Molecular Dynamics Science Gateway with Vine Toolkit providing Unicore middleware support

Piotr Dziubecki, Piotr Grabowski, Tomasz Kuczyński, Krzysztof Kurowski, Dawid Szejnfeld, Dominik Tarnawczyk, Małgorzata Wolniewicz

PSNC
Agenda

- Vine Toolkit introduction
- Vine Toolkit – features
- Web portal with Vine Toolkit
- Bundled components
- UNICORE 6 plugin
- OGF Standards (JSDL, BES)
- Molecular Dynamics Science Gateway
- Application registry
- Software stack and technologies
- User interface features
- Nano-Science Gateway
- Towards modern web applications (Facebook, iGoogle)
- Future plans

UNICORE Summit 2011
Vine Toolkit introduction

• A modular and extensible Java/Flex based framework

• Derived from Grid Portlets for GridSphere

• Developed within EU funded projects: OMII-Europe, BEinGRID, HPC-Europa2

• Currently is being applied to the Polish NGI portal during the PL-Grid - Polish infrastructural project
Vine Toolkit - features

• Integration with different portal frameworks, including well known web products like GridSphere and Liferay

• Ant based installer; automatic, simple installation integrated with svn

• Installation support for Tomcat 5/5.5/6.0

• Advanced BlazeDs data services improve the client-server communication to develop highly interactive and dynamic web applications
Vine Toolkit - features

- **Adobe Flex/Flash** technology allows creating advanced and sophisticated web interfaces similar to many stand-alone GUIs
- Other web GUI technologies could be supported (i.e. HTML, JavaScript with Ajax support)
- Uniform common API exposed to the end user which **abstracts** various middleware implementations
- Extensible model for executing tasks (every action is persisted as **task**)
- Generic **resource based model** - any services and data sources can be integrated with web applications using high-level APIs
Web portal with Vine Toolkit

**GUI** (layout + portlets)

- Server: Apache Tomcat
- Portal: Liferay, Gridsphere etc.
- Users, Authentication, Authorization, Layout management, CMS

**Vine Toolkit**
- Business logic,
  - Integration with portal framework,
  - Extensible API

**BlazeDs**
- Req/Res API

**API**
- Flash (Flex), HTML, CSS, JSP

**Plugins for different low-level, external services:**
- advanced database sources,
- data management services – i.e. UNICORE SMS,
- web services,
- HPC services – i.e. UNICORE/X or BES
Bundled components

- User / Roles / Application / Properties Manager - administrative tools

- Login / Registration components plugin based with a support for i.e. Active Directory, MyProxy etc.

- Resource manager – configuration tool for the domain registry

- File browser component (support for the internal Portal File System and various file system protocols like GridFTP, LFC, SRM, UNICORE SMS, iRODS)
Bundled components

- Job manager (submission, monitoring)
  generic component (based on JSDL) – support for **UNICORE 6**, Globus GT4, gLite 3, QocCosGrid, GRIA

- Credential manager (including MyProxy support)
  – proxy certificates used in portal

- Resource browser – information services client for MDS in GT4

- GSI-SSHTerm applet – SSH console available in the portal

- Set of example components demonstrating Vine's UI features
UNICORE 6 plugin

- Proxy certificates used so proxy certificate UNICORE Gateway extension is required – other middlewares also use them
- It is possible to configure also common portal certificate instead
- SAML Assertions could be possibly used in future – Java Web Start or Java Applet to generate user SAML Assertion (lifetime? How often user should create it and upload? It is passed to portal – weak point?)
- Plugin uses UCC jars – UNICORE/X and BES access possible
- Some fixes related to inner use of Ehcache, some classes overloaded, there were some multiuser access problems
OGF standards

• **JSDL v1.0 support**
  – Application extensions:
    • **POSIXApplication** (default JSDL extension)
    • **HPCProfileApplication** (used with BES services)
    • **SPMDApplication** (translated with XSLT to define mpi jobs in GT4)

• **BES v1.0 (Basic Execution Service)**
  – developed during OMII-Europe project and reimplemented later to support full HPCProfileApplication JSDL extension
  – tested with many BES compliant services like Unicore BES, CREAM-BES Computing Element, GT4 BES service, Chinese CROWN BES metascheduler service

• Vine Plugins use XSLT translation if JSDL is not supported (Glite WMS, Globus GT4, proprietary CE resource)
Molecular Dynamics Science Gateway

