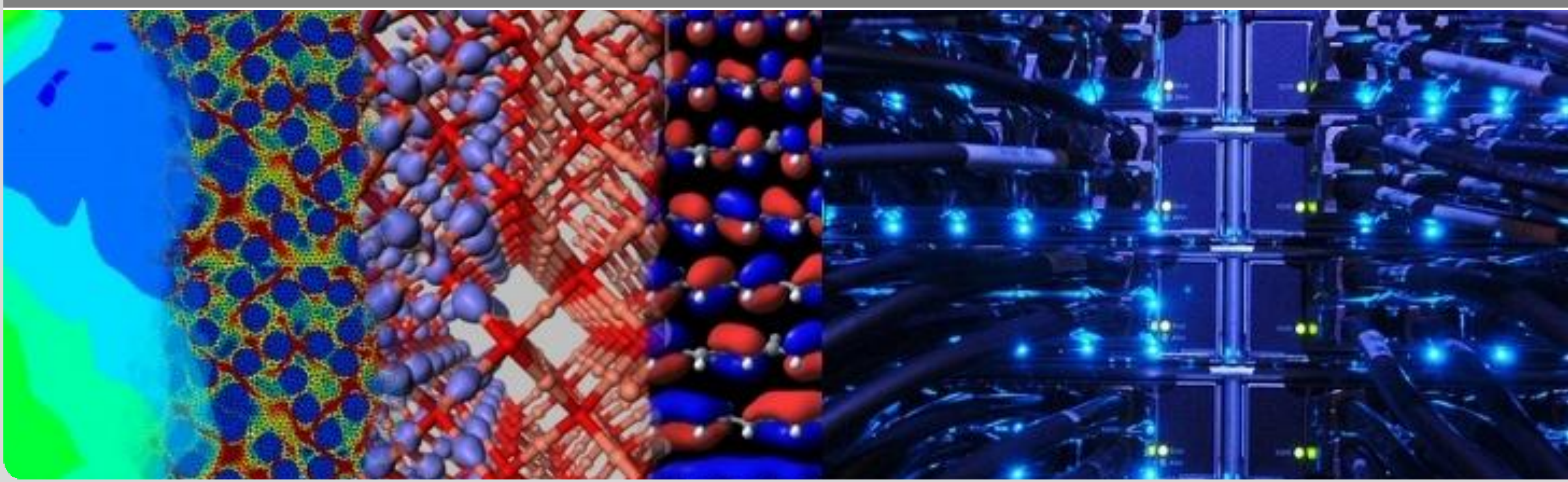
Karlsruhe Institute of Technology

A service-oriented approach for multiscale materials modelling

Ivan Kondov, **Stefan Bozic, Velimir Meded and Wolfgang Wenzel**

STEINBUCH CENTRE FOR COMPUTING - SCC



Outline


Source: Nick Piggott / www.flickr.com

Introduction

- Application area and importance
- Challenges
- Service Oriented Architecture
- Solution strategies

Implementation

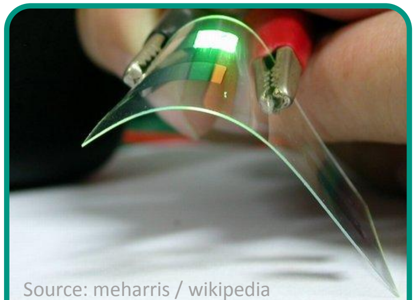
- Implementation strategies
- Methodology – technologies used

Discussion

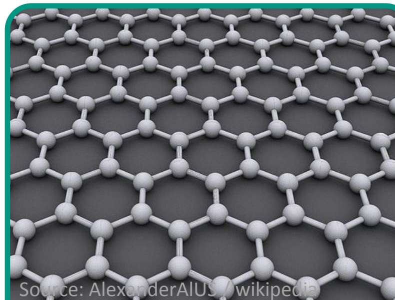
- Key results: Proof of principle
- Conclusions and outlook

