Simulations at the nanoscale on the GRID using Quantum ESPRESSO

P. Giannozzi
Università di Udine and Democritos CNR-IOM Trieste, Italy

Hands on Training School on Molecular and Material Science GRID Applications, Trieste, 2010/03/31

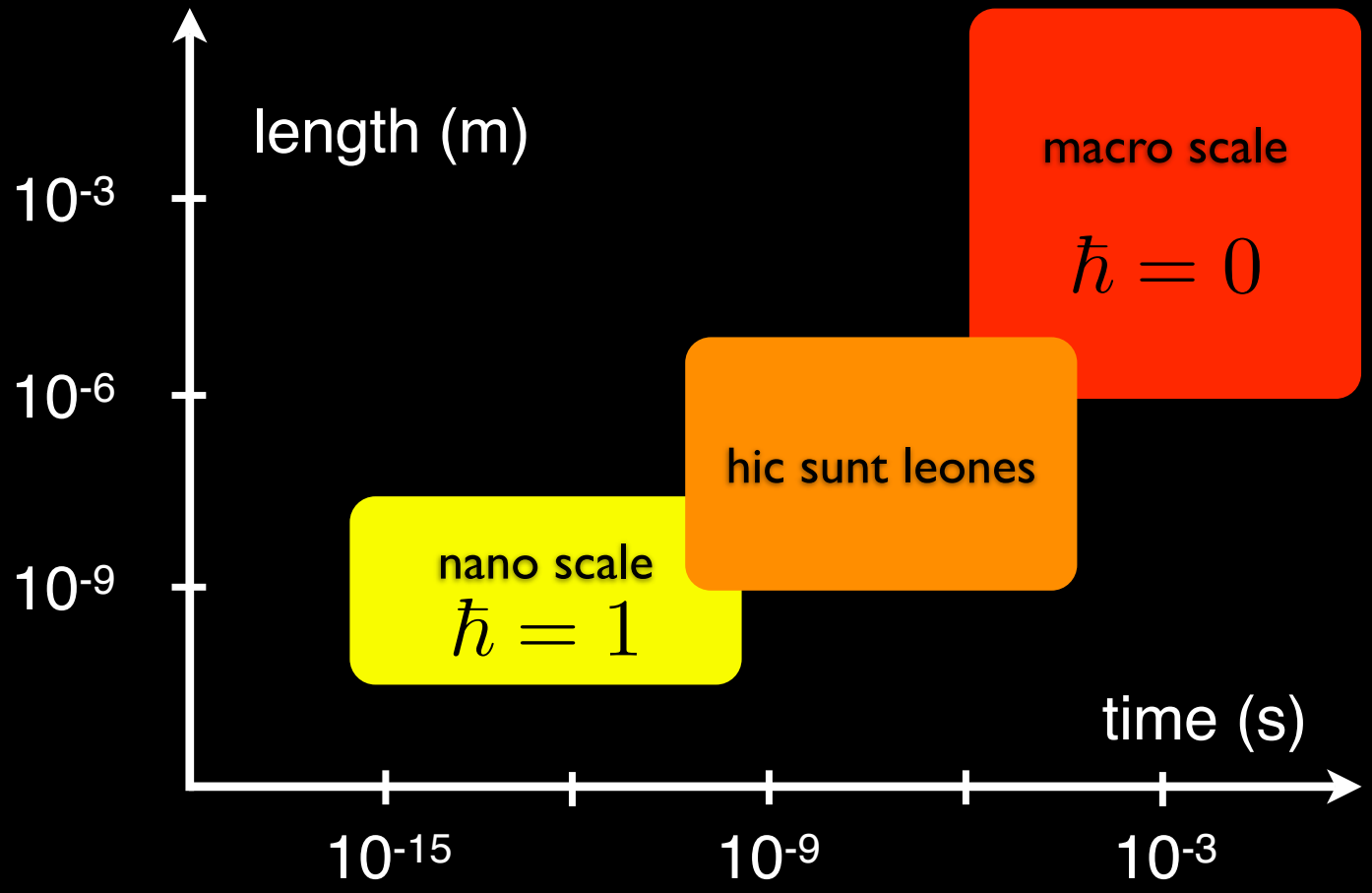
DEMOCRITOS
DEmocritos MOdeling Center for Research In aTOmistic Simulation

Quantum simulation of matter at the nanoscale

- Density-Functional Theory (DFT) (P. Hohenberg, W. Kohn, and L. Sham, 1964-65)
- Pseudopotentials (J.C. Phillips, M.L. Cohen, M. Schlüter, D. Vanderbilt and many others, 1960-2000)
- Car-Parrinello and other iterative techniques (SISSA 1985, and many other places since)

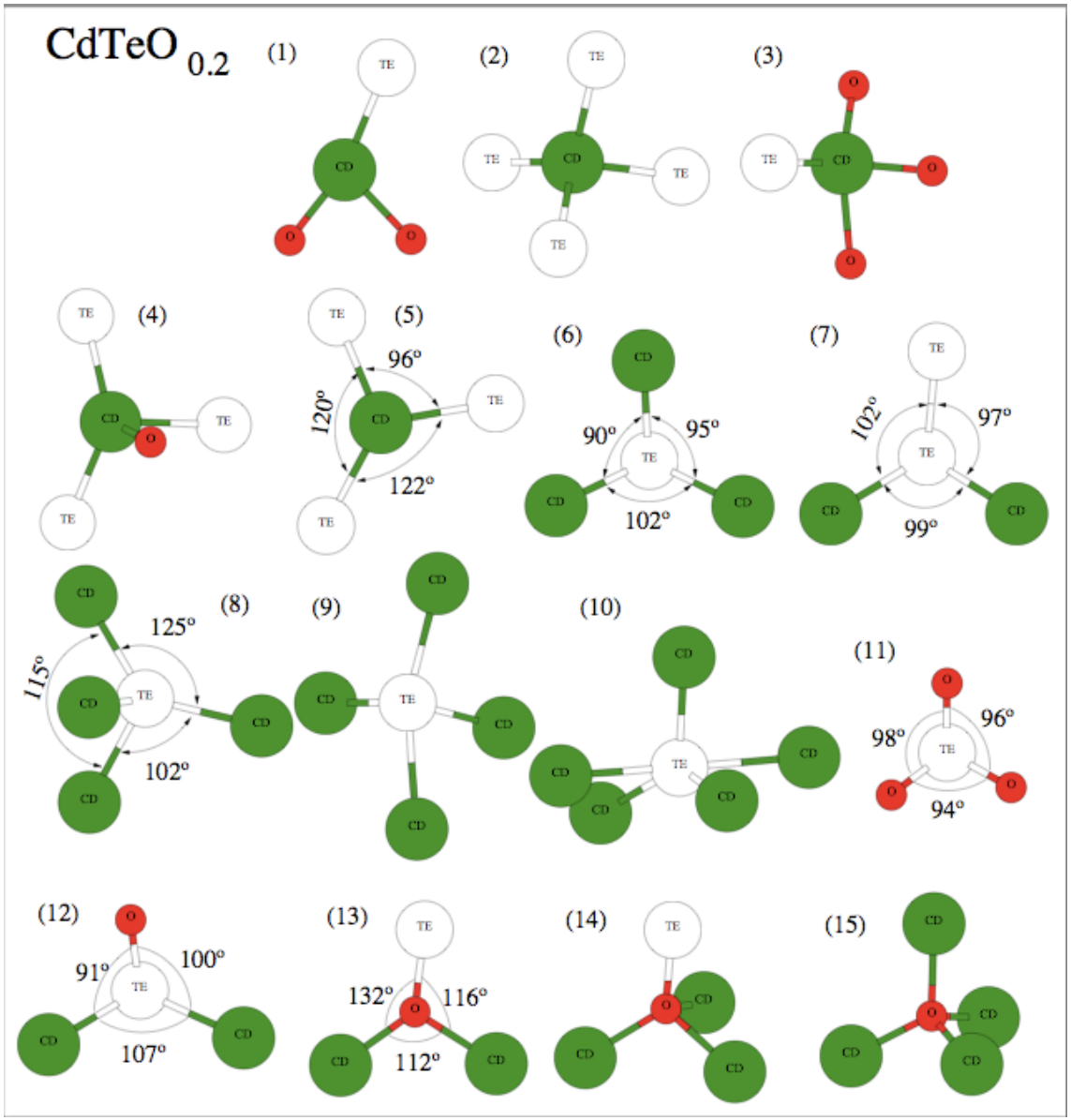
Sometimes referred to as *The Standard Model* of materials science

