HPC/Exascale Centre of Excellence in Personalised Medicine

PhysiCell-X: an HPC-ready multiscale simulation tool for personalised medicine

Arnau Montagud

Barcelona Supercomputing Center (BSC)

CASTIEL2 webinar series "Code of the Month"

The PerMedCoE project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement Nº951773

PerMedCoE aims to integrate PerMed into the new European HPC/Exascale ecosystem

- Physiological-level models (Fluid dynamics – CompBioMed)

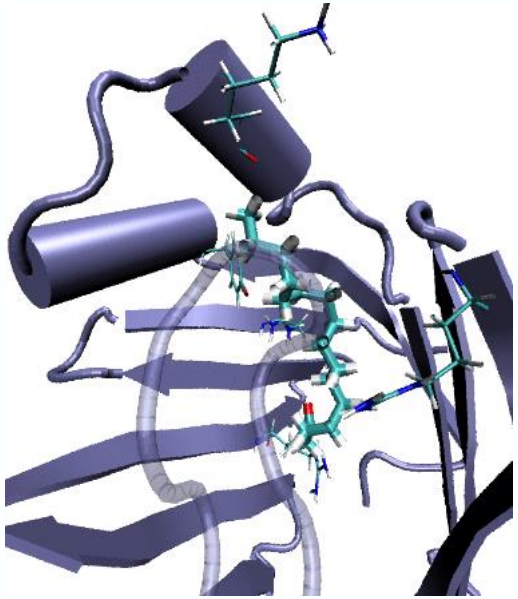
Cell-environment interactions

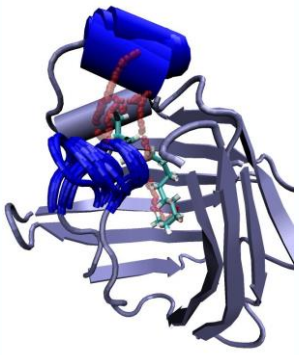
Cell-cell interactions

Metabolic pathways

Signalling pathways

Cell-level models



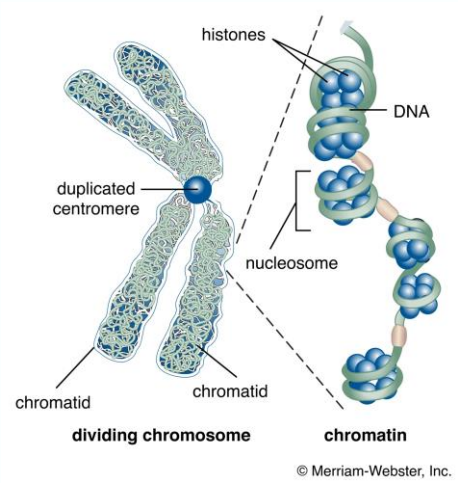


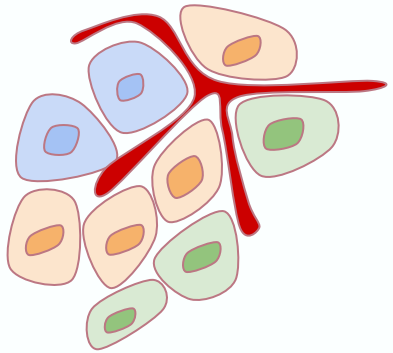
- Atomic-level models (Molecular dynamics – BioExcel)

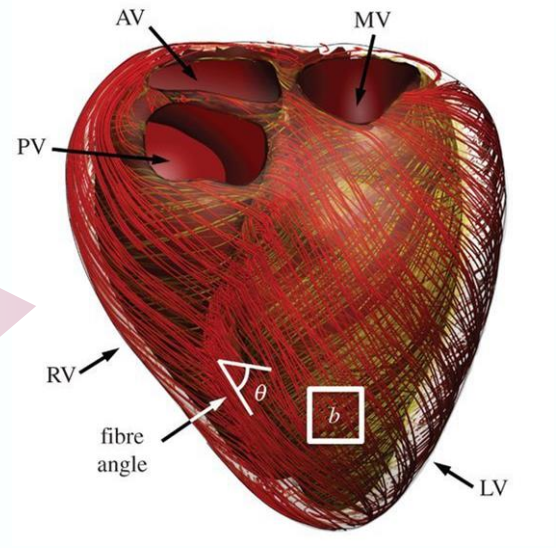
Europace. 2019 May, 21(5): 822–832.
J. Chem. Theo. Comp. 2005, 6, 1304-1311

Simulations in Life Sciences

From cells to tissues and organs



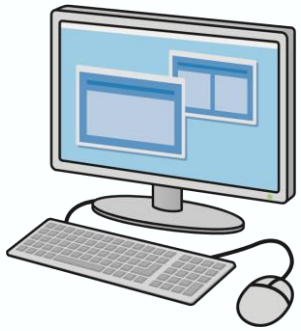




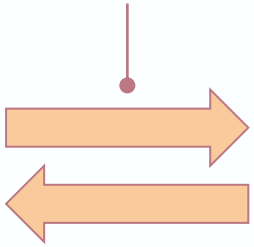
From Chabiniok R et al. Interface Focus. 2016

Designing appropriate user interfaces

Workstation



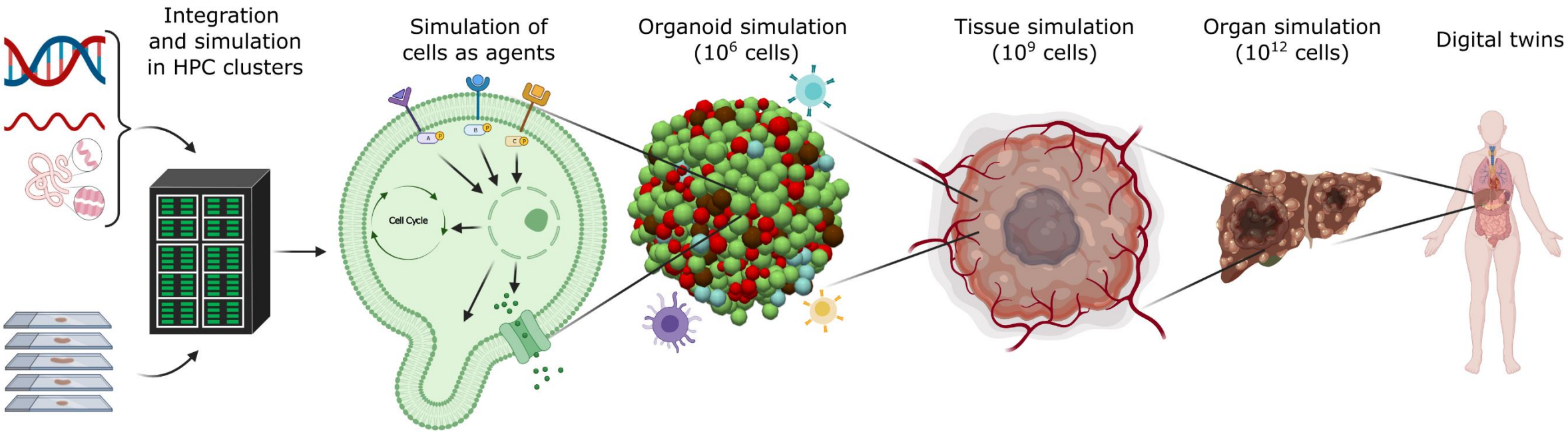
User interface (profile-based)



HPC facility

Simulations as proxies of pre-clinical trials

- USA's FDA allows (and promotes) animal-free pre-clinical tests.
  - EU is expected to follow.
- HPC is needed for real-sized tumour simulations.



Montagud, A. *et al.* (2021) Systems biology at the giga-scale: Large multiscale models of complex, heterogeneous multicellular systems. *Current Opinion in Systems Biology*, **28**, 100385.

PerMedCoE **optimises key software** for cell-level simulations and molecular pathway modelling to the **new HPC pre-exascale platforms**

PhysiCell

Agent-based modelling framework for multi-scale level simulations

COBREXA

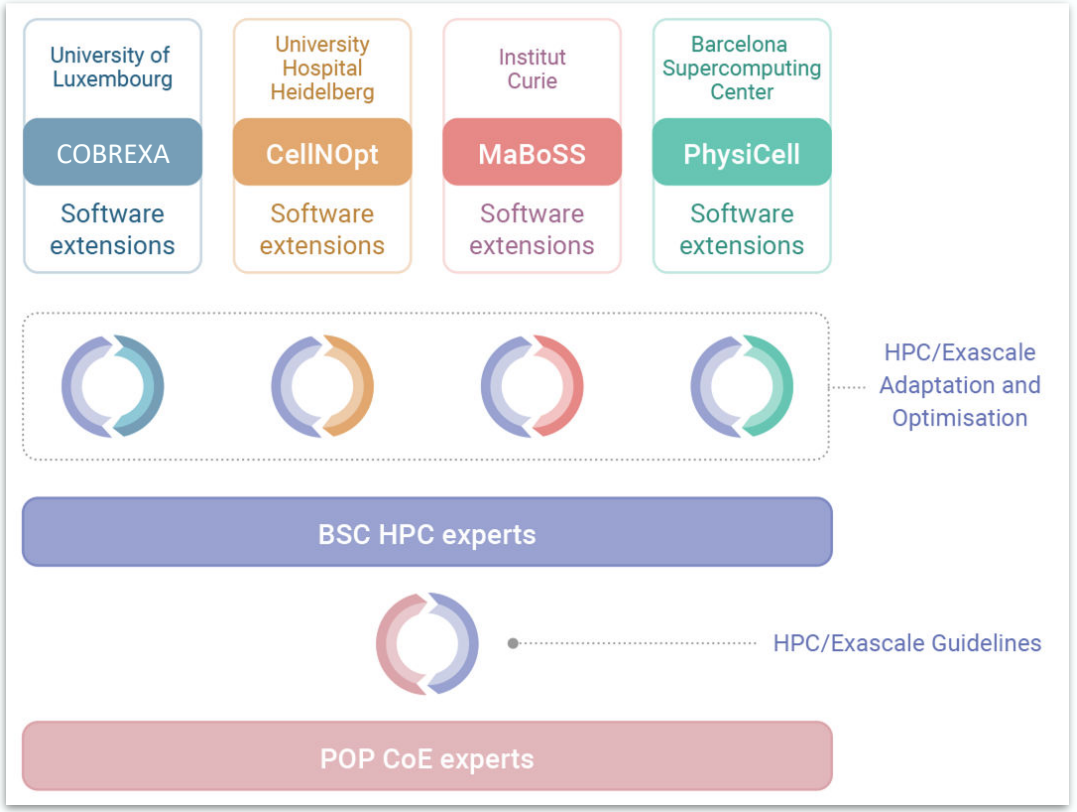
modelling of cellular metabolism at genome-scale

CellNOpt

modelling of signal transduction networks

MaBoSS

Stochastic simulations of Boolean models



PerMedCoE uses **building blocks and workflows** to solve relevant biomedical **use cases**

