HPC in Poland: 2014 update

Marek Niezgódka

ICM, University of Warsaw

marekn@icm.edu.pl

Seattle, HPC User Forum, September 16, 2014

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

HPC in Poland: 2014 update

- Funding 2008-2014:
  - HPC for research $200M+
  - Other e-infrastructures for academia $400M+
  - e-infrastructures for government and administration $1.5-2B
- 2015-2022 programs: continuation focused on operations and R&D

Priorities (academia):

- **Capacity infrastructure:**
  - Cracow (Cyfronet)
  - Poznan (PSNC)
  - Gdansk (TASK)
  - Wroclaw (WNSC)
  - Swierk (NNRC, energy)

- **Capability facility:**
  - Warsaw (ICM)

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

Some background

- 1993: National Research ICT Infrastructure program launched
- 1994: 5 HPC centers operational
- 1996: National optical networking infrastructure program launched
- 1997: Virtual Library of Science started
- 2007: Research Infrastructures: Operational Programs (cohesion, budget)
- 2013: Competence centers founded

- **Current aggregated capacity** (selection)**:**
  - Cyfronet: .5 PF (2+ PF coming)
  - ICM: .5 PF (2+ PF coming)
  - PSNC: 150 TF (subject expansion)
  - In addition, TASK (150 TF), WNSC (100 TF), NCNR (180 TF)

Coordinated concept: competence centers

- High Performance Networking: *PSNC*
- Capacity computing services: *Cyfronet*
- Capability computing and data-based services: ***ICM***

- Foundation: POWIEW and PL-GRID R&D programs (*completed*):
  - Complementary HPC infrastructures
  - Joint job management concept
  - National services

- e-Infrastructure roadmap 2014-2024

OSIRIS: ICM's roadmapping 2022

- Distributed, scalable e-infrastructure for data-driven open science
- National partnership coordinated by ICM
- Core : ICM's OCEAN infrastructure for data research
- Multiple e-infrastructure layers:
  - Storage and retrieval services
  - HPD Analytics feasibility
  - Big data based HPC
- Integrating role of the competence centers: operations, development, consulting, research and education

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

OSIRIS concept outline

- Academia:
  - Data science programs
  - Open science models
  - Special focus: digital humanities and social sciences
  - Partnership with public sector and industry
- Non-academic partners:
  - government agencies, public administration
  - business and industry
  - NGOs
- Industry benefits:
  - access to open data repositories, including open government data deposited
  - business intelligence solutions,
  - access to comprehensive research information systems (OpenAIRE, Infona, PBN)
  - access to open research data, tools and solutions

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

OCEAN program (2013+)

- *Centre for Data, their Analysis and Computational Modelling*
- National data infrastructures:
  - OSIRIS roadmap
  - Research: INFONA integrated platform
  - Government: CRIP public data platform
- HighPerformance data analytics:
  - large-scale data-sets
  - time-critical dependable services
  - development of specialized solutions and infrastructure for big-data analytics
- Capability computing:
  - time-critical applications
  - data-driven extreme computing
  - specialized solutions for big-data driven models
- Shared infrastructure resources
- Competence center: R&D, consulting

OCEAN: concept foundations

- Secure big data operations facilitated: multi-site set-up
- Motivation:
  - rapid growth of experimental and observational data, computer simulation results
  - broad opening of governmental public data
  - challenge of inter-sector collaboration in public sector
  - concertation of High-Performance infrastructures within public sector
- Technology-gap induced challenges:
  - need for novel concepts in algorithmics and computational models construction

OCEAN: initial physical infrastructure

- Distributed location:
  - Current site – future backup facility
  - New site:
    - Under construction
    - Operations: Q3, 2015
    - Total floor: 5000+ sq.m
    - Power supply: 4MW (10MW provided)
  - Set-up:
    - Data storage services: initial 30+ PB fast storage
    - HPData Analytics: 850+ GTeps
    - Capability computing: 2+ PF
    - Technology development site

OCEAN: specific features

- Full processing chain encompassed in knowledge retrieval from big data:
  - Data acquisition, from experiment and observation results through HPC simulations
  - Preprocessing and structuring of big data: selection, reduction, visual data analysis and computing
  - Post-processing: data-driven HPC, feedback loop to real systems
- New generation of algorithms, computational models and highly scalable software implementations

OCEAN: selected priorities

Methodology:

- Mathematical fundations of data science and computational
- Statistical methods
- Visual data analysis and scientific visualization
- Computational software: new drections , optimization
- Complex-structure systems: characterization and dynamics

Application areas:

- Numerical weather prediction and its applications to critical decision planning and supporting systems
- Complex systems in natural environment
- Medicine, healthcare and life sciences
- Materials sciences and engineering
- e-governance and industry
- Complex logistics and transportation networks (aviation)
- Social processes associated with introduction of new technologies and digital societal transformation

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

ICM: some milestones

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

ICM

Founded in 1993, as a center for:

- HPC research infrastructure: operations and development programs on national scale
- national information infrastructure
- research in computational and information sciences

Initiated programs and activities, e.g.:

- national academic applications software system (1996)
- national research database accessibility (1997)
- national w3caching program (1996)
- open-access multi-scale numerical weather prediction system (1997)
- national virtual library of science (1998)
- a group of leading-edge experimental labs (2005), extended into a university centre for new technology (2012)

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

Large-scale research infrastructures and their development: some of ICM contributions

- National networked HPC infrastructure: applications software, capability computing
- National virtual library of science: content, software system
- Unified national academic information infrastructure: integrated system
- Polish Research Bibliography and Polish Citation Index
- EU: DRIVER and OpenAIRE open repository infrastructures
- EU: EuDML (European Digital Mathematics Library)
- EU: UNICORE – grid infrastructure (security functionalities)

ICM: selected HPC-driven research areas

- **Human physiology:**
  - Towards personalised medicine:
    - cardiac system, blood circulation
    - cancerogeneous developments
  - Multiscale developments:
    - Molecular level
    - Cell and tissue scale
    - Functional effects at macro-scale
- **Materials science and engineering**:
  - Design of new functional materials and biomaterials down to nanoscale:
    - characterization
    - process modelling

ICM: HPC-driven research areas, cont.

Methodology:

- New concepts of computational algorithms and data structures for future architectures (towards exascale)
- Nonlinear process dynamics in systems of high complexity
  - spatial structure formation in systems over complex, possibly variable geometry/topology
  - applications to population dynamics
  - irreversible and nonlocal phenomena
  - decision making and control over multiple time-scales
  - stochastic dynamical networks

ICM: main computational infrastructure

- As of August, 2014:
  - x86-based compu-clusters: ~20K cores
  - Blue Gene Q: 16K cores
  - Power 775: 2.4K cores
  - Blue Gene P: 4K cores (R&D & education)
- 2014 (pending):
  - Capacity compu-server: .3 PF
- 2015 (pending):
  - HighPerformance data analytics system
  - Capability computing system
  - HighPerformance data storage

ICM: an evolution of the concept

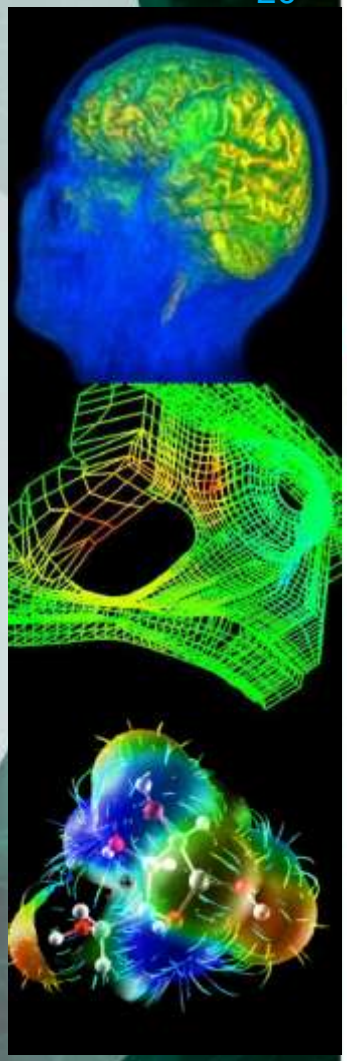
- Center for computational sciences (1993+):
  - Mathematical modelling
  - Foundations: physics, chemistry, biology, …
  - Algorithms
- Promotion and implementation of open publishing models (2004+) – Poland and Europe:
  - software
  - Publishing and scholarly communication
  - Research data
  - e-infrastructures
- Data Science and data-driven sciences centre (2013+):
  - R&D
  - partnerships

