POZNAŃ SUPERCOMPUTING AND NETWORKING CENTER

New approaches for advanced computing and data processing in climate modeling

Marek Błażewicz, Milosz Ciżnicki, Krzysztof Kurowski, Michał Kulczewski, Piotr Kopta

Applications Department

POZNAŃ SUPERCOMPUTING AND NETWORKING CENTER

PSNC overview

Center for Development of e-Infrastructure

- Research Metropolitan Area Network - POZMAN
- HPC Center
- National Research and Education Network PIONIER
- Digital Libraries Federation

Center for R & D

- New Generation Networks
- Clouds & Grids
- Digital Libraries and Portals
- Technology, Applications and Services
- Cyber Security

PIONIER – Polish Optical Network

- **Area**
  312k sq km
- **Population**
  38M
- **Main academic centers**
  22
- **State universities**
  165+
- **Students**
  2M+
- **R&D institutions and Univ. interconnected via PIONIER network**
  700+

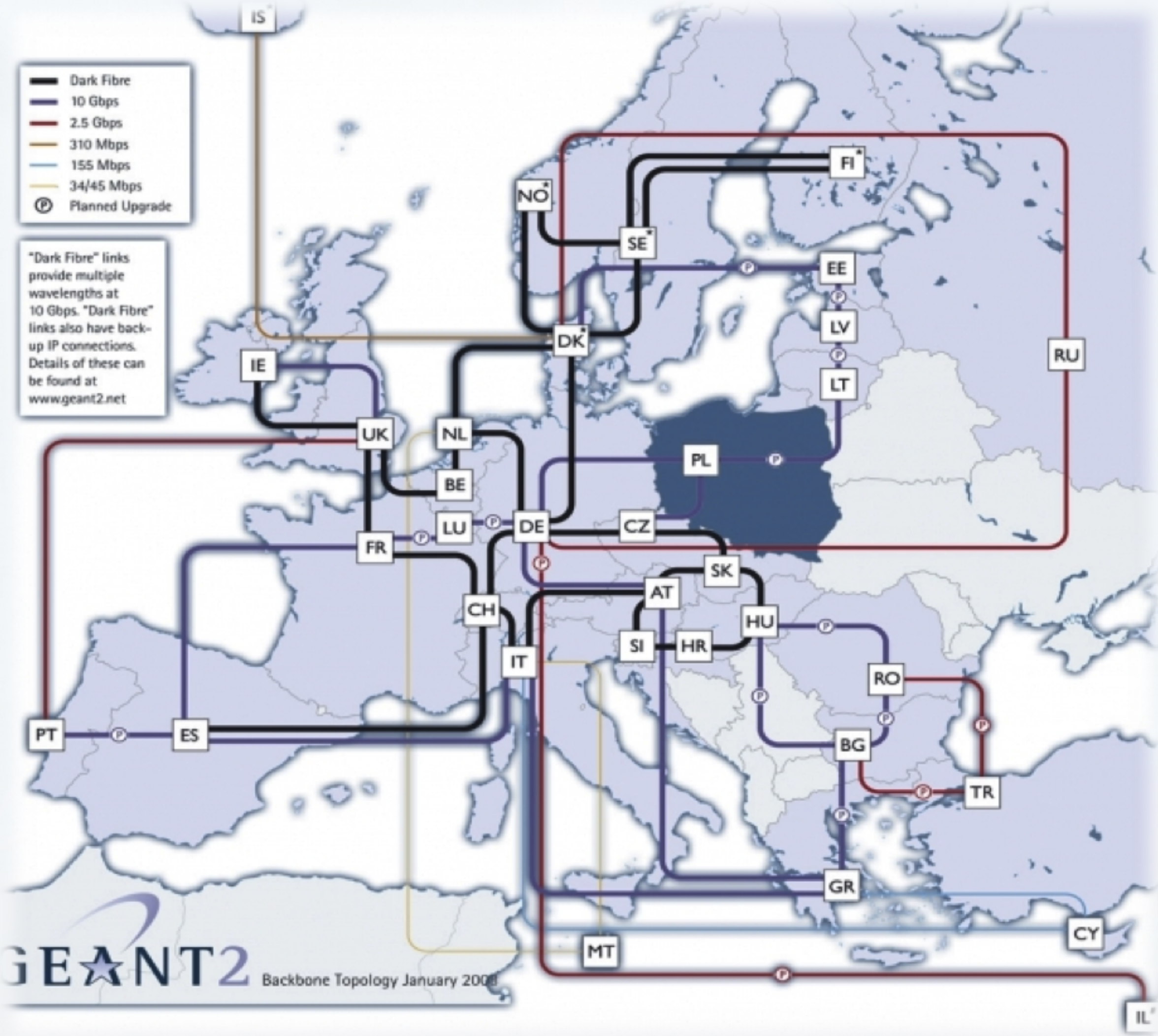
6494 km of fiber infrastructure in Poland

763 km of fiber in Germany (IRU)