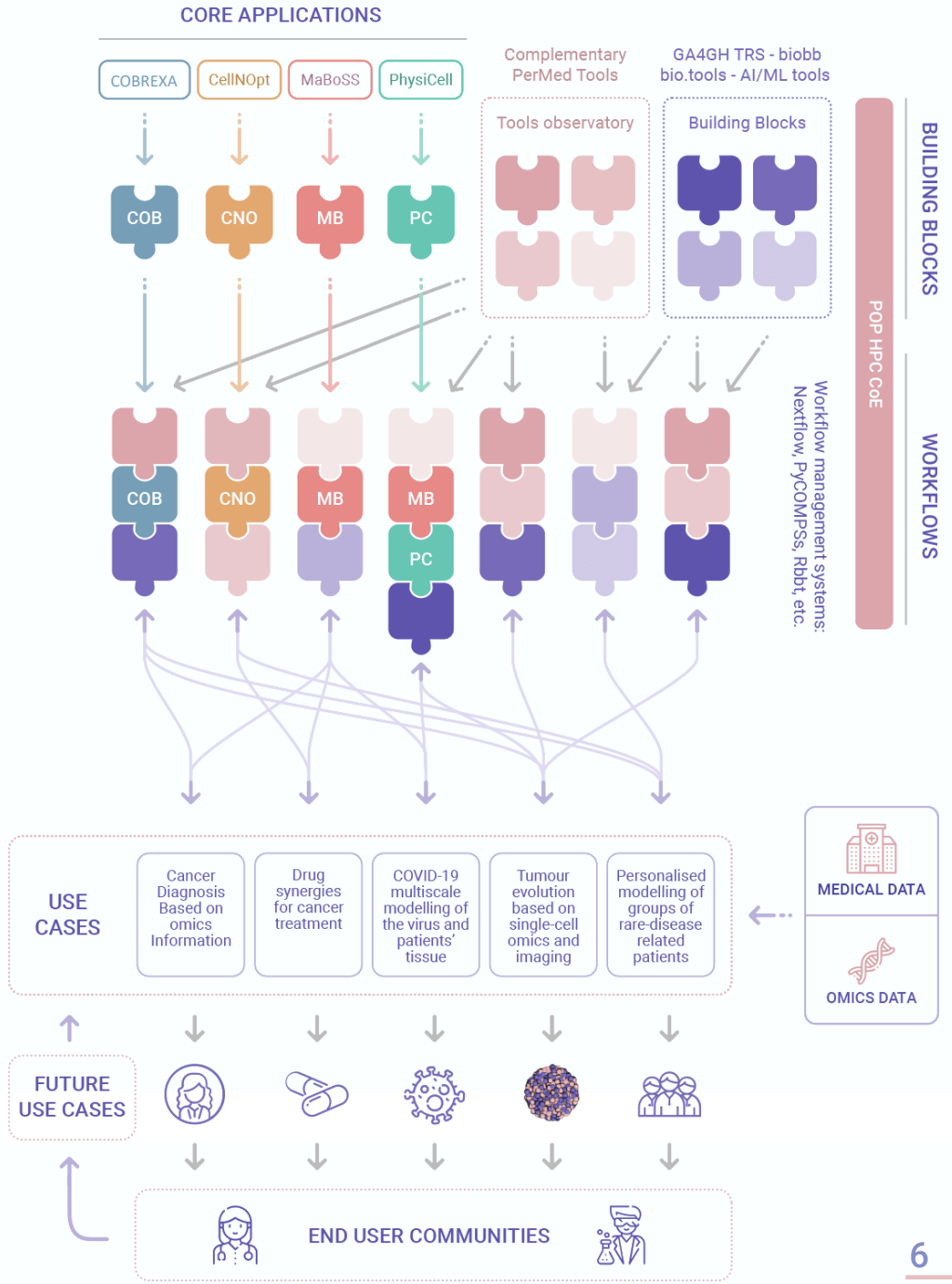
- Cancer Diagnosis Based on Omics Information
- Drug Synergies for Cancer Treatment
- Tumour Evolution Based on Single-Cell Omics and Imaging
- COVID-19 Multiscale Modelling of the Virus and Patients' Tissue

Containers

Workflow managers

Orchestrator

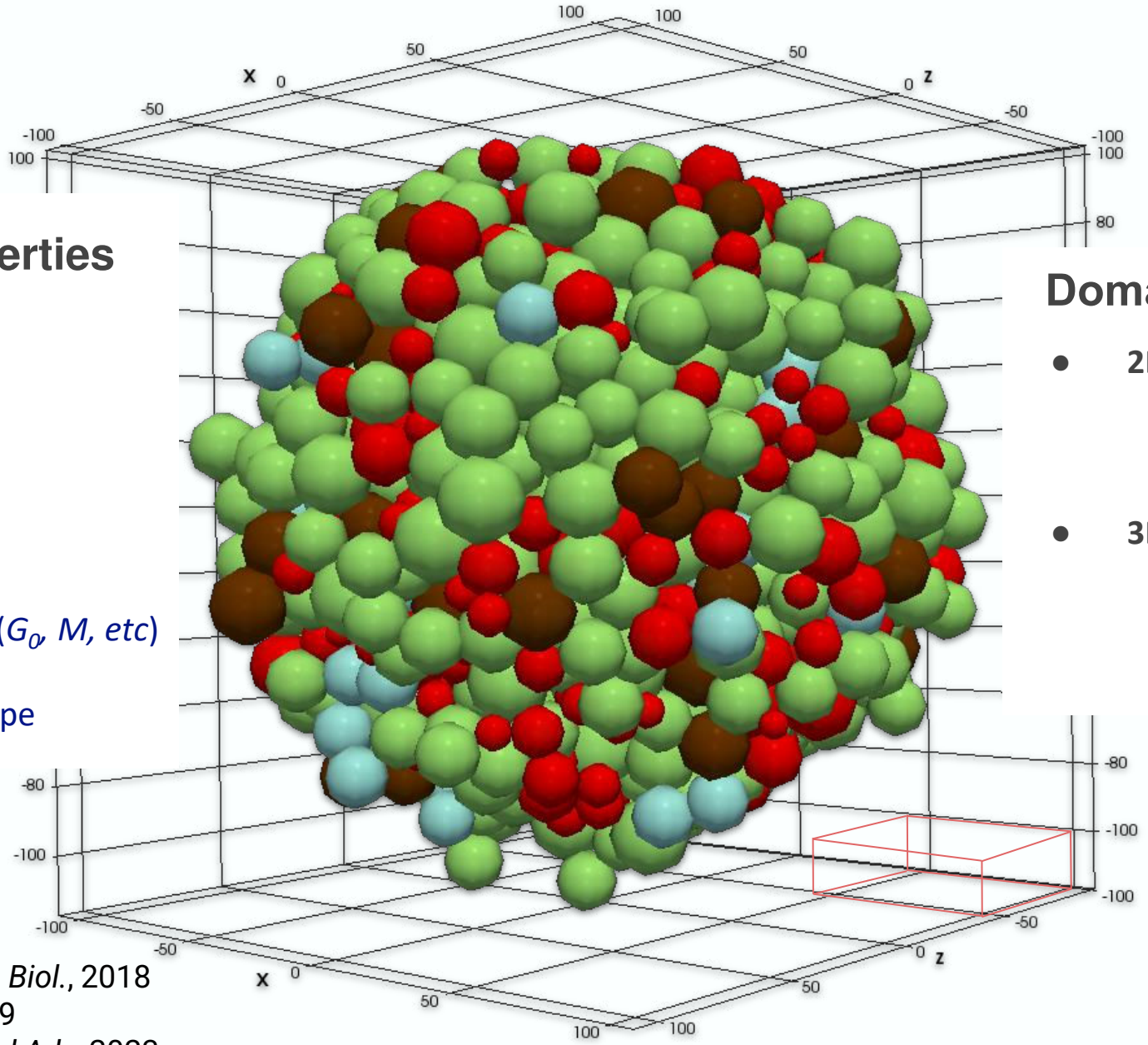
Croupier



Agent-based is a flexible, multiscale modelling framework

Cell agent properties

- **Cell Volume**
  - nucleus
  - cytoplasm
- **Position** (x, y, z)
  - Neighborhood
  - Environment
- **Cell internal state**
  - Cell cycle phase (G_0, *M, etc*)
  - Growth rate
  - Custom phenotype



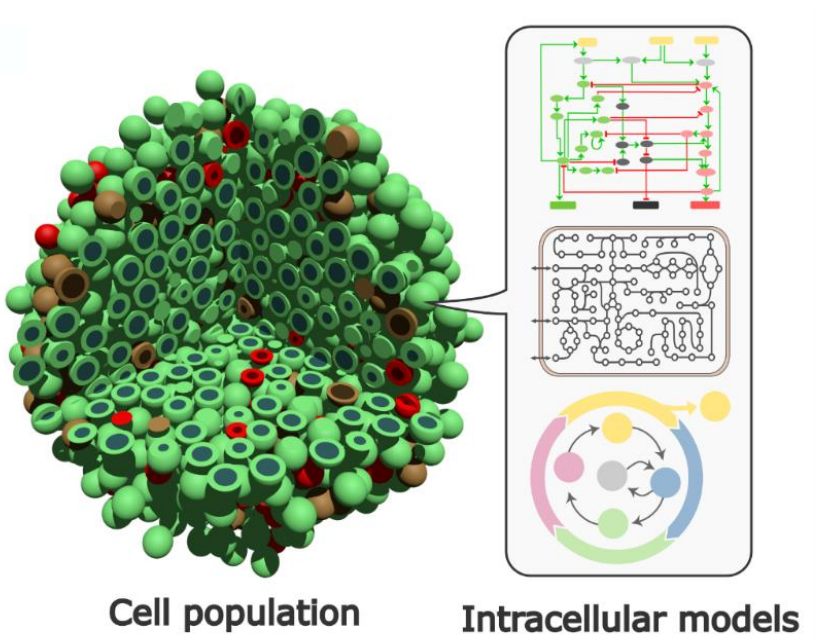
Domain = Voxels' grid

- **2D-Monolayers**
  - petri-dish
  - epitelia
  - bio-film
- **3D-Shapes**
  - spheroid
  - ductal
  - more complex shapes

PhysiCell, PhysiBoSS

Ghaffarizadeh et al, *PLOS Comp Biol.*, 2018
Letort et al, *Biofinformatics*, 2019
Ponce-de-Leon et al. *npj Syst Biol Adv*, 2023

Multiscale because we consider different time scales

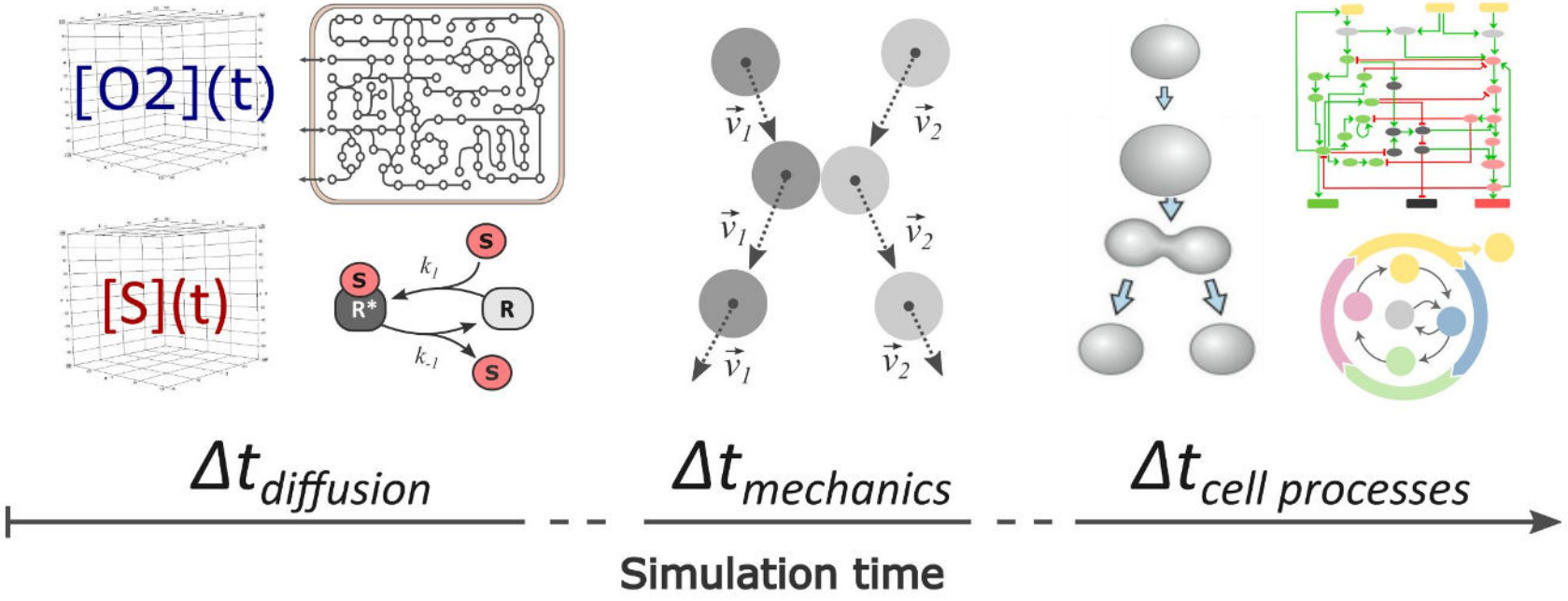


Simulation's main loop

```
while t_current < tend
    update_difussion()
    if Δt % Δtmech == 0
      update_cell_mechanics()
    if Δt % Δtcell == 0
      update_cell_processes()
       Δt = 0
    Δt += t_step
    t_current += t_step
```

