The Energy Efficiency Research of Code for Numerical Simulation of Plasma Physics Problems

Igor Chernykh, Igor Kulikov, Vitaly Vshivkov,
Tatyana Liseykina

Institute of Computational Mathematics and Mathematical Geophysics SB RAS

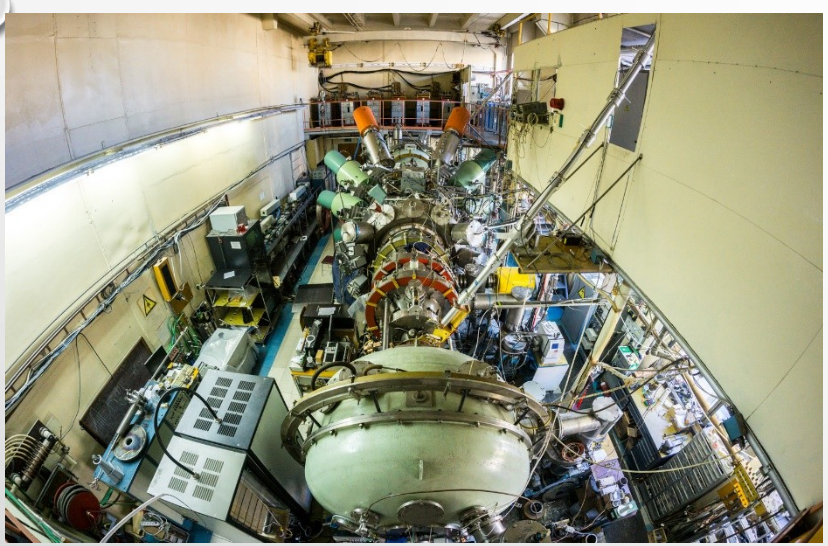
This work was supported by RSF grant No. 19-71-20026

SIBERIAN SUPERCOMPUTER CENTER ICMMG SB RAS

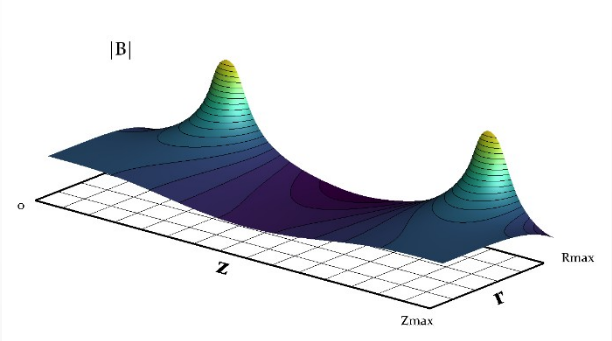
Overview of presentation

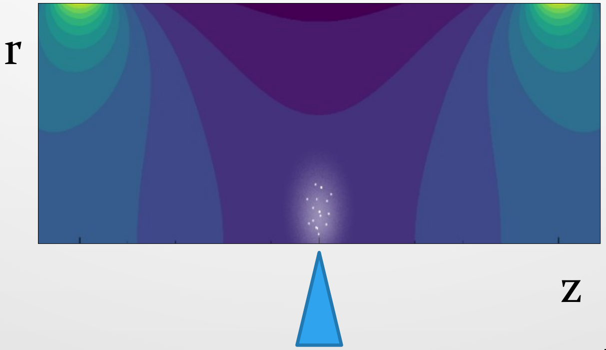
1. Overview of the problem
2. Motivation
3. Mathematical model
4. Numerical method
5. Parallel realization
6. Roofline model
7. Performance evaluation
8. Conclusion and future work

