Massively parallel PSATD solver for PIC codes

An arbitrary scalable parallelization technique

Haithem Kallala

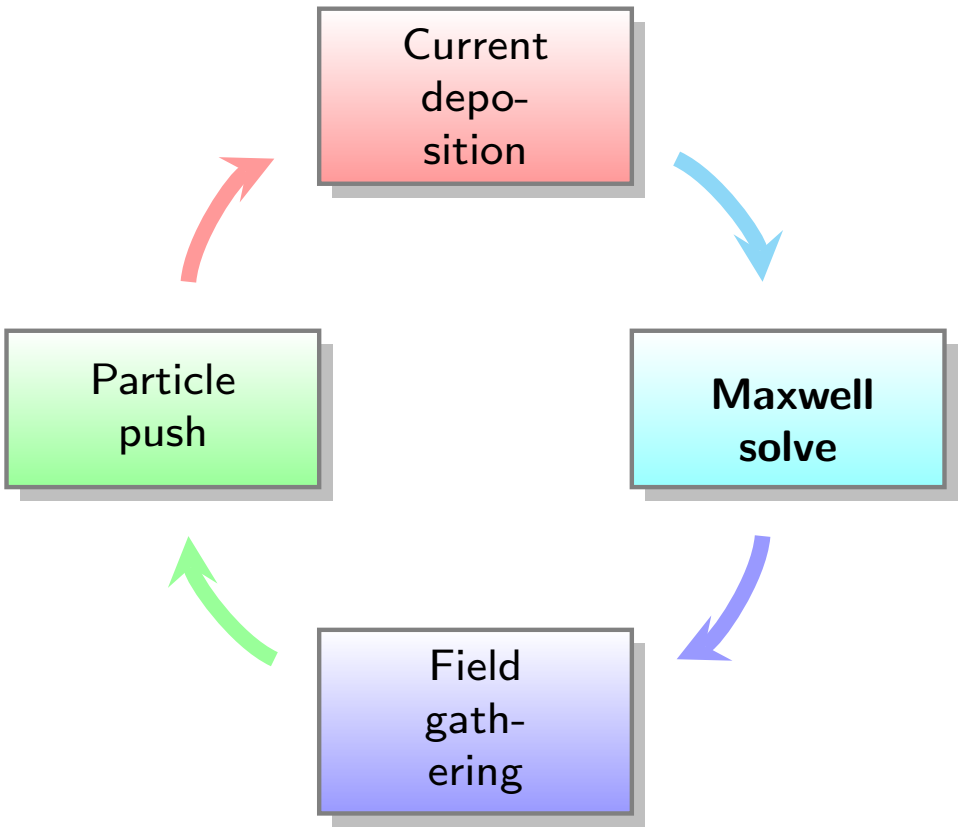
CEA Saclay

BLAST workshop 2018

May 7, 2018

Overview of the PIC algorithm

- The PIC algorithm is an essential numerical tool to model kinetic effects in plasmas.
- Efficient PIC codes need to be highly scalable to take advantage of massively parallel architecture.



Maxwell's equations solvers

Maxwell's equations

- $\vec{\nabla} \wedge \vec{E} = -\frac{\partial \vec{B}}{\partial t}$
- $\vec{\nabla} \wedge \vec{B} = \mu_0 \vec{J} + \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t}$

Numerical resolution:

- Finite difference method.
- Pseudo Spectral method.

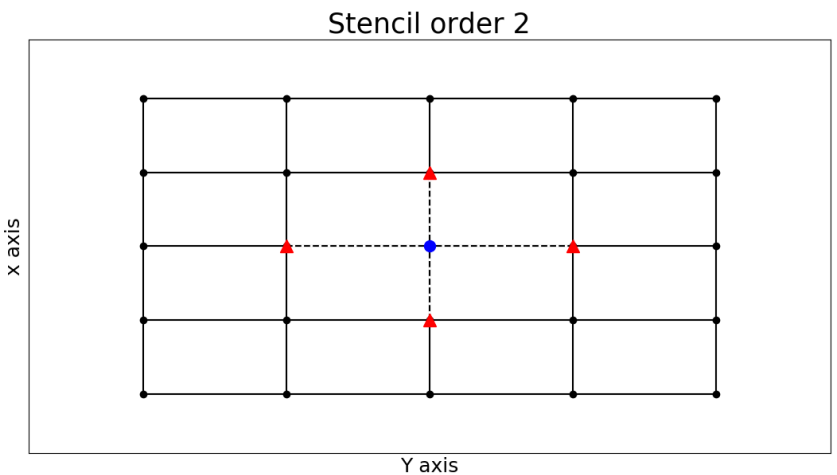
Finite Difference Time Domain

Maxwell's equations

- $\vec{\nabla} \wedge \vec{E} = -\frac{\partial \vec{B}}{\partial t}$
- $\vec{\nabla} \wedge \vec{B} = \mu_0 \vec{J} + \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t}$