the saga of time and length scales

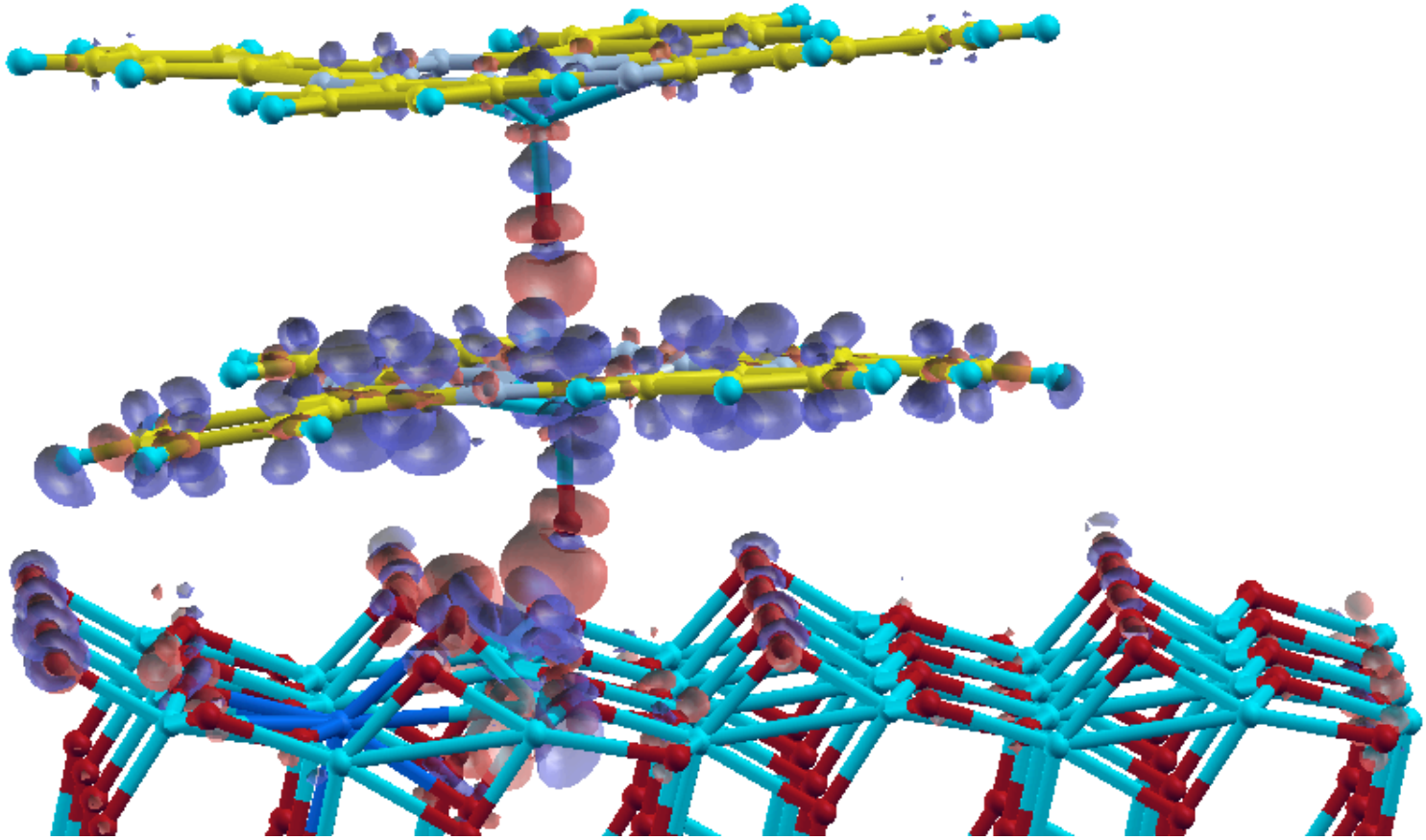


New materials

Most common atomic configurations in amorphous $CdTeO_x$, $x = 0.2$; work done in collaboration with E. Menendez

