The Quantum-ESPRESSO Software Distribution

P. Giannozzi
Università di Udine and Democritos National Simulation Center, Italy

COST Training School on Molecular and Materials Science GRID applications, Trieste, 2008/09/17

DEMOCRITOS
DEmocritos MOdeling Center for Research In aTOmistic Simulation INFM

Outline

- Democritos and scientific software
- What is Quantum-ESPRESSO?
- What can Quantum-ESPRESSO do?
- Quantum-ESPRESSO and the GRID:
  an experiment with realistic phonon dispersions calculation

Democritos and scientific software

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 several other institutions (CINECA Bologna, Princeton University, MIT Boston)

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

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

Motivations for Quantum-ESPRESSO

- to merge several pre-existing packages, used by different groups
- to provide access to several techniques whose usefulness has traditionally been hindered by the limited availability of software:
  - *linear response*
  - *ultrasoft PP/PAW*
  - *Car-Parrinello Molecular Dynamics*
- to provide a software distribution that is suitable for easy addition of new packages, developments, extensions, features. Such a goal requires:
  - a *modular* structure, easy to maintain, easily extensible
  - a *collaborative* environment: free/open-source license

Free software and GNU license

GNU (Gnu's Not Unix) General Public License (GPL) is probably the most common free-software license. Basically:

- The source code is available.
- You can do whatever you want with the sources, but if you distribute any derived work, you have to distribute under the GPL the sources of the derived work.

Advantages:

- Everybody – including commercial entities – can contribute.
- Nobody can "steal" the code and give nothing back to the community.

Quantum-ESPRESSO as a distribution

Quantum-ESPRESSO aims at becoming a *distribution* of packages, rather than a single, monolithic, tightly integrated package.

Main packages presently in the Quantum-ESPRESSO distribution:

- `PWscf`: self-consistent electronic structure, structural optimization, dynamics, linear-response calculations (phonons, dielectric properties), ballistic conductance

- `CP/FPMD`: variable-cell Car-Parrinello molecular dynamics

They share a common installation method, input format, PP format, data output format, large parts of the basic code.

Quantum-ESPRESSO as a distribution (2)

Additional packages:

- `GIPAW`: Gauge-Independent PAW method for EPR and NMR chemical shifts
- `W90`: Maximally Localised Wannier functions
- `atomic`: auxiliary code for PP generation
- `PWGui`: a Graphical User Interface for production of input files

plus other auxiliary codes for data postprocessing. Coming soon:

- `XSpectra`: simulation of X-ray adsorption spectra

More packages planned for addition.

Organization

The distribution is maintained as a single CVS (Concurrent Version System) tree. Available to everyone anytime via anonymous access.

Web sites:

- `http://www.quantum-espresso.org` (for Quantum-ESPRESSO in general)
- `http://www.pwscf.org` (specific to PWscf)

Mailing list:

- `pw_users`: used by developers for announcements about Quantum-ESPRESSO
- `pw_forum`: for general discussions (all subscribed users can post)

Technical characteristics (coding)

- mostly written in Fortran-90, with various degrees of sophistication (i.e. use of advanced f90 features) – no "dusty decks" any longer
- use of standard library routines (lapack, blas, fftw) to achieve portability – Machine-optimized libraries can be used if available
- C-style preprocessing options allow to keep a single source tree for all architectures
- various levels of parallelization via MPI calls, hidden into calls to very few routines – (almost) unified serial and parallel versions; parallel code can (usually) be written without knowing the details

Easy (or not-so-difficult) installation via the GNU utility `configure`

What can Quantum-ESPRESSO do?

