

Modeliranje nanostruktura na gridu - iskustva i izazovi

Dr. sc. Vjeran Gomzi

Zavod za primijenjenu fiziku,
Fakultet elektrotehnike i računarstva Sveučilišta u Zagrebu,
Zagreb

CRO-NGI : putem projekata N. Došlić, D. Smith (IRB), J. Sabolović (IMI)

Biofizika/teorijska kemija: Gaussian - većina uporabivog rada do sada.
(geometrije, prijelazna stanja, prijenos elektrona, pokušaj modeliranja reakcija, pobuđena stanja, ionizacijski potencijali i afiniteti za e-);

Metalne nanočestice: ReactiveFF – MM/MD + dinamička premještanja
vezna/protuvezna orbitala
primijenjen na modeliranje nanoklastera (vlastita iterativna rutina);

Periodične strukture: abinit -struktura i IR/raman (fononski) spektri spinela /u planu.
osnovna geometrija da, RF račun (još) ne.

Outline izlaganja:

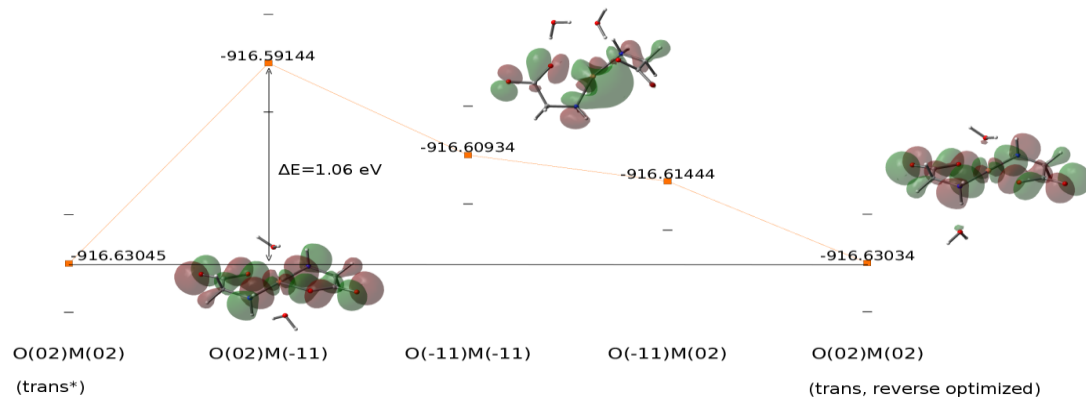
1. - primjeri ovog
2. - primjedbe i iskustva o radu na CRO-NGI-ju
3. - zaključci
4. - prijedlozi, ideje, molbe

1. primjer:

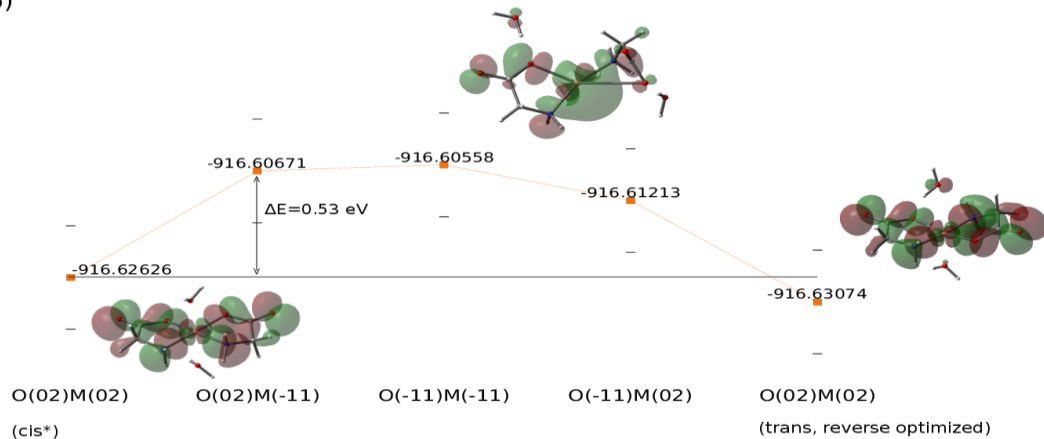
Proučavanje trans-cis izomerizacije $\text{Cu}(\text{gly})_2$ (projekt dr. J. Sabolović, IMI)

račun EA i IP vode i bis-glicinato bakra(II) da se prouči utjecaj prijelaza naboja na okolinu na trans-cis izomerizaciju spoja