Time scales

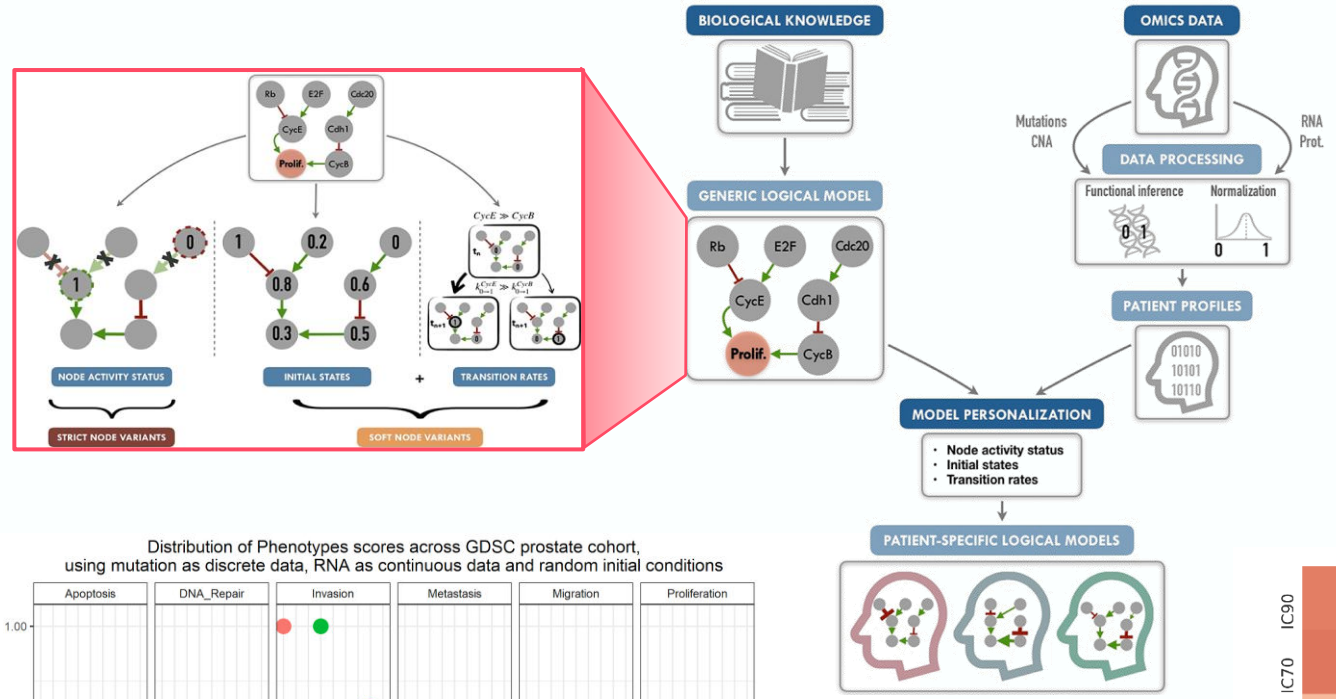
- $\Delta t_{\mathbf{diff}}$: (diffusion/transport): 0.01 min
- $\Delta t_{\mathbf{mech}}$: (cell movement): 0.1 min
- $\Delta t_{\mathbf{cell}}$: (cell processes): 6 min
- $\Delta t_{\mathbf{signalling}}$: (Boolean simulation): 10 min

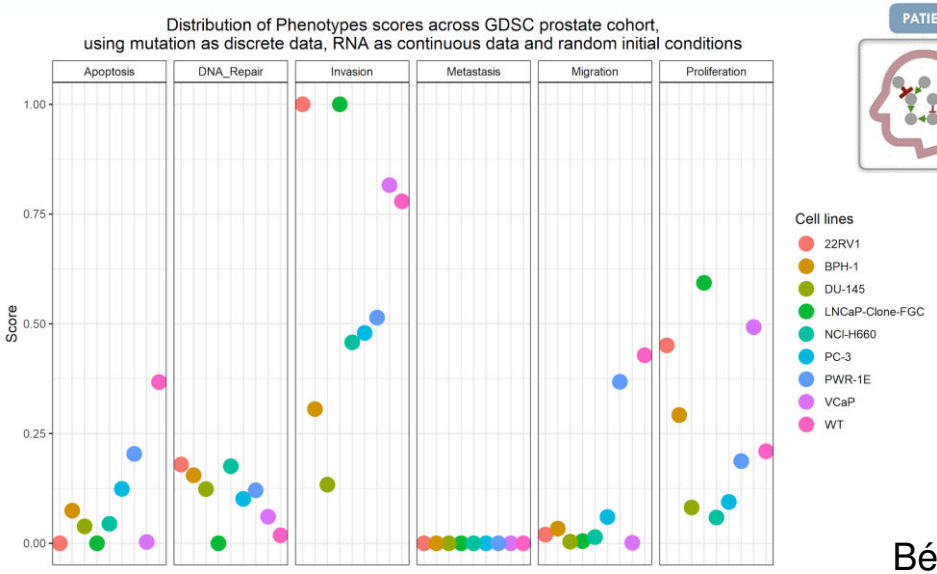


Montagud, A. *et al.* (2021) *Current Opinion in Systems Biology*, **28**, 100385