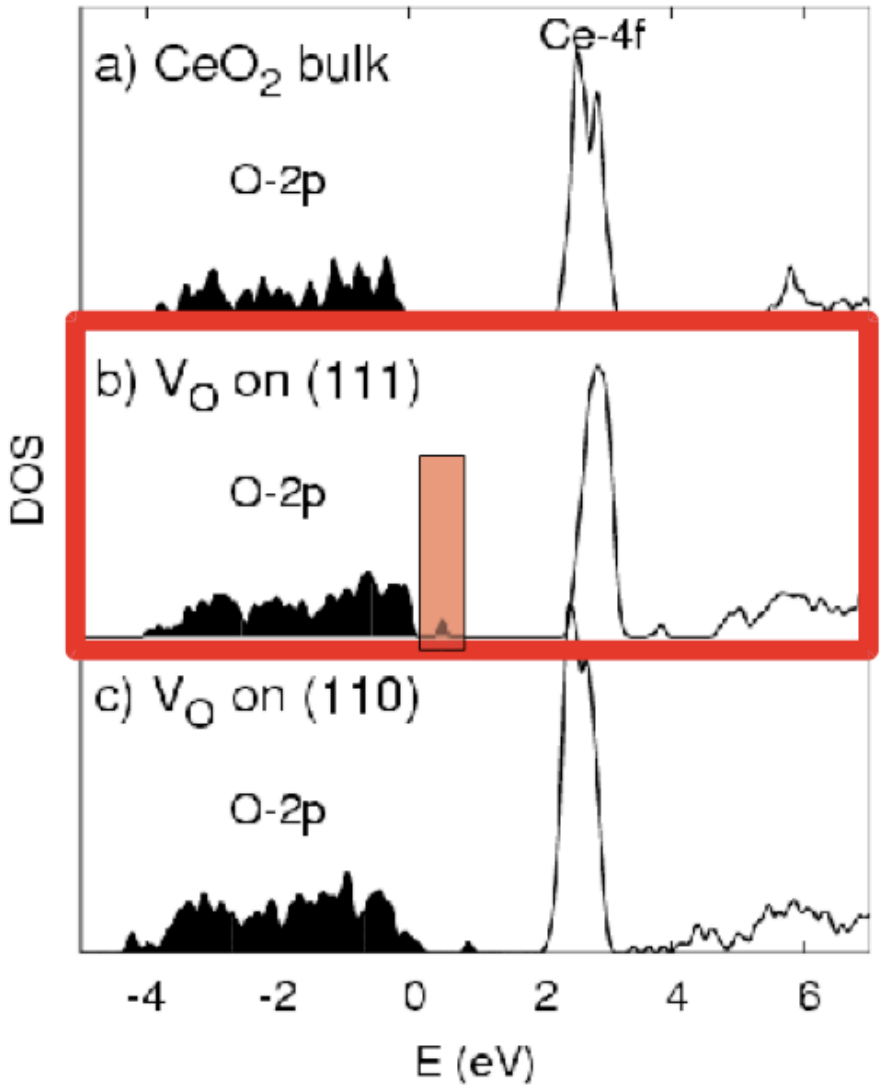
Quantum-ESPRESSO can perform *structural optimization* and find the corresponding *ground-state electronic structure* in

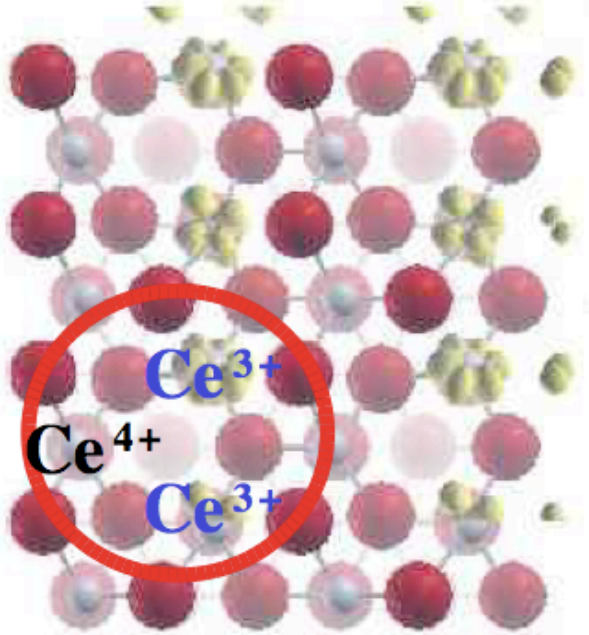
- molecules, clusters, surfaces, perfect and defective crystals
- insulators and metals
- magnetic as well as nonmagnetic systems