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

ICM: research, solutions and services

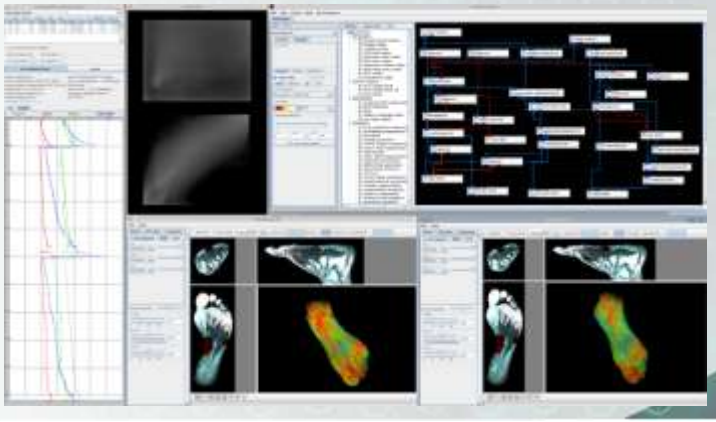
UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

Visual Computing and Data Analysis



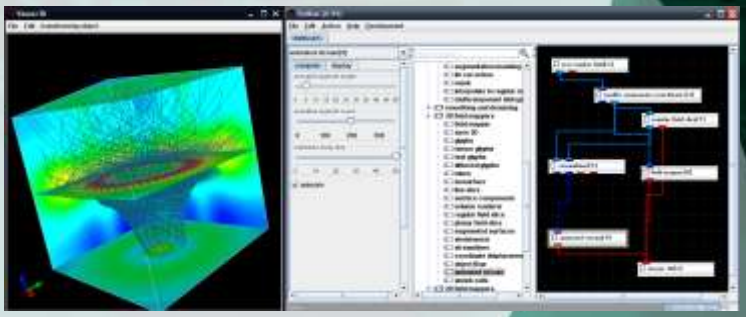
New solutions: VisNow Visual Analysis Software

- Distributed Visualization Engine (MPP,SMP, GPGPU)
- In situ Visualization
- **Visualization of simulation results**
  - Cosmology
  - Turbulent flows
  - Neurobiology
  - Biomedicine
- http://visnow.icm.edu.pl (Open Source license)

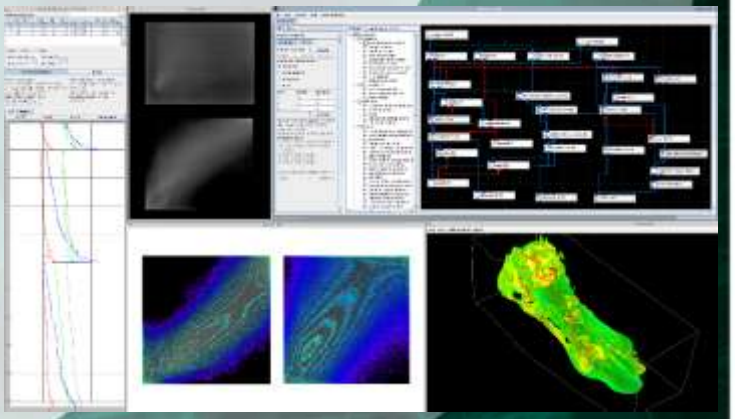


HPC visualization software

- **VisNow** http://visnow.icm.edu.pl
  - Generic **visualization and visual analysis platform**
  - Developed in **Java** at ICM

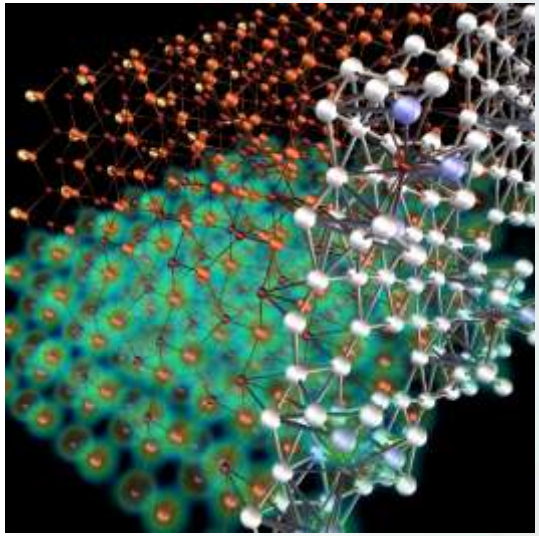


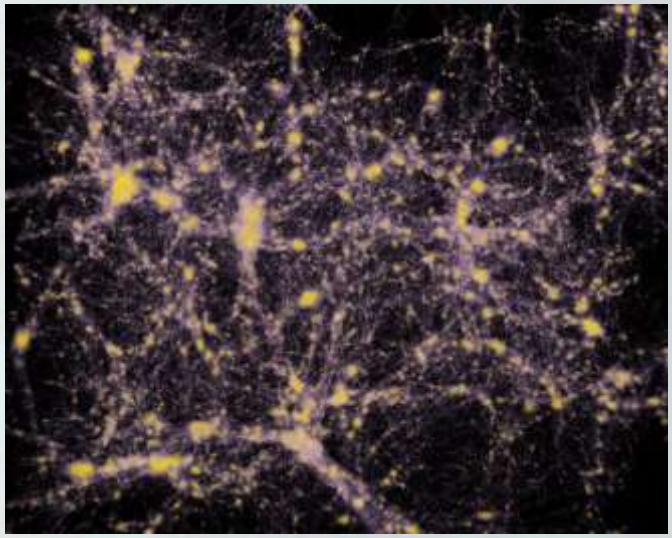
  - **Features**
    - Modular
    - Data flow driven
    - Pluggable
  - **Philosophy**
    - Read-And-Watch – instant visualization
    - Multifunctional modules
    - Module-object-interface connection
    - Reasonable default values
  - **New developments:**
    - Large datasets support
    - Distributed resources
    - Batch processing

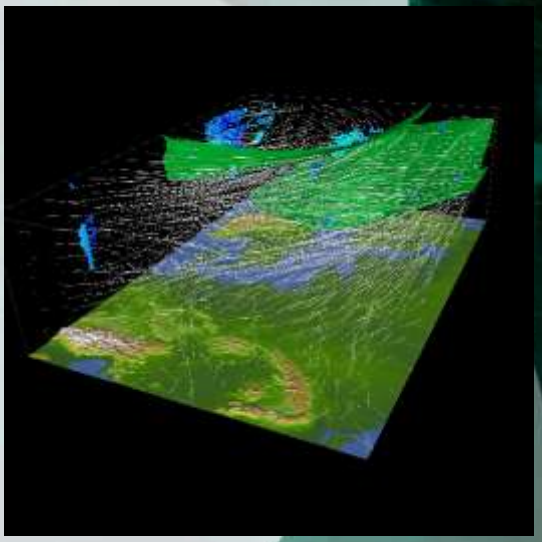


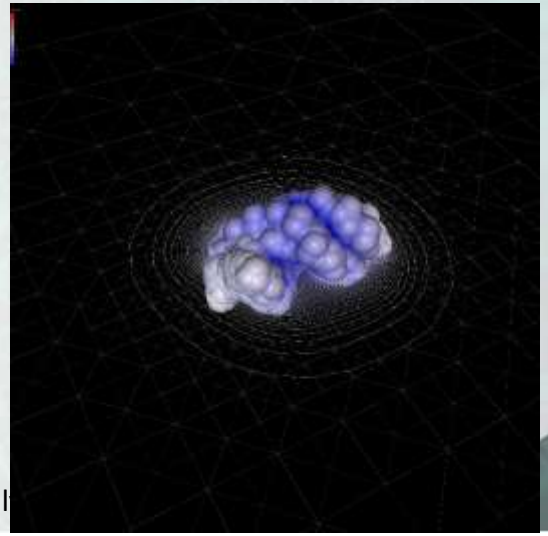
UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

