Optimizing Large-Scale Plasma Simulations on Heterogeneous Memory with Effective Data Placement Across Memory Hierarchy

Jie Ren[1] Jiaolin Luo[1] Ivy Peng[2] Kai Wu[1] Dong Li[1]

[1] University of California, Merced

[2] Lawrence Livermore National Laboratory

Plasma simulation & Particle-In-Cell (PIC) method

- Plasma simulations are critical for understanding plasma dynamics in many fields.

- Particle-In-Cell (PIC) technique is one of the most popular algorithms in plasma simulation.

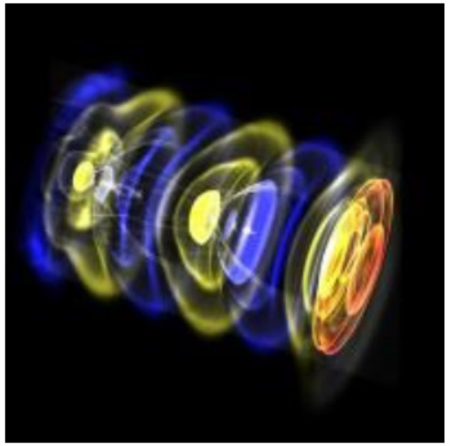


Fig 1. Plasma simulation

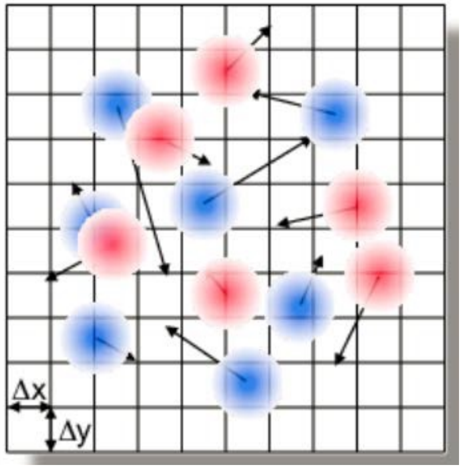


Fig 2. PIC method

WarpX

- A state-of-the-art plasma accelerator simulation code to be deployed in the upcoming exascale HPC system

- Can have large memory consumption
  - E.g., a production run on the Cori supercomputer simulating 62 billions of particles consumes up to 8.9 TB memory

Persistent memory enables WarpX simulation in larger scales

Supercomputer	Mem capacity per node	Largest input problem (in terms of particles)
Sierra	320GB DRAM	10.6 trillions
Summit	608GB DRAM	18.9 trillions
Aurora	256GB DRAM	8.8 trillions
Taihu Light	32GB DRAM	1.1 trillions
Optane-based	1692GB (1.5TB PM plus 192GB DRAM)	58.6 trillions

Assuming optane-based supercomputer use 4,092 nodes

- Compared with Summit and Sierra (the top 2 and top 3 supercomputers), the Optane-based supercomputer increases the simulation scale by 3.1x and 5.5x respectively.