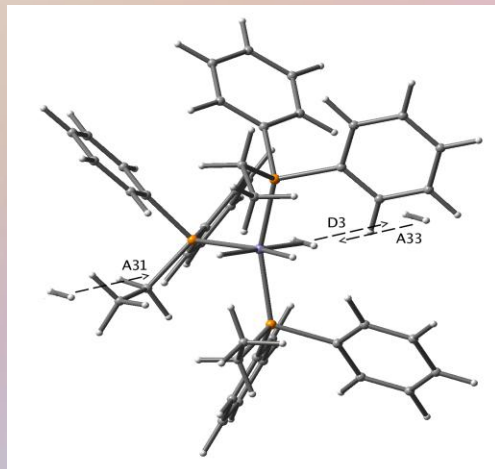
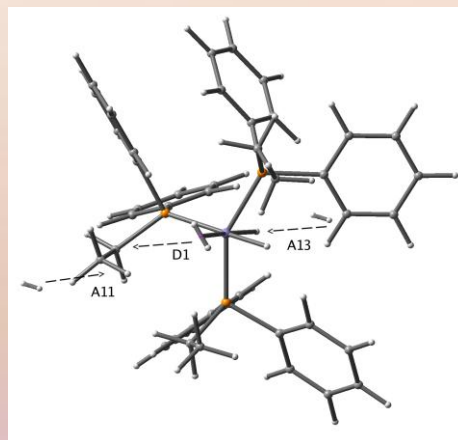
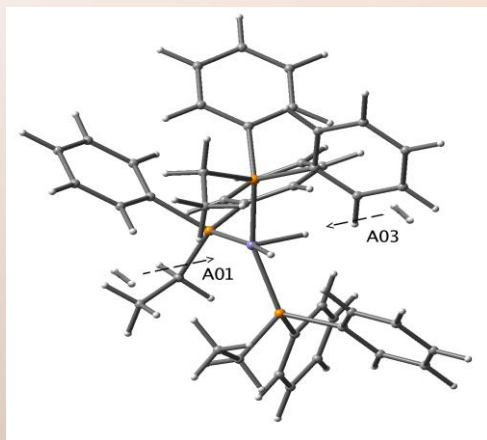
(a)



(b)



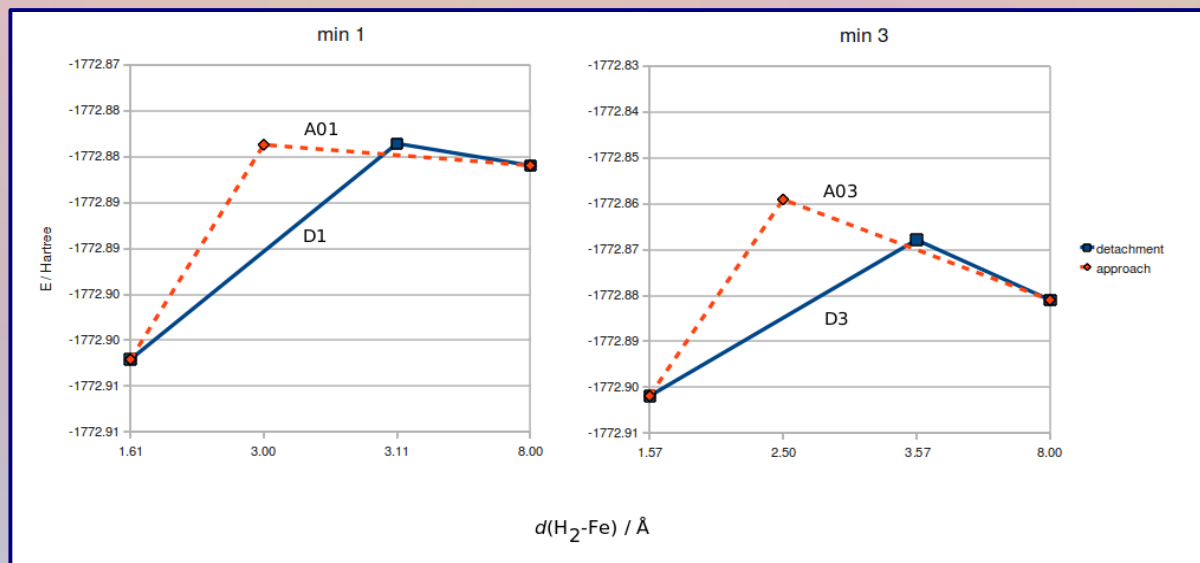
* Optimized isomer structures from previous work [9].



