DL_POLY’s 25th Anniversary

Bill Smith

Tim Forester, Maurice Leslie
Ian Bush, Ilian Todorov
Michael Seaton, Laurence Ellison
Duncan Harris, Christos Kartsaklis
Pierre-André Cazade
Ruslan Davidchak, Andrey Brukhno
Henry Boateng, Alin Elena
Alice’s Wonderland (1865)
Lewis Carroll (Charles Lutwidge Dodgson)
General Purpose Software for Large Computers
Body of Knowledge

- Proteins solvation & binding
- DNA strands dynamics
- Membranes’ processes
- Drug polymorphs & discovery
- Crystalline & Amorphous Solids – damage and recovery
- Dynamic processes in Metal-Organic & Organic Frameworks
- Dynamics at Interfaces & of Phase Transformations
Spinoffs & Further Software Projects @ DL

- FF mapping via DL_FIELD
- Coarse graining via DL_CGMAP
- CRYSTAL (L. Bernasconi)
- CASTEP (K. Refson)
- MC via DL_MONTE
- MS & MD via DL_POLY
- DPD & LB via DL_MESO
- KMC via DL_AKMC
- QM/MM bridging via #ChemShell

Diagram with a graph showing accuracy vs. capability, along with various software projects and techniques.
Software is a critical part of modern research...

SOFTWARE CITATION PRINCIPLES

IMPORTANCE
Software should be considered a legitimate and citable product of research. Software citations should be accorded the same importance in the scholarly record as citations of other research products; they should be included in the metadata of the citing work, such as a reference list. Software should be cited on the same basis as any other research product such as a paper or a book.

UNIQUE IDENTIFICATION
A software citation should include a method for identification that is machine actionable, globally unique, interoperable, and recognized by at least a community of the corresponding domain experts, and preferably by general public researchers.

ACCESSIBILITY
Software citations should facilitate access to the software itself and to its associated metadata, documentation, data, and other materials necessary for both humans and machines to make informed use of the referenced software.

CREDIT AND ATTRIBUTION
Software citations should facilitate giving scholarly credit and normative, legal attribution to all contributors to the software, recognizing that a single style or mechanism of attribution may not be applicable to all software.

PERSISTENCE
Unique identifiers and metadata describing the software and its disposition should persist—even beyond the lifespan of the software they describe.

SPECIFICITY
Software citations should facilitate identification of, and access to, the specific version of software that was used. Software identification should be as specific as necessary, such as using version numbers, revision numbers, or variants such as platforms.

... yet there is little support for its acknowledgement and citation...
DL_POLY Trivia

- **General purpose parallel (classical) MD simulation software**
- It was conceived to meet needs of CCP5 - The Computer Simulation of Condensed Phases (academic collaboration community)
- Written in modularised Fortran90 (NagWare & FORCHECK compliant) with MPI2 (MPI1+MPI-I/O) & fully self-contained
  - 2003 – 2010: DL_POLY_3 (DD) by I.T. Todorov & W. Smith (funded for 4 years by NERC at Cambridge). **Up-licensed** to DL_POLY_4 in 2010 – free of charge to academic researchers and at cost to industry (**provided as source**).
- ~ 16,000 licences taken out since 1994 (~1,600 pa since 2007)
- ~ 4,000 e-mail list
Written in modularised free formatted F90 (+MPI) with rigorous code syntax (FORCHECK and NAGWare verified) and **no external library dependencies**

- **DL_POLY_4** (version 8)
  - Domain Decomposition parallelisation, based on domain decomposition (no dynamic load balancing), limits: up to $\approx 2.1 \times 10^9$ atoms with inherent parallelisation
  - Parallel I/O (amber netCDF) and radiation damage features
  - Free format (flexible) reading with some fail-safe features and basic reporting (but not fully fool-proofed)

- **DL_POLY_Classic** (version 1.9)
  - Replicated Data parallelisation, limits up to $\approx 30,000$ atoms with good parallelisation up to 100 (system dependent) processors (running on any processor count)
  - Hyper-dynamics, Temperature Accelerated Dynamics, Solvation Dynamics, (Path Integral MD)
  - Free format reading (somewhat rigid)
DL_POLY_DD Project Evolution