Performance characterization on WarpX

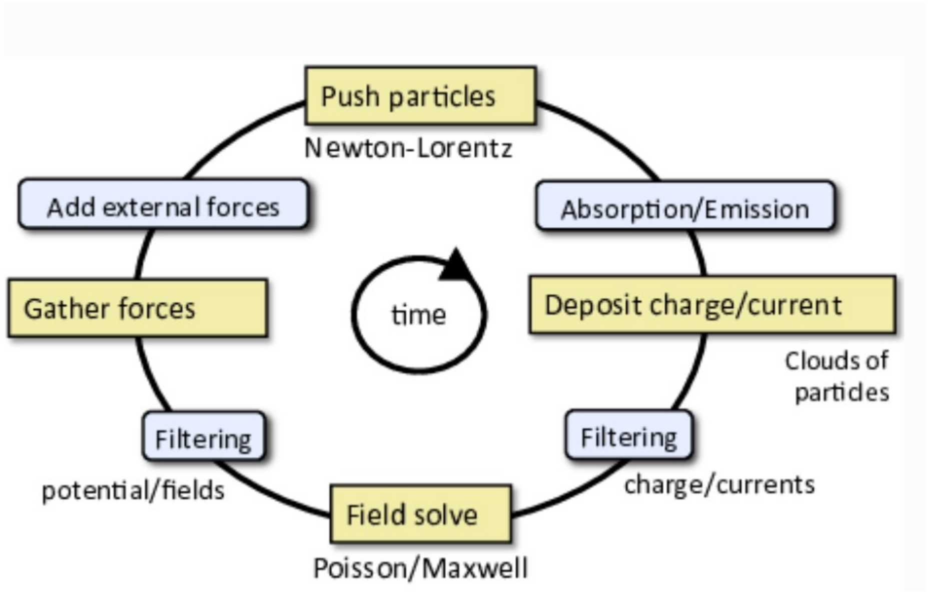


Fig 1. The computation routine of PIC method

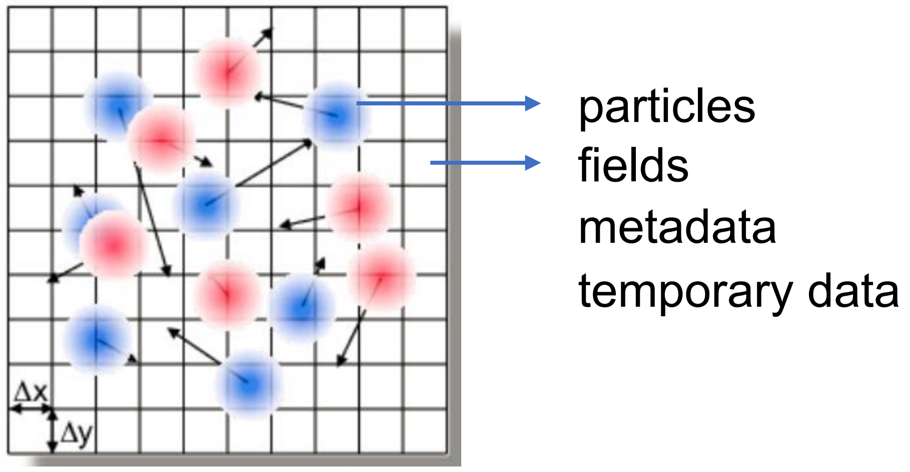
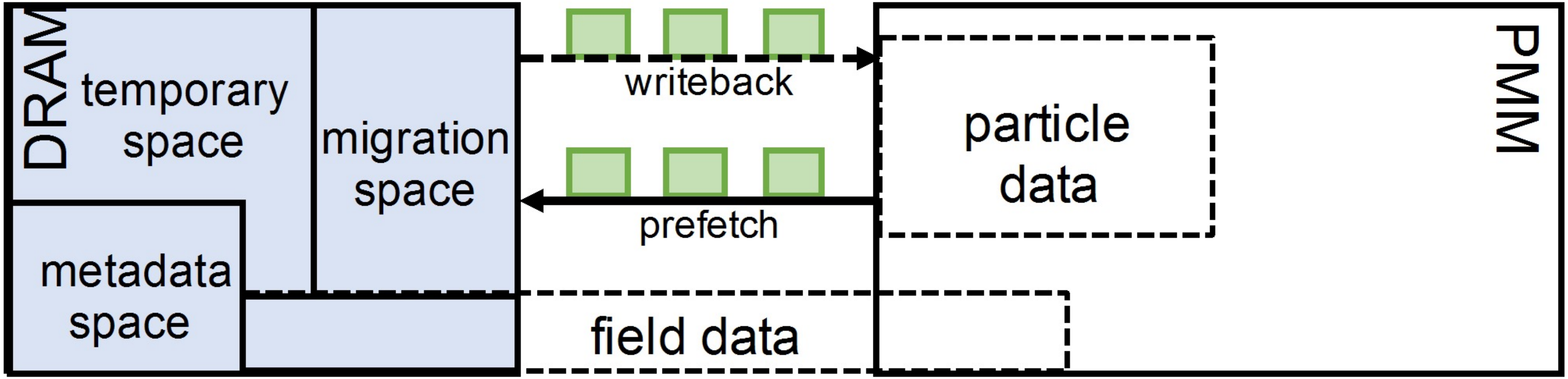