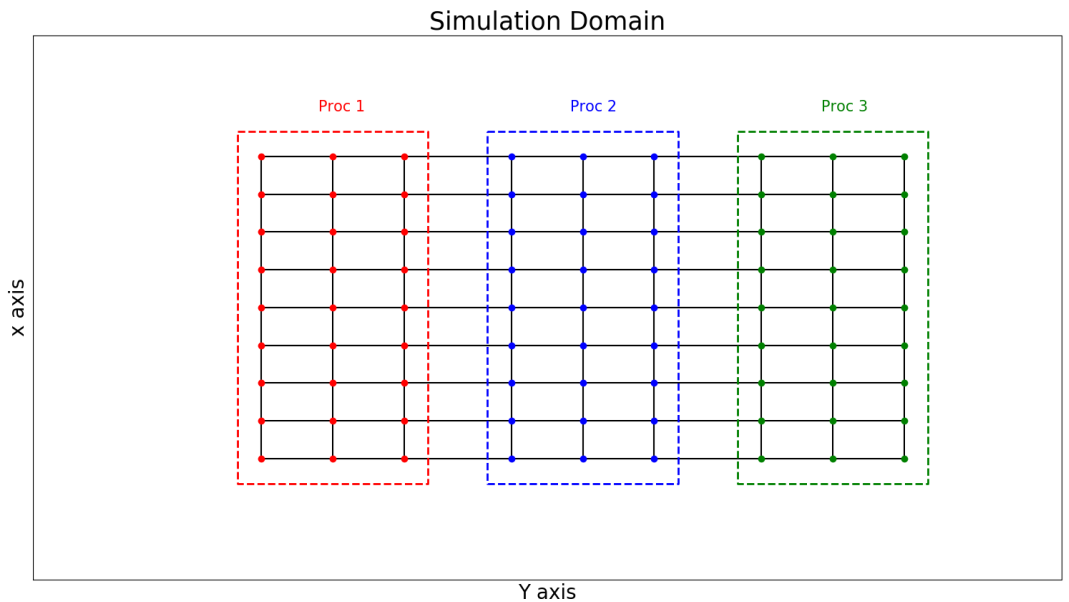
Finite Difference Time Domain

- $\partial t =>$ 2^{nd} order scheme
- $\vec{\nabla} =>$ 2^{nd} order scheme



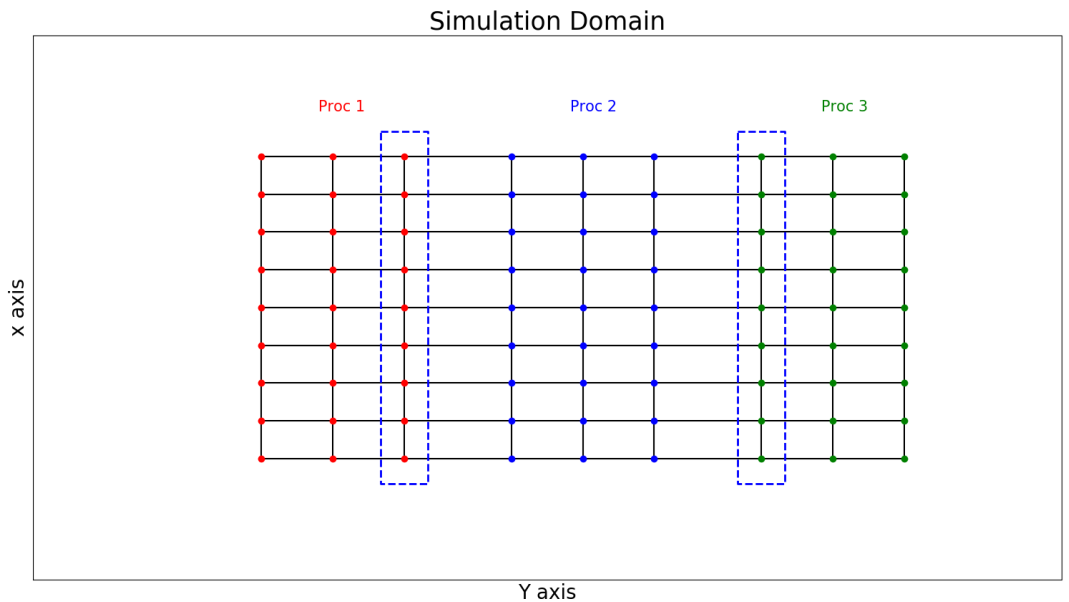
Finite Difference Time Domain

- Local computations.
- Allows parallel implementation



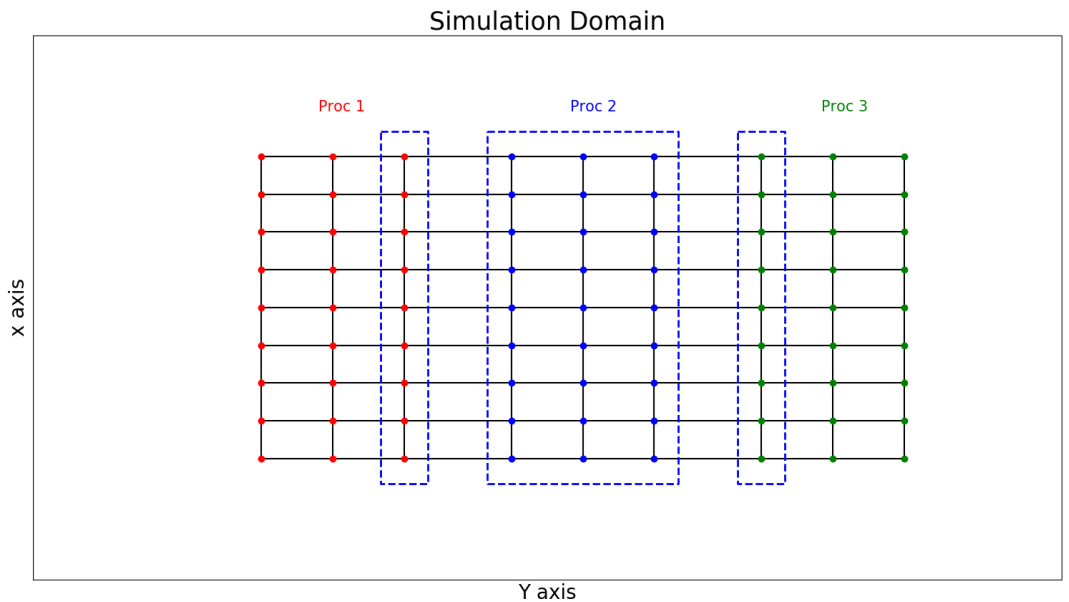
Finite Difference Time Domain

- Local computations.
- Allows parallel implementation



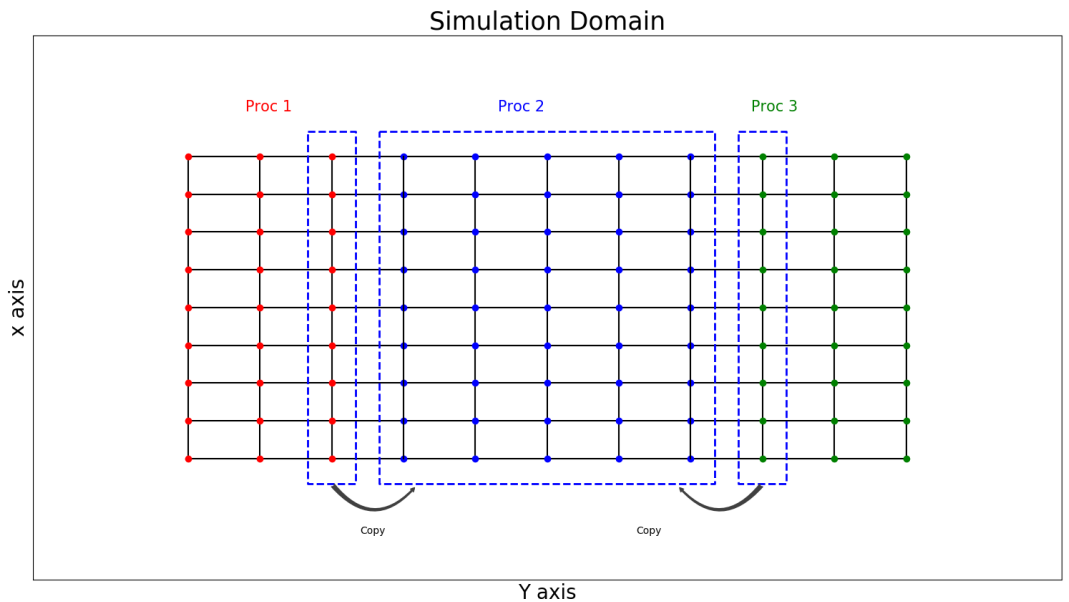
Finite Difference Time Domain

- Local computations.
- Allows parallel implementation



Finite Difference Time Domain

- Local computations.
- Allows parallel implementation



Finite Difference Time Domain

- Low MPI communications.
- Highly scalable.