Application area and importance

Organic electronics



OLEDs

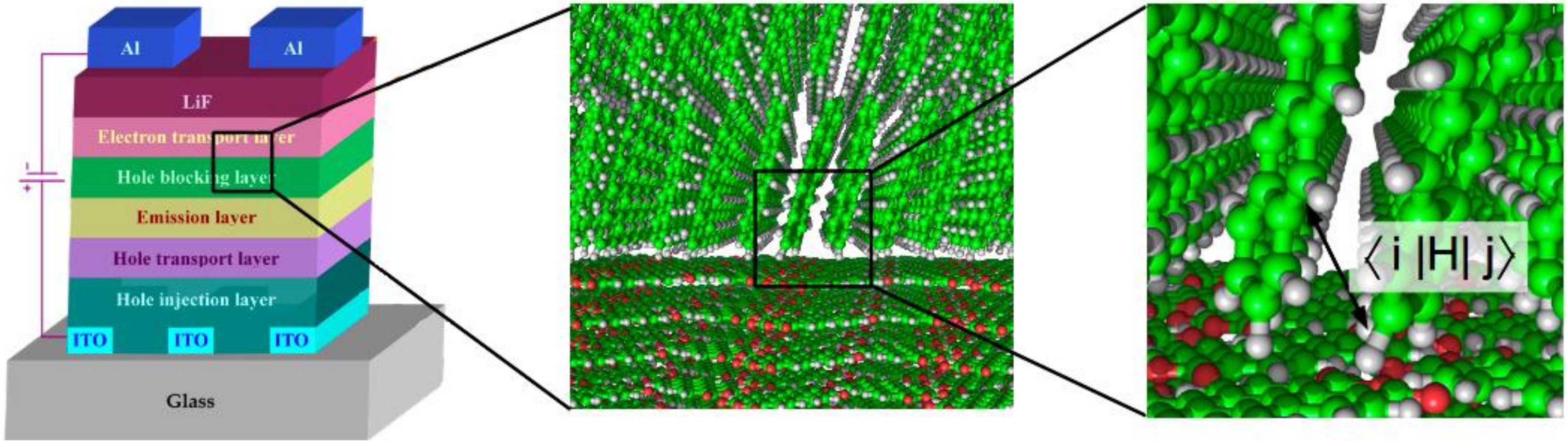


Carbon electronics

Li ion batteries

- Modelling and computer simulation essential
  - Reduce time-to-market
  - Reduce product development costs
  - Increase agility of industrial R&D
- Very complex models and environments for simulation
  - Accessible only for few experts
  - Low effectiveness and general applicability

Example: Multiscale modelling of OLEDs



macroscopic scale ~ 10^{-6} m | molecular scale ~ 10^{-8} m | electronic scale ~ 10^{-10} m

Continuum model (FEA)	Coarse-grained model (CG)	Atomistic model (MM)	Quantum mechanical model (QM)
Elmer	ToFeT (KMC)	DEPOSIT	MOPAC
FEAP	End-bridging MC	DL_POLY	TURBOMOLE
	Transporter	LAMMPS	BigDFT

Challenges and proposed solution strategies

Karlsruhe Institute of Technology

Computing capacity and capability

- High Performance Computing on PRACE resources
- High Throughput Computing on EGI resources

Security and reliability

- Grid Security Infrastructure
- UNICORE: x.509, SSL, SAML

License issues

- Open Source Licenses
- UNICORE: UVOS, SAML,VOMS

In this talk

Application integration and reusability

- Service Oriented Architecture (SOA)
- Application interfaces: GridBeans
- UNICORE Workflows

Data interoperability

- OpenMolGRID
- Chemical Markup Language (CML)
- Data model

In this talk

Service Oriented Architecture (SOA)

"A set of components which can be invoked, and whose interface descriptions can be published and discovered" (W3C)

High reuse, no customization!