2. primjer:

Istraživanje vezanja molekule H_2 na $Fe(H)_2(H_2)(PEtPh_2)_3$ (Nastavak rada temeljenog na projektu dr. N. Došlić, IRB)

Proučavanje energijskih učinaka vezanja molekule H_2 na Fe-strukturu. Preliminarni rezultati na sl. vidljivi kao histereza.



ReaxFF

Uobičajeno se u MM/MD parametrizira polje sila preko duljina veze pojedinih atoma, kuta i torzijskog kuta.

Parametrizacija je ovisna o obliku veze (π , σ , ...).

Reactive FF – reakcijsko polje sila - očekuje se dobar opis reakcija jer za razliku od standardnog polja sila, ne pretpostavlja oblik (π , σ , jednostruka, višestruka) pojedine veze unaprijed.

ReaxFF

```
#!/bin/bash
# setting up directories and starting the main script (exe4clustCNGI)

# extract all files to current temp dir
# tar -xjpf inputfilesG.tar.bz2

# directories and files for reax operation

#TO DO/DEBUG - which folders? which files? in changes at reax clu-genea
```

```
mkdir src
cp reac src
cp ffopt.o src
cp poten.o src
cp shanno.o src
cp vibra.o src
cp minEg.out src
cp geo4clug.out src
cp blas.o src
cp version.o src
cp *.blk src
cp bgf2xyzig.out src
cp divedit.out src
cp infilezadivcf.dat src
cp bin.DIVCF.1.3.2.linux32.intel src
# cp bin.DIVCF.1.3.2.linux64.intel src !
```

ReaxFF izvršni kod

određivanje nove početne strukture,
nalaženje strukture minimalne energije

ReaxFF izvršni kod

priprema podataka i obrada rezultata
pretrage
različitih struktura (grupiranje sličnih
struktura)

```
# directories and files for exe4clusta@CNGI multiple structures calc's
```

```
mkdir cluster
```

```
# starting the main script
```

```
./exe4clug
```

pokretačka skripta koja raspodjeljuje podatke,
pokreće izvršne programe, kreira i briše
mape,

valjana kompilacija (gfortran)
na bilo kojoj 64-bit linux mašini

sh/bash skripta – jednostavno
pronalaženje pogrešaka

**Ni ReaxFF niti rekurzivni kod
trenutno ne podržavaju
višeprocesorsko računanje**

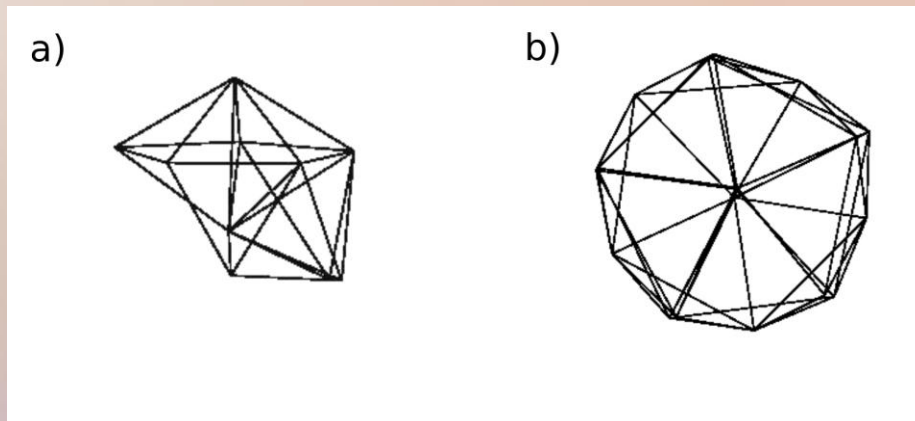
```
Executable=exe4CNGIG
TransferExecutable=True
should_transfer_files=yes
WhenToTransferOutput=ON_EXIT
transfer_input_files=inputfilesG.tar.bz2
transfer_output_files=MultiReaxG.tar.bz2,debug.log
Log=condorG.log
Output=stagr.gt2.out
Error=stagr.gt2.err
universe=grid
environment = "PE_MODE=single"
globusrsi = (jobType=single)(count=1)
x509userproxy = /tmp/x509up_u556
grid_resource = $$ (gridtype_and_args)
requirements = TARGET.Machine == "ce.gradri.cro-ngi.hr"
queue
```