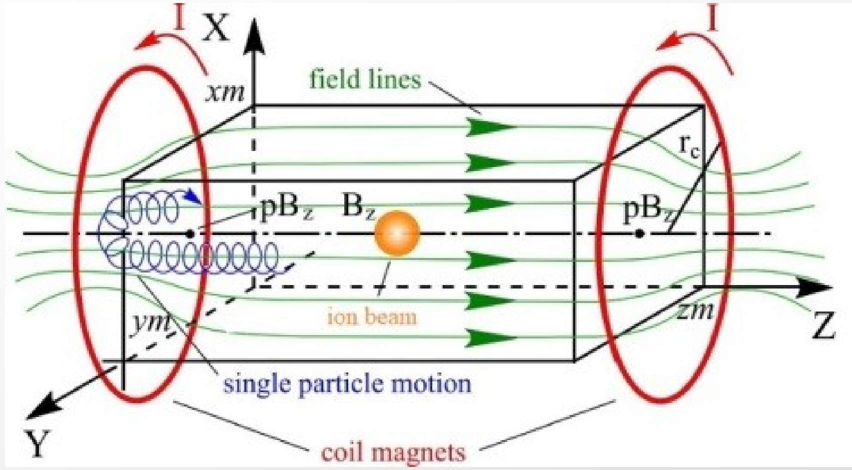
Overview of the problem



GOL3 – Plasma confinement facility with the external magnetic field generation



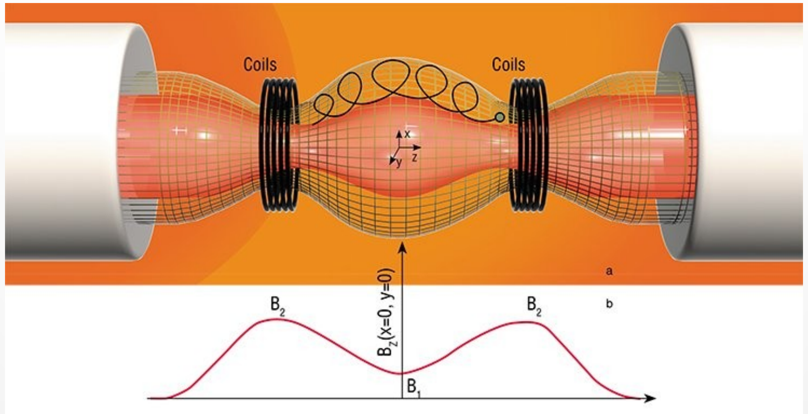




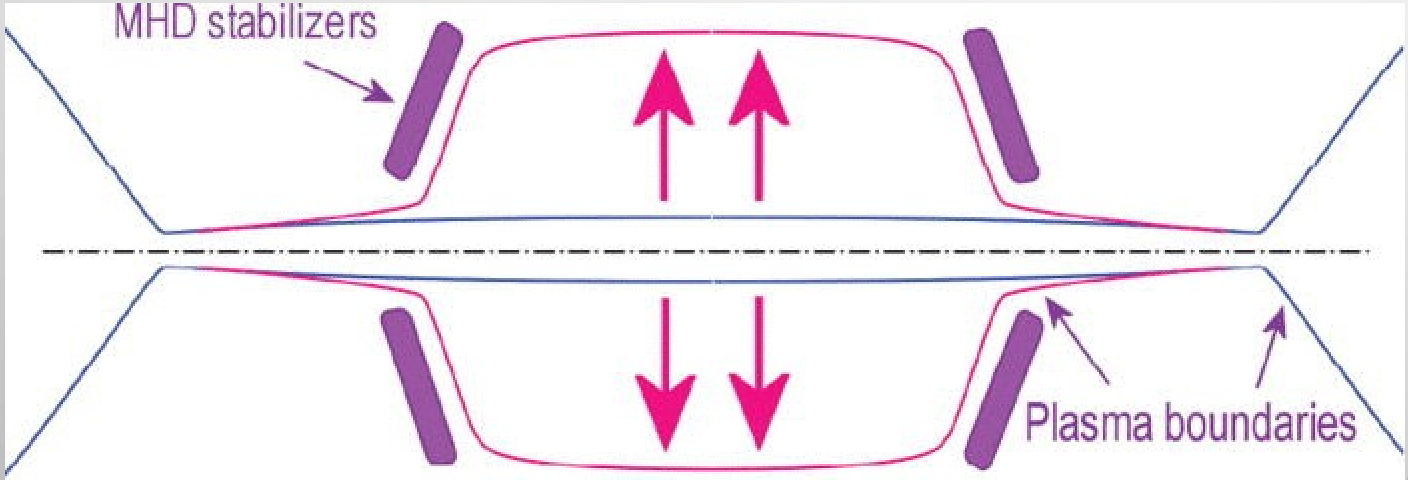
1. Cylindrical volume filled by ions
2. Two magnetic field sources
3. Injection of electrons at the center

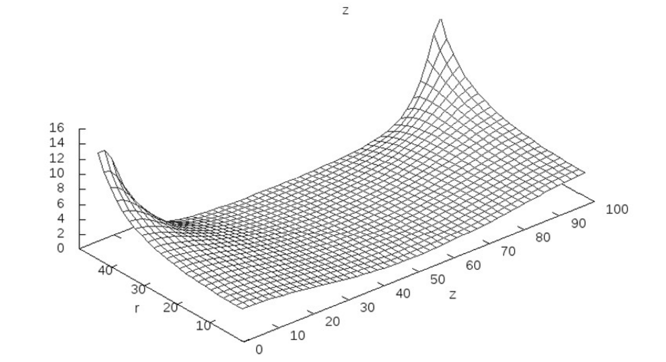
https://en.wikipedia.org/wiki/Budker_Institute_of_Nuclear_Physics

Overview of the problem



A. Shosin 23 Apr 2018 , Budker's Universe , volume Special Issue, N1





beam injection

A. D. Beklemishev· **Diamagnetic "bubble" equilibria in linear traps** Phys. Plasmas, 23, 082506 (2016)

Motivation

Simulate all parameters (magnetic field, power of electrons injection, etc) to find stable mode of GOL plasma confinement facilty.

↓

Diamagnetic equilibrium for thermonuclear fusion

Mathematical model

Vlasov kinetics equation (ions)

$$\frac{\partial f_i}{\partial t} + \vec{v}\frac{\partial f_i}{\partial \vec{r}} + \frac{\vec{F}_i}{m_i}\frac{\partial f_i}{\partial \vec{v}} = 0$$

$$\vec{F}_i = e\left(\vec{E} + \frac{1}{c}\left[\vec{v},\vec{B}\right]\right) + R_i$$

$$n_i(\vec{r}) = \int f_i(t,\vec{r},\vec{v})\,d\vec{v}$$

$$\vec{V}_i(\vec{r}) = \frac{1}{n_i(\vec{r})}\int \vec{v} f_i(t,\vec{r},\vec{v})\,d\vec{v}$$

