

vl·e



virtual laboratory for e·science

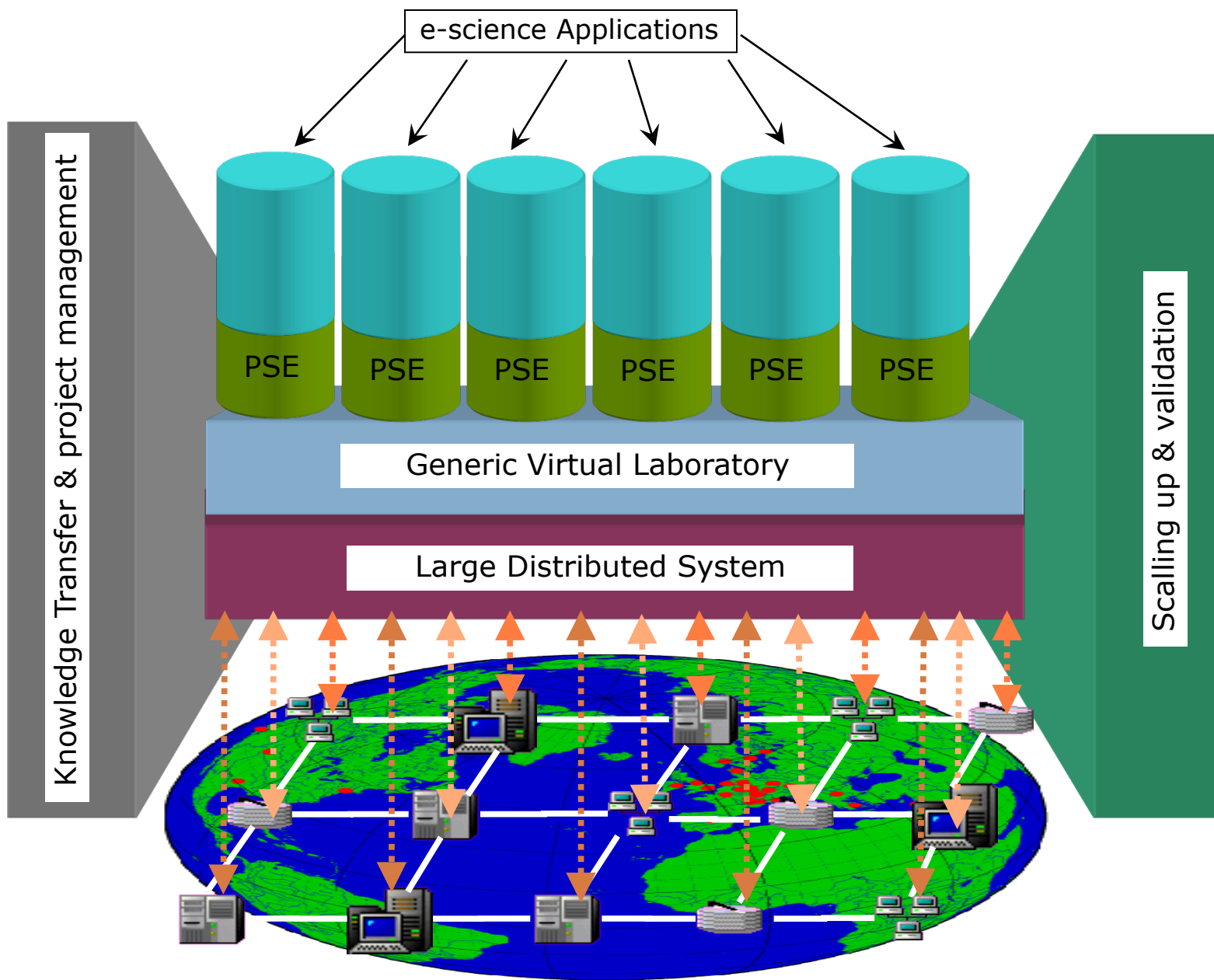
VL-e eScience Framework: from design to sharing

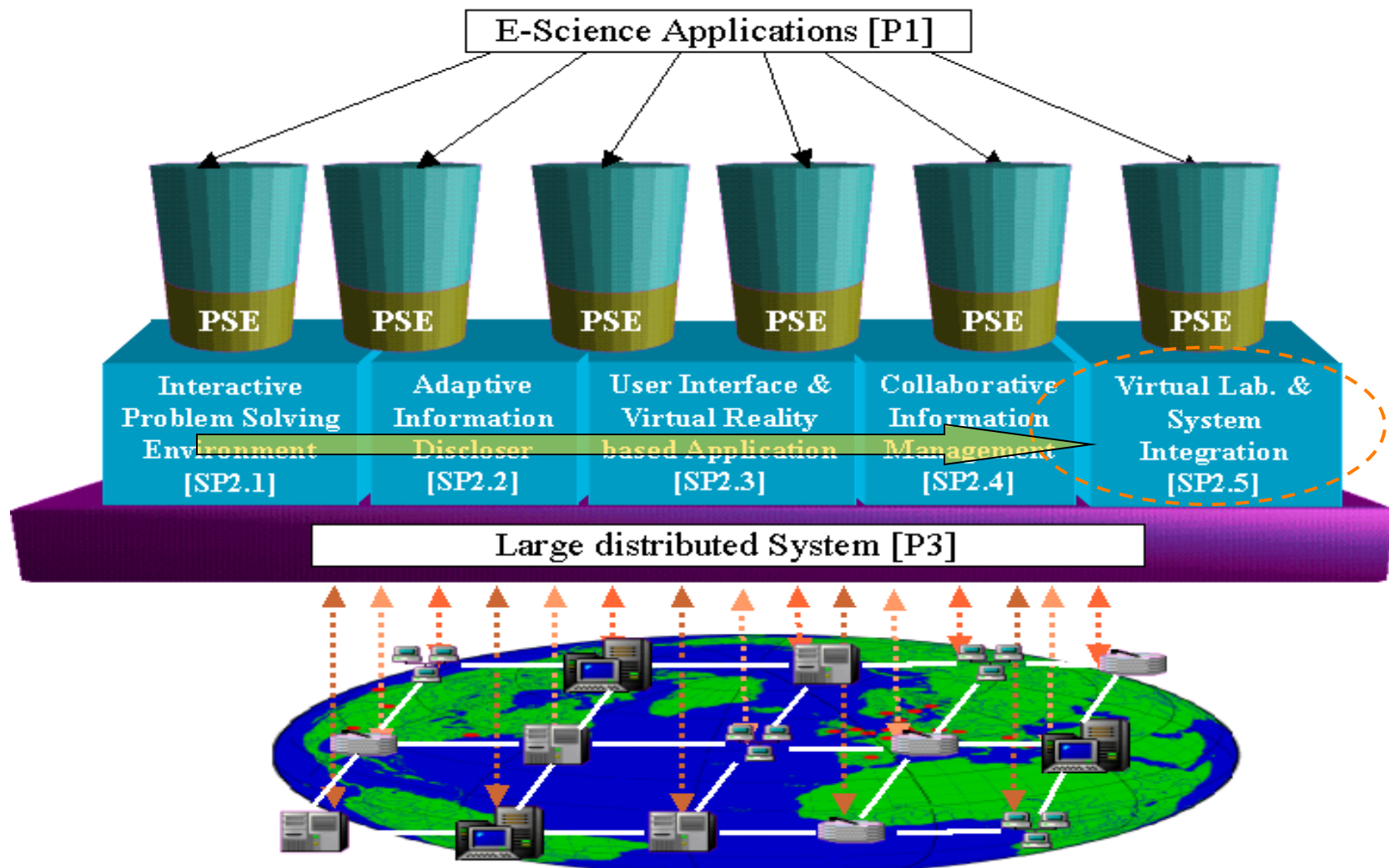
Adam Belloum
Institute of Informatics
University of Amsterdam