ReaxFF

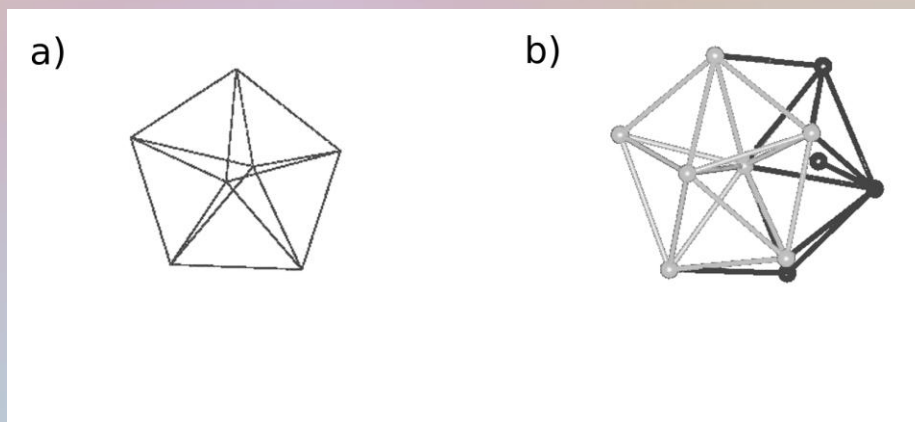
Primjer:

Modeliranje strukture metalnih nanočestica

Proučavanje strukture i energije metalnih nanočestica nastalih dodavanjem po jednog atoma metala u prethodno optimiziranu strukturu. Rekurzivna primjena ReaxFF koda.



Ni a) 10-atomska i b) 19-atomska struktura predviđena ReaxFF računom



a) Pt 7-atomska struktura najniže energije predviđena ReaxFF računom dodavanjem po jednog atoma u optimiziranu strukturu.

b) Pt 11-atomska struktura najniže energije predviđena ReaxFF računom. Svijetla podstruktura pokazuje način nastanka ove strukture iz 7-atomske strukture. Generirajući strukturu dodavanjem po jednog atoma, konačno se za 13-atomsku strukturu dobiva očekivana ikozaedarska struktura.

abinit

“Open-source” projekt

Primjena za – DFT račune molekula i struktura s periodičnim rubnim uvjetima, optimizirana i dinamička svojstva struktura u čvrstom stanju

Primjer:

Proučavanje fononskog odziva (predviđanje IR/Raman spektara) spinela

(suradnja s N. Biliškov, A. Knežević, IRB)

Optimizacija periodične strukture s 14 ili više atoma/j.ć.
Predviđanje fononskih spektara takvih struktura.

Spinel

kristalna tvar, opća formula: AB_2X_4 , A – divalentni kation, B – trivalentni kation, X – anion