MHD equations (electrons)

$$-e\vec{E} - \frac{e}{c}\left[\vec{V}_e,\vec{B}\right] - \frac{\nabla p_e}{n_e} + \vec{R}_e = m_e\frac{d\vec{V}}{dt}$$

$$\vec{R}_e = -m_e\frac{\vec{V}_e - \vec{V}_i}{\tau_{ei}}$$

Maxwell equations

$$\frac{1}{c}\frac{\partial \vec{E}}{\partial t} = rot\vec{B} - \frac{4\pi}{c}\vec{j}$$

$$\frac{1}{c}\frac{\partial \vec{B}}{\partial t} = -rot\vec{E}$$

$$div\vec{E} = 4\pi\rho$$

$$div\vec{B} = 0$$

$$n_i = n_e$$

$$R_i = -R_e$$

$$\vec{j} = e\left(n_i\vec{V}_i - n_e\vec{V}_e\right)$$

$$\rho = e\left(n_i - n_e\right)$$

Heat equations

$$n_e\left(\frac{\partial T_e}{\partial t} + \left(\vec{V}_e\nabla\right)T_e\right) = (\gamma - 1)\left(Q_e - div\vec{q}_e - p_e div\vec{V}_e\right)$$

$$Q_e = \frac{j^2}{\sigma}$$

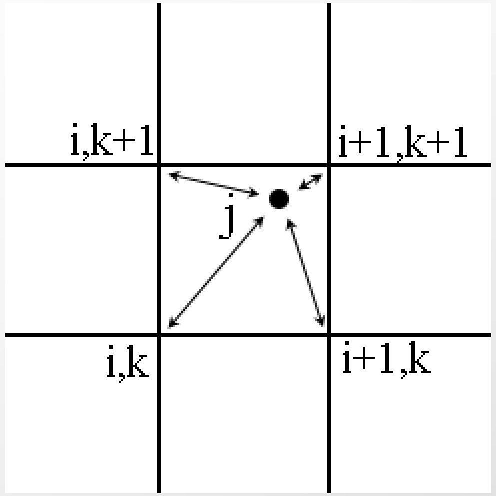
$$\vec{q}_e = -k\nabla Te$$

State equation

$$p_e = n_e T_e$$

Numerical method (Finite-Difference Time-Domain (FDTD))

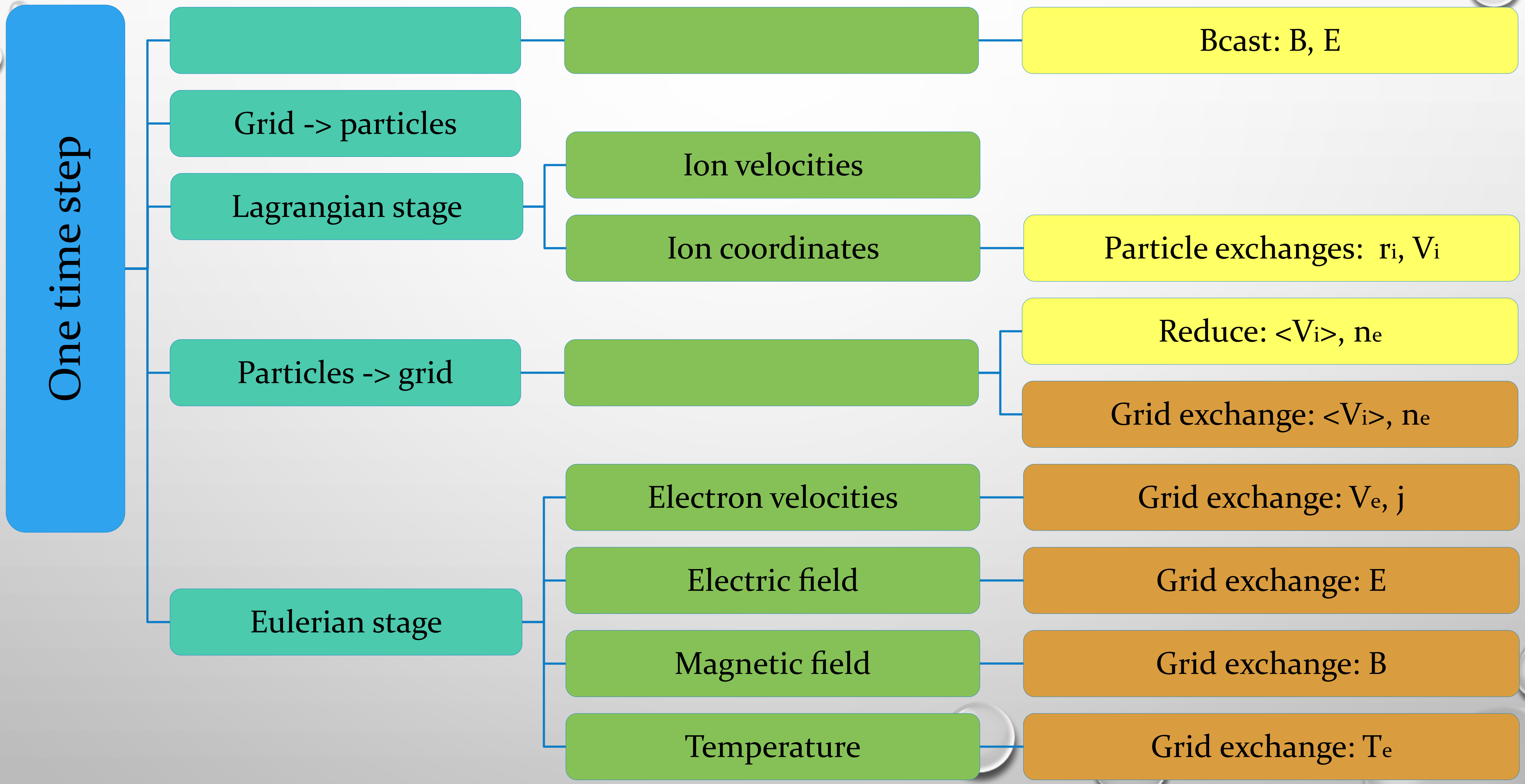
- ✓ Cylinder coordinate system due to the axial symmetry, r=0 is singularity
- ✓ Hybrid model, the ions are described kinetically, the electrons by MHD
- ✓ Particles-in-cell method with PIC form-factor
- ✓ Grids shifted by half a step
- ✓Mixed Eulerian-Lagrangian decomposition for the parallelization



