

vl·e



virtual laboratory for e·science

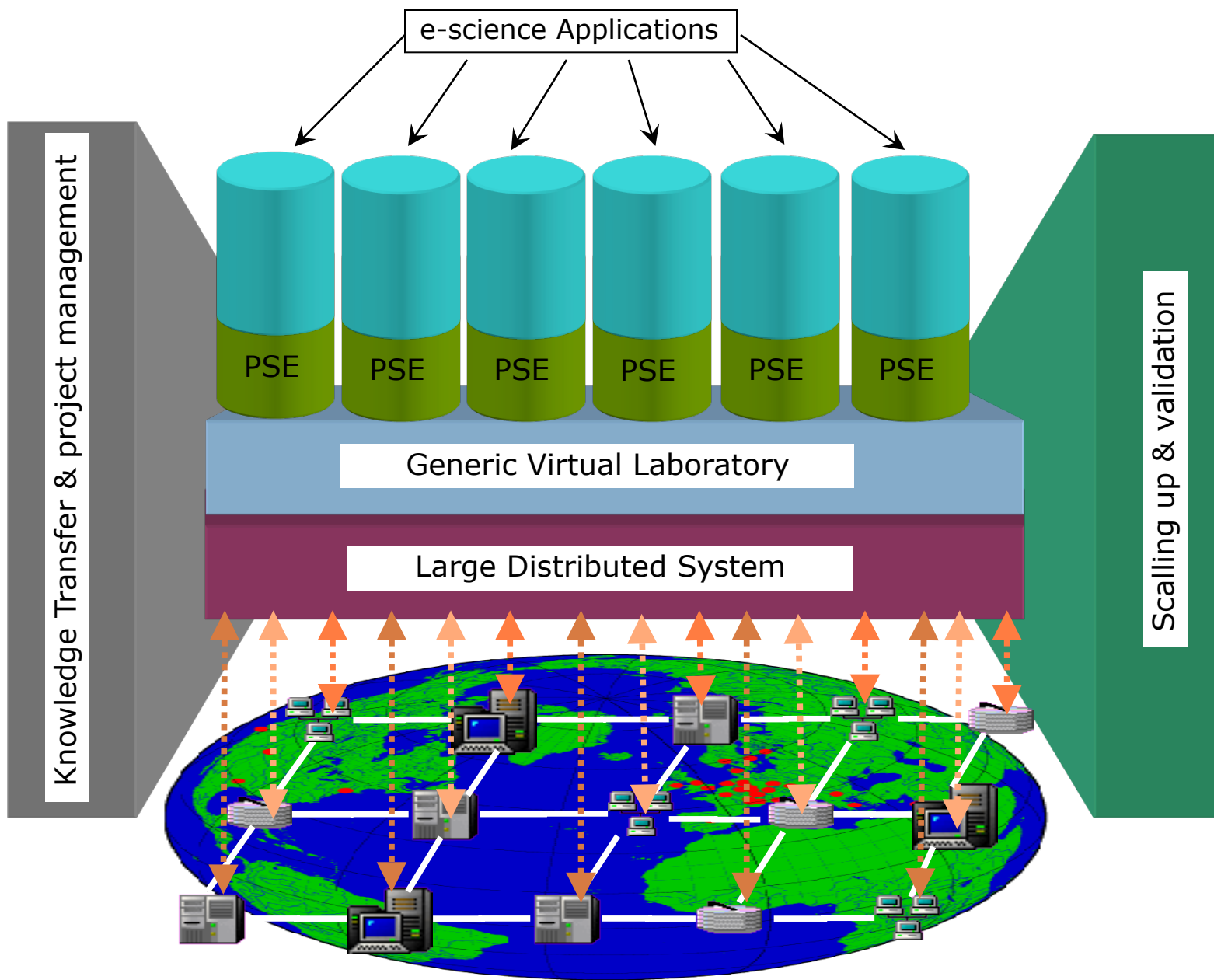
VL-e eScience Framework: from design to sharing

