

PDC NEWSLETTER

Opening of PDC

On March 1:st PDC was inaugurated as the national facility for high performance computing by the Mayor of Stockholm Mats Hulth.

PDC was formed in 1990 to act as a focal point and national forum for both research on and use of parallel computers. The activities during the first years were centered around the Connection Machine. In 1994 an IBM SP system was also installed.

For many years, Sweden lagged behind other industrialized countries in the utilization of supercomputers. The Swedish Council for High Performance Computing (HPDR) was formed in July 1994 to rectify the situation. After careful evaluation, the Council decided to locate the main part of the activity to PDC with a funding of 38 MSEK until December 1996.

This made a substantial upgrade of the SP-system possible with more than a doubling of the capacity. The number of nodes has been increased from 55 to 96 and the existing nodes were upgraded with faster processors and more memory and disk. This makes the system the largest IBM SP system outside the US.

The "new PDC" with funding from HPDR was formally opened March 1:st by the President of KTH Janne Carlsson and the Mayor of Stockholm Mats Hulth. The ceremony was followed by a full day conference on high-performance computing with international well known speakers in several areas. One of the most spectacular pre-



The Mayor of Stockholm, Mats Hulth at the opening of PDC

sentation was given by Professor Chris Johnsson from University of Utah who showed a number of visualizations from medicine.

Sup'Prize to PDC users.

This year's Sup'Prize was awarded to Martin Schütz and Roland Lindh from the University of Lund. The price was presented at SUP'EUR 96 in Krakow, Poland. Martin Schütz and Roland Lindh received the price for their "integral direct, distributed-data parallel MP2 algorithm".

Roland Lindh describes their result as an apparent paradox. "If it takes one grave-digger ten hours to dig ten graves, it would take ten grave-diggers one hour to dig ten graves. But our algorithm means that it takes even less!" The reason for this being that the parallelization increases not only the computational power but also the total memory for the application

See page 3 for more.

In This Issue

Opening of PDC	1
Sup'Prize to PDC users	1
Editorial	2
An integral direct distributed-data, parallel MP2 algorithm	3
Calendar of Events	6
New Web pages	7
How to Contact PDC	8

EDITORIAL

Welcome to the first issue of the PDC Newsletter. We intend to publish four newsletters per year, one each quarter. The purpose of the PDC Newsletter is to follow the development at PDC, present examples of research activities using the PDC facilities, give practical information and announce future directions and events. We are also interested in a dialogue with our readers and are inviting letters to the editor and general comments.

During the last year we have seen a spectacular development at PDC. After an evaluation in the Spring of 95, by the Swedish Council for High Performance Computing (HPDR) and an international team of experts, PDC was awarded a grant to take on the responsibility as the leading high performance computing center serving the Swedish academic community. In the Summer of 95, KTH and HPDR then signed a contract describing the service which PDC will provide.

The Fall and Winter were used to build the required level of support. The IBM SP was extended to 96 nodes, new system software was developed, the networking and storage capabilities were drastically enhanced and the staff was increased to better support the users. Gradually new application software systems have been installed, based on requests from research groups around the country.

All this was leading up to the reopening of PDC as an HPDR facility, on schedule, March 1, 1996. See the article on page 1 for more details.

Since then, the operational routines at PDC are continued to be updated. The help desk program, for answering questions, has been enhanced and the allocation subcommittee of the PDC science council now meets quarterly to evaluate applications for computer resources.

Hardware and software have also been upgraded during the Summer in order to respond to increasing user demand. PDC has acquired a Fujitsu

VX with two vector processors having 2.2 Gflops and 2 GB capacity each. Furthermore, the PDC users now have another option for vector computing. Based on an agreement with KTH there is access to the KTH Cray J932 with 8 GB of memory. This computer also runs the HSM system DMF from Cray for the 6 TB IBM tape storage system. Finally as this editorial is being written the SP2 Strindberg is being upgraded to 110 processors.

We see education as an integral part of the PDC service. A number of short courses have been given and in August we had the two week Summer school: "Introduction to HPC" with 40 participants. These educational efforts and, of course, the increase in service and computer resources have enabled a large number of new users and new projects to open as well as convenient access to high performance computing.

The goal of PDC is to provide the best possible service to our user community. We are, therefore, very pleased to report on the recent Sup'Prize to Martin Schütz and Roland Lindh for their work on their integral direct, distributed-data parallel MP2 algorithm. An important part of the research was performed on the PDC SP, see the article starting on page 3

We congratulate them warmly to the award.

