High-Performance Computing in Molecular Dynamics-based Drug Discovery

CS455: Graduate Project GP00

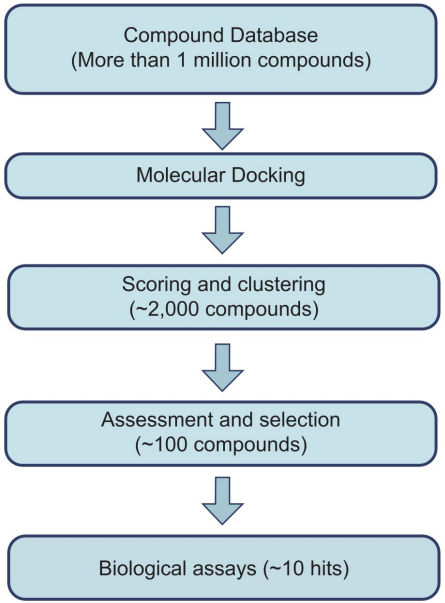
Siddhant Agarwal

HPC Across Domains

Various applications of HPCs:

- AI
- Healthcare
- Defense
- Finance
- And so on…

HPC for Drug Discovery



Source: [5]

- A key application area -

  Drug Discovery through **Molecular Dynamic** (MD) simulations. [1]

- Run simulations on millions of compounds to search for inhibitors against a target.

Limiting Factor in MD Simulation: Compute. [4]

- MD simulations demand high compute power
- GPU clusters now enable simulations on usable timescales
- Moderate-sized clusters ≈ effective, cost-efficient

TABLE 1

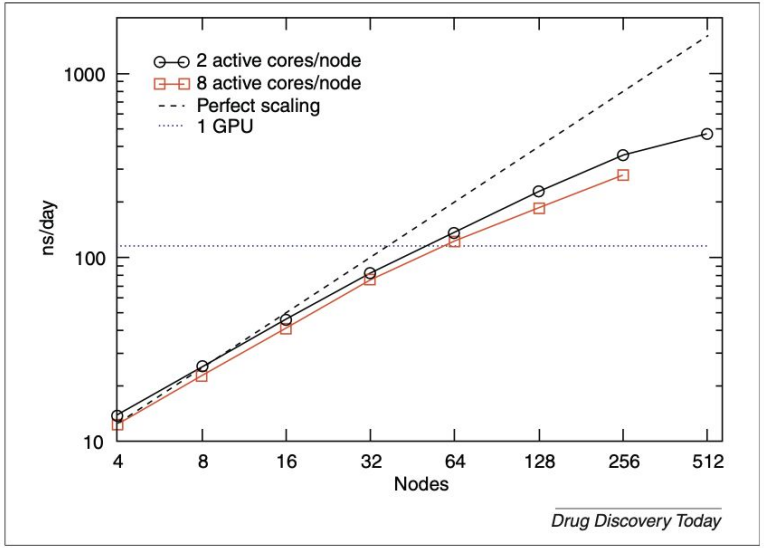
Representative timescales for significant processes in protein motion and associated computation times on 1 GPU and a reasonable GPU cluster. The computational time is estimated for a protein system in explicit solvent (24k atoms), simulated on an Nvidia Geforce 580 GPU with ACEMD [1]. Timescales derived from Fig. 1 of [5]

Timescale	Process	1 GPU compute time	100 GPU compute time
1 fs–1 ps	Bond vibrations	1–1000 ms	
1 ps–1 ns	Side-chain rotations Hinge bending	1–1000 s	Seconds
1 ns–1 us	Loop motions Helix coil transitions	10 min–10 days	Seconds to hours
1 us–1 ms	Allosteric modulation Molecular recognition	10 days–30 years	Hours to months
1 ms–1 s	Protein folding	30 years	Months to years

Source: [4]

