Sandia National Laboratories

SAND2019-9518C

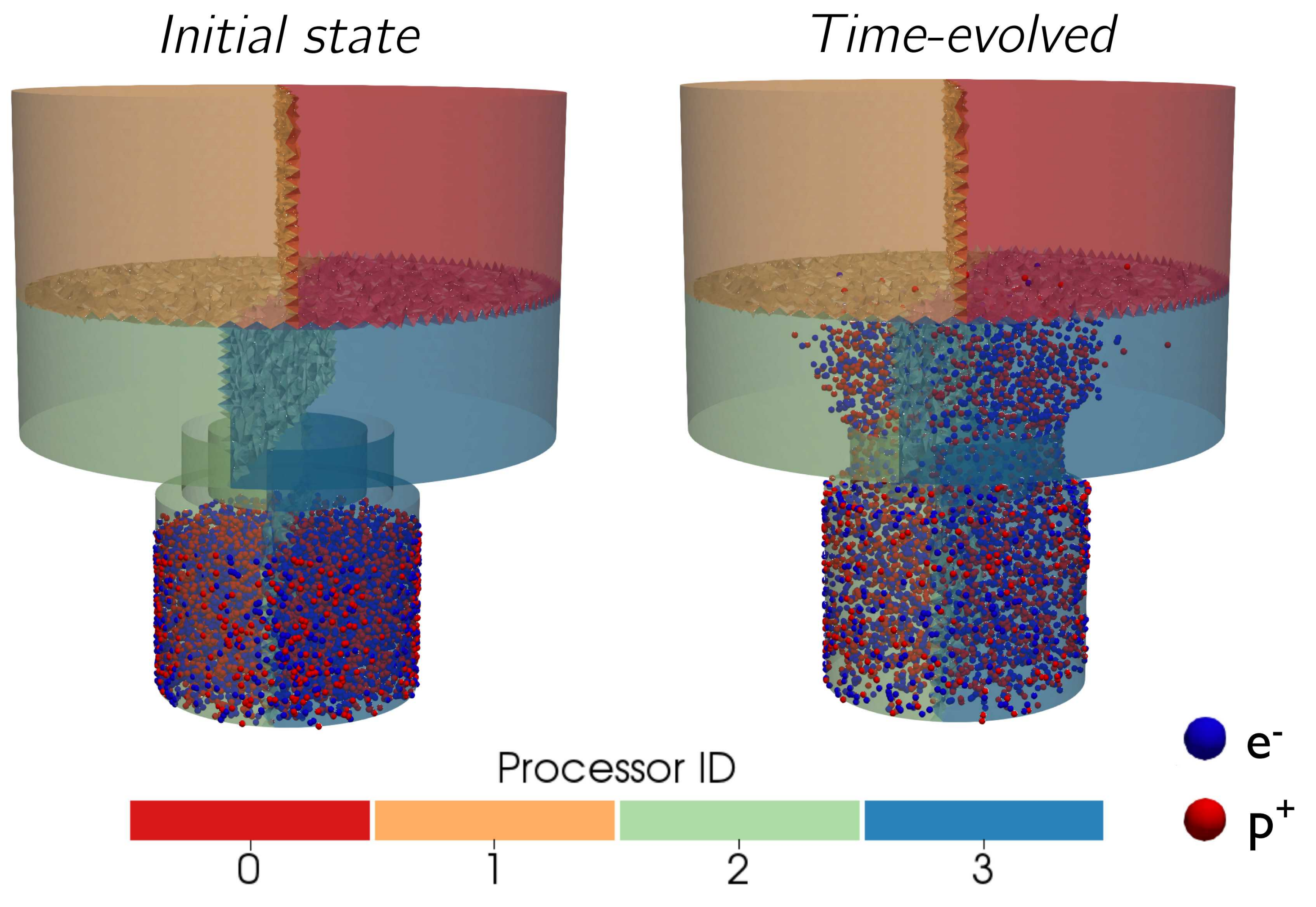
Dynamic, Task-based Load Balancing using DARMA

Presented by: *Jonathan Lifflander & Nicole Slattengren*   Contributors: *Philippe Pébaÿ & Robert Clay*

Problem

EMPIRE is an ATDM plasma physics application that includes a Particle-In-Cell (PIC) algorithm:

- Initial particle distributions can be spatially concentrated, creating **heavy load imbalance**
- Particles may move rapidly across the domain, inducing **workload variation over time**
- Existing MPI-based EMPIRE code **does not support load balancing (LB)**
- Future **Hybrid PIC/Fluid** configurations present a difficult challenge for LB (multi-objective)