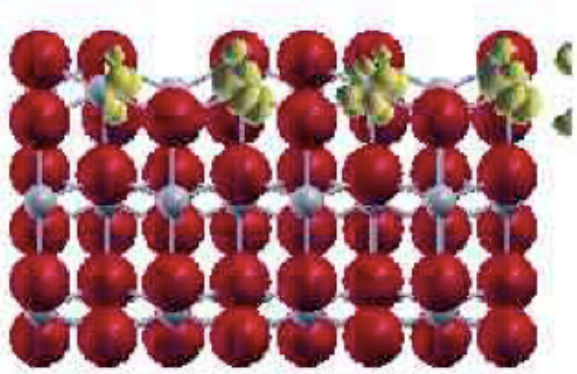
Quantum-ESPRESSO can perform more complicated calculations such as e.g. finding *transition states and minimum-energy pathways* for chemical reactions (next two slides courtesy of Stefano de Gironcoli, SISSA)

DEMOCRITOS

Reduced ceria surfaces

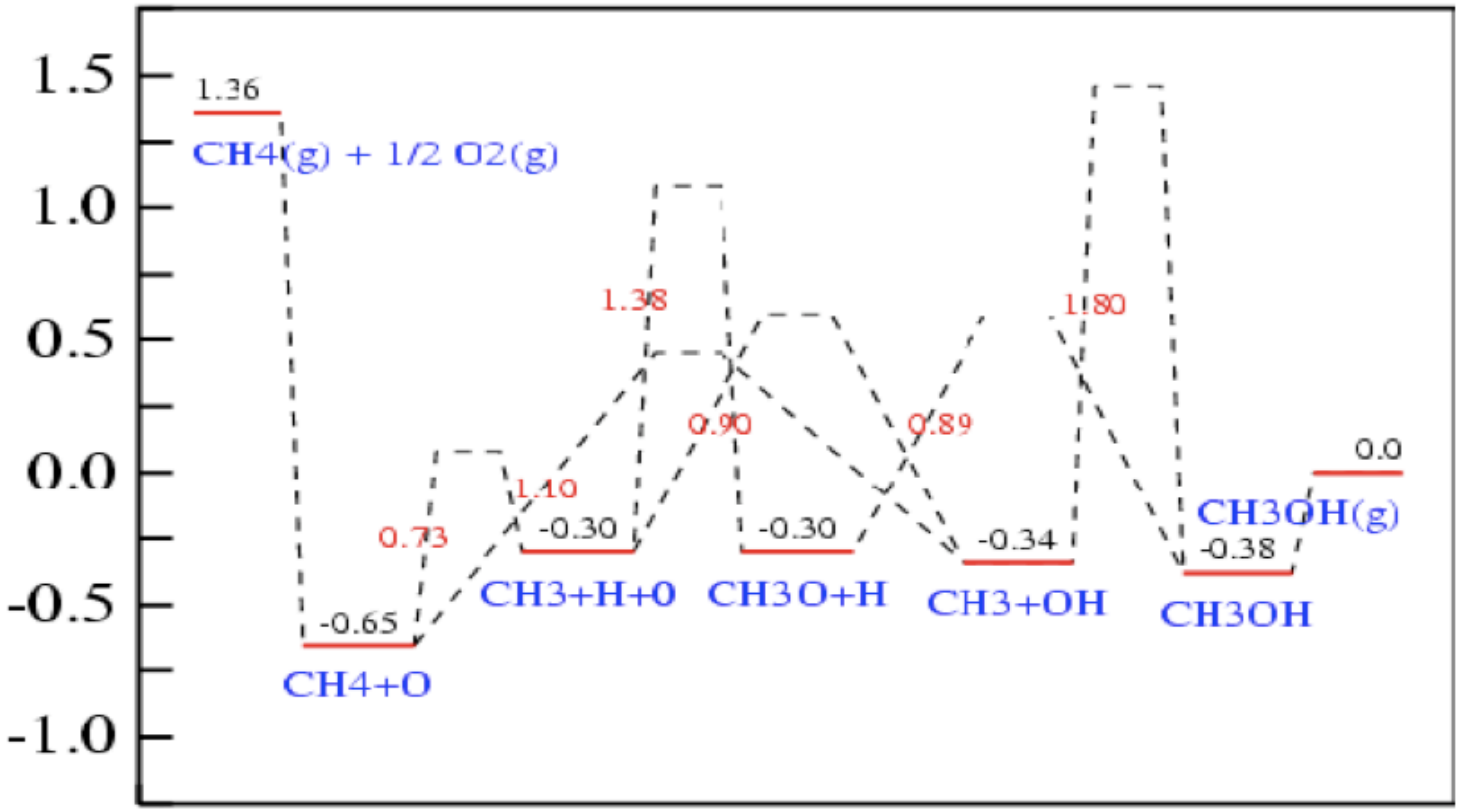






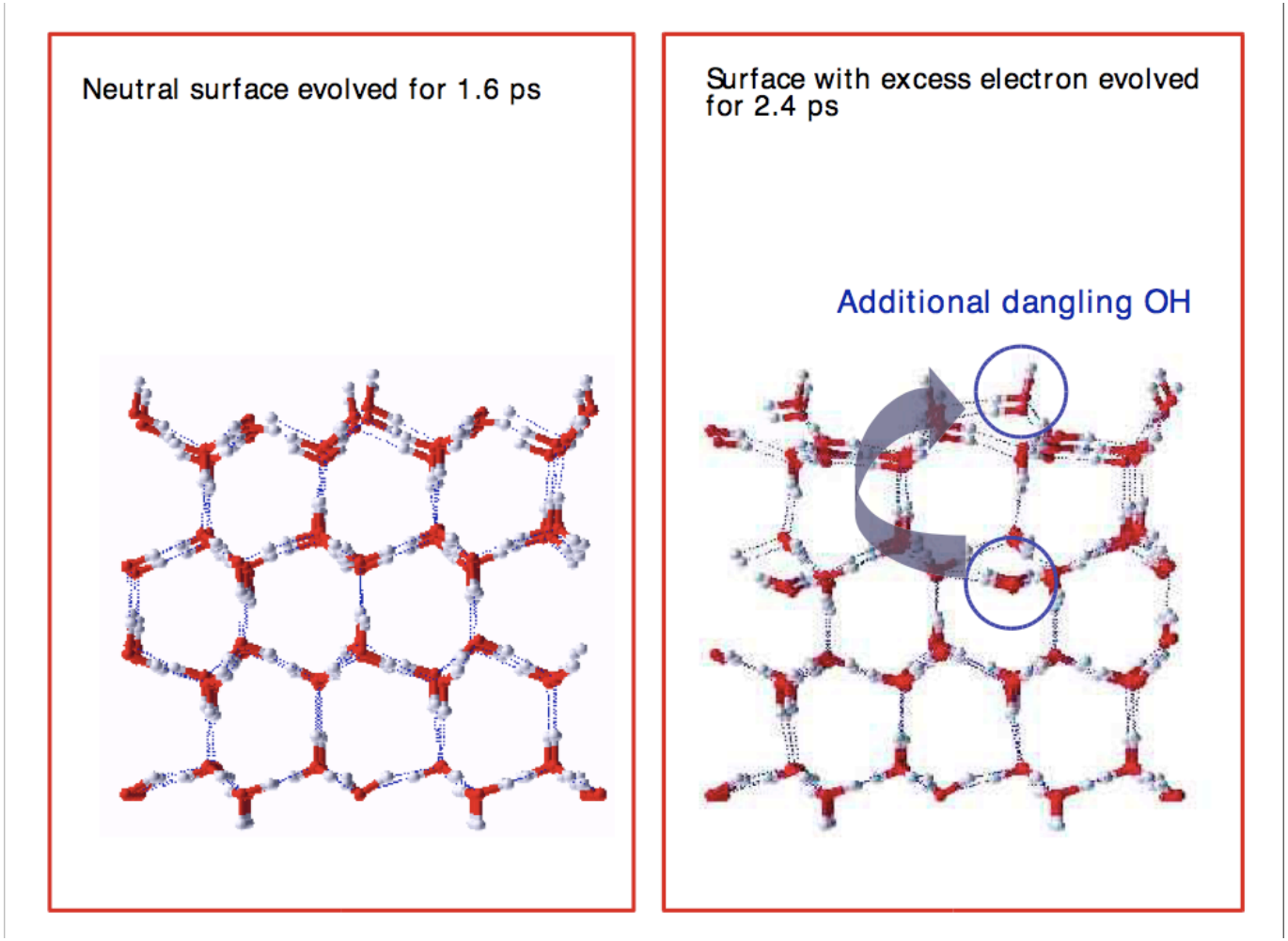
deGironcoli@sissa.it

Methane to Methanol pathway on Rh(111)