- Induces important numerical dispersion.

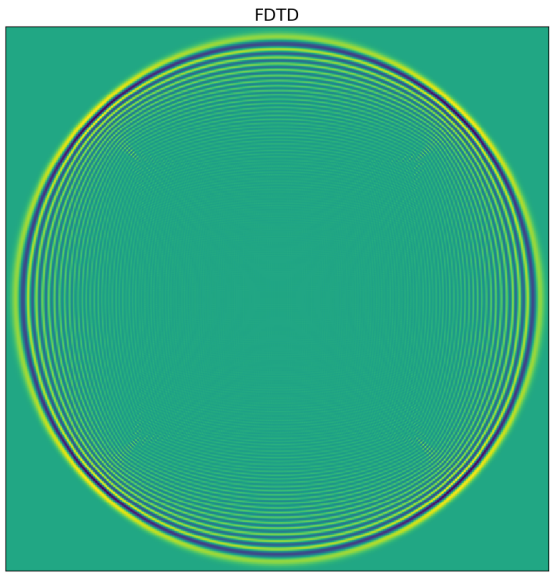


Figure: Dirac pulse propagation with FDTD scheme

Pseudo Spectral Analytical Time Domain

- The PSATD algorithm is a dispersion-free **FFT-based** algorithm that solves Maxwell's equations in Fourier space.

$$i\vec{k} \wedge \hat{\vec{B}} = \mu_0 \hat{\vec{J}} + \frac{1}{c^2} \frac{\partial \hat{\vec{E}}}{\partial t}$$
$$i\vec{k} \wedge \hat{E} = -\frac{\partial \hat{\vec{B}}}{\partial t}$$

PSATD

- $\partial t =>$ Analytical integration
- $\vec{\nabla} => i\vec{k}$ Infinite order in space

Pseudo Spectral Analytical Time Domain

- Allows analytical integration over time under weak assumptions.
- No CFL condition.
- Dispersion free.