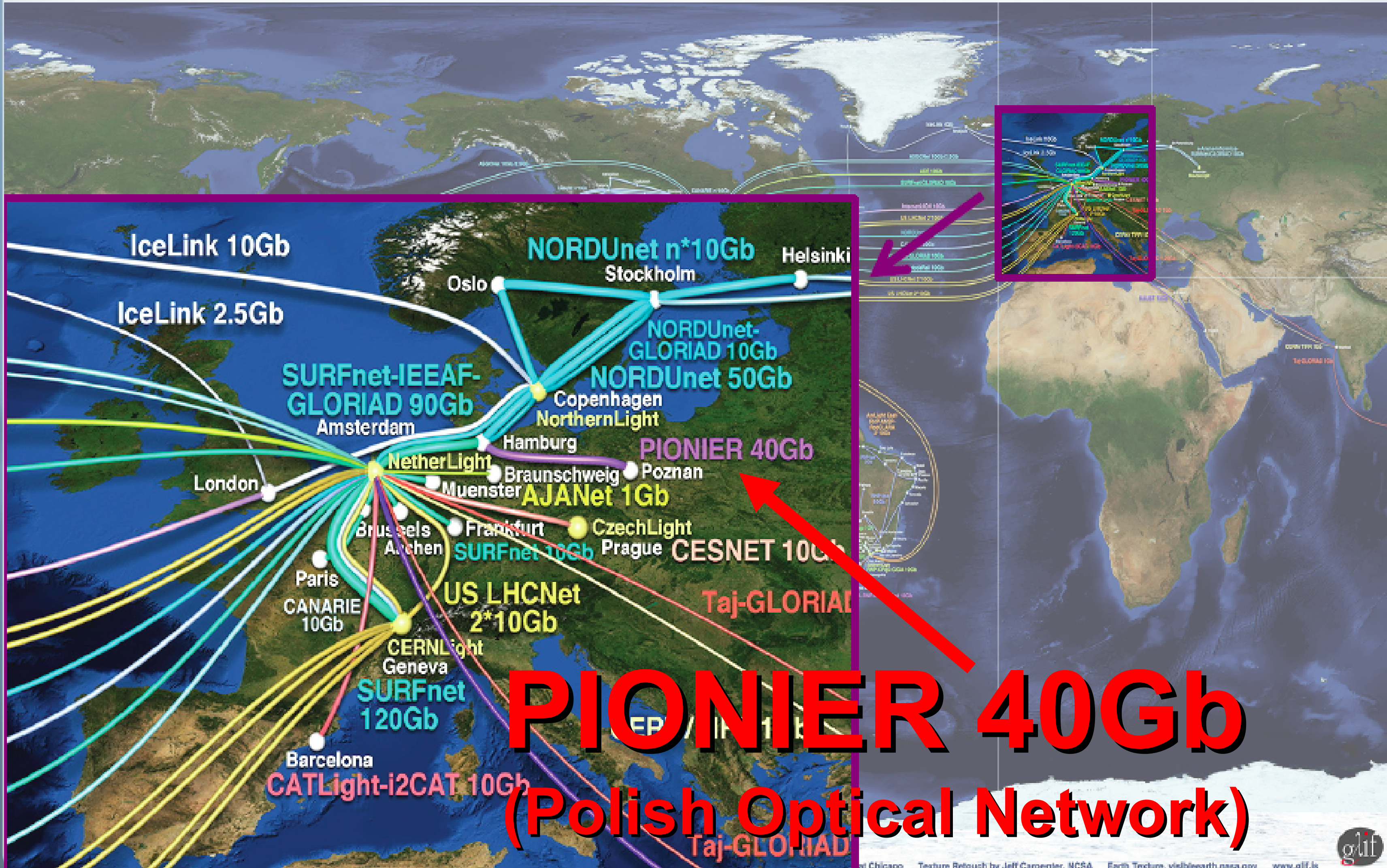
7257 km of fiber in total

22 MANs and 5 HPC Centers (Gdańsk included) in PIONIER Consortium

European optical networks



International optical networks



Networking + Storage + Computing

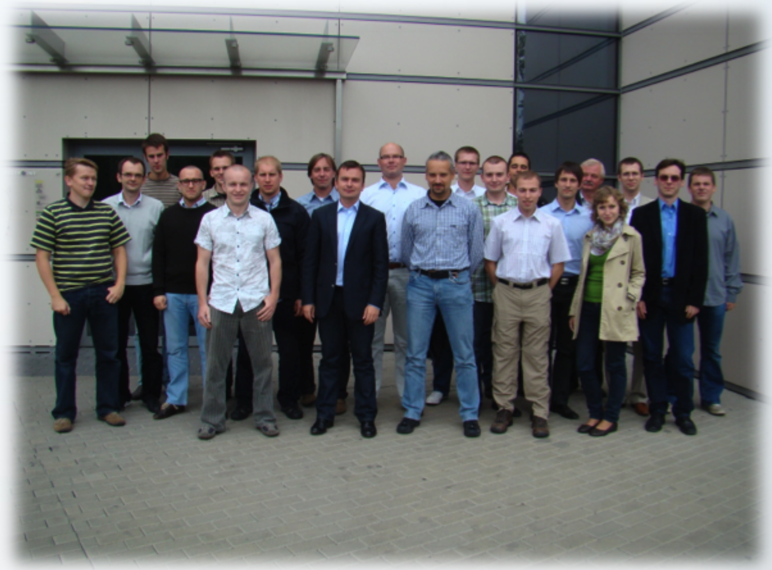
Key services for e-Science

- HD Videoconference services
- Deployment of eduroam services
- Applications on Demand for campuses
- Archiving on Demand
- Science HDTV services

Application Department @ PSNC

- 28 people + 14 students
- Advanced applications and computing, including accelerated computing, masive data analysis & high-resolution advanced visualization
- R&D acitivities in areas:
  - Computational science
  - Energy & environment science
  - Medicine
  - Material science
  - Biology
  - Chemistry
  - **Numerical weather prediction**

NVIDIA CUDA™ RESEARCH CENTER



POZNAŃ
SUPERCOMPUTING AND
NETWORKING
CENTER

HPC & key services

HPC Center

High speed network

- GEANT 2/3 (10 x 10Gb/s)
- National Research and Education Network PIONIER
- 10 Gb/s dedicated optical fiber connection to USA
- HD videoconference services

HPC Center

- Total peak performance: 304 TFlops (increasing)
- European Grid Initiative / National Grid Initiative (PL-GRID)
- European HPC infrastructure (PRACE)
- Applications on demand

Data Center

- National Data Storage
- Disk storage 4.6 PB (increasing)
- Archiving and backup on demand

Cane – GPU cluster (Tier-1 resource)	
Performance	224.3 Tflops
CPU	5448 x AMD Opteron 6234
GPU	334 x Nvidia Tesla M2050
RAM	10.64 TB
HDD	55.42 TB