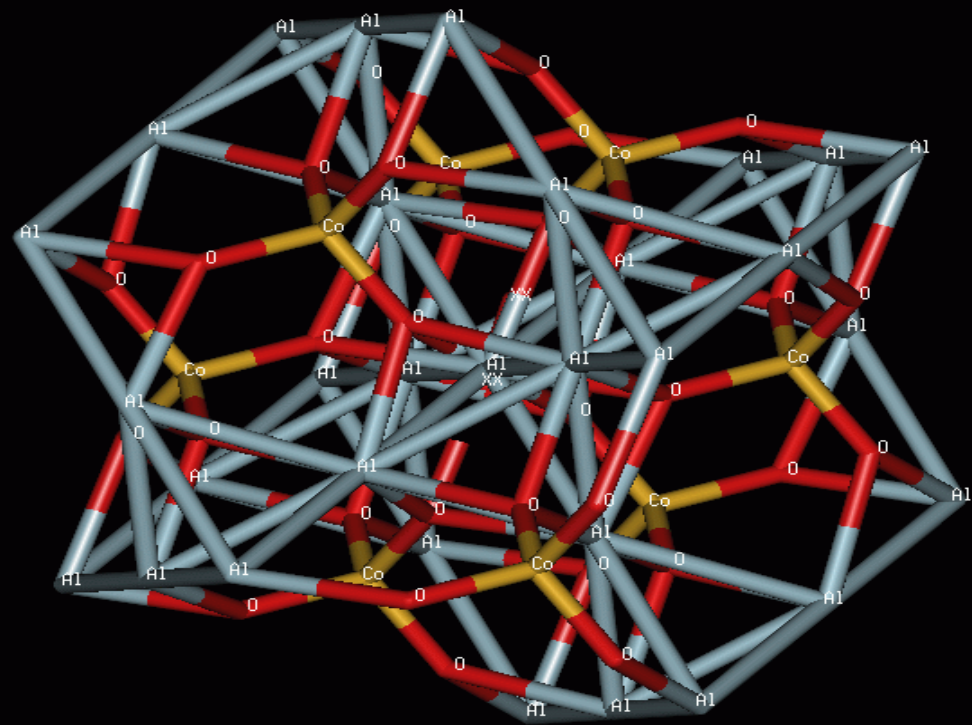
moguća zamjena A i B kationa u mjestima određene simetrije u kristalu, o čemu jako ovise svojstva materijala.

raspodjela kationa u mjestima (tzv. faktor inverzije) ovisi o temperaturi i tlaku

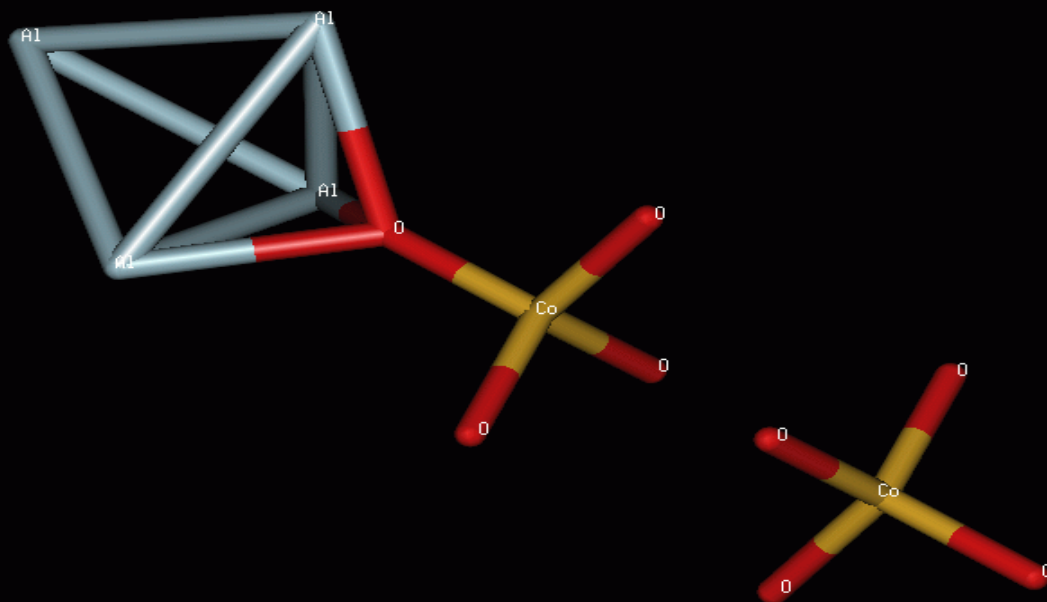
bit - energijski najpovoljnija struktura je s pojedinim faktorom inverzije - to karakterizira pojedine TD uvjete / karakterizacija pomoću fononskih spektara

ponavljanje poznatog računa (Tielens i sur., J. Phys. Chem. B 2006, 106, 988-995.) / teorijsko “baždarenje” tj. učenje zanata

zahtjevni računi - plane-augmented waves - oscilatorne kontinuirane f-je - rešetka, + opis lokaliziranim članovima za bolje modele oko centara



