Computational Quantum Chemistry Tools

Chun-Min Chang

Research Consultant

Institute for Cyber-Enabled Research

MICHIGAN STATE UNIVERSITY

ICER

Outline

- Overview of Quantum Mechanical simulation & tools
- Build molecules & run calcuations in interactive mode
  - – gaussview and gaussian tutorial for computational chemistry
- Calculate electronic structures with more resource by job submission in batch mode
- Checkpoint run your jobs with more time in HPCC
- Summary & Questions

Overview of Quantum Mechanical Simulation & Tools

MICHIGAN STATE UNIVERSITY

ICER

Electron behaving as waves

The wave-like nature of the electron allows it to pass through two parallel slits simultaneously, rather than just one slit as would be the case for a classical particle.

In quantum mechanics, the wave-like property of one particle can be described mathematically as a complex-valued function, the wave function, commonly denoted by the Greek letter psi (ψ). When the absolute value of this function is squared, it gives the probability that a particle will be observed near a location—a probability density $\rho(\mathbf{x})$

Hartree–Fock method

In atomic units, and with r denoting electronic and R denoting nuclear degrees of freedom, the electronic SchrÄodinger equation is

$$\left[-\frac{1}{2}\sum_i \nabla_i^2 - \sum_{A,i}\frac{Z_A}{r_{Ai}} + \sum_{A>B}\frac{Z_A Z_B}{R_{AB}} + \sum_{i>j}\frac{1}{r_{ij}}\right]\Psi(\mathbf{r};\mathbf{R}) = E_{el}\Psi(\mathbf{r};\mathbf{R})$$

with a wavefunction of the general form which is known as a Hartree Product:

$$\Psi_{HP}(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N) = \phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)\cdots\phi_N(\mathbf{r}_N)$$

The generalization to N electrons is

$$\Psi = \frac{1}{\sqrt{N!}}\begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix}$$

Electrons are indistinguishable

Slater determinant

MICHIGAN STATE
UNIVERSITY

ICER

Hartree–Fock method

One-electron operator h as follows

$$h(i) = -\frac{1}{2}\nabla_i^2 - \sum_A \frac{Z_A}{r_{iA}}$$

Coulomb operator as

$$\mathcal{J}_j(\mathbf{x}_1) = \int d\mathbf{x}_2 |\chi_j(\mathbf{x}_2)|^2 r_{12}^{-1}$$

Exchange operator

$$\mathcal{K}_j(\mathbf{x}_1) = \left[\int d\mathbf{x}_2 \chi_j^*(\mathbf{x}_2) r_{12}^{-1} \chi_i(\mathbf{x}_2)\right]$$

Fock operator, as

$$f(\mathbf{x}_1) = h(\mathbf{x}_1) + \sum_j \mathcal{J}_j(\mathbf{x}_1) - \mathcal{K}_j(\mathbf{x}_1)$$

Introducing a basis set transforms

$$\chi_i = \sum_{\mu=1}^{K} C_{\mu i} \tilde{\chi}_\mu$$

Hartree-Fock-Roothaan equations can be written in matrix form

$$\mathbf{FC} = \mathbf{SC}\underset{\uparrow}{\epsilon}$$

Eigenvalues

$$\begin{aligned} S_{\mu\nu} &= \int d\mathbf{x}_1 \tilde{\chi}_\mu^*(\mathbf{x}_1) \tilde{\chi}_\nu(\mathbf{x}_1), \\ F_{\mu\nu} &= \int d\mathbf{x}_1 \tilde{\chi}_\mu^*(\mathbf{x}_1) f(\mathbf{x}_1) \tilde{\chi}_\nu(\mathbf{x}_1) \end{aligned}$$

MICHIGAN STATE
UNIVERSITY

ICER

Density Functional Theory

DFT is the most popular and versatile quantum mechanical modeling method used to investigate the electronic structure (principally the ground state) in physics and chemistry.

Despite recent improvements, there are still difficulties in using density functional theory to properly describe intermolecular interactions, especially van der Waals forces (dispersion); transition states, some other strongly correlated systems; and in calculations of the band gap and ferromagnetism in semiconductors.

DFT Calculation

In density functional theory, the total energy of a system is expressed as a functional of the charge density as

$$E[\rho] = T_s[\rho] + \int d\mathbf{r}\ v_{\text{ext}}(\mathbf{r})\rho(\mathbf{r}) + V_H[\rho] + E_{\text{xc}}[\rho]$$

where *Ts* is the Kohn–Sham kinetic energy which is expressed in terms of the Kohn–Sham orbitals as

$$T_s[\rho] = \sum_{i=1}^{N} \int d\mathbf{r}\ \phi_i^*(\mathbf{r}) \left(-\frac{\hbar^2}{2m}\nabla^2 \right) \phi_i(\mathbf{r}),$$

v_{ext} is the external Potential acting on the interacting system, V_H is the Hartree (or Coulomb) energy,

$$V_H = \frac{e^2}{2} \int d\mathbf{r} \int d\mathbf{r}'\ \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}.$$

MICHIGAN STATE UNIVERSITY

ICER

DFT Method

and E_{xc} is the exchange-correlation energy. The Kohn–Sham equations are found by varying the total energy expression with respect to a set of orbitals:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + v_{\text{eff}}(\mathbf{r})\right)\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r}) \qquad \rho(\mathbf{r}) = \sum_i^N |\phi_i(\mathbf{r})|^2.$$

to yield the Kohn–Sham potential as

$$v_{\text{eff}}(\mathbf{r}) = v_{\text{ext}}(\mathbf{r}) + e^2 \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{\text{xc}}[\rho]}{\delta\rho(\mathbf{r})}.$$

where the last term is the exchange-correlation potential:

$$v_{\text{xc}}(\mathbf{r}) \equiv \frac{\delta E_{\text{xc}}[\rho]}{\delta\rho(\mathbf{r})}$$

How we treat electrons in different simulation systems?

Electron wave function described by a basis set:

*Local Atomic Orbital Basis Set (with Gaussian functions):

Non-periodic molecular systems

(periodic system? Looking for fast method DFTB, QEM)

*Plane Wave Basis Set : Periodic systems (solid state)

All-electron:

system with small atoms (need to use LAO Basis Set)

Pseudo-potentials:

system with heavy atoms. (mostly use Plane Wave Basis Set)

Electrons are divided in two groups: valence electrons and inner core electrons.