- The whole system is meant to solve molecular dynamics (MD) simulation problems
- Offers a profiled web interface to the NAMD application
- User uses external tool to build molecules like peptides for example (to create pdb, psf input files)
- Ability to load ready to use models from external databases like Protein Data Bank
- Input files like psf should be editable directly in the portal
- Number of parameters to set before simulation
- NAMD application exposed for end users on grid infrastructure
- Parallel optimization (in general limited) vs multiple instance of problems (currently desired by end users)
- Goal: system geometry optimization (or) and simulation by molecular dynamics method (hours to days)
- Outputs: charts from generated log files: total energy (ETOTAL), kinetic energy (EKINETIC), bond, angle, ELEC, VDW etc.
- Still at early stage...but already tested with UNICORE
Application registry

- Vine Application Registry and UI component
- Scientific applications and simulations are mapped to logical entities
- Full history of simulation runs with parameters and job descriptions – problem library
- Log scrapping mechanism, which let us visualize the partial result in the realtime
Software stack and technologies

<table>
<thead>
<tr>
<th>Software stack</th>
<th>Technologies</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAMD Web Client</td>
<td>Adobe Flex / BlazeDS</td>
</tr>
<tr>
<td>Portal</td>
<td>Java / Portlets</td>
</tr>
<tr>
<td>Vine Toolkit framework</td>
<td>Java / Web Services</td>
</tr>
<tr>
<td>UNICORE middleware / other HPC resources and data services</td>
<td>Web Services / product specific implementations</td>
</tr>
</tbody>
</table>
User interface features

Input Files
File browser component

MD Simulations
New task
Create new task based on the existing simulation

Simulation
Sample 1
Sample 2
Sample 3
Sample 4
Sample 5
Sample 6
Sample 7
Sample 8
Sample 9
Sample 10

Status
Sample 1
Sample 2
Sample 3
Sample 4
Sample 5
Sample 6
Sample 7
Sample 8
Sample 9
Sample 10

MD Simulations - new task
Upload files (pdb/psf)
Start simulation

alanina.pdb
alanina.png
alanina.psf
alanina_opt.log
alanina_opt.namd

Optimization has been set to the 5000 cycles, dynamics to the 10000 - they are very short simulations.

Get optimization to the 5k 6000 steps in the 298K temperature T=298K
Accordingly for the molecular dynamics case symbol md10k, stands for the 10000 steps performed by the simulation.
NAMD Client screenshots - .namd file configuration
# NAMD Client screenshots – further configuration

## Input files

- **coordinates**: alanina.pdb
- **structure**: alanina.psf
- **parameters**: par_all27_prot_lipid_na.inp

## Output files

- **set output**: alanina_o5k_t298
- **output name**: $output
- **dcdFile**: $output.dcd
- **xstFile**: $output.xst
- **dcdFreq**: 50
- **xstFreq**: 50
- **binaryoutput**: no
- **binaryrestart**: no
- **outputEnergies**: 100
- **restartfreq**: 100
NAMD Client screenshots – .namd file ready to use

# input
coordinates /home/qba/Pulpit/przyklady/mala_mol/alanina.pdb
structure /home/qba/Pulpit/przyklady/mala_mol/alanina.psf
parameters /usr/local/lib/mdPlugins/noarch/tdr/read CharmmPar1.1/par_all27_prot_lipid_na.inp
paratype Charmm on

# output
set output /home/qba/Pulpit/przyklady/mala_mol/alanina.o5k.t298
outputname $output
dcdfile $(output).dcd
xstFile $(output).xst
dcdfreq 50
xstFreq 50
binaryoutput no
binaryrestart no
outputEnergies 100
restartfreq 1000
### NAMD Client screenshots – pdb input file