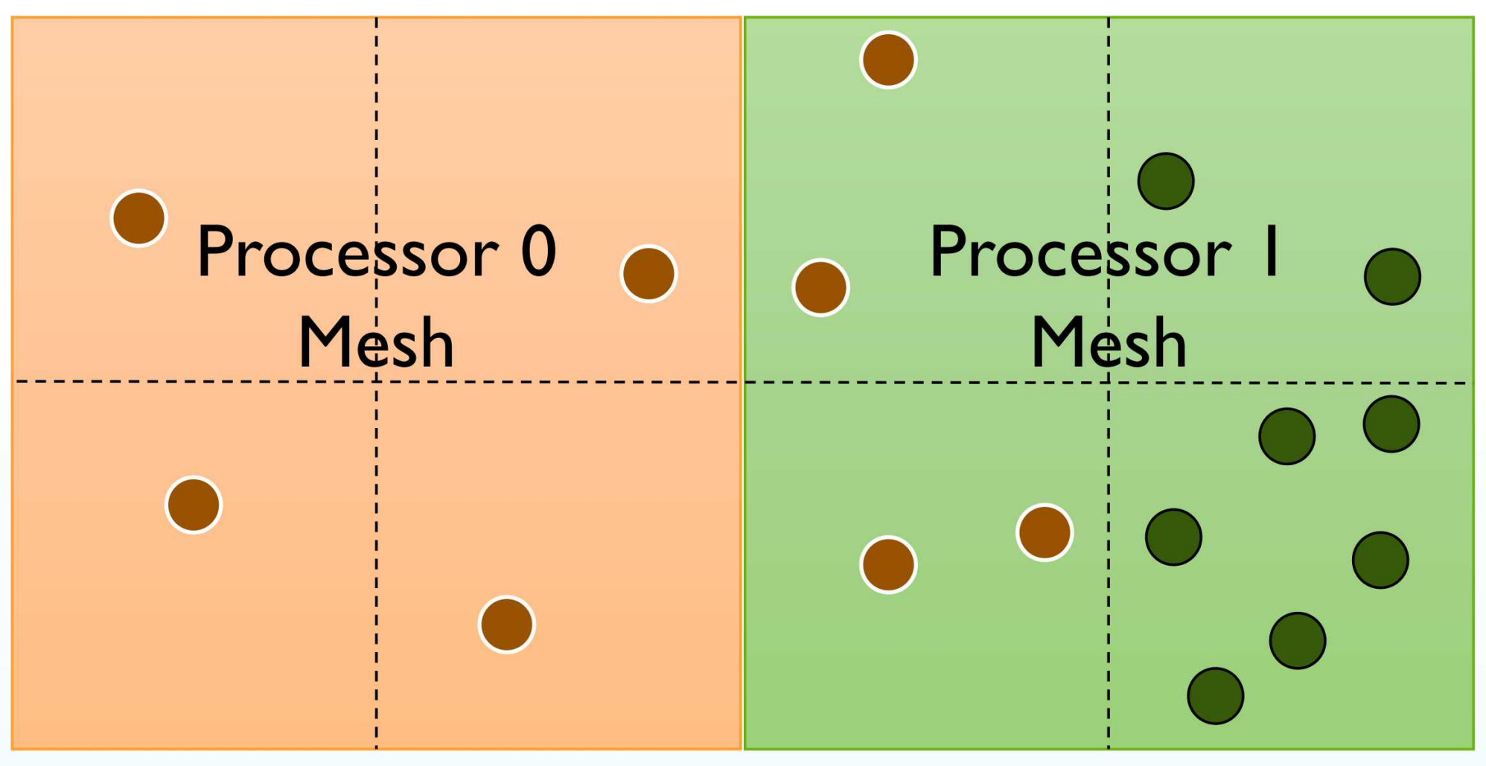
Particle (e^-,p^+) workload imbalance varies over time, necessitating dynamic load balancing

Approach

Conventional approach: infrequently change the mesh decomposition to offset particle imbalance.

− Synchronous process
− Large volumes of data must be migrated to new processors or recomputed from the new mesh

Our approach: maintain the **static, balanced mesh decomposition**, but **split the particles** on *each* rank-decomposed mesh block into k subsets.



Balance achieved with six subsets processed by processor 0.

