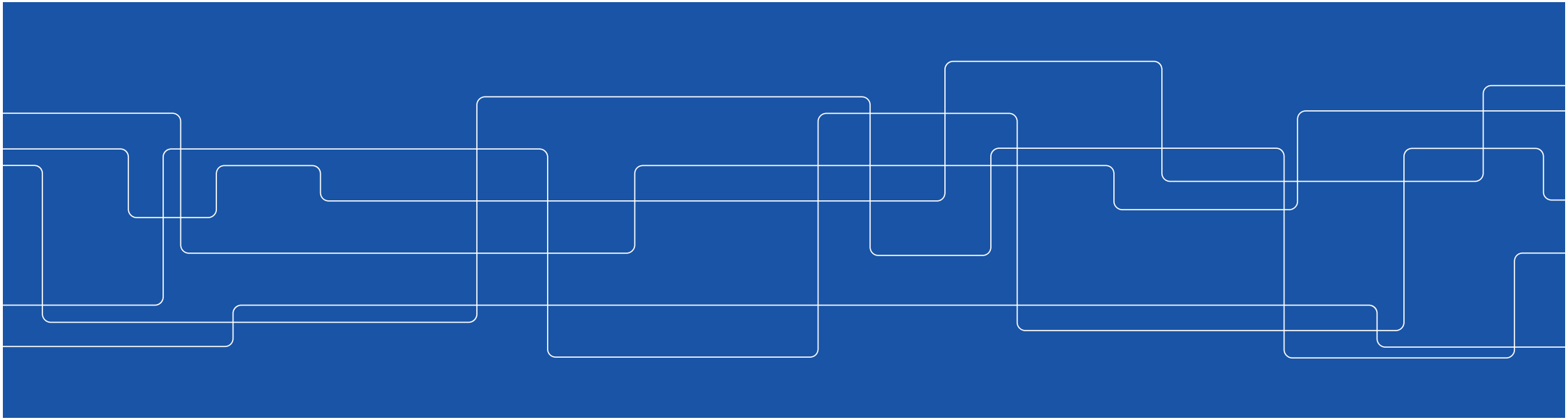
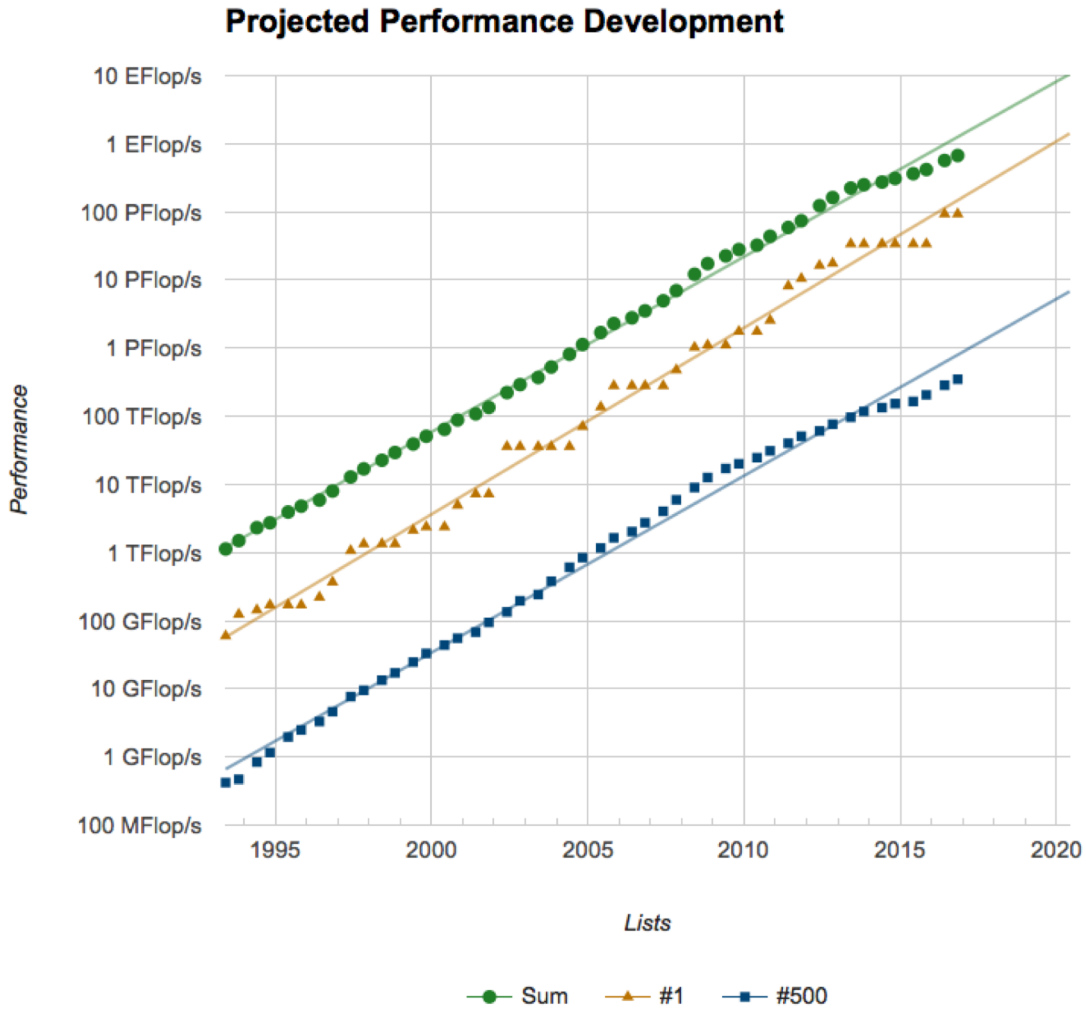

# Preparing Applications for the New Era of Computing

Erwin Laure

Department for Computational Science and Technology &
PDC Center for High Performance Computing, KTH



# We are approaching the Exascale Era




# Why do we Care?

• Exaflop performance is required for specific, important problems

• Yet, most science and engineering is done (and will continue to be done) on the Tera- and Petascale, sometimes even Gigascale

• So, why do we care?

• The quest for exascale is cruelly exposing key problems of our current approach towards computing

- The new technologies will be pervasive

# The End of Historic Scaling

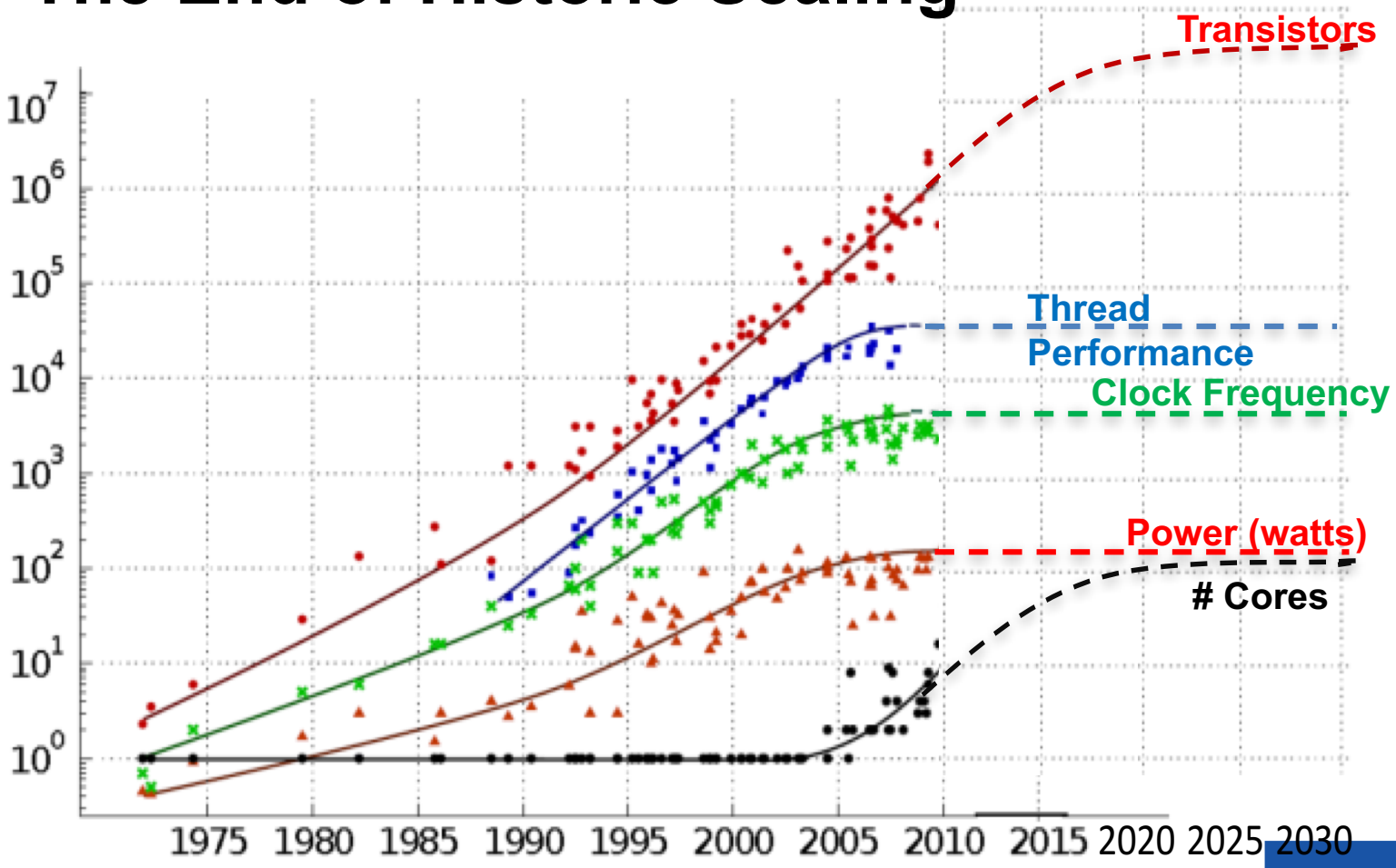




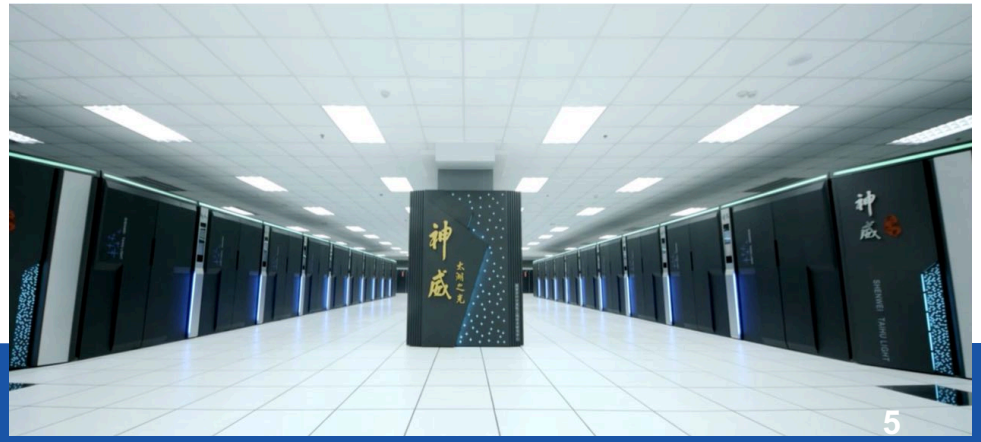

Figure courtesy of Kunle Olukotun, Lance Hammond, Herb Sutter, and Burton Smith

# The Power Problem

TaihuLight: 100 PF@15 MW -> 1 EF@150 MW

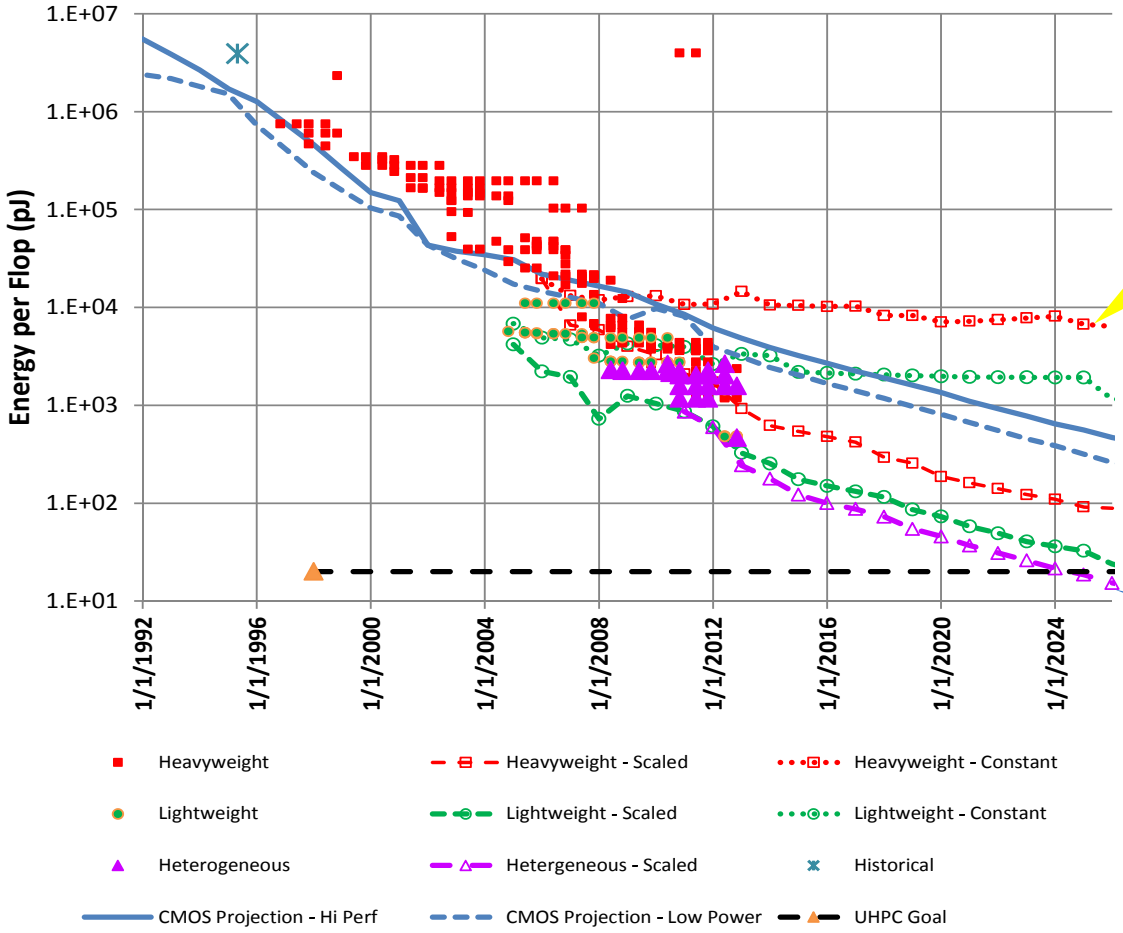
Beskow: 1.3 PF@600kW -> 1EF@461 MW

EGI: 850.000 cores - ~20 PF, ~10 MW

COMPUTER ARCHITECTURE LABORATORY
EXASCALE DESIGN SPACE EXPLORATION

# But Mere Multi-Core is NOT good enough! *(need to go to simpler cores)*




**Can continue with conventional x86 architectures if you want.**

Lightweight cores **OR** Hybrid is the only approach that crosses the exascale finish line

Slide curtesy John Shalf

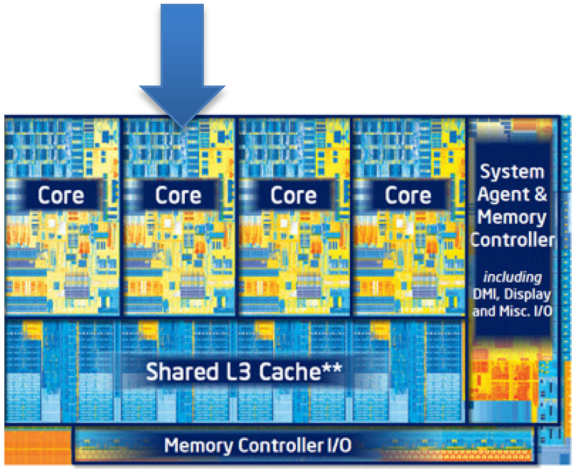


BERKELEY LAB

U.S. DEPARTMENT OF ENERGY | Office of Science

# Heterogeneous Future (LOCs and TOCs)

Big cores (very few)

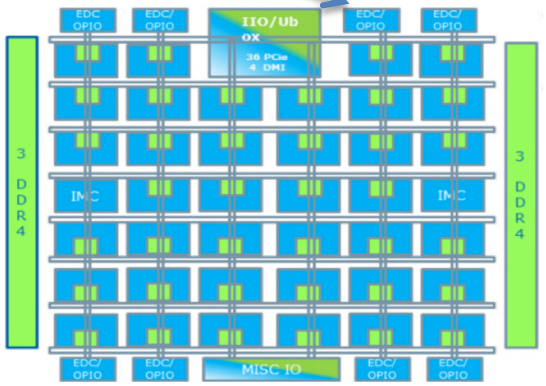



Tiny core

0.23mm

0.2 mm

Lots of them!