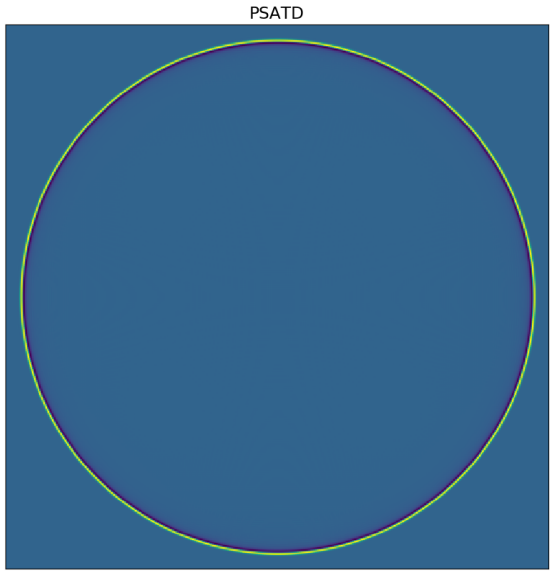
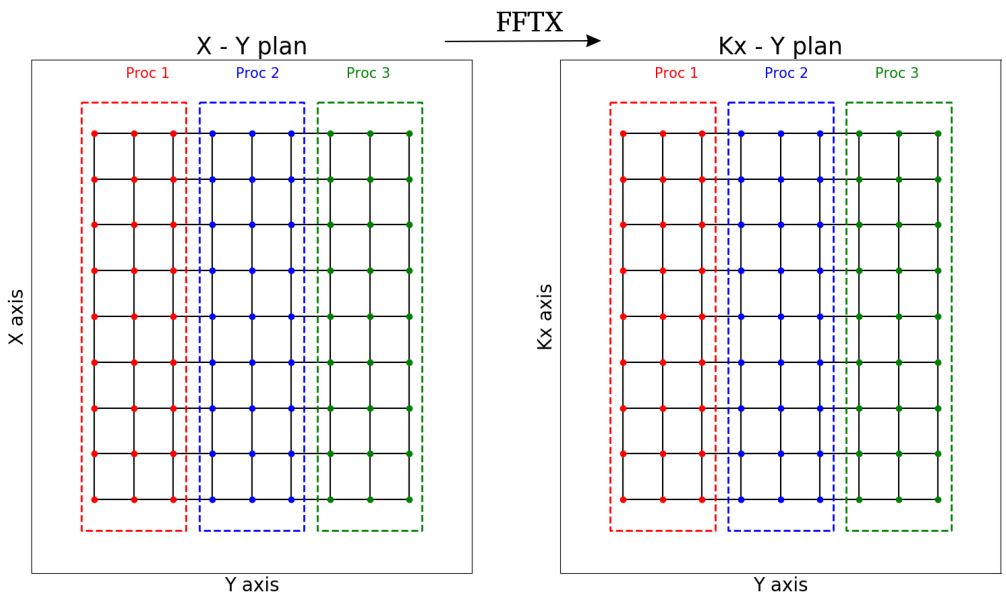


Figure: Dirac pulse propagation with PSATD scheme

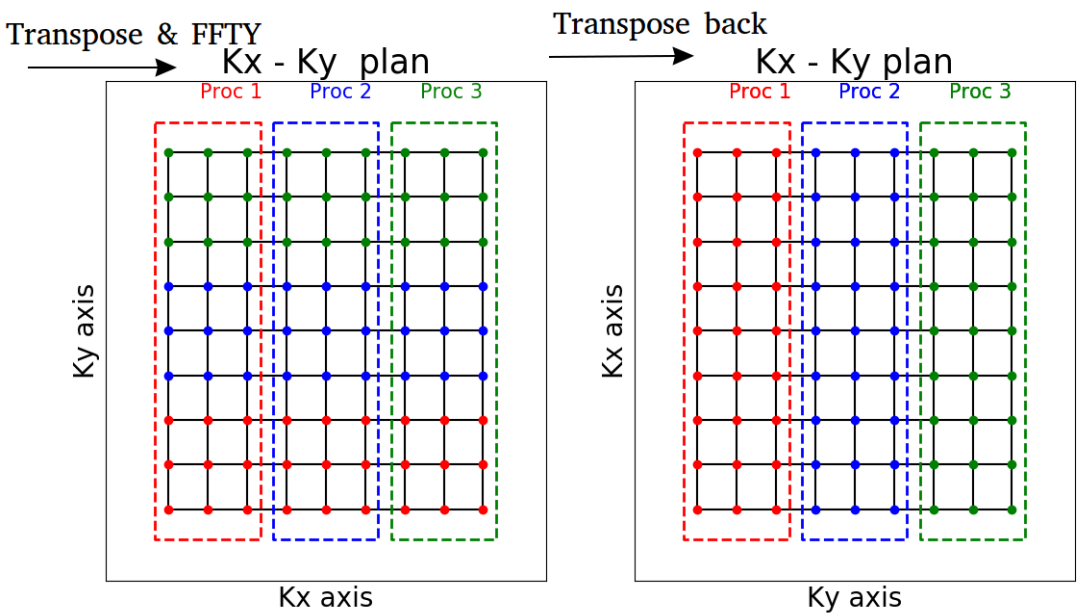
Pseudo Spectral Analytical Time Domain

- Voluminous MPI communications are required to perform distributed-memory FFTs.
- Global communications involving MPI_ALLTOALL calls.
- Extremely hard to scale. (Above few thousands cores)



Pseudo Spectral Analytical Time Domain

- Voluminous MPI communications are required to perform distributed-memory FFTs.
- Global communications involving MPI_ALLTOALL calls.
- Extremely hard to scale. (Above few thousands cores)



Pseudo Spectral Analytical Time Domain

PSATD with standard domain decomposition [1]

- Induces a truncation error at subdomain boundaries [2].
- This error is decreased when using a finite (but arbitrarily high) order stencil in space
  $j\vec{k} = \hat{\nabla}_p = \sum_{i=1}^{p/2} 2jc_i sin(2\pi ik/N)$
- Decreases quickly with the number of guardcells.
- Benchmarked over different physical regimes[3].