Fig 2. Data objects in PIC method

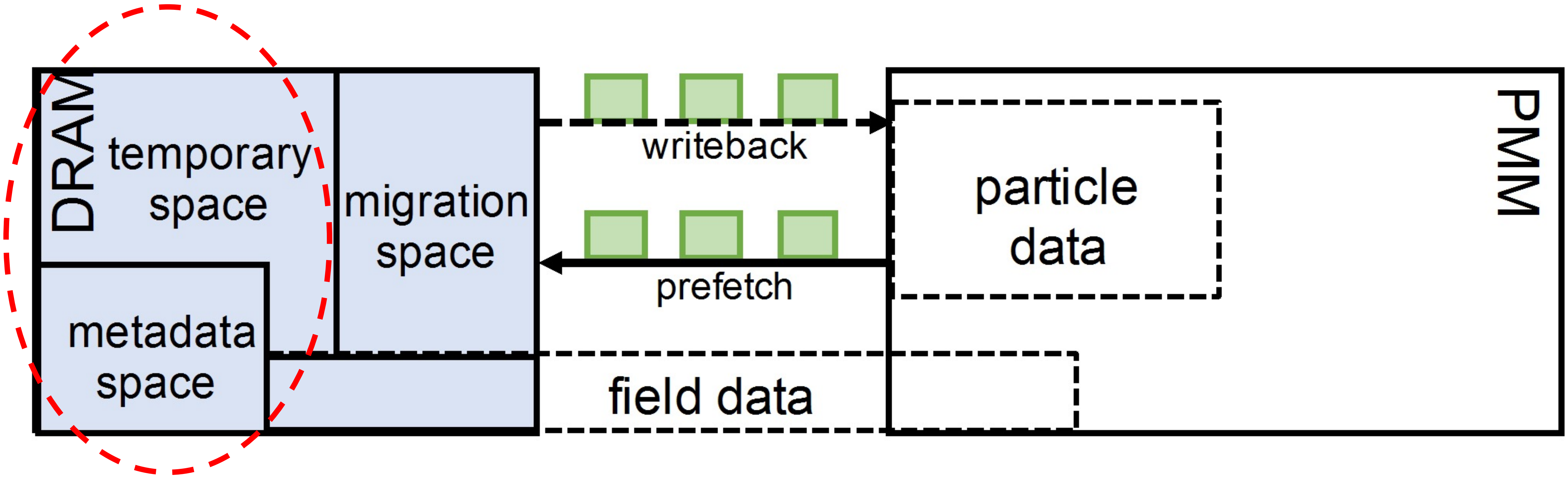
- Streaming-like memory access patterns
- Particles and fields dominate memory consumption
- The execution of WarpX is not bounded by DRAM/persistent memory bandwidth in most of times.

WarpX-PM: an automatic data placement solution for WarpX



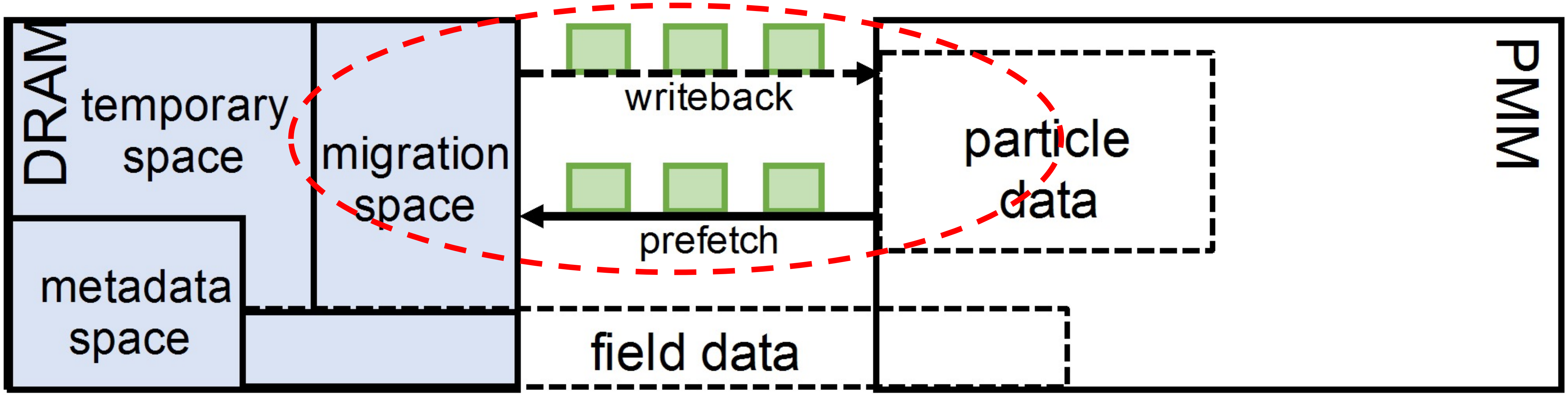
- The DRAM space is partitioned based on the functionality and access patterns of data objects in WarpX.

