



PHOSPHORUS

Rapid Deployment of VS Workflows on PHOSPHORUS using Meta Scheduling Service

M. Shahid, Bjoern Hagemeyer

Fraunhofer Institute SCAI, Research Center Juelich.
(TNC 2009)

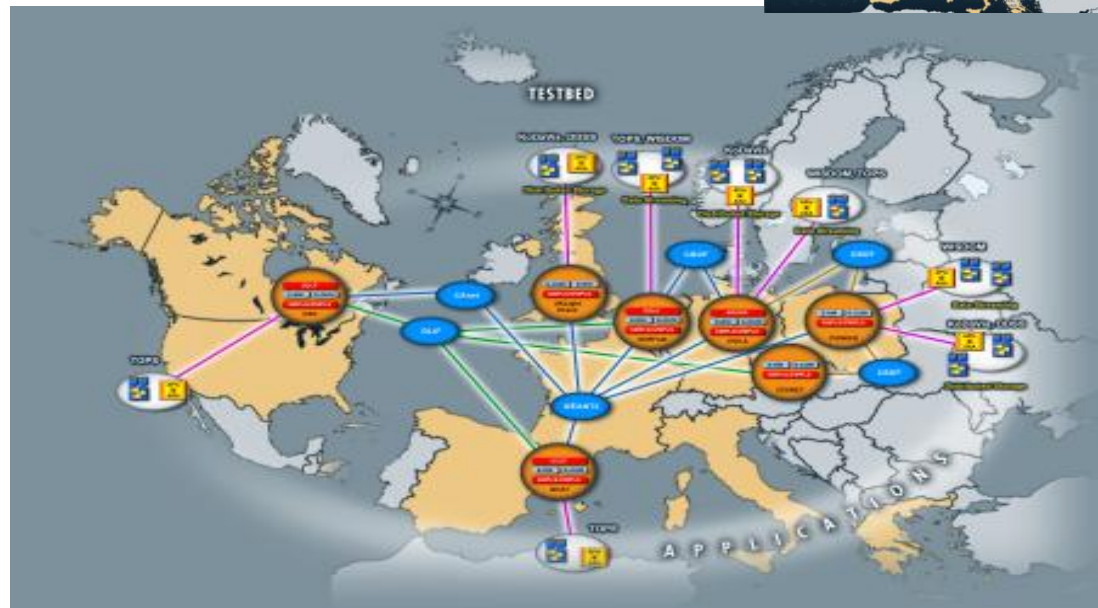


- Introduction and Motivation
 - Bioinformatics Applications on the Grid
 - PHOSPHORUS Testbed
 - Objectives within PHOSPHORUS
- Issues related to Grid VS deployment
- Development of a VS Framework in PHOSPHORUS
 - Virtual Screening Applications
 - vHTS Framework Design
- High throughput Virtual Screening Workflows
- Conclusions
- Example VS Deployment using UNICORE/MSS



- High computational and data storage demands
- Data security/privacy restrictions
- High Throughput data management & deployment on large-scale Grids
- Need of high level tools to enable e-Scientists to use Grids in an easy and transparent way

- Large capacity optical networks
- Satisfying e-Science HPC applications' (high computational & networking) demands...
- Needed: Advanced Grid-aware tools



Objectives within PHOSPHORUS



- Rapid deployment of VS applications to PHOSPHORUS environment
- Using user-friendly workflows based on UNICORE6 / MSS integration
- Efficient Stage-in and stage-out: easy distribution of input and output data
- Data distribution and collection without any data loss
- Support for post-processing and analyses



- Management and deployment of large number of jobs
- Scheduling policy: maximum resource utilization in the available time and space
- Pre/Post processing, management of huge amount of output data
- The production environment should provide automated and fault tolerant jobs and files management



- The amount of transferred data impacts on the overall Grid performance
- Efficient Data distribution on the Grid storage
- Speedy transfer between computing elements and Grid storage.
- Automated post processing: data mining

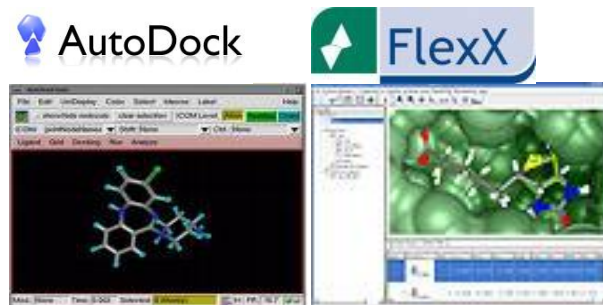


- Support large-scale deployment of virtual screening services in PHOSPHORUS
- Addressing the issues involved in deployment of complex VS workflows
- An extensible framework making the Grid more accessible to end users

