DL_POLY_3: Additional Functionality and Experiences on HPC systems

I.T. Todorov

Computational Science and Engineering Department
CCLRC Daresbury Laboratory
Warrington WA4 4AD

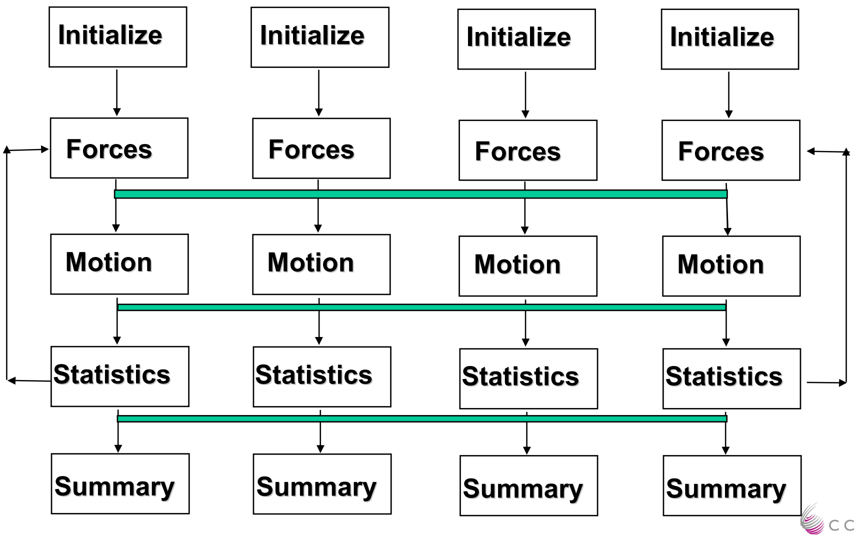
DL_POLY_2&3 Versions

- DL_POLY_2 (version 17)
  - **Replicated Data** parallelisation, limits up to ≈30,000 atoms with good parallelisation up to 32 (system dependent) processors (running on any processor count)
  - Full force field and molecular description
  - Written in modularised F90 (+MPI), no external dependencies
  - Free format reading with somewhat rigid semantics
- DL_POLY_3 (version 07)
  - **Domain Decomposition** parallelisation, limits up to ≈ 2.1×10^9 atoms with inherent parallelisation (any high processor count, 2^k for SPME)
  - Full force field and molecular description but no rigid body description
  - Written in modularised free formatted F90 (+MPI) with rigorous code syntax (FORCHECK and NAGWare verified), no external library dependencies
  - Free format semantically approached reading with some fail-safe features (but fully fool-proofed)

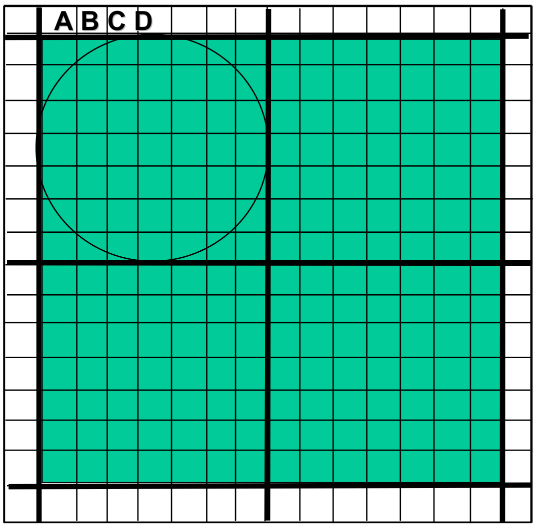
Parallel Computing

- Load Balancing
  - Equal division of work among processors
  - Equal division of memory requirement
  - Concurrent processor use
- Communications
  - Minimise number of messages
  - Maximise information per message
  - Local communication rather than global
  - Asynchronous communication