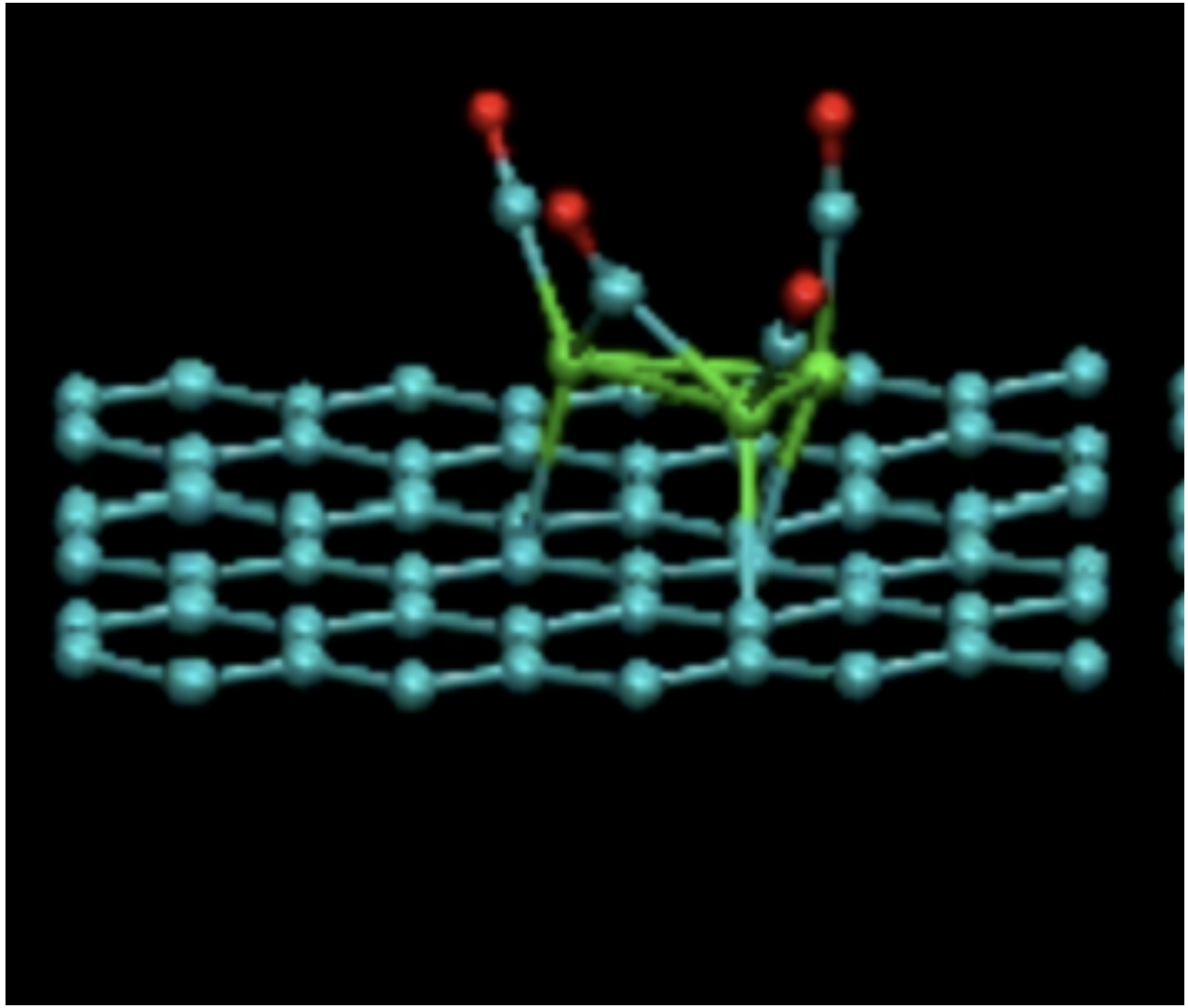


New devices



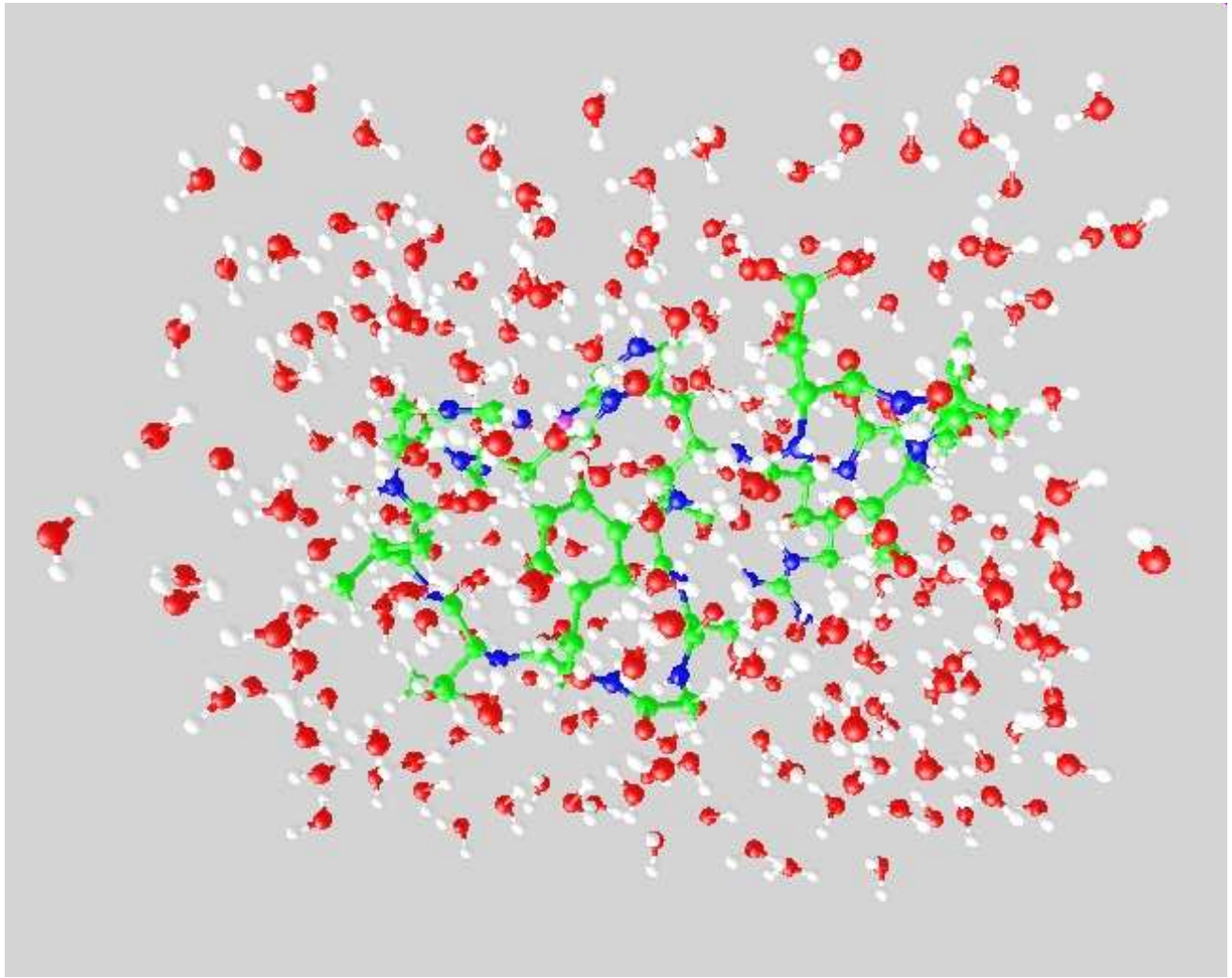
(organic-inorganic semiconductor heterojunction, phtalocyanine over TiO_2 anatase surface; with G. Mattioli, A. Amore, R. Caminiti, F. Filippone)

