CloudBroker

High Performance Computing
Software as a Service

Integration in KNIME

CloudBroker GmbH
Technoparkstrasse 1, CH-8005 Zurich, Switzerland
Phone: +41 44 633 79 34
Email: info@cloudbroker.com
Web: http://www.cloudbroker.com
HPC Applications in KNIME

KNIME works great for constructing, executing and exchanging workflows:

• Focus is on high throughput computing and rich client experience

Current support for batch-oriented, number-crunching, high performance computing (HPC) applications:

• Application-specific nodes to connect to particular external HPC software
• KNIME Cluster Execution to run existing nodes on an SGE-based cluster
• Build your own interface to HPC software and resources using SSH, Java, Python, web service or similar nodes
• Implement your own solution
Problems with HPC Applications

You have to:
• Build, operate and update your own cluster infrastructure
• Buy long-term licenses and support external software
• Make sure that the needed KNIME nodes or implementations are available or developed

This means:
• Complex and time-consuming setup, maintenance and change procedures
• High initial costs for hardware and software
• Probably some own programming efforts
• Need for specialized personnel
• No flexibility for special projects, testing, varying load, etc.
Cloud Computing

Computer resources (infrastructure, platforms, software) provided as a service via the internet, on demand in a self-service manner

This means:
• Minimal initial investments in cost and time
• You only pay for what you actually use
• Nearly unlimited scalability

Your advantages:
• Increased flexibility
• Concentration on core competencies
• CapEx substituted by OpEx
• Over- or underutilization replaced by elastic scalability
Integration into KNIME

- KNIME workflow
- CloudBroker KNIME node
- CloudBroker Java API
- CloudBroker Platform
- Application A
- Application B
- Application C

Inputs: Cloud 1 → Cloud 2
Outputs: Cloud 2 → Cloud 3
Jobs: Cloud 1 → Cloud 2 → Cloud 3
Example Workflow

• Goal: Compute the quantum chemical dipole moments for a number of molecules
• Application: GAMESS (General Atomic and Molecular Electronic Structure System, http://www.msg.chem.iastate.edu/gamess)
• Data set: Some ligands from DUD (Directory of Useful Decoys, http://dud.docking.org), provided by KNIME
Advantages

- Access to a number of different HPC applications from various domains (life sciences, chemistry, engineering, etc.)
- Easy integration into any type of KNIME workflow
- No need to buy, install or operate any HPC hardware or software
- Scalable cloud infrastructure
- Pay per use charging model
- Just registration in the CloudBroker Platform needed
- Deployment of your application into the platform also possible
Availability

- CloudBroker Platform preview version available to selected users (http://www.cloudbroker.com/index.php/platform)
- Applications: GAMESS, BLAST, AutoDock, Gromacs, X! Tandem, OpenFOAM, Rosetta (upcoming)
- Clouds: Amazon Web Services, IBM Cloud (upcoming)
- CloudBroker KNIME node to be released together with next platform version (expected April 2011)
- Future plans: Possibility to execute whole KNIME workflows in the cloud via the platform
Thanks! – Questions?

Thanks to the CloudBroker Team and to KNIME!

Contact:
Wibke Sudholt, wibke.sudholt@cloudbroker.com