<table>
<thead>
<tr>
<th>REMARK</th>
<th>original generated coordinate pdb file</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATOM 1</td>
<td>N ALA X 2 0.000 0.000 0.000 1.00 0.00 P1 N</td>
</tr>
<tr>
<td>ATOM 2</td>
<td>HT1 ALA X 2 -0.326 -0.946 -0.000 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 3</td>
<td>HT2 ALA X 2 -0.326 0.473 -0.819 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 4</td>
<td>HT3 ALA X 2 -0.326 0.473 -0.819 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 5</td>
<td>CA ALA X 2 1.450 0.000 0.000 1.00 0.00 P1 C</td>
</tr>
<tr>
<td>ATOM 6</td>
<td>HA ALA X 2 1.827 -0.479 0.897 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 7</td>
<td>CB ALA X 2 1.969 -0.670 -1.262 1.00 0.00 P1 C</td>
</tr>
<tr>
<td>ATOM 8</td>
<td>HB1 ALA X 2 1.629 -1.727 -1.289 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 9</td>
<td>HB2 ALA X 2 1.581 -0.163 -2.172 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 10</td>
<td>HB3 ALA X 2 3.080 -0.665 -1.301 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 11</td>
<td>C ALA X 2 2.010 1.413 0.000 1.00 0.00 P1 C</td>
</tr>
<tr>
<td>ATOM 12</td>
<td>O ALA X 2 2.911 1.748 0.767 1.00 0.00 P1 O</td>
</tr>
<tr>
<td>ATOM 13</td>
<td>N ALA X 3 1.488 2.280 -0.853 1.00 0.00 P1 N</td>
</tr>
<tr>
<td>ATOM 14</td>
<td>HN ALA X 3 0.756 2.066 -1.509 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 15</td>
<td>CA ALA X 3 1.981 3.643 -0.909 1.00 0.00 P1 C</td>
</tr>
<tr>
<td>ATOM 16</td>
<td>HA ALA X 3 3.033 3.656 -1.629 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 17</td>
<td>CB ALA X 3 1.147 4.464 -1.880 1.00 0.00 P1 C</td>
</tr>
<tr>
<td>ATOM 18</td>
<td>HB1 ALA X 3 1.236 4.036 -2.901 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 19</td>
<td>HB2 ALA X 3 0.072 4.449 -1.595 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 20</td>
<td>HB3 ALA X 3 1.488 5.521 -1.920 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 21</td>
<td>C ALA X 3 1.985 4.326 0.444 1.00 0.00 P1 C</td>
</tr>
<tr>
<td>ATOM 22</td>
<td>O ALA X 3 2.801 4.963 0.924 1.00 0.00 P1 O</td>
</tr>
<tr>
<td>ATOM 23</td>
<td>N ALA X 4 0.710 4.211 1.093 1.00 0.00 P1 N</td>
</tr>
<tr>
<td>ATOM 24</td>
<td>HN ALA X 4 -0.084 3.708 0.753 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 25</td>
<td>CA ALA X 4 0.651 4.841 2.398 1.00 0.00 P1 C</td>
</tr>
<tr>
<td>ATOM 26</td>
<td>HA ALA X 4 0.659 5.915 2.303 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 27</td>
<td>CB ALA X 4 -0.809 4.482 2.976 1.00 0.00 P1 C</td>
</tr>
<tr>
<td>ATOM 28</td>
<td>HB1 ALA X 4 -1.621 4.826 2.312 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 29</td>
<td>HB2 ALA X 4 -0.911 3.358 3.064 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 30</td>
<td>HB3 ALA X 4 -0.959 4.913 3.981 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 31</td>
<td>C ALA X 4 1.591 4.371 3.391 1.00 0.00 P1 C</td>
</tr>
<tr>
<td>ATOM 32</td>
<td>O ALA X 4 2.212 5.167 4.085 1.00 0.00 P1 O</td>
</tr>
<tr>
<td>ATOM 33</td>
<td>N ALA X 5 1.818 3.963 4.463 1.00 0.00 P1 N</td>
</tr>
<tr>
<td>ATOM 34</td>
<td>HN ALA X 5 1.346 2.370 2.918 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 35</td>
<td>CA ALA X 5 2.809 2.556 4.392 1.00 0.00 P1 C</td>
</tr>
<tr>
<td>ATOM 36</td>
<td>HA ALA X 5 2.539 2.814 5.410 0.00 0.00 P1 H</td>
</tr>
<tr>
<td>ATOM 37</td>
<td>CB ALA X 5 2.970 1.855 4.290 1.00 0.00 P1 C</td>
</tr>
</tbody>
</table>
Nano-Science Gateway