Nanocatalysis



(3 Rh atoms and 4 CO molecules on graphene; with S. Furlan)

Biological systems



Metal-β-amyloid interactions; with V. Minicozzi, S. Morante,G. Rossi

ab initio simulations

$$i\hbar\frac{\partial\Phi(r,R;t)}{\partial t}=\left(-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial R_I^2}-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r_i^2}+V(r,R)\right)\Phi(r,R;t)$$

the Born-Oppenheimer approximation (M>>m)

$$M\ddot{R}_I=-\frac{\partial E(R)}{\partial R_I}$$

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r_i^2}+V(r,R)\right)\Psi(r|R)=E(R)\Psi(r|R)$$

density functional theory

$$V(r,R) = \frac{e^2}{2}\frac{Z_I Z_J}{|R_I - R_J|} - \frac{Z_I e^2}{|r_i - R_I|} + \cancel{\frac{e^2}{2}\frac{1}{|r_i - r_j|}}$$

$$V(r,R) \rightarrow \frac{e^2}{2}\frac{Z_I Z_J}{|R_I - R_J|} + v_{[n(r)]}(r)$$

Kohn-Sham Hamiltonian

$$n(r) = \sum_v |\phi_v(r)|^2$$

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2} + v_{[n(r)]}(r)\right)\phi_v(r) = \epsilon_v \phi_v(r)$$

Kohn-Sham equations from functional minimization

$$E[\{\phi\}, R] = -\frac{\hbar^2}{2m}\sum_v \int \phi_v^*(r)\frac{\partial^2\phi_v(r)}{\partial^2 r}dr + \int v(r,R)n(r)dr +$$

$$+\frac{e^2}{2}\int \frac{n(r)n(r')}{|r-r'|}drdr' + E_{xc}[n(r)]$$

$$E(R) = \min\big(E[\{\psi\}, R]\big)$$

$$\int \phi_v^*(r)\phi_u(r)dr = \delta_{uv}$$

Kohn & Sham

$$H_{KS}\phi_v = \epsilon_v\phi_v$$

Helmann & Feynman

$$\frac{\partial E(R)}{\partial R_I} = \int \frac{\partial v(r,R)}{\partial R_I}n(r)dr$$

The tricks of the trade

- expanding the Kohn-Sham orbitals into a suitable *basis* set turns DFT into a multi-variate minimization problem, and the Kohn-Sham equations into a *non-linear matrix eigenvalue problem*
- the use of *pseudopotentials* allows one to ignore chemically inert core states and to use *plane waves*
- plane waves are *orthogonal* and the *matrix elements* of the Hamiltonian are usually easy to calculate; the *completeness* of the basis is easy to check
- plane waves allow to efficiently calculate *matrix-vector products* and to solve the Poisson equation using *Fast Fourier Transforms* (FFTs)

The tricks of the trade II

- plane waves require *supercells* for treating finite (or semi-infinite) systems
- plane-wave basis sets are usually large: *iterative diagonalization* or *global minimization*
- summing over occupied states: *special-point* and *Gaussian-smearing* techniques
- *non-linear extrapolation* for self-consistency acceleration and density prediction in Molecular Dynamics
- choice of fictitious masses in Car-Parrinello dynamics
- . . .

Accuracy vs. Approximations

Theoretical approximations / limitations:

- the Born-Oppenheimer approximation
- DFT functionals (LDA, GGA, ...)
- pseudopotentials
- no easy access to excited states and/or quantum dynamics

Numerical approximations / limitations

- finite/limited size/time
- finite basis set
- differentiation / integration / interpolation

Requirements on effective software for quantum simulations at the nanoscale

- Challenging calculations stress the limits of available computer power: software should be **fast and efficient**
- Diffusion of first-principle techniques among non-specialists requires software that is **easy to use and error-proof**
- Introducing innovation requires new ideas to materialize into new algorithms through codes: software should be **easy to extend and to improve**
- Complex problems require a mix of solutions coming from different approaches and methods: software should be **interoperable with other software**

The Quantum ESPRESSO distribution

The Democritos National Simulation Center, based in Trieste, is dedicated to atomistic simulations of materials, with a strong emphasis on the development of high-quality scientific software

Quantum ESPRESSO is the result of a Democritos initiative, in collaboration with researchers from many other institutions (SISSA, ICTP, CINECA Bologna, Princeton, MIT, EPF Lausanne, Oxford, Paris IV...)

Quantum ESPRESSO is a distribution of software for atomistic calculations based on electronic structure, using density-functional theory, a plane-wave basis set, pseudopotentials.

Quantum ESPRESSO stands for *Quantum opEn-Source Package for Research in Electronic Structure, Simulation, and Optimization*

Computer requirements of quantum simulations

Quantum ESPRESSO is both CPU *and* RAM-intensive.
Actual CPU time and RAM requirements depend upon:

- *size of the system under examination:* CPU $\propto N^{2\div3}$, RAM $\propto N^2$, where $N =$ number of atoms in the supercell or molecule
- *kind of system:* type and arrangement of atoms, influencing the number of plane waves, of electronic states, of **k**-points needed
- *desired results:* computational effort increases from simple self-consistent (single-point) calculation to structural optimization to reaction pathways, molecular-dynamics simulations

CPU time mostly spent in FFT and linear algebra.
RAM mostly needed to store wavefunctions (Kohn-Sham orbitals)

Typical computational requirements

Basic step: self-consistent ground-state DFT electronic structure.

- Simple crystals, small molecules, up to ~ 50 atoms – CPU seconds to hours, RAM up to 1-2 Gb: *may run on single PC*
- Surfaces, larger molecules, complex or defective crystals, up to a few hundreds atoms – CPU hours to days, RAM up to 10-20 Gb: *requires PC clusters or conventional parallel machines*
- Complex nanostructures or biological systems – CPU days to weeks or more, RAM tens to hundreds Gb: *massively parallel machines*

Main factor pushing towards parallel machines is excessive CPU time; but when RAM requirements exceed the RAM of single machine, one is left with parallel machines as the only choice

Quantum ESPRESSO and High-Performance Computing

A considerable effort has been devoted to Quantum ESPRESSO parallelization. Several parallelization levels are implemented; the most important, on *plane waves*, requires fast communications.