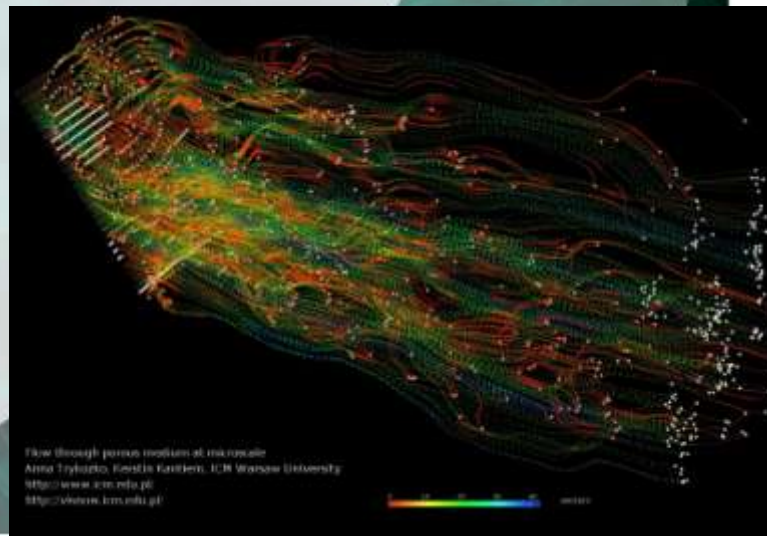
ICM: software and services



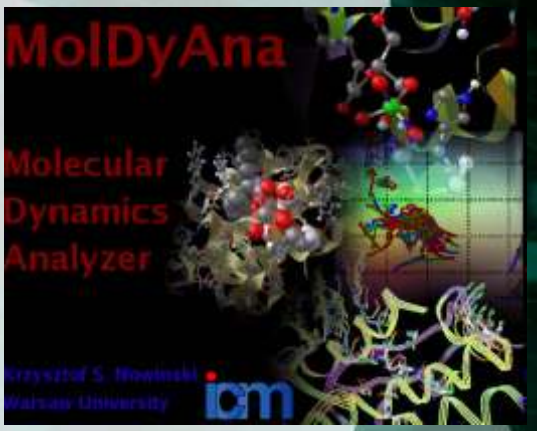








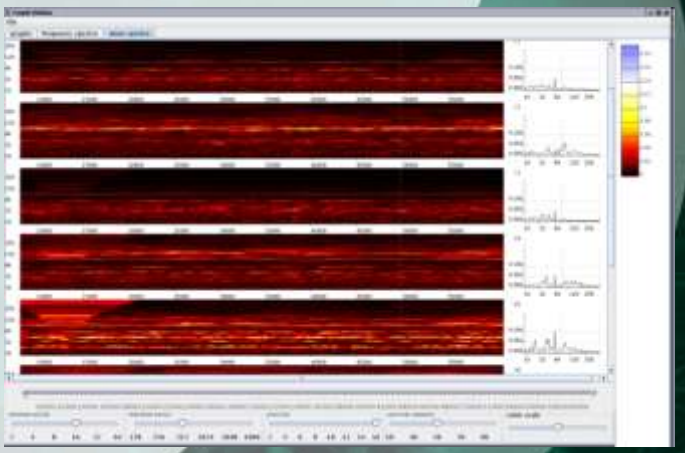
VisNow for molecular design

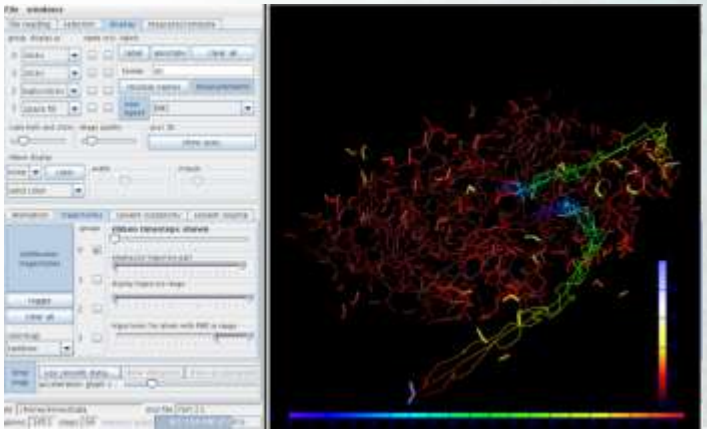


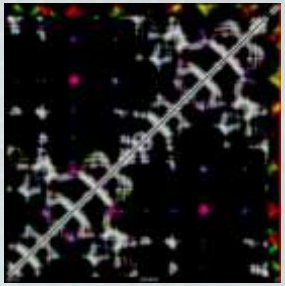
MolDyAna

- Interactive tool for **molecular dynamics** data analysis and visualization
- Dynamics animations
- Geometry measurements
- Density charts
- Trajectory analysis
- Spectral analysis

http://moldyana.icm.edu.pl







UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

Numerical weather forecasting and dependent services

- Multi-model multi-grid processing, mesoscale
- Top horizontal resolution in development 1km
- Vertical resolution: 70+ layers

- Services:
  - Severe weather warning systems
  - Energy sector
  - Transportation and logistics
  - Agriculture
  - Space program
- R&D solutions:
  - Energy smart grids
  - Precise agriculture
  - Airspace management

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

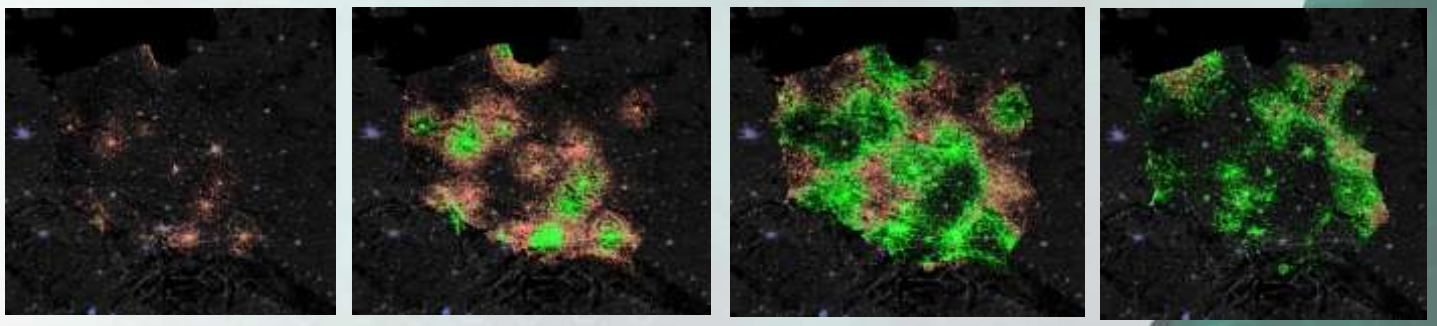
Personalized medicine and healthcare

- Cardiac interventions:
  - Non-invasive diagnostics
  - Cardio-intervention design
  - Remote monitoring and diagnostics
- Cellular and molecular scale computational modeling:
  - Disease profiling
  - Personalized drug design
- Healthcare:
  - Remote monitoring
  - Remote diagnostics
  - Epidemiology (structured populations over networks)

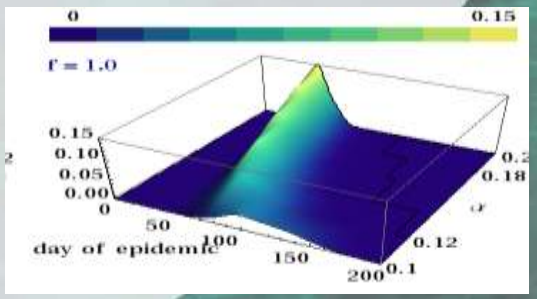
UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

Infectious disease spread over Poland: computational Agent Based Model and its predictions

- Simulations based on detailed, statistically relevant artificial social structure and mathematical model of the infection
- Model predicts space-distributed epidemic outbreak:



- and location of the epidemic peak in time, dependent on different socio-biological circumstances:



UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

Marek Niezgódka

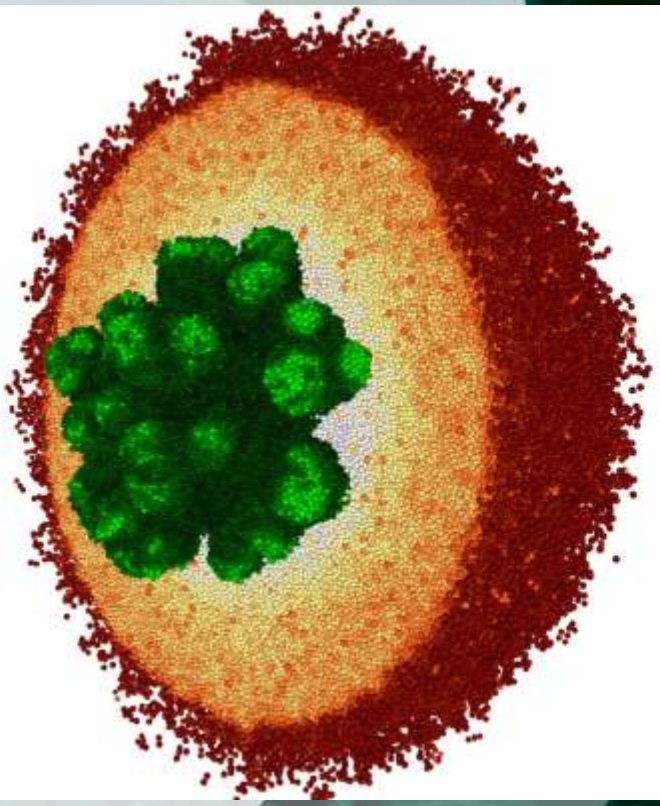
marekn@icm.edu.pl

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

ICM: some more R&D references

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

Case study: biological cellular processes at tissue scale



- **Main objective**: set up a methodology and powerful computational algorithms to enable large scale modelling of cellular biosystems
- **Multiscale approach**: simulation of cellular biosystems dynamics, individual cell processes and environment
- **Hybrid computational model**: solving off-lattice many body system, PDE's describing cellular environment and their interactions
- **Challenge**: enabling simulations on clinically detectable scales (10^9 cells = $1cm^3$ tissue)

Ref.: Large Scale Parallel Simulations of 3-D Cell Colony Dynamics, M.Cytowski, Z.Szymańska, IEEE Computing in Science & Engineering, 2014

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

POWIEW program

Coalition: ICM, Cyfronet, PSNC
http://wielkiewyzwania.pl

Large scale scientific computing projects:

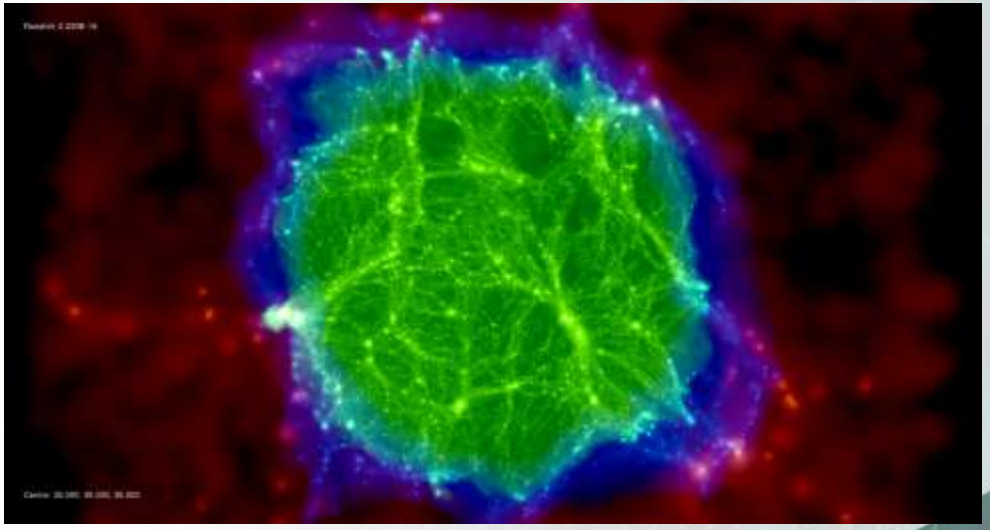
- Numerical Weather Prediction
- Semiconductor modelling
- Modelling and Visualization of RNA Structures
- Neuroinformatics Simulations
- Modelling of the Structure of the Universe
- Molecular Modelling
- Reservoir Modelling
- Astrophysics and Radio Astronomy
- Visual Analysis

Main computing architectures:

- MPP systems – IBM BlueGene/P & follow-up
- Fat Node systems – IBM POWER7 IH
- SMP systems – HP Blade Center Versatile SMP (vSMP)
- GPU-based hybrid systems – HP SL390s nVidia Fermi

Case study: Copernicus Complexio

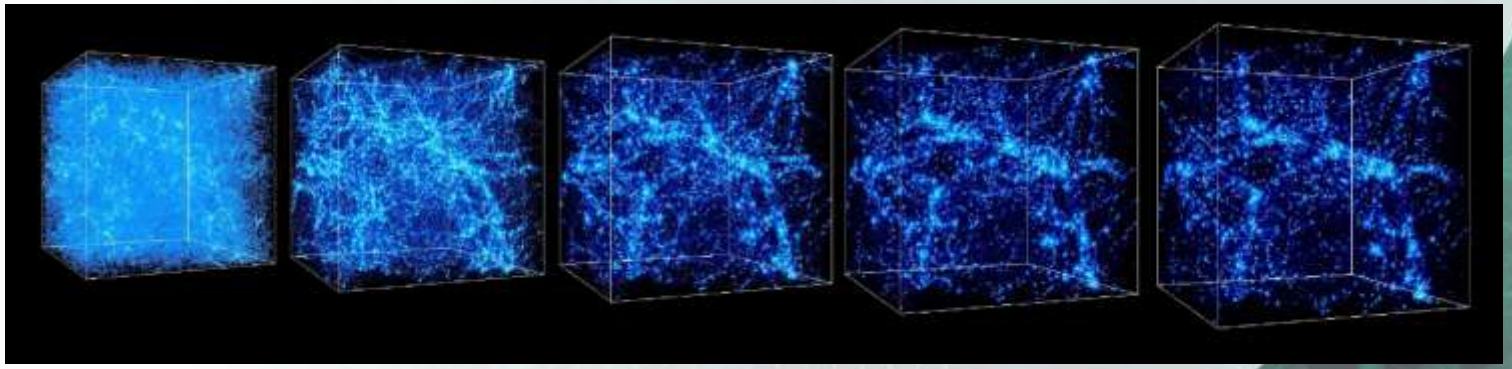
- **Zoom-in type simulation:**
  - **Ultra high-resolution**: over 13 bilion particles in the middle of computational domain (green)
  - **Medium resolution**: transition zone (blue)
  - **Lowest resolution**: sufficiently large segment of the Universe (red)
- **Technical details:** 70 Tflop/s & 10 TB computational partition, approx. 4 weeks wall clock time
- **Analysis**: results of the CoCo simulation are being analysed by international cosmological consortium
- **Visualization available on Vimeo**: http://vimeo.com/76812335



UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

Modelling of the Structure of the Universe

- Large scale simulations with N-body codes (Gadget3, GotPM)
- **Warsaw Universe Simulation**, 2048^3 particles
- MPP-based results analysis:
  - Statistical methods
  - Topological classification
  - Geometrical classification
  - Delaunay Tesselation, Alpha-shapes and Betti numbers