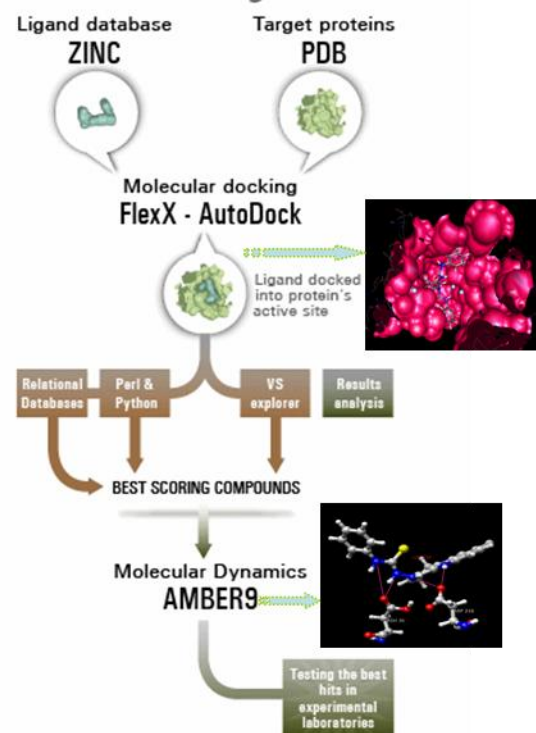
Virtual Screening by Molecular Docking



- CPU and Data intensive applications:
- FlexX (BioSolveIT)
 - Predicts geometry/binding free energy of protein ligand complex
- AutoDock (Scripps Research Institute)
 - Comparatively time-consuming, a single docking job takes ~ 30-60 minutes on a standard CPU.
- Amber (MD Simulations)
 - Package for simulations of biomolecules.
- INPUT/OUTPUT data:
 - Several GB to TB.



Virtual screening workflow





- Using UNICORE-6 Middleware Technology
 - Seamless, secure and transparent access to Grid
 - Ease of use, reduced complexity, increased security
- Using UNICORE Client extensions/plugins
 - Simplifying interactions with applications on the Grid
- Using Meta Scheduling Service
 - Further support to utilize Grid resources