PhysiCell allows for personalised drug studies

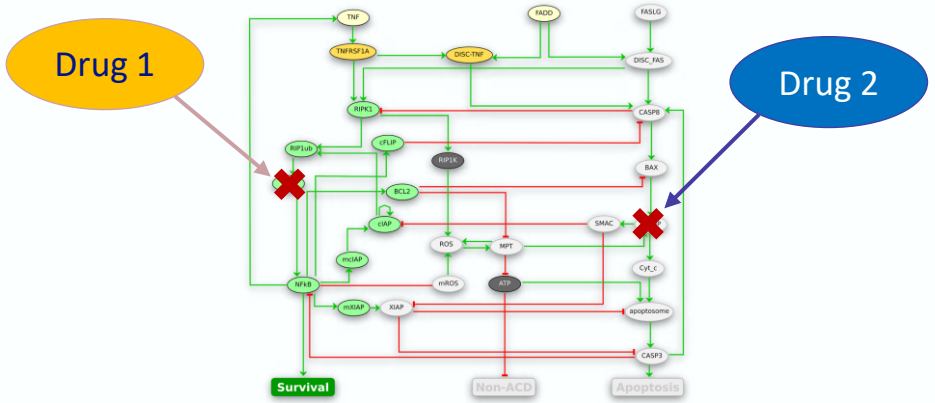
Personalisation of intracellular models

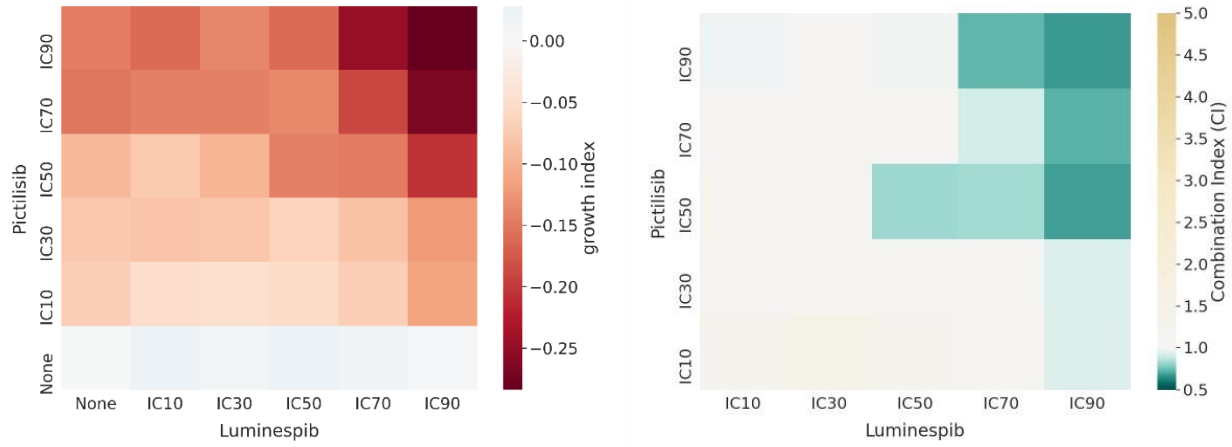




Béal et al. *Frontiers in Physiology*, 9:1965, 2019

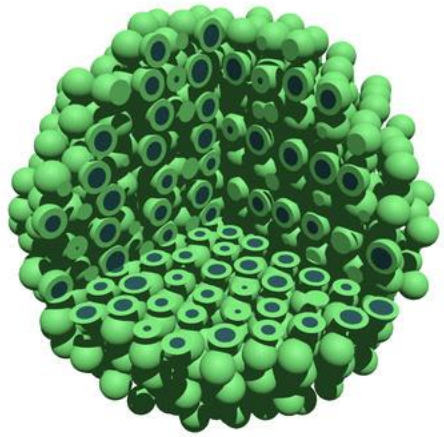
Different combinations of drugs

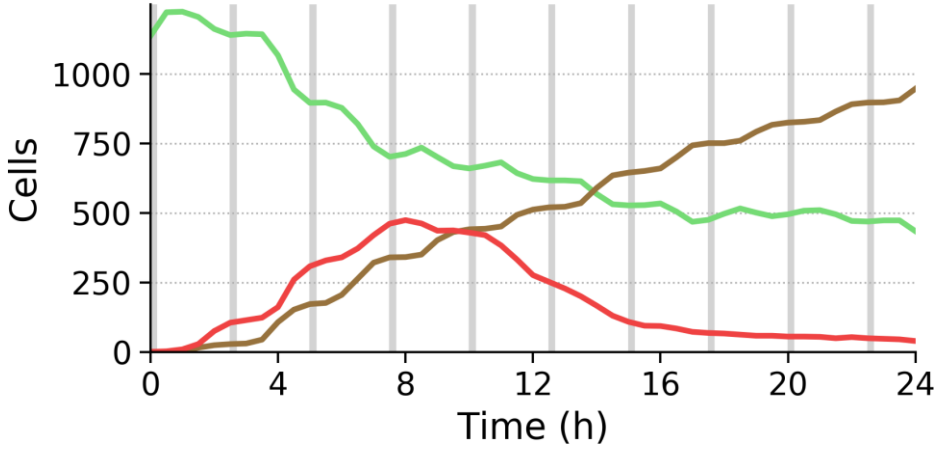


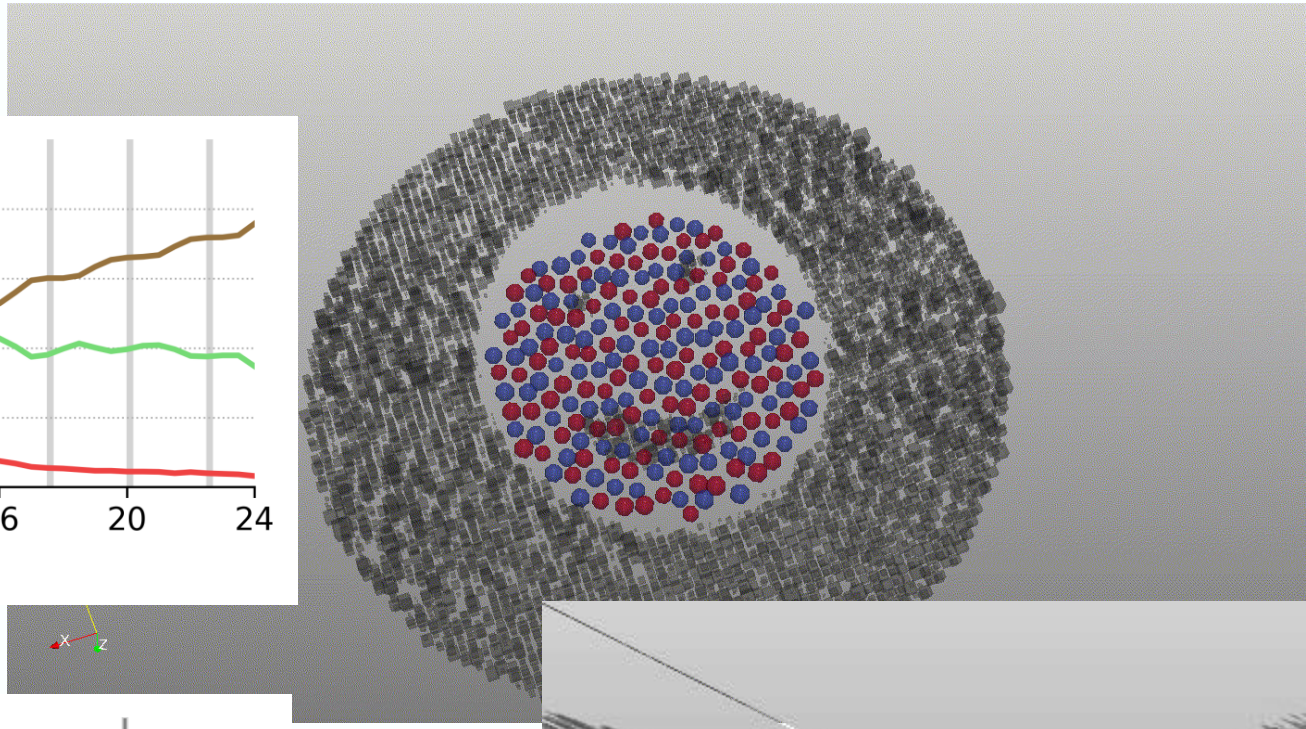


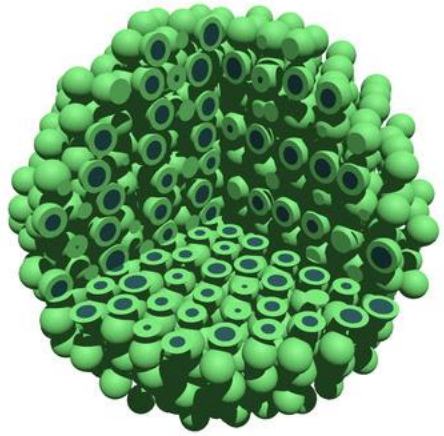
Montagud et al. *eLife* 2022;11:e72626, 2022

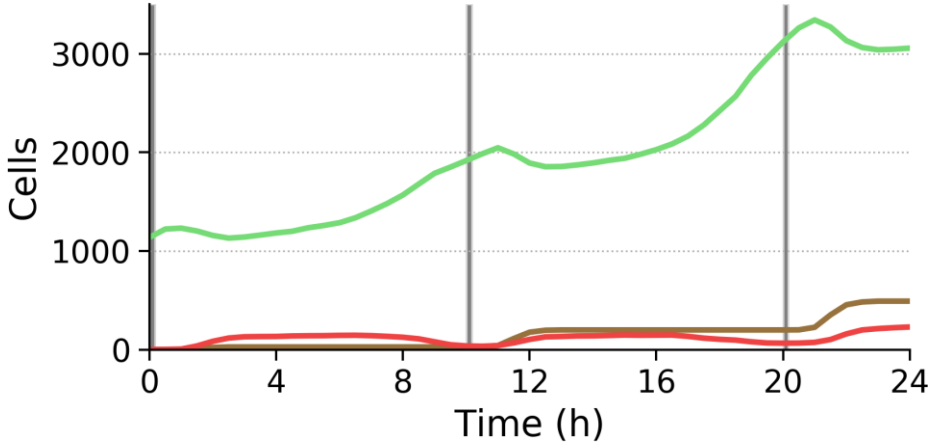
PhysiCell allows to study complex environments and drug regimes

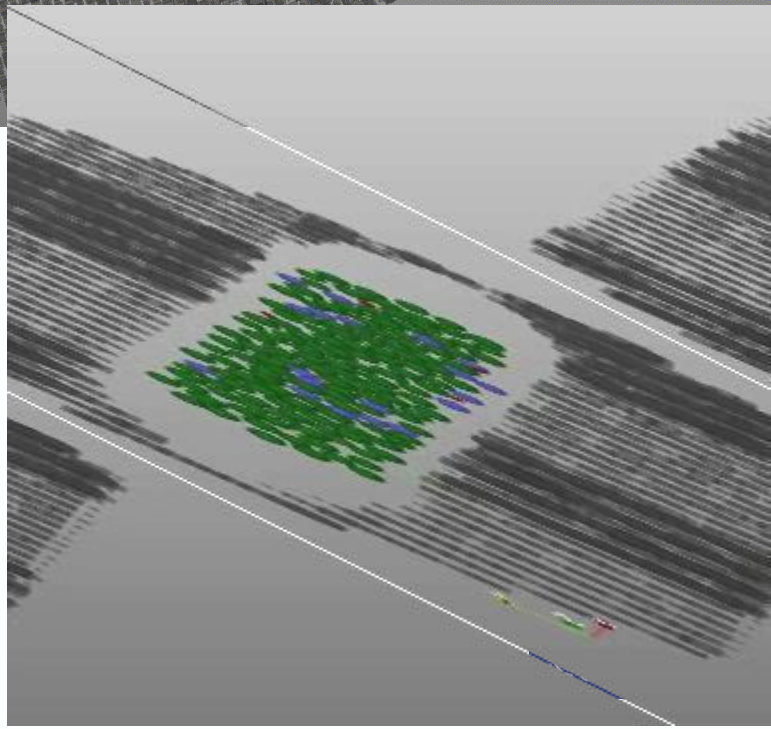










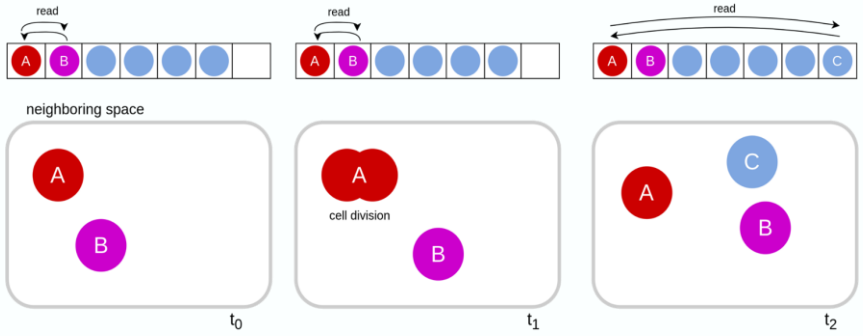


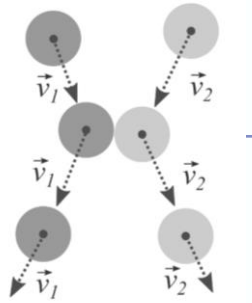
Miguel Ponce de León, BSC

We use traces to guide the performance analysis using BSC's extrae and paraver tools

Simulation's main loop

```
while t_current < tend
    update_difussion()
    if Δt % Δtmech == 0
      update_cell_mechanics()
    if Δt % Δtcell == 0
      update_cell_processes()
      Δt = 0
    Δt += t_step
    t_current += t_step
```





2 threads

4 threads

8 threads

16 threads

24 threads

48 threads

Marc Clascà,
Marta Garcia Gasulla, BSC

Runtime improvement: Changing the library improves PhysiCell's scalability

PhysiCell vanilla *malloc* library

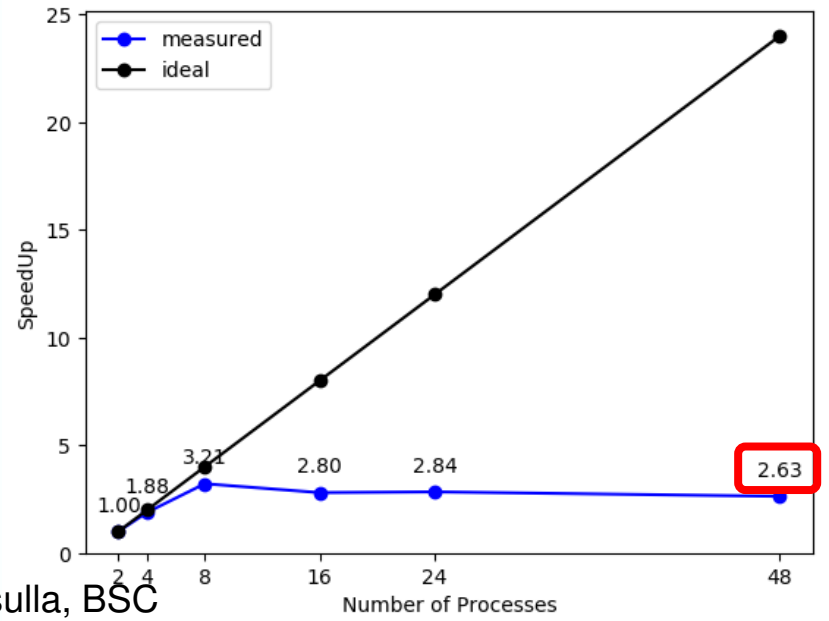
	2	4	8	16	24	48
Global efficiency	96.66	91.10	77.63	33.83	22.84	10.59
-- Parallel efficiency	96.66	93.85	86.16	93.98	91.24	91.62
-- Load balance	97.36	94.70	88.66	95.45	92.19	94.39
-- Communication efficiency	99.28	99.10	97.18	98.45	98.97	97.06
-- Serialization efficiency						
-- Transfer efficiency						
-- Computation scalability	100.00	97.07	90.10	35.99	25.03	11.55
-- IPC scalability	100.00	97.18	90.33	86.43	83.28	78.81
-- Instruction scalability	100.00	99.87	99.89	99.66	99.50	99.86
-- Frequency scalability	100.00	100.01	99.85	41.79	30.21	14.68

Percentage(%): 100, 80, 60, 40, 20, 0

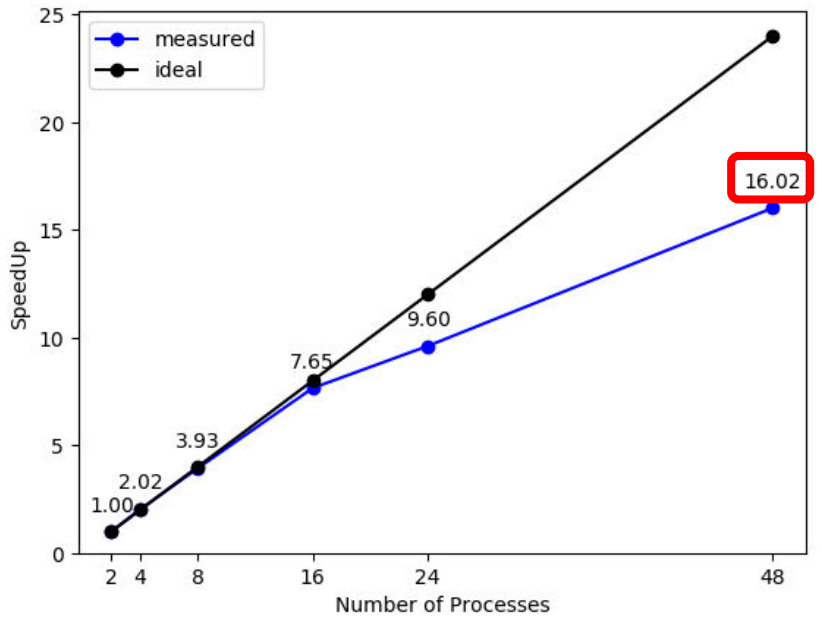
PhysiCell using *jemalloc* library

	2	4	8	16	24	48
Global efficiency	97.01	98.13	95.27	92.82	77.64	64.73
-- Parallel efficiency	97.01	96.05	90.91	88.04	76.35	62.71
-- Load balance	97.79	97.81	93.36	91.17	79.82	73.25
-- Communication efficiency	99.21	98.20	97.37	96.56	95.65	85.62
-- Serialization efficiency						
-- Transfer efficiency						
-- Computation scalability	100.00	102.17	104.80	105.44	101.69	103.22
-- IPC scalability	100.00	102.17	104.83	105.77	102.31	106.13
-- Instruction scalability	100.00	100.03	100.11	99.96	99.78	99.86
-- Frequency scalability	100.00	99.97	99.86	99.73	99.60	97.40