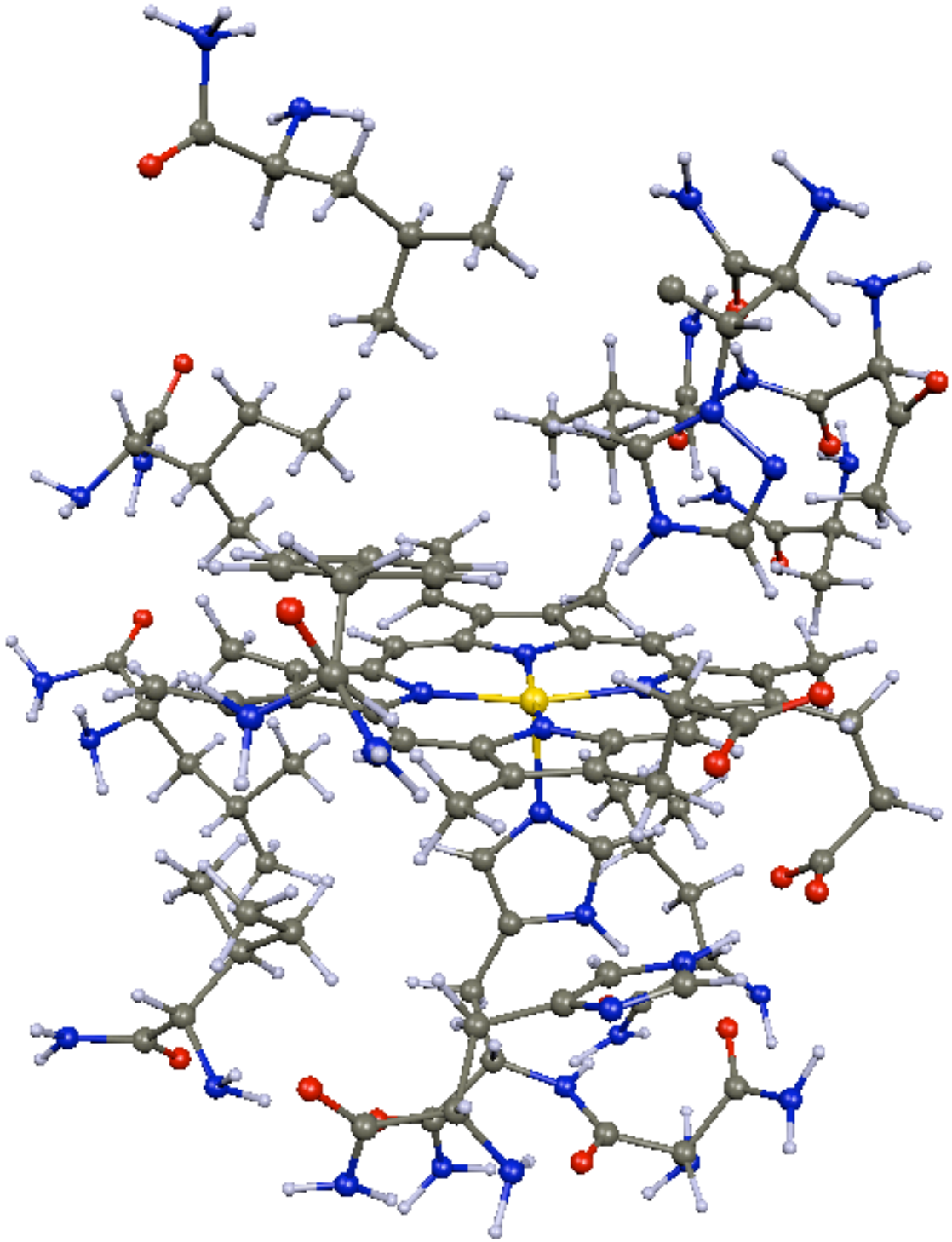


DEMOCRITOS

Quantum-ESPRESSO can perform *first-principle molecular dynamics*.



An example: reconstruction of the surface of Ice with an excess electron (courtesy of Sandro Scandolo, ICTP)



Quantum-ESPRESSO can deal with complicated organometallic compounds and models of active sites in proteins.