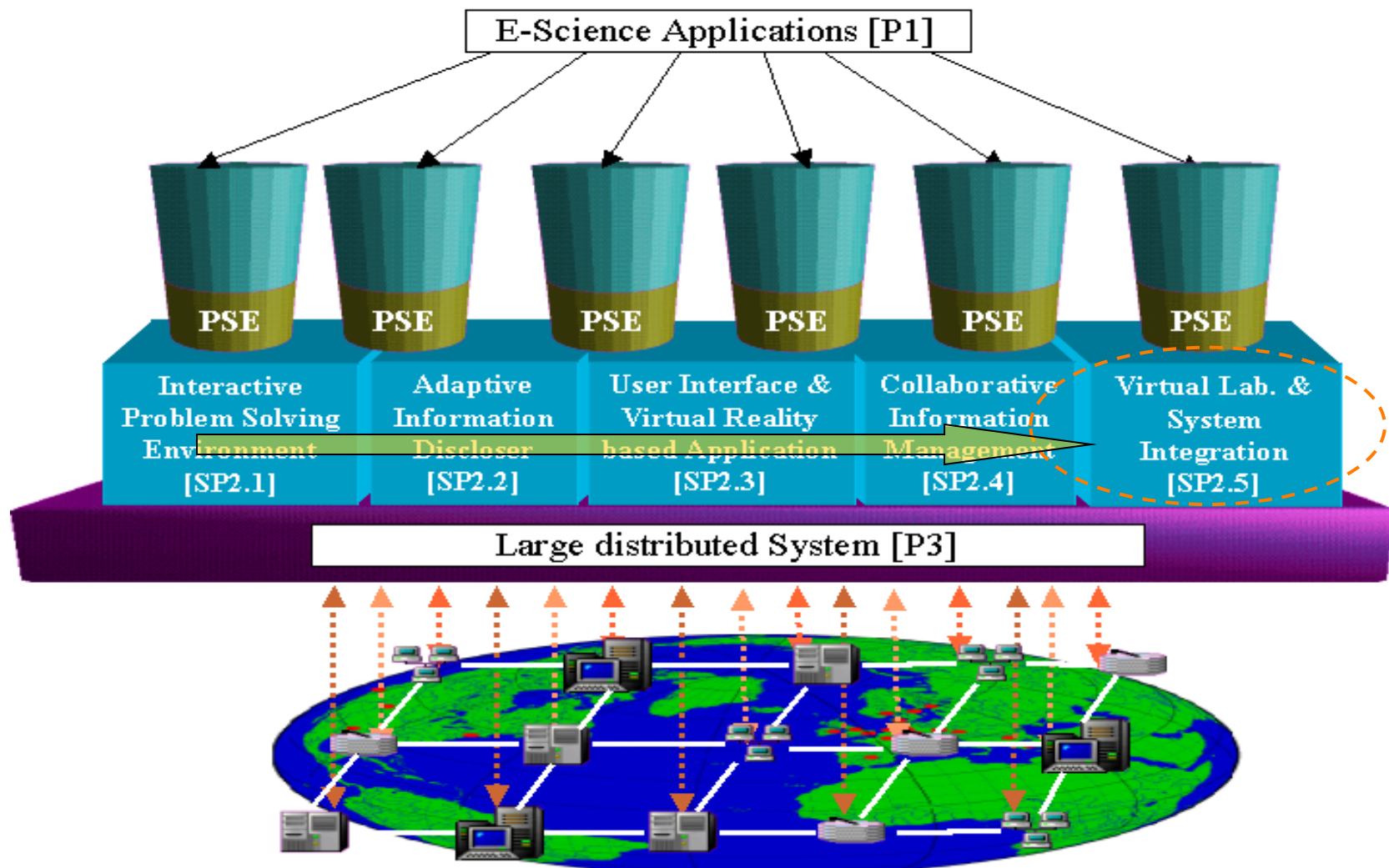
Adam Belloum
Institute of Informatics
University of Amsterdam