Some Comparison between QM and MM Calculation

Molecular Mechanics	Quantum Mechanics
Electrons and nuclei are usually considered as point charges or multipole moments.	Electrons are treated as charge density distributions (probability) within certain ranges or PBC.
Electron correlation interactions are replaced by force field.	Consider electron correlation interactions.
No matrix diagonalization or SCF.	Diagonalization of Hamiltonian and self consistent field (SCF) iterations to derive the stationary energy state.
Less accurate but fast, for big molecules, large systems & statistical behaviors.	Accurate but slow, good for small molecular systems, charge transfer, bond forming and breaking.

Quantum Mechanical Simulation Packages

Package	License[†]	Basis	Periodic[‡]	Mol. mech.	Semi-emp.	HF	Post-HF	DFT
CP2K	GPL	Hybrid GTO / PW	3d	Yes	Yes	Yes	Yes	Yes
CRYSTAL	Commercial	GTO	Any	Yes	No	Yes	Yes[10]	Yes
GAMESS	Academic	GTO	No	Yes[2]	Yes	Yes	Yes	Yes
Gaussian	Commercial	GTO	Any	Yes	Yes	Yes	Yes	Yes
MOLPRO	Commercial	GTO	No	No	No	Yes	Yes	Yes
NWChem	ECL v2	GTO, PW	Yes(PW) No(GTO)	Yes	No	Yes	Yes	Yes
PSI	GPL	GTO	No	No	No	Yes	Yes	Yes
Q-Chem	Commercial	GTO	No	Yes	Yes	Yes	Yes	Yes
Quantum ESPRESSO	GPL	PW	3d	Yes	No	Yes	No	Yes
Spartan	Commercial	GTO	No	Yes	Yes	Yes	Yes	Yes
VASP	Commercial	PW	3d	Yes	No	Yes	Yes	Yes

Molecule Editor Software

Program	Developer(s)	License	Platforms	Info
Avogadro	Avogadro project team	GPL	Linux, Mac OS X, Windows	3D molecule editor and visualizer
ChemDraw	CambridgeSoft	proprietary	Mac OS X, Windows	Chemical structure and reaction editor
Discovery Studio	Accelrys	proprietary	Windows	freeware version available; includes name2structure and structure2name, InChI naming, and canonical SMILES
GaussView	Gaussian	Commercial	Linux, Mac OS X, Windows	-
PyMOL	Schrödinger	GPL	Linux, Mac OS X, Windows	PyMOL is an open-source molecular visualization system.
MOE	Chemical Computing Group	Commercial	Linux, Mac OS X, Windows	-
SPARTAN	Wavefunction, Inc.	proprietary	Linux, Mac OS X, Windows	
VMD	UIUC	proprietary	Linux, Mac OS X, Windows	-

Building Molecular Geometry and Computing Electronic Structure Interactively

MICHIGAN STATE
UNIVERSITY

ICER

Calculations with Gaussian

An electronic structure package capable of predicting many properties of atoms, molecules, and reactive systems, e.g.

- single point energy and properties (electron density, dipole moment, …)
- geometry optimization
- frequency
- reaction path following

utilizing ab initio, density functional theory, semi-empirical, molecular mechanics, and hybrid methods.

Calculations with Gaussian

Levels of Theory Available:

- semi-empirical: AM1, PM3, MNDO, …
- density functional theory: B3LYP, MPW1PW91, …
- ab initio: HF, MP2, CCSD, CCSD(T), …
- hybrid: G2, G3, …

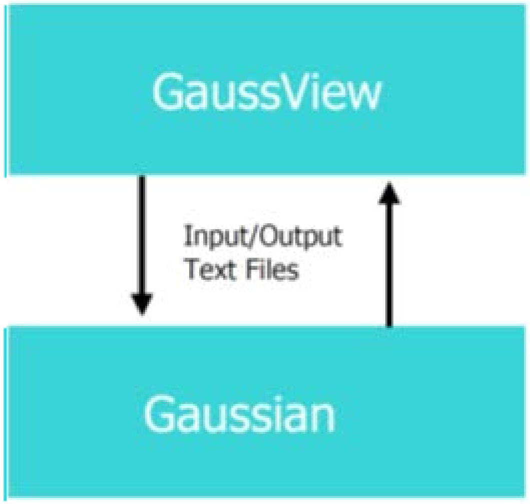
Basis Sets Available:

- Pople-type: 3-21G, 6-31G, 6-311G(d,p), …
- Dunning: cc-pVDZ, aug-cc-pVTZ, …
- Huzinaga and Others: MIDIX, …
- User-defined or General Basis Set

Molecule Building with GaussView

GaussView --- graphical interface for Gaussian 09 ---

- build molecules or reactive systems
- setup Gaussian 09 input files
- graphically examine results



GaussView

Input/Output
Text Files

Gaussian

Visualization

Electronic Structure Calculations

Building with GaussView

- Instead of typing all the coordinates, theory, basis set, etc., we can use GaussView.
- The calculation is specified by pointing and clicking to build the molecule, and using pull-down menus to select the calculation type, level of theory and basis set.
- GaussView generates the Gaussian input file, and can run Gaussian without ever returning to the Unix prompt.
- GaussView can also be used to read Gaussian output files and visualize the results.