Recent achievements (mostly due to Carlo Cavazzoni, CINECA):

- realistic calculations (e.g 1532-atom porphyrin-functionalized nanotube) on up to ~ 5000 processors
- initial tests of realistic calculations on up to ~ 65000 processors using mixed MPI-OpenMP parallelization

Obtained via addition of more parallelization levels and via careful optimization of nonscalable RAM and computations.

Quantum ESPRESSO and the GRID

Large-scale computations with Quantum ESPRESSO require large parallel machines with *fast communications*: unsuitable for GRID. BUT: often many smaller-size, loosely-coupled or independent computations are required. A few examples:

- the search for transition pathways (Nudged Elastic Band method);
- calculations under different conditions (pressure, temperature) or for different compositions, or for different values of some parameters;
- the search for materials having a desired property (e.g. largest bulk modulus, or a given crystal structure);
- **full phonon dispersions in crystals**

Hand-made GRID computing

Accepted Manuscript

$Si_xC_{1-x}O_2$ alloys: A possible route to stabilize carbon-based silica-like solids?

Assa Aravindh, Artoto Arkundato, Sonali Barman, Stefano Baroni, B.L. Bhargava, K.R.S. Chandrakumar, Wei Chen, Roby Cherian, Andrea Dal Corso, Soumendu Datta, Stefano de Gironcoli, Suman S. Dhayal, Alok Kumar Dixit, Sudipta Dutta, Pavel D'yachkov, Calin Gabriel Floare, Nirmal Ganguli, Shreemoyee Ganguly, Ralph Gebauer, Saurabh Ghosh, Paolo Giannozzi, Govind, Alison J. Hatt, K.P.S.S. Hembram, Mighfar Imam, V. Jayalakshmi, C.S. Jayanthi, Tuhina Kelkar, Anil Kumar, Jun Hee Lee, Mal-Soon Lee, Dayana Lonappan, Priya Mahadevan, Sairam Swaroop Mallajosyula, Madhura Marathe, Nicola Marzari, Brent Melot, Nicholas Miller, Joseph Morrone, Sachin Nanavati, Asiri Nanayakkara, Prithwish Kumar Nandi, Shobhana Narasimhan, Bhaarati Natarajan, Fahmida Parvin, Sujata Paul, Kalpataru Pradhan, G. Praveena, Dasari L.V.K. Prasad, Himanshu K. Poswal, Bhalchandra Pujari, Raghani Pushpa, K. Hari Krishna Reddy, Srijan Kumar Saha, Carlo Sbraccia, Sandro Scandolo, Prasenjit Seal, Ghazal S. Shafai, K.V. Shanavas, James O.H. Simrall, Aarti Srirangarajan, Vipul Srivastava, Mina K. Talati, Yuthana Tantirungrotechai, Kartick Tarafder, Tiju Thomas, T. Uthayathasan

PII: S0038-1098(07)00661-8
DOI: 10.1016/j.ssc.2007.09.011
Reference: SSC 9271

To appear in: *Solid State Communications*

Received date: 10 May 2007
Revised date: 13 August 2007
Accepted date: 13 September 2007

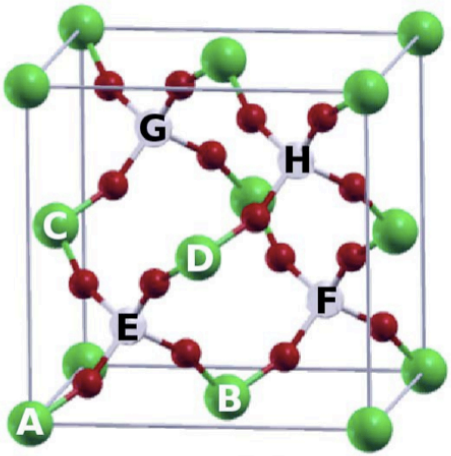
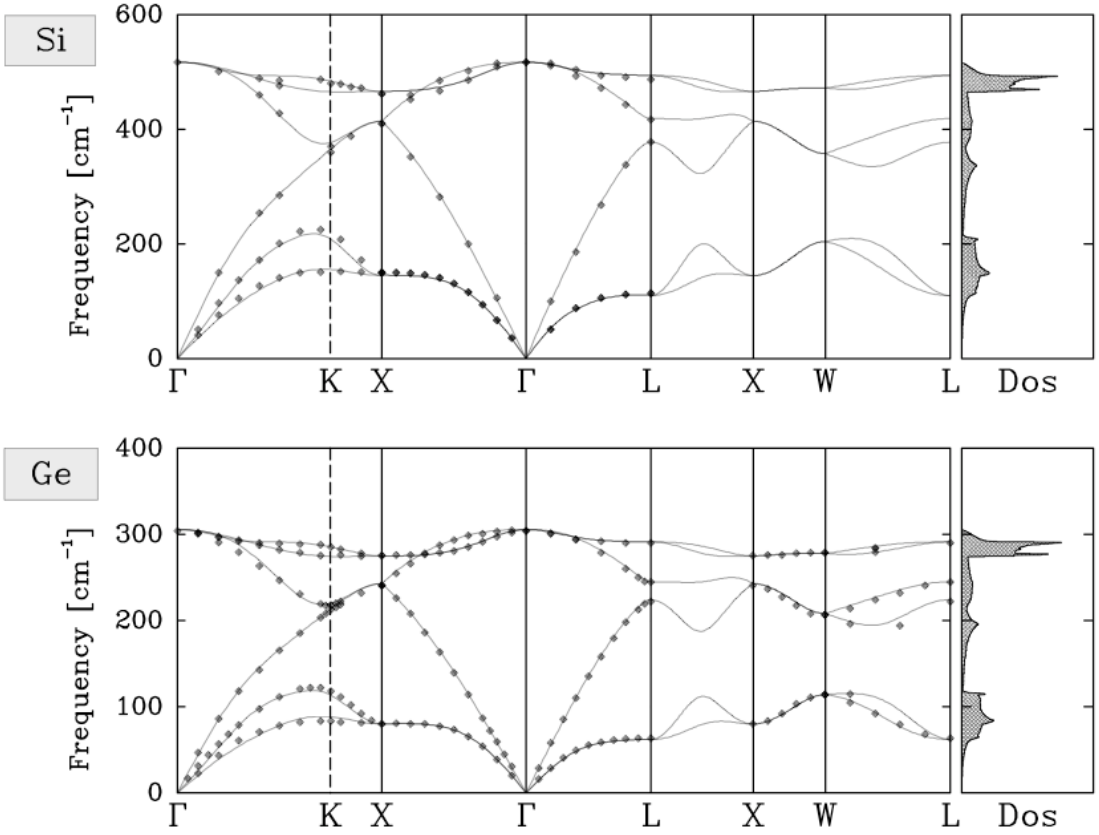


Figure 1: Example of a conventional cubic unit cell for I-42d cristobalite. The small dark spheres (red online) indicate O atoms, while the large grey (green online) and small light (white online) spheres represent Si and C atoms respectively. The positions of C/Si atoms are labeled A to H (see also Table I); the figure shows a representative structure (g) at a 50-50 Si-C concentration.