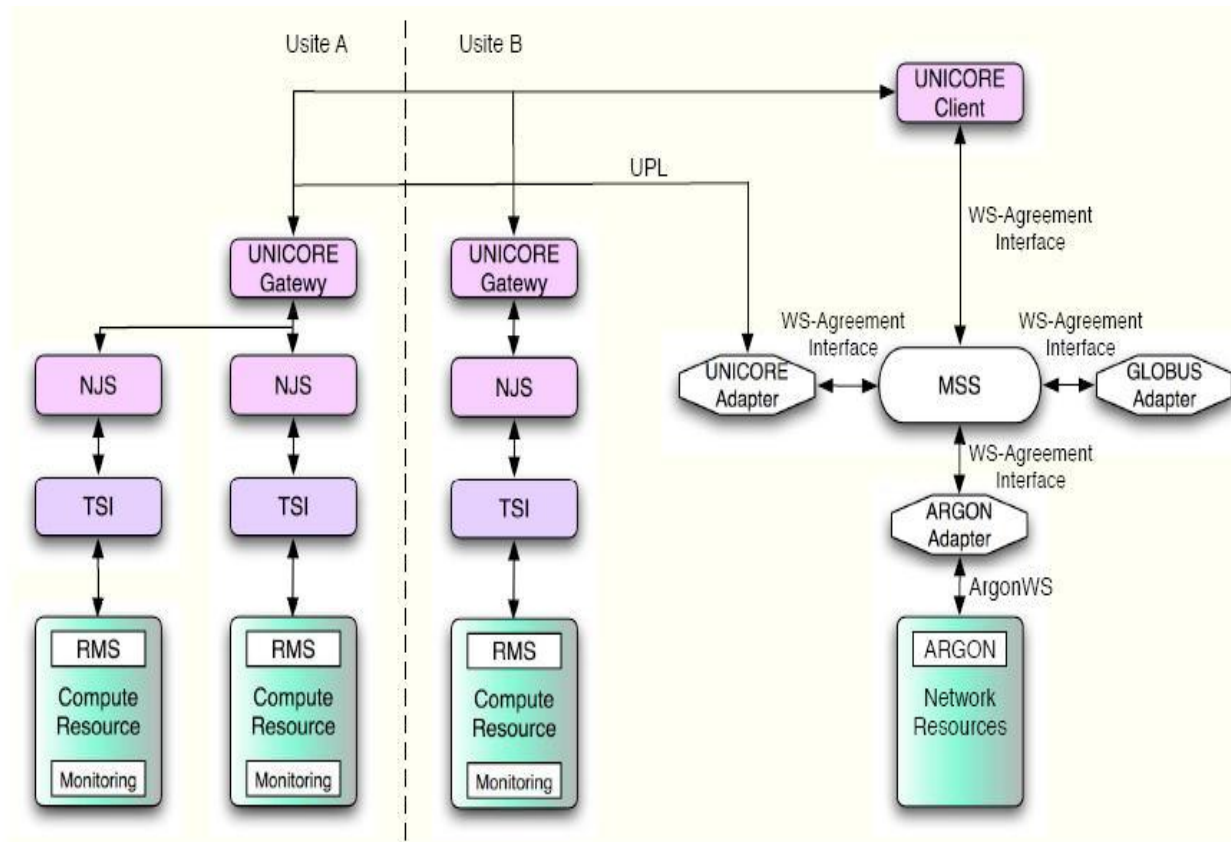


- UNICORE Expert Client uses eclipse RCP mechanism
- RCP is based on plug-in architecture that provides development of extendible components
- UNICORE Client: communication to UNICORE services
- The client extension provides interface to the Meta Scheduling Service
- Exploits job/data management mechanism of UNICORE client

Meta Scheduling Service



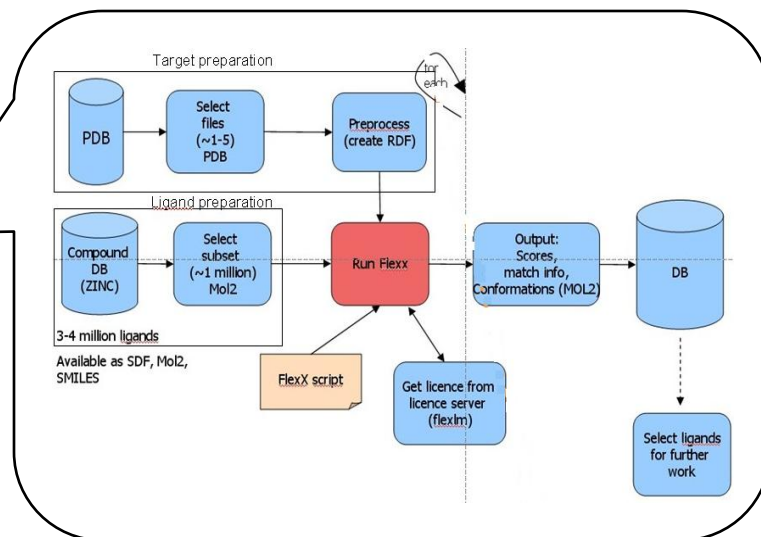
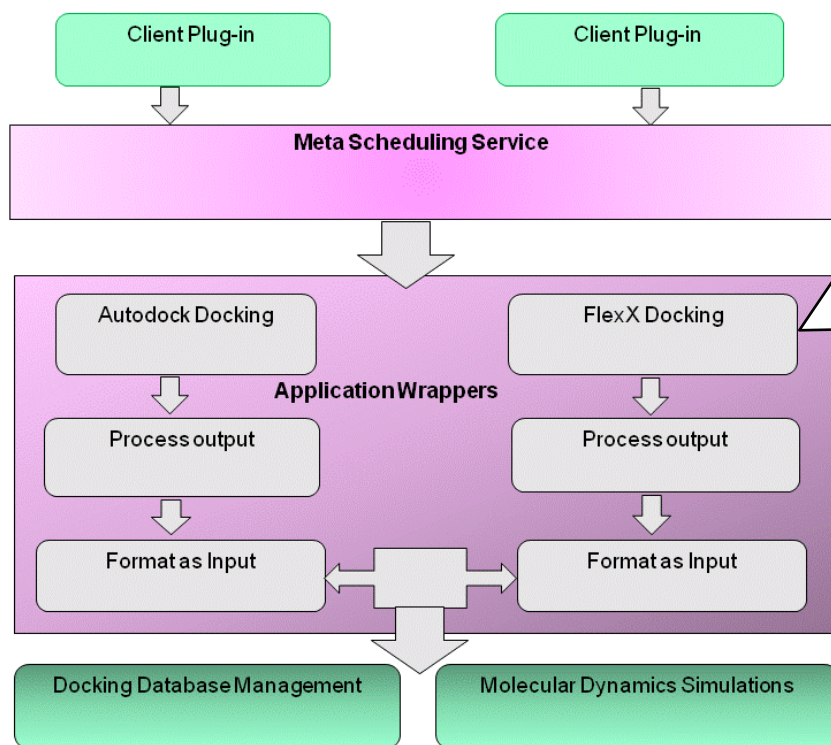
- Orchestration & co-ordination of Grid resources, SLA
- Supporting users to utilize maximum available resources
- Workload distribution...