Björn Engquist

Director

An integral direct, distributed-data, parallel MP2 algorithm

Martin Schütz, Roland Lindh, Department of Theoretical Chemistry, University of Lund

In recent years, advances in computer technology together with substantial improvements in quantum chemical algorithms have enabled *ab initio* electronic structure calculations on chemical systems of increasing complexity. Common to all of these *ab initio* algorithms which aim at large scale problems is that they are integral direct in the sense that the electron repulsion integrals (ERIs) are reevaluated whenever needed, rather than computed once, stored on disk and read from disk when required. Hence, the bottleneck of storing the results which precluded many applications, is avoided. Furthermore, the heavy input/output load inflicted by conventional algorithms is also avoided. On the other hand, the CPU time required for a given problem, increases drastically. However, it was also recognized, that the number of integrals one

has to recompute can be reduced substantially by using proper integral pre-screening techniques.

Integral direct methods were first used in SCF theory, but have since been extended to methods which include electron correlation. Integral direct methods are especially powerful in combination with efficient parallelization strategies, moving problems of biochemical interest within reach of quantum chemical methods.

In this article, results are presented from a scalable, distributed-data parallel implementation of direct integral transformation for a restricted set of molecular orbitals indices. The code is capable of dealing with chemical systems beyond 1000 basis functions and 200 correlated electrons, including low-symmetry cases. Highly superlin-

ear speedups were observed due to efficient use of both the compute power and the bulk memory of the MPP system. The algorithm is currently implemented in an MP2 code, yet it is easily adopted for other electron correlation methods, where only ERIs over molecular orbitals integrals with two indices in the virtual orbitals space are required: actually, distributed-data parallel implementations for MP2-R12 and CASSCF, based on the same transformation code, are currently under development.

Performance

The parallel efficiency of the algorithm is demonstrated by calculations on the phenantrene molecule (C_{2v} symmetry, with 762 primitive Gaussians, contracted to 412 basis functions). For molecular systems of