[1] J.-L. Vay, I. Haber, and B. B. Godfrey. J. Comput. Phys., 2013.
[2] H. Vincenti and J.-L. Vay. Comput. Phys. Comm
[3] G. Blaclard, H. Vincenti, R. Lehe, and J. L. Vay Phys. Rev. E 96, 033305

Pseudo Spectral Analytical Time Domain

Local Pseudo-Spectral Time domain

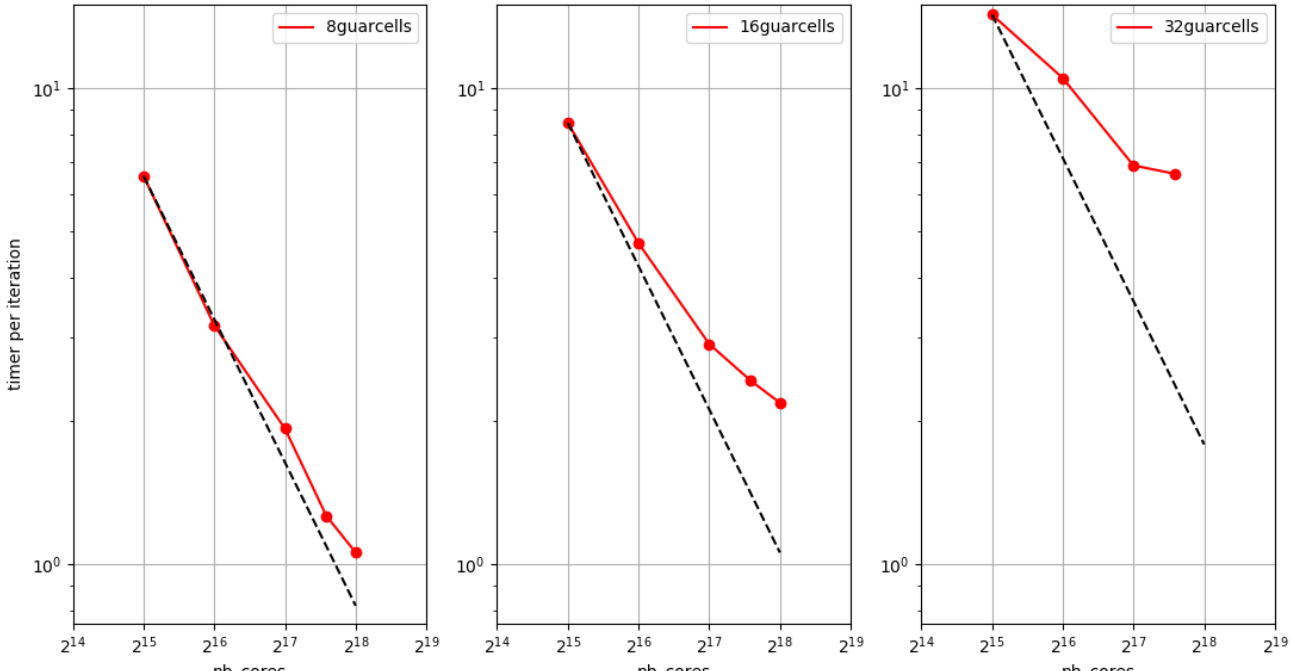
- Highly scalable (weak scaling). [4]
- Nearly dispersion free with a high order stencil.
- Available in WARP+PXR

[4] H. Vincenti and J-L. Vay, Comp. Phys. Comm 2018

Pseudo Spectral Analytical Time Domain

Local Pseudo-Spectral Time domain

- Memory footprint increases quickly due to data redundancy (up to x27 in 3D)
- Strong Scaling is hard to achieve with high number of guardcells.