High Throughput Virtual Screening Workflow



- VS Applications: FlexX, AutoDock...
- Fully automated workflow



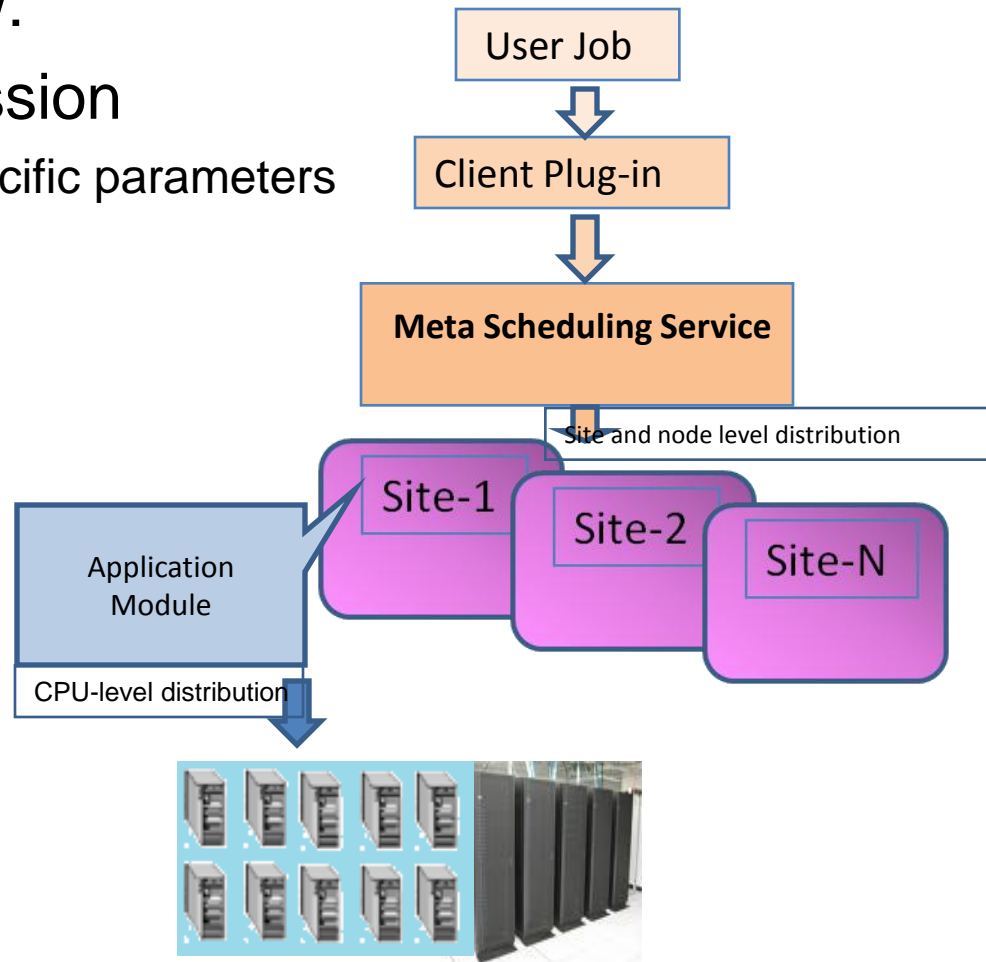
High Throughput Virtual Screening Workflow