WarpX-PM: an automatic data placement solution for WarpX



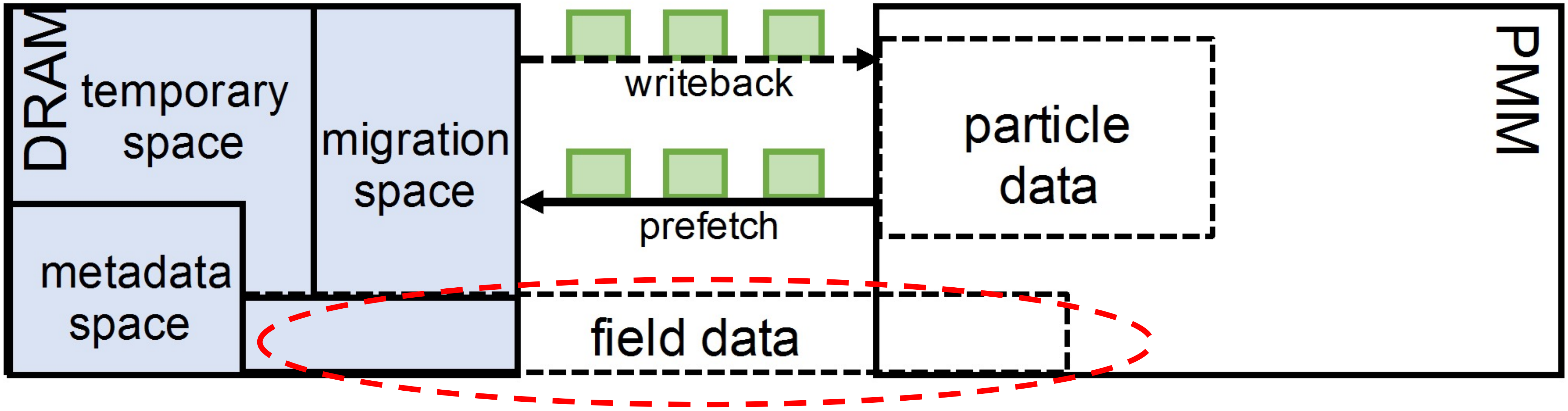
- Static assignment for performance-critical data objects.

WarpX-PM: an automatic data placement solution for WarpX



- Dynamic migration (copy) of particles batch by batch using helper threads.

WarpX-PM: an automatic data placement solution for WarpX



- Prefetch field data to processor-cache directly.