UvA



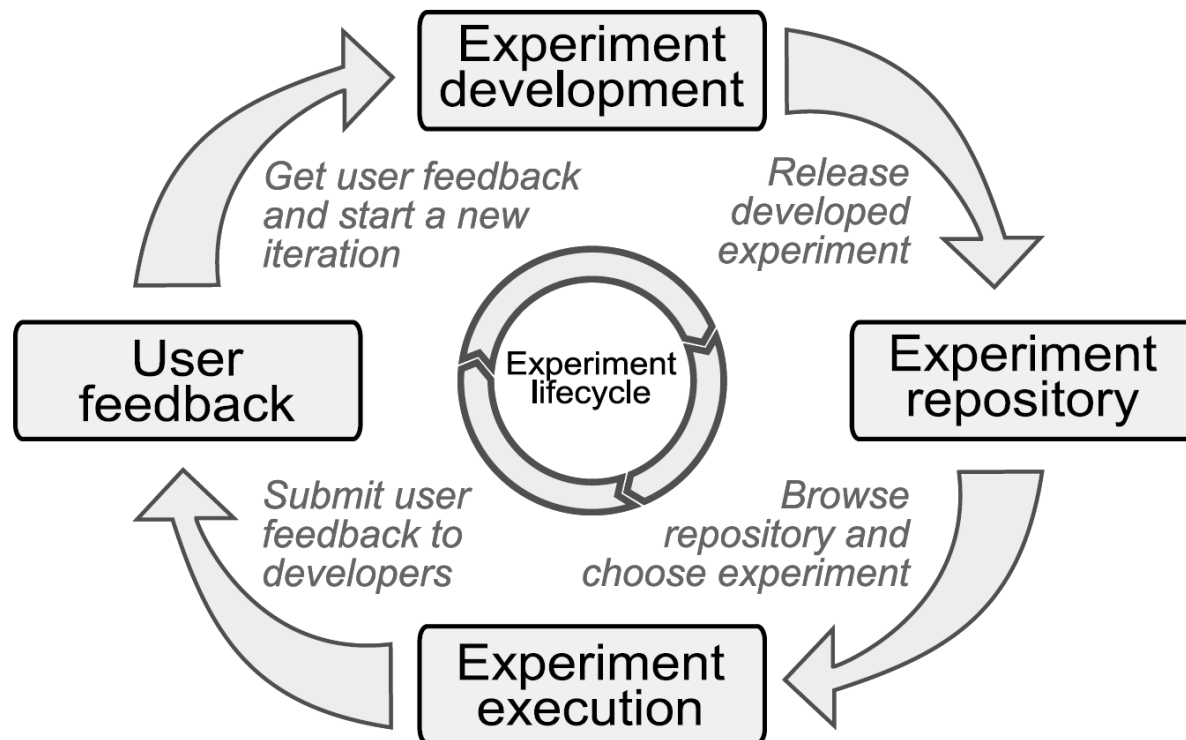
UNIVERSITEIT VAN AMSTERDAM



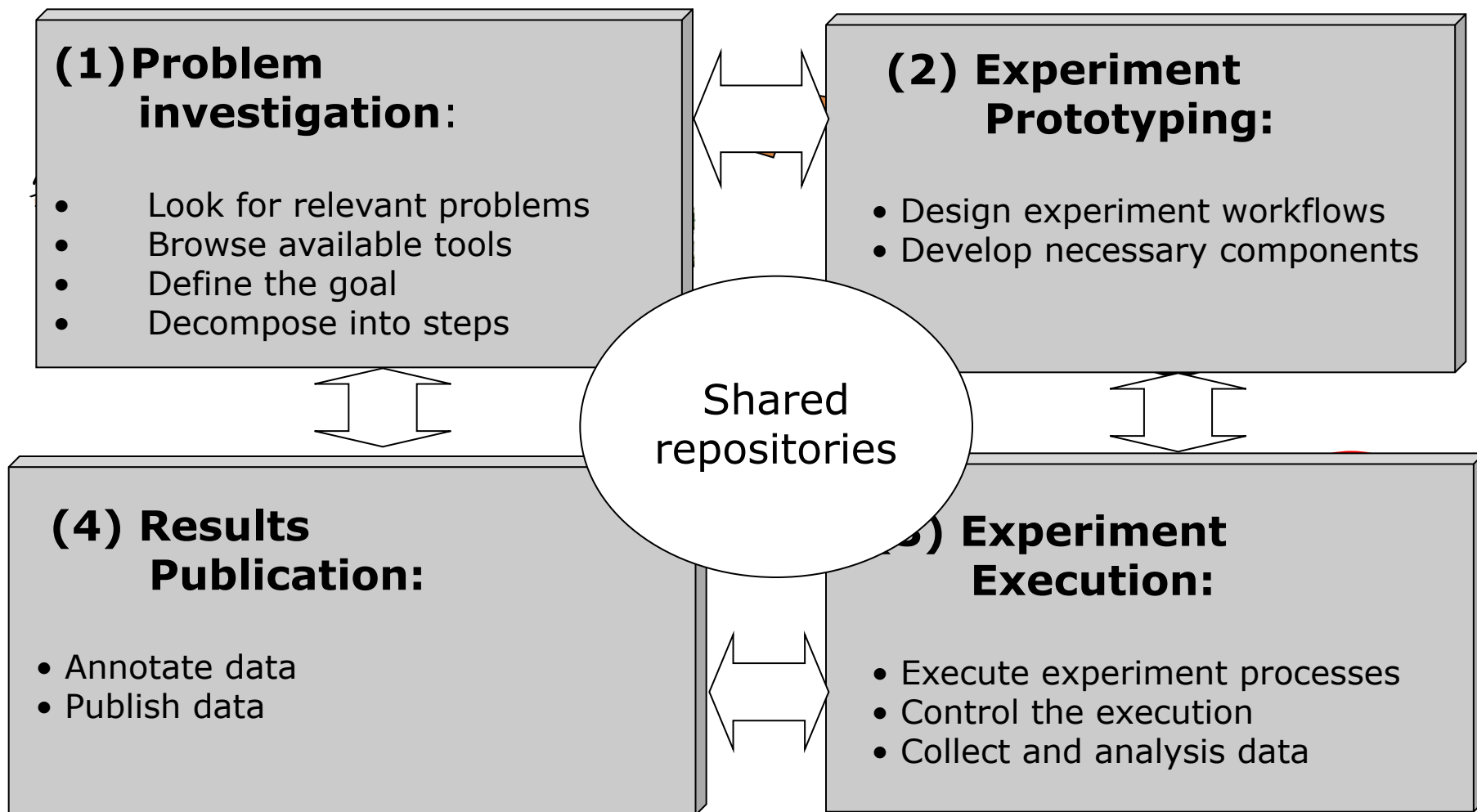


Experiment Lifecycle

- *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

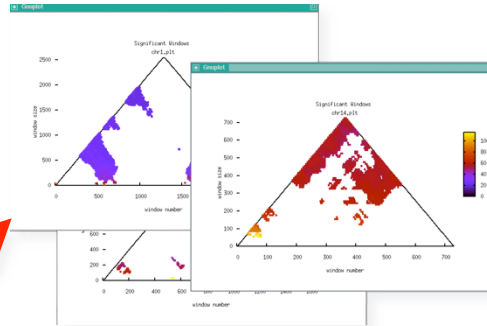