UvA



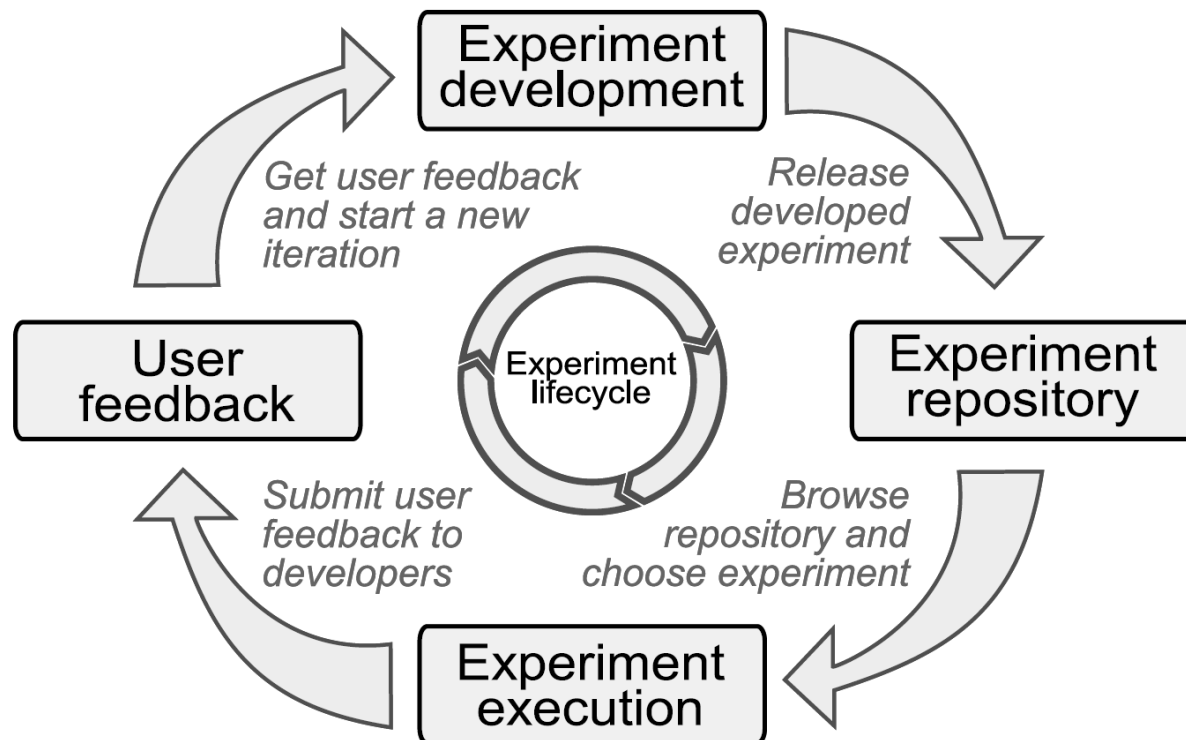
UNIVERSITEIT VAN AMSTERDAM



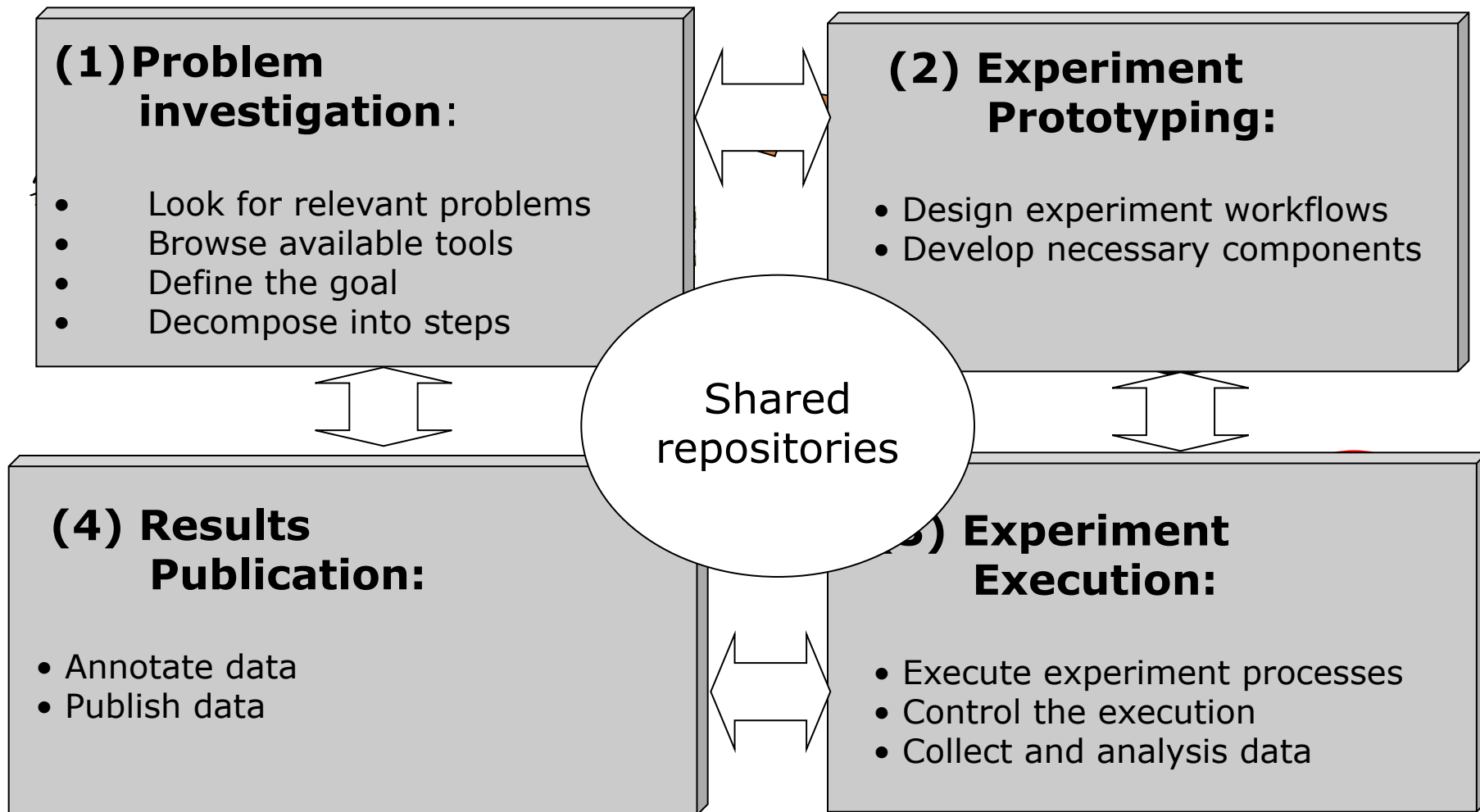


Experiment Lifecycle

- *Experiment Pipeline* - a collaborative planning and execution process that may create a new experiment