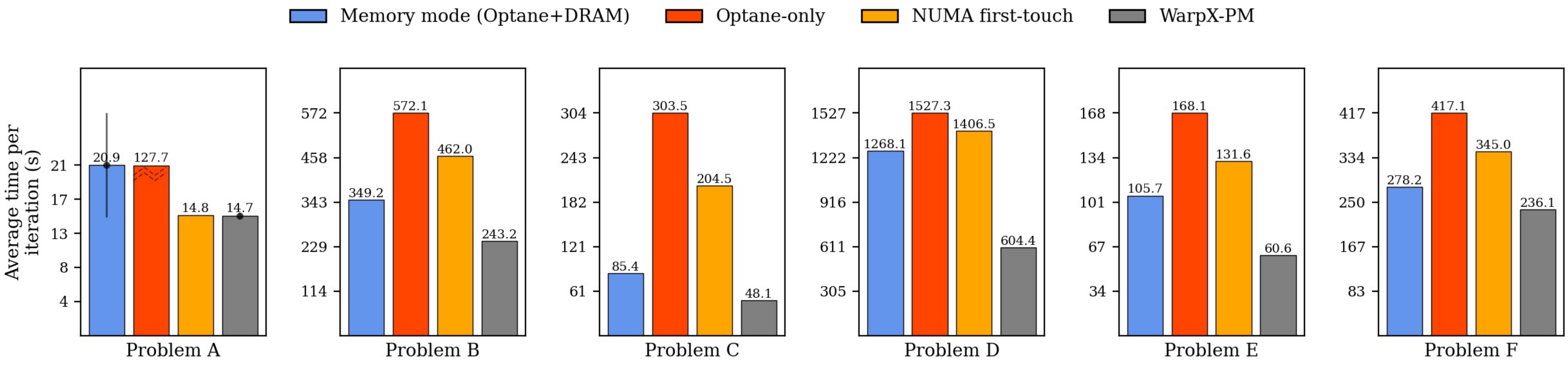
Evaluation setup

- Testbed
  - 2 x Intel Xeon Scalable processor
  - 35.75 MB last level cache
  - 192GB DRAM + 1.5TB PM

- Input problems

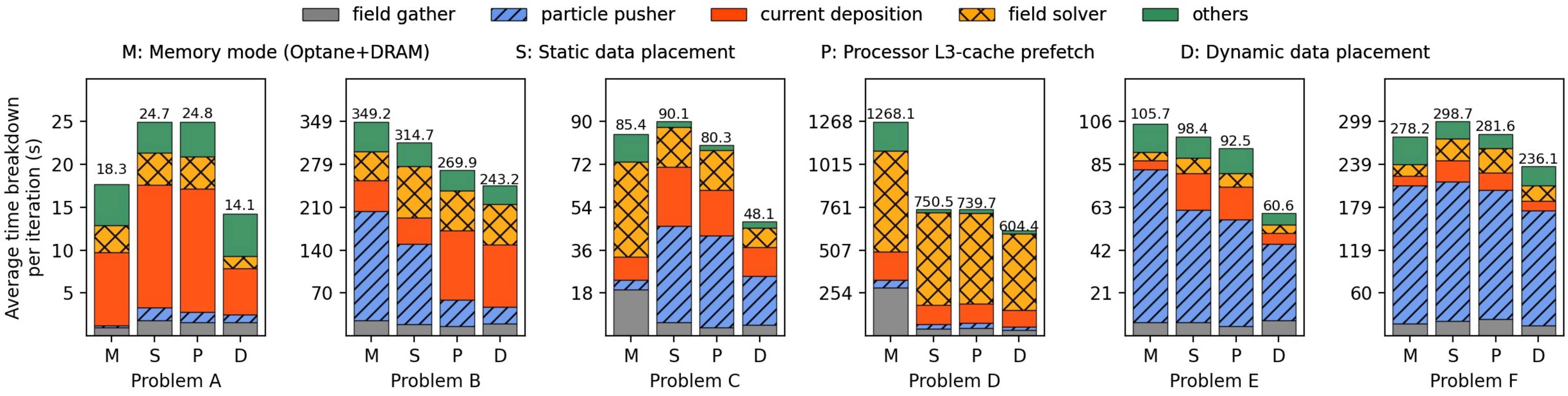
ID	Type	# of particles	Memory consumption
A	Laser-driven	1.1 billion	228.5 GiB
B	Laser-driven	8.4 billion	1.2 TiB
C	Beam-driven	2.1 billion	306 GiB
D	Beam-driven	10.7 billion	960 GiB
E	Uniform-plasma	3.7 billion	525 GiB
F	Uniform-plasma	8.6 billion	1.2 TiB

Overall performance



- WarpX-PM performs the best in all cases.
  - On average, WarpX-PM outperforms memory mode, Optane-only, and NUMA first-touch by 38.8%, 66.4%, and 45.1%.

Performance breakdown



- WarpX-PM with the three proposed techniques achieves the best performance in all problems.

Conclusions

- The emerging large-capacity persistent memory (PM) enables high-resolution large-scale scientific simulations.
- We use a combination of static and dynamic migrations and cache prefetching to make best use of the PM-based memory hierarchy.
- With a set of performance optimization strategies driven by detailed performance analysis, WarpX-PM outperforms DRAM-cached, the NUMA first-touch policy by 38.8% and 45.1% respectively.