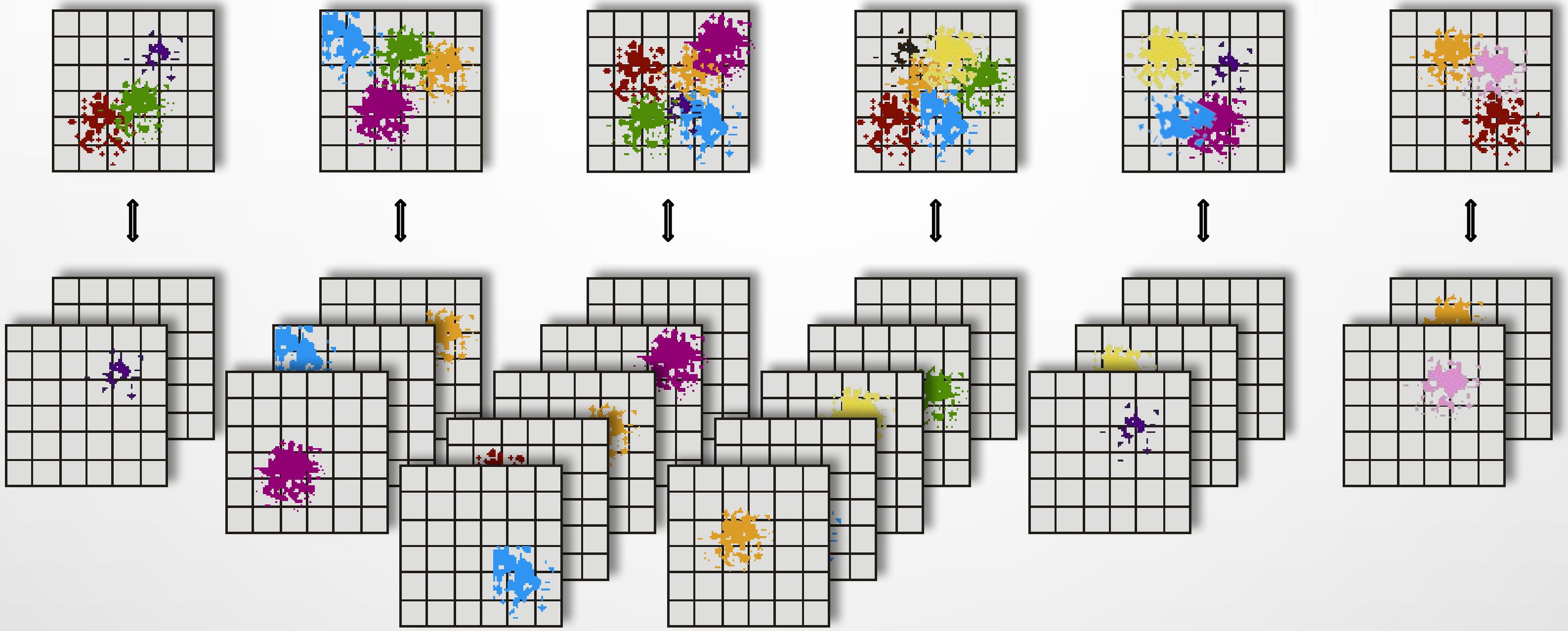
particles >> cells

https://warpx.readthedocs.io/en/latest/theory/picsar_theory.html

Parallel realization



Parallel realization



At the initial moment, the background particles are distributed within the group.

The injected particle is written to the next nucleus in its group.

When leaving the subarea, the array of particles is sent to one of the processors of the neighboring group.