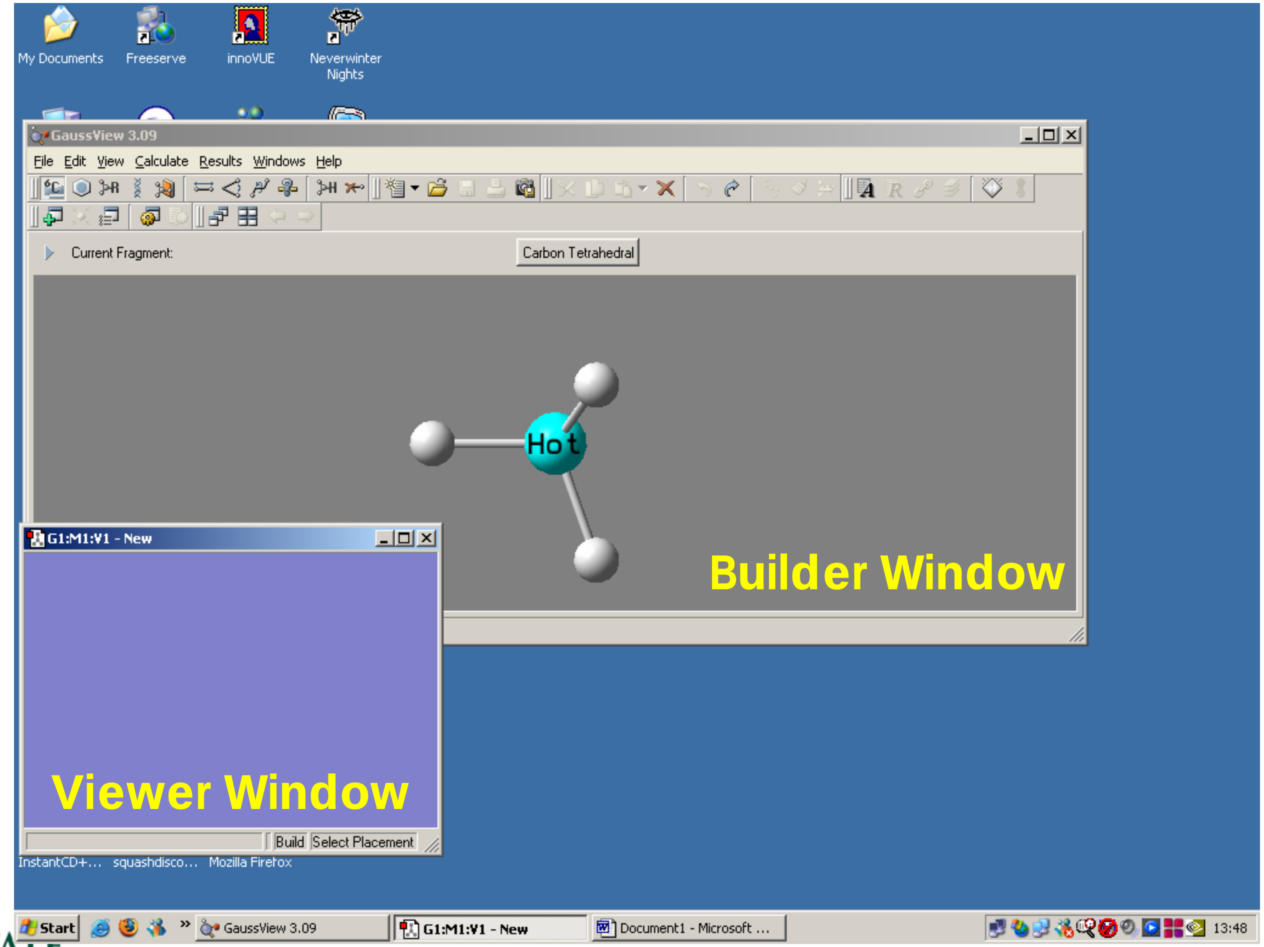
Getting Started with GaussView

- Open a terminal (or MobaXterm) and Connect to HPCC: **ssh –X <NetID>@hpcc.msu.edu** and ssh to a development node. (see Connect to HPCC wiki page.)
- For a fast graphical interface, log into the web site: https://webrdp.hpcc.msu.edu. Use a Mate Desktop session. Open a terminal and ssh to a development node.

  (See Web Site Access to HPCC)
- Use GaussView to create molecular geometry.
- In the command line, please do

```
$ module load Gaussian/g16
$ gview
```