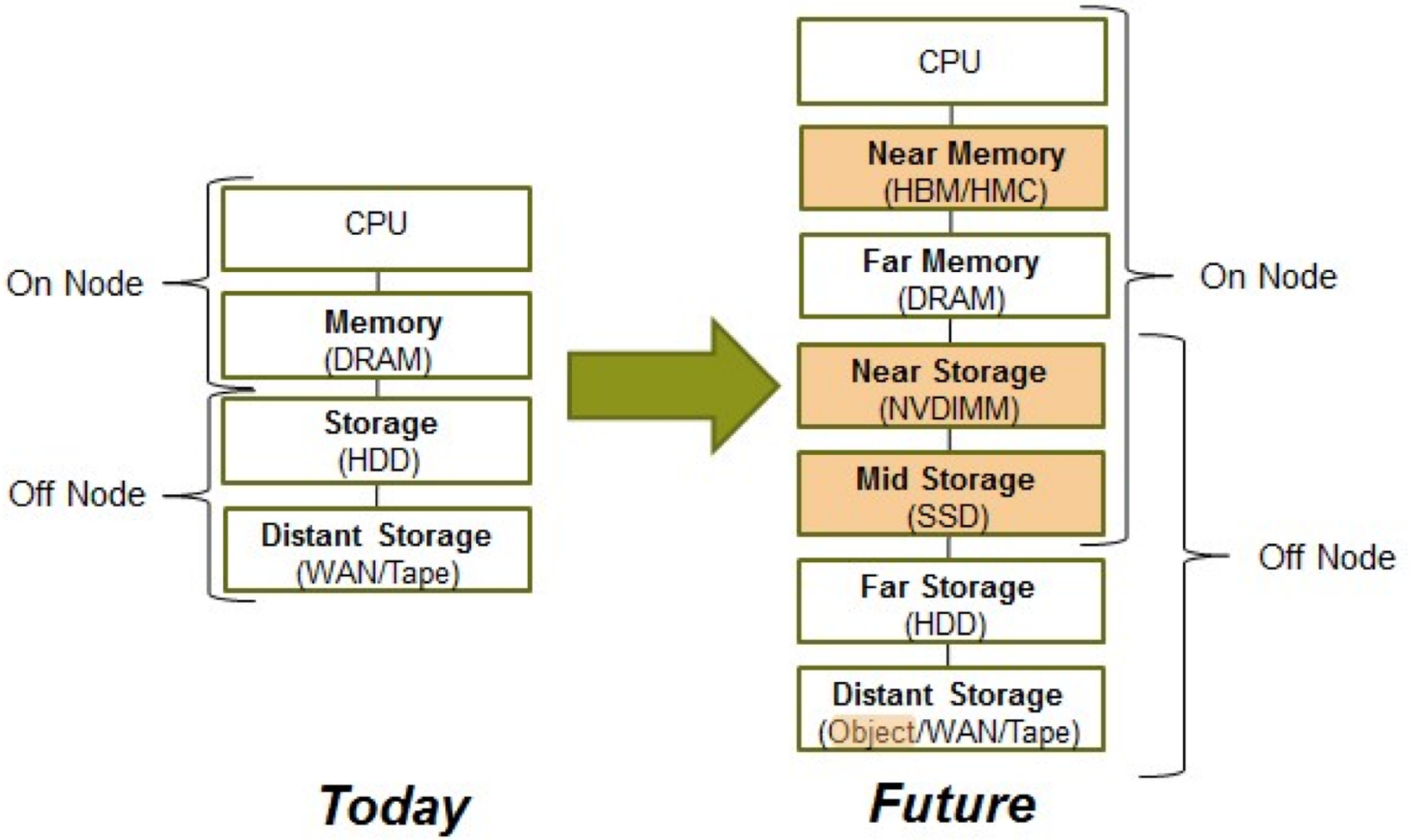

## Latency Optimized Core (LOC)

Most energy efficient if you don't have lots of parallelism

## Throughput Optimized Core (TOC)

Most energy efficient if you DO have a lot of parallelism!

# Trends in the Memory/Storage Subsystem



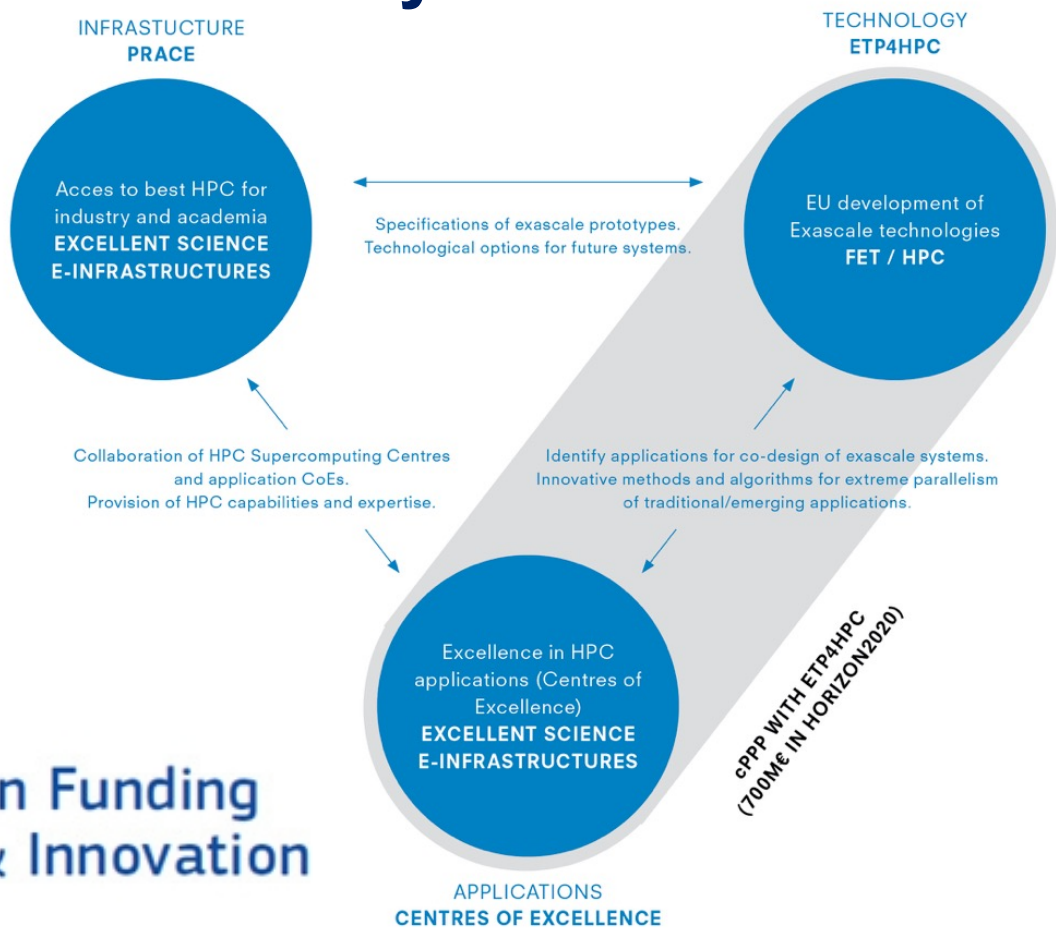

# The New Computing World

- Increased performance only through parallelism
- Data locality more important than reduction of operations
- Computers are increasingly dynamic
- Faults may become an increasing issue – or not


# What does this mean for Applications

- Cannot simply wait for the next generation of CPUs
- Parallelism needs to be treated as a first class problem
  - Not as “implementation detail”
- Algorithms may need to be revised to exploit massive parallelism and reduce data movement
- Implementation/Programming models need to be revised
- Co-Design
  - Collaboration of hardware, algorithms, programming environments, applications

# Ecosystem: cPPP




Horizon 2020
European Union Funding
for Research & Innovation

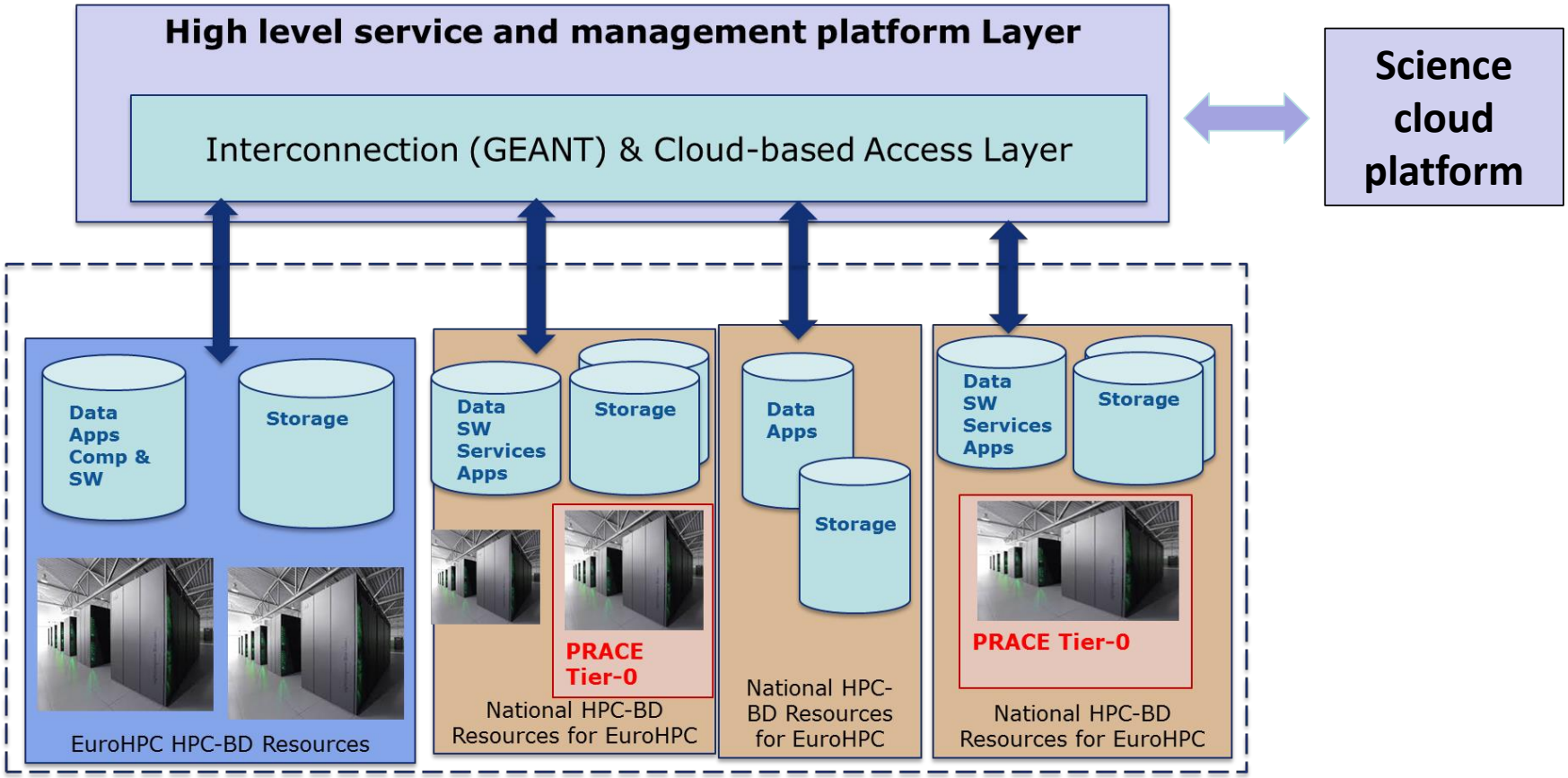
# The new European HPC research landscape



# EuroHPC JU: Overall activities


- **Infrastructure & Operations**
  - Acquisition of infrastructure (linked to Research and Innovation)
  - Installation, deployment and operation via hosting entities
  - providing and managing access to users
- **R&I, Applications & Skills**
  - Supporting technologies and systems developed in Europe
  - Excellence in HPC applications; Supporting Industry (incl. SMEs); Training and Outreach

# The European Data Infrastructure Implementation (vision)




# Centers of Excellence (CoE)

- Improve important applications towards the Exascale
- Provide training and support
- Cover full workflow (including data handling)