Complex Scientific experiments model





Objective

- The Objective of this work is to address the research issues related to building an **e-Science Framework** which will **support scientists to share their knowledge** and to perform specific experiments on grid resources

Keywords:

- Grid, Scientific workflow, SOA, provenance, interoperability

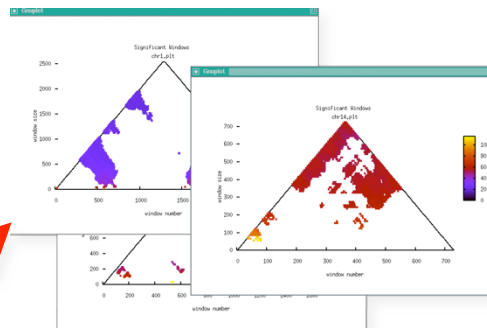
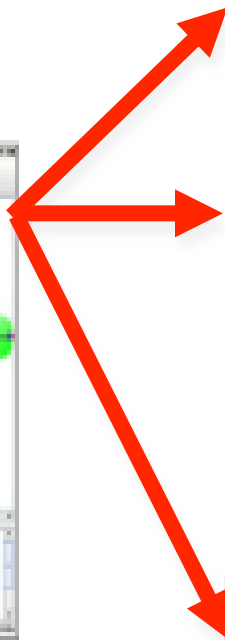
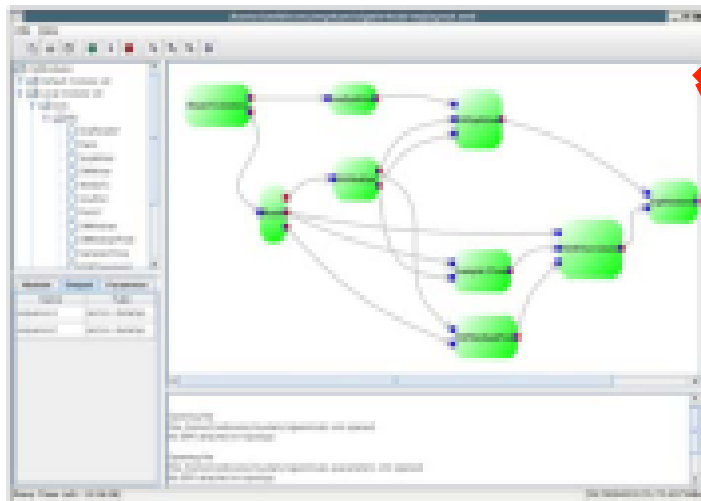
WS-VLAM e-Studio

- So far it is **simple program**, used to start the different **tools developed** in the VL-e project by the *Virtual Laboratory and system integration* group

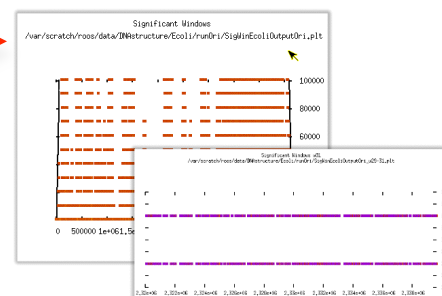


The Application Workflow ...

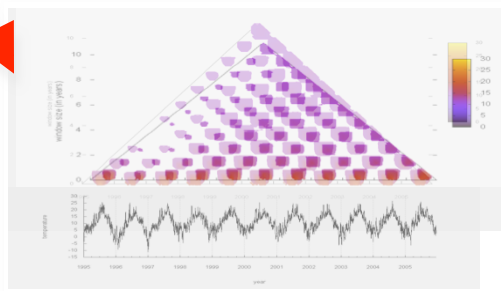
Developed in the context of the
Ridge-O-grammer Application use case
by Micro-array department UvA



Human
transcriptome
map



DNA curvature
of the
Escherichia coli
chromosome



Temperature in
Amsterdam

(1) Problem investigation

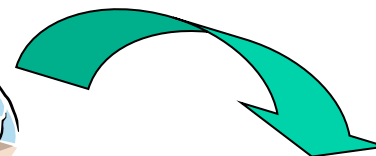
- Tools to improve **re-usability**, and **share of expertise**. These tools should allow:

- Advanced **search capabilities**

- HAMMER: **H**ybrid-b**A**sed **M**atch-**M**aker for **E**-**S**cience **R**esources
- Web Service Harvester (RPC Style WSDL)