Percentage(%): 100, 80, 60, 40, 20, 0



Strong scaling speedup with **2 processes as baseline**

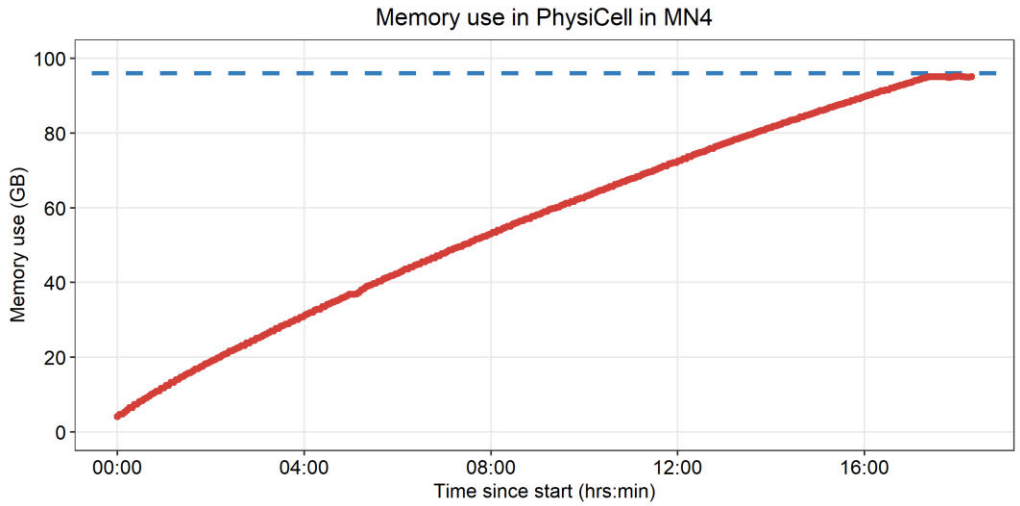


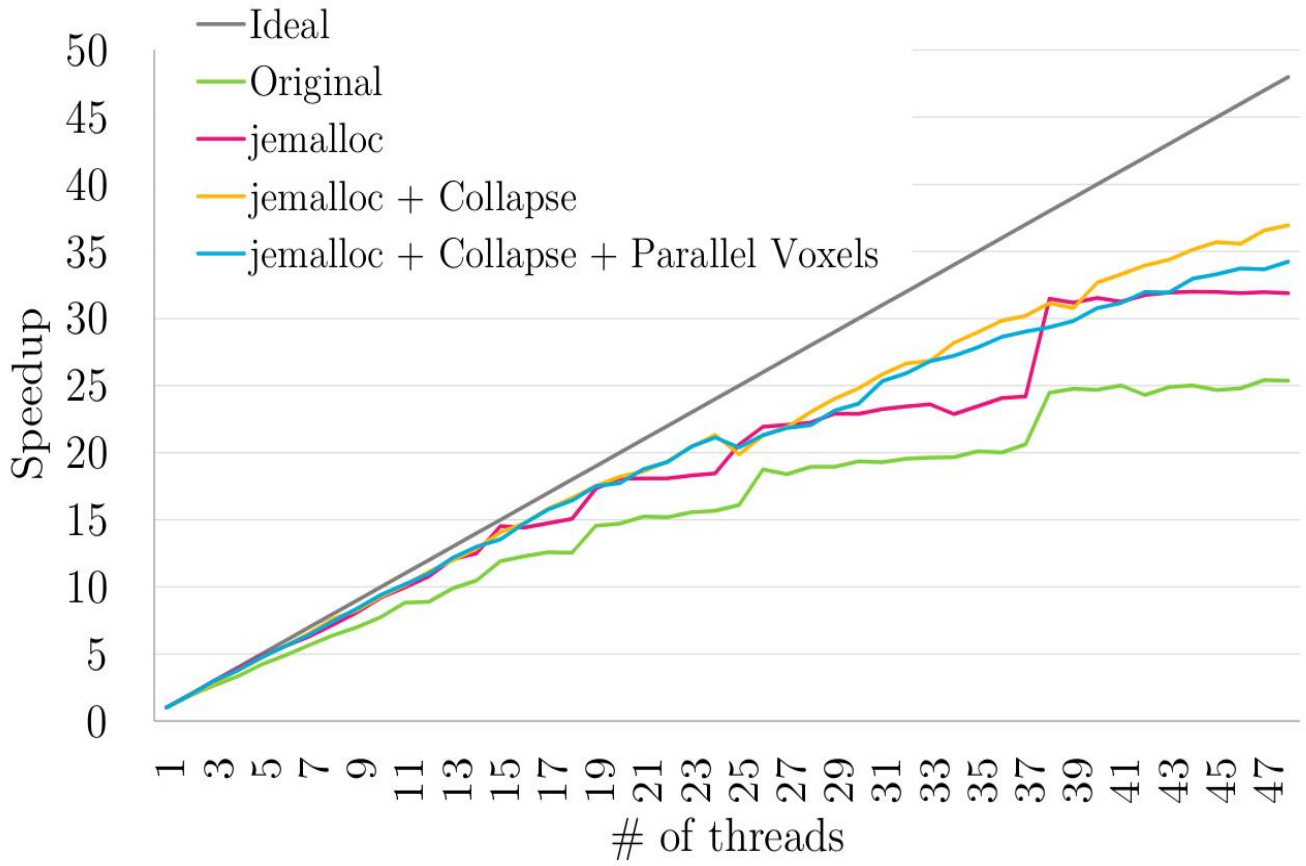
6x faster!

Marc Clascà,
Marta Garcia Gasulla, BSC

Code improvement: A family of PhysiCell versions

- **Original**: PhysiCell vanilla
- **Jemalloc**: memory allocation library
- **Collapse**: adding a collapse clause in nested loops
- **Parallel voxels**: leaving space in between the cells vector

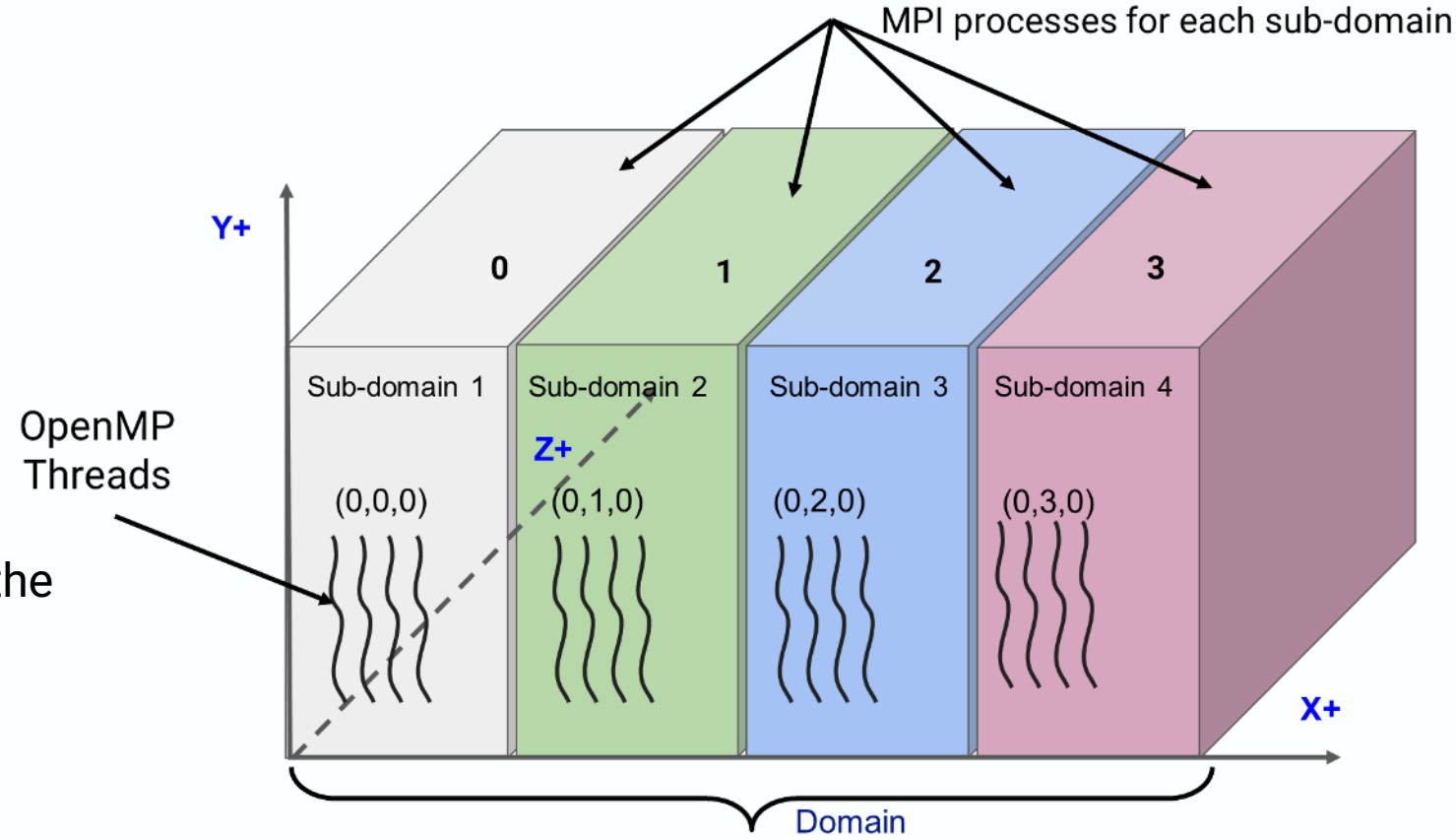




Clascà, et al. *Proceedings of the PASC Conference,* 2023

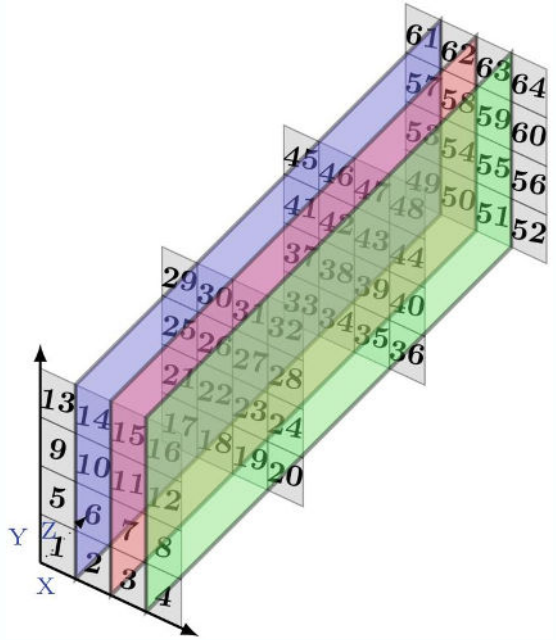
Code refactoring: In PhysiCell-X we use OpenMP + MPI

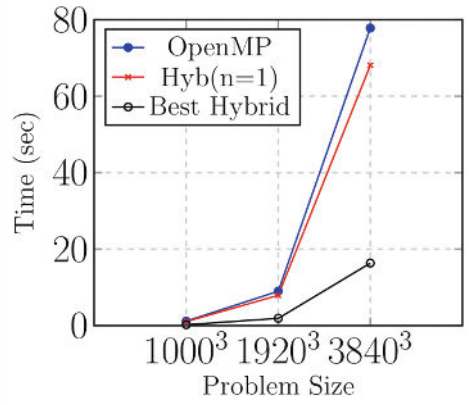
- 1-D Domain Partitioning:
  - The domain is partitioned solely in the X-direction among MPI processes.
    - Each partition is sent to a different computation node.
  - Think of this as slices of bread in a single dimension.