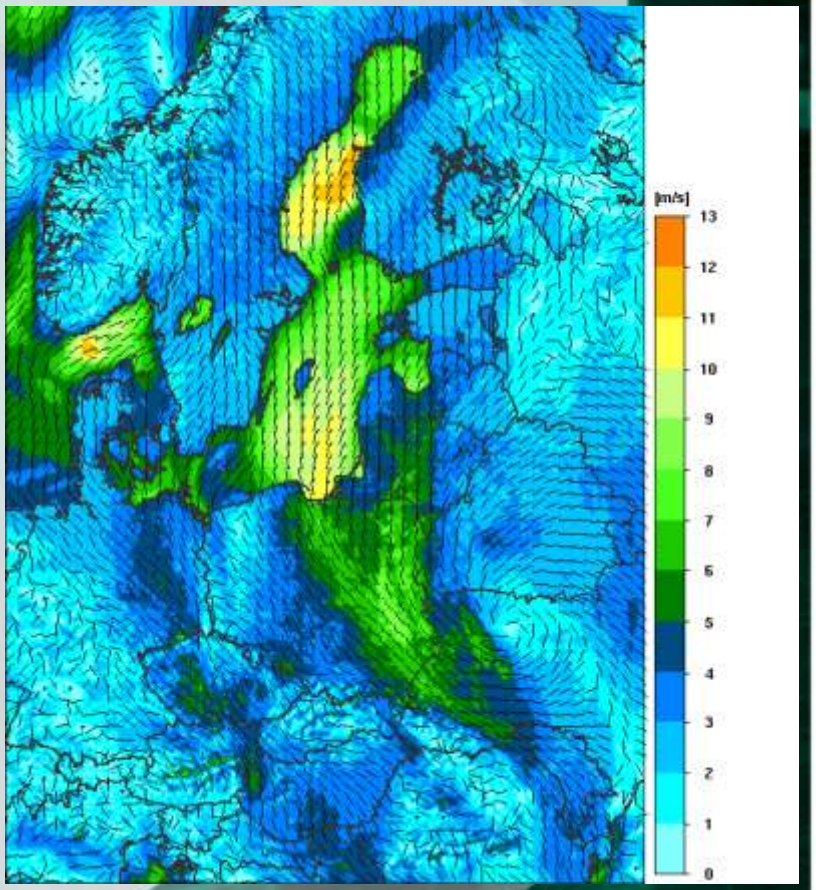
Case study: Numerical Weather Forecasting

- **Operational numerical weather forecasting for Central Europe**

  http://meteo.pl

- **Unified Model 60h forecasting,** current resolution approx. 4km, planned resolution approx. 1km
- **IBM Power 775** – 5x speedup over x86 cluster

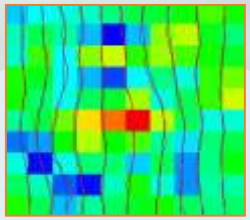
- **Technical details:**
  - 4 runs per day
  - 24 nodes, 768 Power7 cores
  - approx. 20’ walltime per run



UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

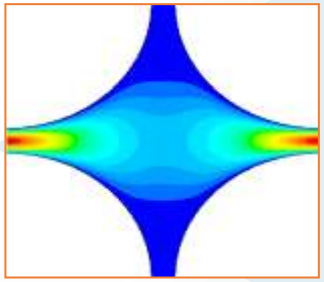
Porous media: multiscale modeling

FLOW AND TRANSPORT AT MESOSCALE



Darcy model

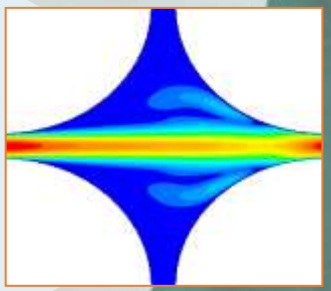
Non-Darcy model



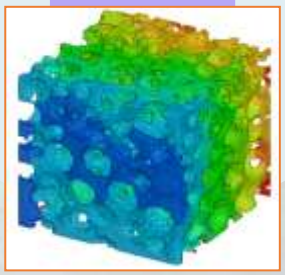
slow flow

UPSCALING

fast flow



Stokes equations



Navier-Stokes equations

FLOW AND TRANSPORT AT PORESCALE

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

Modeling flows through porous media

- Computational challenges:
  - Flows at pore scale for large range of velocities
  - Upscaling to core scale: effective parametres and model fitting
  - Inertia effects at pore scale ⇒ non-Darcy models at core scale
  - Modeling processes resulting in pore clogging and their impact on core scale parameters
  - Computations based on realistic geometries obtained from microimaging (X-ray computed microtomography)

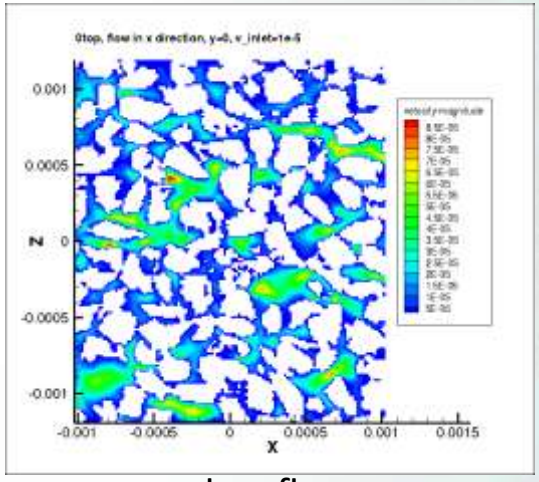
UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

Porous media: a virtual laboratory set-up

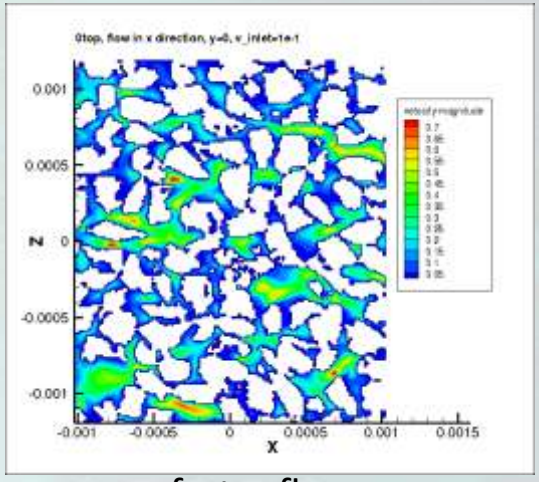
- Pore-scale geometries
  - Synthetic: periodic, regular patterns, random
  - Realistic data: micro imaging
- Grids at pore-scale
  - 2D and 3D models
  - Triangular and quadrilateral (voxel-based) unstructured grids
- Computations at pore-scale
  - Direct Numerical Simulations by Finite Volumes
- Upscaling and model fitting

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

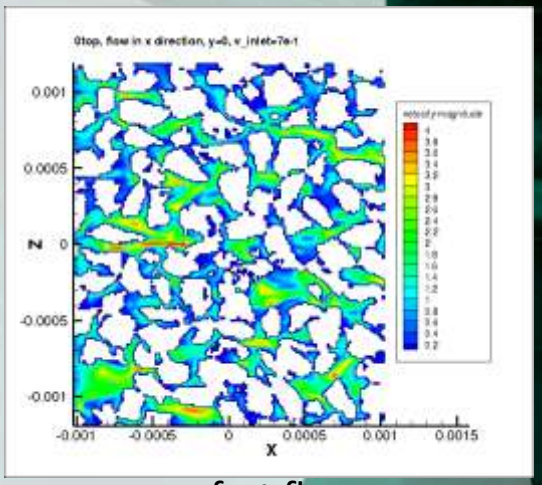
Upscaling from pore to core scale



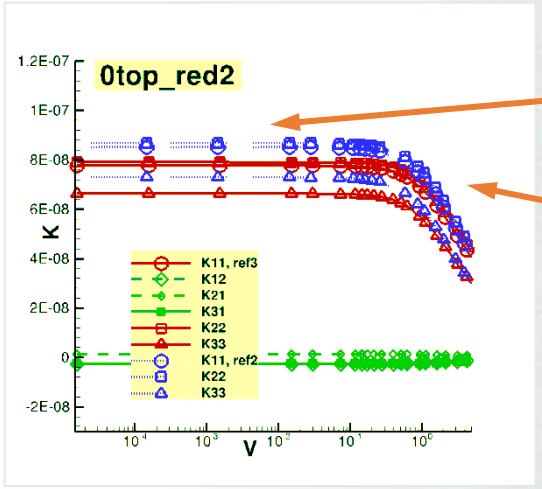
slow flow



faster flow



fast flow



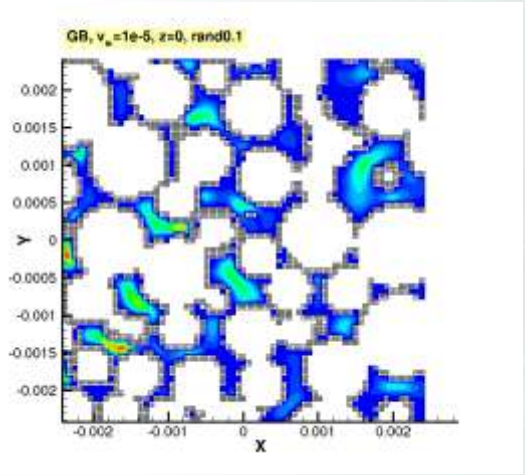
SUMMARY in terms of upscaled values

- Darcy's regime: constant ratio of P (upscaled pressure) and V (upscaled velocity)
- Inertia effects – non-Darcy models
- Accounting for anisotropy