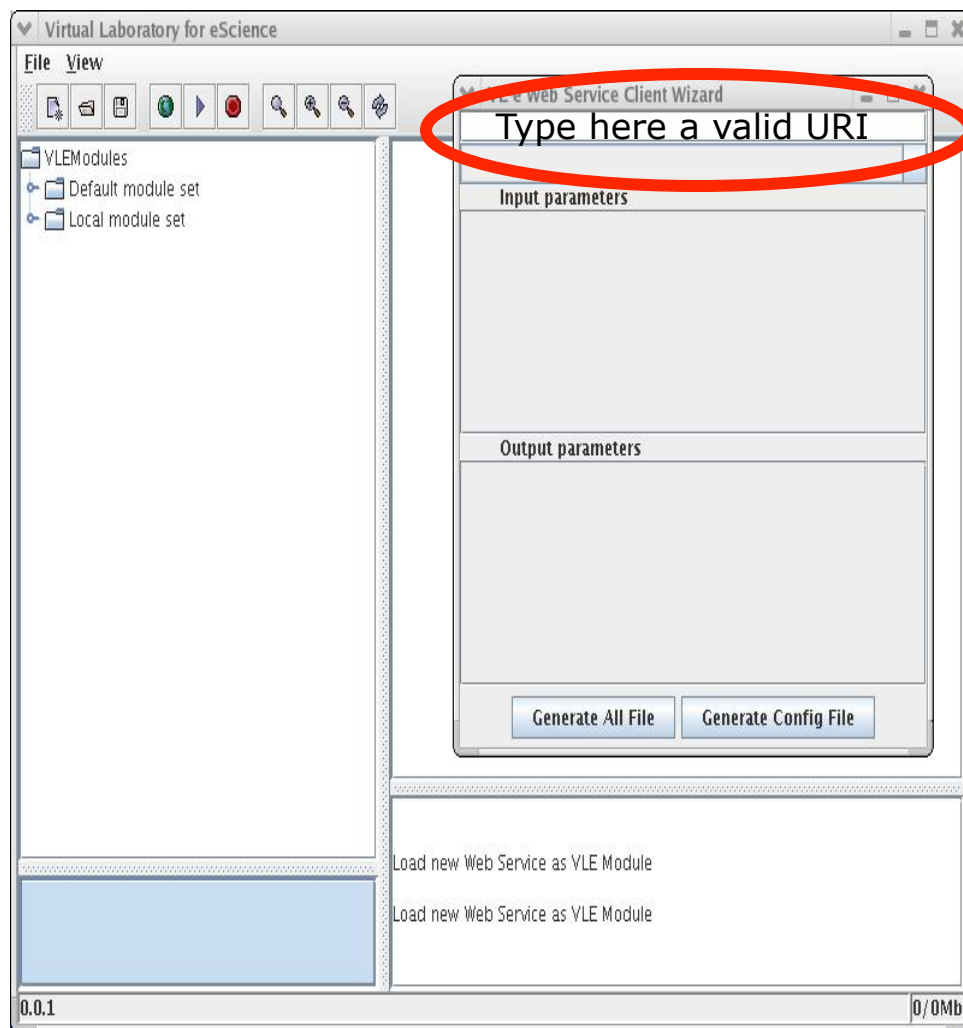


designing the experiment



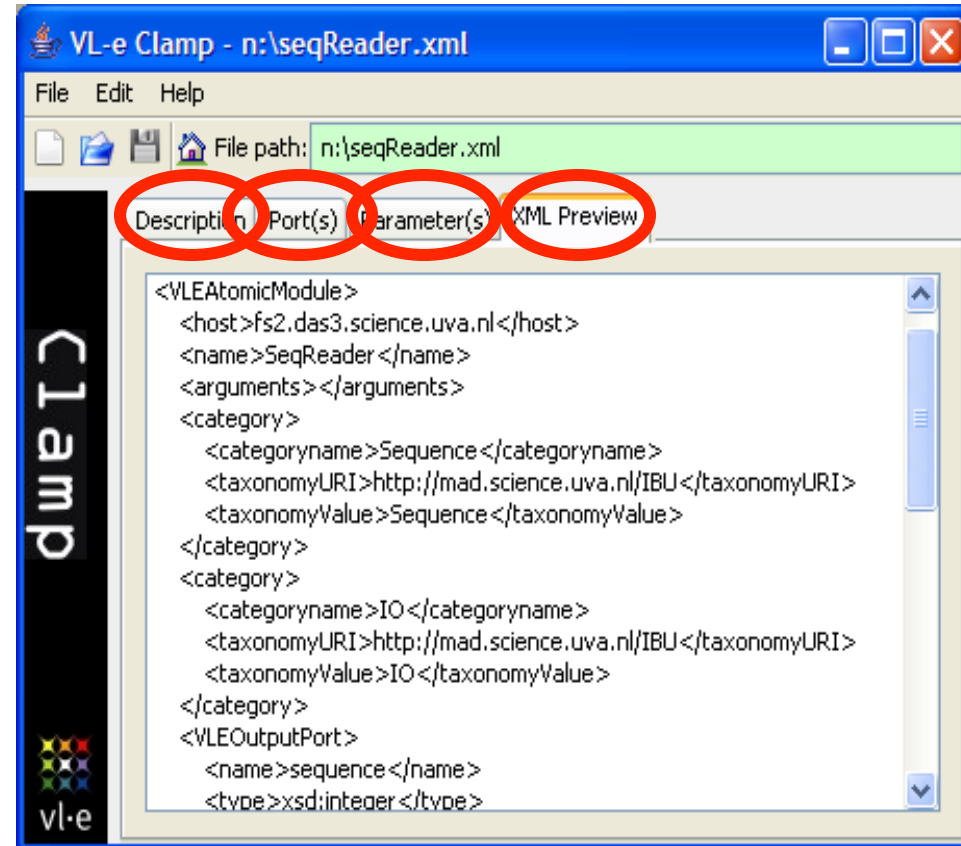
(1) Problem investigation

- Tools to improve **re-usability**, and **share of expertise**. These tools should allow:
 - Advanced **search capabilities**
 - HAMMER: **H**ybrid-**b**Ased **M**atch-**M**aker for **E**-Science **R**esources
 - Web Service Harvester (RPC Style WSDL)



(1) Problem investigation

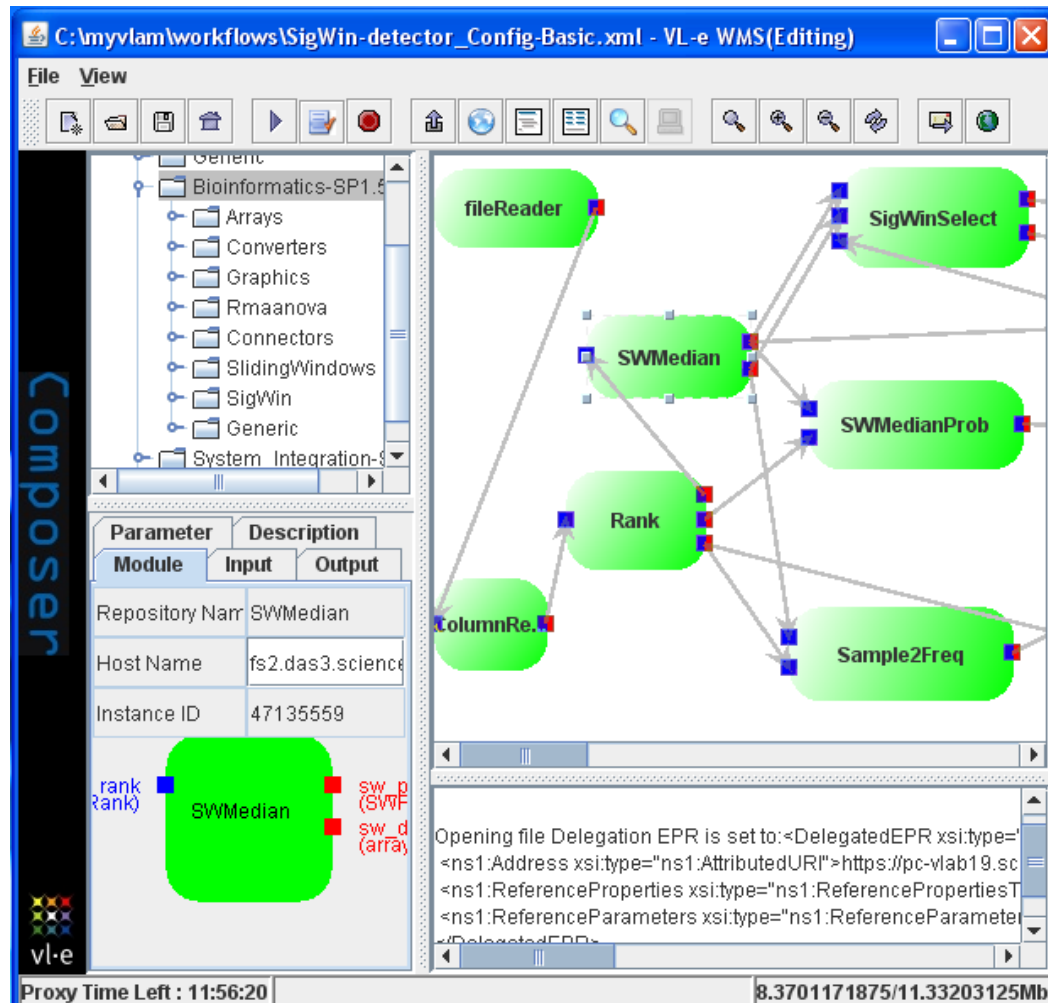
- Tools to design **new components**. These tools should allow:
 - **Speedup** and **simplify** the design of the new
 - CLAMP: **C**onnecting **L**anguage for **M**odules and **P**rograms