Example VS Workflow:

One single job submission

- Setting up Application specific parameters
- Input/output specification



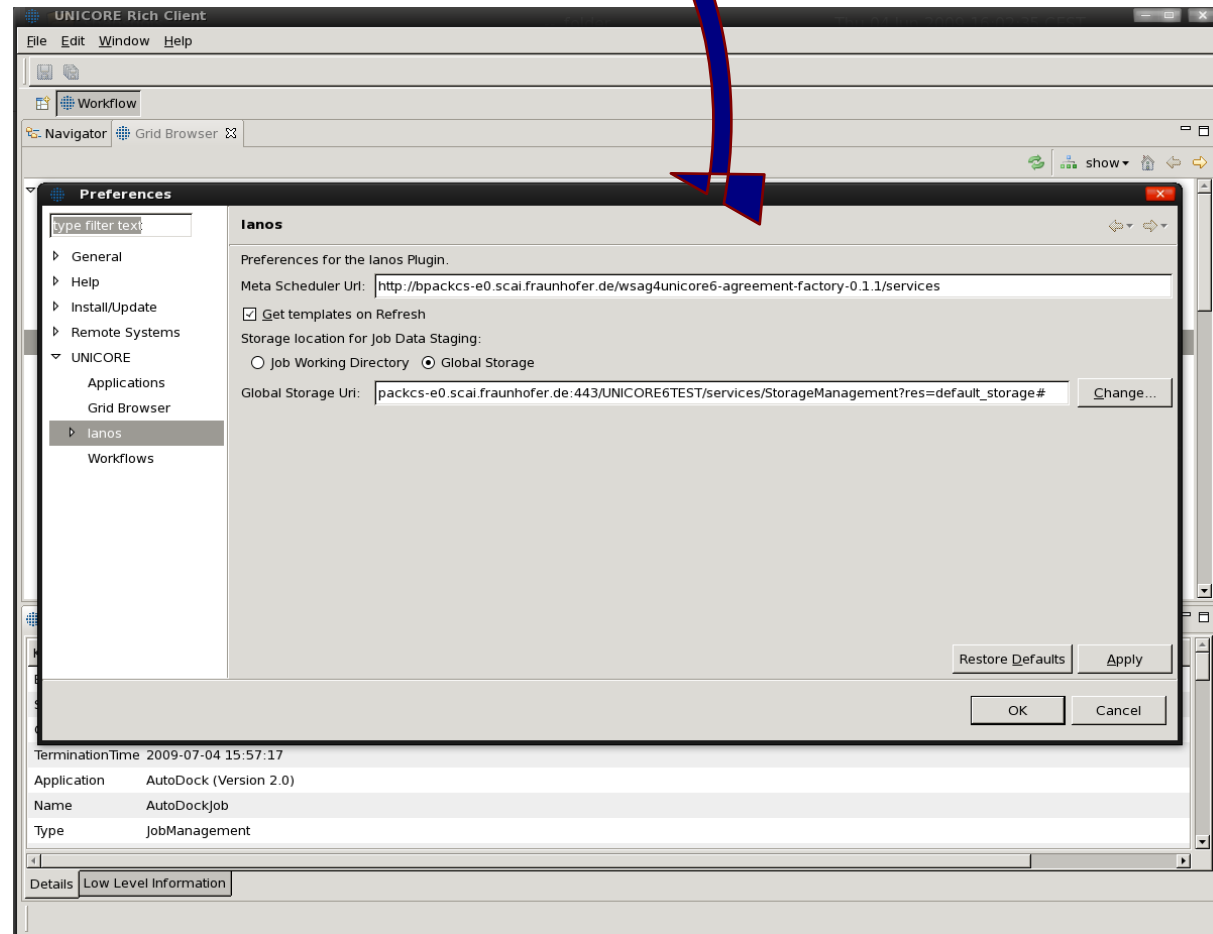


- Rapid deployment
 - High bandwidth input/output (PHOSPHORUS network layer)
- Better deployment efficiency with MSS than using UNICORE client plug-ins alone
- Simplifies, speeds up the complex virtual screening process in the Grid environment
- Flexibility and extensibility
- Allow easy integration of other tools & services

Example VS Deployment: UNICORE/MSS (1/9)



- Load MSS Client Plugin
- Provide MSS URL, Global Storage URL
 - Input/Output



Example VS Deployment: UNICORE/MSS (2/9)