- Large-scale simulation studies based on **Density Functional Theory (DFT)** and **Many-Body Perturbation Theory**
  - Command line applications fired through **CosQosGrid** middleware (other could be used like UNICORE 6) as MPI application in batch mode
  - **ABINIT** and **Quantum Espresso** both allows:
    - finding the total energy
    - finding charge density and electronic structure of systems made of electrons and nuclei within Density Functional Theory (DFT), using pseudopotentials and a planewave basis
    - optimize the geometry, perform molecular dynamics simulations, generate dynamical matrices, Born effective charges, and dielectric tensors according to the DFT forces and stresses
- To hide the complexity and provide a web-based collaborative access to **ABINIT** we created many **new rich web applications** using **Vine Toolkit** and **Adobe Flex**
  - **Simple and advanced mode** (user manually edits **ABINIT** input file)
  - **Dynamic charts** during computations and after completion (convergence of the relative difference between subsequent computation iterations, density of electronic states - DOS)
- Successfully presented at the **NANO 2010 workshop** attached to the 4th National Conference on Nanotechnology, another workshop just took place in July 2011
Nano-Science Gateway main view
NanoEditor

- Enables creation of nano structures, their visualization and modifications
- Consists of visual editor (further called NanoEditor) and interface that allows users to enter input data (further called NanoBuilder)
- NanoBuilder functionality covers:
  - creating crystal geometry using such parameters like: space group by its Hermann-Mauguin symbol, lattice vectors lengths, lattice angles, number of unit cells in each direction and telling whether unit cell should be found
  - Atoms could be added in a simple way, user can define their positions in Angstrom units
  - uses Spacegroup application from Elk software package (managed locally by the Vine itself and is made available through standard Job Management API)
  - generates output files which are used later in NanoEditor
- NanoEditor is a graphical tool that enables creating and editing super-cells and generating output files that can be further used in the Abinit package
  - cell replication
  - changing positions of certain atoms or groups of atoms
  - addition of atoms
  - defining the vacuum between cells
  - based on Papervision3D library
NanoEditor
Towards modern web solutions

- Other Liferay based portals
- Social networks like Facebook
- Content aggregators, personalized web pages like iGoogle
- Vine web applications could be integrated into third part web infrastructures almost seamlessly
- We have tested Vine applications in a range of external web systems (starting from pure HTML pages, through Wiki, various CMS solutions, other instance of the Liferay, ending on Facebook and iGoogle)
- *RemoteVineGeneratorApp* - generate ready to use integration packages for various platforms (war packages)
- It is possible to set up Vine to automatically create an account for the new users (Single sign-up - SSU)
- Single sign-on (SSO) also supported
- SSO and SSU are implemented for Facebook and iGoogle currently
- Science Gateway components could be exposed in many places but controlled by central web server; other configurations possible
Facebook example
Future plans

• **PL-Grid National Grid Infrastructure** project
  – Main Liferay portal – production deployment, integration and support in 2011
  – Software validation through operational and security audits
  – More science gateways planned based on new end users requirements in domains like biotechnology or chemistry, etc.
  – Production SG deployments with QCG middleware developed by PSNC (although other middlewares also possible to use)

• Molecular Dynamics Gateway
  – Development continuation and support for new applications is planned (in case more end users will be involved)

• Beyond proxy certs - UNICORE SAML Assertion support in the future… (depends from ongoing actions and projects)

• Vine will be shipped with Serpens within Kepler release as an abstraction web service for grid infrastructures
Future plans

• Posible portals with QosCosGrid middleware support under the ongoing Infrastructure EU funded MAPPER project (Multiscale Applications on European e-Infrastructures) 2010-2013
  – closer collaboration with PRACE and EGI/EGI-InSPIRE partners possible
  – training sessions and workshops could be organized for PRACE and EGI users

• Production deployment of QCG middleware within infrastructural project PL-Grid on Polish NGI resources: PSNC, Cyfronet, TASK, (ICM maybe?)

• QCG production deployment during infrastructural MAPPER project (MoU with EGI, part of Task Force MAPPER-PRACE) – target sites LRZ (currently heavily tested), SARA

• Big impact from MAPPER communities from such domains: fusion, hydrology, physiology, nano-material science, computational biology

• Talks began with EMI, possible distribution of QCG middleware, Vine possibly could support EMI services
Any questions, comments or remarks are very welcome.

contact: dejw@man.poznan.pl

Vine Toolikt: http://vinetoolkit.org/

Vine users mailing list: vine-users@gforge.man.poznan.pl


Nano-Science Gateway: http://nano.man.poznan.pl

MAPPER: http://www.mapper-project.eu

QosCosGrid: http://www.qoscosgrid.org

Serpens: https://kepler-project.org/developers/interest-groups/distributed/serpens-suite-in-kepler