(2) Experiment Prototyping

Grid Enabled workflow management system

- Tools to **speedup** and **simplify** building prototypes. These tools should allow:
 - **Combine** existing software components
 - **Seamless access** to computational and storage resources.



Composer

Parameter	Description	Module	Input	Output
Repository Name	SWMedian			
Host Name	fs2.das3.science			
Instance ID	47135559			

rank (rank)

SWMedian

sw_p (SWF)

sw_d (arfa)

fileReader

ColumnRe

Rank

SWMedian

SWMedianProb

Sample2Freq

SigWinSelect

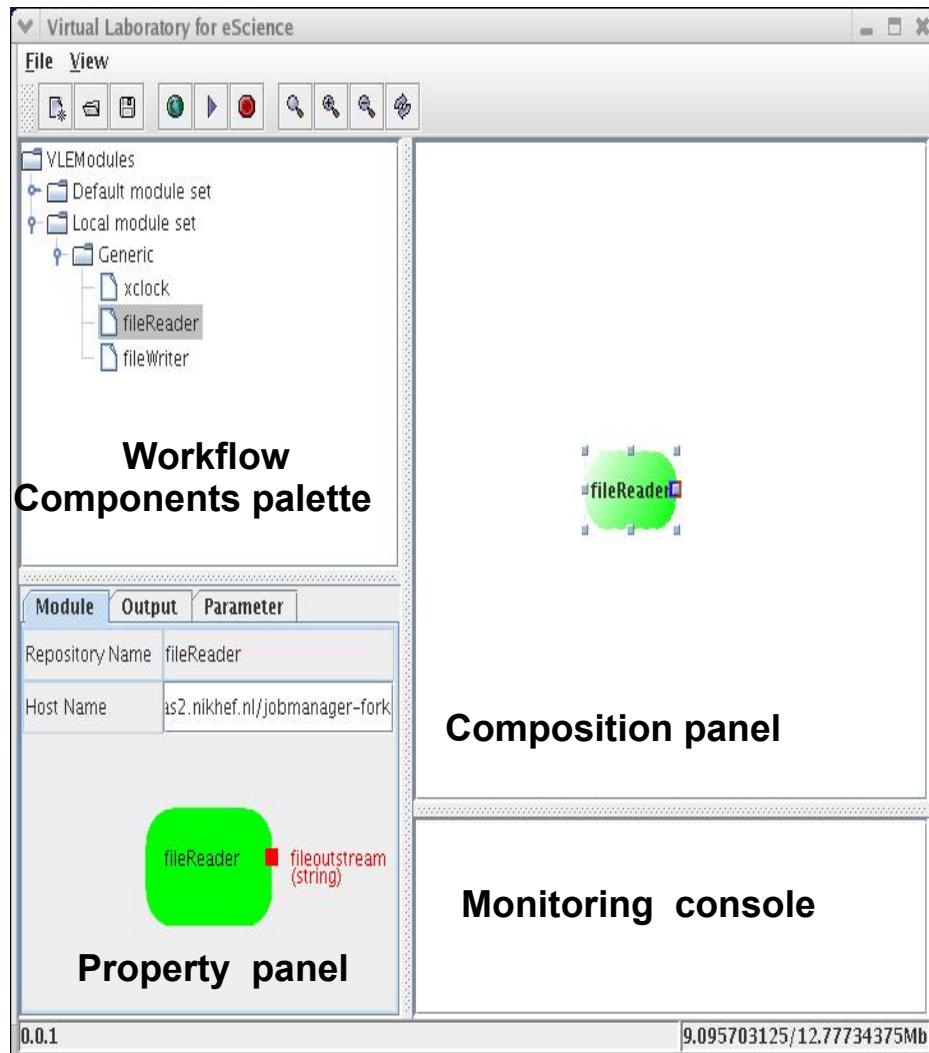
Proxy Time Left : 11:56:20

8.3701171875/11.33203125Mb

(2) Experiment Prototyping

Grid Enabled workflow management system

- For Application developers
 - API for JAVA, C/C++, PYTHON,
 - Data Streaming capabilities
- End-users (**scientists**)
 - Workflow monitoring
 - Drag-and-drop composition
 - Hierarchical workflow creation
 - User in the loop
 - Detach/re-attach capability



The screenshot displays the Virtual Laboratory for eScience interface, which is divided into several functional panels:

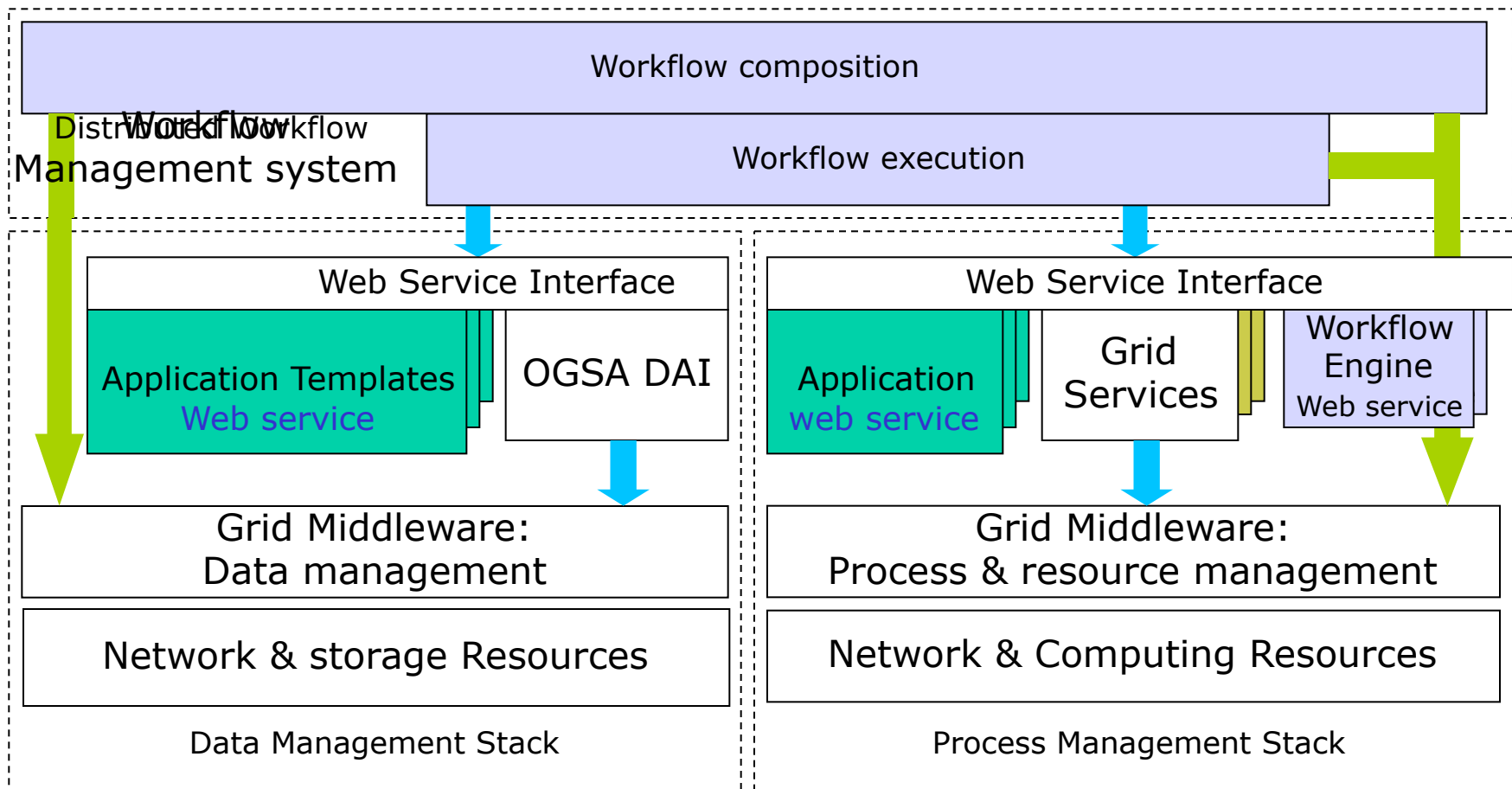
- Workflow Components palette:** A tree view on the left showing the hierarchy of modules: VLEModules > Default module set > Local module set > Generic > xclock, fileReader, and fileWriter. The 'fileReader' module is currently selected.
- Composition panel:** A central workspace where the 'fileReader' module is being dragged and dropped, indicated by a green dashed border around the module icon.
- Property panel:** A panel at the bottom left showing the configuration for the selected 'fileReader' module. It includes a table with the following data:

Module	Output	Parameter
Repository Name	fileReader	
Host Name	as2.nikhef.nl/jobmanager-fork	

 Below the table, there is a green rounded rectangle representing the module icon, with a red square next to it labeled 'fileOutputStream (string)'.
- Monitoring console:** A panel at the bottom right, currently empty, intended for monitoring the workflow's execution.

The interface also features a standard menu bar (File, View) and a toolbar with various icons for file operations and workflow control. The status bar at the bottom shows the version '0.0.1' and a file path '9.095703125/12.77734375Mb'.

A WSRF enabled workflow engine



(2) Experiment Prototyping Semantic tools

- Tools to improve **sharing** and **re-usability**. These tools should allow:

- **Annotate** and **publish** the designed experiment prototype

- SAW: Semantic Annotation for Workflow and workflow components
- HAMMER: **H**ybrid-**b**Ased **M**atch-**M**aker for **E**-Science Resources

VL-e SAW 1.1 - n:\seqReader.xml

File Edit Help

Base URI <http://vle.wtcw.nl/SeqReader.owl>

Profile Service Model OLW-S Graphic OWL-S Source

```
<service:Service rdf:about="http://vle.wtcw.nl/SeqReader.owl#SeqReader"
  <service:presents>
    <profile:Profile rdf:about="http://vle.wtcw.nl/SeqReader.owl#SeqReader"
  </service:presents>
  <service:describedBy>
    <process:AtomicProcess rdf:about="http://vle.wtcw.nl/SeqReader.owl#SeqReader"
  </service:describedBy>
</service:Service>
<profile:Profile rdf:about="http://vle.wtcw.nl/SeqReader.owl#SeqReader"
  <profile:serviceClassification rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  ><profile:serviceClassification>
  <profile:serviceName>SeqReader</profile:serviceName>
  <profile:serviceCategory>
  <profile:ServiceCategory rdf:about="http://vle.wtcw.nl/SeqReader.owl#SeqReader"
  <profile:categoryName>IO</profile:categoryName>
  <profile:taxonomy>http://mad.science.uva.nl/IBU</profile:taxonomy>
```

SAW

vl-e

(2) Experiment Prototyping Publishing tools

- Tools to improve **sharing** and **re-usability**. These tools should allow:

- **Annotate** and **publish** the designed experiment prototype

- SAW: Semantic Annotation for Workflow and workflow components
- HAMMER: **H**ybrid-**b**Ased **M**atch-**M**aker for **E**-Science Resources

WS-VLAM HAMMER: An semantic-based workflow broker - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://pc-vlab19.science.uva.nl:8081/ws- Google

Most Visited Getting Started Latest Headlines

DESCRIPTION

- overview
- HOW-TOs
- support
- about

WS-VLAM HAMMER
Hybrid-bAsed Match-Maker for E-science Resources

WS-VLAM

Quick Description

This page provides an interface to a semantic workflow matchmaker, developed within the VL-e project. The interface allows scientist to advertise and discover workflow descriptions. In order to advertise, the scientist must create a OWL-S Profile (version 1.1) of the workflow and, then, 'publish' it in the HAMMER repository

Two query functionalities are available to discover resources from the repository. It is possible to create simple keyword queries or, define an OWL-S Profile to describe the capabilities expected in soughted resources. The discovery page submits the request to matchmaker, which will attempt to locate the appropriate descriptions.

> **Publishing Page**

> **Discovery Page**

Copyright © 2005-2008. UvA and Contributors. All rights reserved.

