ParslBox: Computational Materials Workflow Executor for AI-Agents on High-Performance Computing Systems

Fakhrul Hasan Bhuiyan,[1] Álvaro Vázquez-Mayagoitia[1]

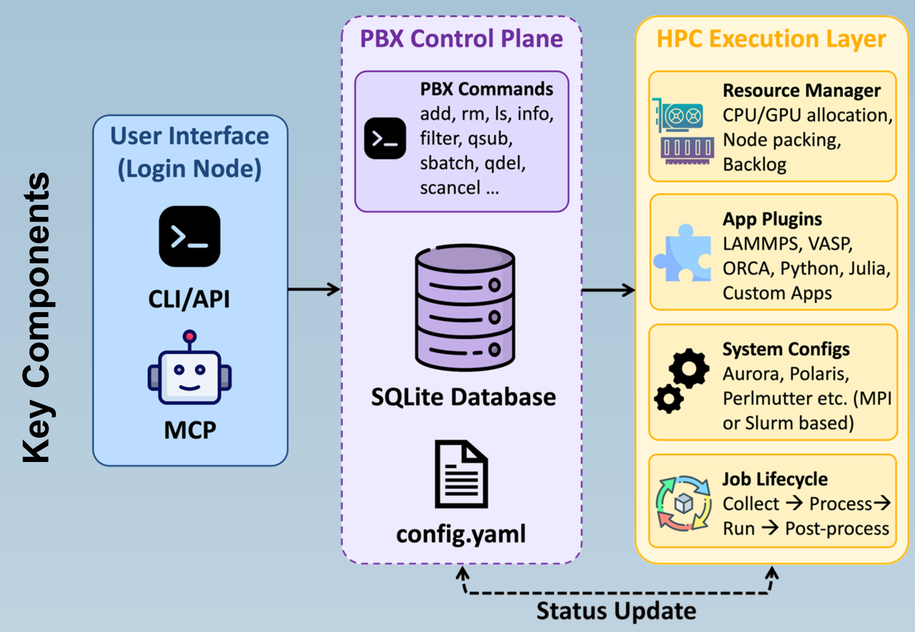
[1]Computational Science Division, Argonne National Laboratory, Lemont, IL 60439, USA