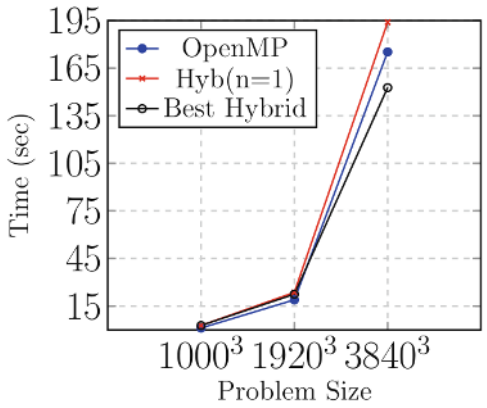
Code refactoring: PhysiCell-X enables the simulation of bigger, more complex problems

Gaurav Saxena,
Thalia Diniaco, BSC





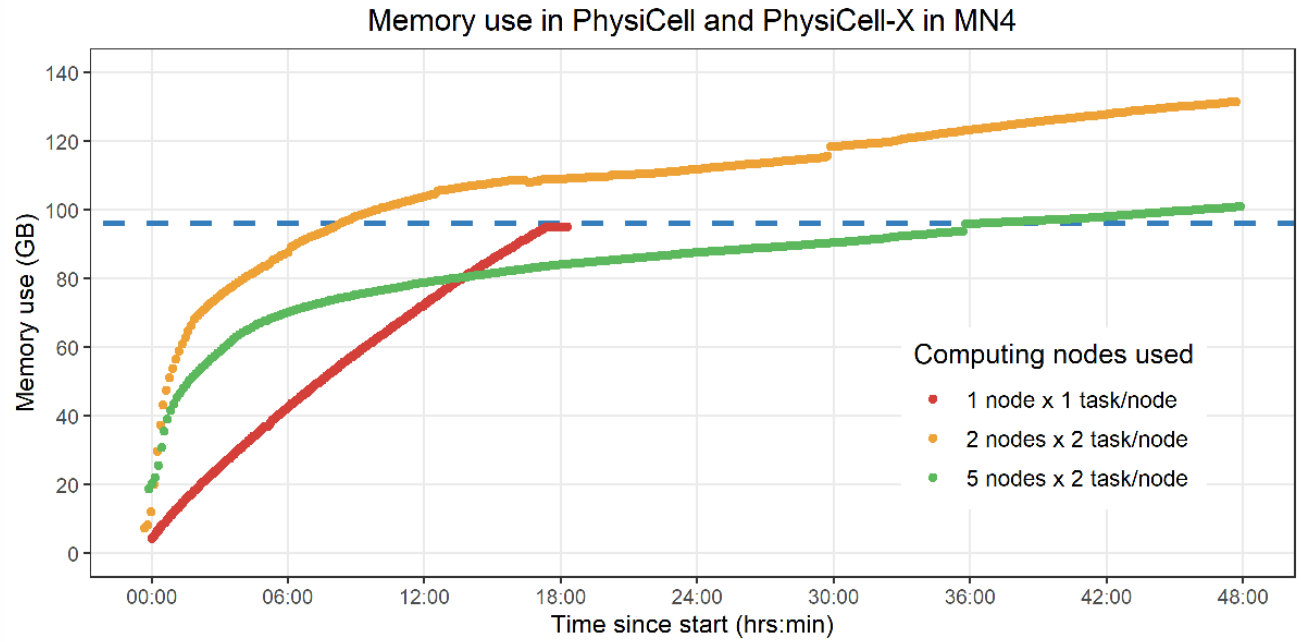
(a) μ-environment



(e) Thomas solver

- Re-factored the diffusion solver
  - Lower scale of multiscale
- Allows to simulate bigger setups
  - Needed to reach huge, complex simulations
- Still efficient vs serial in smaller setups

Domain: 7680x7680x7680 ≈ 0.5 billion voxels, 1 substrate	OpenMP	Hybrid (n=4)	Hybrid (n=8)
Build Microenvironment	x	141.98 s	67.81 s
Gaussian profile	x	0.92 s	0.45 s
File I/O	x	7.30 s	7.40 s
Agent Generation	x	0.11 s	0.0023 s
Source/Sink Diffusion Solver	x	1109.69 s	1210.41 s

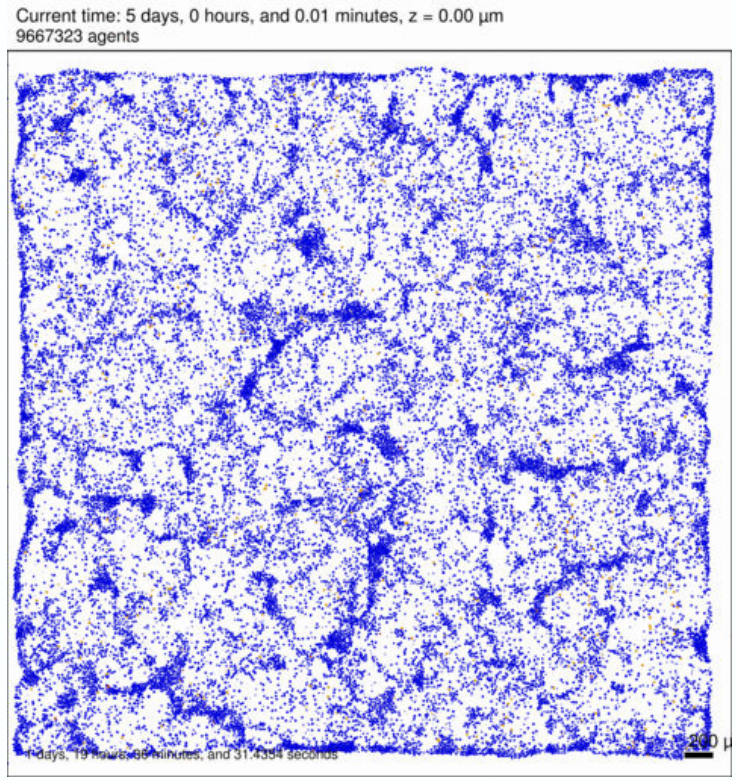


Saxena, G. *et al.* (2021) BioFVM-X: An MPI+OpenMP 3-D Simulator for Biological Systems., *Computational Methods in Systems Biology*, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 266–279.

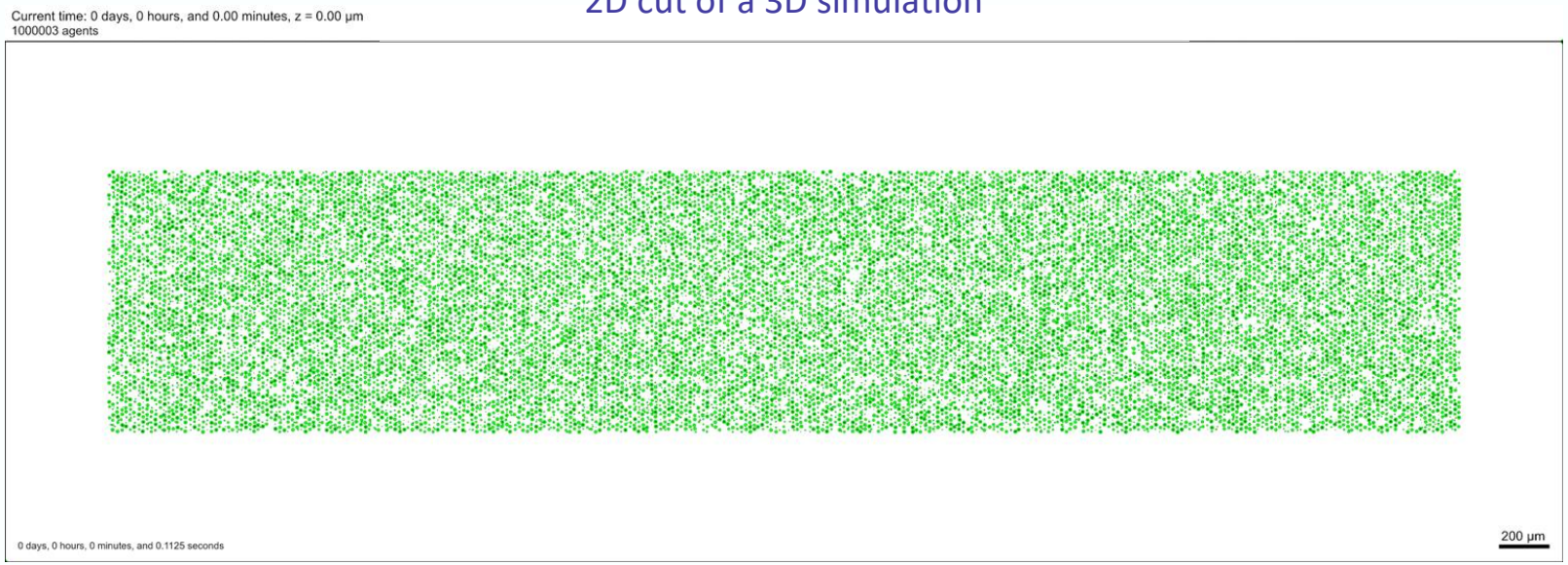
Some examples of PhysiCell-X capabilities

Predator-Prey PhysiCell-X example

- Predators (0.1 M) hunt Preys (1 M)
- Domain: 4200^3 voxels
- 84 mm-side cube
- 10 nodes used (480 cores)
- Serial 177 h vs Parallel 43.5 h
  - Speed-up: 4.06x
  - Efficiency: 40.3%



2D cut of a 3D simulation



Cancer growth PhysiCell-X example

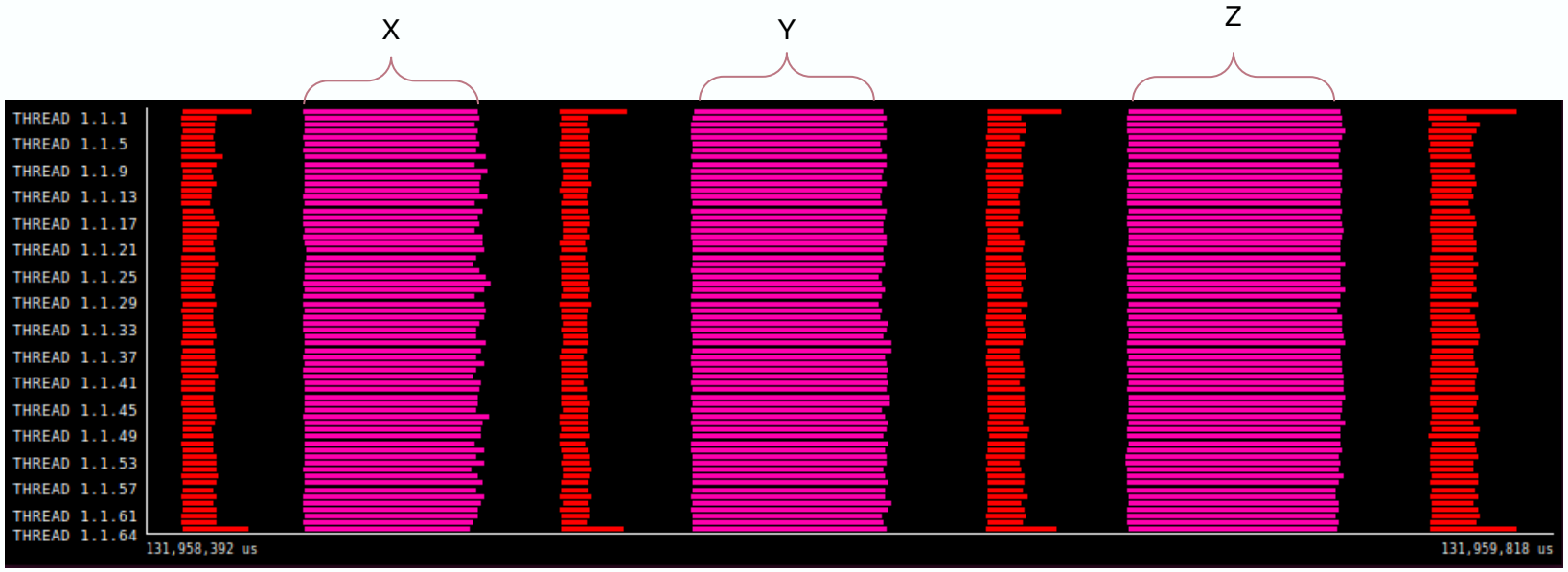
- 1 M initial cells, no TNF
- Domain: 2500 x 500 x 500 voxels
  - 50 x 10 x 10 mm-side box
- Drug diffusion from the 3D walls
- 10 nodes used (480 cores)