Replicated Data Schematic

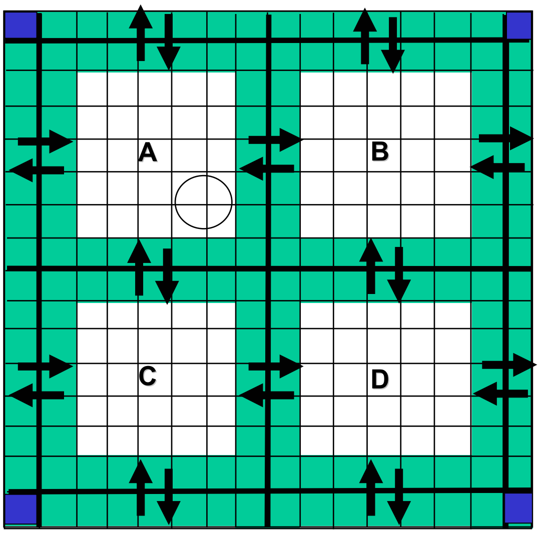


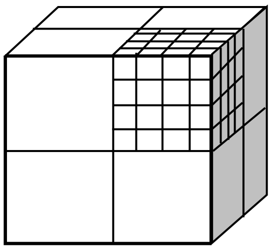
Replicated Data MD Cell

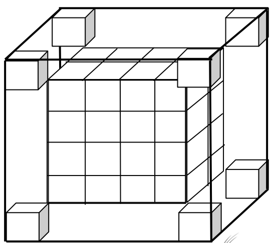


- No memory distribution (performance overheads and limitations increase with increasing system size)
- Functional decomposition of the workload
- Cutoff ≤ 0.5 min system width
- Extensive global communication (extensive overheads increase with increasing system size)

Domain Decomposition







CSE Computational Science & Engineering Department

Major MD cycle expenses

- Construction of the Verlet neighbour-list using Linked Cell (LC) lists, taking up ≈ 40% of the CPU time per MD cycle.
- Evaluation of the coulombic interactions in ionic systems using Smoothed Particle Mesh Ewald (SPME) summation, taking up ≈ 40% of the CPU time per MD cycle. SPME uses 3D FFT procedures (DAFT) implemented to fit in the DD and LC splitting of the MD box so that overall scaling is $\approx O(N \log N)$ rather than $\approx O(N^3)$. 90% of the time spent on SPME evaluation is for the generation of particles' B-splines and construction of a DD 3D charge array.

CCLRC

CSE Computational Science & Engineering Department

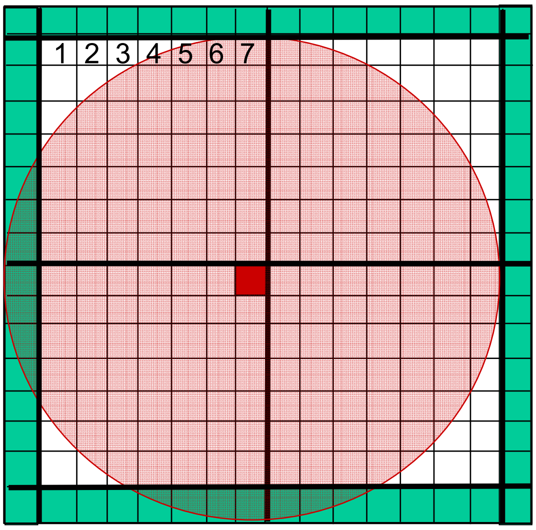
Linked Cells

- Linked lists provide an elegant way to scale short-ranged two body interactions from $O(N^2/2)$ to $\approx O(N)$. The efficiency increases with increasing link cell partitioning – as a rule of thumb best efficacy is achieved for cubic-like partitioning with number of link-cells per domain ≥ 4 for any dimension.
- Linked lists can be used with the same efficiency for 3-body (bond-angles) and 4-body (dihedral & improper dihedral & inversion angles) interactions. For these the linked cell halo is double-layered and since $cutoff^{3/4\text{-}body} \leq 0.5{*}cutoff^{2\text{-}body}$ which makes the partitioning more effective than that for the 2-body interactions.
- The larger the particle density and/or the smaller the cutoff wrt the domain width, (the larger the sub-selling and the better the spherical approximation of the search area), the shorter the Verlet neighbour-list search.