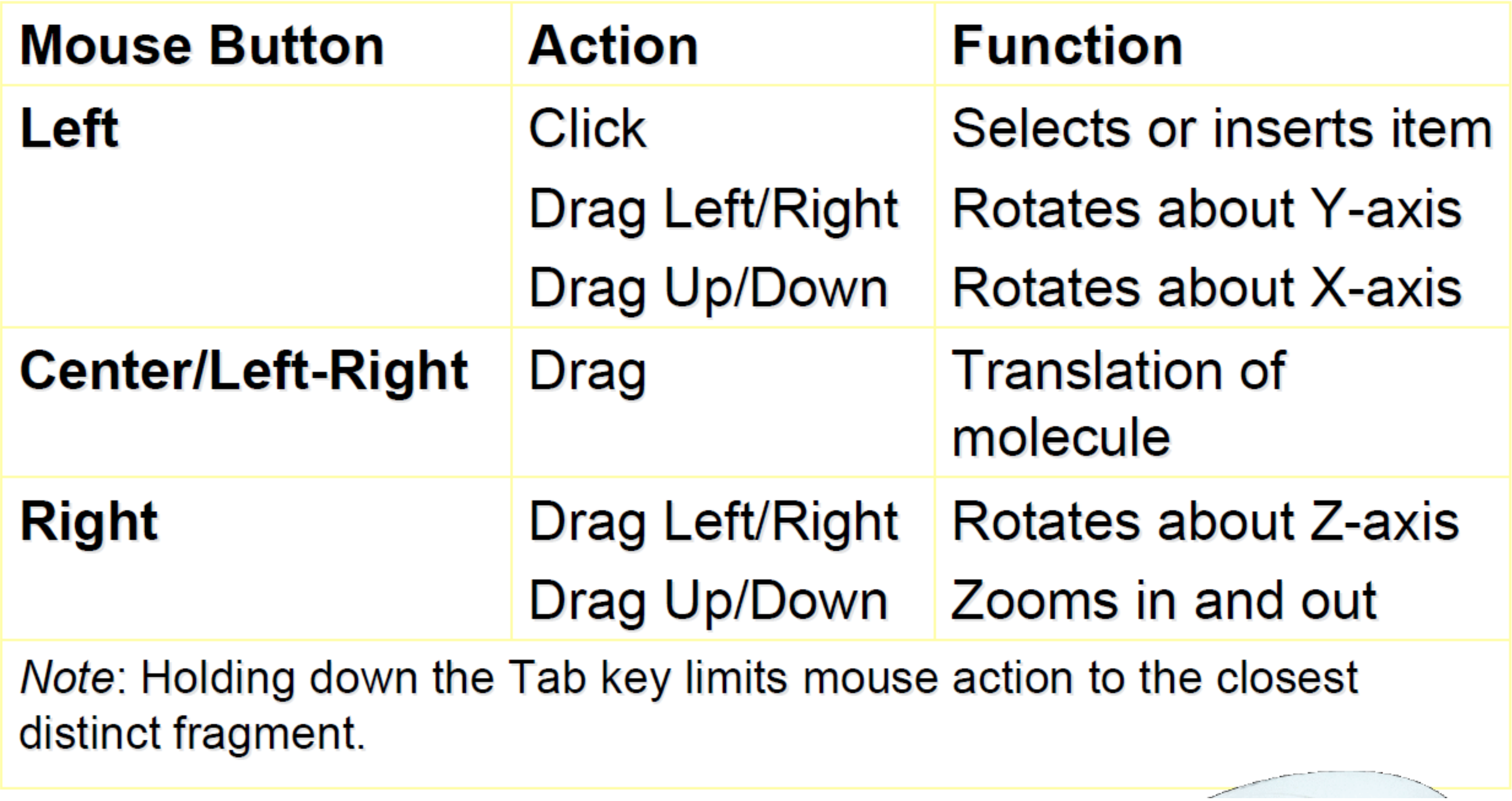
Start GaussView



Builder Window

Viewer Window

Use Your Mouse

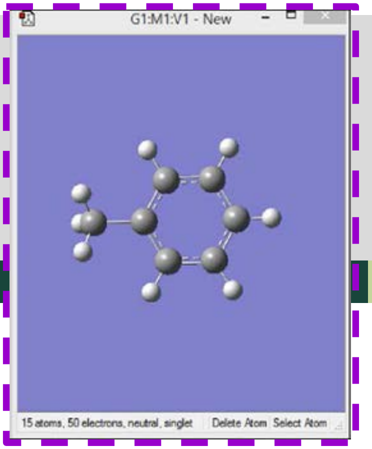


Mouse Button	Action	Function
Left	Click Drag Left/Right Drag Up/Down	Selects or inserts item Rotates about Y-axis Rotates about X-axis
Center/Left-Right	Drag	Translation of molecule
Right	Drag Left/Right Drag Up/Down	Rotates about Z-axis Zooms in and out

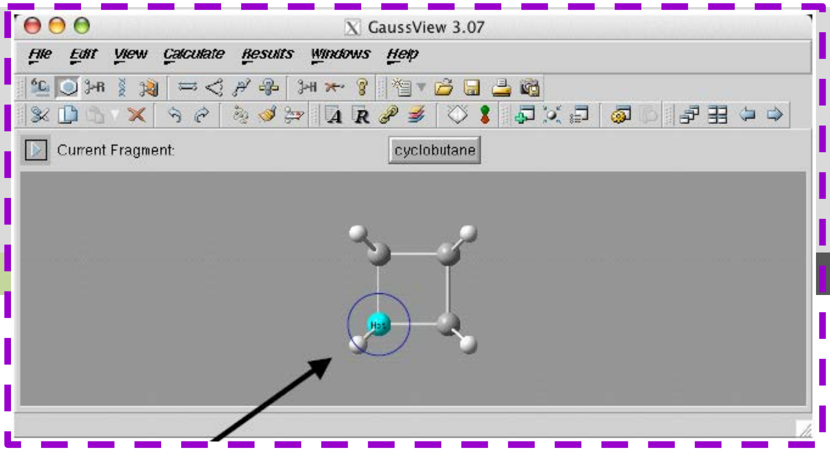
Note: Holding down the Tab key limits mouse action to the closest distinct fragment.

Steps to Building

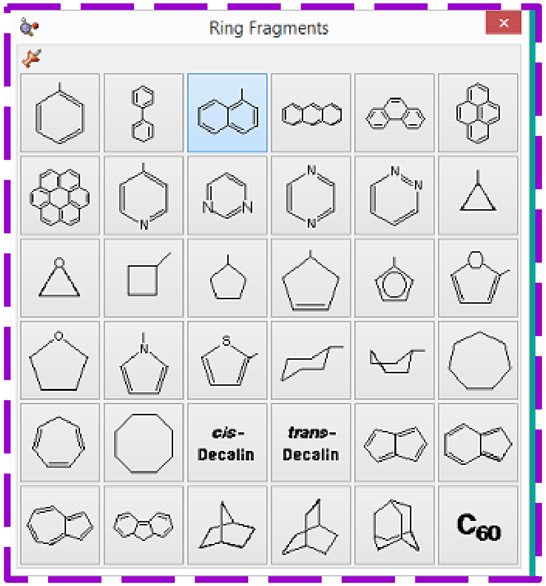
- Choose an atom or fragment.
- Choose the location of the fragment that will attach to your molecule in the builder window.
- Select/add the fragment to your molecule in the viewer window.
- Repeat until molecular building is done.
- Click the Clean button (the broom icon 🧹). This option will use Molecular Mechanics calculations to optimize the geometry (very fast).



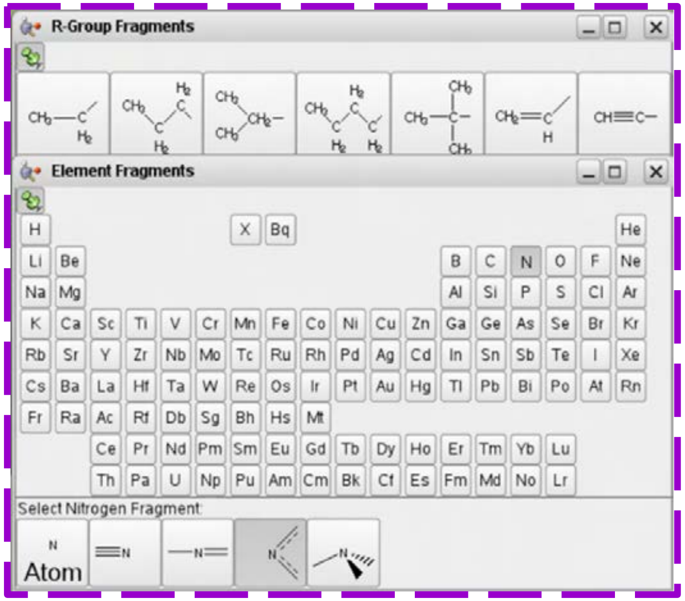
Molecule is put together here.



The atom labeled “Hot” is is where the fragment will attach to the system you are building.