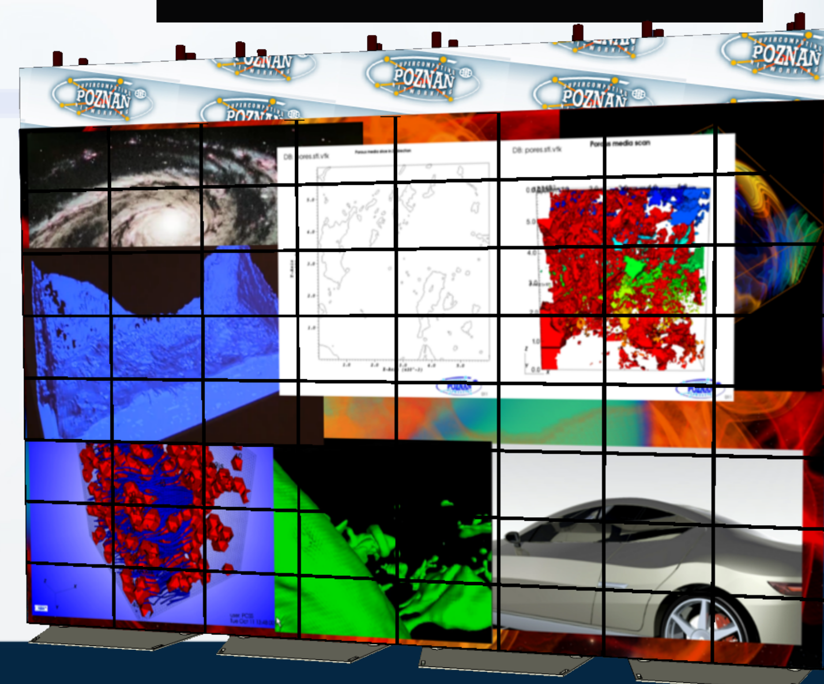
Chimera - SMP	
Performance	21.8 Tflops
CPU	2048 x Intel Xeon E7-8837
RAM	16 TB
HDD	1.2 TB

Reef - cluster	
Performance	54.27 Tflops
CPU	6176
RAM	10.208 TB
HDD	163.668 TB

Advanced visualization (up to 16K)

Vitrall

- Distributed web-based visualization system
- Multi-GPUs support
- Remote (web browser) or local (attached screens) rendering
- Support for sensor fusion and multi-touch based user interfaces



POZNAŃ
SUPERCOMPUTING AND
NETWORKING
CENTER

Advanced computing

New approaches for advanced computing

Motivation
- Code adoption to new architectures
- Increase in efficiency, robustness, model resolution

Architectures
- MPI nodes with GPUs
- MPI with OpenMP nodes
- Upcoming CPU+GPU hybrids

Approaches
- Task-based decomposition
- GPU (CUDA, OpenCL, Accelerators)
- Hybrid codes (MPI + OpenMP + GPU)

CUDA Research Center

- PSNC is CUDA research center from 2011
- CaKernel exa-scale GPGPU application
- CAVE visualization

MultiGPU Applications

Motivation:

- Common PC are often equipped with two or more GPUs
- HPC centers offer access to supercomputers with a large number of GPUs

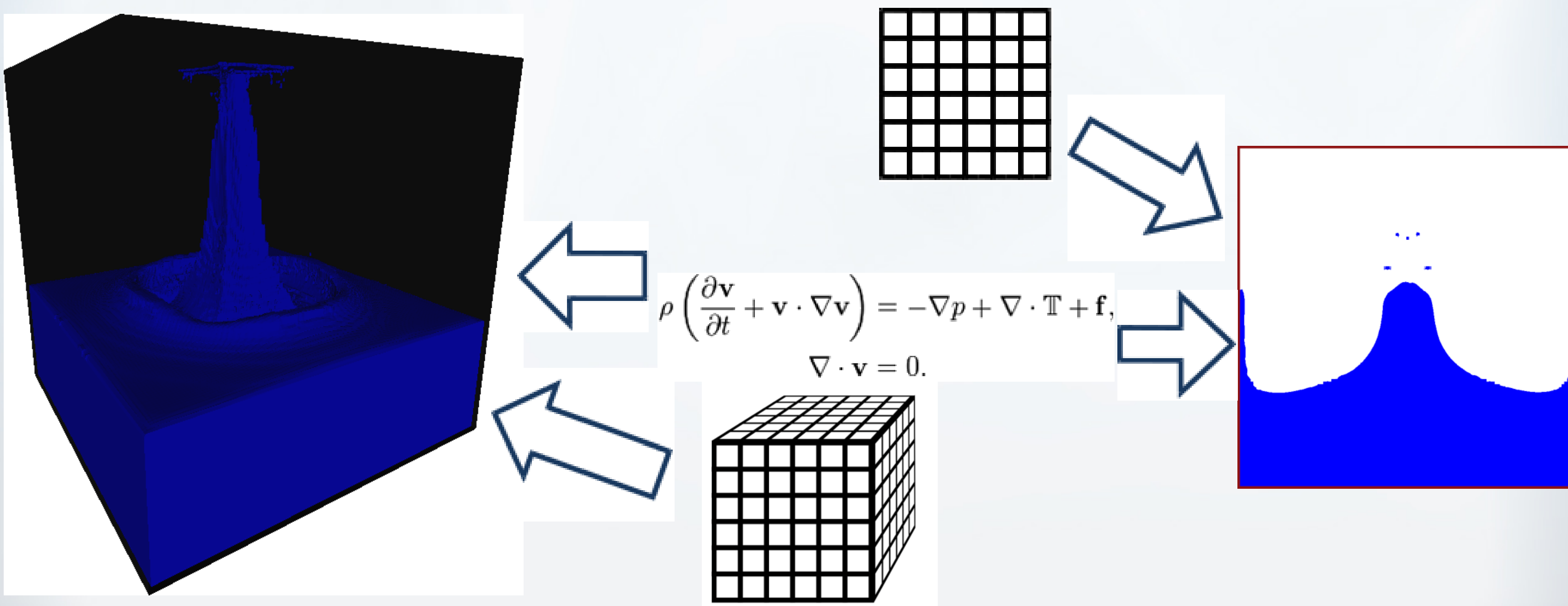
Communication issues:

- High parallelism in multi-GPU systems requires effective dealing with communication on many levels, e.g. between threads, blocks of threads, GPUs etc.
- Different memory types with its properties (bandwidth, latency) should be taken into account when planning communication
- Data distribution patterns have different pros and cons

Computational Fluid Dynamics (1)

Algorithm:

- Two and three-dimensional CFD simulation with a free surface approximation (VOF method);
- Navier-Stokes equations discretized using finite-difference method;
- Fluid propagation over an Eulerian grid;



$$\rho\left(\frac{\partial \mathbf{v}}{\partial t}+\mathbf{v}\cdot\nabla\mathbf{v}\right)=-\nabla p+\nabla\cdot\mathbb{T}+\mathbf{f},$$