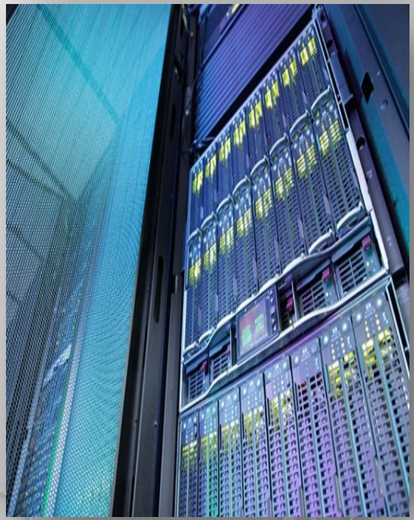
Siberian Supercomputer Center

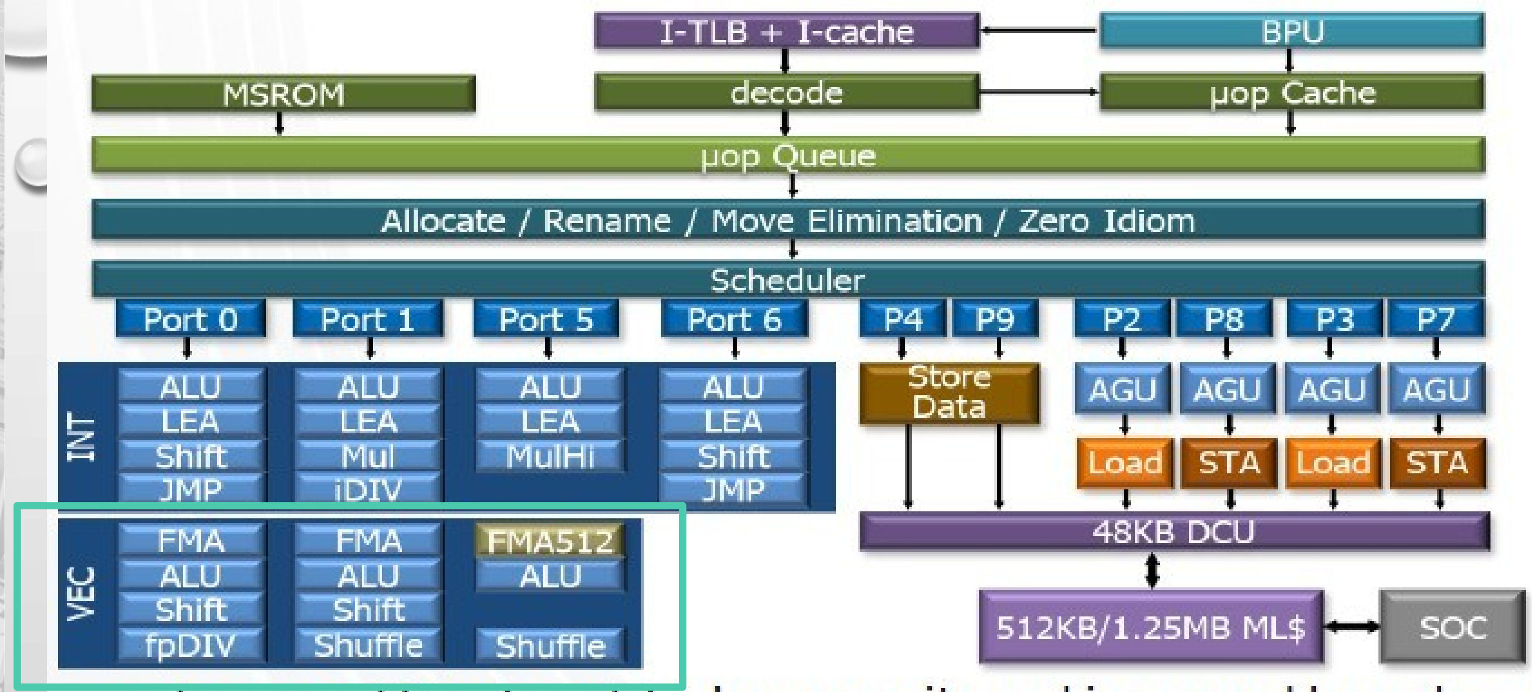
NKS-1P (RSC, hot water cooling, 2448 cores, ~182TFLOPS Rpeak):

- 27 nodes: 2 CPU Intel Xeon E5-2697v4 [128 GB DDR4, 256 GB DDR4] (864 cores, 2.6GHz) (1 узел 2x375GB Intel Optane [IMDT])
- 16 nodes: 1 CPU Intel Xeon Phi 7290 KNL [16 GB MCDRAM+96 GB DDR4] (1152 cores, 1.5-1.7 GHz)
- 1 node: 2 CPU Intel Xeon Platinum 8268 [192 GB DDR4] (48 cores, 2.9 GHz)
- 8 nodes: 2 CPU Intel Xeon Gold 6248R [192/384/768 GB DDR4] (384 cores, 2.9 GHz)
- Intel OmniPath 100 Gb/s
- Intel Lustre – 200 TB + NFS 100TB

NKS-30T (HP, air cooling, ~1500 CPU cores (2.9GHz), ~30000 GPU cores, 85TFLOPS (GPU) + 22TFLOPS (CPU):

- 576 CPU Intel Xeon E5450/E5540(2688 cores)
- 80 CPU Intel Xeon X5670(480 cores)
- 120 GPU NVIDIA Tesla M 2090(61440 cores)
- Infiniband QDR 40 Gb/s
- HP Ibrix – 90 TB