Fine-grained, **dynamic** LB of particles:

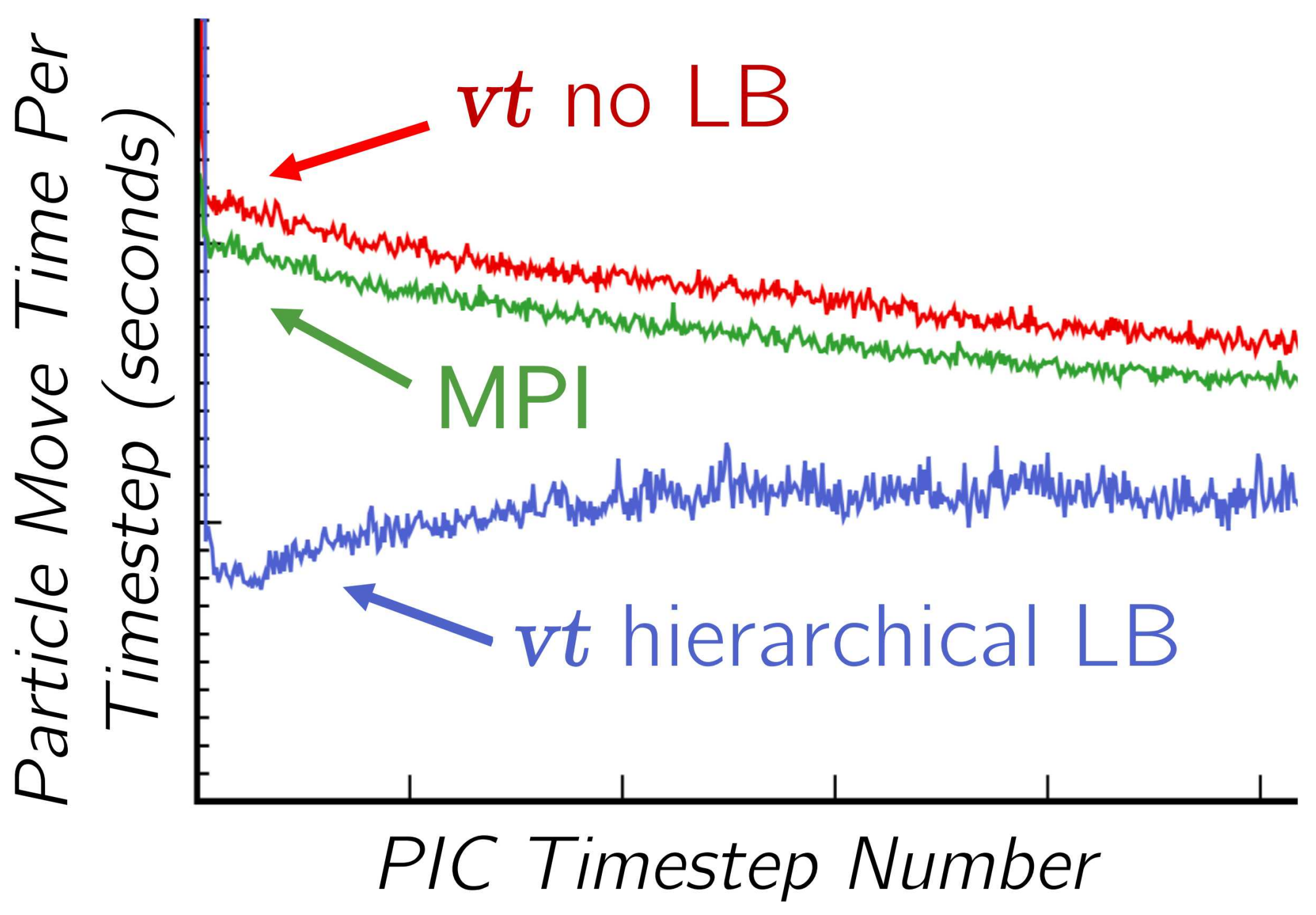
+ Decreases data migration cost
+ Facilitates communication/computation overlap

DARMA's Virtual Transport (*vt*) Tasking Library:

- Interoperable with MPI
- Incremental adoption model for C++ "taskification"
- **Dynamically** migrate data and work off-processor
- Includes scalable load balancers
- Developing a fully-distributed, measurement-driven, communication-aware LB
- Development and tuning are driven by EMPIRE-PIC