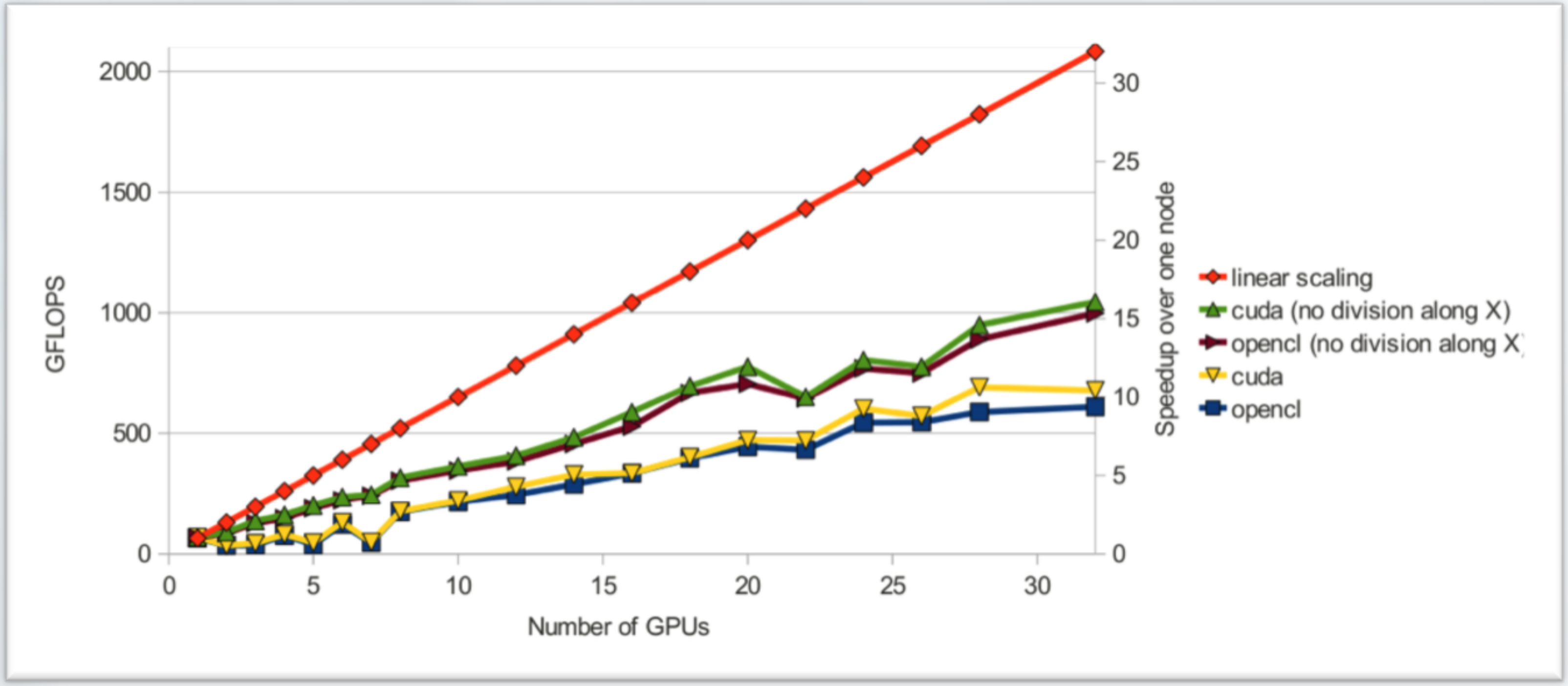
$$\nabla\cdot\mathbf{v}=0.$$

Volume of Fluid (VOF) Method for the Dynamics of Free Boundaries, C.W. Hirt, B.D. Nichols, 1979.

Computational Fluid Dynamics (2)

Results:

- MPI-based, Multi-GPU heterogeneous application;
- 80x speedup comparing to sequential CPU implementation;
- Good scaling on 32 GPUs.



Motion Tracking

Motion

- A need to track object efficiently in high resolution movies in real-time

Approach:

- Good points to track are selected by Harris corner detector
- Lucas-Kanade method is used to track points
- All steps are performed on GPU

Results:

- Real-time motion tracking in movies up to **4K**

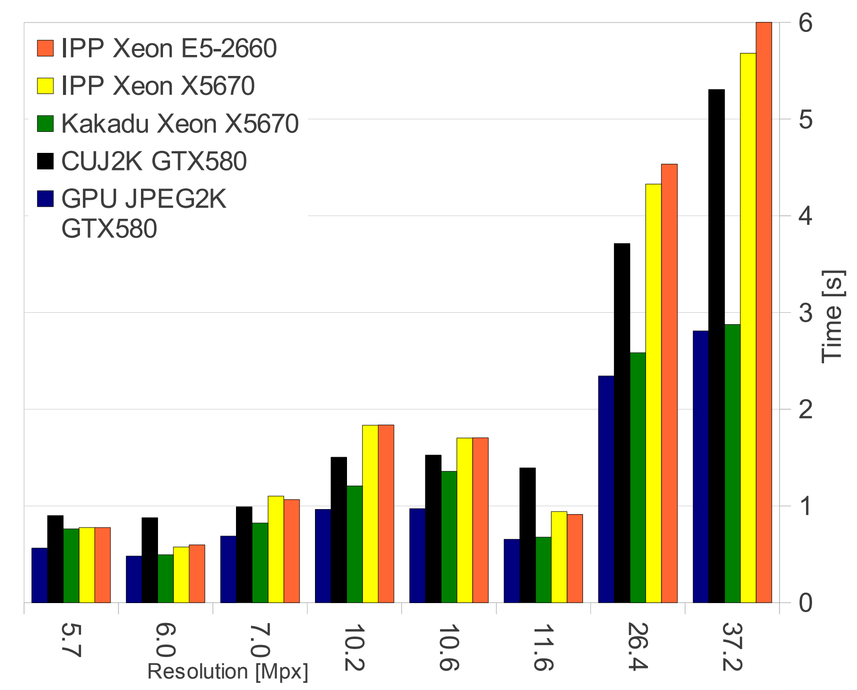
JPEG 2000

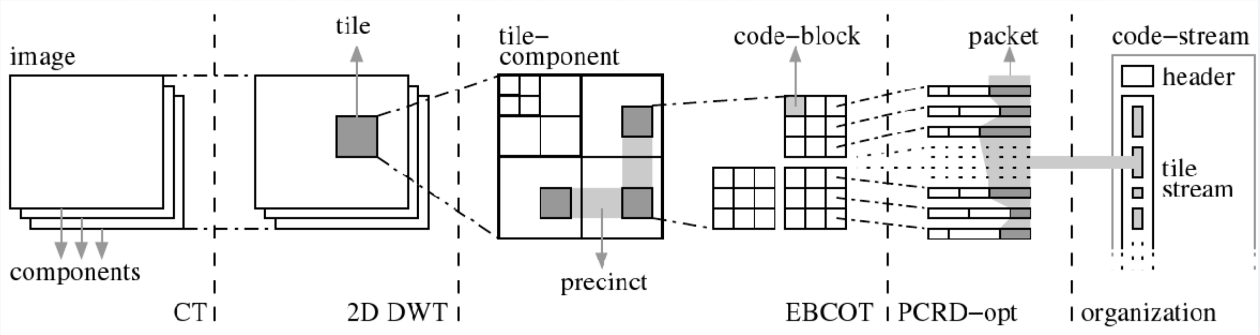
Motivation

- New image compression algorithm
- Better quality than JPEG
- Digital cinema standard
- Lossy hyperspectral image compressic
- Benchmarking

Approach

- Parallel implementation on GPU
- The fastest available implementation





JPEG2000 Data partitioning, coding and code-stream organization.

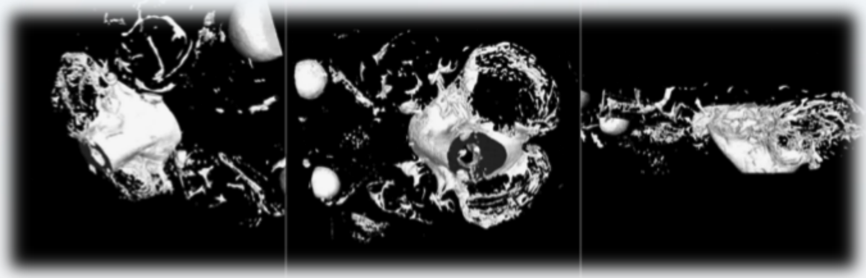
3D Tomography images (1)

Motivation:

- Existing implementations are capable of processing small domains only
- There's no application dealing with real-life models efficiently

PSNC's software:

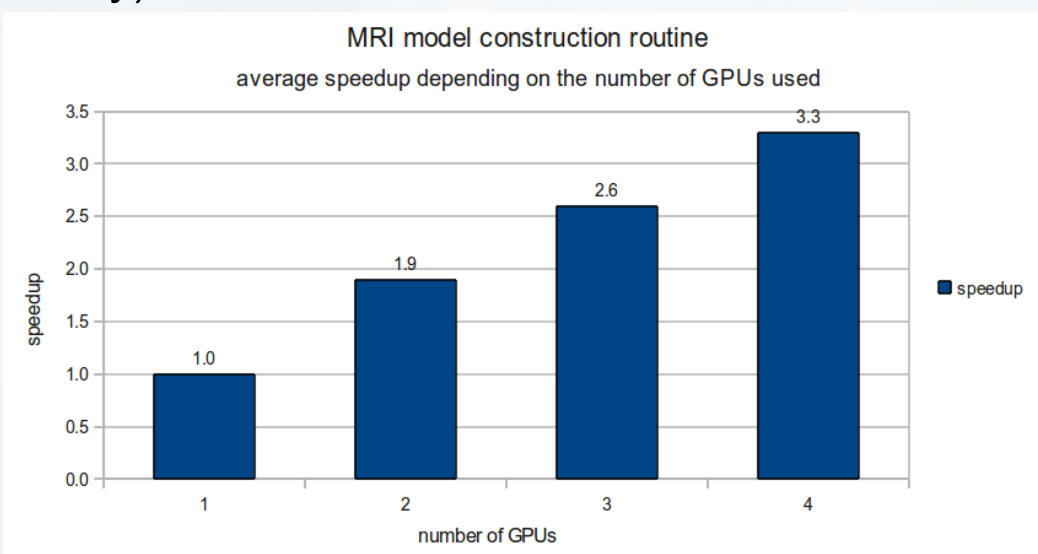
- Merging 2D slices of CT/MRI images into a 3D model
- Application: medical tomography (body scans), petroleum engineering (rock scans)
- Algorithm: Marching Tetrahedra creating isosurface from neighbouring pixels representing the same density

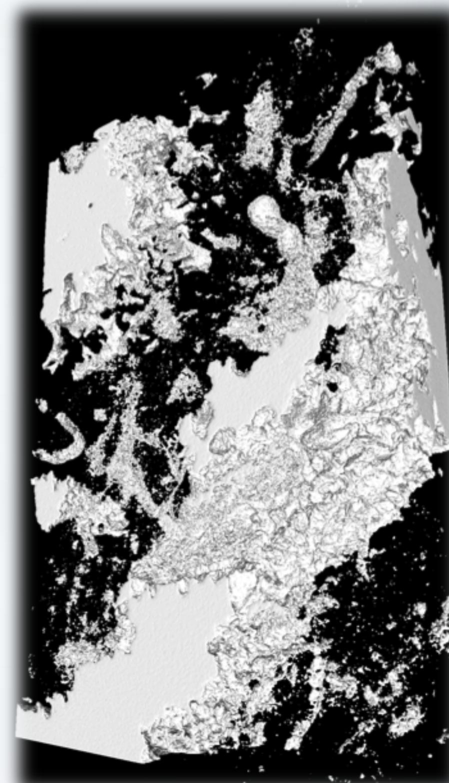


3D Tomography images (2)

Results:

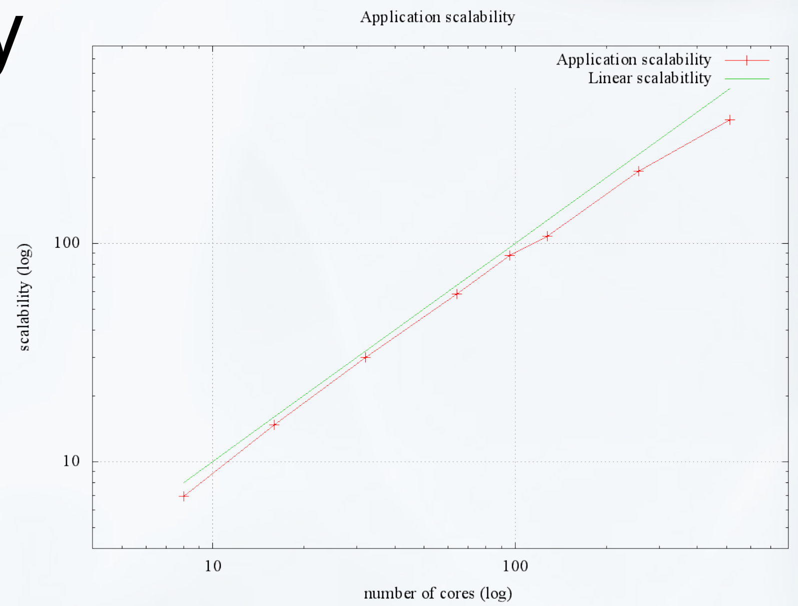
- ~ 50-fold speedup over single threaded CPU version
- Multi-GPU support scaling up well
- *isorock* – command line version designed to create rock structure models (computes open and closed porosity)

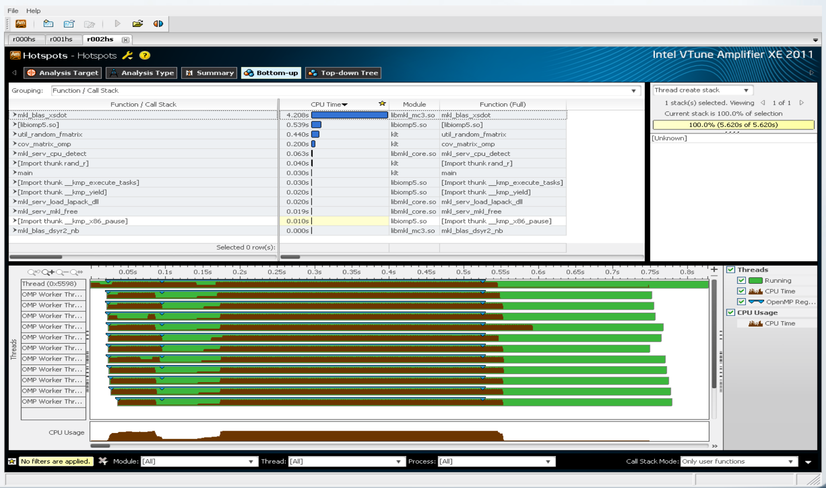




Quantum chemistry

- Code rewrite Fortran -> C
- Use of parallel numerical libraries (BLAS/LaPACK)
- Memory usage minimization
- Hybrid parallel implementation (for large number of cores optimization)
- Scalability up to 500 cores (limitations of the algorithm)
- Up to 6x speedup
- PRACE grant given





CaKernel/Cactus/Kranc

Cactus

- Framework for developing portable, modular applications solving 3D PDE time evolutions
- HPC
- Hardware architecture independence
- Programming language independence

CaKernel

- Integrated with Cactus, fully compatible with legacy applications
- Execution of routines on GPU and/or CPU without modifications to applications
- Keeps track of recent data changes and copies minimum amount required for consisted computations

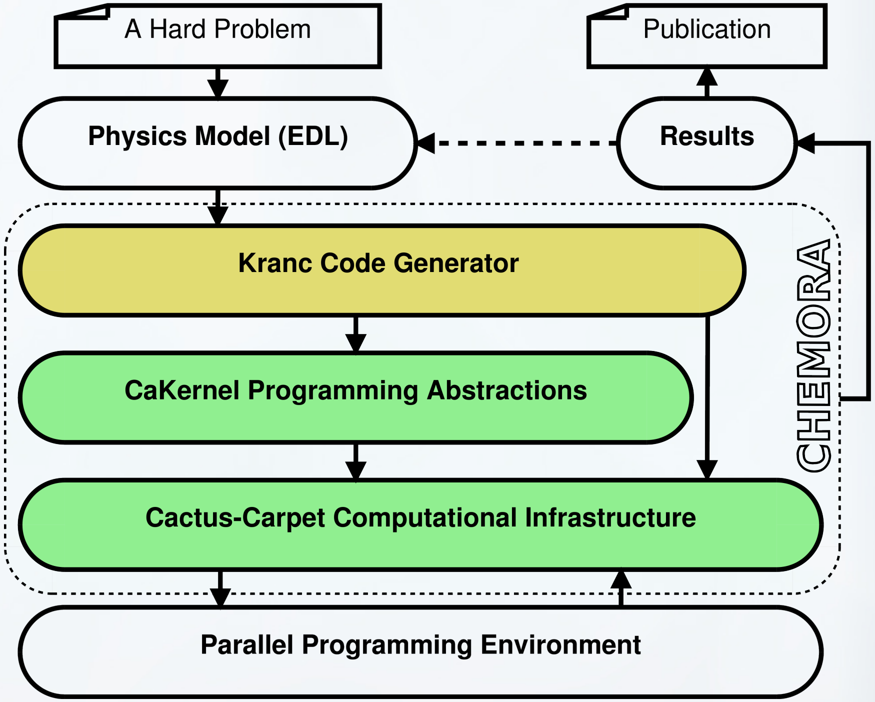
Kranc

- Mathematica application
- Translates tensorial description of a time dependent partial differential equation into a Cactus module
- Very efficient time-to-solution factor
- Mathematica translates and simplifies equations for their more efficient compilation/execution

Chemora = CaKernel +Cactus + Kranc

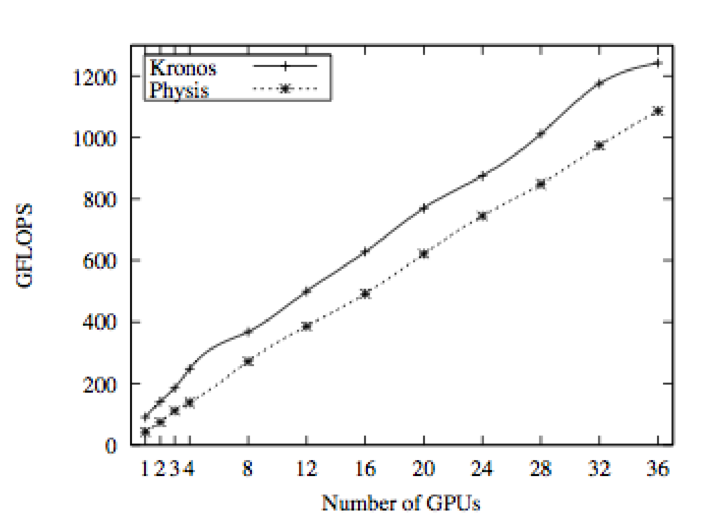
Results:

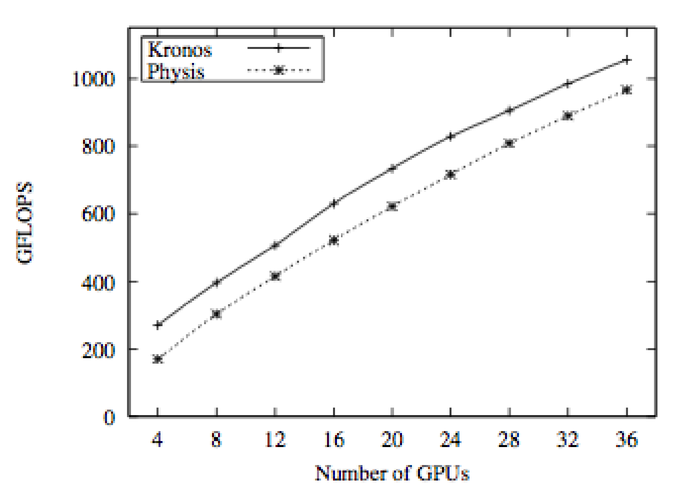
- Define problem using a **domain-specific language**: EDL (Equation Description Language)
- Compile this specification into a C/C++/CaKernel source code module Cactus thorn) – **Kranc**
- Launching the same kernel definition in different environment (CUDA, OpenCL, CPU, OpenMP) by just changing templates - **CaKernel**
- **Cactus** code compiled and run on laptop/ workstation/ **supercomputer /heterogeneous environment**
- Very short time to solution



Chemora

- Simple CFD problem
- Chemora vs. Physis
- GPU for computations
- CPU for data transfer
- Single precision