Tested in MareNostrum4, Kunpeng, LUMI-C, HAWK, Marconi-100

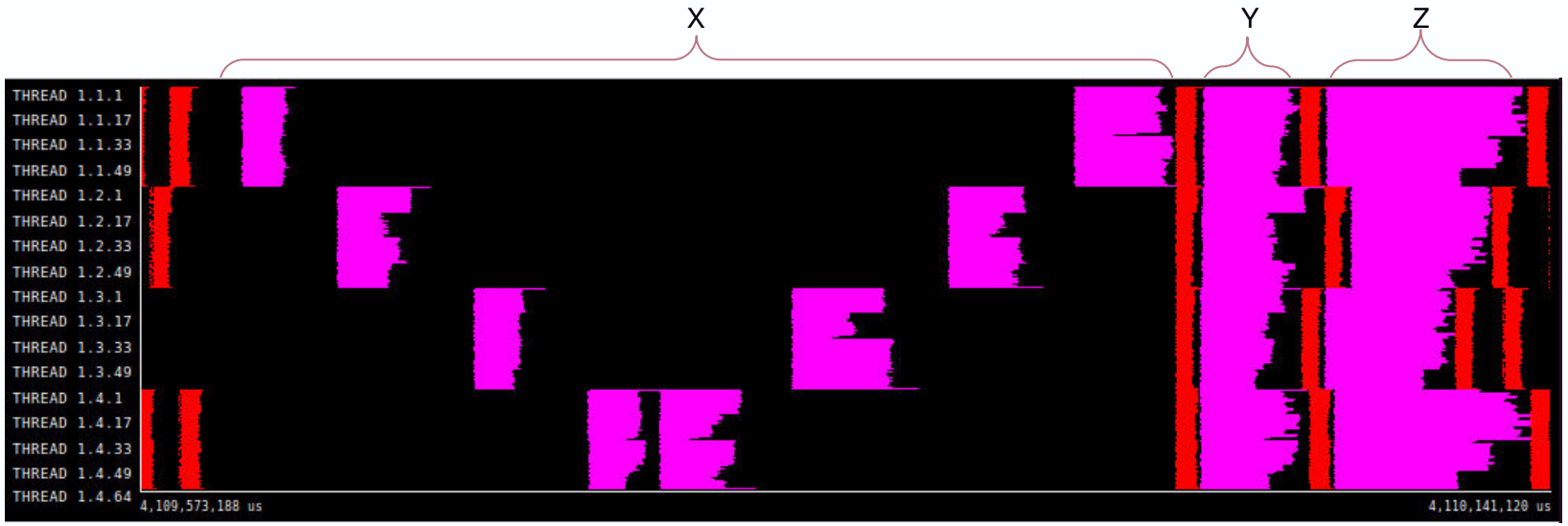
Jose Estragués,
Thalia Diniaco, BSC

PhysiCell-X serialises the X-dim diffusion iteration

PhysiCell (OpenMP)

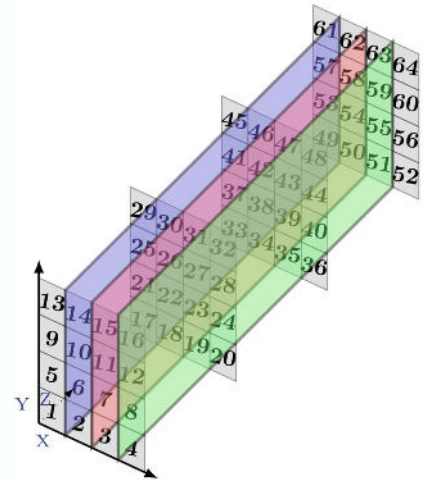


PhysiCell-X (MPI + OpenMP)



Tridiagonal matrix method

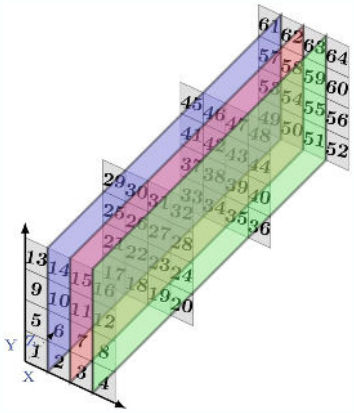
$$\begin{bmatrix} b_1 & c_1 & & & 0 \\ a_2 & b_2 & c_2 & & \\ & a_3 & b_3 & \ddots & \\ & & \ddots & \ddots & c_{n-1} \\ 0 & & & a_n & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_n \end{bmatrix}$$

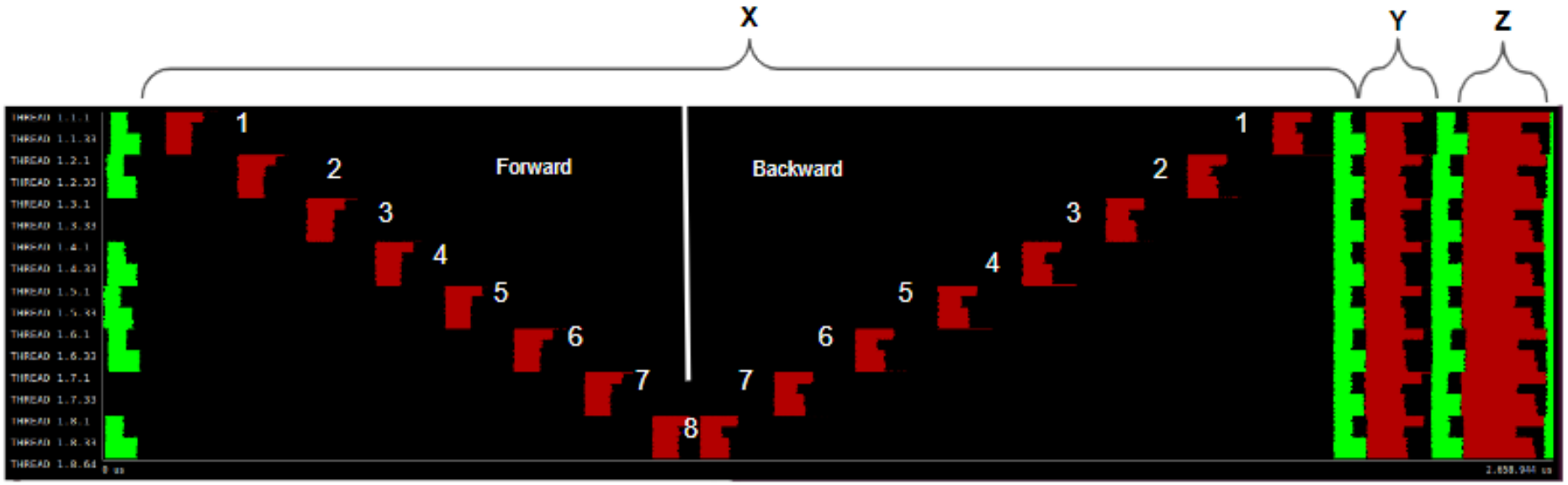


Code refactoring: Overlapping communication and computation improves PhysiCell-X performance

PhysiCell-X vanilla

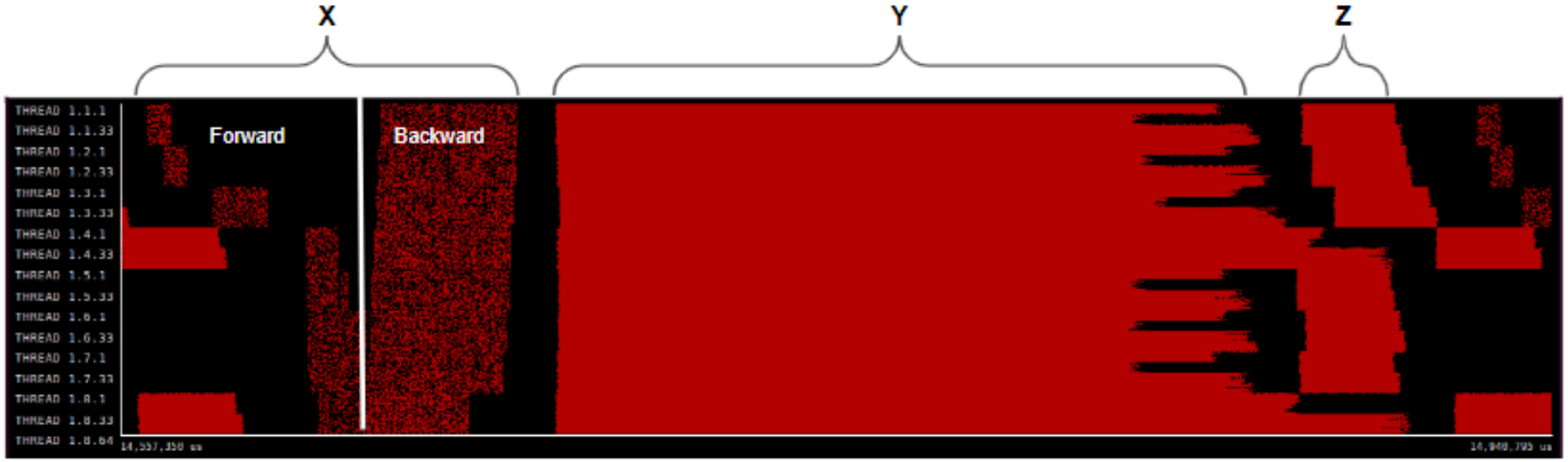
2.6 s



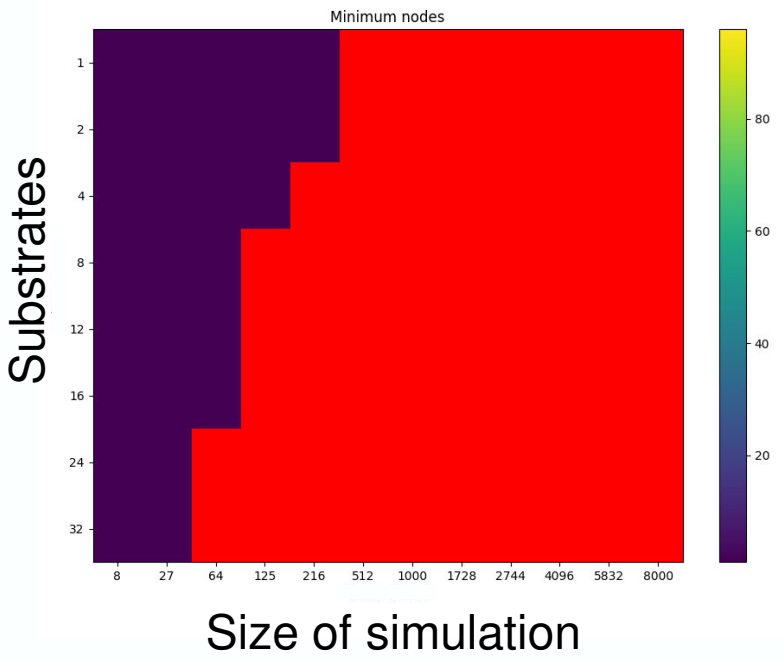


PhysiCell-X overlap

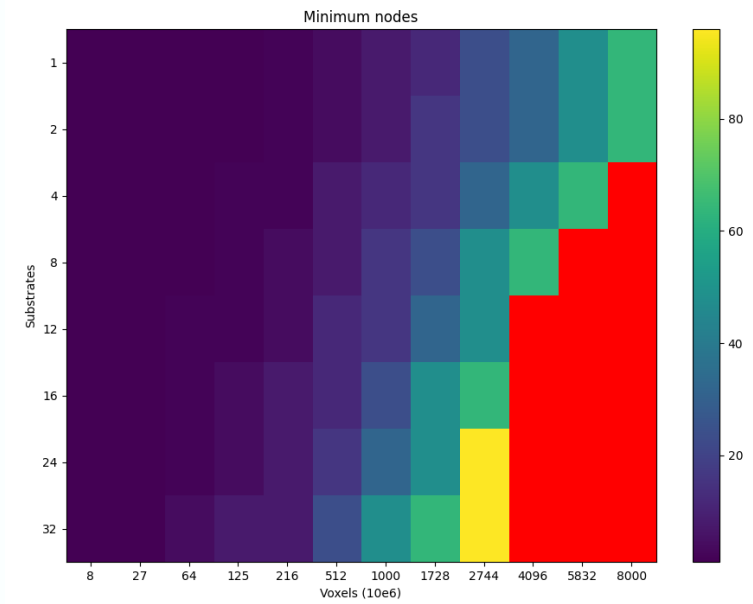
384 ms
64 blocks



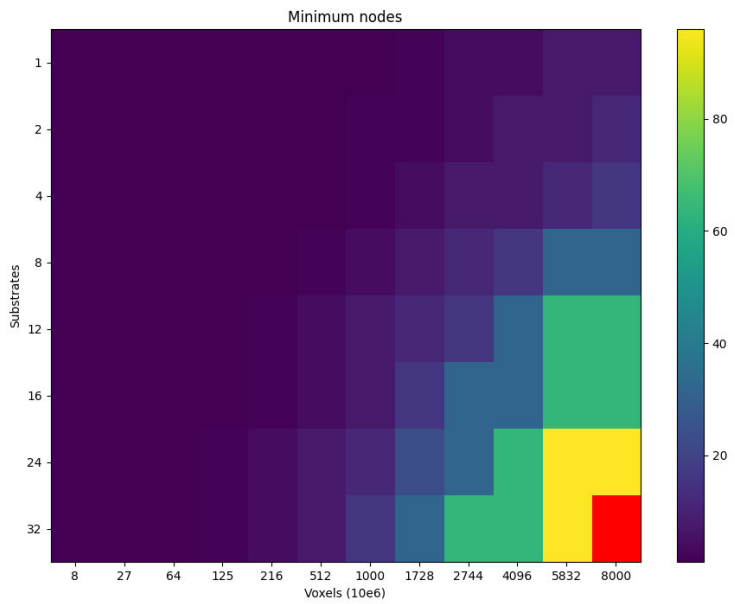
Code refactoring: new data structures allows PhysiCell-X for bigger, more complex diffusions



