximation ((5), (10)) (Fig. 1) and with account of electron correlation (14) (Fig. 2). The dependence of solution CPU time from the number of nodes in the cluster were achieved using sockets libraries in Beowulf cluster and with MP-1 in SCore system. As an egzample we used earlier investigated trinitritoluene molecule [10]. The AO basis 6–31G was used. Fig. 1 shows almost the same speed up using sockets libraries and SCore. It can be explained taking into account the fact that in HF approximation network is not much loaded. Separate nodes just evaluate blocs of multicenter integrals.

In the case of HF + MP2 all many electron integrals should be transformed from AO to MO basis. In that case we expect more load of the network and we expect beter results from SCore computations. Fig. 2 shows that in that case sockets library preferable against SCore in speedup evaluation and even in absolute times evaluations.

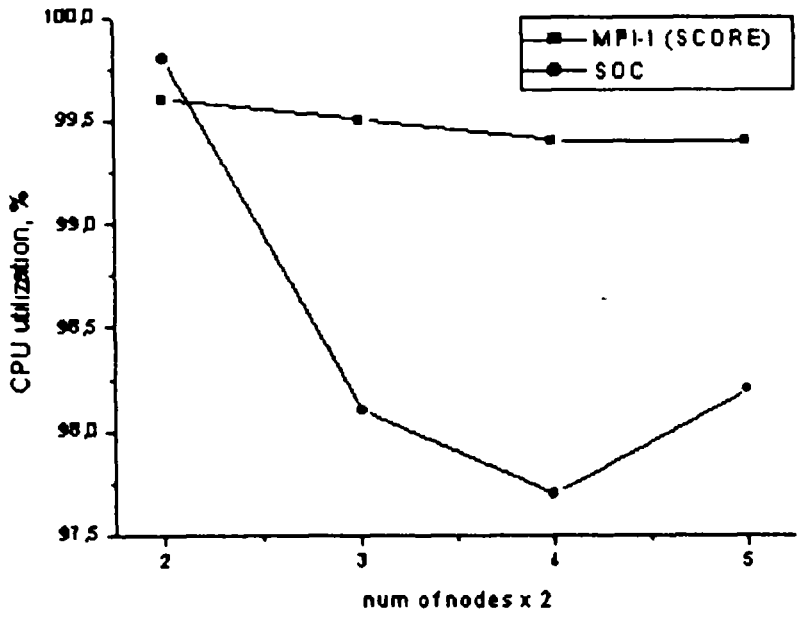

Fig. 1. Dependence of computation time from number of nodes in the cluster. Trinitrotoluene molecule single point computations in HF approximation.

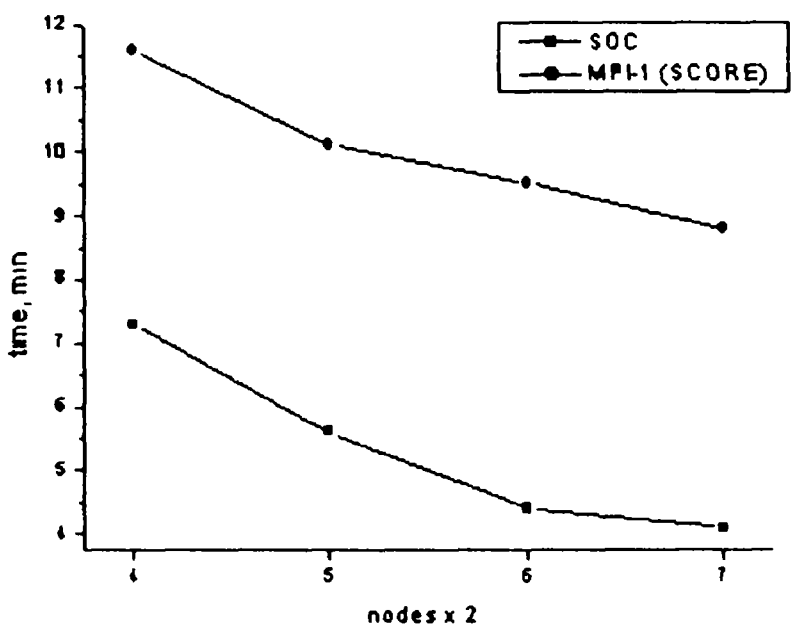

Fig. 2. Dependence of computation time from number of nodes in the cluster. Trinitrotoluene molecule single point computations in HF + MP2 approximation.

It can be explained by the fact that in parallel GAMESS runs two processes which are generated for each computer in the cluster. It is all correct in sockets case: on each CPU we have one service process and one computing process. As SCore regards all processes equaly it happens that SCOre runs more that one computational process on one CPU. We think this slows computations regadles of faster network communications in SCore system. For solving of that problem we plan to change the GAMESS parallelization code in order to have only one computation process on one CPU.

References

[1] `http://clusters.top500.org/db`.
[2] The Beowulf Project: `http://www.beowulf.org//`
[3] M.W. Schmidt *et al.*, General atomic and molecular electronic structure system, *J. Comput. Chem.*, **14**, 1347–63 (1993).
[4] H. Tezuka, A. Hori, Y. Ishikawa, M. Sato, PM: An operating system coordinated high performance communication library, in: *High-Performance Computing and Networking*, Peter Sloot Bob Hertzberger (Ed.), vol. 1225 of *Lecture Notes in Computer Science*, (1977), Springer–Verlag, pp. 708–717.
[5] Y. Ishikawa *et al.*, In: *Proceedings of the 5th Annual Linux Expo 55* (1999).
[6] E. Lusk, A. Skjellum, *Using MPI W. Gropp*, The MIT Press (1999).
[7] *High Performance Computing Linpack Bentchmark (HPL) HPL 1.0* `www.netlib.org/bentchmark/hpl`, September 27, 2000.
[8] Wilson S., *Electron Correlation in Molecules*, Moscow, Mir, (1987) (in Russian).
[9] Fudzinaga S., *Molecular Orbitals Method*, Moscow, Mir (1983) (in Russian).
[10] S. Cicėnas, R.J. Rakauskas, S. Vošterienė, J. Šulskus, Theoretical investigation of the vibrational spectra of trinitrotoluene and 2,4,6-trinitrophenol molecules, *Lithuanian J. of Physics*, **41** (2001).

Asmeninių kompiuterių klasterio našumo tyrimas, sprendžiant molekulių kvantinės mechanikos uždavinius

E. Pankevičius, R.-J. Rakauskas, J. Šulskus, S. Vošterienė

Asmeninių kompiuterių dirbančių lokaliam tinkle Linux operacinėje sistemoje, klasteriai yra tapę pigiausiais "superkompiuteriais". Mūsų darbe ištirtas LKA Taikomųjų mokslų katedroje sukurto klasterio TAURAS našumas, sprendžiant neempirius molekulių kvantinės mechanikos uždavinius. Darbe parodytos naudojamos Score programinės įrangos galimybės ir pranašumai, sprendžiant tokius uždavinius.