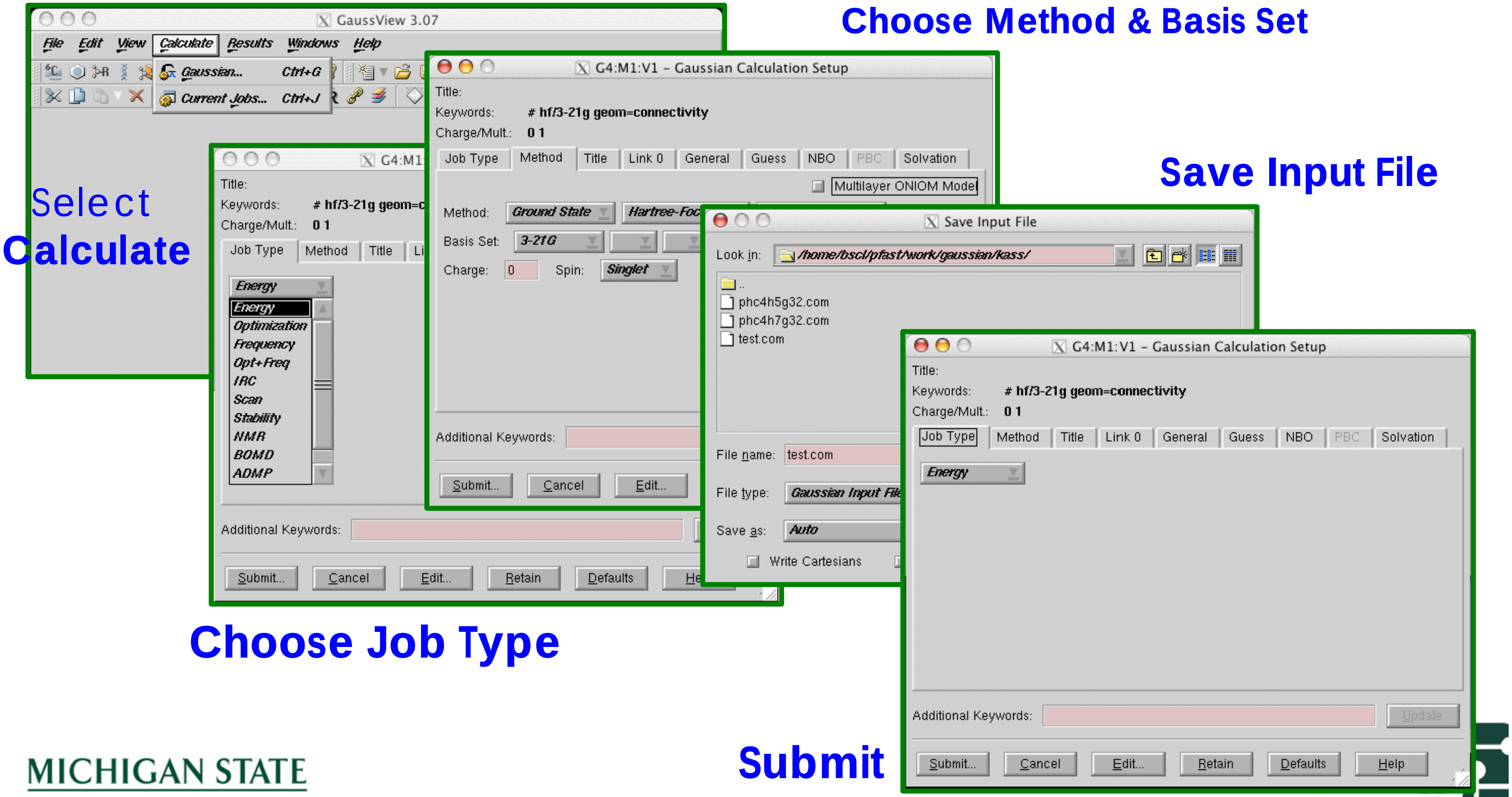


Fragments are selected here.


R-Group Fragments
Element Fragments
H X Bq He
Li Be B C N O F Ne
Na Mg Al Si P S Cl Ar
K Ca Sc Ti V Cr Mn Fe Co Ni Cu Zn Ga Ge As Se Br Kr
Rb Sr Y Zr Nb Mo Tc Ru Rh Pd Ag Cd In Sn Sb Te I Xe
Cs Ba La Hf Ta W Re Os Ir Pt Au Hg Tl Pb Bi Po At Rn
Fr Ra Ac Rf Db Sg Bh Hs Mt
Ce Pr Nd Pm Sm Eu Gd Tb Dy Ho Er Tm Yb Lu
Th Pa U Np Pu Am Cm Bk Cf Es Fm Md No Lr
Select Nitrogen Fragment
Atom

Elements are selected here

Submitting jobs through GaussView



Title:

Keywords: # hf/3-21g geom=connectivity

Charge/Mult.: 0 1

Method: Ground State, Hartree-Foc

Basis Set: 3-21G

Charge: 0

Spin: Singlet

Job Type: Energy

Look in: /home/bscl/pfast/work/gaussian/kass/

..
phc4h5g32.com
phc4h7g32.com
test.com

File name: test.com

File type: Gaussian Input File

Save as: Auto

□ Write Cartesians

Select Calculate

Choose Job Type

Choose Method & Basis Set

Save Input File

Submit

MICHIGAN STATE UNIVERSITY

ICER

Run your Gaussian job

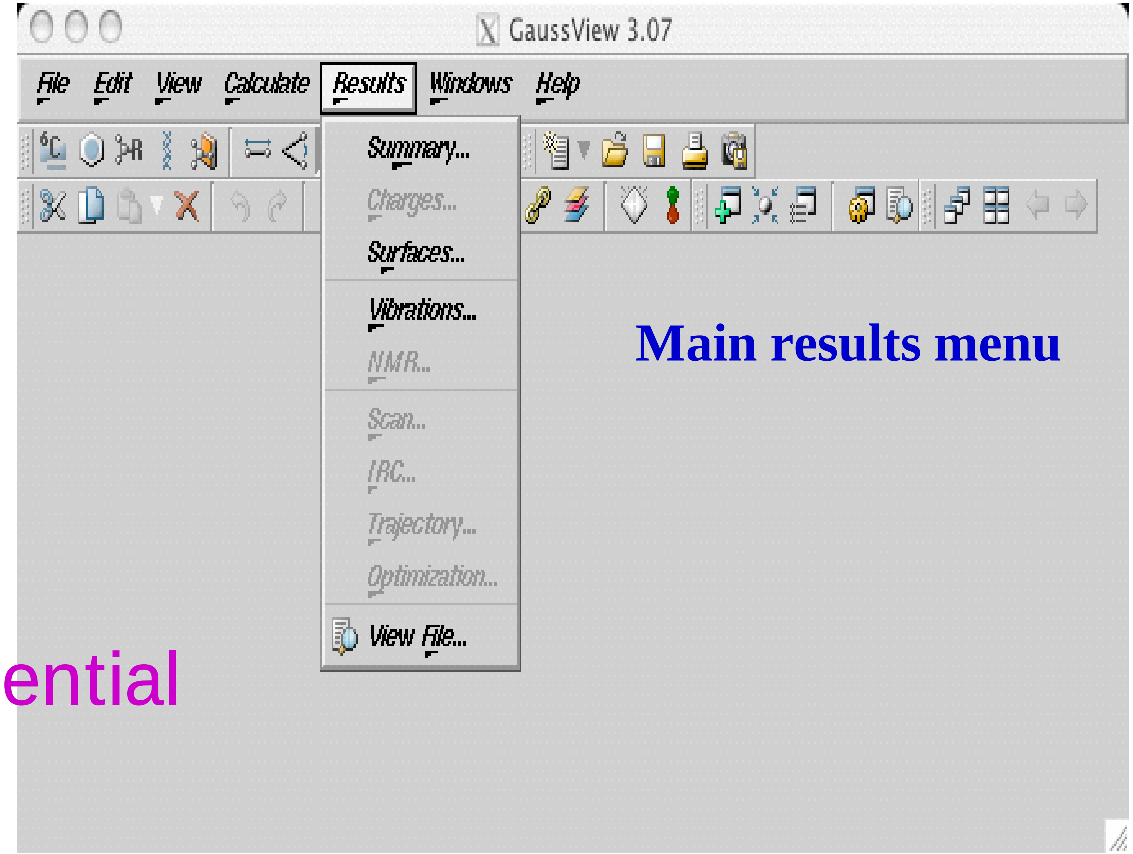
- If you haven’t saved your input file yet, Gaussview will Prompt you to save your input file.
- You will name the input file, and then submit the calculation.
- When Gaussian is finished running, you will receive a message in Gaussview.
- **BE CAREFUL how long you run interactively.**

How to view your output

- Your submitted calculation will run in the background.
- When it is complete, Gaussview will inform you and ask you if you wish to view an output file.
- From the list of files, you can pick your output (.log or .chk file).
- To open an output file again, click **File** on the menu then click **Open**.

What can be visualized

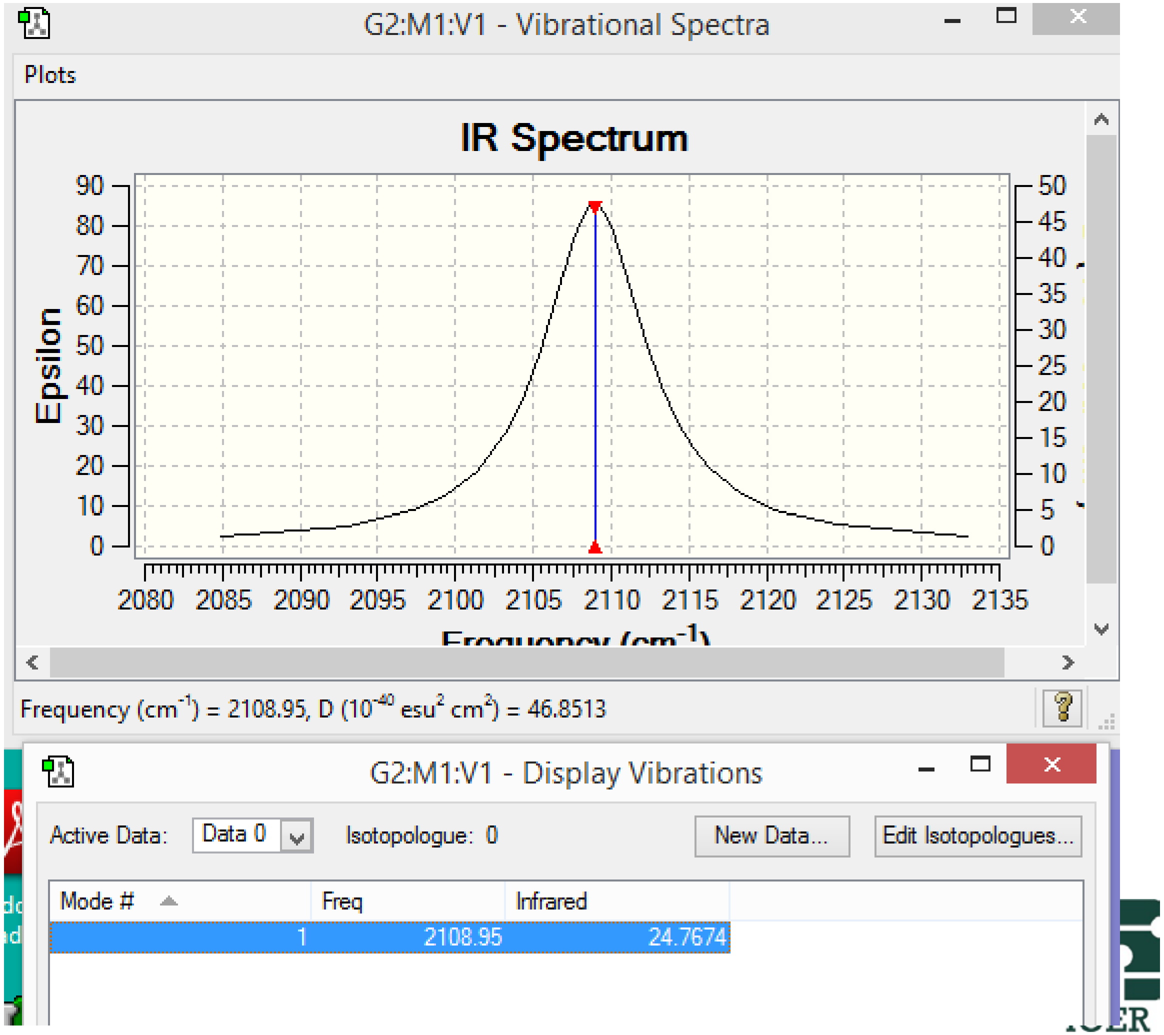
- Geometry
- Vibrations
- Orbitals
- Electron density
- Electrostatic potential



Main results menu

Vibration Frequency & IR Spectrum

Selecting Vibrations from the main-window Results menu will calculate and then display the vibration frequency, plus a calculated IR spectrum.