- Low effort to create new "ad hoc" composite applications from existing services
- Low effort for changing application

- **SOA principles:** standardized service contract, loose coupling, abstraction, reusability, autonomy, statelessness, discoverability and composability
- Standards for **SOA implementation:** Web Services
  - WSDL or WADL for describing the service
  - SOAP or REST for messaging

Implementation strategies

Application integration and reusability

- UNICORE middleware
- Application interfaces with GridBeans
- Application wrappers with OpenMolGRID
- Sequential modelling with UNICORE workflows

Data interoperability

- Pre- & post-processing with OpenMolGRID
- Data exchange with CML
- Data model
- Dataflow management

UNICORE

- UNICORE: **UN**iform **I**nterface to **CO**mputing **RE**sources
- Grid computing technology supported by the European Middleware Initiative (EMI)
- Seamless, secure and intuitive access to distributed grid resources
- Used in daily production at numerous supercomputer centres worldwide (for example in PRACE)
- Open source under BSD license
- Implements SOA using standards from the Open Grid Forum (OGF), W3C and OASIS

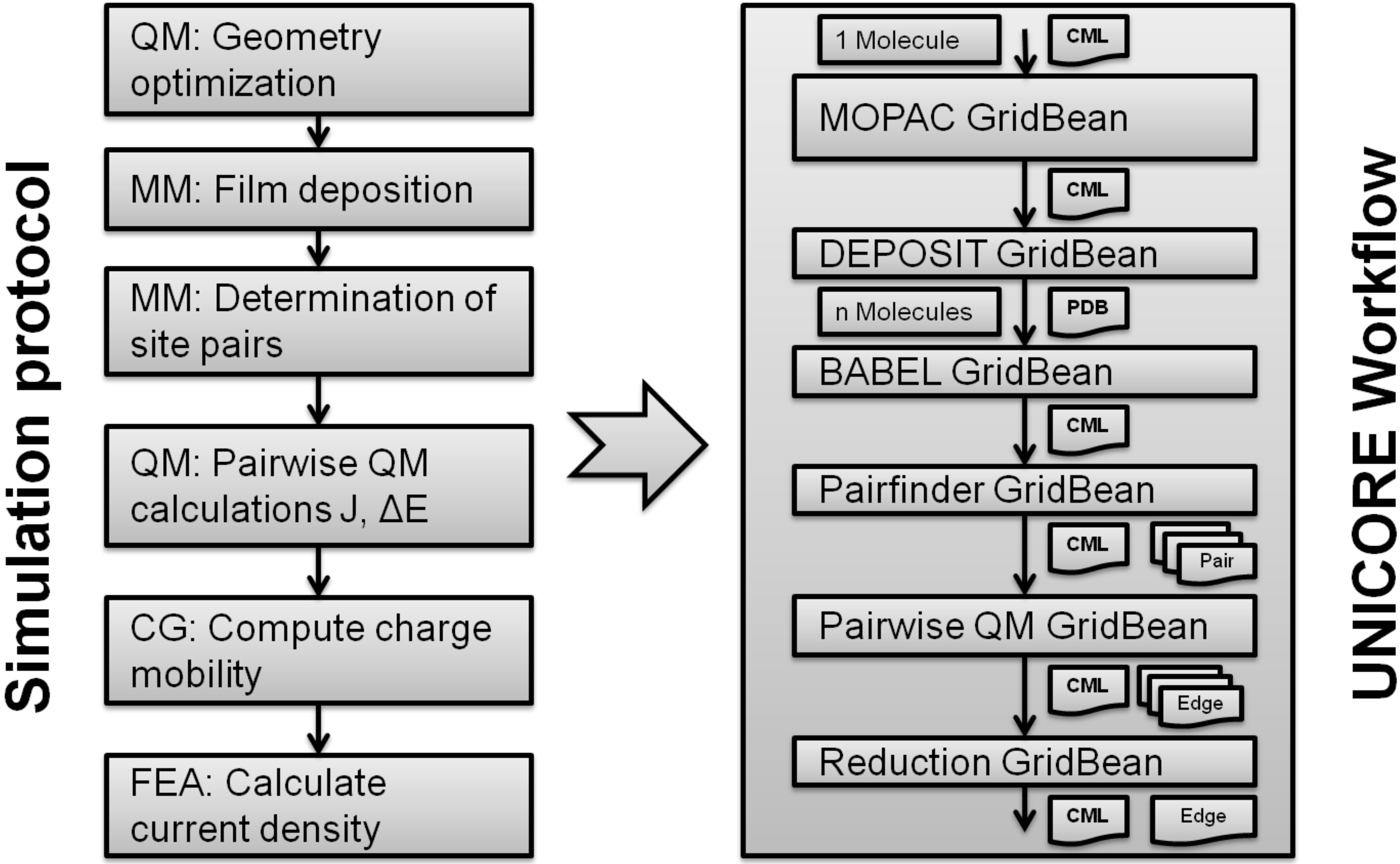
A. Streit et al., UNICORE 6 - Recent and Future Advancements
Annals of Telecommunications 65 (11-12), 757-762 (2010).

