# Pilot 2: MAPPER Use Case

**Mariusz Mamonski**

On behalf of Derek Groen, James Suter, Peter Coveney and the MAPPER consortium

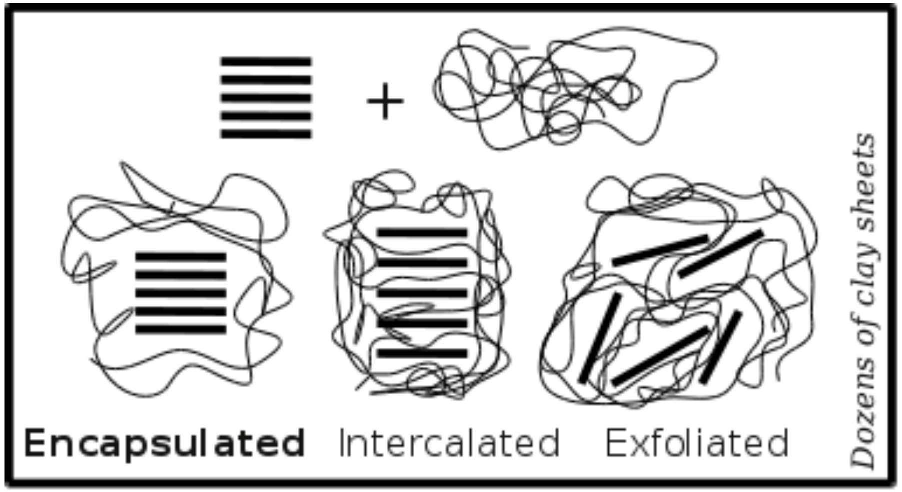
The Mapper project receives funding from the EC's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° RI-261507.

# Clay-polymer nanocomposites

- We develop quantitative coarse-grained models of clay-polymer nanocomposites to predict materials properties, e.g.:
  - The thermodynamically favourable state of the composites.
  - Their elasticity.
- Wide range of potential applications:

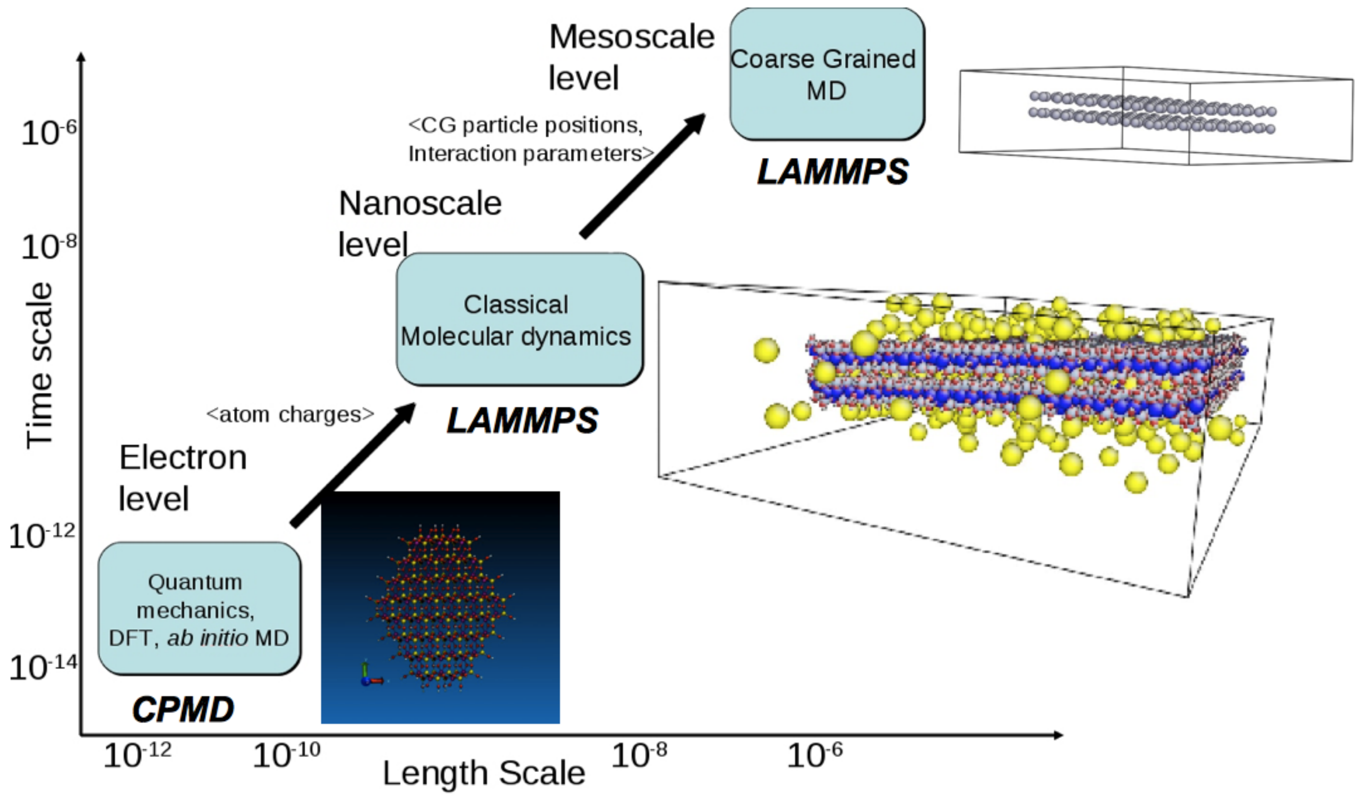
# Nanocomposites

- Main ingredients:
  - Montmorillonite clay, both "charged" and "uncharged".
  - Polymers, such as polyvinyl alcohol and polyethylene glycol.
- Simulations start off in an encapsulated state.
- We are assessing the properties of these composite systems to find cases where the materials exfoliate.