Frequency (cm^{-1}) = 2108.95, D (10^{-40} esu^{2} cm^{2}) = 46.8513

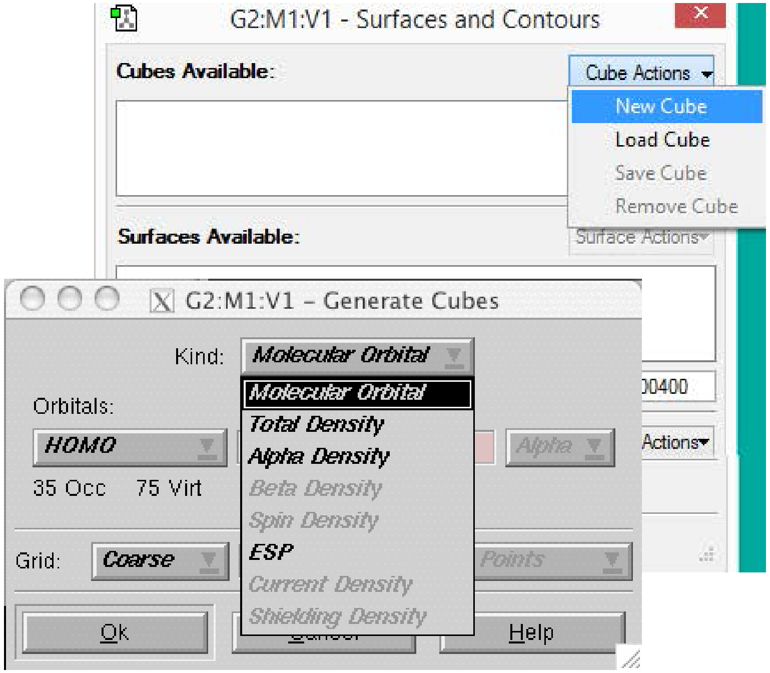
Active Data: Data 0 Isotopologue: 0

Mode #	Freq	Infrared
1	2108.95	24.7674

MICHIGAN STATE
UNIVERSITY

Display Molecular Orbitals

To view molecular orbitals, choose Surfaces/Contours from the main-window Results menu. Begin by generating "cube" files for the HOMO and LUMO to display orbitals, electron densities, electrostatic potentials, etc.



G2:M1:V1 - Surfaces and Contours

Cubes Available:

Surfaces Available:

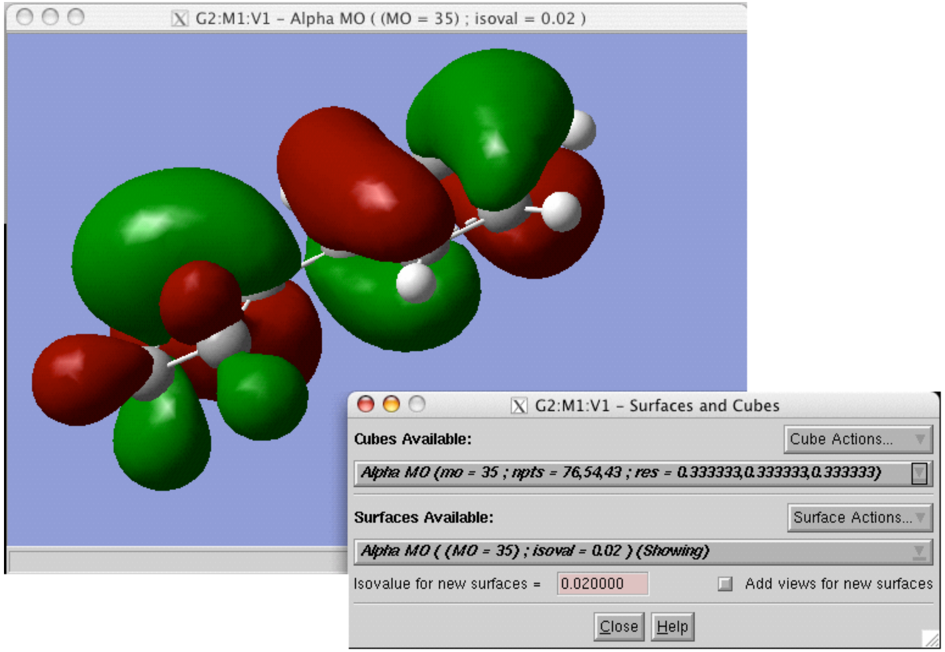
G2:M1:V1 – Generate Cubes

Kind: Molecular Orbital

Orbitals: HOMO

35 Occ 75 Virt

Grid: Coarse



G2:M1:V1 – Alpha MO ((MO = 35) ; isoval = 0.02)

G2:M1:V1 – Surfaces and Cubes

Cubes Available: Alpha MO (mo = 35 ; npts = 76,54,43 ; res = 0.333333,0.333333,0.333333)

Surfaces Available: Alpha MO ((MO = 35) ; isoval = 0.02) (Showing)

Isovalue for new surfaces = 0.020000

□ Add views for new surfaces

Submitting Calculations by Command Line

Interactively use **command:**

```
g16 < input_file > output_file
```

Notes:

- **A path can be specified in front of the input and output file name if they are not in the current directory.**
- **Gaussian temporary files can be redirected by setting the `GAUSS_SCRDIR` environment variable, i.e.**

```
GAUSS_SCRDIR=/mnt/scratch/$USER/GAUSS_SCRDIR
```

- **See an example in** Gaussian wiki page**.**
- **Schedule more time to run your jobs interactively:** https://wiki.hpcc.msu.edu/display/ITH/Interactive+Job

Quantum Mechanical Simulations with More Resource Usage by Job Submission in Batch Mode

MICHIGAN STATE UNIVERSITY

ICER

More Resource Request

- In Gaussian input file, you can specify more cores and memory for your calculation in link 0 section (lines begun with % sign).
- For job submission, create a job script with the **g16** command line to execute your Gaussian calculations.
- The wall-time of your job needs to be less than or equal to 7 days.

