Present and Future Computing Requirements

Rational Catalyst Design for Energy Production

Andreas Heyden

Department of Chemical Engineering

University of South Carolina, Columbia

NERSC BES Requirements for 2017

October 8-9, 2013

Gaithersburg, MD

Project Description

- Molecular understanding and design of heterogeneous catalysis at solid-liquid interfaces (Applications: Biomass-to-Fuels, Biomass-to-Chemicals)
  - Multi-scale, mixed-resolution modeling techniques
  - Planewave DFT coupled with
    - Continuum solvation models
    - Classical molecular dynamics

Other Projects

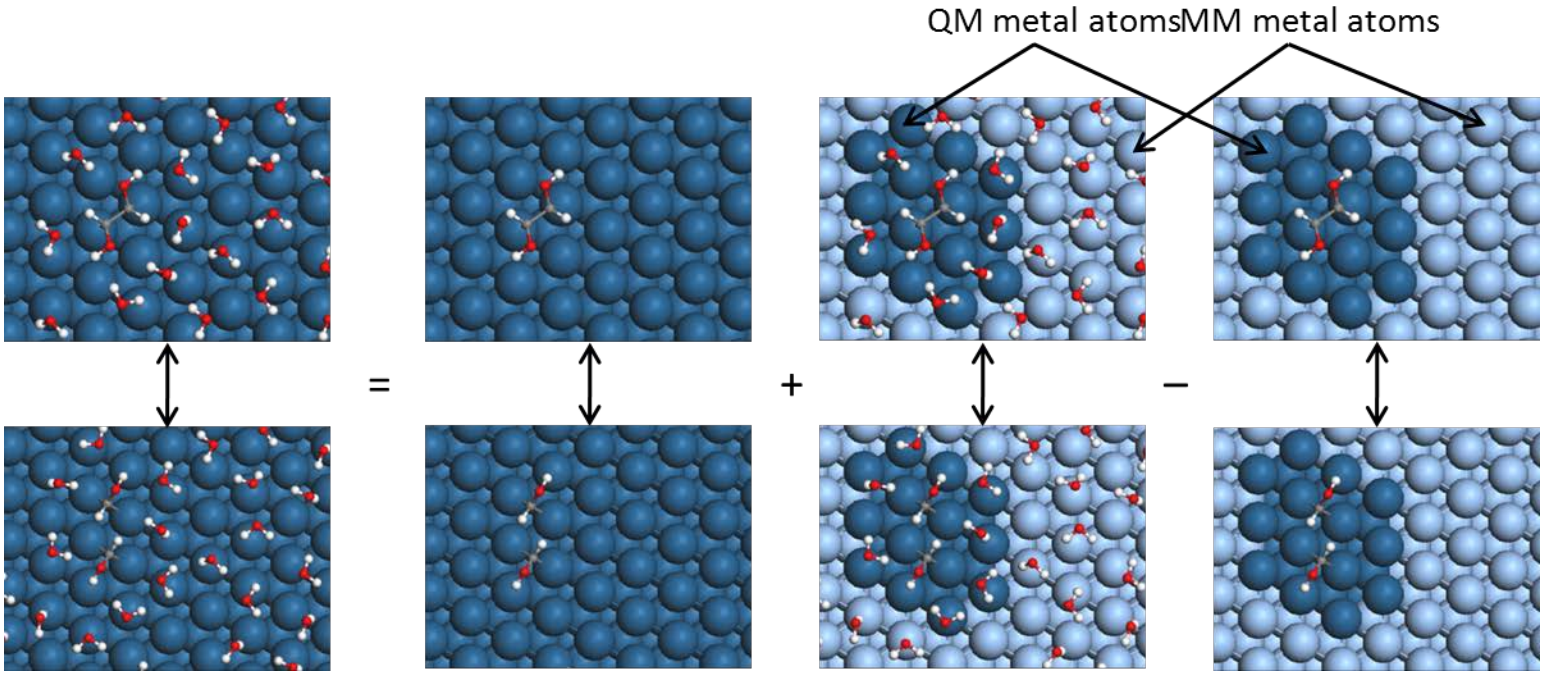
- Molecular understanding and design of heterogeneous catalysis at solid-solid interfaces (Applications: Water-Gas Shift at metal-oxide interface sites)
- Molecular understanding and design of electrodes in solid oxide fuel/electrolyzer cells (Applications: Natural gas utilization / Renewable energy storage)
- Uncertainty Quantification in computational studies for improving the connection to experiments

Scientific Objectives for 2017

- Framework for simulating heterogeneously catalyzed reactions in liquid-phase environments
  - Successive improvement of the description of the effect of a liquid environment on reaction equilibria and kinetics
    - Planewave DFT (reaction network)
    - Implicit solvation (preliminary screening)
    - Explicit solvation (important elementary steps)
    - “Few” solvent coordinates part of reaction coordinate (rate-controlling steps)

