Material Science/Chemistry at Exascale

Mat. Sci & Chem apps like VASP, Quantum ESPRESSO, NWChem, GAMESS, QMCPACK, BerkeleyGW, and CP2K are some of the most heavily used apps at DOE facilities.

They are being used to design and understand the fundamental components of Quantum Computers, Solar Cells, OLEDs, Batteries, Catalysts, Bio-Energy, Semiconductors, Sensors, Hydrogen Storage, Carbon Sequestration
What is GW?

Green’s function/screened Coulomb interaction

State-of-the-art to **accurately** describe many-body **excited-state** phenomena in complex materials:

- What happens when you add or remove an electron from a system?
- How do electrons behave when you apply a voltage?
- How does the system respond to light or X-rays?

BerkeleyGW

https://berkeleygw.org
Computational motifs:
- Large matrix multiplications (100k’s x 10m’s!)
- Fourier transforms
- Large low-rank reductions
- Eigen problems
- Matrix inversions

Scaling for computation vs memory:
- Epsilon: $O(N^4)$ vs $O(N^3)$
- Sigma: $O(N^3)$ vs $O(N^2)$

<table>
<thead>
<tr>
<th>Kernel</th>
<th>Computation</th>
<th>Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTXEL</td>
<td>$O(N_v N_c N_G^2 \log N_G^2)$</td>
<td>$O(N_v N_c N_G)$</td>
</tr>
<tr>
<td>CHI-O</td>
<td>$O(N_v N_c N_G^2)$</td>
<td>$O(N_v N_c N_G + N_G^2)$</td>
</tr>
<tr>
<td>Inversion</td>
<td>$O(N_G^3)$</td>
<td>$O(N_G^3)$</td>
</tr>
<tr>
<td>Epsilon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTXEL</td>
<td>$O(N_{\Sigma} N_b N_G^2 \log N_G^2)$</td>
<td>$O(N_b N_G)$</td>
</tr>
<tr>
<td>GPP</td>
<td>$O(N_{\Sigma} N_b N_G^2)$</td>
<td>$O(N_G^2 + N_b N_G)$</td>
</tr>
</tbody>
</table>
**BerkeleyGW**

**Initial Code:**  
~100k LOC; Fortran; MPI/OpenMP on CPU

**GPU porting and optimization:**
- CUDA/C++ and OpenACC branches
- cuBLAS/cuFFT libraries and custom codes
- non-blocking cyclic communication scheme (comp./comm. overlapping)
- CUDA streams (comp./D-H transfers overlapping)
- batching mechanism
- pre-computing

---

Isosurface for one of the in-gap states associated with a divacancy defect in Silicon

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Si-214</th>
<th>Si-510</th>
<th>Si-998</th>
<th>SiC-998</th>
<th>Si-3742</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{\text{spin}}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2 (↑/↓)</td>
<td>1</td>
</tr>
<tr>
<td>$N_{\Psi}$</td>
<td>31,463</td>
<td>74,653</td>
<td>145,837</td>
<td>422,789</td>
<td>363,477</td>
</tr>
<tr>
<td>$N_{G}$</td>
<td>11,075</td>
<td>26,529</td>
<td>51,627</td>
<td>149,397</td>
<td>141,505</td>
</tr>
<tr>
<td>$N_{b}$</td>
<td>6,397</td>
<td>15,045</td>
<td>29,346</td>
<td>16,153</td>
<td>80,394</td>
</tr>
<tr>
<td>$N_{\nu}$</td>
<td>428</td>
<td>1,020</td>
<td>1,996</td>
<td>1,997/1,995</td>
<td>5,484</td>
</tr>
<tr>
<td>$N_{c}$</td>
<td>5,969</td>
<td>14,025</td>
<td>27,350</td>
<td>14,156/14,158</td>
<td>75,210</td>
</tr>
<tr>
<td>$N_{\Sigma}$</td>
<td>Variable, up to 120 per spin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Epsilon PFLOPs  
2.5 80.5 1164 10,091 66,070