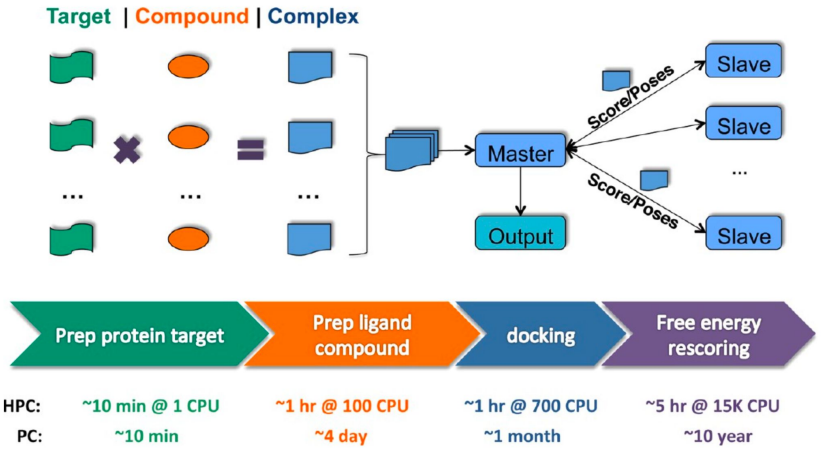
Speedup with HPC [3]

- Conventional MDVS takes months/years
- HPC (e.g., Tianhe-1A) ≈ 600x speedup [3]
- Example: 6 years → 7 hours
- Speed (through GPUs/HPCs)does not compromise accuracy



Source: [4]

Massively Parallel MPI Framework



Source: [6]

- Automated molecular dynamic virtual screening pipeline [6] implemented with MPI [2]
- Thousands of CPUs via MPI in master-slave model
- Master is responsible for dispatching jobs, inputs and outputs
- slave process are busy running the simulation calculations

Strengths

- Huge time efficiency gains
- Scalability across diseases & compound libraries
- Automation = reproducibility and less bias

Weaknesses

- High computational costs still a barrier
- Limited access for small labs/startups
- HPC usage often restricted by job queues and quotas

Future Directions

- Precision Medicine: patient-specific virtual screening
- Deep Learning + HPC: smarter compound suggestions
- Cloud HPC: democratized access (e.g., AWS, Google Cloud)

Conclusion

- HPC is transforming drug discovery via MD simulations
- Feasibility of screening millions of compounds
- Still facing accessibility issues
- Future: HPC as a central pillar of modern pharmacology

Thanks! Questions?

References

[1] Jacob D. Durrant and J. Andrew McCammon. Molecular dynamics simulations and drug discovery. BMC Biology, 9(1):71, October 2011.
[2] Message P Forum. Mpi: A message-passing interface standard. Technical report, USA, 1994.
[3] Hu Ge, Yu Wang, Chanjuan Li, Nanhao Chen, Yufang Xie, Mengyan Xu, Yingyan He, Xinchun Gu, Ruibo Wu, Qiong Gu, Liang Zeng, and Jun Xu. Molecular dynamics-based virtual screening: Accelerating the drug discovery process by high-performance computing. Journal of Chemical Information and Modeling, 53(10):2757–2764, 2013. PMID: 24001302.
[4] Matthew J. Harvey and Gianni De Fabritiis. High-throughput molecular dynamics: the powerful new tool for drug discovery. Drug Discovery Today, 17(19):1059–1062, 2012.
[5] Tingting Liu, Dong Lu, Hao Zhang, Mingyue Zheng, Huaiyu Yang, Yechun Xu, Cheng Luo, Weiliang Zhu, Kunqian Yu, and Hualiang Jiang. Applying high-performance computing in drug discovery and molecular simulation. National Science Review, 3(1):49–63, 2016.
[6] Xiaohua Zhang, Sergio E. Wong, and Felice C. Lightstone. Toward fully automated high performance computing drug discovery: A massively parallel virtual screening pipeline for docking and molecular mechanics/generalized born surface area rescoring to improve enrichment. Journal of Chemical Information and Modeling, 54(1):324–337, 2014. PMID: 24358939.