An example: 332-atom model of a heme site, including all atoms in a sphere within 8 Å radius centered around the Iron atom of the porphyrin

PG, F. de Angelis, R. Car, JCP 120, 8632 (2004)

Computer requirements

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 PWs, of electronic states, of **k**-points needed
- *desired results:* increasing effort from simple self-consistent (single-pont) 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 $\sim$ 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*

Excessive RAM requirements for single PCs leave no choice, but main factor pushing towards parallel machines is excessive CPU time.

Quantum-ESPRESSO and High-Performance Computing

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

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

- porting to the BlueGene architecture (BG/L and BG/P)
- CP has been run on up to 4800 Cray XT4 processors, reaching 10 TFlops, for a 1532-atom porphyrin-functionalized nanotube

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:

- 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);
- transition path search (NEB);
- **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

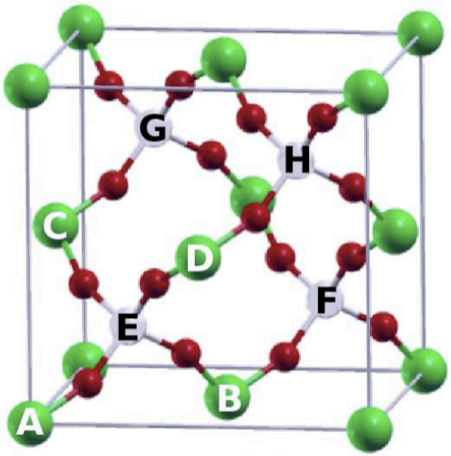
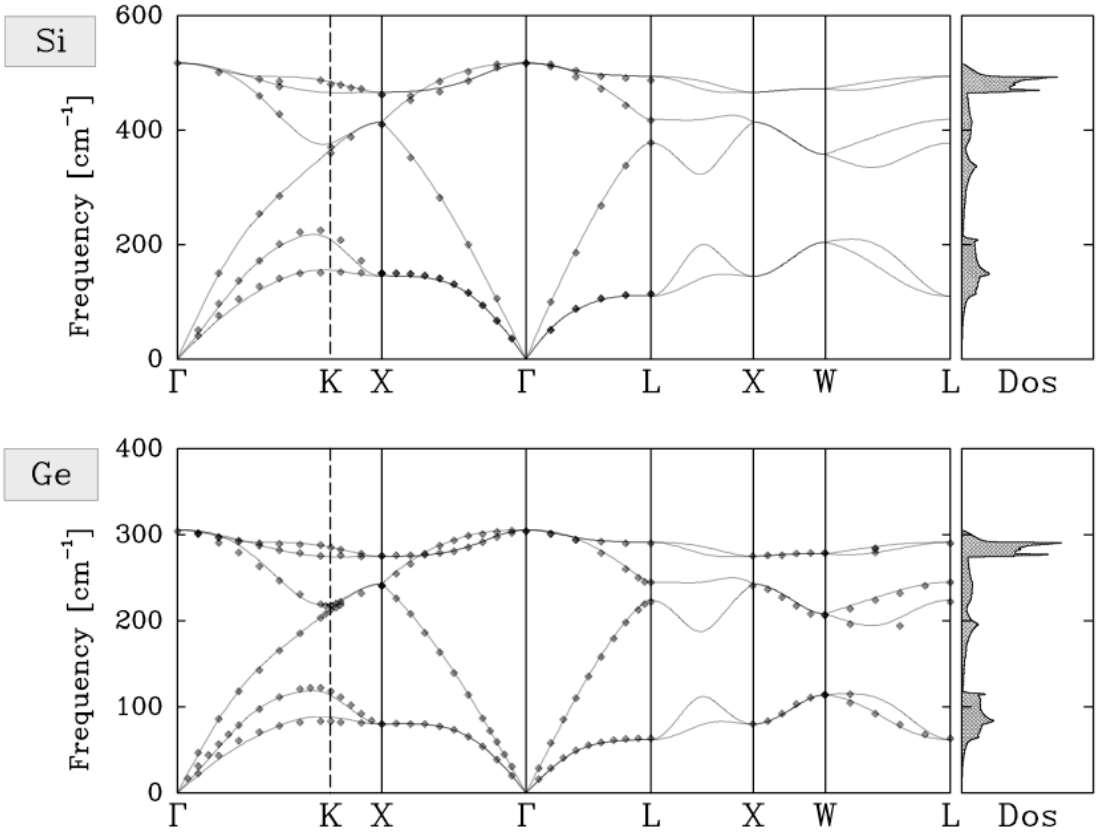


Figure 1: Example of a conventional cubic unit cell for I-42d cristobalite. The small dark spheres (red online) indicate O atoms, while the large grey (green online) and small light (white online) spheres represent Si and C atoms respectively. The positions of C/Si atoms are labeled A to H (see also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

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 4)

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 collected

Perfect for execution on the GRID!