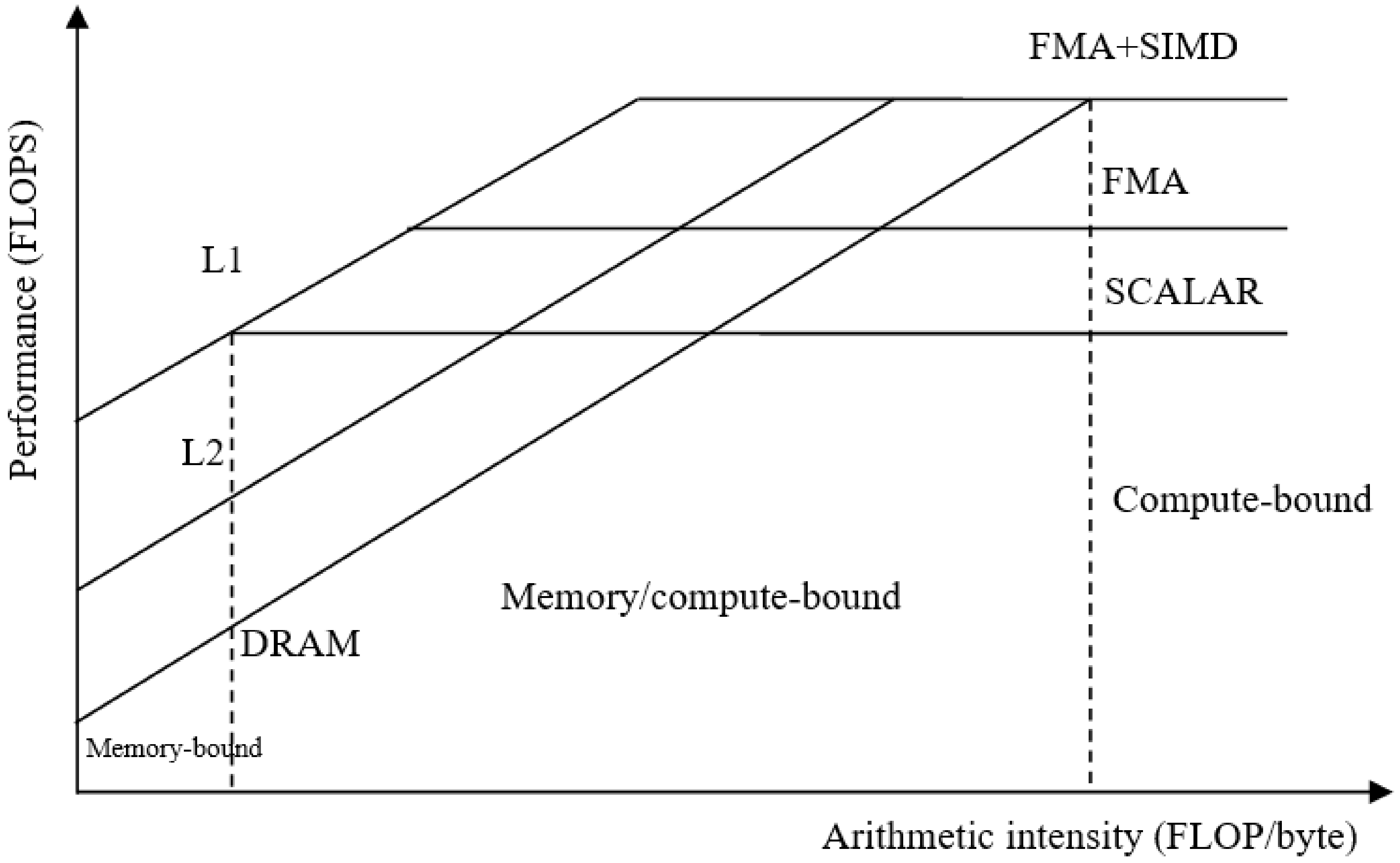


	Cascade Lake (per core)	Ice Lak (per core
Out-of-order Window	224	384
In-flight Loads + Stores	72 + 56	128 + 7
Scheduler Entries	97	160
Register Files – Integer + FP	180 + 168	280 +2
Allocation Queue	64/thread	70/threa 140/1 thr
L1D Cache (KB)	32	48
L1D BW (B/Cyc) – Load + Store	128 + 64	128 + 6
L2 Unified TLB	1.5K	2K
Mid-level Cache (MB)	1	1.25

- Improved Front-end: higher capacity and improved branch predictor
- Wider and deeper machine: wider allocation and execution resources + larger structures
- Enhancements in TLBs, single thread execution, prefetching
- Server enhancements – larger Mid-level Cache (L2) + second FMA