# 1st Generation of CoE


**EoCoE** - Energy oriented Centre of Excellence for computer

**BioExcel** - Centre of Excellence for Biomolecular Research


**NoMaD** - The Novel Materials Discovery Laboratory


**MaX** - Materials design at the eXascale



**ESiWACE** - Excellence in SImulation of Weather and Climate in Europe

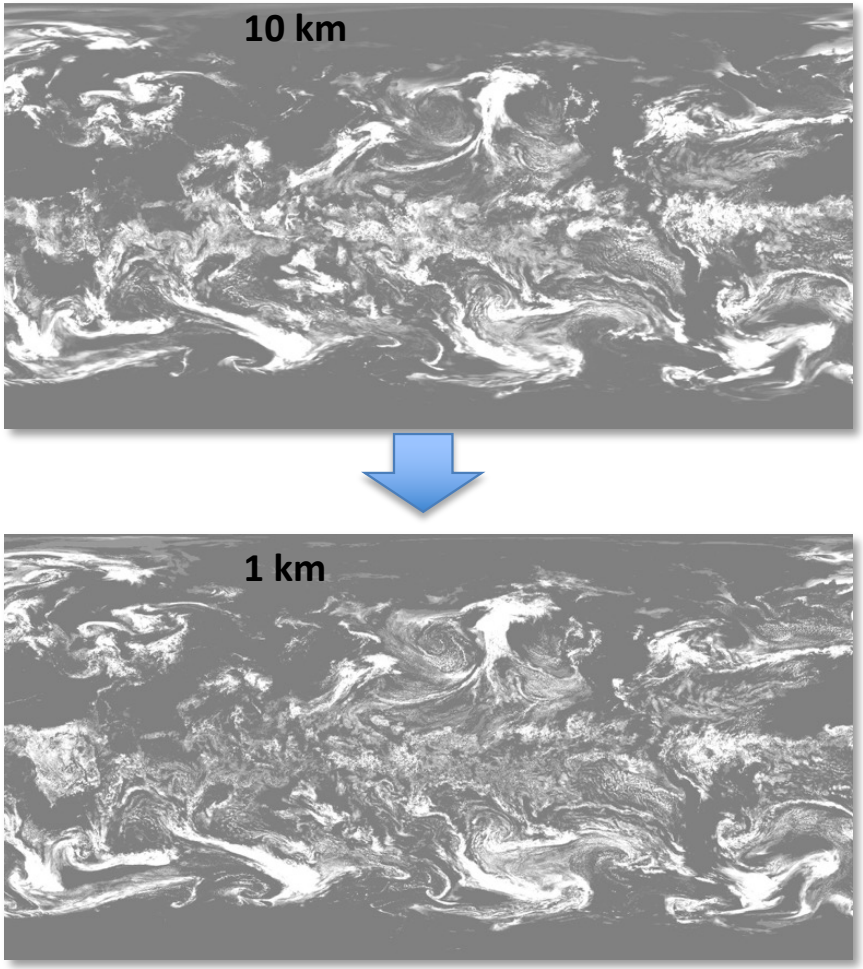

**E-CAM** - An e-infrastructure for software, training and consultancy in simulation and modelling


**POP** - Performance Optimisation and Productivity


**COEGSS** - Center of Excellence for Global Systems Science

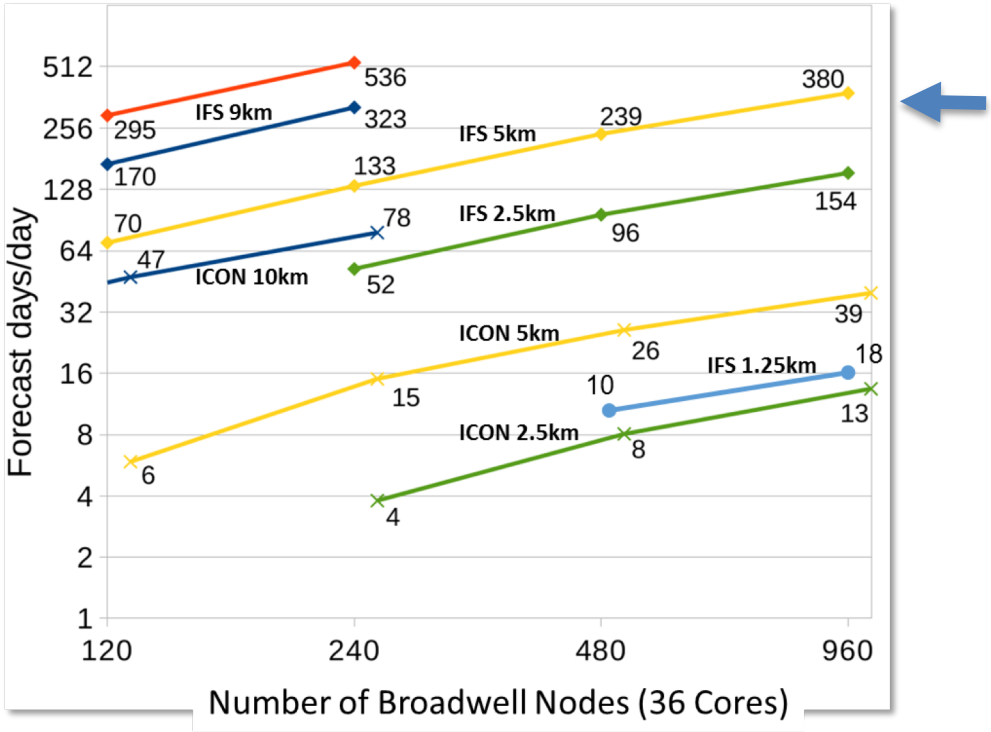

**CompBioMed** - A Centre of Excellence in Computational Biomedicine

# Example: ESiWACE Science Challenge



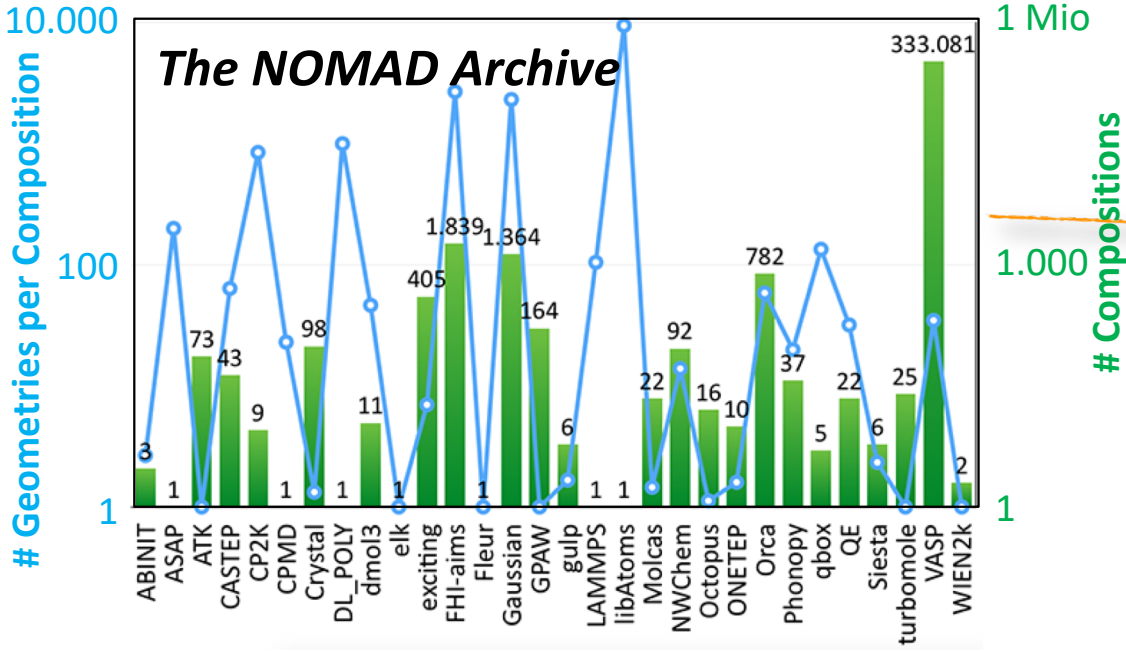

**Target for addressing key science challenges in weather & climate prediction:**

Global 1-km Earth system simulations @ ~1 year / day rate



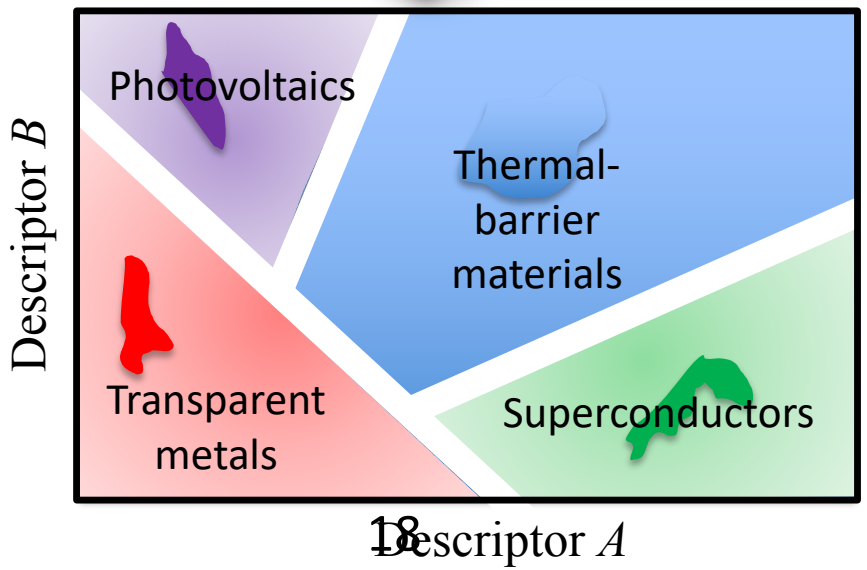
# Example: NOMAD Science and Data Handling Challenges

**Data is the raw materials of the 21st century**




NOMAD supports ***all*** important codes in computational materials science. The code-independent Archive contains data from ***many million calculations*** (billions of CPU hours).

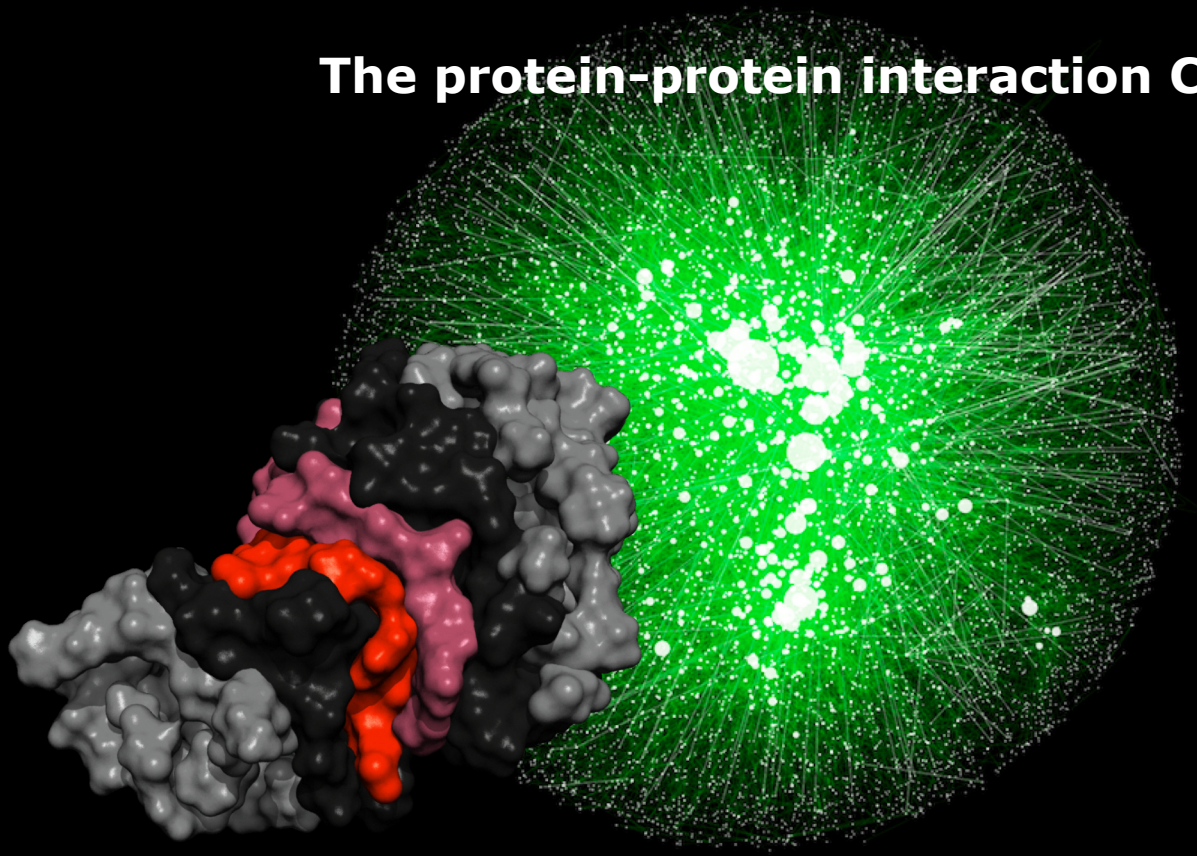
***The NOMAD challenge: Build a map and fill the existing white spots***




***Discovering interpretable patterns and correlations*** in this data will

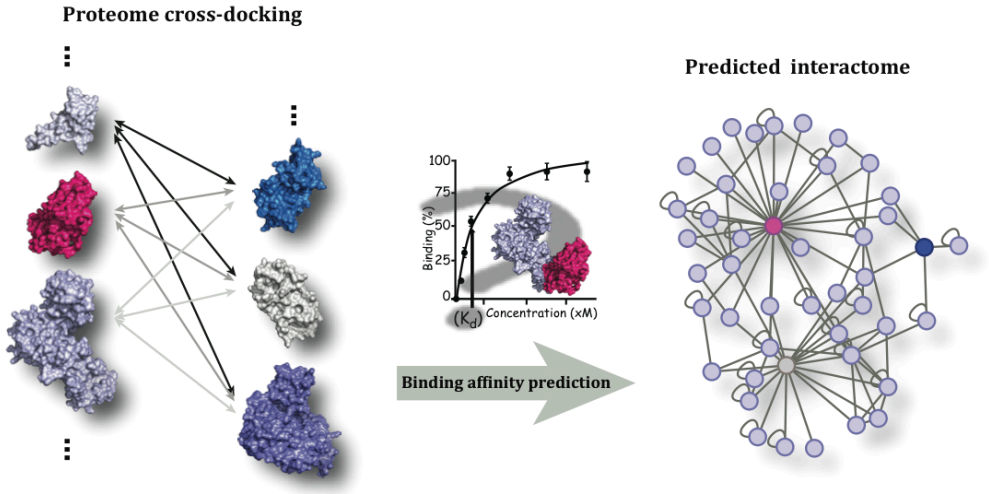
- create knowledge
- advance materials science,
- identify new scientific phenomena, and
- support industrial applications.

# The protein-protein interaction Cosmos



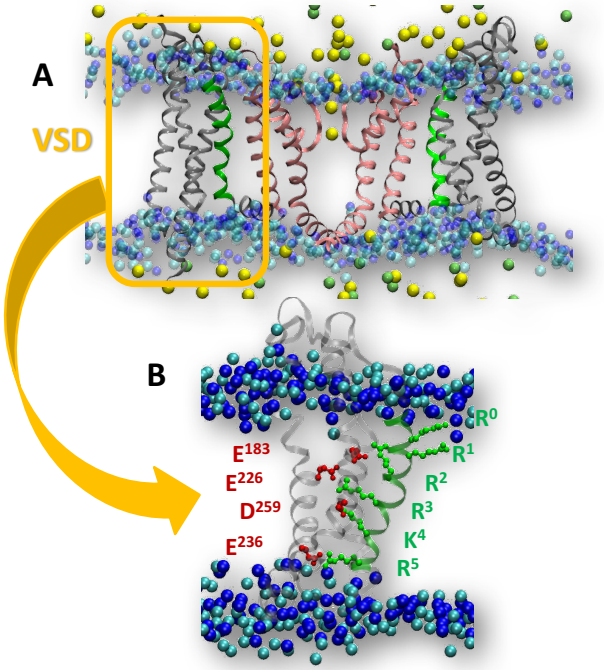
- ~20’000 human proteins
- 400’000 interactions
- Adding the 3D structural dimension to those will require > 10 million CPU hours and generate exabytes of data

# Predicting interactomes by docking… a dream?



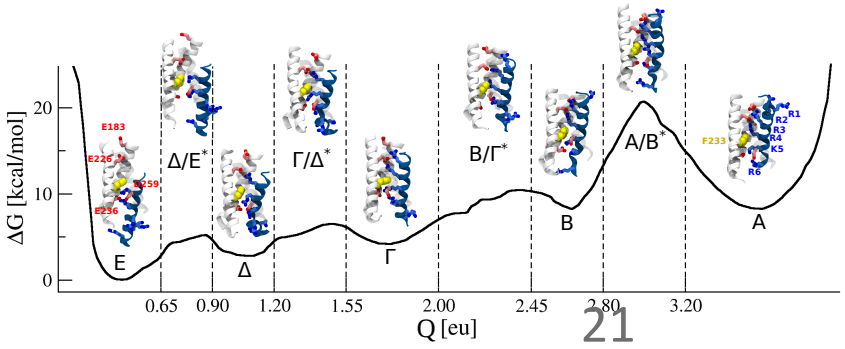

- ~20’000 human proteins
- Interactome prediction will require 20’000$^2$ docking runs
- Which will require > 10 billions CPU hours and generate about 100 exabytes of data
- Interest in simulating/understanding the impact of disease-related mutations that affect/alter the interaction network

# Molecular Dynamics on the exascale




- Understanding proteins and drugs
- A 1 μs simulation: 10 exaflop
- Many structural transition: many simulations needed
- Study effect of several bound drugs
- Study effect of mutations
- All this multiplies to >> zettaflop
- Question: how far can we parallelize?

Example: ion channel in a nerve cell.
Opens and closes during signalling.
Affected by e.g. alcohol and drugs.
200 000 atoms



# Top 3 Challenges

- HPC System Architecture and Components
  - Efficient use of memory and I/O hierarchies - Balance Compute, I/O and Storage Performance
  - Efficient interaction between “fat” and “thin” (GPU) cores
- System Software and Management
  - Software standards (C++17 and Fortran 2015 in particular, but also OpenMP 4.5, MPI 3.1, OpenCL 2.2,...)
- Programming Environment
  - (Dynamic) environments for task parallelism.