Ovisno o faktoru inverzije koji se želi modelirati, jedinična ć. sadrži 14, 28 ili 42 atoma



***.files**

```
TielensCoAl2O4-gp.in  
TielensCoAl2O4-gp.out  
abi  
abo  
tmp  
Co.atompaw.gga.paw  
13al.paw  
08o-gga.paw
```

<pokretačka skripta za condor_submit>

```
Executable=ruabinit.sh  
#Arguments=<TielensCoAl2O4.files >&log  
Log=condorG.log  
Output=run.$(cluster).out  
Error=run.$(cluster).err  
transfer_input_files = TielensCoAl2O4-gp.in, 08o-gga.paw, 13al.paw, Co.atompaw.gga.paw,  
TielensCoAl2O4.files  
transfer_output_files = TielensCoAl2O4-gp.out, log  
should_transfer_files = yes  
WhenToTransferOutput = ON_EXIT  
# ovo je nuzno kako bi se posao izveo na jednom cvoru  
x509userproxy = /tmp/x509up_u556  
grid_resource = $$ (gridtype_and_args)  
environment = "PE_MODE=single"  
globusrsi = (jobType=single)(count=8)  
requirements = TARGET.Machine == "ce.etfos.cro-ngi.hr"  
queue
```

<drugi dio pokretačke skripte za abinit run na CE - Emir>

```
#!/bin/sh
currDir=`pwd`
cp TielensCoAl2O4-gp.in 08o-gga.paw 13al.paw Co.atompaw.gga.paw TielensCoAl2O4.files $TMPDIR/
cd $TMPDIR
sed -i s:TielensCoAl2O4-gp.out:${currDir}/TielensCoAl2O4-gp.out: TielensCoAl2O4.files
mpirun -np 8 -machinefile $TMPDIR/machines /usr/local/bin/abinip < TielensCoAl2O4.files >& $currDir/log
ls
mv log $currDir/
```

Pomoć -

Igor Lukačević za abinit
Darko Babić za G03/G09
Emir Imamagić za MPI

(zajednica korisnika nije prevelika pa se nalaženje pomoći često sastoji od komunikacije s drugima koji "vrte" slične vrste računa)

Modeliranje nanostrukture na gridu



iskustva i izazovi

Zašto računati na gridu?

G03/G09 - Porast memorijskih, posebno *scratch disk* i procesorskih zahtjeva s brojem stupnjeva slobode (broj atoma tj. broj elektrona u simulaciji i baza koja se rabi. Većina naprednih metoda traži ovisnost o gradijentu potencijala. Najzahtjevniji koraci su transformacije integrala.)

standardna primjena na - 100 atoma s jednim teškim (Fe, Cu npr.) reda veličine tjedan-dva s laganom metodom (B3LYP tipično)

MP2 računi s teškim elementom i oko 40 atoma - baza od oko 600 baznih f-ja (1000 primitivnih gaussian-f-ja) traje više mjeseci.

Moguća poboljšanja - prije svega SW - bolja uporaba raspodijeljene okoline tj. velikog prostora na disku mogla bi skratiti vrijeme računa?

Poboljšanja na gridu - problem s MP2 računima - raspodjeljivanje poslova na čvorove s nedostatnim prostorom na disku (Route defs tj u G03 environment varijablama)

Zašto računati na gridu?

ReaxFF - "laki" računati. Kompajlirano na gridu: gfortran, 64 bit okruženje, robustan source SW. Na žalost nije razvijen za rad u multiprocesorskom okruženju pa je to moguće poboljšanje.

Razvijene rutine i fortranski moduli koji rekurzivno pozivaju ReaxFF SW pri tome dodajući ili oduzimajući po jedan atom iz nanostrukture. Primjena na proučavanje rasta i strukture nanočestica (Au, Ag, Pt npr.). Primjena CRO-NGI okruženja prije svega zbog količine podataka te uporaba kao "dedicated" računala - izbjegavanje stalnog opterećenja računala na kojem želimo još nešto raditi kroz dugi period, bez obzira sto su računati na gridu na samo 1 procesoru. Moguće bi bilo i samo rekurzivnu proceduru razviti za rad na gridu.

Ako nekog zanima više o ovom (reactive FF koncept i ideja), rado dijelim savjete, a još bolje suradnju.