~18% Increase In IPC On Existing SPECcpu2017(est) Integer Rate Binaries

Roofline model

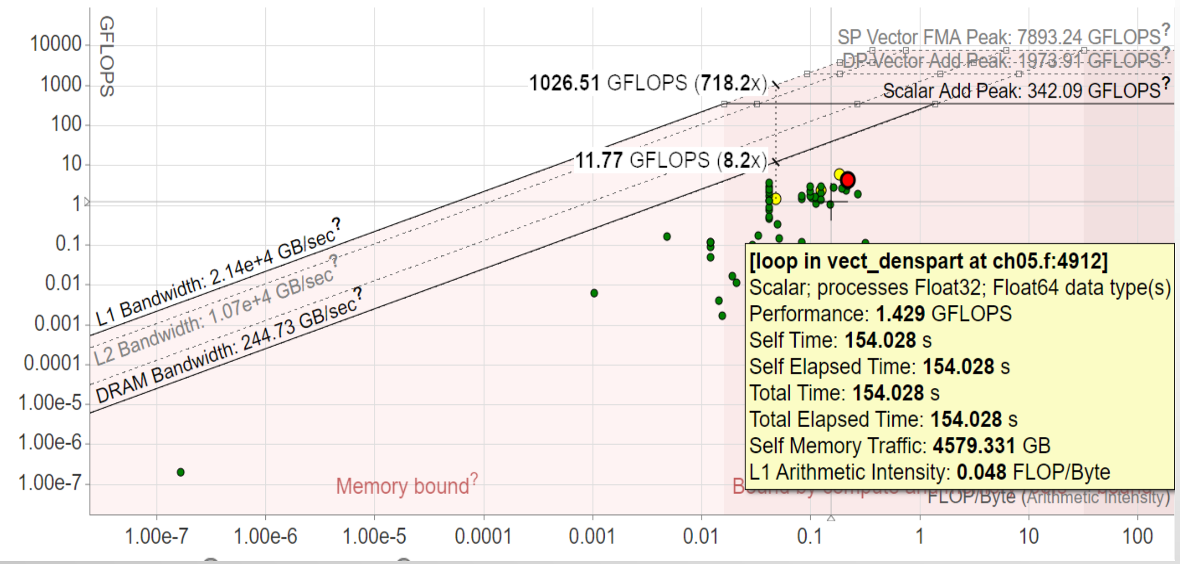


Roofline model

Diamagnetic trap

Roofline model of parallel code

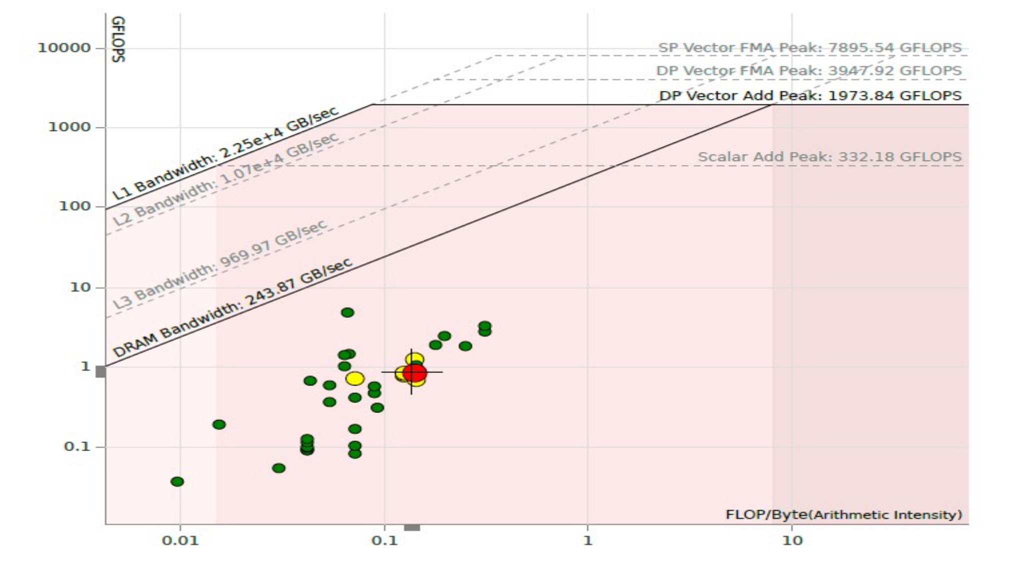
2x6248R, non-SIMD, 1.23 GFLOPS



Laser-plasma interaction

Roofline model of parallel code

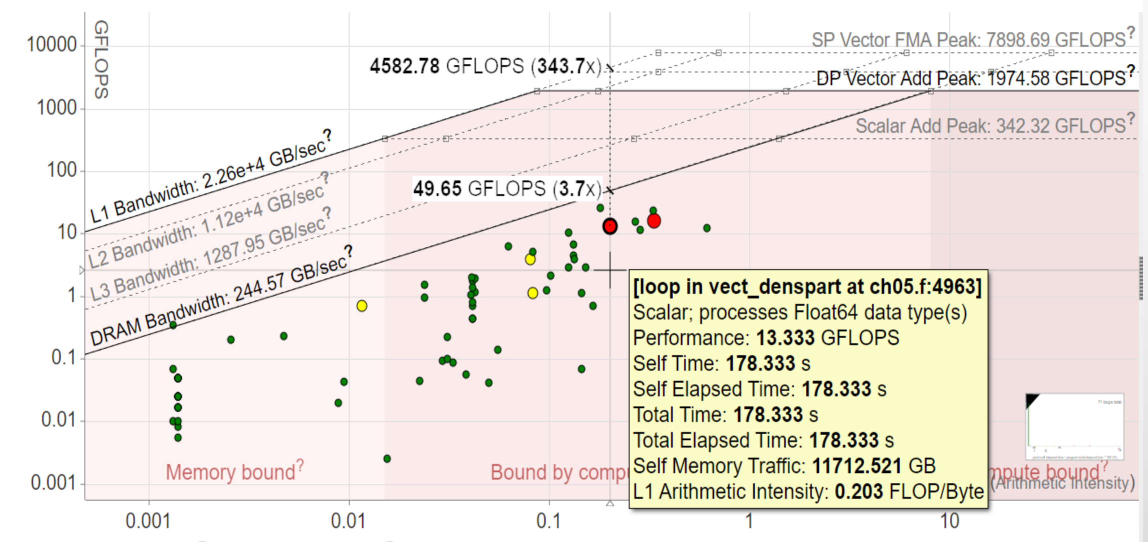
2x6248R, non-SIMD, 0.87 GFLOPS



Diamagnetic trap

Roofline model of parallel code

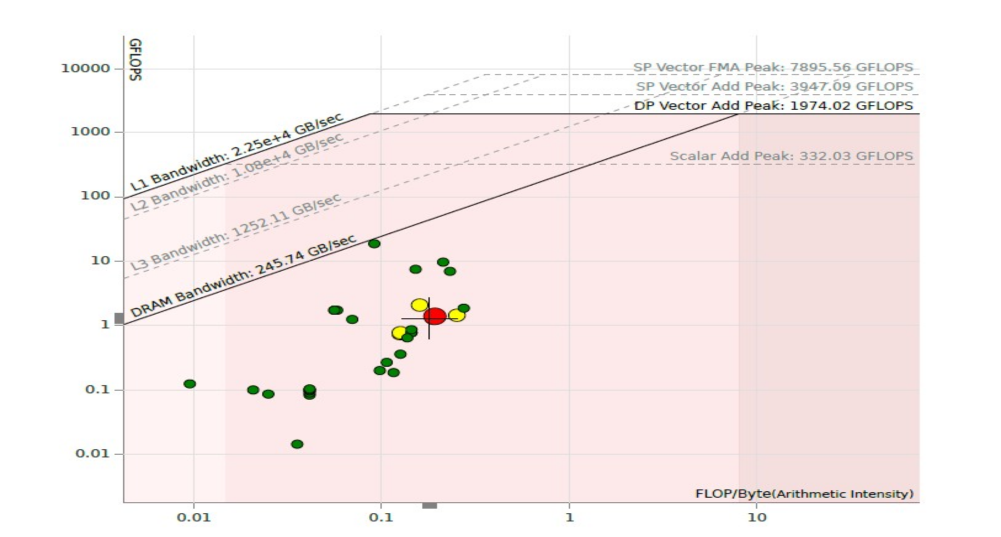
2x6248R, AVX512, autovec, 2.7 GFLOPS



Laser-plasma interaction

Roofline model of parallel code

2x6248R, AVX512,autovec, 1.31 GFLOPS



SIBERIAN SUPERCOMPUTER CENTER ICMMG SB RAS

Energy efficiency evaluation

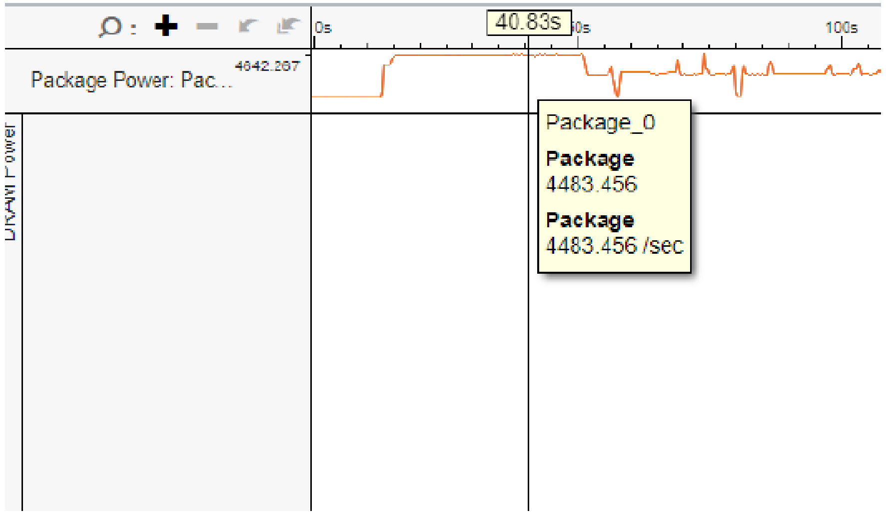


Fig. 5. Maximum package power consumtion during initial data distribution calculations (mW/sec).

Table 1. Energy efficiency results for new PIC code

CPU	Vectorization	FLOPS/mW
1xIntel Core i9-10980XE	AVX2	0.15
1xIntel Core i9-10980XE	AVX512	0.21
2xIntel Xeon 8268	AVX2	0.21
2xIntel Xeon 8268	AVX512	0.30

Grouping: Profiled Entity Hierarchy

Profiled Entity Hierarchy	Package Power Rate (mW): Self by Type	Package Po...
	Power	Power
Total	0.000	0.000