Chemical Markup Language (CML)

- Website: www.xml-cml.org
- Developed by Peter Murray-Rust and Henry Rzepa since 1995
- Provides semantics for molecules, compounds, reactions, spectra, crystals and computational chemistry
- Infrastructure includes **legacy converters**, **dictionaries** and **conventions**

```
<list dictRef="compchem:basisSetContractions" id="basisSetContractions">
      <scalar dataType="xsd:string" id="atomType" dictRef="compchem:atomType">h1</scalar>
      <scalar dataType="xsd:string" id="elementType" dictRef="compchem:elementType">H</scalar>
      <scalar dataType="xsd:string" id="basisSetType" dictRef="compchem:basis_set_type">orbital</scalar>
      <scalar dataType="xsd:string" id="basisSetLabel" dictRef="compchem:basisSetLabel">cc-pvdz</scalar>
      <scalar dataType="xsd:string" id="basisSetHarmonicType" dictRef="compchem:basisSetHarmonicType">cartesian</scalar>
      <list dictRef="compchem:basisSetContraction" id="basisSetContraction">
        <scalar dataType="xsd:string" dictRef="compchem:basisSetShell">s</scalar>
        <array size="3" dataType="fpx:real" dictRef="compchem:basisSetExponent">1.301000000000e1 1.962000000000e0
4.446000000000e-1</array>
        <array size="3" dataType="fpx:real" dictRef="compchem:basisSetCoefficient">1.968500000000e-2 1.379770000000e-1
4.781480000000e-1</array>
      </list>
      <list dictRef="compchem:basisSetContraction" id="basisSetContraction">
        <scalar dataType="xsd:string" dictRef="compchem:basisSetShell">s</scalar>
        <array size="1" dataType="fpx:real" dictRef="compchem:basisSetExponent">1.220000000000e-1</array>
        <array size="1" dataType="fpx:real" dictRef="compchem:basisSetCoefficient">1.000000000000e0</array>
      </list>
      <list dictRef="compchem:basisSetContraction" id="basisSetContraction">
        <scalar dataType="xsd:string" dictRef="compchem:basisSetShell">p</scalar>
        <array size="1" dataType="fpx:real" dictRef="compchem:basisSetExponent">7.270000000000e-1</array>
        <array size="1" dataType="fpx:real" dictRef="compchem:basisSetCoefficient">1.000000000000e0</array>
</list>
```