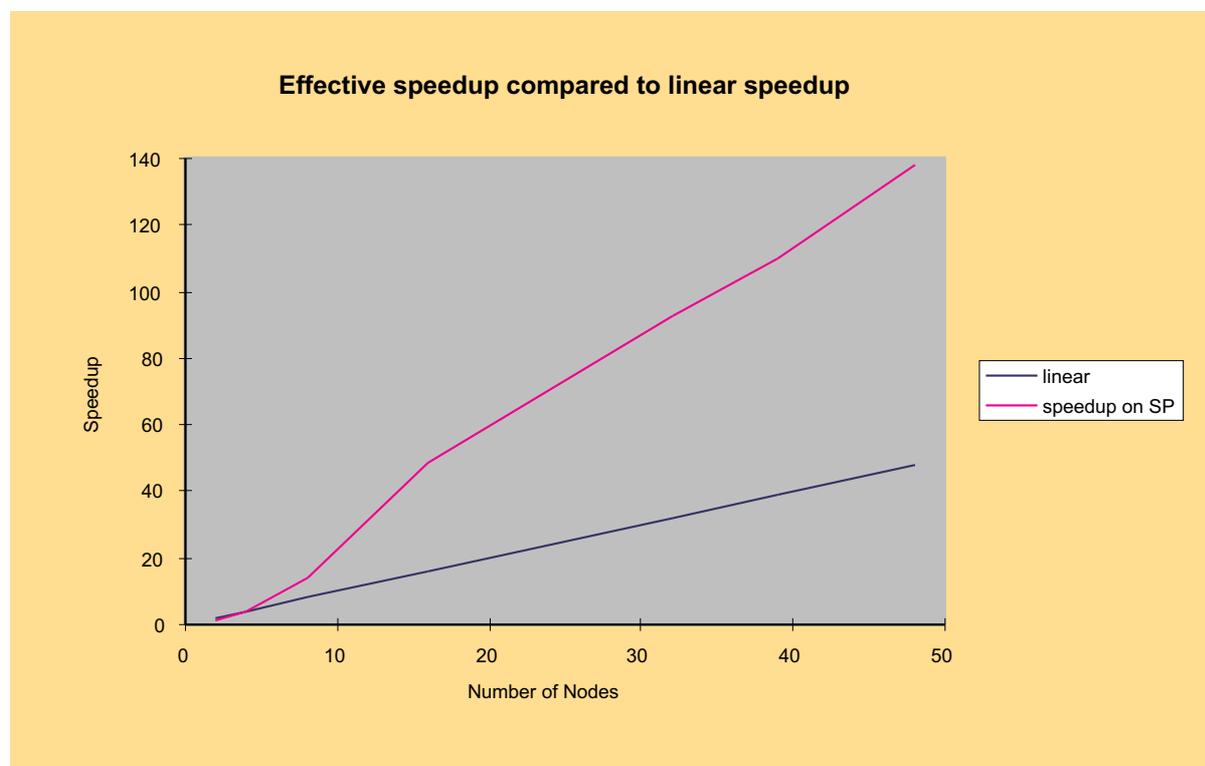
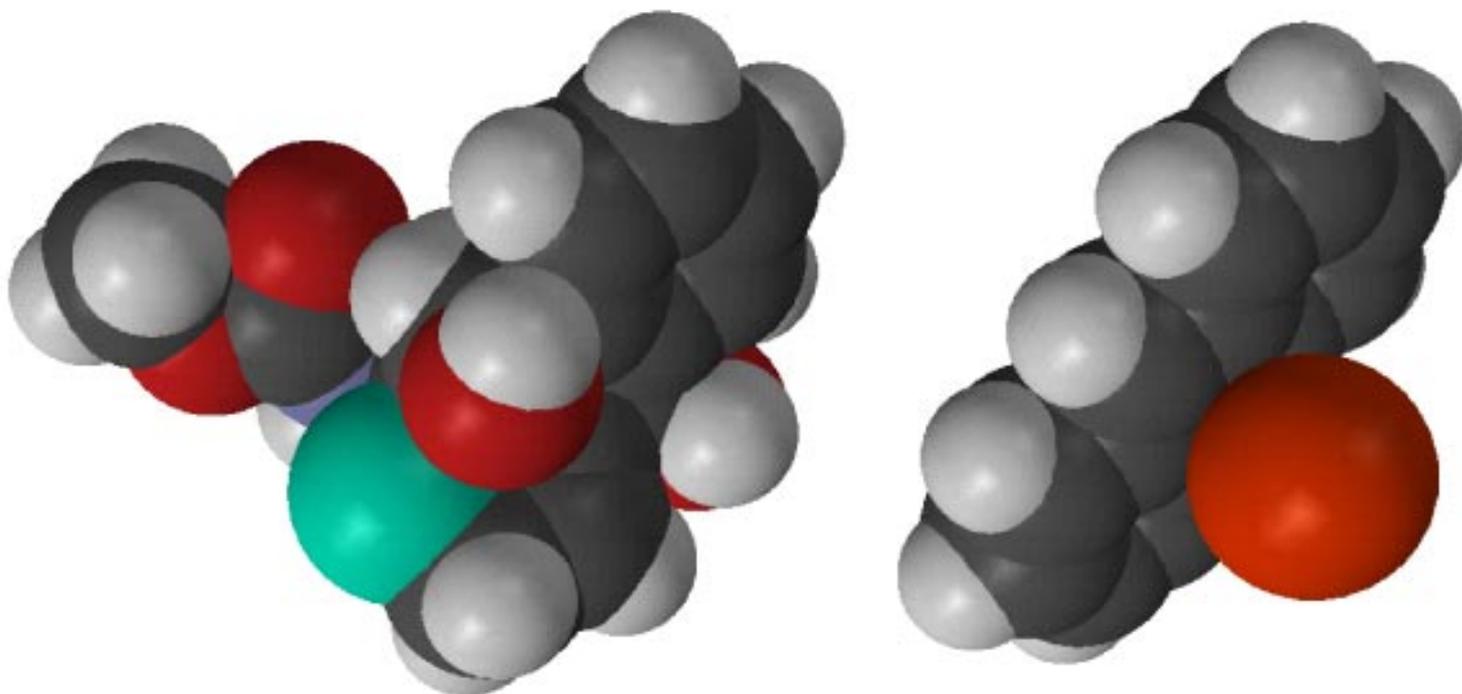


Figure 1. Graph showing the speedup with parallel processors.



Examples of molecular systems currently studied at PDC: On the left side the van der Waals complex anthracene-Ar in its supposed minimum-energy configuration, with the Ar atom adsorbed on the inner ring of the catacondensed aromatic ring system.

The right side shows a subunit of the anti-cancer drug calicheamicin in its configuration after the trigger reaction and ring closure.

this size, integral direct methods start to become the only possible route; hence this test case may have some practical significance, although it is still small enough to be well suited for scalability measurements. The number of symmetry-unique shells for this system is 31, forming 496 individual tasks. The calculations were performed on an IBM SP with 48 nodes at the Center for Parallel Computers (PDC) at the Royal Institute of Technology (KTH) in Stockholm. Figure 1 displays the speedup gained when the number of nodes increases. For comparison a line showing linear speedup is also added to the figure.