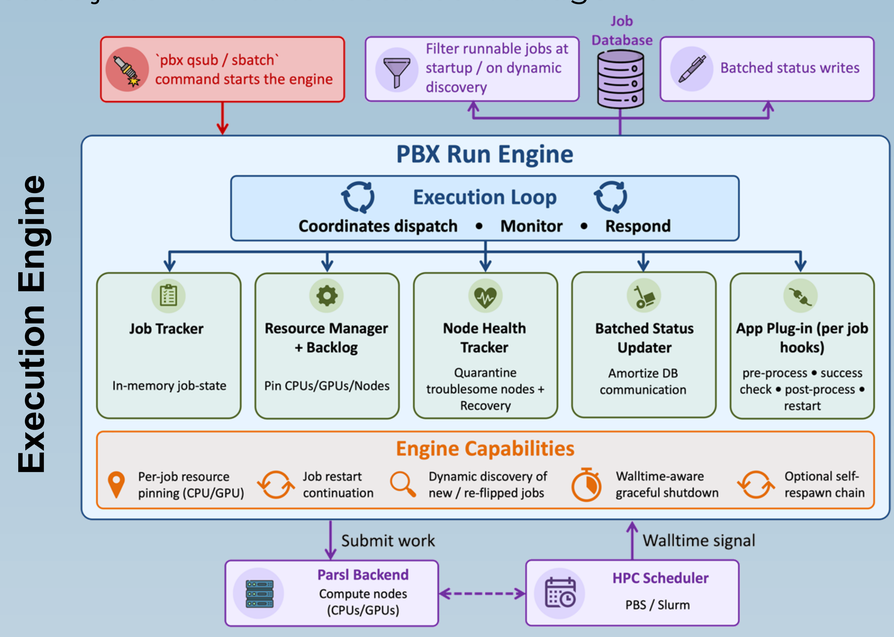
Introduction

- Artificial intelligence (AI)-based agentic systems can orchestrate large computational materials workflows comprising thousands of tasks, including molecular dynamics (MD) simulations, density functional theory (DFT) calculations, and data analysis
- However, high-performance computing (HPC) systems differ widely in their schedulers, queue policies, software environments
- To become system-agnostic, agents need a flexible and reliable executor toolkit
- We present **ParslBox**, a plug-and-play Python package for humans and AI agents to run complex workflows on **any** HPCs

System Overview

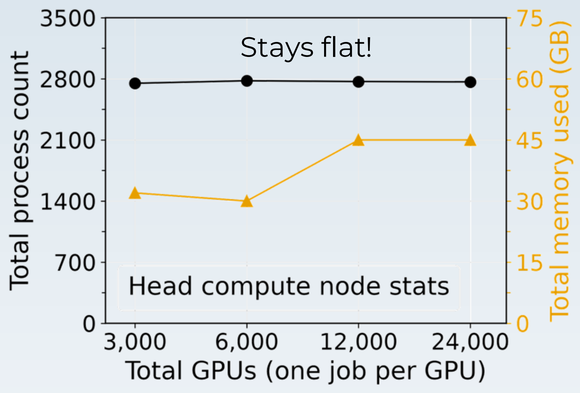
- ParslBox provides CLI, API, and MCP interfaces with access to various commands and SQLite job databases
- The execution engine runs on the compute node and executes jobs with auto resource management

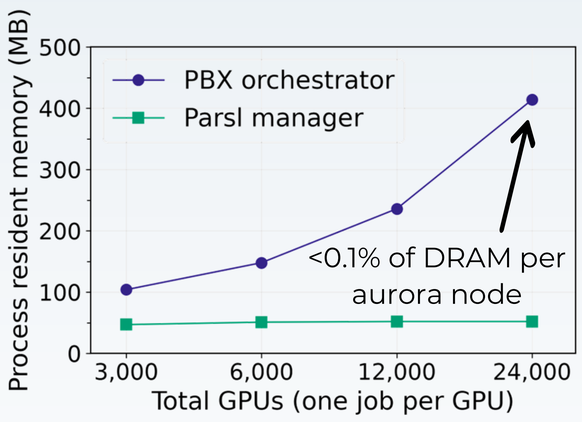


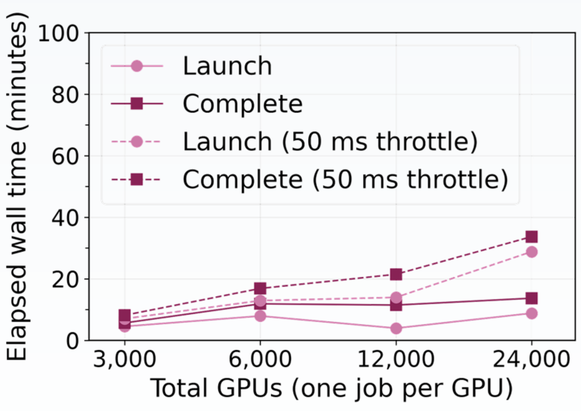


Performance Evaluation

- Stress tested on ALCF Aurora, from 3,000 GPUs (250 nodes) up to 24,000 GPUs (2,000 nodes)
- Head-node memory usage remains constant at any scale
- ParslBox execution engine takes only 400 MB (0.04% DRAM per Aurora node) when running 24,000 jobs
- Time to launch all jobs remains flat at scale