Workflow and dataflow

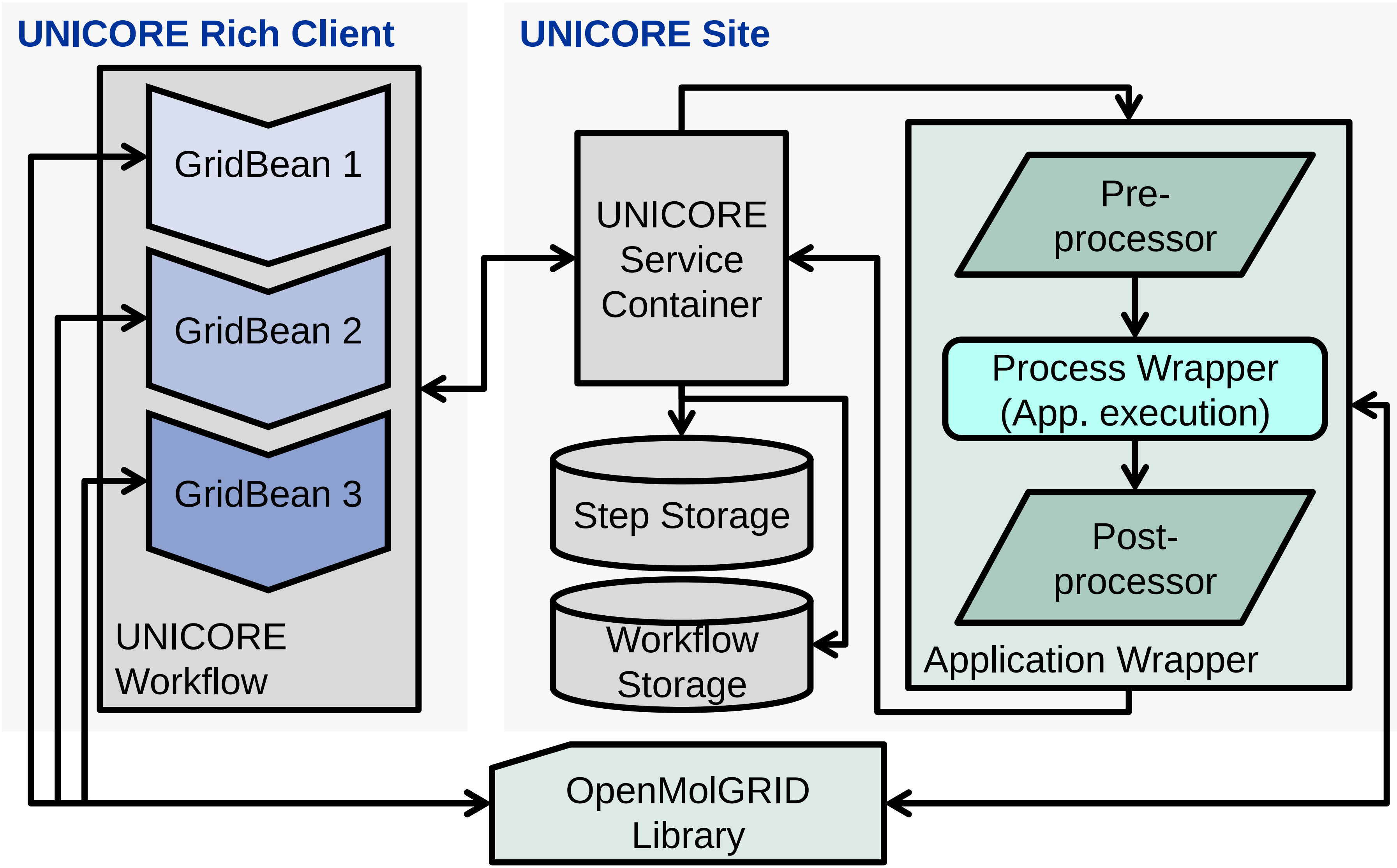


OpenMolGRID library

- The OpenMolGRID project (http://www.openmolgrid.org) has provided solutions for
  - Chemical data management and process automation
  - Automatic QSAR and molecular engineering
- Our developments:
  - Data model implemented on top of OpenMolGRID library
  - Extension with classes in the process and format packages
  - Access classes written for different individual applications
  - Classes for CML as storage backend developed