The measured wall clock time on 48 nodes is less than 15 minutes for this calculation, and the speedup relative to single node execution is estimated to 527. Speedup factors substantially

Speedup factors substantially above linear speedup are observed.

above linear speedup are observed. On 48 nodes, the observed speedup is 5.8 times larger than linear speedup. This superlinear speedup is a result of exploiting both the compute power

and the aggregate memory of the parallel computer, reducing the number of necessary integral passes and hence saving not only wall clock but CPU time relative to single node execution.

Outside of the superlinear regime, i.e. between 16 and 48 nodes when only a single pass is necessary, the algorithm still shows satisfying performance, with a parallel efficiency on 48 vs. 16 nodes of 95%. Calculations in excess of 1000 basis functions with reasonable elapsed times are possible with our code. The largest calculations performed so far comprised 1266 pri-

mitives contracted to 702 basis functions, and 82 correlated electrons. On 64 SP nodes, each with 128 MByte memory, this calculation took about 90 min wall clock, and two integral passes.

Applications

Currently, two application projects based on the new integral direct, parallel MP2 code are in progress at PDC:

(i) Weakly bound van der Waals complexes between aromatic molecules and rare gas atoms: Solvent clusters M-R between aromatic molecules (M) and rare-gas atoms (R) have become prototype systems for the investigation of solvation at the microscopic level. The combination of molecular beam techniques, laser and mass spectroscopy enables the synthesis of such clusters in collisionless environment, and their mass-specific detection and spectroscopy. Recently, the stimulated emission pumping/resonant two-photon ionization (SEP-R2PI) method was introduced, a novel spectroscopic technique, which allows for the mass-selective determination of van der Waals binding energies of such systems in the ground and 1st electroni-

cally excited state with very high accuracy.

On the theoretical side, *ab initio* calculations on such system are very demanding, since in order to model the predominant dispersive interactions properly, large basis sets have to be used. In the present PDC project the intermolecular potential energy surfaces (IPES) of the anthracene-Ar and anthracene-Ar vdW complexes are studied at the MP2 level, employing large ANO basis sets, i.e. 700 basis functions and more. Different bonding sites of the Ar atom on top of the cata-condensed aromatic ring system are investigated. The computed *ab initio* IPES will be used later to construct an analytical model potential for the M-R solute-solvent interaction, which in turn can be employed in Monte Carlo and dynamics studies to investigate microsolvation effects, occurring in larger anthracene-Ar_n ($n=10..40$) and similar clusters.

(ii) the significance of the trigger reaction of the calicheamicin anti cancer drug: The new enediyne class of anti cancer drugs acts by cleaving the DNA strand after undergoing a Bergman like autoaromatization and

forming a highly reactive biradical. This autoaromatization process is initiated by a so called trigger reaction. The calicheamicin is a molecule of this class, which all have in common the enediyne moiety. Whereas the formation of the biradical is fully understood the significance of the preceding trigger reaction still needs to be investigated. The favored hypothesis is that the effect of the trigger reaction is to alter the enediyne structure itself, leading to a decrease in the activation barrier of the biradical formation.

Another theory is that the trigger reaction enables the biradical formation by saturating a double bond in the vicinity of the enediyne moiety. The presence of this double bond would make the biradical formation energetically unfavorable. i.e. the trigger reaction is reducing the reaction energy of the biradical formation. The present study investigates the latter hypothesis at the MP2 level of theory employing small to large ANO basis sets (400-800 basis functions). The calculations are extremely expensive since the molecules under study are large and do not have any symmetry.

Calendar of Events

September

- 3 Recent Advances In Many Body Quantum Mechanics: Applications to Chemical and Materials Systems
PDC
<http://www.pdc.kth.se/events/pdc-sem-960903.html>
- 8-11 SUPEUR 96
High Performance Computing in Europe on IBM Platforms
Krakow, Poland.
<http://www.cyf-kr.edu.pl/supeur96>
email: supeur96@cyf-kr.edu.pl
- 9 Wavelets and two-dimensional turbulence
NADA
Marie Farge, Ecole Normale, Paris
<http://www.pdc.kth.se/events/nada-sem-60909.html>
- 17-18 SHPCnet
Swedish High Performance Computing Network
Nova Park Hotell, Knivsta
http://www.nsc.liu.se/SHPCnet_inbjudan.html
email: msa@nsc.liu.se
- 17-19 Fifth NASA/Goddard Space Flight Center Conference on Mass Storage Systems and Technologies
University of Maryland.
<http://esdis.gsfc.nasa.gov/msst/msst.html>
email: whetzel@ltpmail.gsfc.nasa.gov
- 17-21 Physics Computing '96
Krakow, Poland
<http://www.cyf-kr.edu.pl/pc96/index.html>
email: pc96@cyf-kr.edu.pl