PhysiCell



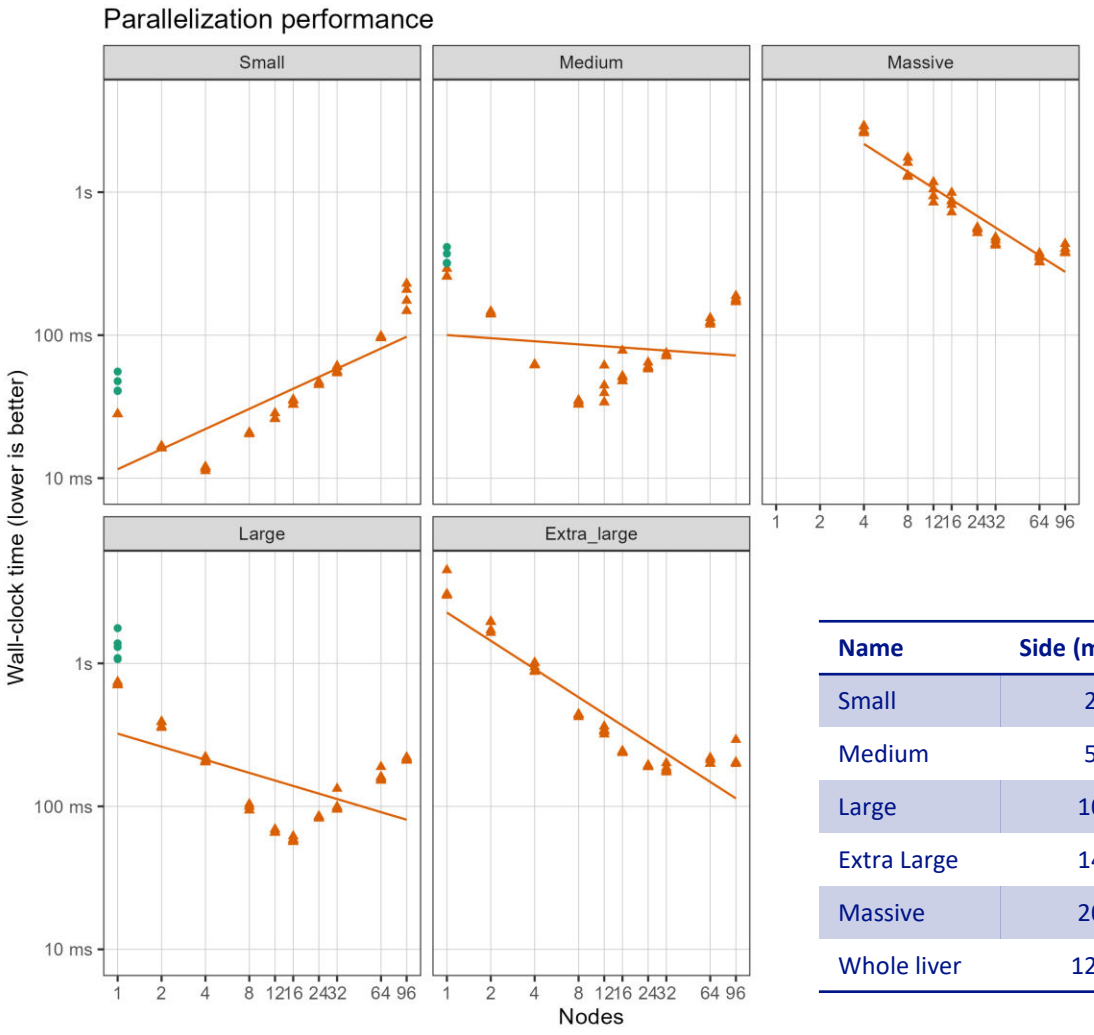
PhysiCell-X
original version



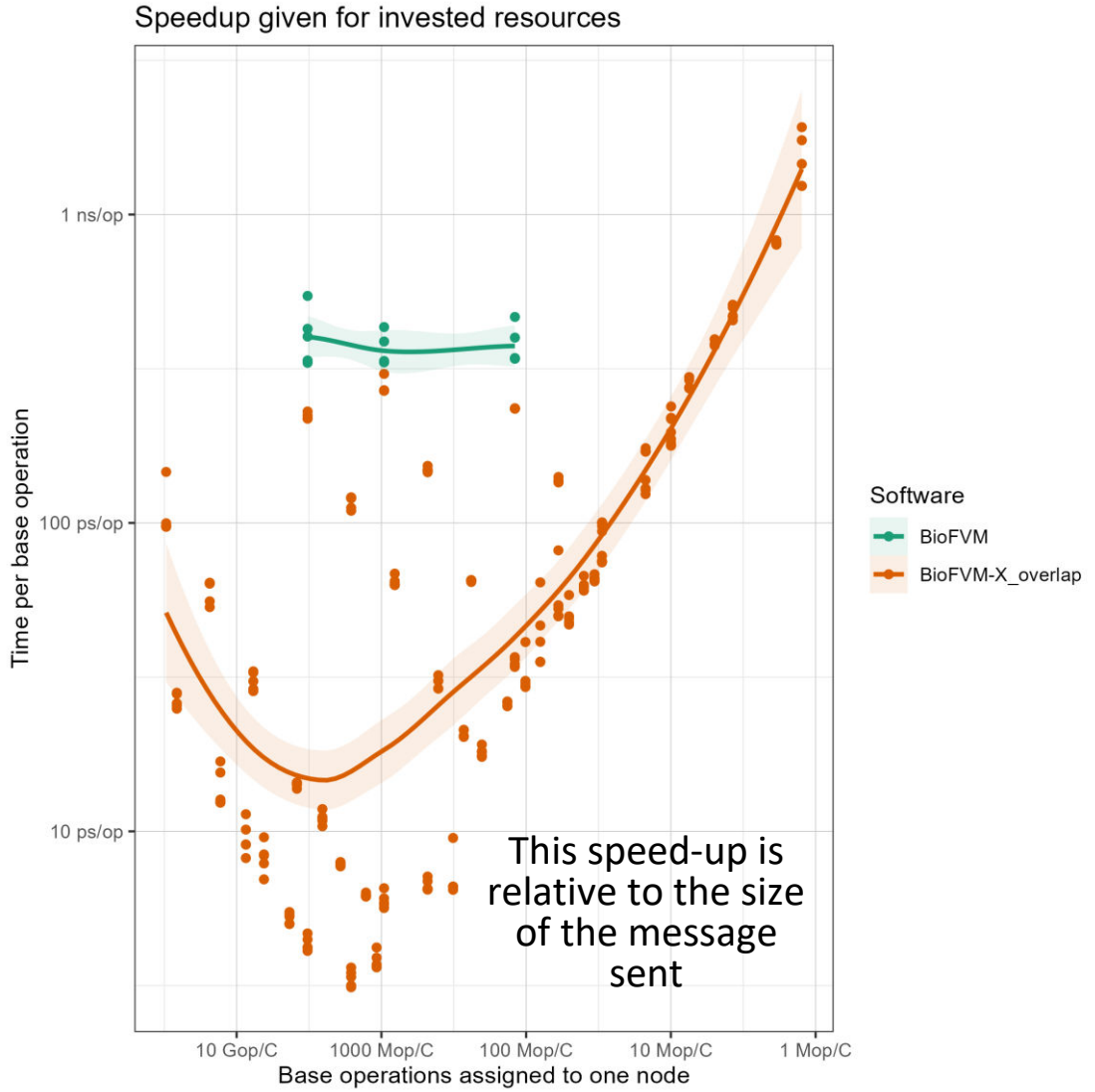
PhysiCell-X
overlap

: computation not possible in MN4

Code benchmark: Scalability tests of the diffusion of PhysiCell-X



Name	Side (mm)	Substrates
Small	2	1
Medium	5	2
Large	10	4
Extra Large	14	8
Massive	20	8
Whole liver	120	>100

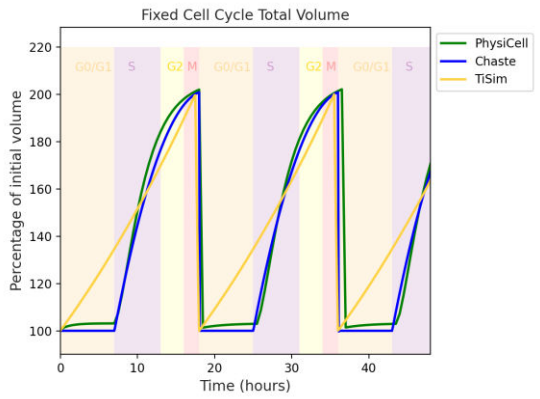


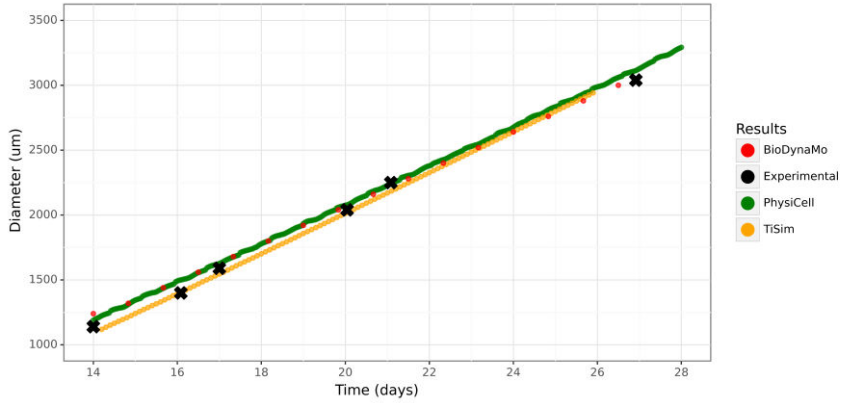
We have tested PhysiCell-X in up to 100 computation nodes (4800 cores)

Perspectives towards a highly-efficient exascale-ready human digital twin

- Extensive tests
  - Scalability
  - Energy-aware
  - Pre-exascale clusters
- Combine with accelerators
  - GPUs
    - Alpha version
  - Vectorisation
    - Refactoring code
  - Stencil computing
    - Change algorithms
- Co-design strategies
- Learn from success cases

- Integration with organ-level simulators
  - Alya from BSC
  - Simulating whole organs
- Benchmark with other multiscale codes
  - Set up unit tests and test cases
  - Comparison of tools
  - Performance tests





Acknowledgments

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Miguel Ponce de León
Gaurav Saxena
Jose Carbonell-Caballero
Thalia Diniaco
Othmane Hayoun-Mya
Alejandro Madrid
Jose Estragués
Alfonso Valencia

institutCurie

Laurence Calzone
Vincent Noël
Marco Ruscone

Computational Biology Group at Barcelona Supercomputing Center (BSC)

Contact

arnau.montagud@bsc.es
@ArnauMontagud

Collaborators

Paul Macklin (Indiana University)
Emmanuel Barillot (Institut Curie)
Rosa Maria Badia (BSC)
Julio Sáez-Rodríguez (Heidelberg U)
Tommi Nyrönen (CSC)

Projects

PerMedCoE, CrexData, Oncologics, EDITH

Arnau Montagud
arnau.montagud@bsc.es
@ArnauMontagud

HPC/Exascale Centre of Excellence in Personalised Medicine

www.permedcoe.eu

Follow us in social media:

www.linkedin.com/company/permedcoe
@permedcoe

The PerMedCoE project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement Nº951773