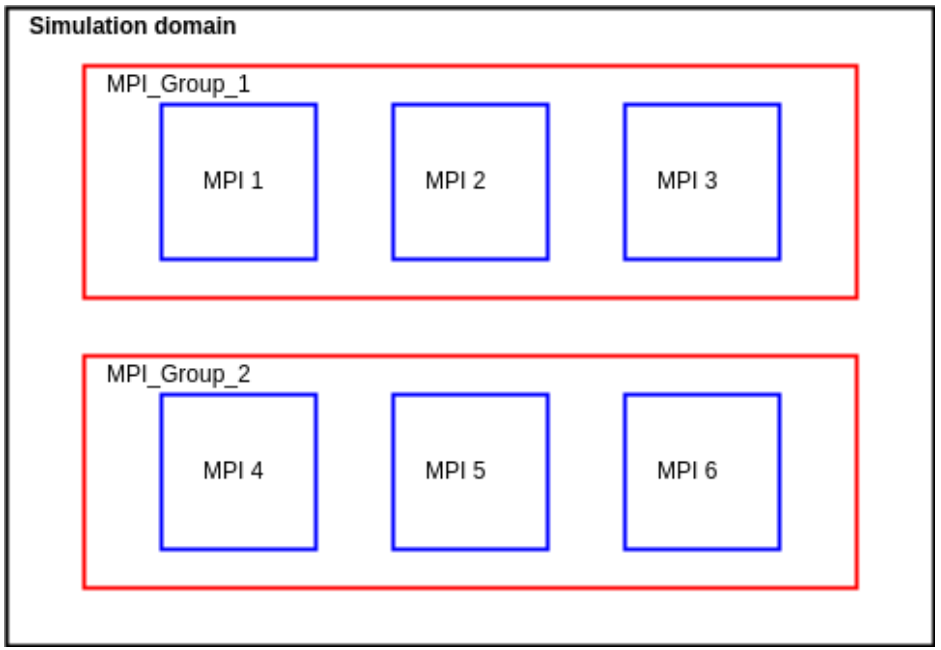


Hybrid Pseudo Spectral Analytical Time Domain

- How to achieve strong scaling on hundreds of thousands of nodes ?
- Novel parallelization technique.

General Idea

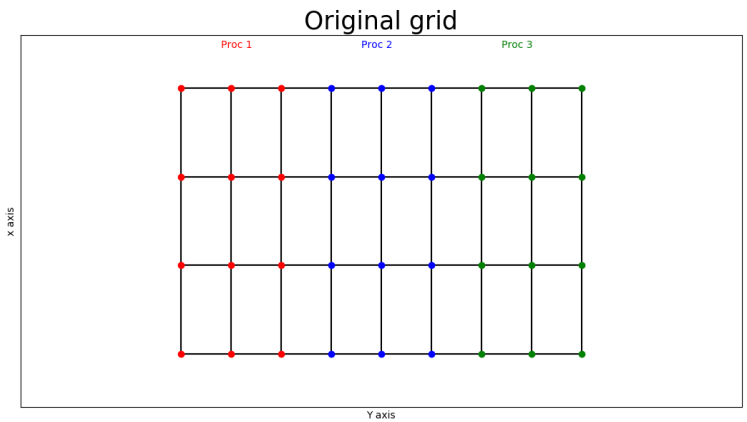
- MPI subdomains are grouped into MPI clusters, called groups.
- Maxwell's equations are solved with PSATD scheme on each group using **distributed memory FFT.**

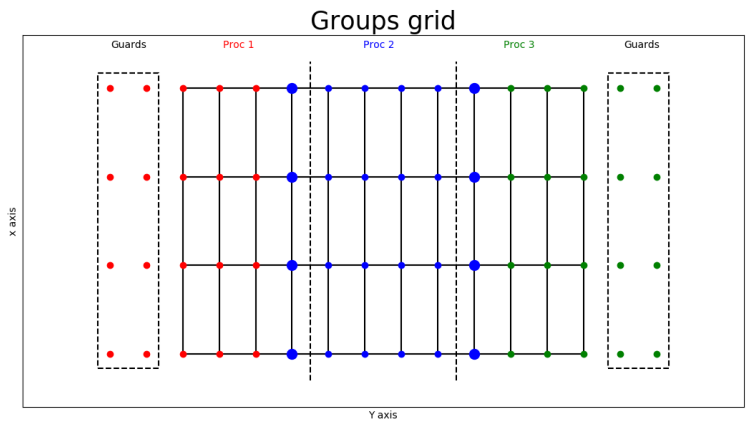


Hybrid Pseudo Spectral Analytical Time Domain

Load Balancing

- Using the hybrid technique requires to define a new grid appended to the groups topology. The two grids need to be linked by an optimized load balancing strategy.

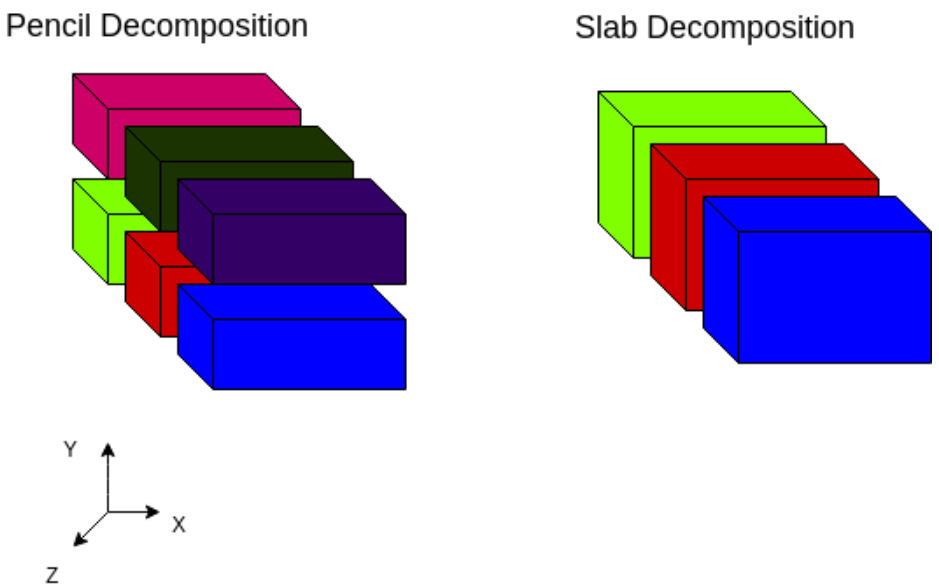




Hybrid Pseudo Spectral Analytical Time Domain

Pencil vs Slab FFT?