# BioExcel

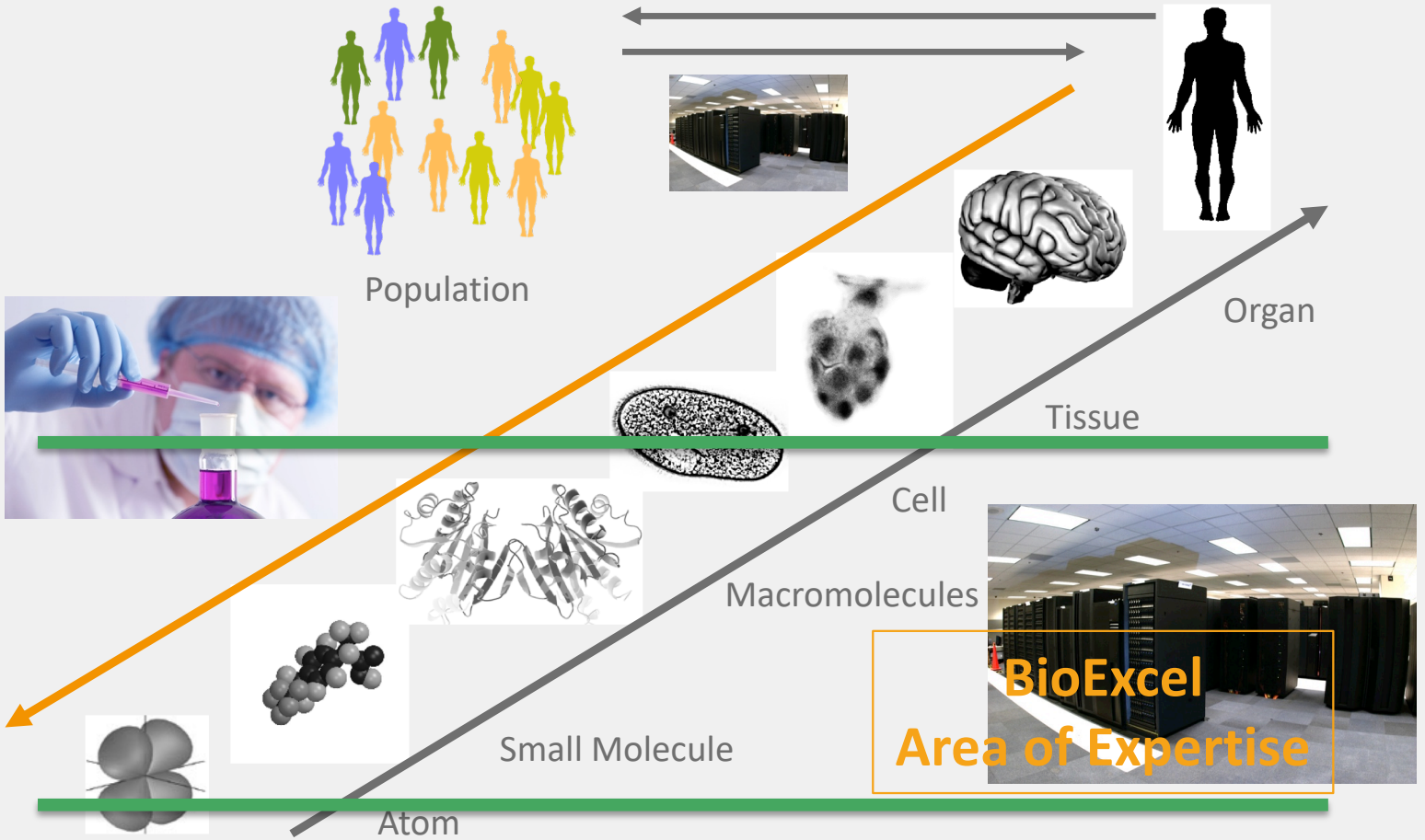
## Center of Excellence for Computational Biomolecular Research

Partners

Funding

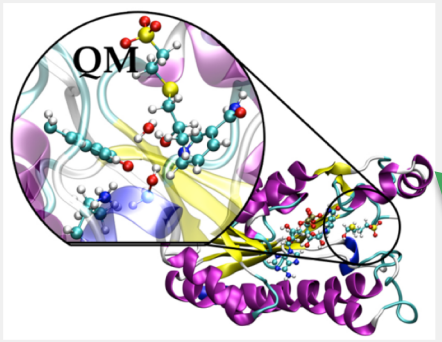
Horizon 2020
European Union Funding
for Research & Innovation

# Life Science and HPC

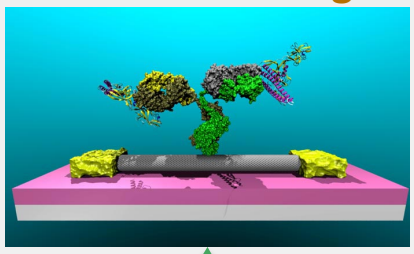


# Biomolecular Modeling and Simulations

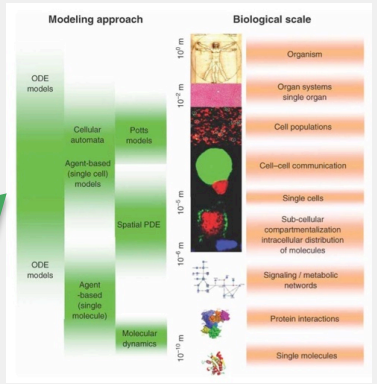
## Electronic structure



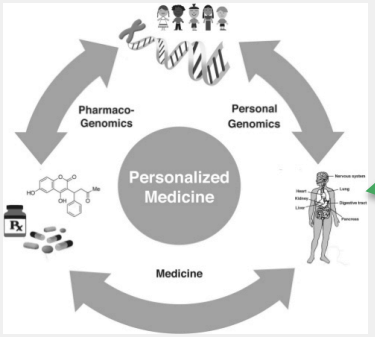

## Biomarkers design



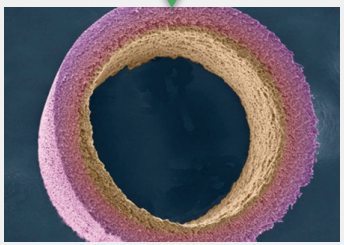
## Physiology



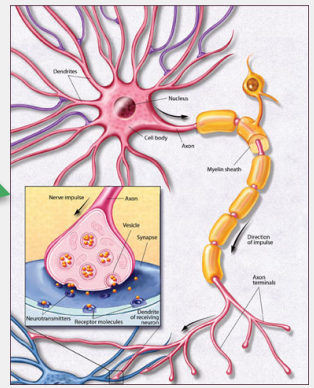

## Personalized medicine



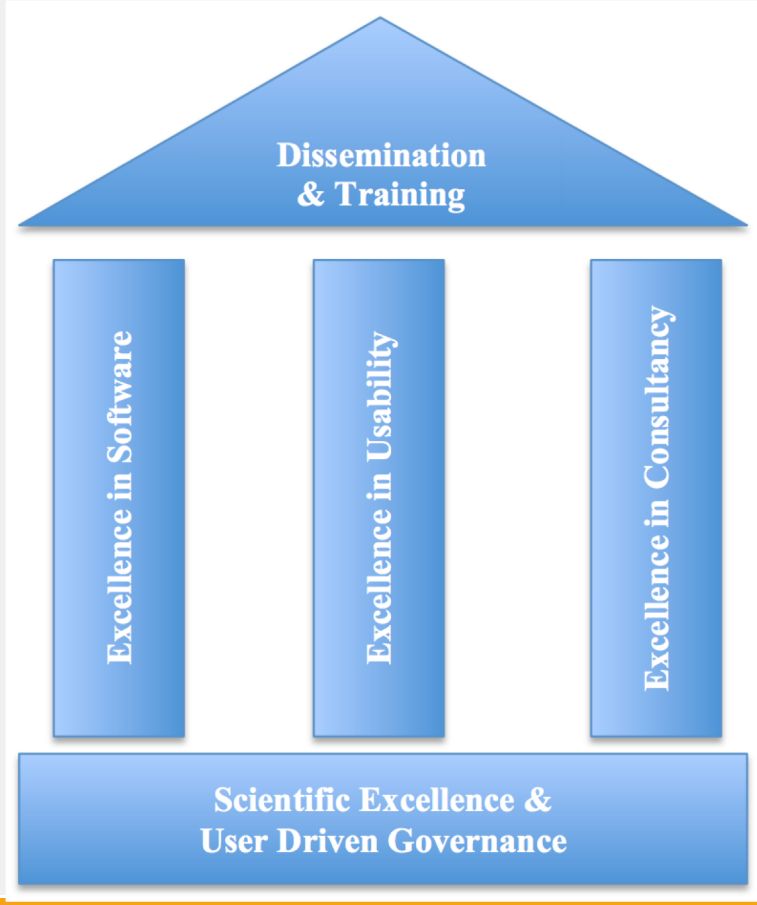

## Biomaterials science and nanotechnology



## Neuroinformatics



# BioExcel Pillars of Excellence

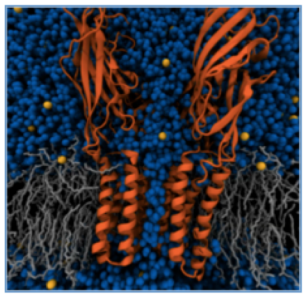


# Objectives of BioExcel

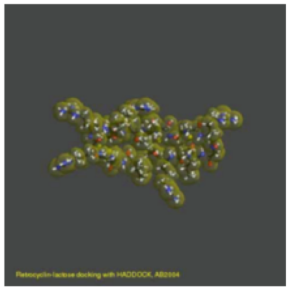
## Excellence in Biomolecular Software

Improve the performance, efficiency, scalability, and maintainability of key codes

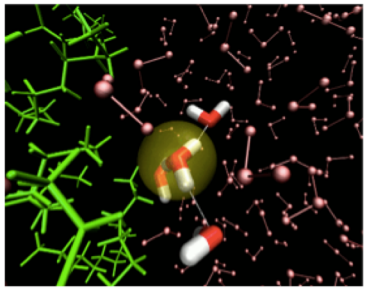
- GROMACS (Molecular Dynamics Simulations)
- HADDOCK (Integrative modeling of macro-assemblies)
- CPMD (hybrid QM/MM code for enzymatic reactions, photochemistry and electron transfer processes)



MD simulations
/GROMACS/

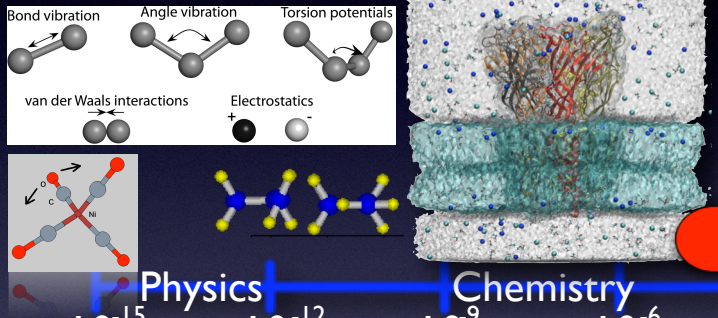


Docking
/HADDOCK/

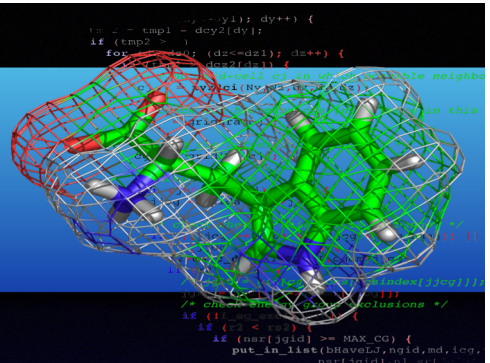


QM/MM
/CPMD/

# Molecular Dynamics





**Experiments**

Efficient averaging

Less detail

Where we need to be

Physics — Chemistry — Biology

$10^{-15}$s $10^{-12}$s $10^{-9}$s $10^{-6}$s $10^{-3}$s $10^{0}$s $10^{3}$s

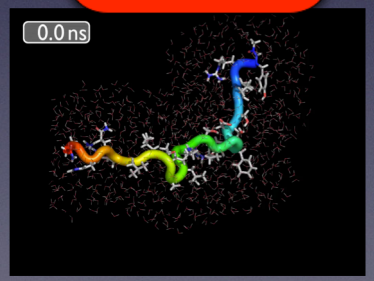
**Simulations**

Extreme detail

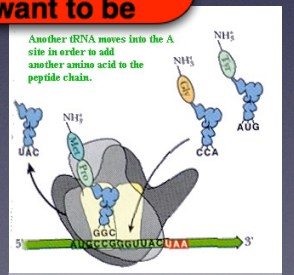
Sampling issues?

Parameter quality?

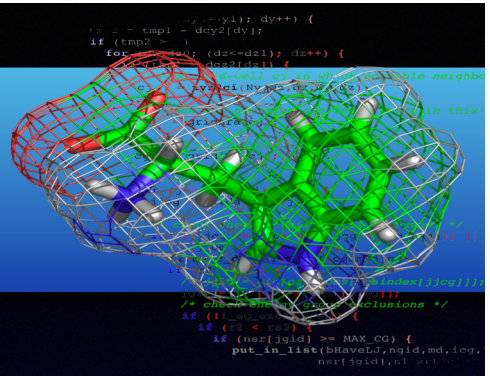
Where we are




Where we want to be



# How do we get to ms range?



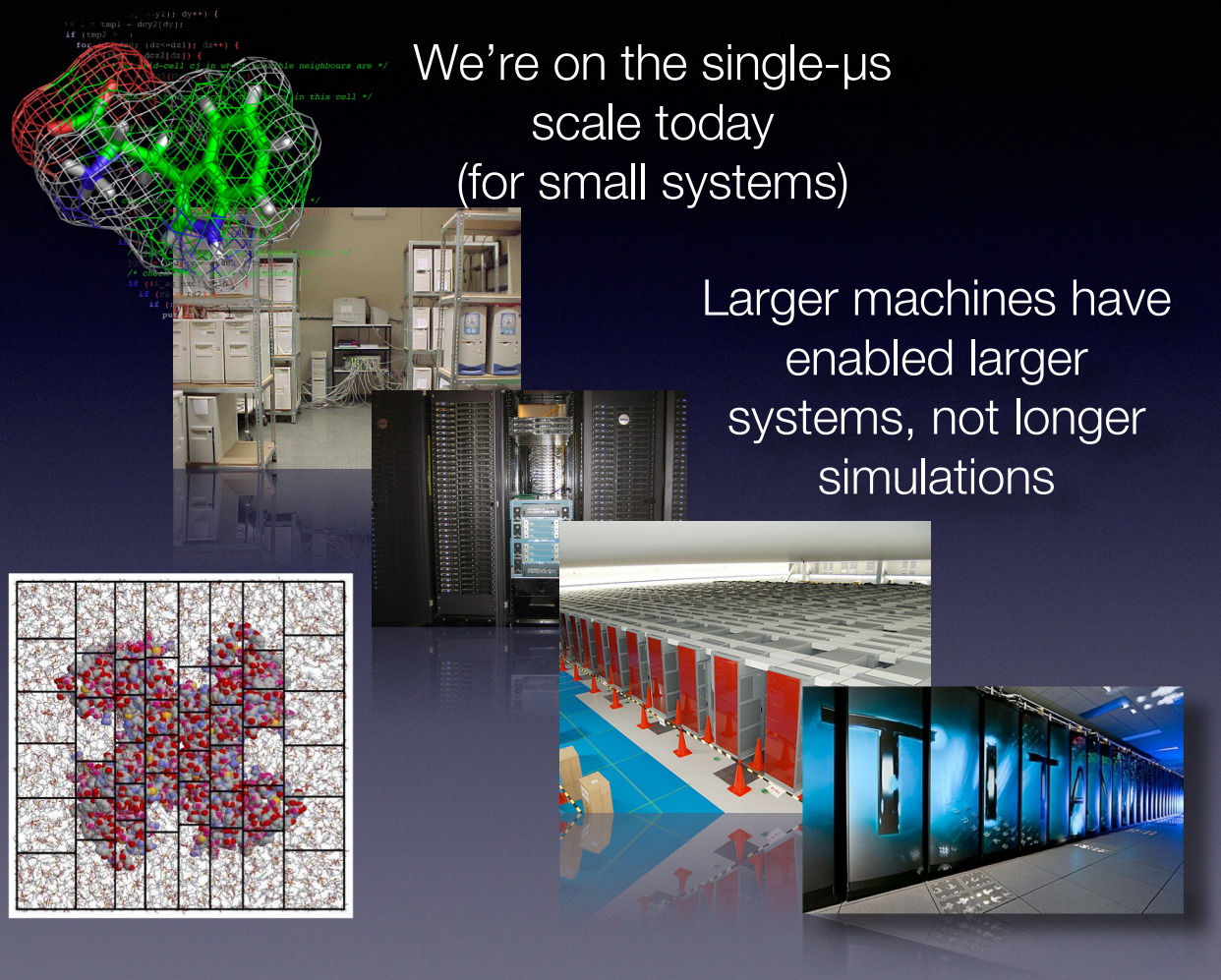
With a time step of 5fs...