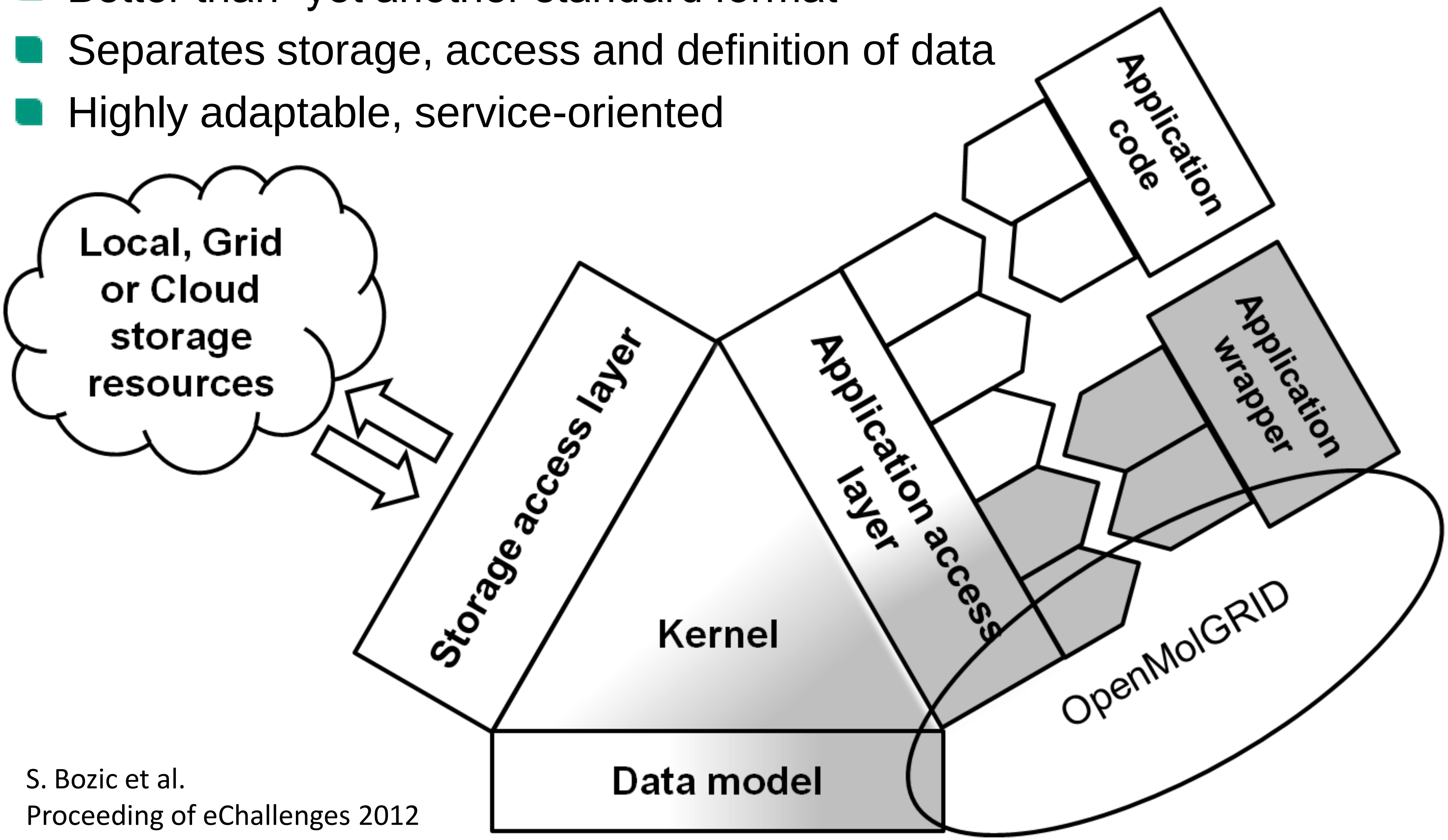
Supported application or format	Class name
CML	CMLReader, CMLWriter
PDB	PDBReader
MOPAC	MOPAC2009Parser, MOPACInputWriter
ToFeT	ToFeTXyzWriter