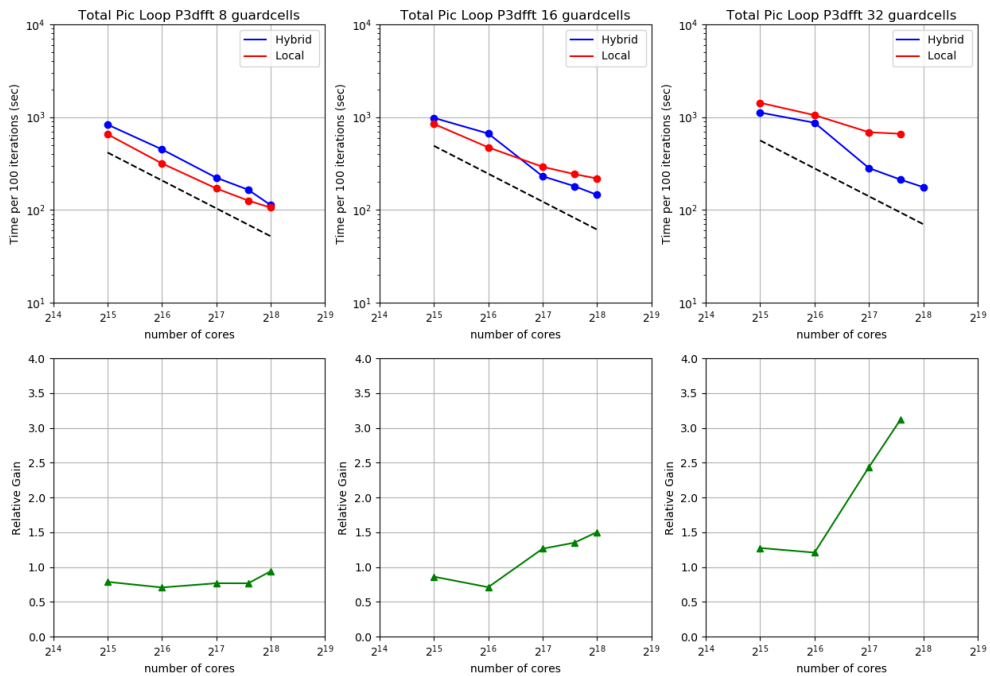
- Grouping MPI depends on the FFT library used.
- Pencil decomposition allows data distribution along two axes.
- Slab decomposition allows only data distribution along one axis.
- Data can always be distributed along X axis as for the local technique (one MPI task per group along X axis).



Hybrid Pseudo Spectral Analytical Time Domain

Strong Scaling benchmark: Pencil technique

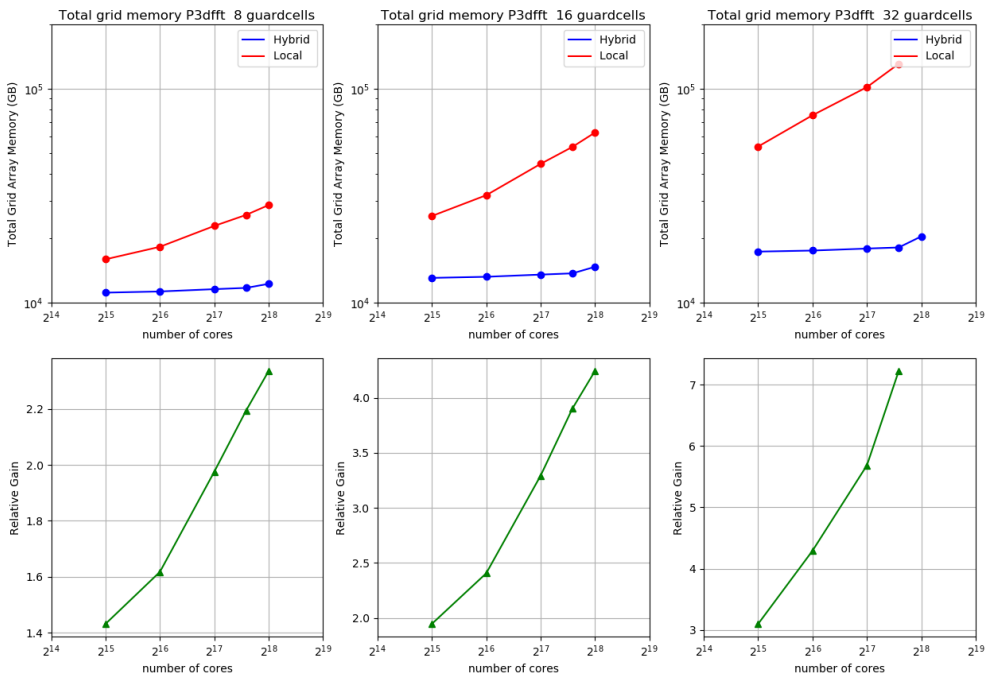
- From 32768 to 262144 cores on THETA (KNL architecture)
- 241 x 6145 x 12289 grid points
- Guard cells number : 8, 16, 32
- 32 MPI per group