Link 0 section

```
%NProcShared=4
%Mem=5GB
%Chk=mychk
#P Freq …

input continues…
```

The Gaussian input file request:

- **4 cores**
- **5GB RAM memory**

to do the calculation

Job Sheduling Script

How to write a job script to schedule a Gaussian simulation job?

See an example in the wik page:

[Gaussian Job Script]

Job Scripts
SLURM options
List my Jobs
Show Job Information
Resource Usage of Jobs
Queuing Policies

A Simple Job Script

Resource Request for HPCC

Tip:

Request more memory in job script than the request from Gaussian input file.

```
#!/bin/bash --login
####### Give the job a name
#SBATCH --job-name=GaussianJob
####### Number of tasks (HPCC Gaussian can only use 1 task.)
#SBATCH --ntask=1
####### Number of CPUs to use
#SBATCH –cpu-per-task=4
####### Memory needed for each node
#SBATCH --mem=7gb
####### Time for job to run (HH:MM:SS format)
#SBATCH --time=2:00:00
####### Setup the clusters available for the job
#SBATCH –constraint=“intel14|intel16|intel18”
####### Send an email when a job is aborted, begins or ends
#SBATCH --mail-type=BEGIN,END
```

scontrol show job $SLURM_JOB_ID   ### write job information to SLURM output file
js -j $SLURM_JOB_ID   ### write resource usage to SLURM output file (powetools command)

A Simple Job Script

Command Lines =>

Gaussian in HPCC system can only run one task in one node but with multiple CPU threads.

```
InputFile=<Gaussian input file name>
OutputFile=<Gaussian output file name>

module load Gaussian/g16 powertools
#GAUSS_SCRDIR=/mnt/scratch/$USER/GAUSS_SCRDIR
#mkdir -p ${GAUSS_SCRDIR}

g16 < ${InputFile} > ${OutputFile}

### write job information to SLURM output file
scontrol show job $SLURM_JOB_ID

# Print out resource usage
js -j $SLURM_JOB_ID          ### powetools command
```

Use Scratch Space

- The I/O in home and research directory is good. For faster job running (especially running with many cores) and more file space usage, try to use scratch space (**/mnt/scratch/$USER**).
- You can create a work directory in scratch space. Copy all Gaussian input and job script files there and submit your jobs.
- You may set up Gaussian Scratch directory to a proper place by pointing it to the variable **GAUSS_SCRDIR** in your job script. You can also use the command: `mkdir –p ${GAUSS_SCRDIR}` in case it has not been created yet.
- After jobs are done, copy your files back to home or research directory. Files on scratch will be purged after 45 days.

What if you need calculation time more than 7 days?

Long Running Gaussian Job with Checkpointing Function

Using Checkpointing Function

- Gaussian can set up checkpointing so a snapshot of the last running state can be saved during a Gaussian calculation.
- To use this function, set rwf file by **`%RWF`** and chk file by **`%Chk`** .
- Once the files are saved, a restart input file can be used to restart a Gaussian calculation where has not finished last time.
- Benefit for using checkpointing : **Job time can be short !!**

HPCC maintains buy-in nodes (about 55% of nodes). If you are not in the buy-in group of the nodes, only jobs with walltime less than 4 hours can run on them, i.e., it is likely that the short jobs can be started fairly quickly.

Restart Previous Gaussian Job

For example, in a frequency calculation you would include following commands:

```
%RWF=myrwf
%NoSave
%Chk=mychk
#P Freq …

input continues…
```

The following input file will restart a previous job:

```
%RWF=myrwf
%NoSave
%Chk=mychk
#P Restart
```

- The checkpoint file can be placed in the directory of the output file, which might have only a moderate amount of free space. However, the read-write file for a job large enough to be worth restarting should be on a large, local scratch file system.
- The file (such as **`myrwf`** from **`%RWF`**) before **`%NoSave`** will be deleted if the job finishes normally.

Job Script for Continuous Submission

- A long running job can break into many 4-hour short jobs so they can start running quickly.
- It is better we can have a job script to keep submitting jobs to restart automatically if the calculation is not finished.
- See example in wiki page: Gaussian Job with Checkpoint Run
- The job script contains two major processes:

  A foreground process to execute Gaussian calculation.

  A background process in sleep until the end of wall-time to submit another.
- If the foreground process is not finished at the end of wall-time the background process will stop the running and submit another job.

How to Specify Job Resource & Check Job Status

- Check HPCC resources of computing nodes:

  https://wiki.hpcc.msu.edu/pages/viewpage.action?pageId=20120131

- Check usage of HPCC resource now by powertools command:

  ```
  node_status   (powertools needs to be loaded.)
  ```

- According to the HPCC resource, specify your job resource.
- Check your job status with commands (load powertools first):

  ```
  qs, sq            (powertools commands)
  scontrol show job <Job ID>
  ```

- Check the resource usage of your job with the powertools command:

  ```
  js –j <Job ID>
  ```

MICHIGAN STATE UNIVERSITY

ICER

Buy-In Opportunities

- We will maintain your computers for you
- Researchers get exclusive use of their nodes within 4 hours of submitting a job
- Buy-in jobs will automatically overflow into the general resources.
- 2020 Buy-In is in place.

Getting Help

- Documentation and User Manual – wiki.hpcc.msu.edu
- Contact HPCC and ICER Staff for:
  - Reporting System Problems
  - HPC Program writing/debugging Consultation
  - Help with HPC grant writing
  - System Requests
  - Other General Questions
- Primary form of contact - http://contact.icer.msu.edu/
- HPCC Request tracking system – rt.hpcc.msu.edu
- HPCC Office – 1400 PBS
- Open virtual helpdesk: https://wiki.hpcc.msu.edu/x/SYFnAg
  with office hours: – 1pm to 2pm Monday and Thursdays

Q & A

Thanks!

MICHIGAN STATE
UNIVERSITY

ICER