... you need 200 billion iterations to reach 1000 μs of simulated time

Let's spend 100 days

For 10 μs per day…

...each iteration must complete in 43.2 (wallclock) μs!



We’re on the single-μs scale today (for small systems)

Larger machines have enabled larger systems, not longer simulations





~2024: 1B 'cores'

2022: ~300M cores

2020: ~100M cores

2018: ~30M cores

2016: ~10M cores

2014: ~3M cores

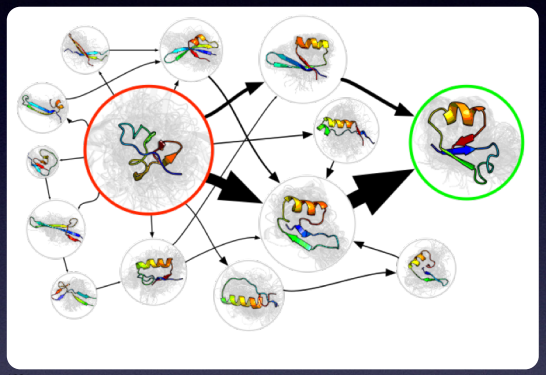
2012: ~1M cores

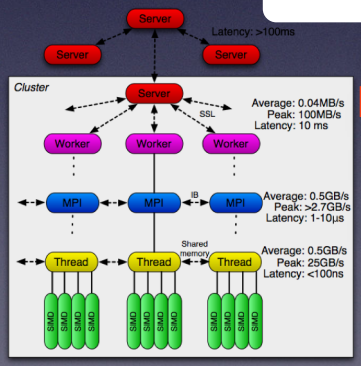
2010: ~300,000 cores

# How will YOU use a billion cores?

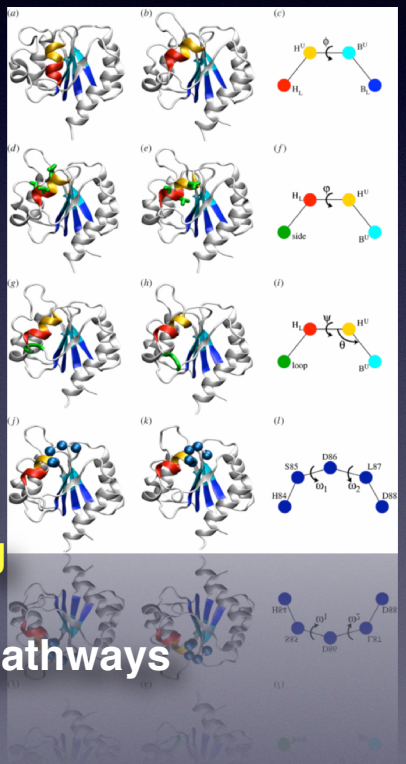
We keep scaling "up" (larger systems) where we should scale "down" (fine-grained parallelism, ensembles)!

# From ~100k cores to Exascale: Ensembles

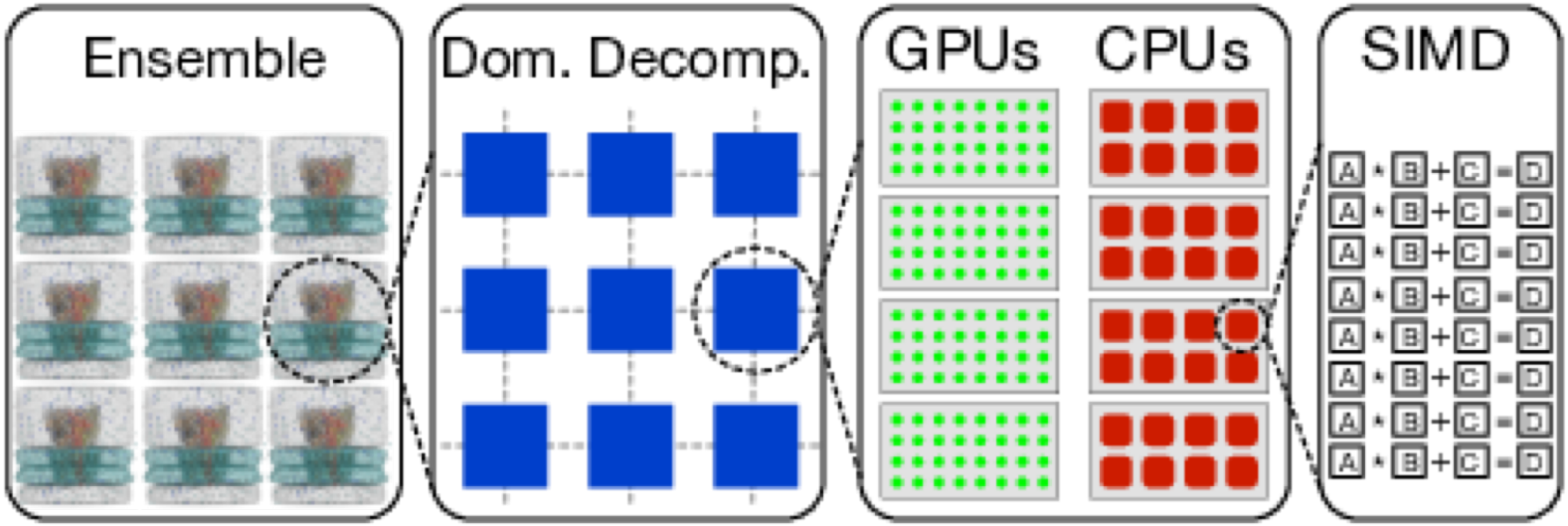




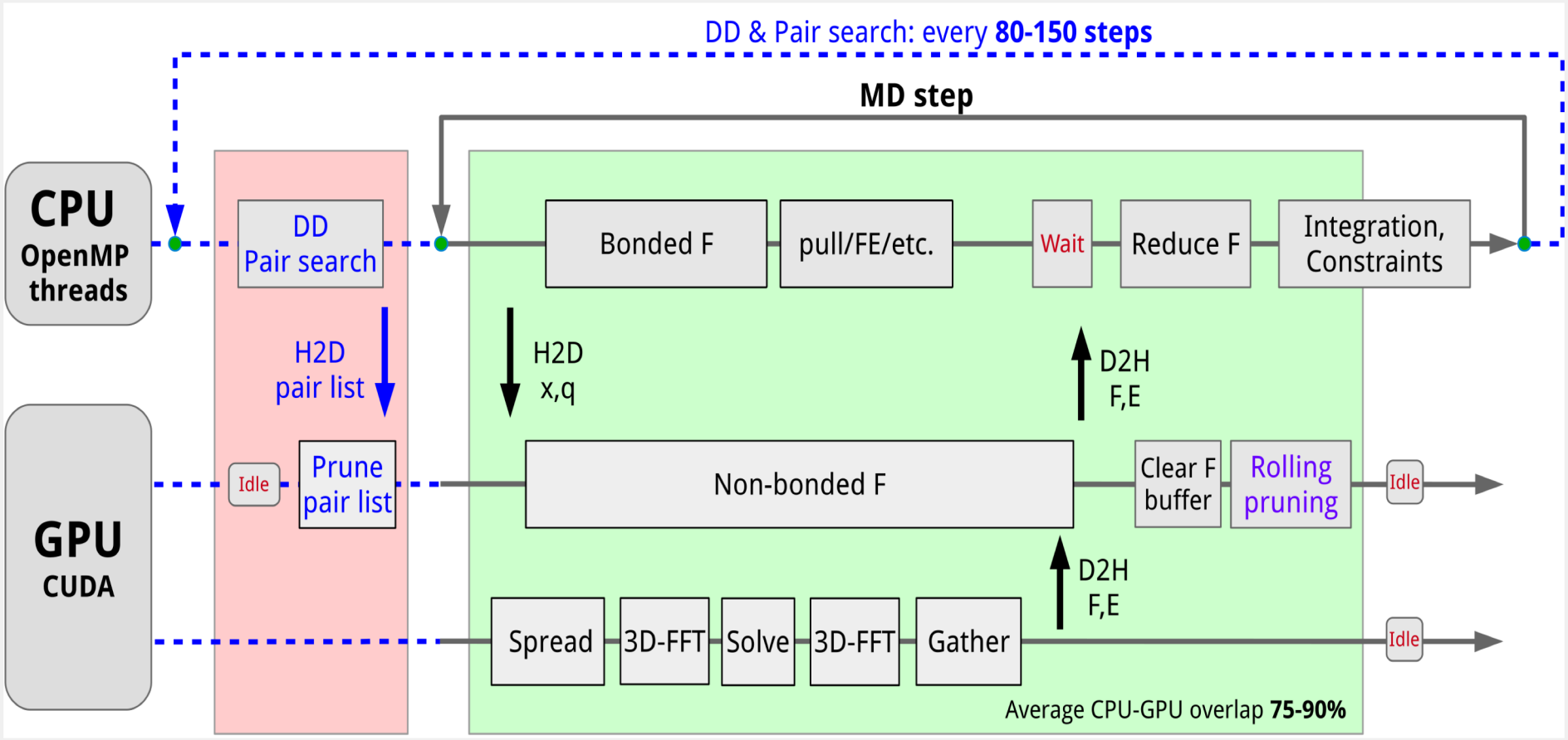

**Milestoning**
**Metadynamics**
**Markov State Models**
**Monte Carlo Sampling**
**Free Energies**
**Swarms / Transition pathways**