Technicalities

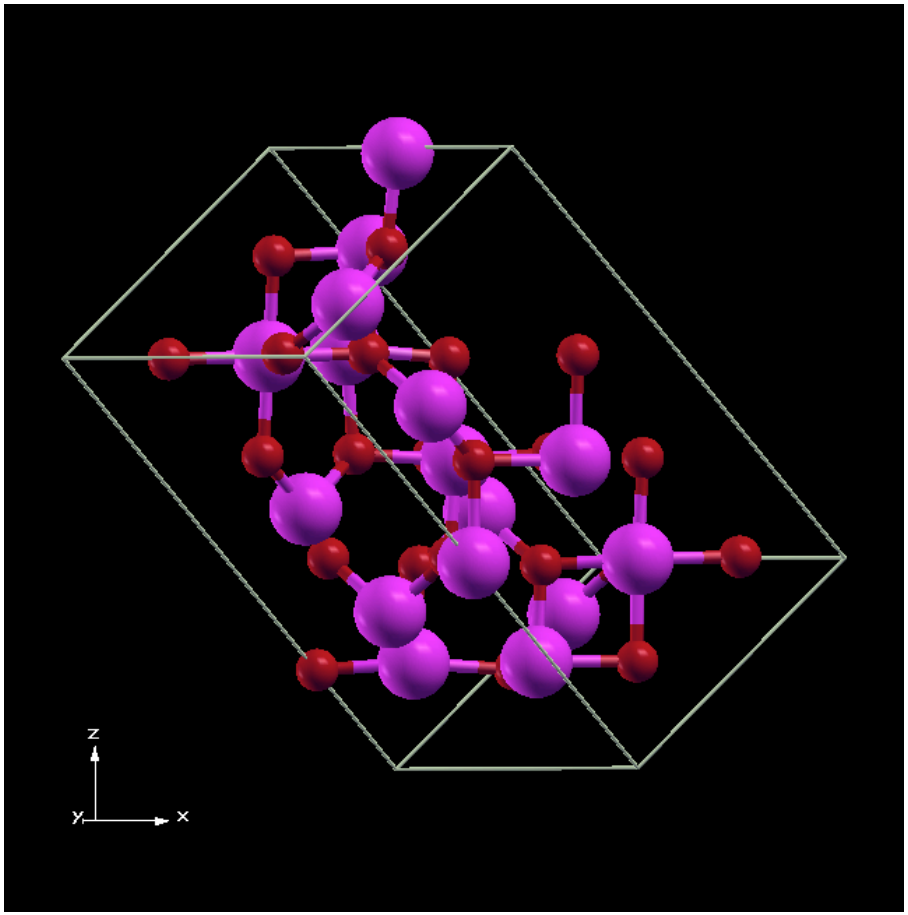
Only minor changes needed in the phonon code, namely

- possibility to run one **q**-vector at the time (already there)
- possibility to run one irrep at the time and to save partial results (a single row of the force constant matrix)

Server-client application (written by Riccardo di Meo) takes care of "dispatching" jobs and of "collecting" results

A realistic phonon calculation on the GRID

γ-Al_2O_3 has a (simplified) unit cell of 40 atoms, i.e. $120 \times n_q$ linear-response calculations, $n_q \sim 10$: a few weeks on a single PC.



Currently running on a lot of machines!

Credits

- Thanks to the many people have contributed to Quantum-ESPRESSO
- to Stefano Cozzini for arising in me the interest in GRID computing with Quantum-ESPRESSO
- to Riccardo di Meo, Riccardo Mazzarello and Andrea Dal Corso who did the real work
- to Eduardo Ariel Menendez Proupin (Santiago, Chile) who suggested phonons in γ-Al_2O_3
- ...and thank you for your attention!