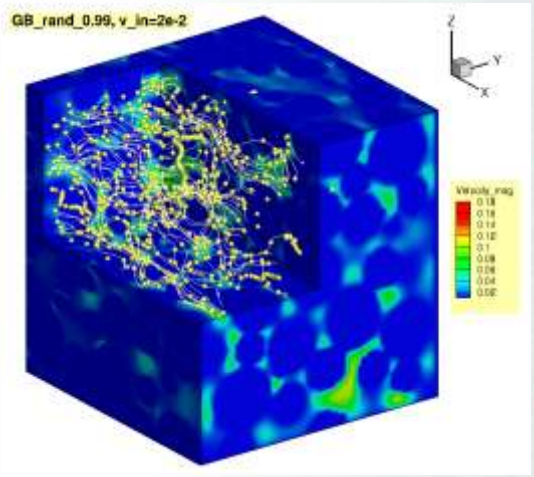
M. Peszyńska, A. Trykozko: *Pore-to-core simulations of flow with large velocities using continuum models and imaging data*, Computational Geosciences, Vol. 17, nr 4, (2013), 623-645

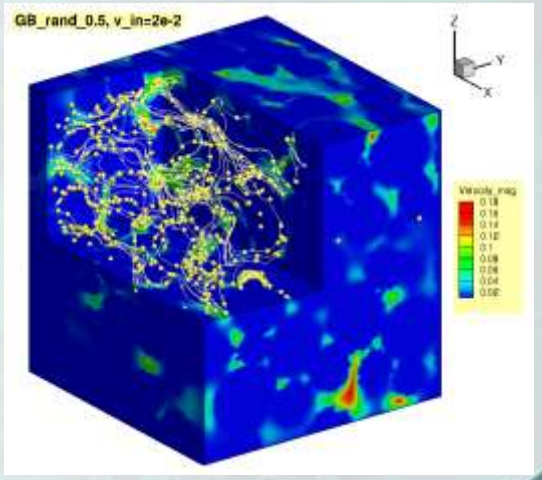
UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

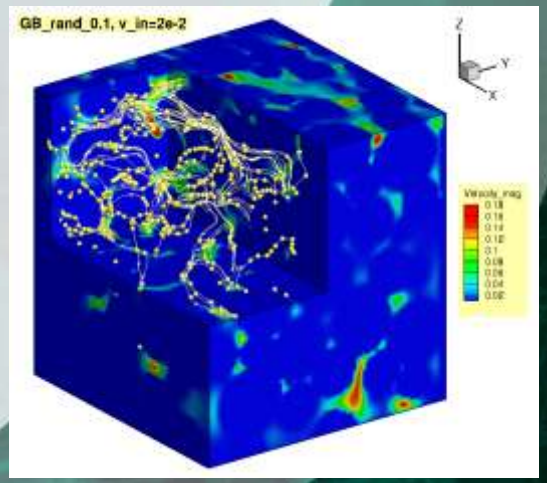
Impact of pore space clogging on core parameters



- Changes in pore geometries as causes of: biofilm growth, sedimentation, reactive flows, antrophogenic actitivite, …
- Significant changes in permeability due to clogging
- Figures: random model of clogging (irregular pore lining)

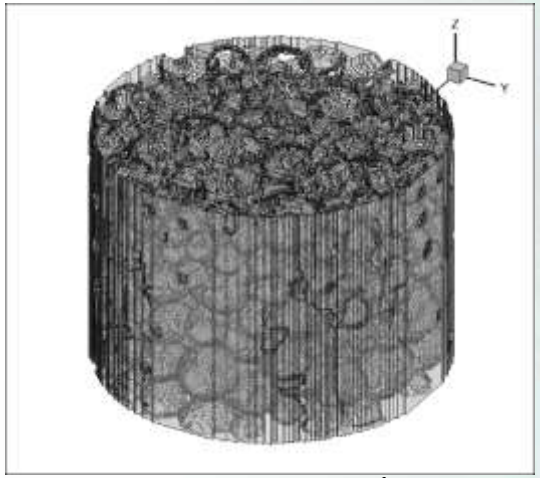




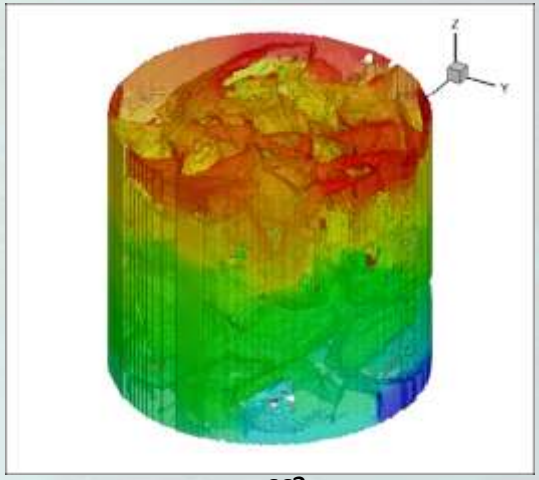


A. Trykozko, M. Peszyńska: *Pore-scale simulations of pore clogging and upscaling with large variations*, GAKUTO International Series, Mathematical Sciences and Applications, Vol. 36 (2013), 277-300.

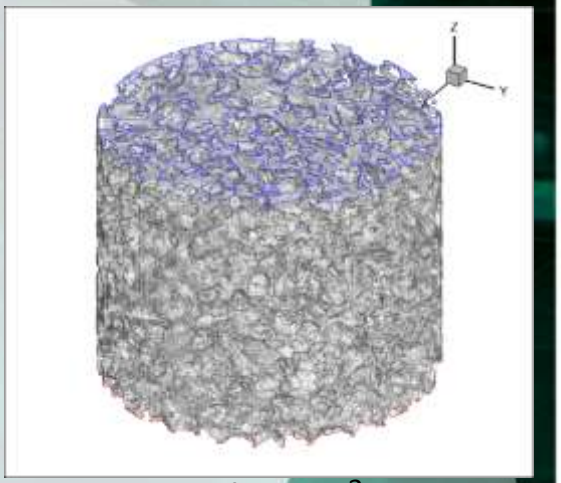
Challenges in 3D pore-scale simulations: SIZE!!!



Glass beads[1]



Tuff[2]



Sandstone[3]

- Complexity ↑30M cells, 10Gb files, challenging visualization and post-processing
- Grid reduction (data coarsening) necessary
- Sampling vs. *Reference Elementary Volume* size
- Mesh refinement

1[DWildenschild], 414x414x300, largest problem: 30Mcells, 9Gb, voxel 34µ.

2[DWildenschild], 431x436x380, 23Mcells, too small to be *REV*

3[BLingquist], 731x731x600, largest problem 8Gb, 27Mcells, voxel=16 µ.