Vibration modes (phonons) in crystals



Phonon frequencies $\omega(\mathbf{q})$ are determined by the secular equation:

$$\| \widetilde{C}_{st}^{\alpha\beta}(\mathbf{q}) - M_s\omega^2(\mathbf{q})\delta_{st}\delta_{\alpha\beta} \| = 0$$

where $\widetilde{C}_{st}^{\alpha\beta}(\mathbf{q})$ is the matrix of *force constants* for a given $\mathbf{q}$

Calculation of phonon dispersions

- The force constants $\widetilde{C}_{st}^{\alpha\beta}(\mathbf{q})$ are calculated for a uniform grid of n_q $\mathbf{q}$-vectors, then Fourier-transformed to real space

- For each of the n_q $\mathbf{q}$-vectors, one has to perform $3N$ *linear-response* calculations, one per atomic polarization; or equivalently, 3ν calculations, one per *irrep* (symmetrized combinations of atomic polarizations, whose dimensions range from 1 to a maximum of 6)

Grand total: $3\nu n_q$ calculations, may easily become heavy. But:

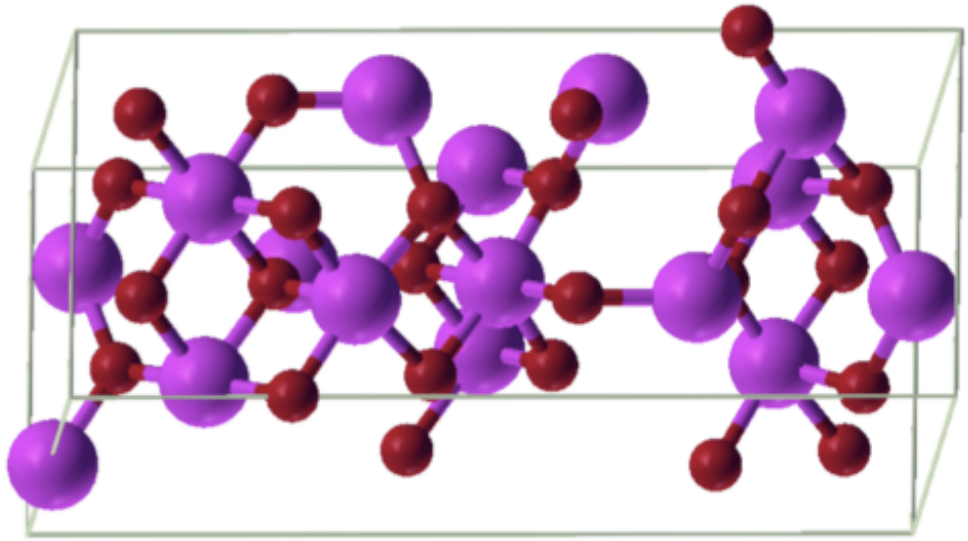
- Each $\widetilde{C}_{st}^{\alpha\beta}(\mathbf{q})$ matrix is *independently* calculated, then collected

- Each irrep calculation is *almost independent* except at the end, when the contributions to the force constant matrix are calculated

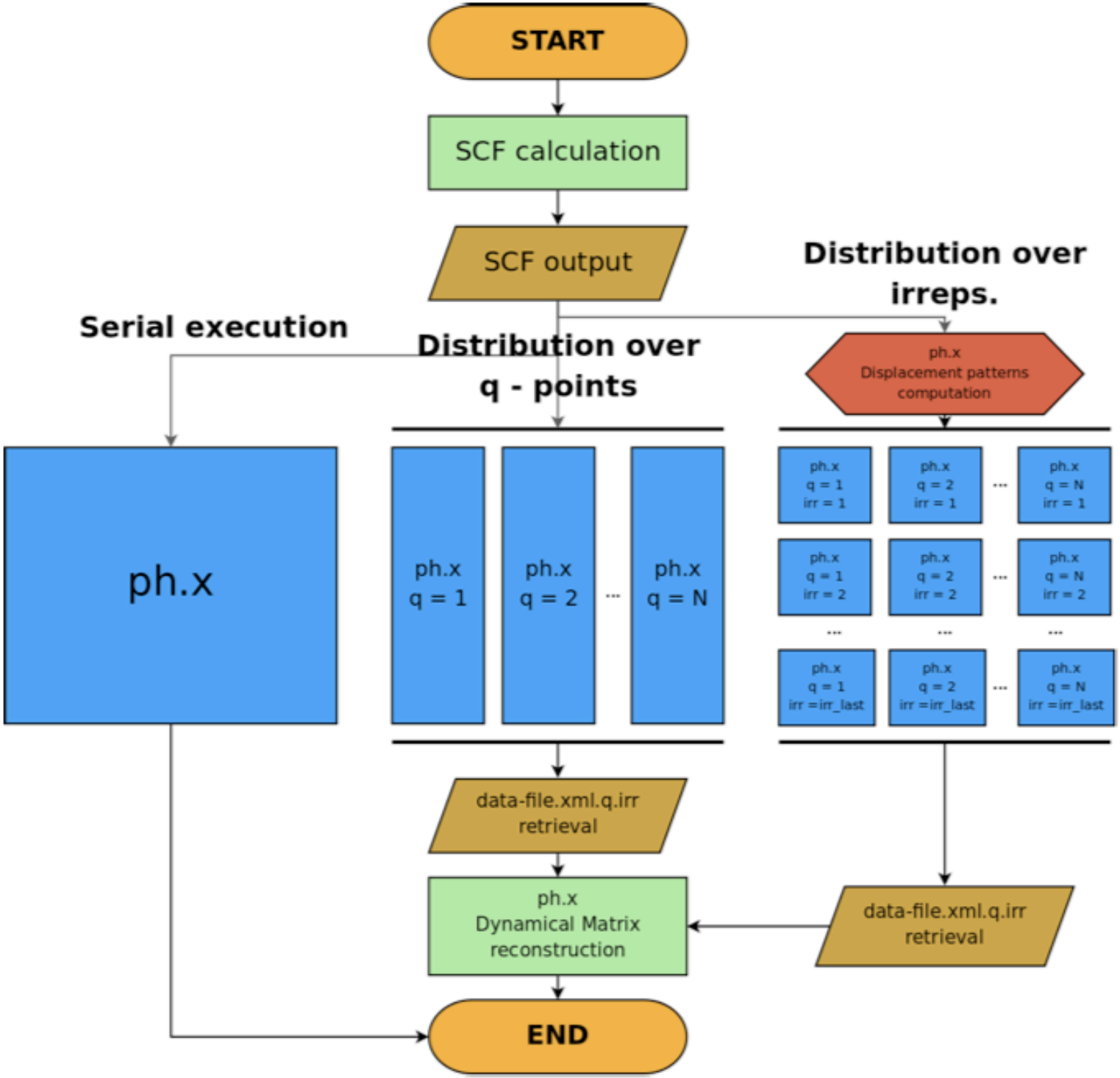
Perfect for execution on the GRID!