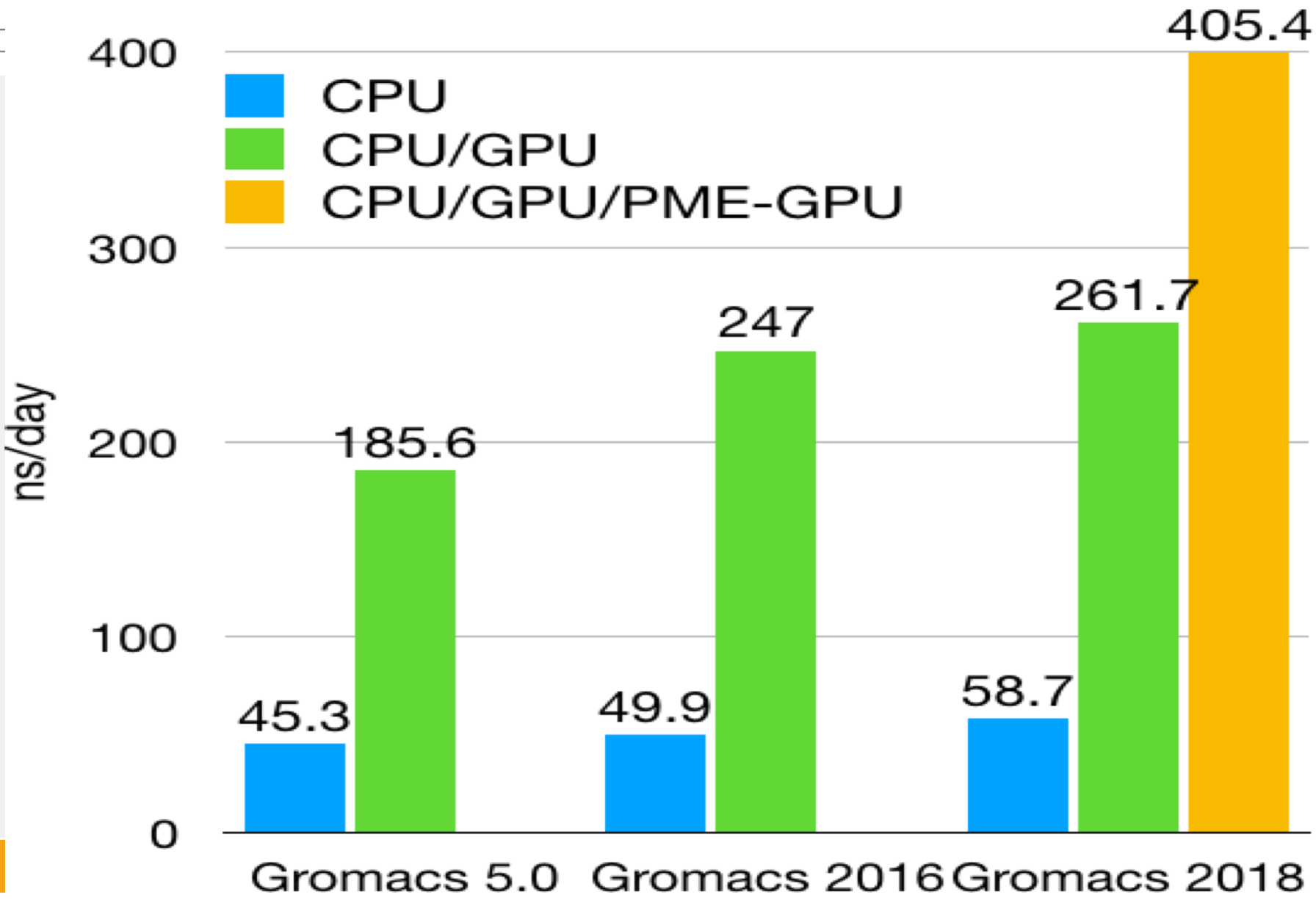


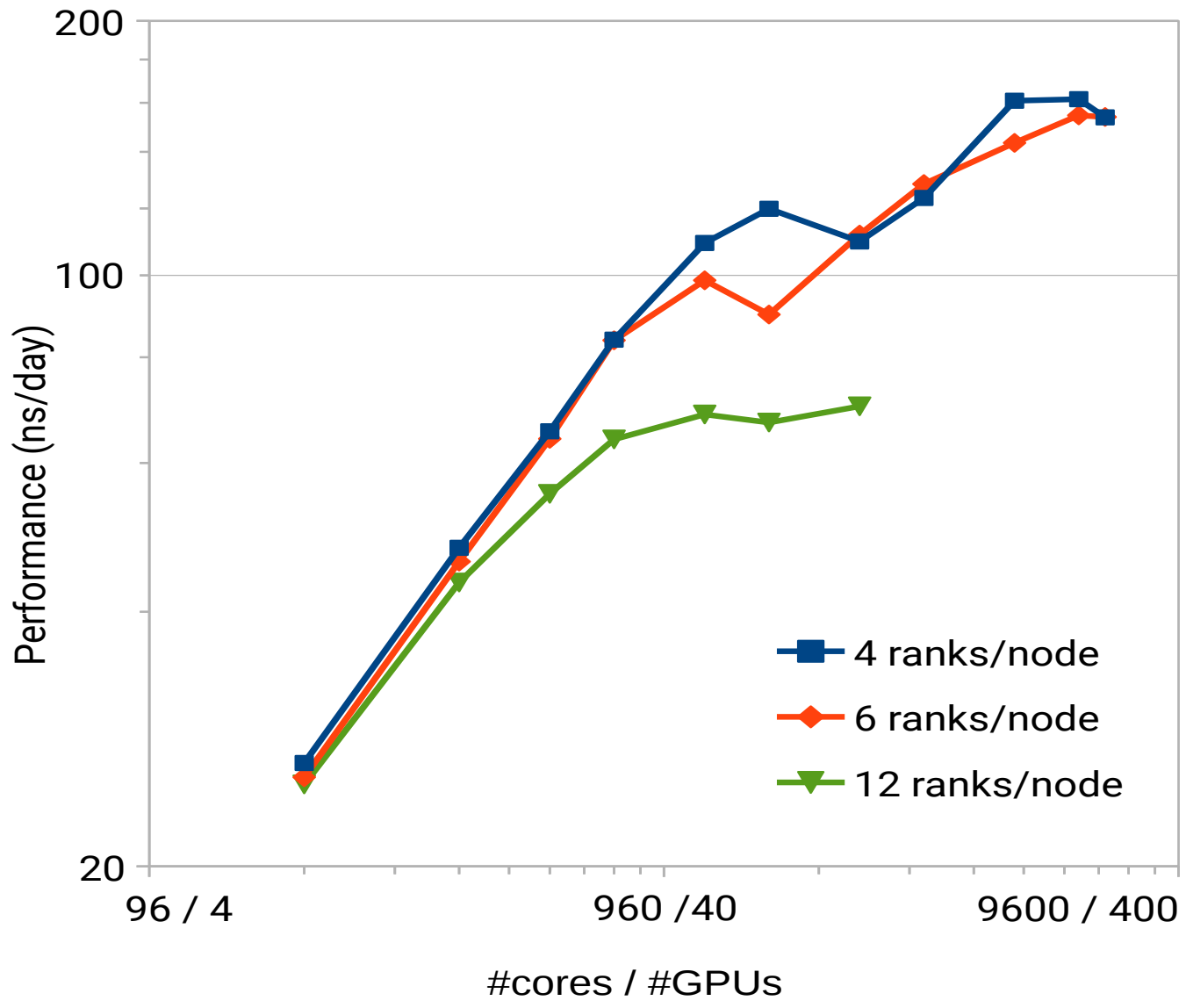
# GROMACS exascale strategy



# GROMACS Heterogeneous Parallelism



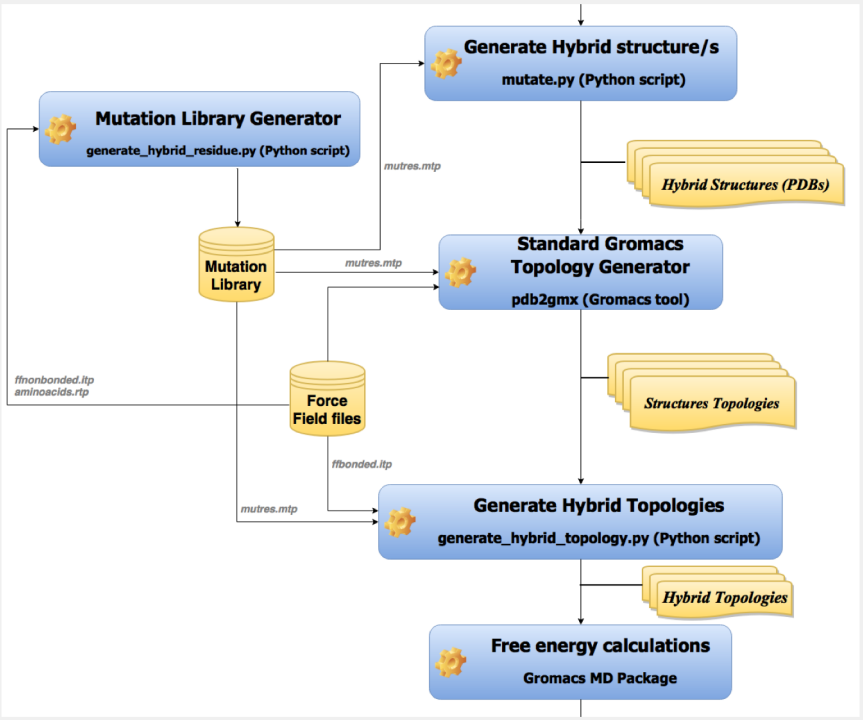

400
300
200
100
0
ns/day
CPU
CPU/GPU
CPU/GPU/PME-GPU
45.3
185.6
49.9
247
58.7
261.7
405.4
Gromacs 5.0
Gromacs 2016
Gromacs 2018


200
100
20
Performance (ns/day)
96 / 4
960 /40
9600 / 400
#cores / #GPUs
4 ranks/node
6 ranks/node
12 ranks/node

# Objectives of BioExcel

## Excellence in Usability

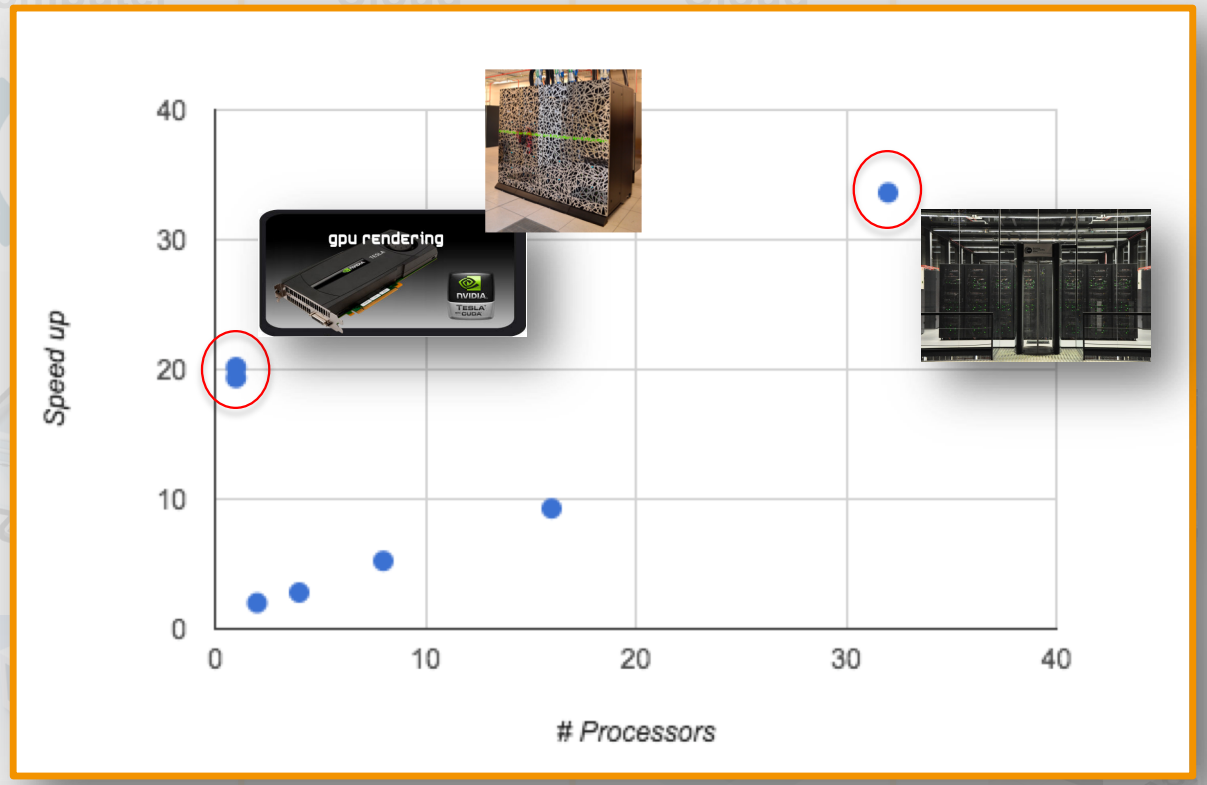
- Make ICT technologies easier to use by biomolecular researchers, both in academia and industry
- Devise efficient workflow environments with associated data integration




**Key Workflows and Platforms**

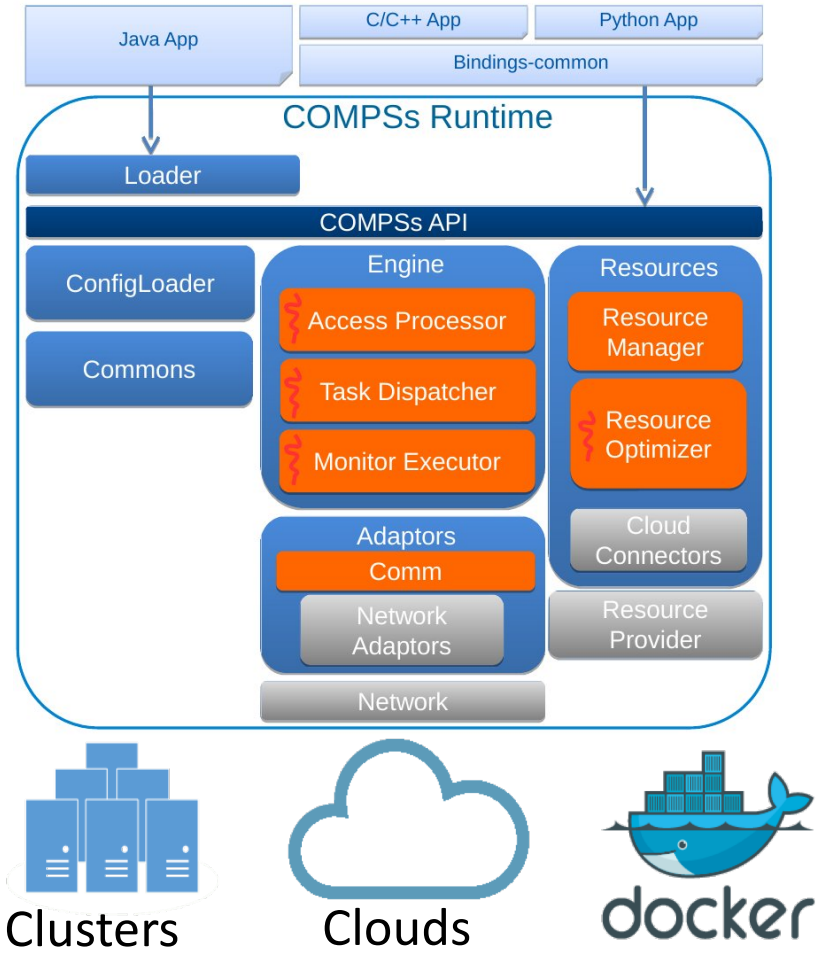
Galaxy PROJECT · Open PHACTS · COMPSs · Open for Innovation KNIME · Apache Taverna

| Personal Computer | Private Cloud | Public Cloud | HPC |
|---|---|---|---|
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |


Personal Computer
Private Cloud
Public Cloud
HPC
gpu rendering
NVIDIA
Speed up
# Processors
0
10
20
30
40

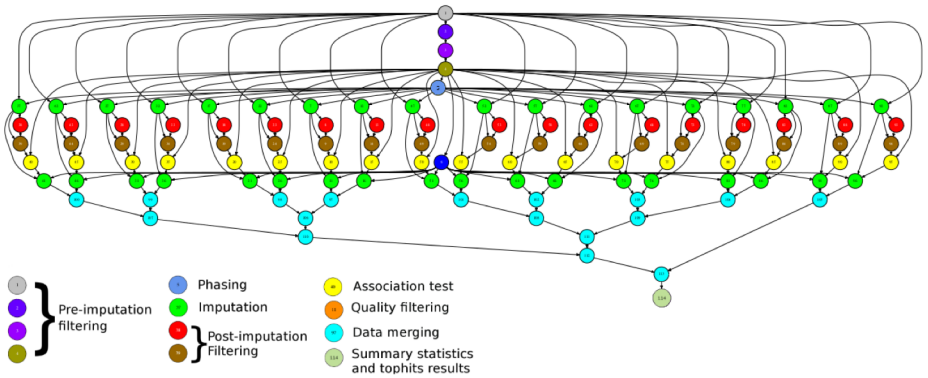
# PyCOMPSs runtime System

- Componentized
- Adaptable
- Extensible
- Interoperable



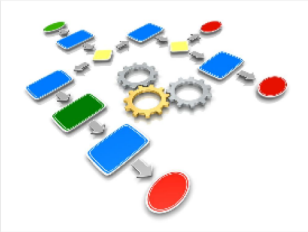
# Programming with PyCOMPSs

- Sequential programming
- General purpose programming language + annotations/hints
  - To identify tasks and directionality of data
- Task based: task is the unit of work
- Simple linear address space
- Builds a task graph at runtime that express potential concurrency
  - Implicit workflow
- Exploitation of parallelism
  - … and of parallelism created later on
- Agnostic of computing platform
  - Enabled by the runtime for clusters, clouds and container managed clusters