CCLRC

CSE Computational Science & Engineering Department

Sub-celling of LC



- Provides dynamically adjustable workload for variable local density and VNS speed up of ≈ 30% (45% theoretically).
- Provides excellent serial performance, extremely close to that of Brode-Ahlrichs method for construction of the Verlet neighbour-list when system sizes are small < 5000 particles.

9

CCLRC

CSE Computational Science & Engineering Department

Functionality Differences

DL_POLY_2

- Rigid bodies
- Standard Ewald, HK_Ewald, Neutral groups
- Multiple timestep (not precise for NED)
- Truncated octahedral and rhombic dodecahedral periodic boundaries (geometries cannot be DD partitioned)

DL_POLY_3

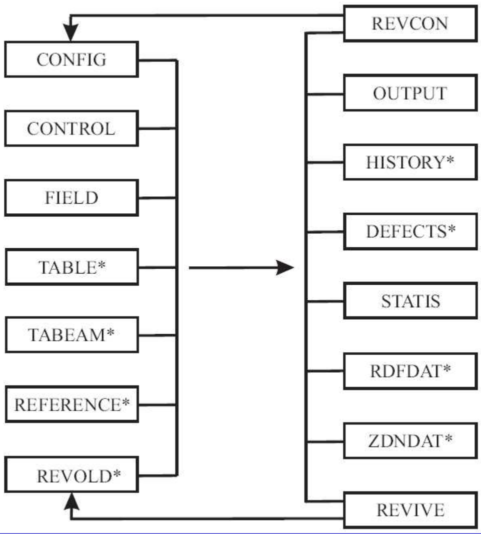
- Variable timestep, Defect Detection, Pseudo Thermostat, Extended Coulombic eXclusion, densvar, rlxtol, mxshak, bisize (RDF,Zden), nfold, PMF to all ensembles (CONTROL)
- MTK NPT and NσT ensembles, Langevin thermostats
- RDF look-up (FIELD)
- Atom displacements from original positions in HISTORY
- MUST run only on 2^k number of processors when SPME

10

CCLRC

CSE Computational Science & Engineering Department

DL_POLY_3 I/O Files



11

CCLRC

CSE Computational Science & Engineering Department

DL_POLY_3.CML

- Benefits
  - Metadata information – data about the data
  - Built in semantics and native hierarchy of data structures allows for faster search of data
  - Avoids fixed data/document formatting
    - Provides common ground for scientific simulation codes and supplementary software – INTEROPERABILITY
    - Avoids the use of parsers
  - Facilitates database incorporation
  - Allows transformation to XHTML and to graphical browsing

- Drawbacks
  - Not (intended to be) human readable
  - Increases the volume of data - storage

12

CCLRC

CSE Computational Science & Engineering Department

Experiences on HPC*x*

- HPC*x* is an IBM proprietary cluster situated at Daresbury Laboratory which is regulated by the HPC*x* consortium (University of Edinburgh, CCLRC and IBM) funded by EPSRC. (www.hpcx.ac.uk)
- HPC*x* is currently the 2^{nd} (4^{th}) best in the UK and 50^{th} best in the world. (www.top500.org)
- Currently HPC*x* consists of 198 nodes each having 8 p5-575 Series (1.5 MHz) dual core CPUs with 4MB memory per core. It is SMT2 enabled allowing for 2 processes per core.
- Largest MD system simulated on HPC*x* with DL_POLY_3 is a 300 million NaCl on 1024 processor units (67 seconds per timestep). L=2μm, λ=0.5μm theoretically visible with the eye.

