ITR Collaborative Research:
NOVEL SCALABLE SIMULATION TECHNIQUES FOR CHEMISTRY,
MATERIALS SCIENCE, AND BIOLOGY

Principal Investigators:


NSF Review Meeting – Urbana, June 19-20, 2002
Motivation

Growing impact of Electronic and Molecular Modeling in several areas of science

Algorithmic developments difficult to share among different groups

Need efficient and systematic incorporation of new software and hardware developments into scientific codes

Methods are only available to experts
Mission

Combine application scientists, algorithm developers, and computer scientists to:

• Improve methodological base and use it in new applications

• Develop algorithms that can be easily/efficiently ported to different platforms and can scale on thousands of processors

• Develop common standards for scientific codes

• Embed new techniques and algorithms in a friendly, extensible, object oriented framework accessible to the community at large (open source environment)
Strategic partnerships

- European CP2K consortium headed by M. Parrinello (ETH/CSCS Switzerland) – open source initiative for ab-initio molecular dynamics and related codes

- Democritos center for scientific computing (Trieste, Italy, INFM center of excellence) directed by S. Baroni (SISSA, Trieste) – open source initiative for PW pseudopotential calculations – DFT linear response codes

- CECAM (Lyon) – schools and workshops

- LLNL - summer school on materials simulation

- .....Others
Challenge

• How to make the diversity of codes and approaches (even within the team) a strength rather than a weakness

• How to overcome barriers separating modeling scientists and algorithm developers from computer scientists

• How to plan today something that may continue to be useful and help to make progress in the years to come (5 year and beyond perspective)
Strategy

• Adopt a democratic approach (open source initiative)
• Devise realistic short, medium and long term goals such as:

  short term - adopt common standards in different codes maintained by different groups (data structures, compatible input/output, etc.) – create a web site for the initiative

  medium term – develop object oriented structures to allow use of same subroutine in different codes

  long term – integrate algorithms and software tools in a user friendly environment – contribute to develop a new paradigm for progress in the area
Preliminary results & examples of application
Algorithms and scientific software

• The CP90 code (first-principles MD code of the Princeton group) and the PWSCF code (Trieste) already have very similar data structures and use compatible input/output – Full external compatibility will be achieved soon and both codes will be made available on the web site of the initiative (in preparation)

• The PINY code (first-principles and classical MD code developed by Martyna, Tuckerman and Klein) and the CP90 codes will share capabilities that are now available only to one code (ex. QM/MM in PINY, ultrasoft pseudopotentials in CP90). This will also be available on the web site
QM/MM in the PINY code (from Martyna and Tuckerman)

Applications: HCA II in Water

- The system consists of a 260-residue HCA-II enzyme (with catalytic zinc), solvated by 8,859 waters, for a total of 30,652 atoms.
- Only the catalytic zinc, the side-chains of active site residues, HIS 94, HIS 96, HIS 119, THR 199, GLU 106 and the seven water molecules in the active site are treated using an ab initio description, is studied (Rothlisberger 1e pseudos) others TM.
- Thus, 204 valence electrons of 68 atoms in the active site are treated at an ab initio level while the remainder of the system is treated using the empirical CHARMM22 all-atom force field which includes TIP3P water model.
Orthonormal localized real-space basis sets

Localized functions on grid points

\[ u_i(x_j) = \delta_{ij} \]

\[ \langle u_i | O(x) | u_j \rangle \sim \omega(x_j) \delta_{ij} \]

Use DVRs (Light et al. 1985) in ab-initio MD (from Yi Liu and Mark Tuckerman)

- Long range forces evaluated in reciprocal space
Ensemble DFT ab-initio MD code (Yudong Wu, Nicola Marzari and R.C.) : oxygen dissociating at a vacancy on a nanotube
- CO $\rightarrow$ CO$_2$ oxidation on reduced rutile (110)

(from A. Selloni)
Electronic and Molecular simulation codes will be interfaced with parallel Charm++ (developed by L. Kale and collaborators). Preliminary work has been made in connection with the PINY code.

Scalability tests and improve portability to PSC in collaboration with N. Nydstrom
work done so far by Kale and Students

A preliminary program working in parallel does multiple 3D FFTs and some other step involved in the ab Initio algorithm. The interface is been restructured to make it a clear interface between application data structures and subroutines and the parallel coordination code.

Some preliminary performance runs for the skeleton code have been made, and communication bottlenecks have been identified. Libraries of algorithms that can optimize these communication steps are built (by using pipelining).
Future Design Goals

- Localized basis sets (discrete variable representations)
- Maximally localized orbital evolution schemes.
- External fields.
- Additional analysis tools (Rahman, NMR, etc.).
- Middleware tools for load-balancing and scaling on 1000s of processors (collaboration with L. Kale).
- Interactive development tools.
- Robust top-level C++ class structure (e.g. avoid deep inheritance trees).
- Portability across industry-standard C++ compilers.
- Component-based design.
People Currently Involved

Princeton: PIs: R. Car and A. Selloni, postdocs: Balazs Hetenyi + new hire, graduate students: Yudong Wu (supported by project) + another graduate student (Xiaofei Wang (supported by other project)), collaboration with Democritos center supports a senior scientist (P. Giannozzi) who collaborates with both Princeton and Democritos groups.

Upenn: PIs: M. Klein, postdocs: Yves Mantz, Troy Whitefield, graduate student: Ivaylo Ivanov.


UIUC: PIs: Sanjay (V.L.) Kale and J. Torellas, graduate students: Ramkumar Vadali, Orion Lawlor (Supported by other grants), Theckla Louchios, Sindhura Bandhakavi, Sameer Kumar (supported by other grants), Chee Wai Lee.

PSC: N. Nydstrom (partially supported by grant).