A realistic phonon calculation on the GRID

γ-Al_2O_3 is one of the phases of Alumina – a material of technological interest, with a rather complex structure. Can be described as a distorted hexagonal cell with a (simplified) unit cell of 40 atoms:



The calculation of the full phonon dispersion requires $120 \times n_q$ linear-response calculations, with $n_q \sim 10$, each one costing as much as a few times a self-consistent electronic-structure calculation in the same crystal: several weeks on a single PC.


START
SCF calculation
SCF output
Serial execution
Distribution over q - points
Distribution over irreps.
ph.x
Displacement patterns computation
ph.x
ph.x q = 1
ph.x q = 2
ph.x q = N
ph.x q = 1 irr = 1
ph.x q = 2 irr = 1
ph.x q = N irr = 1
ph.x q = 1 irr = 2
ph.x q = 2 irr = 2
ph.x q = N irr = 2
ph.x q = 1 irr =irr_last
ph.x q = 2 irr =irr_last
ph.x q = N irr =irr_last
data-file.xml.q.irr retrieval
ph.x Dynamical Matrix reconstruction
data-file.xml.q.irr retrieval
END

Practical implementation

Only minor changes needed in the phonon code, namely

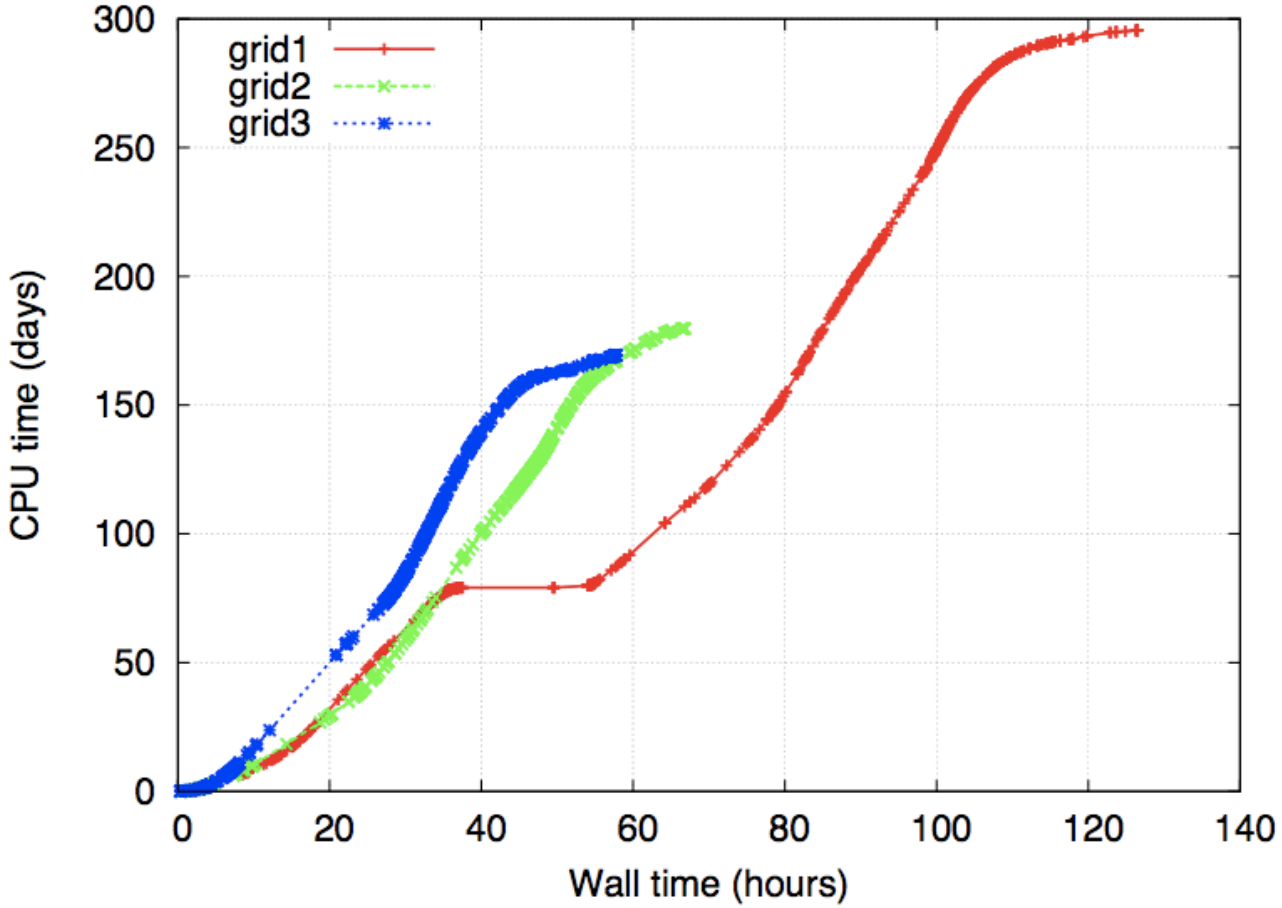
- possibility to run one $\mathbf{q}$-vector at the time (already there)
- possibility to run one irrep (or one group of irreps) at the time and to save partial results: a single row or a group of rows of the force constant matrix (a few Kb of data)

Python server-client application, written by Riccardo di Meo, takes care of dispatching jobs and of collecting results (uses XMLRPC).