The screenshot displays the UNICORE Rich Client interface. The main window is titled "UNICORE Rich Client" and features a menu bar with "File", "Edit", "Window", and "Help". Below the menu bar is a toolbar with icons for "Workflow", "Navigator", and "Grid Browser". The "Grid Browser" tab is active, showing a tree view of the grid resources. A red circle highlights the context menu for the "Grid" resource, which includes options: "add Registry", "refresh", "Add Meta Scheduling Service", "Global Storage at packages-eo.sca.fraunhofer.de", "Meta Scheduling Service", "JUCC", "Phosphorus_FZJ", "KoDaVisServiceFactory?res=default_kodaservice_factory", and "fraenkel".

The "Details" panel at the bottom shows the following information:

Key	Value
Name	Grid
URI	Grid
State	Ready
Type	{http://www.unicore.fzj.de}Grid

At the bottom of the interface, there are tabs for "Details" and "Low Level Information".

Example VS Deployment: UNICORE/MSS (3/9)



UNICORE Rich Client

File Edit Window Help

Workflow

Navigator Grid Browser

Grid

- packcs
- Meta Scheduling Service
- JUGGLE
 - add Registry
 - configure Credentials
 - edit address
 - details
 - remove Bookmark
 - create job
 - refresh
 - load job
 - rename
- Phosphor
- KoDAV
- fraenkel

Details Log Monitor Used Credentials

Key	Value
Number of Templates	5
Name	Meta Scheduling Service
Type	MetaSchedulingService
State	Ready
URI	http://bpackcs-e0.scai.fraunhofer.de
CurrentTime	unknown
TerminationTime	unknown

Details Low Level Information

Select an application

AutoDock v.2.0

FlexX v.2.0

Generic v.1.0

POVRay v.3.51

Script v.2.0

Download Applications

OK Cancel

Example VS Deployment: UNICORE/MSS (4/9)



UNICORE Rich Client

File Edit Window Help

Workflow

Navigator Grid Browser

Grid

- ▼ bpackcs
 - packcs-e0.scai.fraunhofer.de:443/UNICORE6TEST
 - Global storage at packcs-e0.scai.fraunhofer.de
- Meta Scheduling Service
- ▼ JUGGLE
 - Phosphorus_FZJ
 - KoDaVisServiceFactory?res=default_kodaservice_factory
 - fraenkel

Details Log Monitor Used Credentials Truststore Keystore Input for AutoDock

Application parameters

Submit job to MSS

Job Name: AutoDock

Receptor pdbqt File (.pdbqt): 1zxb_a

Multi-mol2 File (mol2): acpligands.mol2

Number of Energy Evaluations (ga_numeval): 25000

Number of GA Population (ga_pop): 50

Number of GA Runs (ga_run): 5

RMS Tolerance (rmstol): 2

☒ Flexible receptor docking

AutoDock-Input-Panel Files Variables Job Properties

Example VS Deployment: UNICORE/MSS (5/9)



UNICORE Rich Client

File Edit Window Help

Workflow

Navigator Grid Browser

Grid

- bpacscs
 - packcs-e0.scai.fraunhofer.de:443/UNICORE6TEST
 - Global storage at packcs-e0.scai.fraunhofer.de
- Meta Scheduling Service
- JUGGLE
 - Phosphorus_FZJ
 - KoDaVisServiceFactory?res=default_kodaservice_factory
- fraenkel

Details Log Monitor Used Credentials Truststore Keystore Input

Imports to job directory:

Name	Source Type	Source File(s)	File(s) in Job Directory
INPUT_FILE	Local_File	/home/shahid/1zxb_a_input.tar.gz	inputdata.tar.gz

Exports from job directory:

Name	File(s) in Job Directory	Destination Type	File(s) at Destination / File ID
OUTPUT_FILE	results.tar.gz	None	results.tar.gz
STANDARD_ERROR	stderr	None	
STANDARD_OUT	stdout	None	stdout

AutoDock-Input-Panel Files Variables Job Properties

Stage in

Stage out

Example VS Deployment: UNICORE/MSS (6/9)