VL-e SP 2.5
Kruislaan 403
1098 SJ, Amsterdam

vl-e virtual laboratory for e-science

Find: problem Next Previous Highlight all Match case

Done

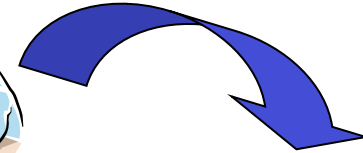
(3) Experiment Execution

Virtual resource Browser

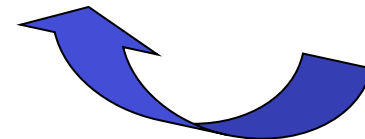
- Tools with **intuitive interface** allowing end users (Scientists):
 - **Find** results: access resources to manipulate data
 - **Re-start** experiment and monitor its execution



*analyzing the
experiment results*



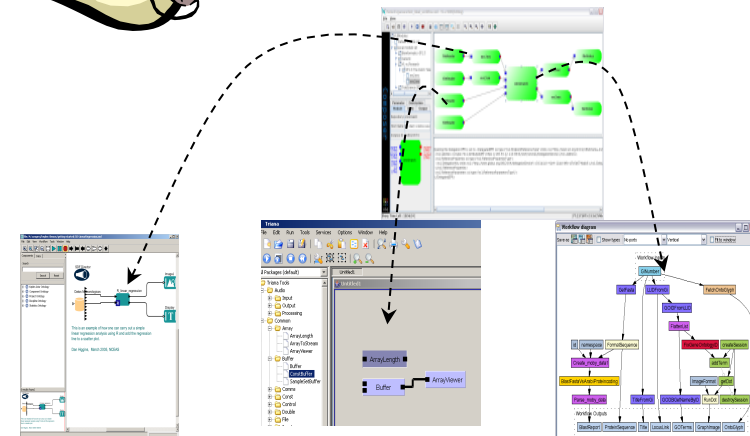
*Perform more
experiments*



(3) Experiment Execution VLE-WFBus

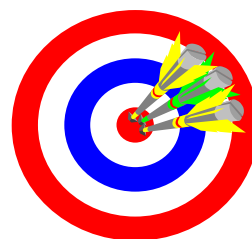
- Tools to **improve interoperability** among various execution platforms.
 - Recognize different workflow descriptions.
 - Coordinate the execution of workflows, and monitor their execution.

*performing
the
experiment*

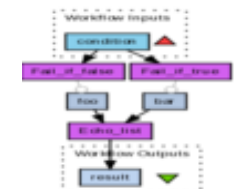


(4) Results Publication

- Tools to improve **share of expertise**. These tools should allow:
 - **Annotate** and **publish** data, workflows, etc
 - HAMMER: **H**ybrid-**b**Ased **M**atch-**M**aker for **E**-Science Resources (at the project level)
 - MyExperiment: (at the community level)



success

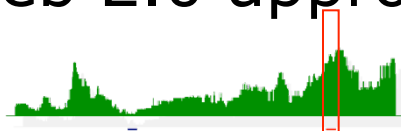


publish

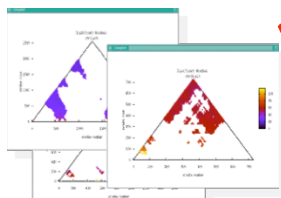
(4) Results Publication

- Workflow can be invoked from other systems
- Workflow can be made available to entire community (using Web 2.0 approach)

• Human transcriptome map

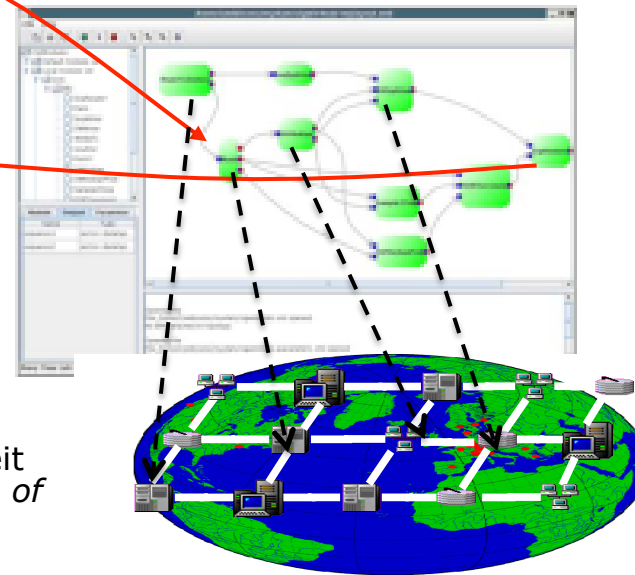


• DNA curvature of the *Escherichia Coli* chromosome



• myexperiment web site

• WS-VLAM composer



Marcia A Inda, Marinus F van Batenburg, Marco Roos, Adam SZ Belloum, Dmitry Vasunin, Adianto Wibisono, Antoine HC van Kampen and Timo M Breit
 SigWin-detector: a Grid-enabled workflow for discovering enriched windows of genomic features related to DNA sequences, BMC Research Notes 2008,

Example of Complex eScience

VLe Studio

- WS-VLAM composer
- VBrowser
- Semantic tools



SAW: Semantic Annotation for Workflow

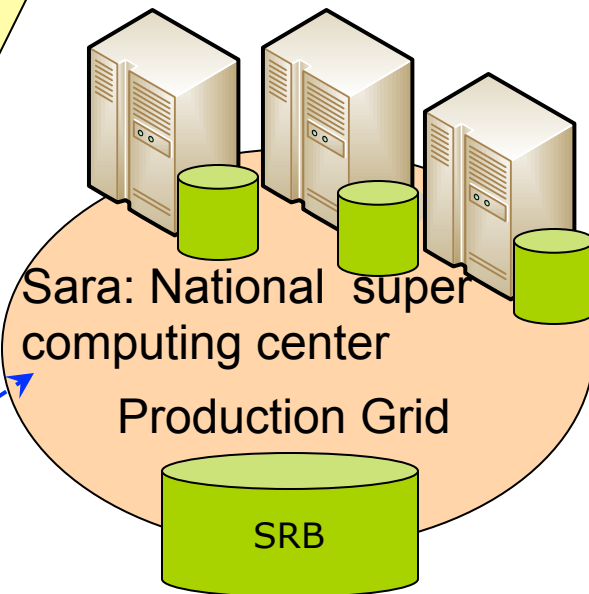
CLAMP: Connecting Language for Modules & Programs