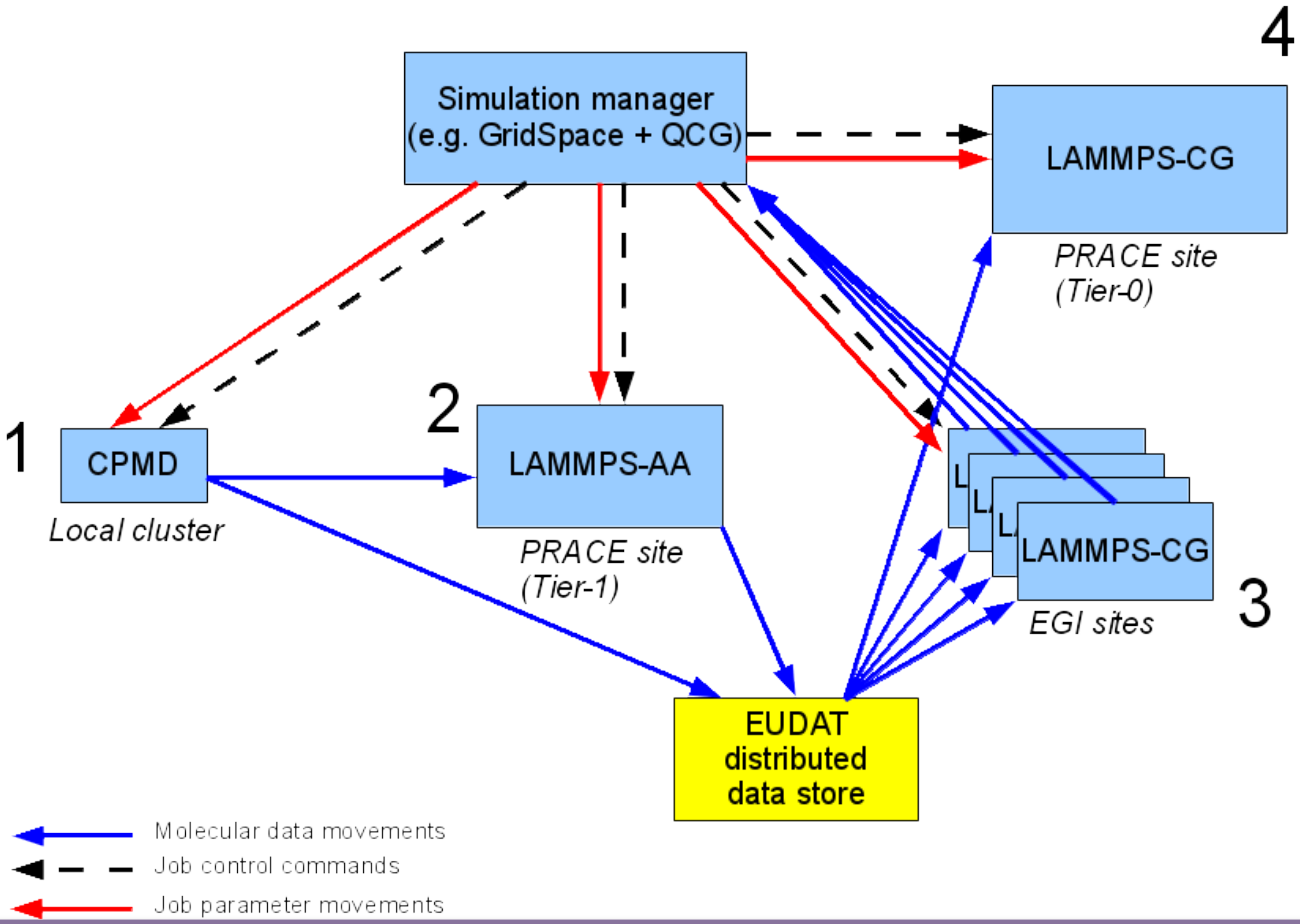


# Scale Separation Map

Scale Separation Map

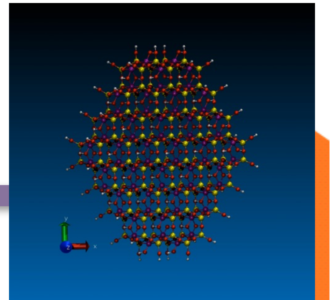


# Nanomaterials use case (extensive)

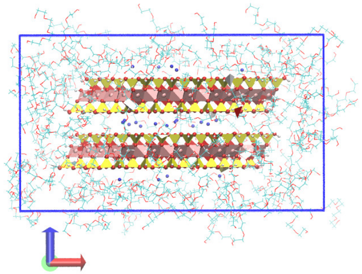


# Step 1: Quantum mechanical

- Goal: calculate energy potentials to be used in step 2.
- Code: CPMD (www.cpmd.org), optionally CASTEP.
- # of simulations: 1 (or a few).
- # of cores per simulation: <64.
- Duration per simulation: ~24 hours.
- Data produced per simulation: typically MBs, although the restart file is ~3GB.
- Data transfer required: MBs before and after the simulations.
- Site type: local cluster.

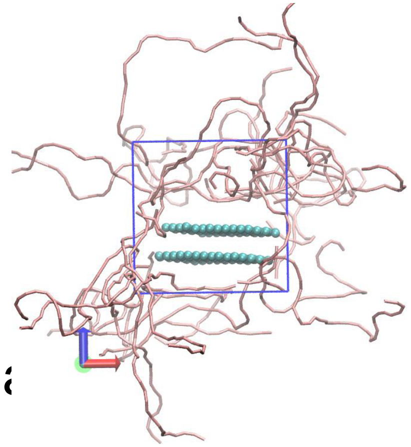


# Step 2: All-atom

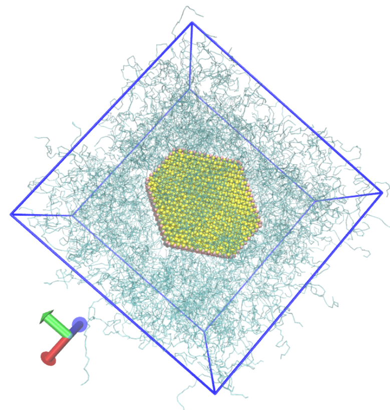


- # of simulations: 1
- # of cores per simulation: 1,024-8,192
- Duration per simulation: ~24h
- Data produced per simulation: ~1GB
- Data transfer required per simulation between F and the manager: ~1GB.
- Access mechanisms required/supported:
  - Required: GridFTP, support for remote job submission using UNICORE (via QCG-Broker)

# Step 3: CG parametrization

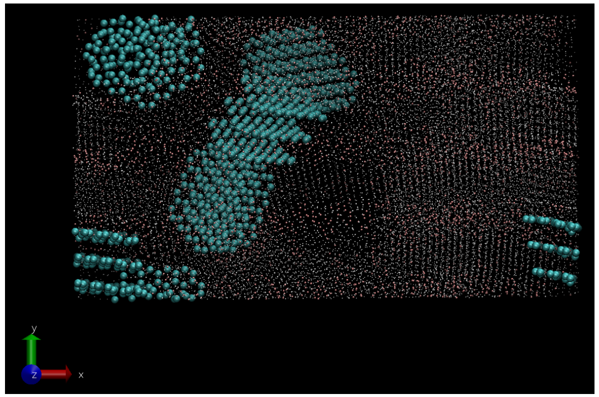


- # of simulations: ~20-40 (one after another)
- # of cores per simulation: 16 – 256
- Duration per simulation: 1h - 4h
- Data produced per simulation: 75MB - 4GB
- Data transfer required per simulation between EGI site a manager: 25 - 1GB.
    - (rest are particle positions, which can be stored for future reference)
- Access mechanisms required/supported:
    - Required: GridFTP, must have support for QCG job submission, preferably have QCG Computing installed.
    - Advance reservation provides a performance benefit here.



# Step 4: CG Large Simulation

- # of simulations: 1
- # of cores per simulation: 8,192-65,536 cores.
- Duration per simulation: ~12h
- Data produced per simulation: 1TB+
- Data transfer required: MBs to start, 1TB+ afterwards.
  - The particle positions of these simulations are to be stored for future reference and analysis.
- Access mechanisms required/supported:
  - Required: GridFTP, support for QCG job submission.



# Tools

- GridFTP.
- UNICORE.
- QosCosGrid Environment.
- AHE.
- CPMD.
- LAMMPS.
- Perl / Python scripts.
- GridSpace (on the user side).

# A few other data aspects

- All data is stored in files.
- Filenames may be non-unique (e.g., "in.lammps")
- Filename+directory tree is unique.
- Position files are typically large (>1GB), other files are much smaller.

MAPPER

# Questions?

d.groen@ucl.ac.uk