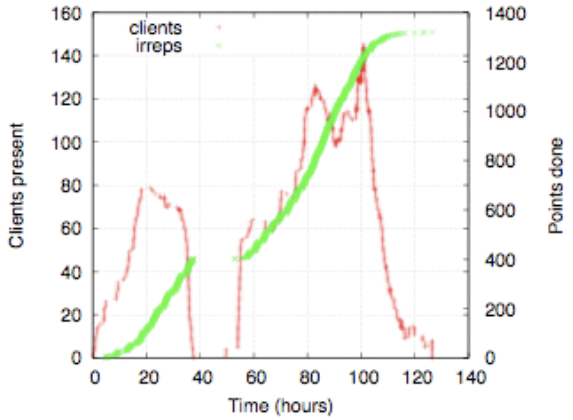
3000 jobs submitted in chunks of 500: clients contact back the server, receive input data and starting data files (hundreds of Mb).

Jobs lost in cyberspace ($\sim$ 60% of all contacted servers! of which 30-40% due to failure in downloading starting data files) are resubmitted.

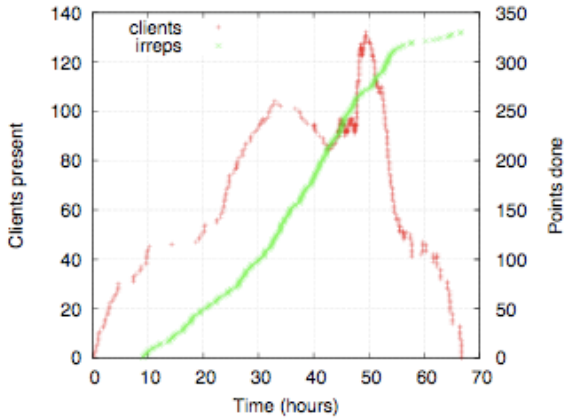
Execution on the GRID



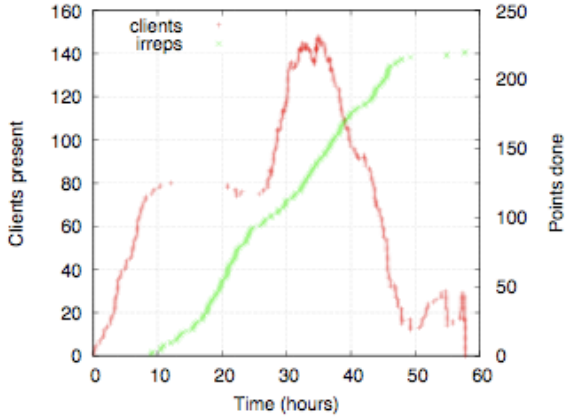
Resources spent on the GRID (`compchem` Virtual Organization): cumulative CPU time as a function of wall time, for three different distributions of irreps per CPU (1, 4, 6 resp. for grid1, grid2, grid3)



(a) grid1



(b) grid2



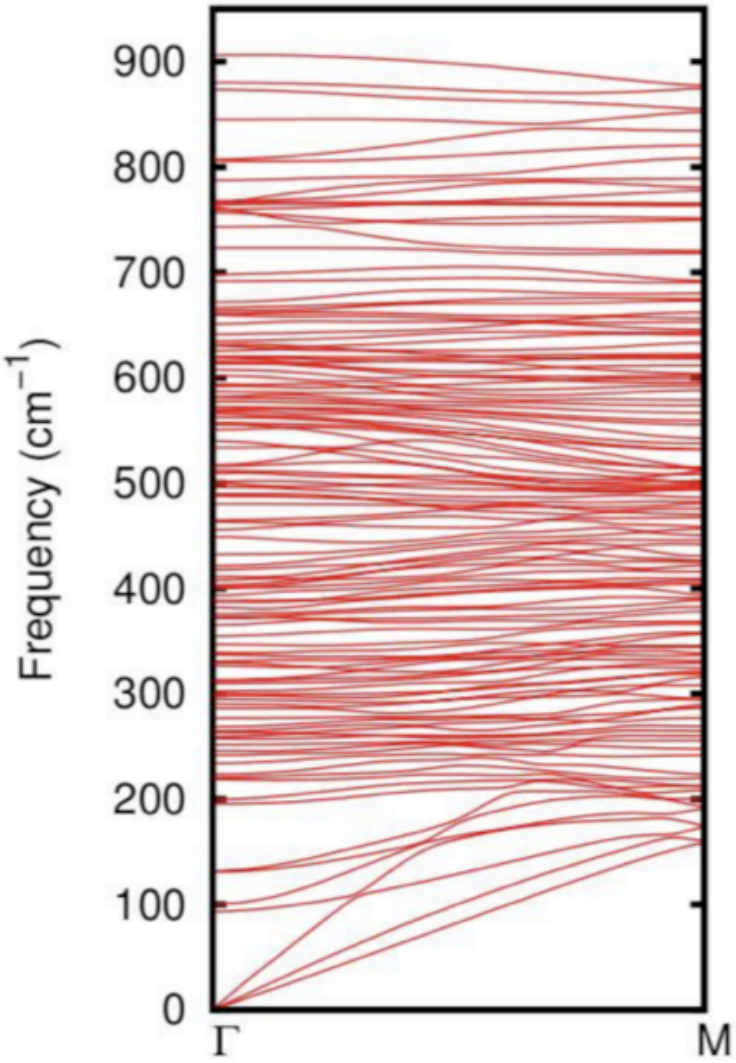
(c) grid3

Number of computed irreps and of clients present over time

Final result

Phonon dispersions, with TO-LO splitting, along special line $\Gamma - M$. $21 \times 1 \times 1$ **q**-vector grid ($n_{\mathbf{q}} = 11$).

Ultrasoft pseudos, 45Ry cutoff for wavefunctions and 360Ry for charge density. Brillouin Zone sampling with 221 Monkhorst-Pack grid. a=5.579Å, b=5.643Å, c=13.67Å. $\widehat{ab} = 120^{\circ}$, $\widehat{ac} = 90^{\circ}$, $\widehat{bc} = 89.5^{\circ}$.



Comments and Conclusion

- A realistic application of Quantum ESPRESSO to first-principle calculations at the nanoscale was demonstrated on the GRID

- Results produced in a relatively short time in spite of a rather high job failure rate: GRID can be competitive with conventional High-Performance Computers on much cheaper hardware

Needed for larger-scale calculations:

- Possibility to select parallel machines (with MPI), or large multicore machines (with OpenMP), to reduce RAM bottlenecks

Credits

- Thanks to Stefano Cozzini for arising in me the interest in GRID computing with Quantum ESPRESSO;
- to Riccardo di Meo and Andrea Dal Corso who did the real work;
- to Riccardo Mazzarello for help in the initial stages of this work;
- to Eduardo Ariel Menendez Proupin (U. de Chile, Santiago) who suggested phonons in γ-Al_2O_3
- ...and thank you for your attention!