HAMMER: Hybrid-based MatchMaker for e-Science

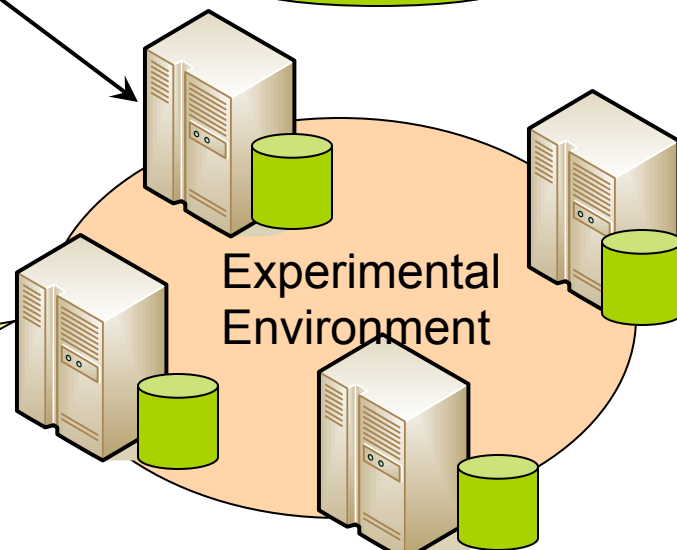
Resources

WSRF Services

- WS-VLAM engine
- workflow component repository

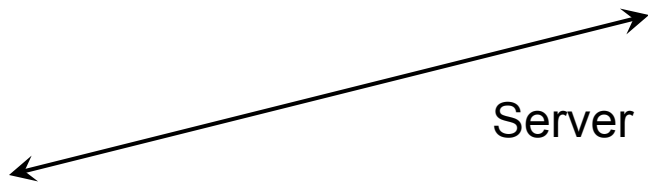
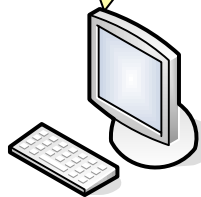
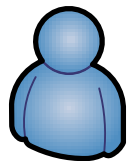


Server host



Computing Nodes

- Workflow components
- Grid Middleware → GT4

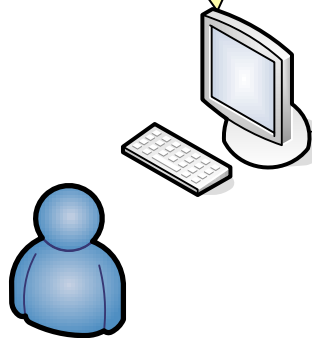


Example of usage Scenario ...

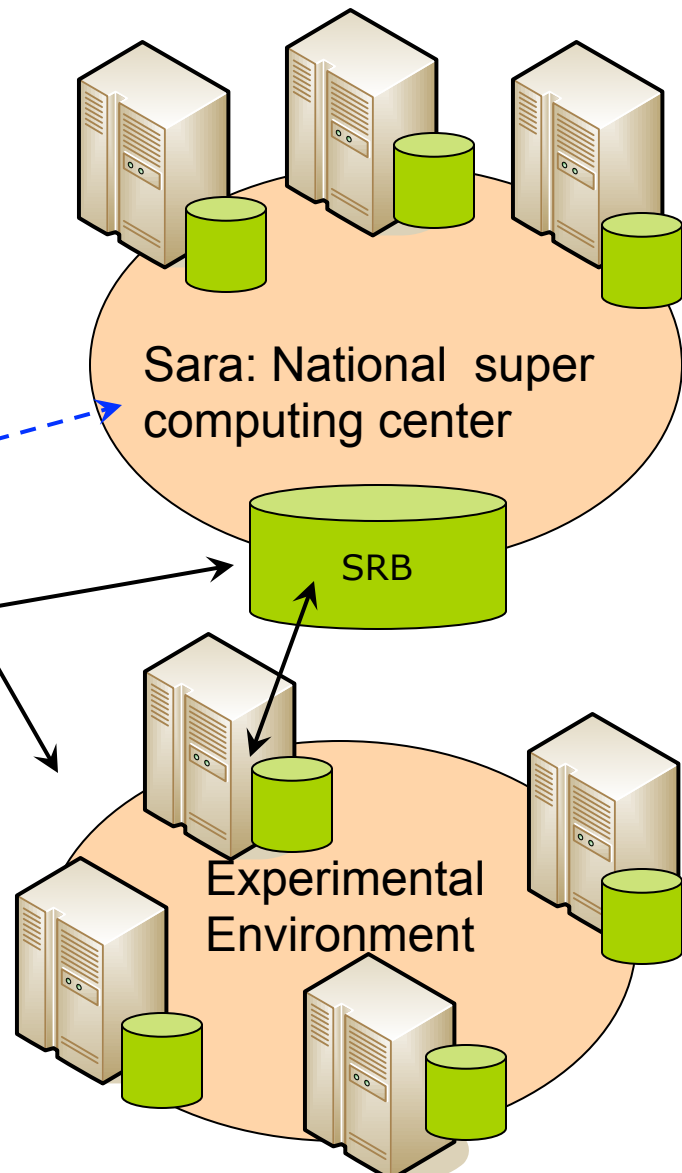


User start VL-e Studio

- Search for services for the workflow
- Upload in the workflow composer
- Compose the workflow
- Search for input data set in the SRB
- Set the workflow parameters
- Execute the workflow
- Monitor the execution



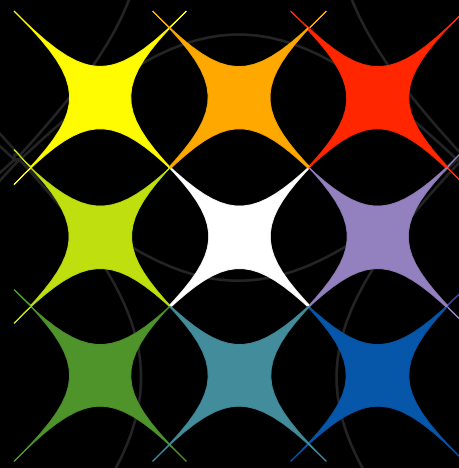
pc-vlab19.uva





References

1. **A.S.Z. Belloum**, V. Korkhov, S. Koulouzis, M. A Inda, and M. Bubak, *Collaborative e-Science experiments: from scientific workflow to knowledge sharing*, IEEE Internet Computing, vol. 15, no. 4, pp. 39-47, July/August, 2011, doi:10.1109/MIC.2011.87.
2. C. Leguy, Bosboom, F.N.V.D Vosse, **A.S.Z. Belloum**, A. Hoeks, *Global sensitivity analysis of a wave propagation model for arm arteries*, Journal of Medical Engineering Physics 2011 Oct, 33(8):1008-16, doi:10.1016/j.medengphy.2011.04.003.
3. M. Gerhards, V. Sander, **A.S.Z. Belloum**, D. Vasunin, A. Benabdelkader, *Hist/PLIER: A two-fold Provenance Approach for Grid-enabled Scientific*, In Proceedings of the 12th IEEE/ACM International Conference on Grid Computing, pp.224-225, 21-23 Sept. 2011, doi:10.1109/Grid.2011.39.
4. Frank Berretz, Sascha Skorupa, Volker Sander, **Adam S.Z. Belloum**, Marian Bubak. *Actor-driven Workflow Execution in Distributed Environments*, Euro-Par 2010 Parallel Processing Workshops, Lecture Notes in Computer Science vol. 6586, 2011, pp. 287-294, doi: 10.1007/978-3-642-21878-1_36.
5. R. Cushing, S. Koulouzis, **A.S.Z. Belloum**, M.T. Bubak, *Prediction-based Auto-scaling of Scientific Workflows*, 9th International Workshop on Middleware for Grids, Clouds and e-Science (MGC'2011), Lisbon Portugal Dec. 2011, doi:10.1145/2089002.2089003



vl-e

<http://www.vl-e.nl/>