Published lines of code [x1,000]

<table>
<thead>
<tr>
<th>Year</th>
<th>Lines [x 1,000]</th>
<th>Comment</th>
<th>Blank</th>
<th>Total</th>
<th>Manual</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>DL_POLY_3.01</td>
<td>4.0</td>
<td>5.6</td>
<td>10.6</td>
<td>178 p</td>
</tr>
<tr>
<td></td>
<td>DL_POLY_4.08</td>
<td>20.1</td>
<td>46.7</td>
<td>66.8</td>
<td>312 p</td>
</tr>
</tbody>
</table>

DL_POLY_3.01

- Single developer
- Reengineered

DL_POLY_4.08

- Reengineered

Projects:
- DL_POLY_3.01
- DL_POLY_4.08

Manual Documentation:
- 178 pages
- 312 pages
DL_POLY Licence Statistics

Annual Downloads & Valid eMail List Size

2010 :: DL_POLY (2+3+MULTI) - 1,000 (list end)
2016 :: DL_POLY_4 - 3,867 (list start 2011)

2016 Downloads
• UK - 23.2%
• EU-UK - 18.7%
• China - 12.1%
• USA - 10.6%
• India - 6.3%
• France - 4.8%
• London - 7.6%
• Beijing - 2.8%
• Cambridge - 1.7%

DL_POLY Licence Statistics
DL_POLY Impact

Google Scholar Citations

Year

1993 1995 1997 1999 2001 2003 2005 2007 2009 2011 2013 2015
DL_POLY Project Challenges

• Routines’ Interfaces (Rewriting for future fitness)
• Extending Code of Coding (becoming stricter)
• Inclusion of tests and user manual information (information)
  • Automatic testing and Coverage Analysis
• Code Reviewing (acceptance of submissions)
• Large projects/challenges
  • EVB
  • Shaped Particles
  • DFTB
  • FFS
  • Beyond Ewald
    • Poisson Solvers
    • Multipolar Electrostatics
  • Open MP
• UI and coupling to other methods (MC, QM/MM, etc.)
Chair: Ilian Todorov  
09.05-09.35 **Prof. Martin Dove (QMUL)** - *Molecular dynamics simulations of carbon capture by porous hybrid materials*  
09.40-10.10 **Dr. Patrice Bordat (University of Pau)** - *Solvation and free energy module implemented in DL_POLY: Study for a preferential CO2/CH4 adsorption in silica monoliths*

10.15-10.35 *Tea/Coffee Break*

Chair: Tim Forester  
10.40-11.10 **Prof. John Harding (University of Sheffield)** - *Understanding biomineralisation: what has DL_POLY ever done for us?*  
11.15-11.45 **Dr. Simone Melchionna (ISC-CNR)** - *Proteins and multiscale biology: the long time legacy of DL_POLY*  
11.50-12.20 **Prof. Richard Catlow (UCL/University of Cardiff)** - *Molecular dynamics in Catalytic systems*

12.25-13.25 *Lunch*

Chair: Maurice Leslie  
13.30-14.00 **Dr. Kostya Trachenko (QMUL)** - *Using DL_POLY to understand radiation damage effects and soft matter (glasses, liquids, supercritical fluids)*  
14.05-14.35 **Dr. P.-L. Chau (Institut Pasteur)** - *General Anaesthetics and Membrane Interactions*  
14.40-15.10 **Dr. David Quigley (University of Warwick)** - *The Hackademic Approach to Simulations with DL_POLY*

15.10-15.30 *Tea/Coffee Break*

Chair: Neil Allan  
15.35-16.05 **Prof. Steve Parker (University of Bath)** - *Atomistic Simulations of Oxide and Mineral Interfaces*  
16.10-16.40 **Prof. Martyn Guest (University of Cardiff)** - *DL_POLY - A Performance Overview; Analysing, Understanding and Exploiting available HPC Technology*  
16.45-17.25 Closing Remarks by **Prof. William Smith** - *A Short History of DL_POLY*

17.30-19.00 Free time to explore the site  
19.00 Official Dinner