Platform Architecture



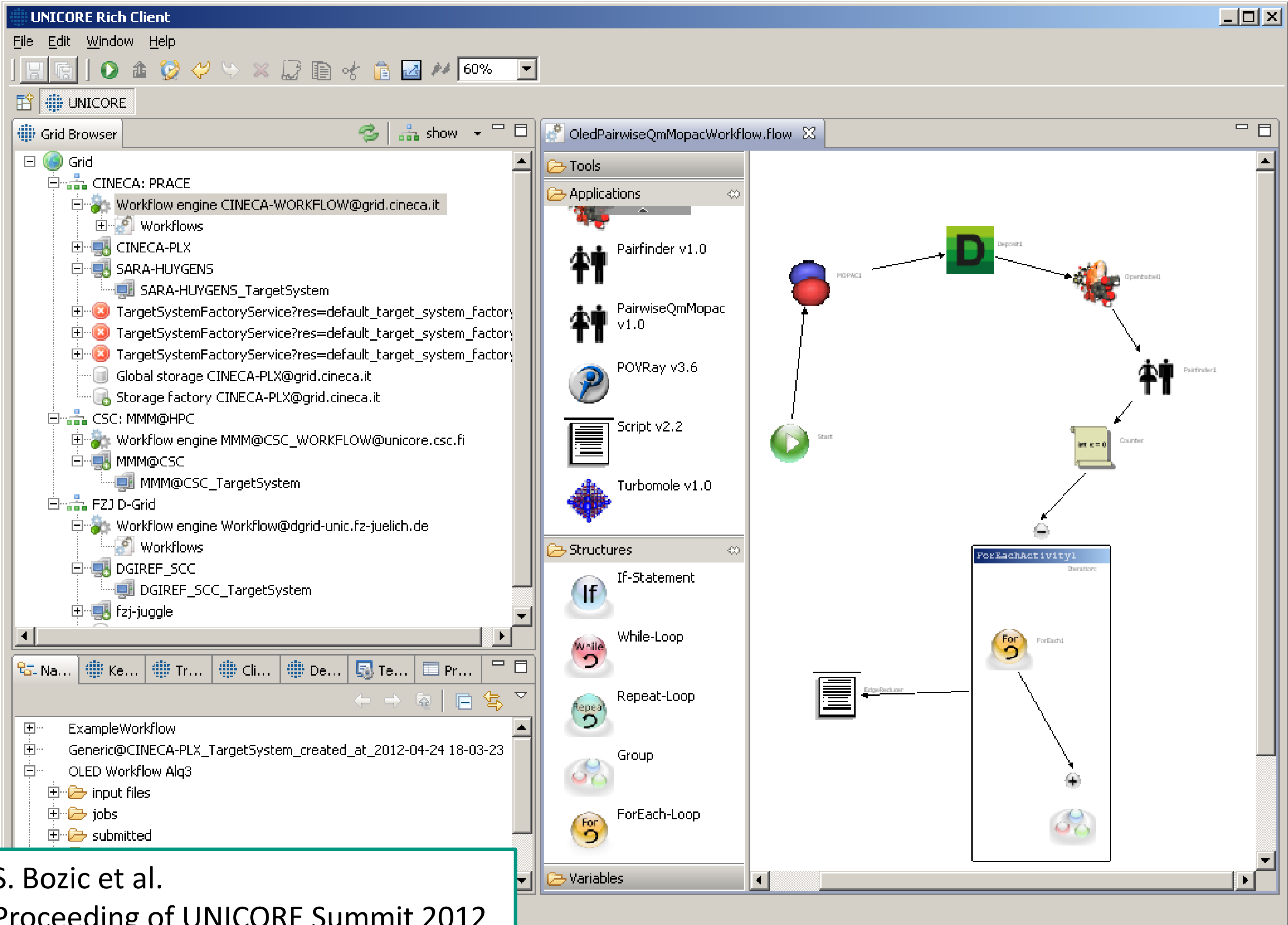
Data model driven approach

- Better than “yet another standard format”
- Separates storage, access and definition of data
- Highly adaptable, service-oriented



S. Bozic et al.
Proceeding of eChallenges 2012

Proof of principle: OLED workflow



S. Bozic et al.
Proceeding of UNICORE Summit 2012

Conclusion and outlook

- SOA based platform for multiscale materials modelling
- Proof of Principle:
  OLED workflow and simulation
- Data model driven dataflow management
- Adopted the CML standard
- Workflows using standard components require common data model
- Solution distributed as a software bundle and deployed as a service

- Realize workflows for further multiscale materials modelling applications
- Extend the CML *compchem* dictionary
- Implement the remaining part of the dataflow manager
- Create cross-links between communities of scientists, industrial customers, software engineers and resource providers and support their collaboration
- Make communities familiar with the solution to increase its acceptance and enable adoption

Acknowledgement

- All consortium partners in MMM@HPC
- Funding from the EC

The project MMM@HPC is funded by the 7th Framework Programme of the European Commission within the Research Infrastructures with grant agreement number RI-261594.

- Partner projects, supporting infrastructures and software providers