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

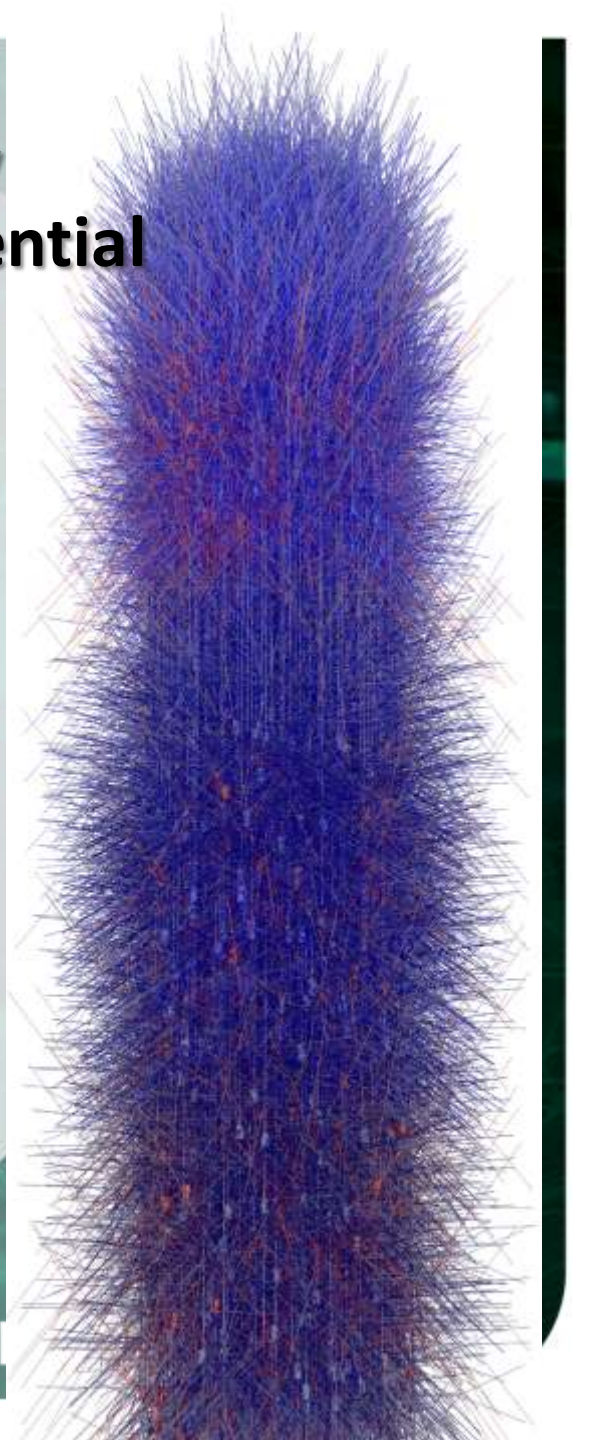
Case study: Activity of the sensory system and extracellular field potential

Main objectives:

- **Simulations of activity** of the sensory and visual neuronal systems
- **Simulation of response** to physiological stimulation and electrical stimulation
- **Computational toolchain:** from computer-aided construction of neuronal systems to large scale simulations and efficient analysis

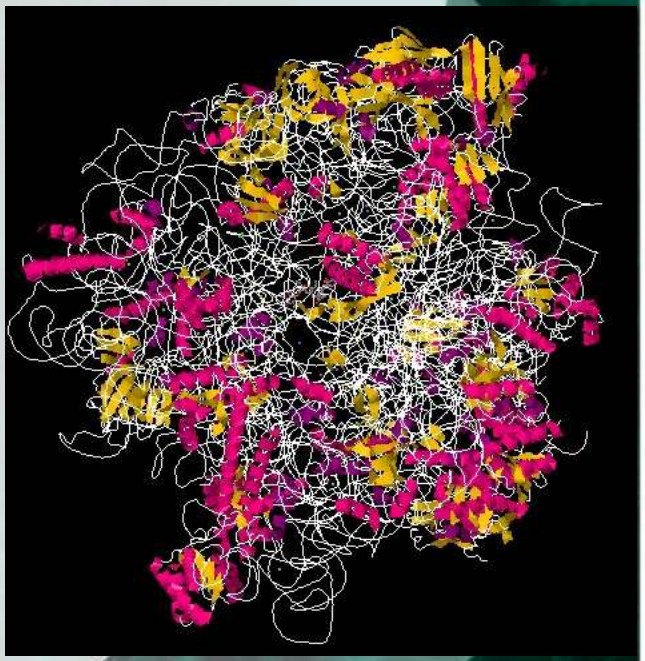
The resulting tool and models allow for the validation of methods for analysis of **experimental data** and the verification of hypotheses posed in the experimental work, which may be difficult to verify in another way.

Example research field: *Analysis of dynamics of thalamic activation in stimulations of somato-sensory pathway in anesthetized rat and activity in the barrel cortex in behaving rats*



High-throughput modelling of functionally and therapeutically relevant spatial RNA structures

- Modelling of spatial RNA structures
- Design of new therapeutic targets aimed at the RNA
- Validation of ICM's RNAComposer software for fully automated spatial RNA models generation from structural fragments



UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

Modeling of growth and structure transformations in heterogeneous systems

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

Mathematics within multiscale modeling framework

Nonlinear dynamics of spatial developments:

- Phase transformations
- Chemical reactions/ process thermodynamics
- Rheology / memory / nonlocal interaction mechanisms
- Structured populations
- Coupled multiscale systems
- Geometric / topological evolution
- Uncertainty components / **sensitivity** aspects

UNIVERSITY OF WARSAW
Interdisciplinary Centre for Mathematical and Computational Modelling
www.icm.edu.pl

Specific applied modeling problems

Phase separation and related phenomena:

- nonlinear variational problems
- process dynamics
- structure formation

Phase transformations, crystal growth, physiology:

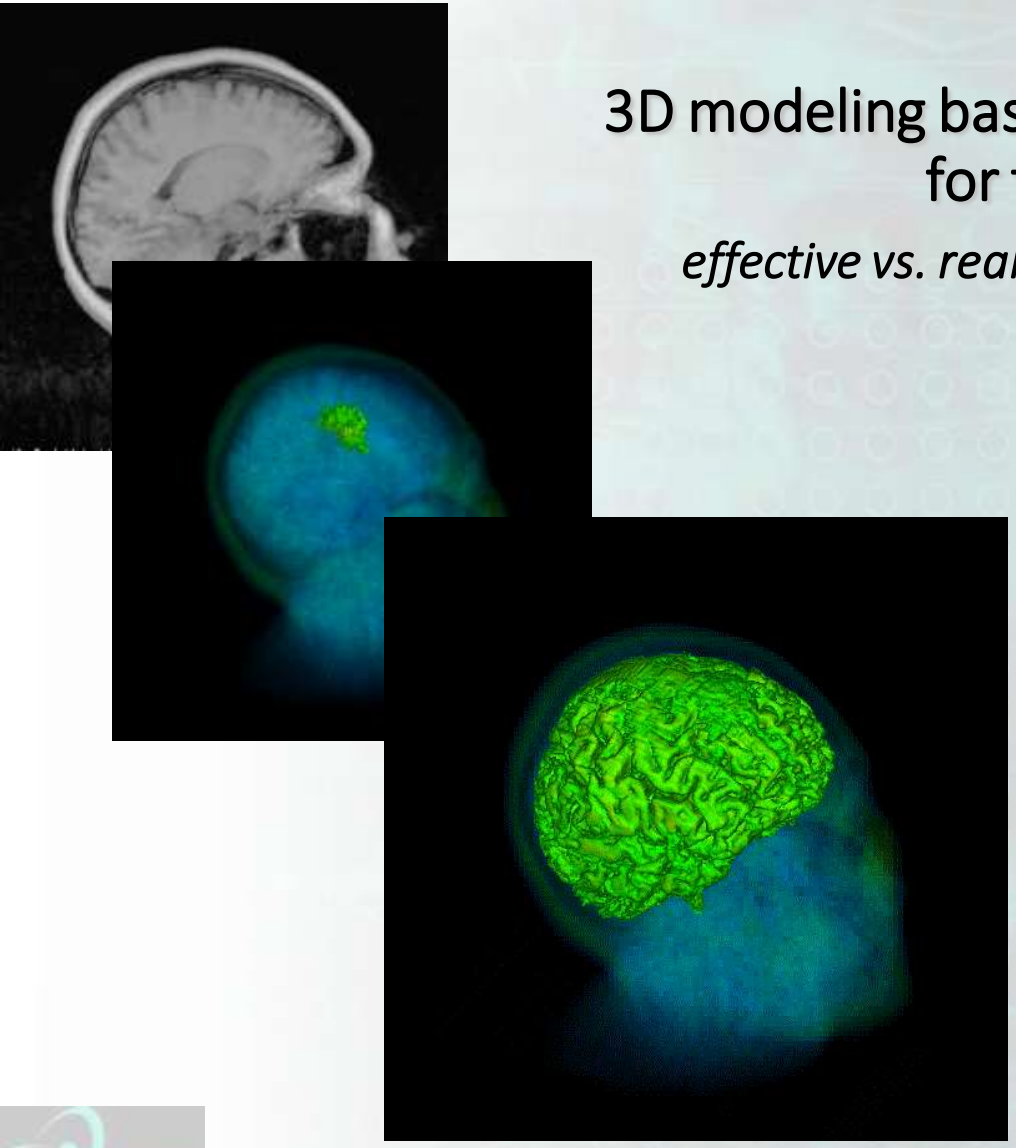
- biomedical processes: blood circulation, tumor growth
- process design & control