Epsilon Memory (TB)  
0.45 6.07 45.1 135 934

Epsilon Comm.Vol. (GB)  
3.92 22.5 85.3 1428 640

Sigma PFLOPs  
0.127 1.71 12.6 58.2 260.7

Sigma Memory (GB)  
6.19 34.3 133.8 791.4 1006

Sigma Comm.Vol. (GB)  
2.27 12.8 48.5 77.2 365.4
Large distributed GEMMs:
- MPI collectives (MPI_Reduce)
  versus
- point-to-point routines (MPI_Isend/Irecv)

Communication Schemes

non-blocking cyclic communication scheme

collective-based communication scheme
Large Low-Rank Reductions

Sigma GPP kernel:
- Large matrices collapsed to a 3x1 vector
- Sum across threads, thread blocks, streams, and MPI ranks
- Shared memory, mixed CPU-GPU work

$$E_n = \sum_{G} \sum_{G'} M_{nG}^* (-G) M_{nG'} (-G') \frac{\Omega_{GG'}^{2}}{\omega_{GG'} (E_n - \omega_{GG'})} v(G')$$

\(\omega\) and \(\Omega\) are complex DP arrays over \(G, G'\) from polarizability
\(M\) are complex DP arrays for transition probabilities
\(E_n\) are DFT orbital energies
\(E\) is an array of “response” energies
\(v\) is the Coulomb interaction in plane-wave basis \(G\)

Useful optimizations:
- Rearrange loops to transition from bandwidth bound to compute bound region (Roofline)
- Replace long latency instructions with shorter ones
- Remove excessive branching
- Increase occupancy

https://www.nersc.gov/users/training/events/roofline-on-nvidia-gpus-hackathon/
GPU vs CPU Speedup

- **[Top]** Runtime comparison of Epsilon on Cori between Skylake CPU and V100 GPU for Si-214
- Case B prepares ZGEMM buffer on the host (Host-Prep.) while Case C on the device (Full-Offload)

  18x node-to-node speedup!

- **[Bottom]** Runtime comparison of Sigma between Cori Haswell CPU and Summit V100 GPU for Si-510

  86x node-to-node speedup!
• **[Left]** Weak scaling of **Epsilon** on Summit
  • The number of GPUs is scaled according to the computational complexity $O(N^4)$.

• **[Right]** Weak scaling of **Sigma** on Summit
  • The number of GPUs is scaled according to the $O(N^3)$ computational complexity in Cases A, B and C, and to the number of quasiparticles in Cases C, D and E.
Strong Scaling and Best Performance

- [Top left] Throughput of Epsilon CHI-0 and Sigma GPP for SiC-998 and Si-2742 on Summit
- [Top right and Bottom] Strong scaling and best performance (PFLOP/s) of Sigma on Summit

<table>
<thead>
<tr>
<th></th>
<th># of GPUs</th>
<th># of Pools</th>
<th>GPUs per Pool</th>
<th>Compute (s)</th>
<th>IO (s)</th>
<th>Throughput (PFLOP/s)</th>
<th>% of Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si-2742</td>
<td>27,360</td>
<td>120</td>
<td>228</td>
<td>401</td>
<td>226</td>
<td>78.0</td>
<td>39.2</td>
</tr>
<tr>
<td>SiC-998</td>
<td>27,360</td>
<td>80</td>
<td>342</td>
<td>142</td>
<td>71</td>
<td>65.3</td>
<td>32.9</td>
</tr>
</tbody>
</table>
Acknowledgement

• This research used resources of the National Energy Research Scientific Computing Center (NERSC), which is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH11231.

• This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory (ORNL), which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

• This work is supported by the Center for Computational Study of Excited-State Phenomena in Energy Materials (C2SEPEM), which is funded by the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.
Thank You