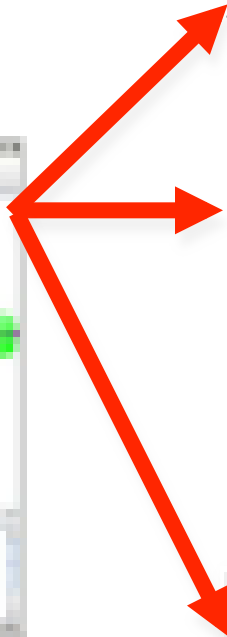
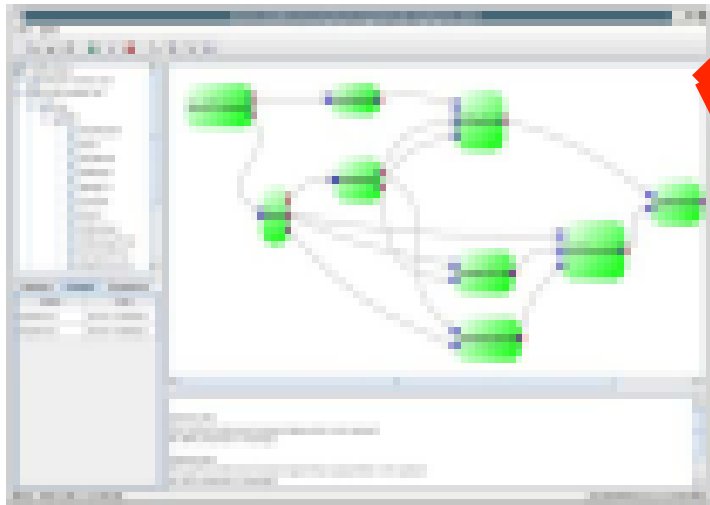
WS-VLAM e-Studio

- 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

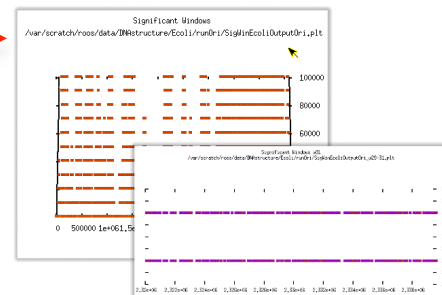


The Application Workflow ...