Hybrid Pseudo Spectral Analytical Time Domain

Strong Scaling benchmark: Pencil technique

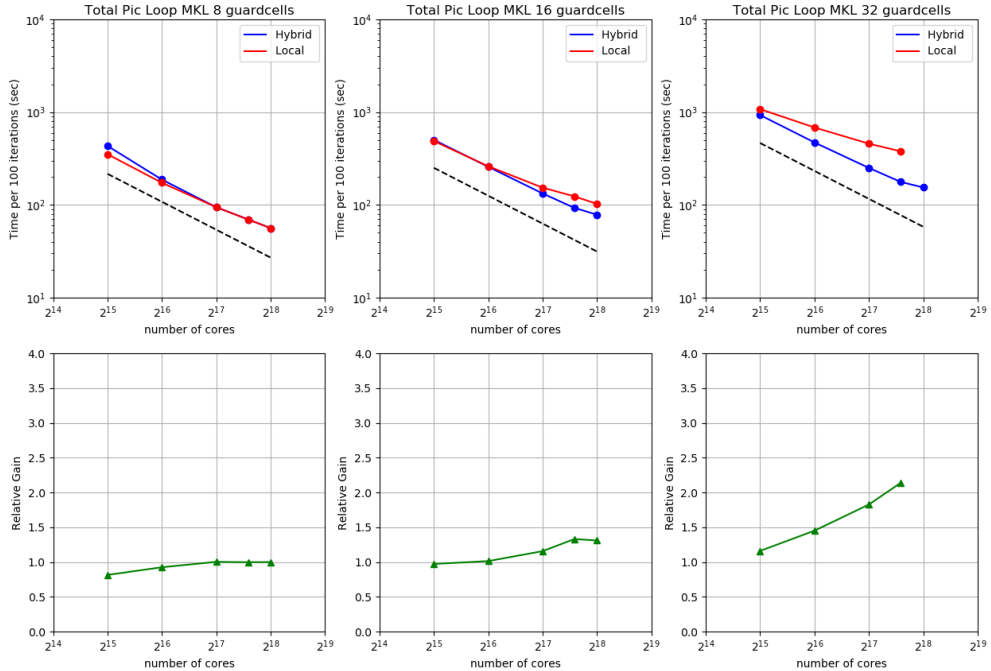
- From 32768 to 262144 cores on THETA (KNL architecture)
- 241 x 6145 x 12289 grid points
- Guard cells number : 8, 16, 32
- 32 MPI per group



Hybrid Pseudo Spectral Analytical Time Domain

Strong Scaling benchmark: Slab technique

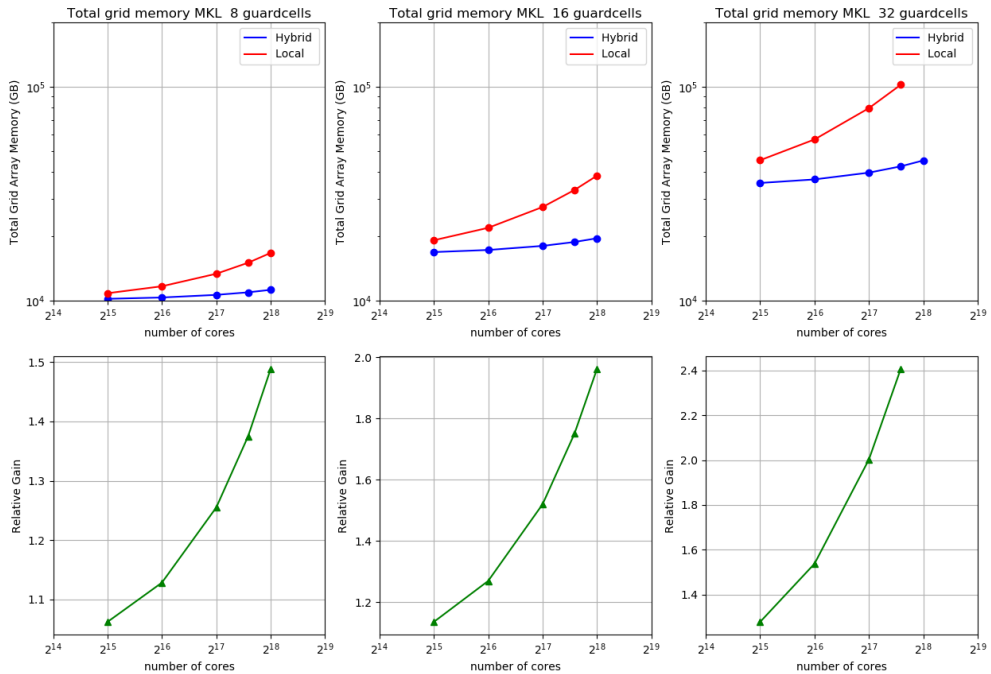
- 161x161x393217 grid points
- Guard cells number : 8, 16, 32
- 8 MPI per group



Hybrid Pseudo Spectral Analytical Time Domain

Strong Scaling benchmark: Slab technique

- 161x161x393217 grid points
- Guard cells number : 8, 16, 32
- 8 MPI per group



Conclusion

- Benchmarks show that the hybrid technique performs better than the local technique especially with an increased number of guardcells.
- There is an optimum number of MPIs per group that depends on the number of grid points and guardcells (optimum between data redundancy and MPI_ALLTOALL overhead).
- The memory footprint and global performance gain is very important (x7 and x4 respectively) with the Pencil technique.

Any questions?