Zašto računati na gridu?

abinit - vrlo zahtjevni za veći broj atoma u jediničnoj ćeliji (SCF u DFT uz PAW pseudopotencijale), što je nužno da se opiše spinelna struktura.

Računi uz PAW baze / pseudopot. i 14 atoma (CoAl_2O_4)/j.ć. Traju oko 2-3 mjeseca za optimizaciju geometrije, za fonone čak moguće manje jer se ne mogu rabiti PAW pseudopotencijali.

Problem - neke značajke su “version-specific”, a abinit-online help za najnoviju verziju. Možda update abinita?

Tu sam najmanje potkovan i prilično početnik.

Podešavanje

abinit i G03/G09 nude mogućnosti za podešavanje MP okoline.

Kod G03/G09 jednostavne - broj procesora, scratch-direktoriji / uglavnom pridijeljeno pokretačkom skriptom.

abinit može zahtijevati određene parametre za pojedine vrste posla. (npr. broj procesora / količinu memorije samo za SCF, samo za optimizaciju i td.) Moguće da bi njihovo napredno korištenje dovelo do bržeg izvršavanja posla. No uporaba tih mogućnosti traži da se korisnik upozna detaljno s radom pokretanih rutina.

Primjedbe

isabella vs. CRO-NGI: interaktivnost i nadzor jednostavniji na clusteru

CRO-NGI - veći broj procesora/diskovni prostor.

dobro za - račune za koje se zna/lako je razlučiti pogreške zbog *softwarea* i multiprocesorske okoline. Gdje ta razlika i izvori mogućih poteškoća nisu jasni, *debugiranje* zna biti problematično (ako je narušena reproducibilnost znanstvenih rezultata, narušena je i paradigma znanstvenog rada).

Jednostavnost nadzora/pouzdanost računa (da je posao uistinu raspodijeljen i da radi ok):

- katkad nejasna veza između I/H stanja i broja slobodnih procesora na čvorovima
- ne reagiranje condorG-a na remove (čak ni s -forcex) naredbu za H - poslove
- lakše razaznavanje grešaka u aplikaciji tj. u submitter/MPI ili pak u CE olakšalo bi sate uzaludnog pokretanja poslova.

Prijedlozi

Update abinita bi olakšao primjenu početnicima. (A sam paket je ionako dosta kriptično pisan iako je open-source.)

Bilo koja promjena u skriptama koja bi omogućila jasniju distinkciju porijekla prekida pojedinih poslova.

Nedavno mi se pojavljuje "pucanje" G03 poslova zbog nedostatka na disku / kad se zahtijeva 30GB scratcha npr.

- Možda neki sustav "najavljenih rezervacija" ? Moguća rješenja:

1 - u komunikaciji s adminom (pošalje se e-mail adminu, pa on provjeri i da informaciju oblika - "taj čvor s toliko je od tad do tad rezerviran za posao x"). Dobro: manje vjerojatna zlouporaba te mogućnosti. Loše: traži dodatni angažman čovjeka.

2 - Možda omogućiti neku procjenu praznog prostora pa da condor-skripta pošalje posao samo onima s takvim *rank*-om?

disk.requirement = X npr. varijabla u init skripti, pa submission env. prije submitanja na CE provjeri ima li dotični disk u tom trenutku više od X slobodnog prostora. Ako da, trebao bi rezervirati prostor za r-w samo za taj posao. Dobro: automatski, do korisnika; lose: lako zlouporabiti / tražiti više od potrebnog, češće od potrebnog.

Završne ideje

Modeliranje nanostruktura na gridu - iskustva i izazovi

- Izazovi su teška riječ. Bliže “neriješeni problemi”.

Heuristička uloga grida: htjeli-ne htjeli znanstvenici se upućuju u metode računanja i grid tehnologija kad rabe takve račune. Koristi dio znanstveničkih mogućnosti i upućuje da _ne_ postoji čista teorija na računalima. Sve u odnosu prema konkretnoj implementaciji. U primjeni se vrlo brzo naiđe na specifičnosti (ograničenja) računalnih paketa.

SW mogućnosti ograničene pa su mnoge već iskorištene drugdje. Primjene na različite sisteme, no konceptualne razlike pokazuju se npr. u veličini istraživanih sistema (a što je katkad moguće tek prelaskom na grid).