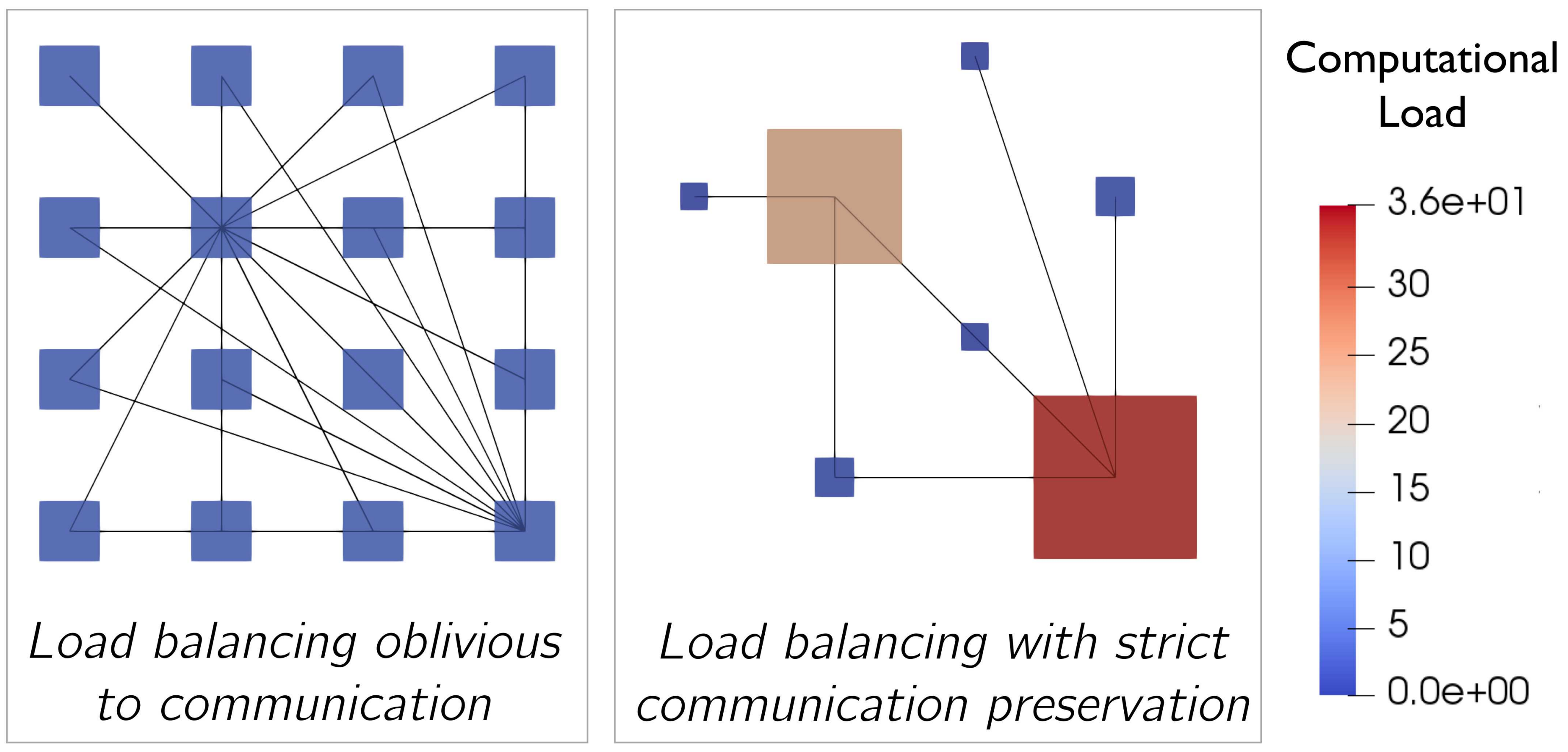
Results

Proof of concept: **demonstrated better than 2x speedup** compared to the MPI baseline by arbitrarily over-decomposing particles on an unbalanced problem.



Novel load balancing algorithm in development: Standalone LB simulation and analysis framework demonstrates the benefit of communication-aware LB:

- **Iteratively refine workloads** with incremental changes
- Preserve **localized communication graphs**
- Optimize load balance by trading off communication vs. computation imbalance



Load balancing oblivious to communication

Load balancing with strict communication preservation

Significance

- Enables dynamic load-balancing for imbalanced, time-varying workloads in codes like EMPIRE
- Mitigates performance imbalances on heterogeneous architectures

Funding

ASC/CSSE: ~3 FTEs, DARMA/EMPIRE integration effort started in FY19

U.S. DEPARTMENT OF ENERGY
NNSA National Nuclear Security Administration

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.
SAND No. ____________

Sandia National Laboratories