Scientific Objectives for 2017

Solvent effects



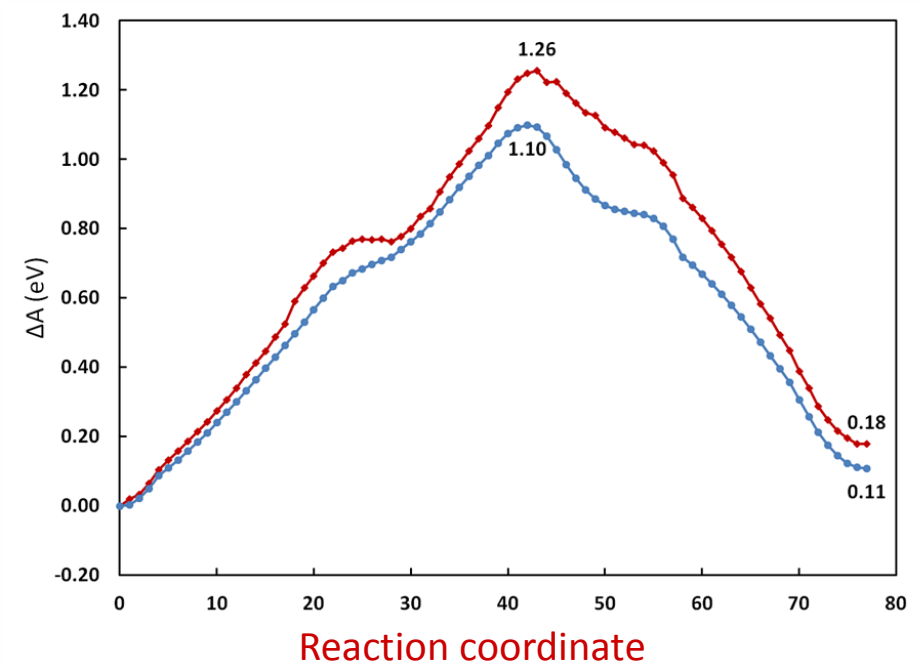
$$\Delta E_{surface/water} = \Delta E_{surface} + \Delta E_{cluster/water} - \Delta E_{cluster}$$

$$\Delta A_{surface/water} = \Delta E_{surface} + \Delta A_{cluster/water} - \Delta E_{cluster}$$

M. Faheem *et al., J. Phys. Chem. C,* 2012, **116**, 22458-22462.

Scientific Objectives for 2017

Validation of implicit solvation by explicit solvation and free energy perturbation



$C_2H_4O_2 \rightarrow 2CHOH$ on Pt(111)

(eV)	Activation Barrier	Reaction Energy
Planewave	1.10	0.11
Implicit Solvation (COSMO)	1.15	0.17
Implicit Solvation (COSMO-RS)	**1.26**	**0.28**
Explicit Solvation	**1.26**	**0.18**

Computational Strategies

- Integration of DFT and MD codes
  - DFT calculations in planewave basis (VASP)
  - DFT calculations in Gaussian-type basis (TURBOMOLE)
  - Classical molecular dynamics (DLPOLY)
  - Interfaces, automated input generation and output processing (FORTRAN programs and BASH scripts)

Computational Challenges

- Convergence of Gaussian-type orbitals for metal clusters + determination of spin states on metal clusters
- Scaling of DFT calculations
- Large number of tasks
- Maximum wall-clock in **high-throughput** queue (on Carver not just Hopper)

Current HPC Usage

Period	Resource	Core-Hours used (million)
January 2013 – December 2013	NERSC (Carver/Hopper)	3.5 (allocation)
January 2013 – December 2013	NERSC (Edison)	1.0 (estimated)
October 2012 – September 2013	XSEDE (Kraken/Stampede/Steele)	3.0
October 2012 – September 2013	PNNL (Chinook)	3.7
January 2013 – December 2013	USC (ACM/Planck/Maxwell)	>1.0

Current HPC Usage

Job Type	QM	MM	QM/MM
Number of cores	8-64	8-16	16-48
Run time (hours)	12-48 (240 for TS)	4-8	12-100+
Jobs per year	>10 k	100-200	100-200
Total data read/written	1 GB	10-50 GB	1-2 TB
I/O bandwidth (per minute)	Write: only at end	Write: 100 MB	Read+Write: 2 GB

HPC Requirements for 2017

- Expected usage: >20 million core-hours
- General shift towards high-throughput computing
- Considerably larger share of QM/MM jobs
- No significant changes expected in HPC requirements on per job basis (already relatively large systems, DFT functionals get more complex, however, processors also get faster)

Strategies for New Architectures

- No plan for modification of any code
- No constraint for switching, if modified codes are available

Recommendations for System Configuration

- TURBOMOLE requires Cluster Compatibility Mode (CCM) on Hopper/Edison.
- For hybrid QM/MM jobs, VASP, TURBOMOLE, and DLPOLY must all run together.
- Before CCM was available (April 2012), Carver was the only choice on NERSC.

Summary

- What new science results might be afforded by improvements in NERSC computing hardware, software and services?

  Coming closer to realizing the aims of the White House MGI, e.g., high-throughput screening and rational design of complex catalytic systems relevant for biomass conversion to fuels and chemicals, energy storage, and energy independence

Summary

- NERSC generally refreshes systems to provide on average a 2X performance increase every year. What significant scientific progress could you achieve over the next 5 years with access to 32X your current NERSC allocation?

Currently our eSMS – explicit solvation method for solid surfaces is ~2 orders of magnitude slower than gas phase computations. As a result, a 32X speed increase would permit the routine application of eSMS for rational catalyst design studies

Summary

- What "expanded HPC resources" are important for your project?
- Recommendations on NERSC architecture, system configuration and the associated service requirements needed for your science

The availability of a cluster system such as Carver & Stampede (UT) with a long max. wall clock time benefits many quantum chemistry codes that currently do not scale to hundreds or thousands of processors