Population dynamics:

- structured populations
- operational (stochastic) networks

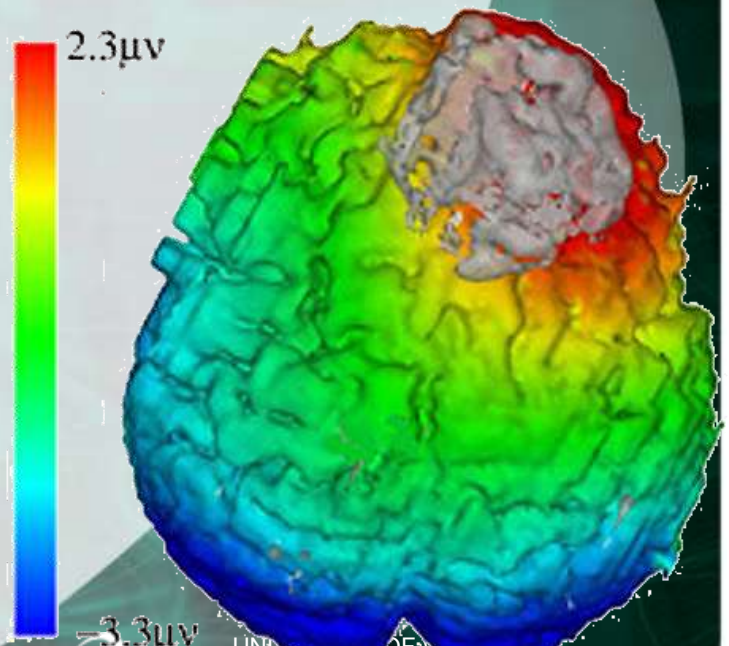
Applications in biomedical process modelling and physiology

Visual neurocomputing:

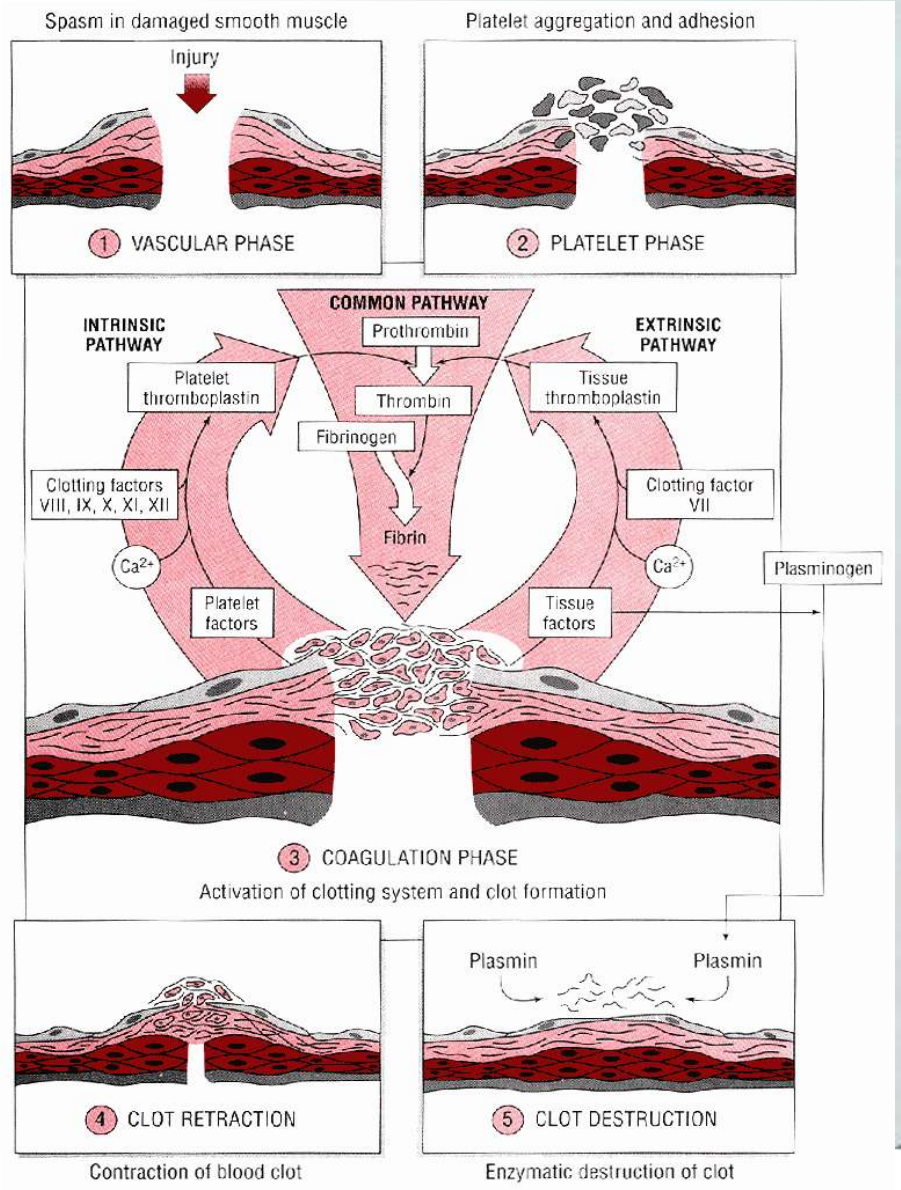


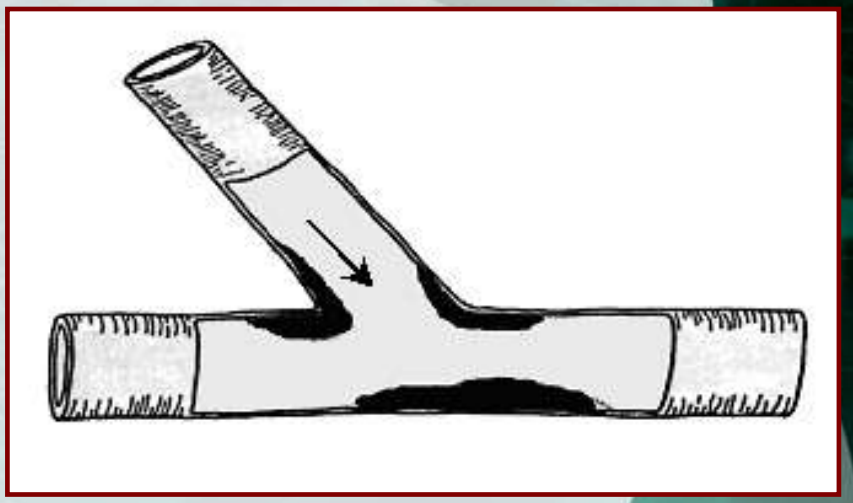
3D modeling based on of multimodal data for therapy/surgery planning:

effective vs. real medium computational model



Thrombosis: an outline of temporal sequencing and effects





Process dynamics in blood circulation

- variable geometry
- free boundaries
- complex flow nature
- chemical reactivity
- multiscale
- granularity
- role of mechanics
- rheological features
- thermal sensitivity **(→HSPs)**

Interscale model consistency: validation

- Cell populations dynamics
- Degenerate tissue growth
- Cell populations dynamics, pathological cell populations growth spreading mechanisms
  - Human agents and their behaviour: transport-related reference patterns
  - Process dynamics in large-scale networked environments: local and global passenger-centric model
  - Information transfer trees: hybrid and continuous models in information spaces