UNICORE Rich Client

File Edit Window Help

Workflow

Navigator Grid Browser

Grid

- bpackcs
 - packcs-e0.scai.fraunhofer.de:443/UNICORE6TEST
 - UNICORE6TEST_TargetSystem
 - UNICORE6TEST_TargetSystem Home
 - UNICORE6TEST_TargetSystem WISDOM
 - AutoDock
 - Global storage at packcs-e0.scai.fraunhofer.de
 - Meta Scheduling Service
 - JUGGLE
 - Phosphorus_FZJ
 - KoDaVisServiceFactory?res=default_kodaservice_factory

Details Log Monitor Used Credentials Truststore Keystore

Key	Value
Installed Operating Systems	
Processor main memory	1024 Megabyte
CurrentTime	2009-06-04 11:44:08
TerminationTime	2012-04-29 15:20:56
Processors per node	4
Number of nodes	8
Processors architecture	<xml-fragment xmlns:idb="http://www.fz-juelich.de/unicore/xnjs/idb" xmlns:jsdl="http://schemas.ggf.org/jsdl/2005/11/jsdl" xmlns:typ="http://unigrids.c...><jsdl:CPUArchitectureName>x86</jsdl:CPUArchitectureName></xml-fragment>
State	Ready
Type	TargetSystemFactoryService
	WISDOM-PACK (Version 1.0)
	WISDOM-UNPACK (Version 1.0)

Details Low Level Information

waiting for job AutoDock to finish

Example VS Deployment: UNICORE/MSS (7/9)



UNICORE Rich Client

File Edit Window Help

Workflow

Navigator Grid Browser

Grid

- bpackcs
 - packcs-e0.scai.fraunhofer.de:443/UNICORE6TEST
 - UNICORE6TEST_TargetSystem
 - UNICORE6TEST_TargetSystem Home
 - UNICORE6TEST_TargetSystem WISDOM
 - AutoDock
 - Global storage at packcs-e0.scai.fraunhofer.de

Meta Scheduling Service

JUGGLE

- Phosphorus_FZJ
- KoDaVisServiceFactory?res=default_kodaservice_factory

fraenkel

Details Log Monitor Used Credentials Truststore Keystore

Key	Value
Number of Templates	5
Name	Meta Scheduling Service
Type	MetaSchedulingService
State	Ready
URI	http://bpackcs-e0.scai.fraunhofer.de/wsag4unicore6-agreement-factory-0.1.1/services
CurrentTime	unknown
TerminationTime	unknown

Details Low Level Information

Example VS Deployment: UNICORE/MSS (8/9)



UNICORE Rich Client

File Edit Window Help

Workflow

Navigator Grid Browser

acpigrans.moiz
AD4_parameters.dat
ligands
ligpdbqts
dpfs
dlgs
status.log
results.tar.gz
UNICORE_SCRIPT_EXIT_CODE
stdout
stderr
Global storage at packcs-e0.scai.fraunhofer.de
Meta Scheduling Service
JUGGLE
Phosphorus_FZJ

show

Details Log Monitor Used Credentials Truststore Keystore

Key	Value
Exit code	0
Submitted at	2009-06-04 16:06:07
CurrentTime	2009-06-04 16:08:06
TerminationTime	2009-07-04 16:06:11
Name	AutoDock
Application	AutoDock (Version 2.0)
Type	JobManagement
State	Ready
Job Status	SUCCESSFUL

Details Low Level Information

Example VS Deployment: UNICORE/MSS (9/9)



UNICORE Rich Client

File Edit Window Help

Workflow

Navigator Grid Browser

Grid

- bpaccks
 - packcs-e0.scai.fraunhofer.de:443/UNICORE
 - UNICORE6TEST_TargetSystem
 - UNICORE6TEST_TargetSystem Home
 - UNICORE6TEST_TargetSystem WISDOM
 - AutoDock
 - add Registry
 - show Outcomes
 - destroy
 - configure Credentials
 - details
 - restore Job definition
 - set termination time
 - refresh

Selection Needed

Select files for download

- ☒ /stdout (444 bytes)
- ☒ /stderr (3552 bytes)
- ☒ /results.tar.gz (1548532 bytes)

Select All Deselect All

OK Cancel

Details Log Monitor Used Credentials

Key	Value
Exit code	0
Submitted at	2009-06-04 16:06:07
CurrentTime	2009-06-04 16:08:06
TerminationTime	2009-07-04 16:06:11
Name	AutoDock
Application	AutoDock (Version 2.0)
Type	JobManagement
State	Ready
Job Status	SUCCESSFUL

Details Low Level Information



Thank You for your attention