Barcelona Supercomputing Center
Centro Nacional de Supercomputación

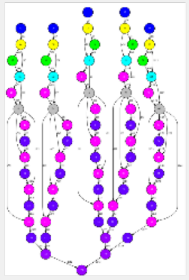
# Workflows: Protein Mutations



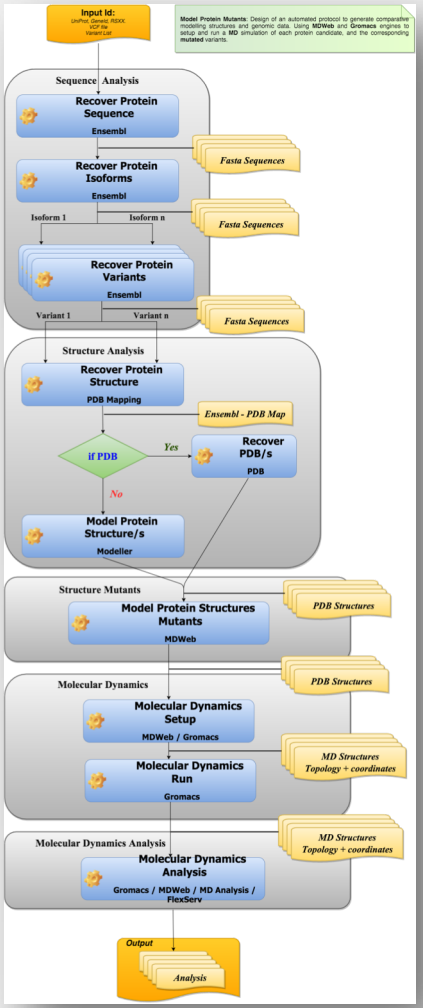
**Workflow**: GROMACS MD Setup

**Building Blocks**:
Gromacs Python wrapper

*Interoperable*
*Workflow manager agnostic*



**Workflow Manager**:
pyCOMPSs

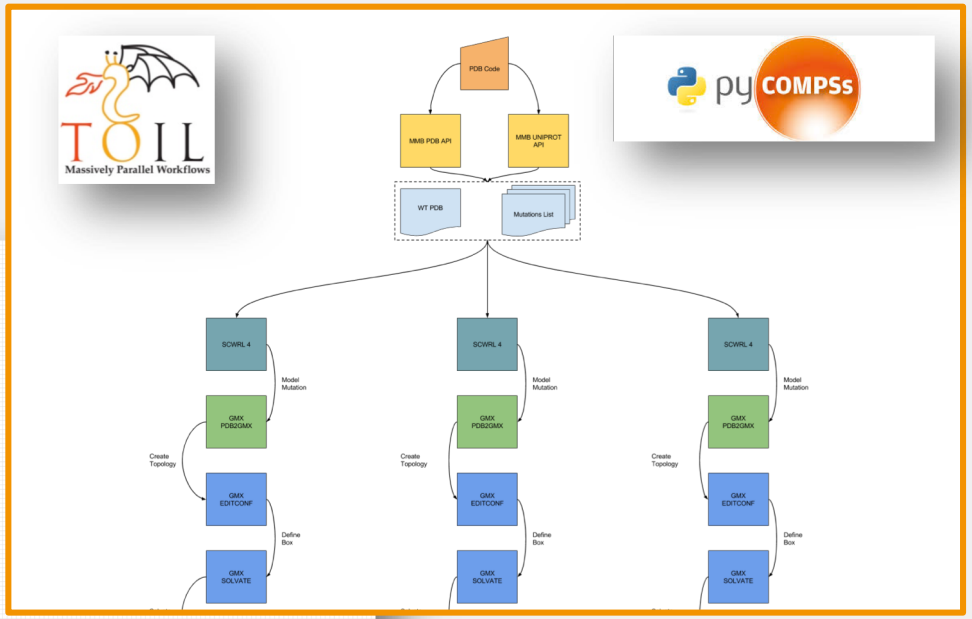


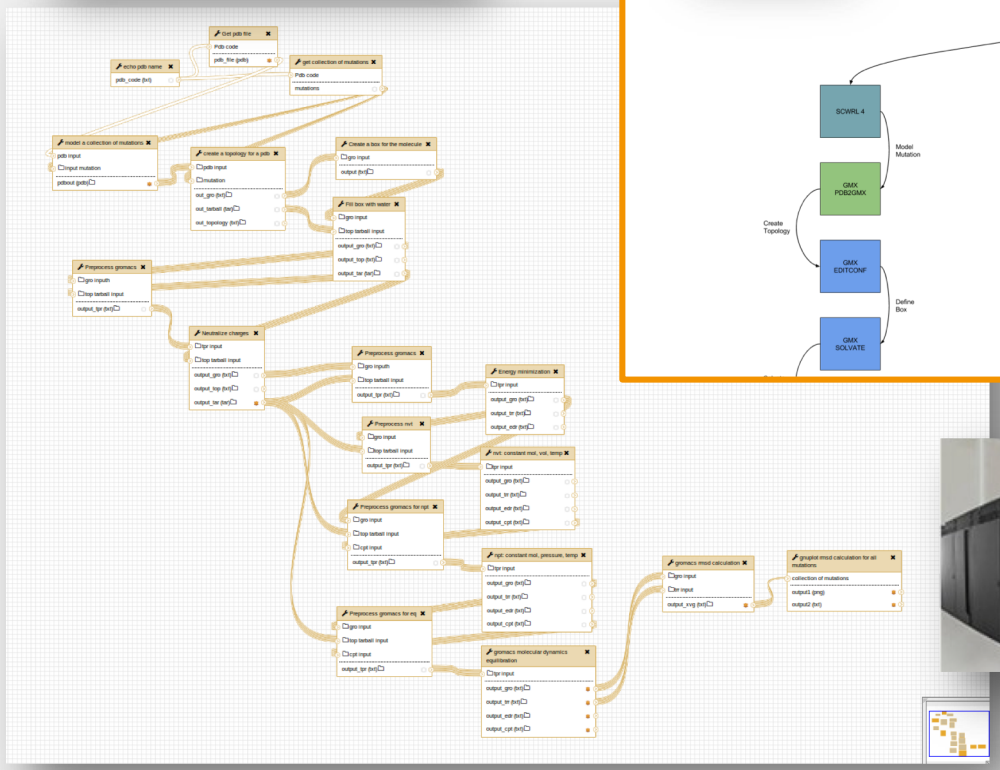
# Workflows: Workflow Managers

Galaxy

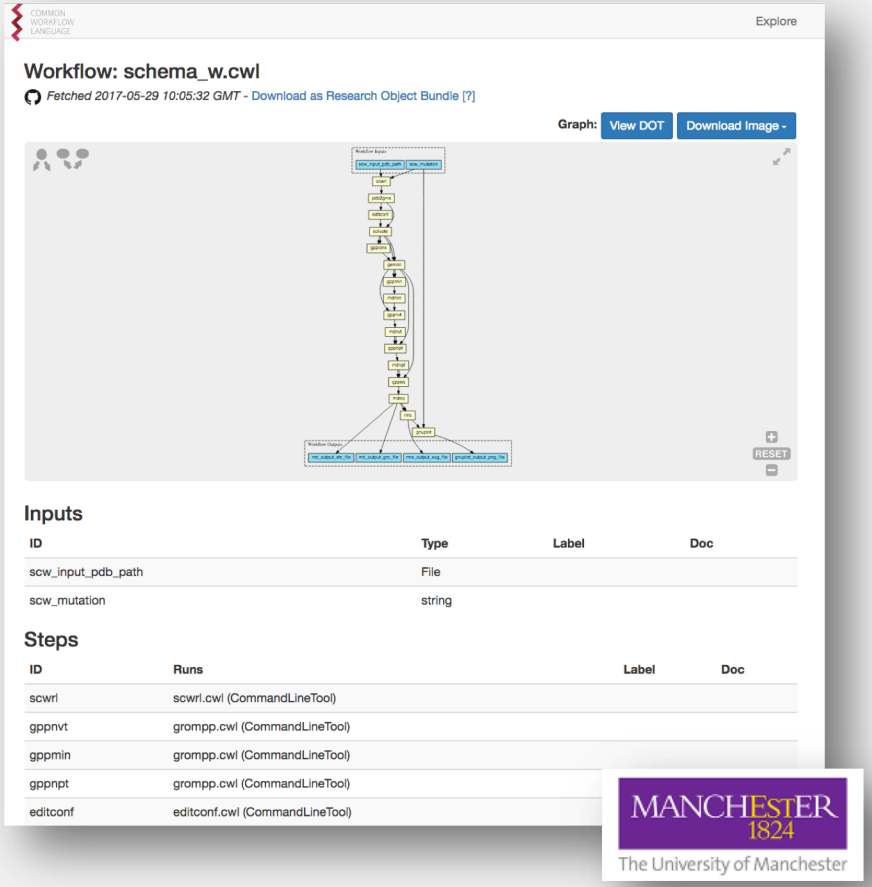
TOIL
Massively Parallel Workflows

pyCOMPSs



# Workflows: Description (CWL)

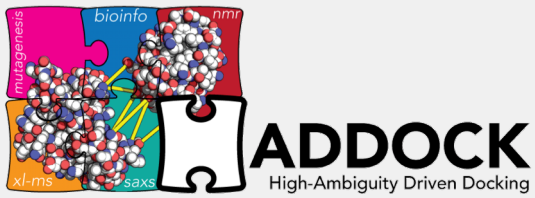



- **CWL** adopted by **ELIXIR** and the **US Food and Drug Administration** (FDA) BioCompute Object project.
- **BioExcel** supported **CWL** since day one.
- **BioExcel** have also led development of the **CWL Viewer**, which will be used to showcase **BioExcel** workflows in interactive graphical diagrams.
- Description + Python wrapper library: **BioExcel as a use case in ELIXIR, implementation study funded**.

MANCHESTER 1824 The University of Manchester

CWL elixir

# Haddock web portal



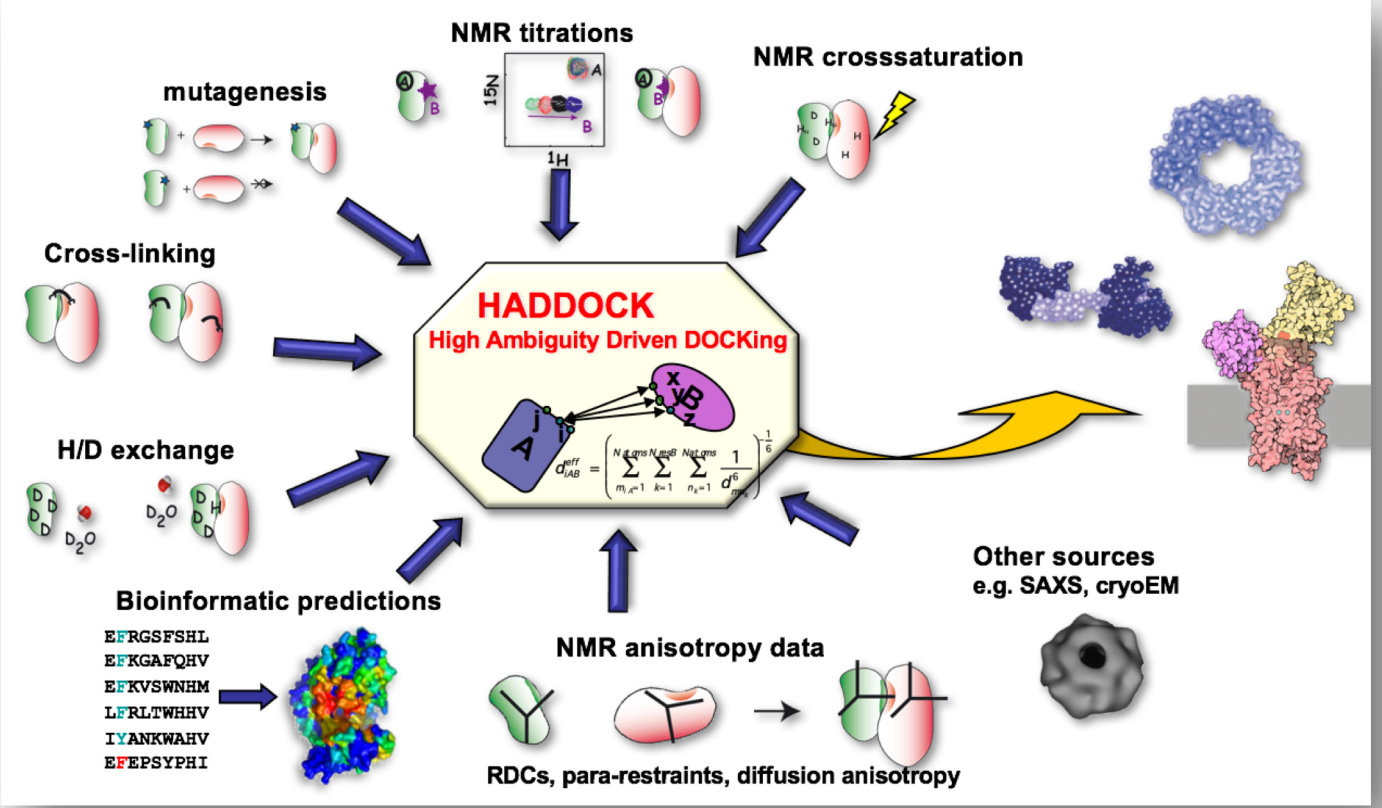

- >10000 registered users
- >190000 served runs since June 2008
- >33% on the EGI HTC resources

De Vries *et al.* Nature Prot. 2010

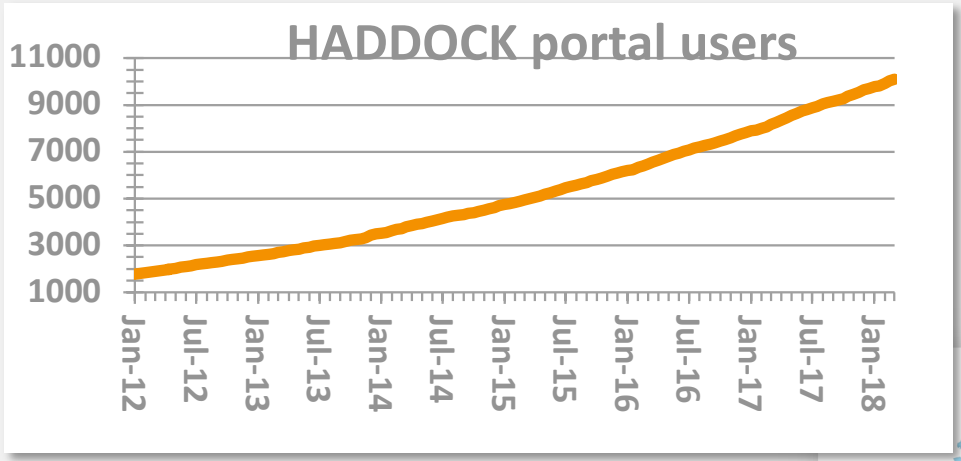
Van Zundert *et al.* J.Mol.Biol. 2016

# HADDOCK:
# An integrative modeling platform

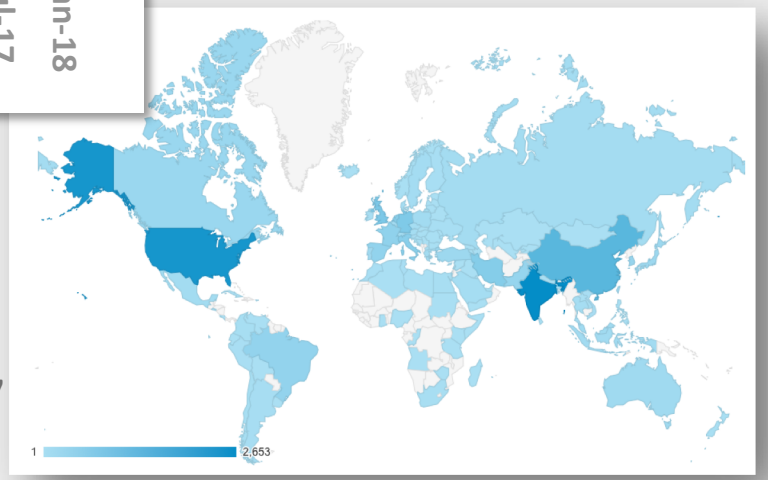
- Incorporates ambiguous and low-resolution data to guide the modelling of biomolecular complexes
- Capable of docking up to 6 molecules
- Symmetries can be leveraged
- Powerful algorithms to handle flexibility at the interface
- Final flexible refinement in explicit solvent
- One of the best performing software in CAPRI



# HADDOCK user community

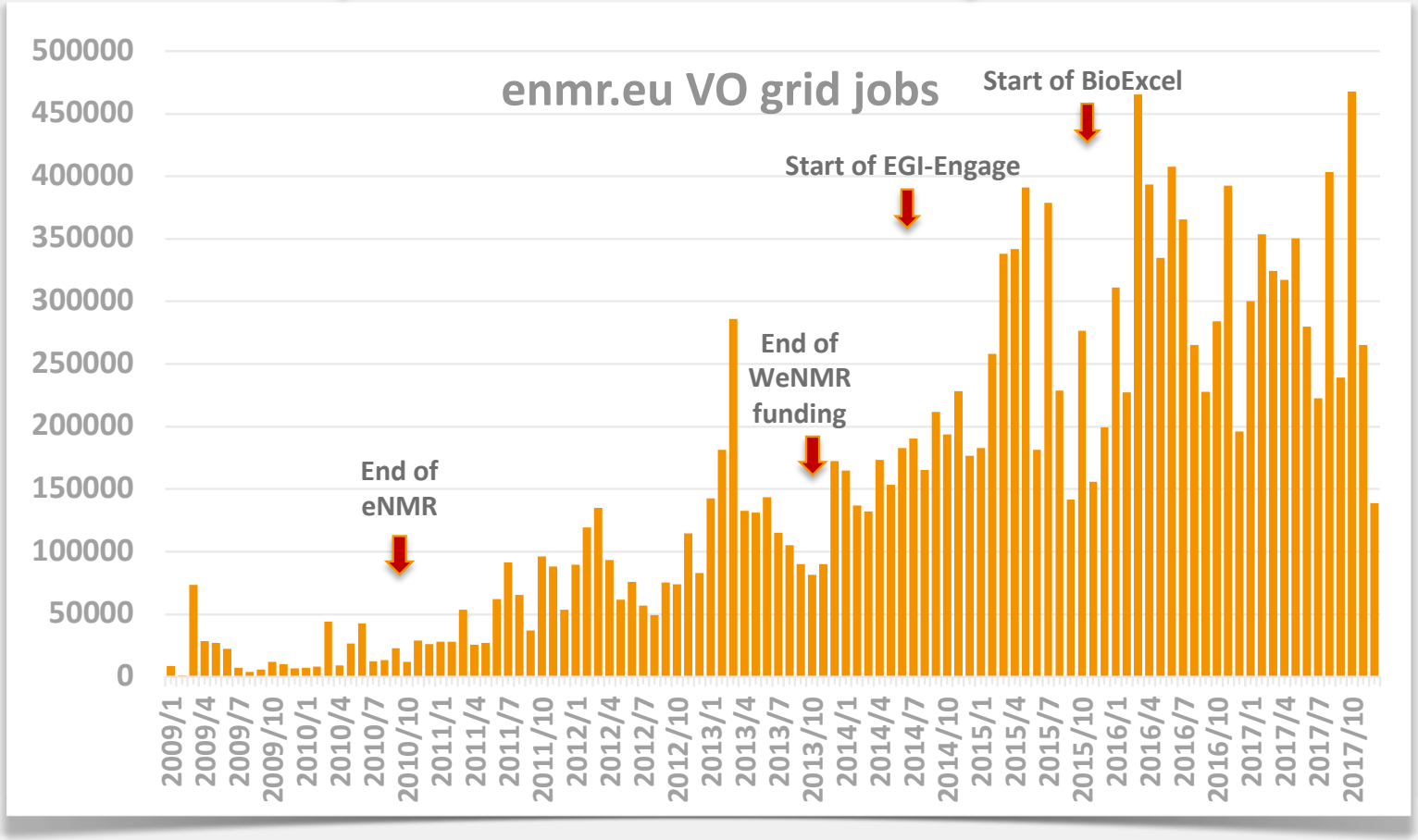



- Steady growth
- Worldwide community
- >1370 active users in 2017
- >43000 submissions to server in 2017