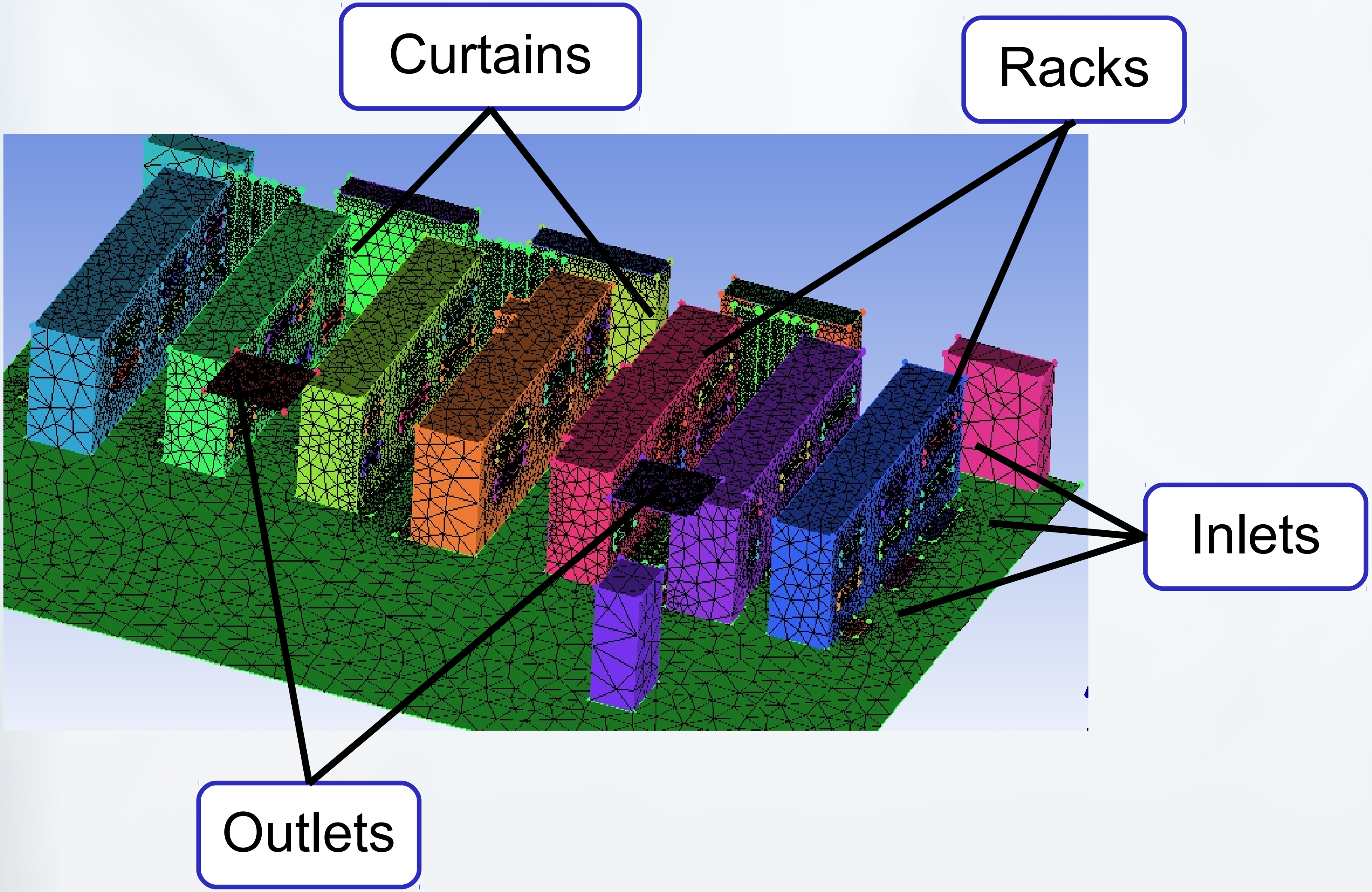




POZNAŃ SUPERCOMPUTING AND NETWORKING CENTER

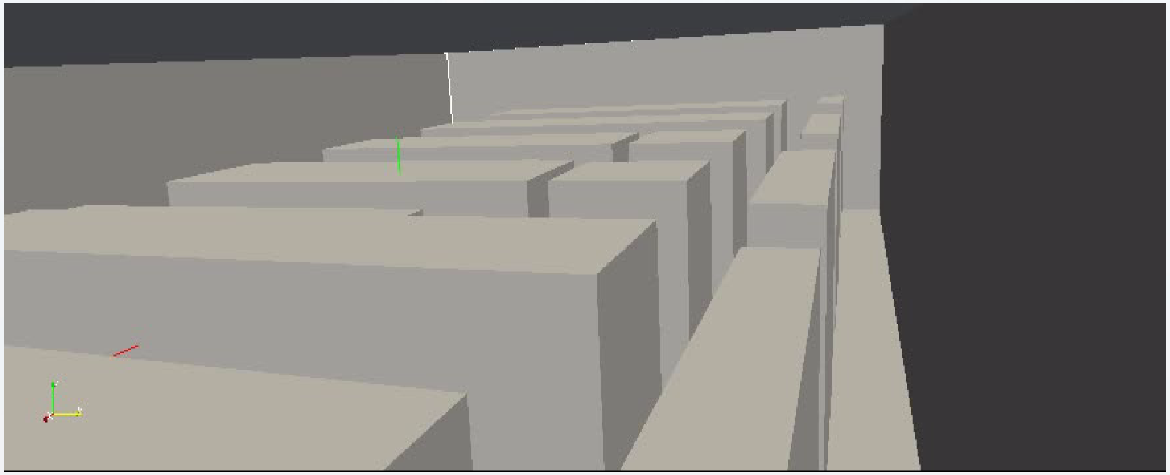
Eulag activities

Airflow and heat transfer in data center (1)

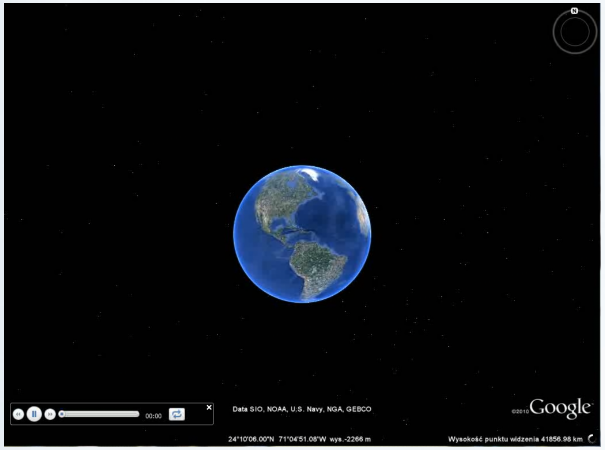


Airflow and heat transfer in data center (2)

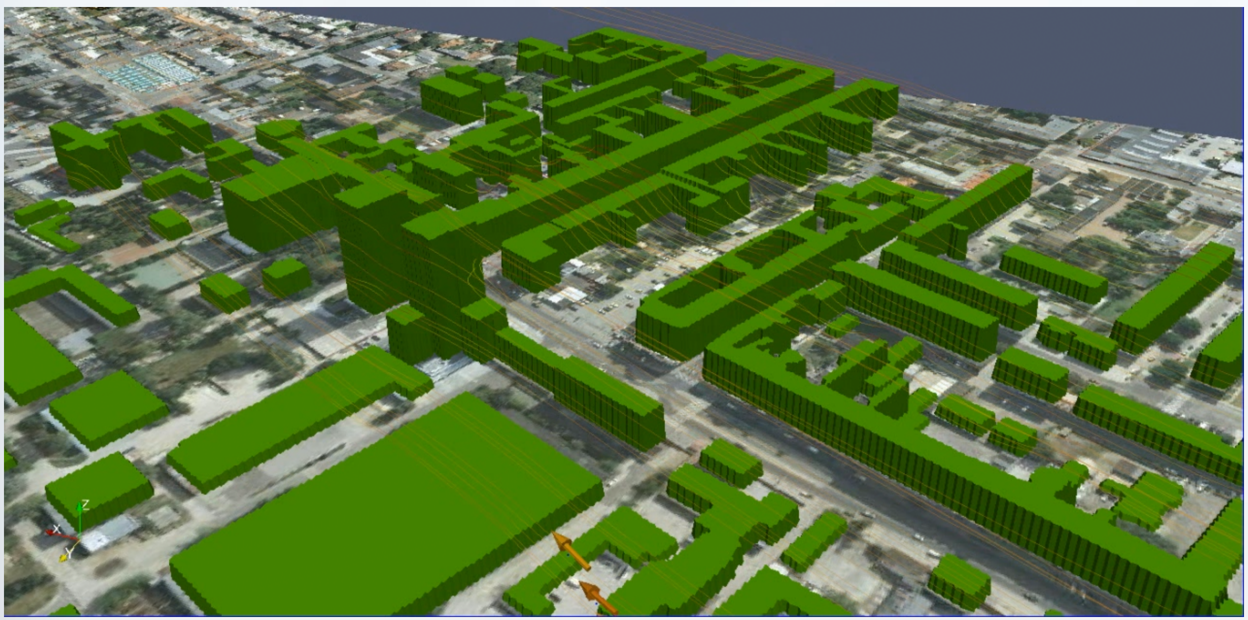
- Realistic wind flow
- Heat transfer
- Temperature
- Power consumption optimization
- Racks placement optimization



Poznan urban flow



- Simulation of ideal and realistic wind flow
- Contamination & dispersion of hazardous substances



Future activities

Motivation
- Code adoption to new architectures
- Speedup the model even more

Focus
- Elliptic solver and preconditioner

Architectures
- GPGPU cluster
- ccNUMA SMP

Approaches
- GPU (CUDA, OpenCL, Accelerators)
- OpenMP + MPI + NUMA awereness
- Load balancing, thread scheduler, framework (CaKernel), …

Backup slides

CaKernel Optimizations

- Dynamic tile size selection (JIT compilation)
  - The stencil size
  - The number of grid variables
  - The shape of the local grid
- Lightweight kernel generation
  - Program parameters and tile shape seen as constants by compiler
  - Fixed load offset
- Fat kernel detection
- Profiling information
  - NVIDIA Cupta API

POZNAŃ SUPERCOMPUTING AND NETWORKING CENTER

Thank you

Poznań Supercomputing and Networking Center
affiliated to the Institute of Bioorganic Chemistry of the Polish Academy of Sciences,
ul. Noskowskiego 12/14, 61-704 Poznań, POLAND,
Office: phone center: (+48 61) 858-20-00,
fax: (+48 61) 852-59-54,
e-mail: office@man.poznan.pl, http://www.man.poznan.pl