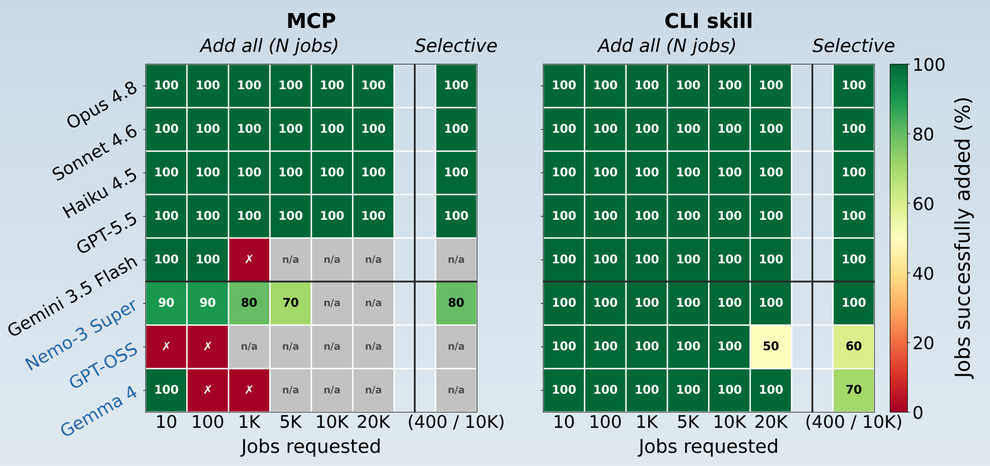


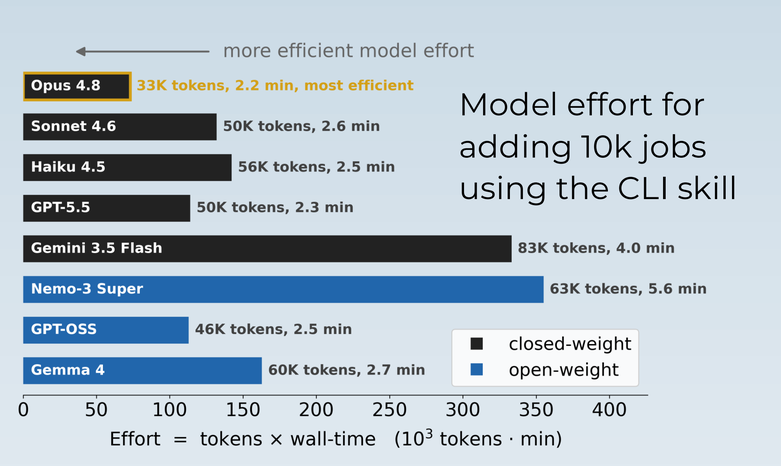




Agentic Usage Benchmarks

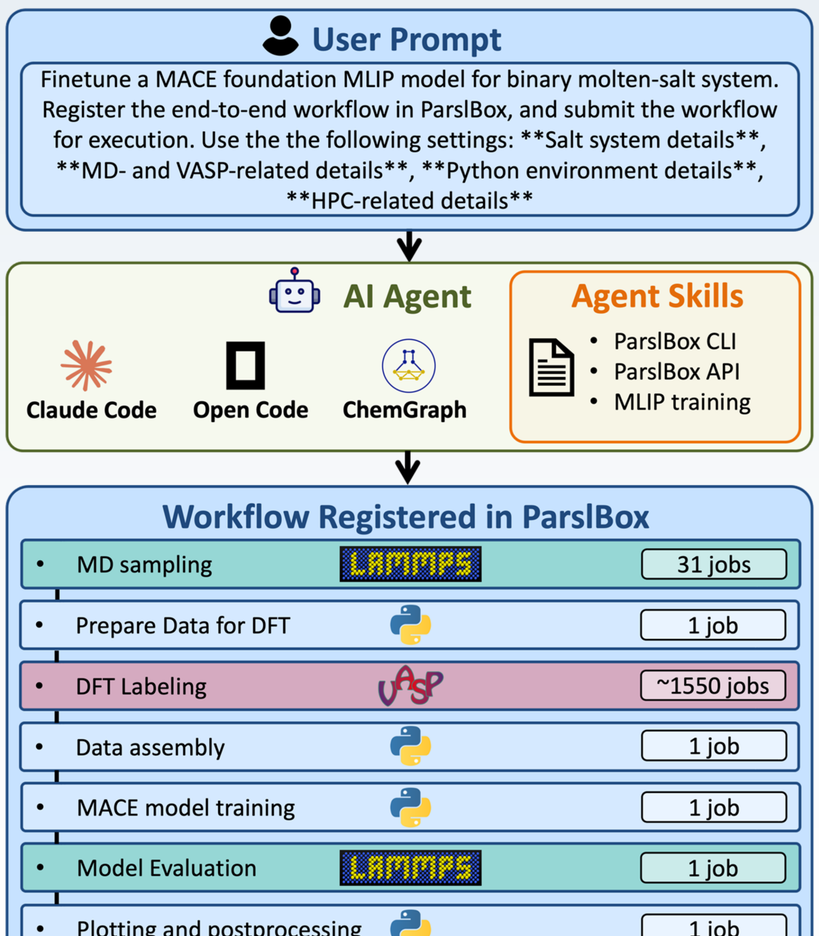
- ParslBox ships with an MCP tool interface and agent skills for the ParslBox CLI and Python API usage
- We evaluated five closed-weight models and three open-weight models as representatives
- All models successfully registered up to 20k jobs using the CLI skill given a **one-shot** prompt

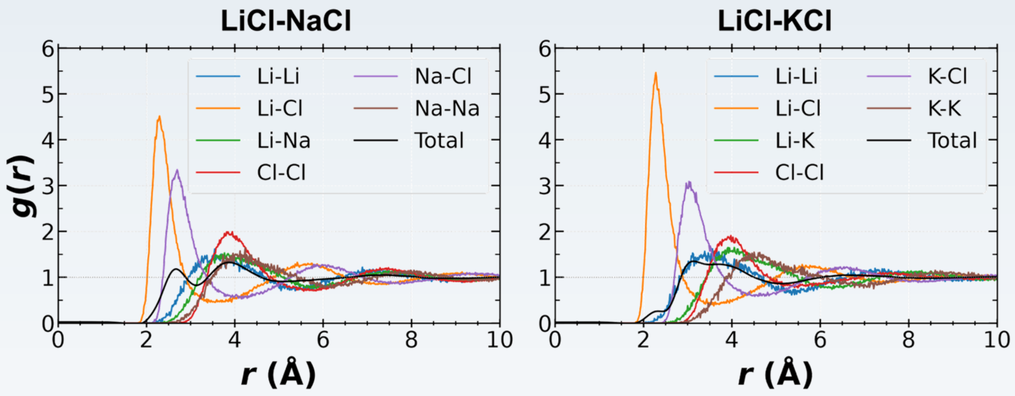




End-to-End Agentic Workflow Execution (MLIP Development)

- Workflow construction was evaluated using Claude Code, OpenCode, and ChemGraph, with Sonnet 4.6 or Haiku 4.5 as the underlying language model





Evaluating the finetuned MACE models using MD simulations

Conclusions

- ParslBox lowers the barrier to managing and executing complex HPC workflows for both humans and AI agents
- Agents can use the CLI skill with ease
- End-to-end MLIP training can be done from a terminal using any agent harness on any HPCs (PBS or Slurm)
- Code, data, and finetuned MACE-mh-0 models available at: https://github.com/fbhuiyan2/parslbox

References

- Finetuning details: Structure of Molten $FeCl_2$ and $FeCl_3$, ACS Mater. Au 2026
- Chemgraph: Commun Chem 9, 33 (2026)

U.S. DEPARTMENT of ENERGY

Argonne National Laboratory is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC.

Questions?
Email: fbhuiyan@anl.gov

Argonne NATIONAL LABORATORY