October

- 7-8 Second DELOS Workshop
Bonn, Germany
<http://www.gmd.de/People/Thomas.Baker/Delos-bonn.call.html>
- 7-18 1996 First Electronic Molecular Graphics and Modelling Society Conference
to be held on the Internet and World Wide Web.
<http://bellatrix.pcl.ox.ac.uk/mgms>
- 22-25 Gaussian Workshop
Mexico City
email: info@gaussian.com

27-November 1

Visualization '96
San Fransisco, California
<http://WWW.ERC.MsState.Edu/conferences/vis96/>
<http://davinci.informatik.uni-kl.de/Vis96/>
email: nejohnston@lbl.gov
email: ribarsky@cc.gatech.edu

November

- 4-5 10th ERCIM Database Research Group Workshop on Heterogenous Information Management
Praha, Czech Republic
http://www-ercim.inria.fr/publication/Ercim_News/events/10th-EDRG.html
email: pokorny@kki.ms.mff.cuni.cz
- 17-22 Supercomputing 96
Pittsburgh, Pennsylvania
<http://scxy.tc.cornell.edu/sc96/index.html>
email: questions96@mail.supercomp.org
- 23-30 SOFSEM '96
XXIII-rd Seminar on Current Trends in Theory and Practice of Informatics
Milovy, Czech Republic
email: sofsem@ics.muni.cz
- 27-31 Frontiers'96
Annapolis, Maryland
<http://www.aero.hq.nasa.gov/hpcc/front96.html>
email: frontiers96@cesdics.gsfc.nasa.gov

December

- 2-5 The 1996 Asian Computing Science Conference
Singapore
<http://www.iscs.nus.sg/~asian96>
email: asian96@iscs.nus.sg
- 3-6 Gaussian Workshop
Liege, Belgium
email: info@gaussian.com
- 3-4 Workshop on Scientific Visualization
PDC
email: johani@pdc.kth.se
- 11-12 Workshop on Computational Electromagnetics
PDC
email: per@pdc.kth.se

16-17 Annual PDC-conference
 PDC
 Software for parallel computing
 email: gert@pdc.kth.se

January 97

5-March 5
 Virtual Workshop
 Parallel Programming with Message-Passing
 Libraries
 Cornell Theory Center
 http://www.tc.cornell.edu/Edu/Upcoming/vw.html
 email: jeanne@tc.cornell.edu

March 97

13 SIAM Short Course on Performance Programming
 for Scientific Computation

14-17 Eighth SIAM Conference on Parallel Processing
 for Scientific Computation.
 Minneapolis, Minnesota
 http://www.siam.org/meetings/pp97/pp97home.htm
 email: meetings@siam.org

April 97

8-11 Fractal 97
 Fractals in the Natural and Applied Sciences
 4th International Multidisciplinary Conference
 Denver, Colorado, USA
 http://www.kingston.ac.uk/%7Eap_s412/
 fract97.html
 email: novak@kingston.ac.uk

9-11 ISADS 97
 The Third International Symposium on
 Autonomous Decentralized Systems
 Berlin, Germany
 http://www.fokus.gmd.de/ws/isads97/
 email: isads97@fokus.gmd



New Web Pages

The web pages of PDC have been redesigned and are now using a graphical interface. We hope that the pages should be easier to navigate and that you find the information you need. All the pages at PDC have been revised and updated. An electronic form of this newsletter will also be added to the pages.

Please have a look at <http://www.pdc.kth.se>.

B

Sverige
Porto
Betalt

PDC NEWSLETTER

Published quarterly by the Center for Parallel Computers at KTH

The center operates leading-edge, high-performance computers as easily accessible national resources. These resources are primarily available for Swedish academic research and education.

PDC receives funding from The Swedish Council for High-Performance Computing, HPDR and KTH.

Editor: Björn Engquist

Editorial staff: Peter Carlsson

ISSN 1401-9671

How to Contact PDC

Visiting address: Lindstedtsvägen 3, Plan 5, KTH, Stockholm
Mail: PDC, KTH, S-100 44 STOCKHOLM
Email: pdc-staff@pdc.kth.se
WWW: <http://www.pdc.kth.se/>
Phone: +46 8 790 78 00
Fax: +46 8 24 77 84