13 CCLRC

CSE Computational Science & Engineering Department

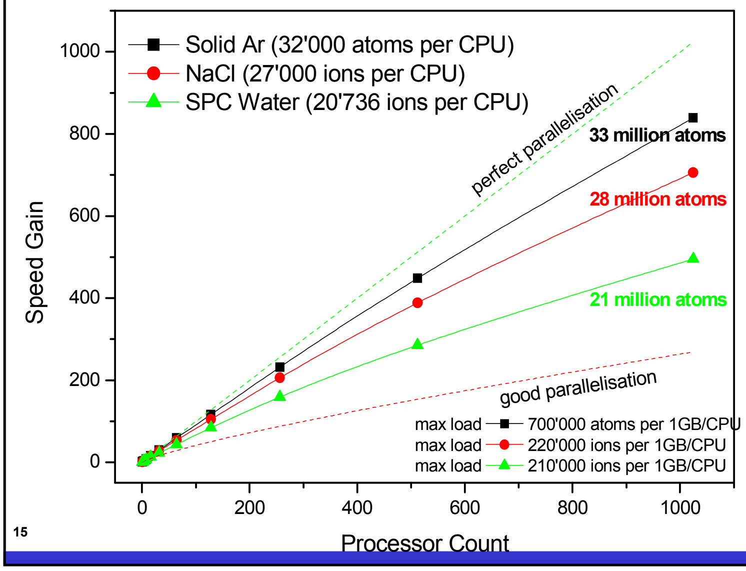
Benchmarking on HPC*x*

- **Solid Ar** : L-J interactions between Ar-Ar
- **NaCl** : vdW interactions between Na^+-Na^+, Na^+-Cl^- and Cl^--Cl^-, and Coulomb forces between the ions.
- **SPC Water** : L-J interactions between O^--O^-, Coulomb forces between the ions and three constraints per water molecule: O-H1, O-H2 and H1-H2.
- To keep the numeric algorithms the same with different processor counts all test cases have been constructed recursively from previously equilibrated initial structures so that the work load and memory distribution per CPU is kept constant.
- Constant Ewald precision and constraint algorithms do not scale linearly with system size.

14 CCLRC

CSE Computational Science & Engineering Department

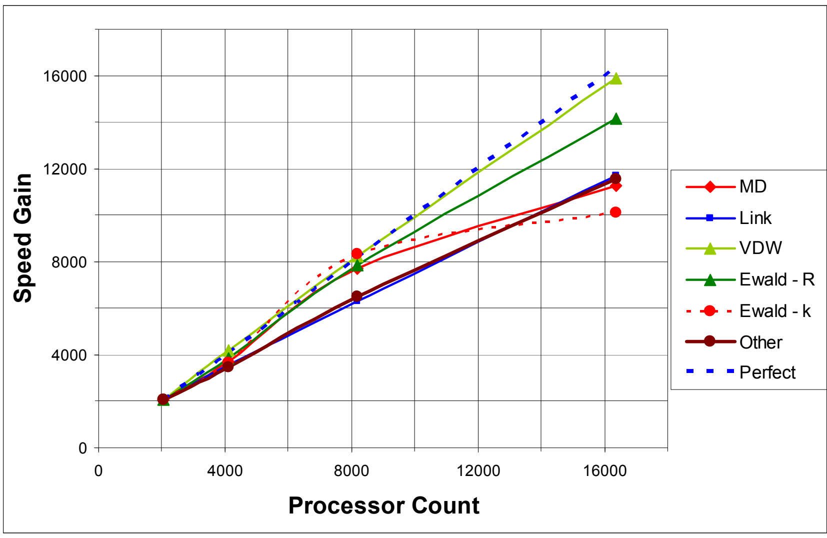
Performance on HPC*x*



15

CSE Computational Science & Engineering Department

Performance on BG - 15×10^6 ions



16

FURTHER INFORMATION & HELP

- Information
  http://www.ccp5.ac.uk/DL_POLY

- Forum
  http://www.cse.clrc.ac.uk/disco/forums.shtml