2017 web site visits (Google analytics)

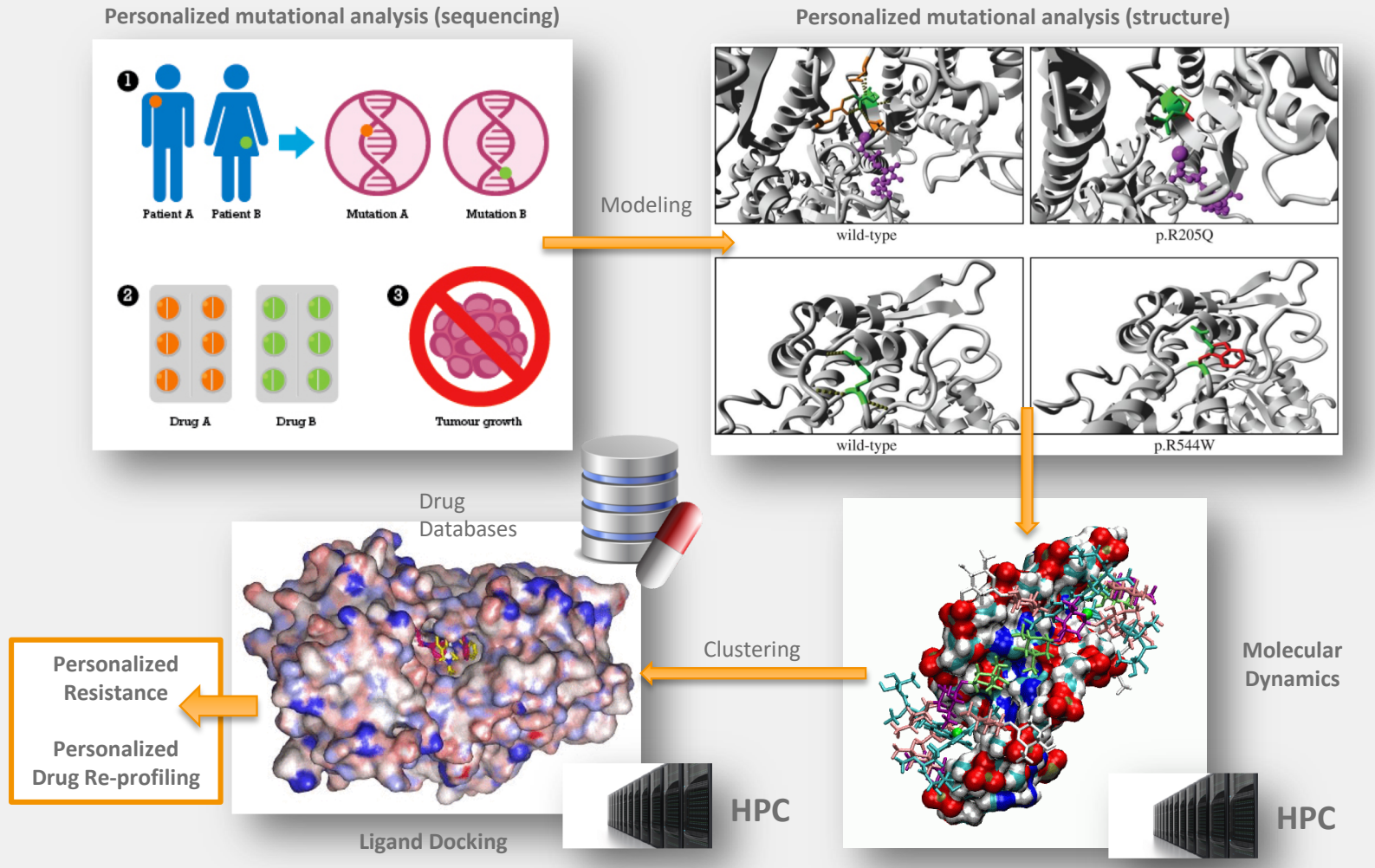
# Operational since 10 years




**~2400 normalized CPU years over 2017**

# WF1: Moving mutational analysis into the structural field for drug design

ar Research






http://www.cancer.ca/en/research-horizons/a/1/b/personalized-medicine-is-transforming-cancer-treatment/

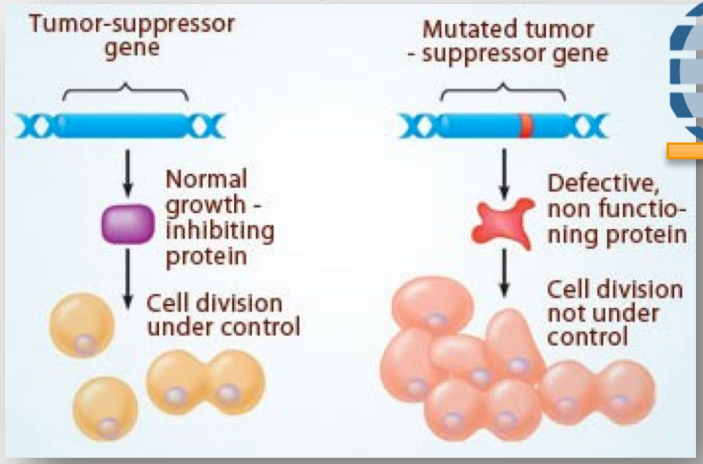
https://www.dockingserver.com/web

http://life.nthu.edu.tw/~lslth/html/ResearchFields_en_computing.html

# WF2: Tackling mutations inactivating tumor suppressors

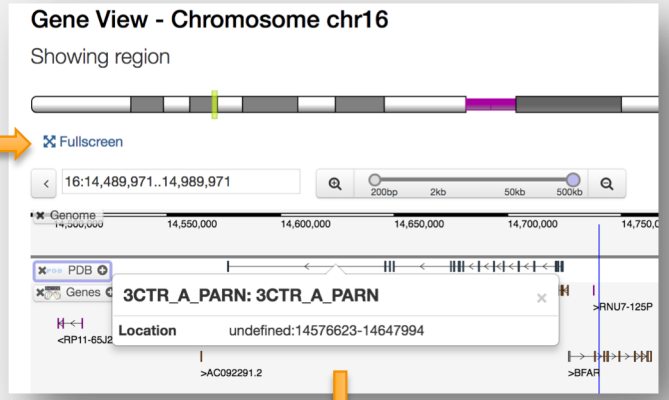
ar Research

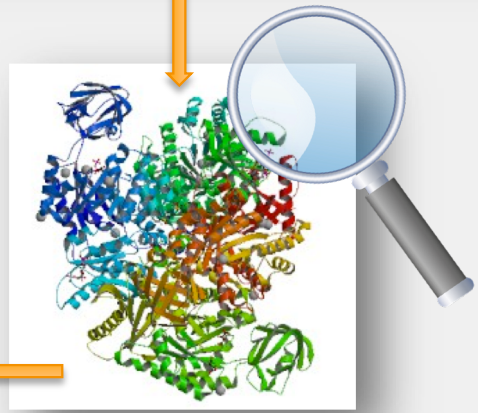
Tumor Suppressor Genes




Identify from cancer genomics projects

Sequence to Structure



Select mutations affecting allosteric signaling or stability.

Mutation modeling + MD simulations + Ligand Docking

Workflow 1

Function Recovery

Offline HPDA: drugs recovering loss of function

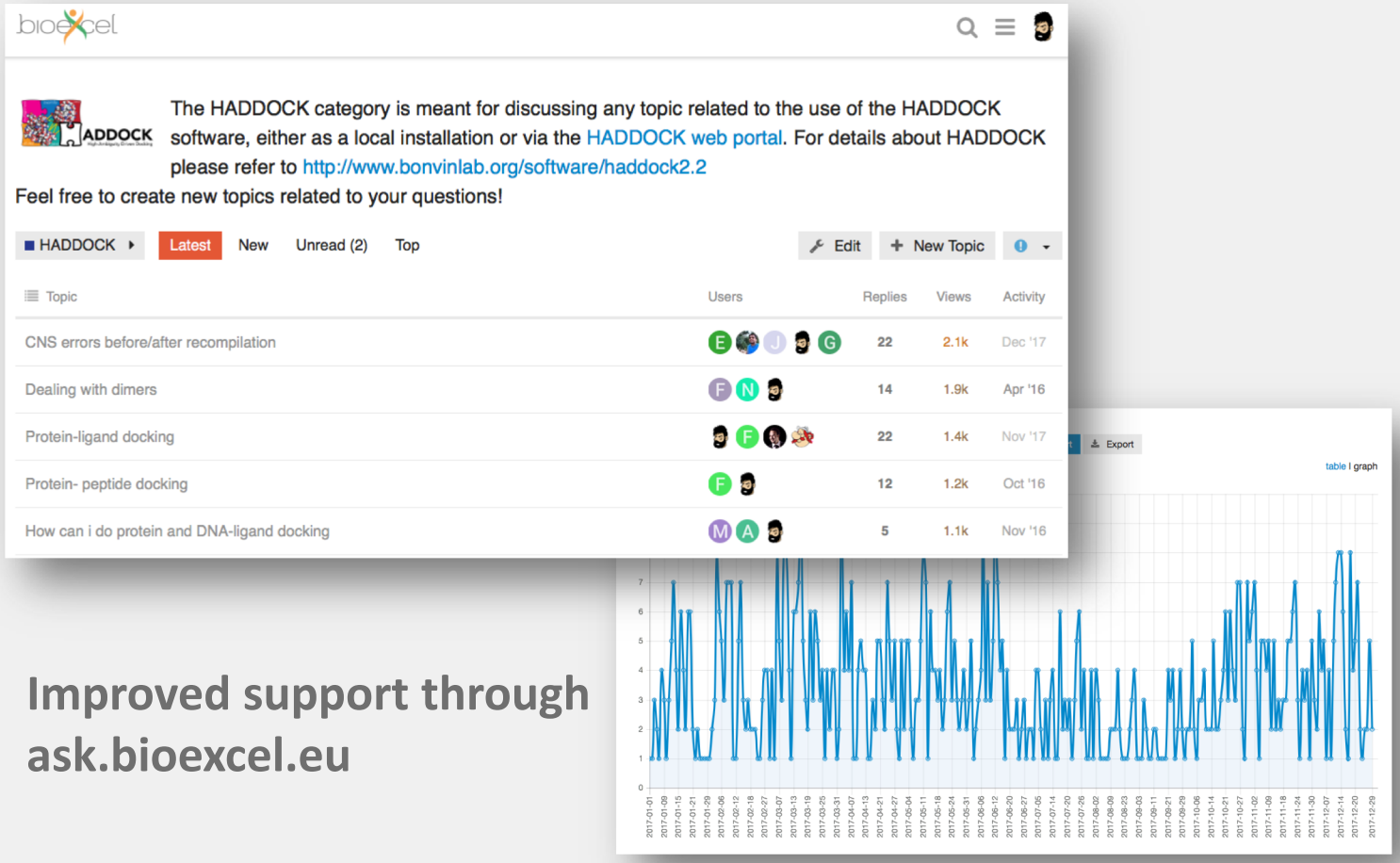
*http://www.rahmacancercare.com/Page/Patients_Support/What%20is%20Cancer*
*http://www.rcsb.org/pdb/chromosome.do*

# Objectives of BioExcel

## Competence-building among academia and industry

Promote best practices and train end users to make best use of both software and computational infrastructure

- academic and non-profit users
- industrial users
- academic and commercial resource providers
- academic code providers of related software and independent software vendors (ISVs)




## Improved support through ask.bioexcel.eu

# Interest Groups

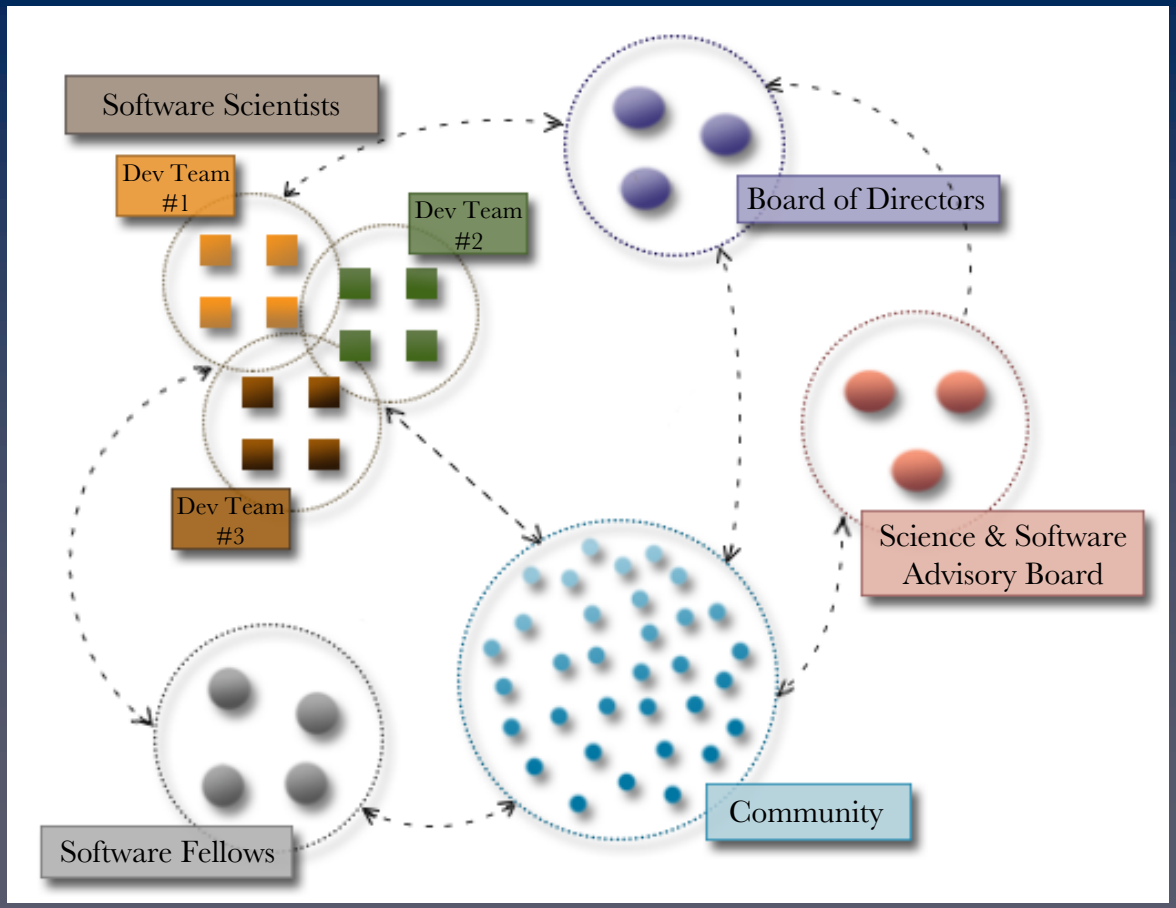
Interest Groups focused on a particular subject

- Mailing List, Webinars, Forums
- Real-Time Chat
  Code Repositories
- Collaboration Platforms
- Face-to-Face Meetings

- **Integrative Modelling**
- Best Practices for Performance Tuning and Optimization
- **Biomolecular Simulations for Entry Level Users**
- Free Energy Calculations
- **Hybrid Methods for Biomolecular Systems**
- Practical Applications for Industry
- **Training**
- Workflows

**http://bioexcel.eu/interest-groups/**

# The Molecular Sciences Software Institute (MolSSI)



# Thanks to Contributions from


University of Edinburgh


KTH Royal Institute of Technology





Universiteit Utrecht




The University of Manchester


# Summary

- The quest towards Exascale is affecting us all
- Efforts on all levels needed to be ready for the hardware to come
  - Will also see alternative hardware (e.g. neuromorphic hardware @ HBP)
- Europe is investing heavily in Exascale technologies
- Centers of Excellence (CoE) are an effective way of supporting scientific communities on their path towards the exascale
  - Needs long-term commitment
  - Critical mass