Total → CPU	0.000	0.000
Package_0	3155.915	1135180.542

Fig. 6. Package power rate (mW/sec) for AVX2 version of PIC code.

Grouping: Profiled Entity Hierarchy

Profiled Entity Hierarchy	Package Power Rate (mW): Self by Type	Package Pow...
	Power	Power
Total	0.000	0.000
Total → CPU	0.000	0.000
Package_0	3397.857	1223261.372

Fig. 7. Package power rate (mW/sec) for AVX512 version of PIC code.

SIBERIAN SUPERCOMPUTER CENTER ICMMG SB RAS

Conclusion

1. Update your system software and compilers.
2. Use advanced vector instructions (AVX2, AVX512, AMX) together with MPI and OpenMP in your code.
3. Use all compiler optimization options to build the fastest code for your CPU architecture. (mpiifort –xCOMMON-AVX512)
4. We achieved 7% growth of FLOPS/Watts value only by changing the auto-vectorization optimization level from AVX2 to AVX512 instructions.

Future work: Vectorized adding and/or multiplying

```
double A[vec_width], B[vec_width];
for(int i = 0; i < vec_width; i++)
A[i]+=B[i];
```

```
double A[vec_width], B[vec_width];
__m512d A_vec = _mm512_load_pd(A);
__m512d B_vec = _mm512_load_pd(B);
A_vec = _mm512_add_pd(A_vec,B_vec);
_mm512_store_pd(A,A_vec);
```

Thank you for your attention