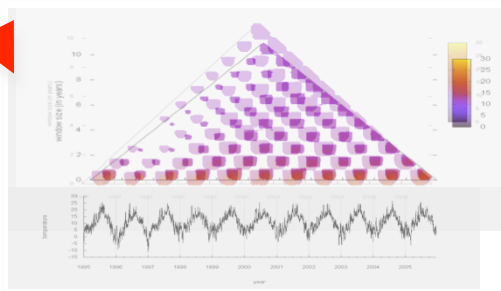
Developed in the context of the
Ridge-O-grammer Application use case
by Micro-array department UvA



Human
transcriptome
map



DNA curvature
of the
Escherichia coli
chromosome



Temperature in
Amsterdam

(1) Problem investigation

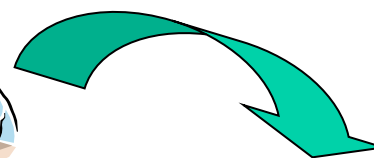
- Tools to improve **re-usability**, and **share of expertise**. These tools should allow:

- Advanced **search capabilities**

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

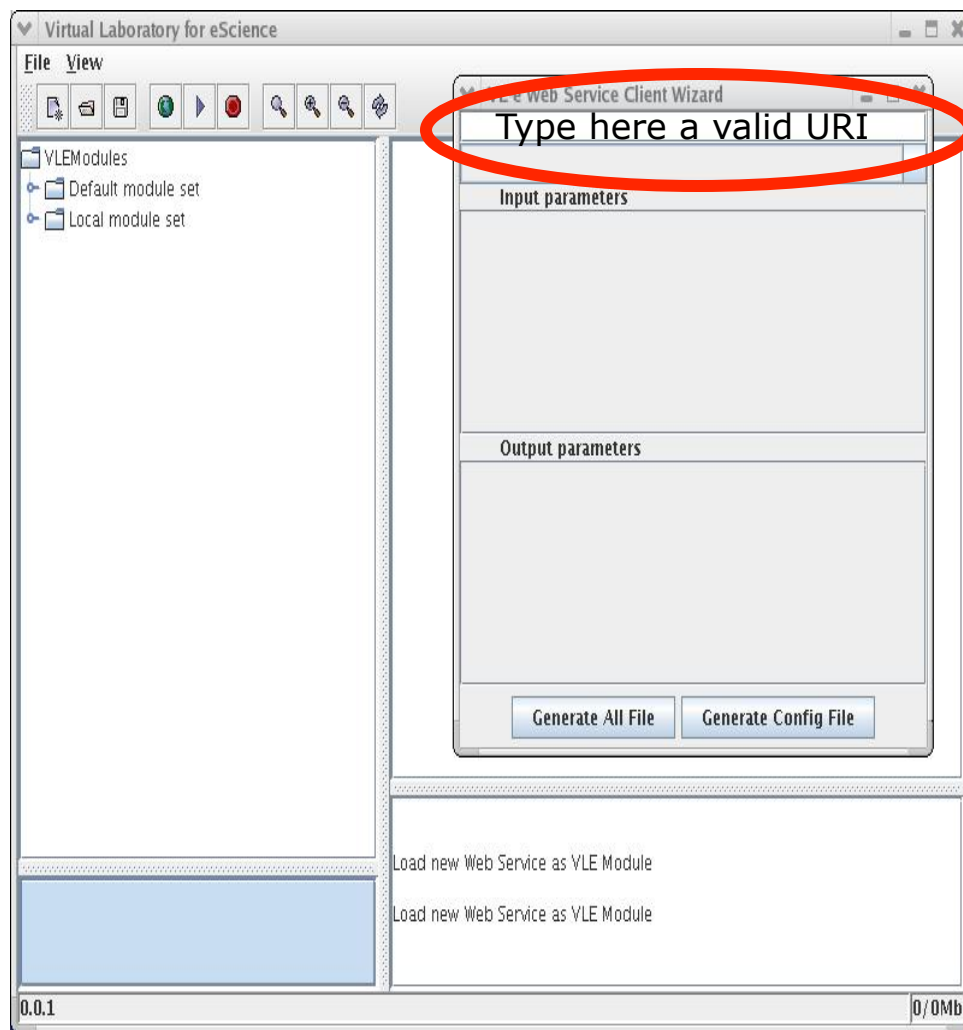


designing the experiment



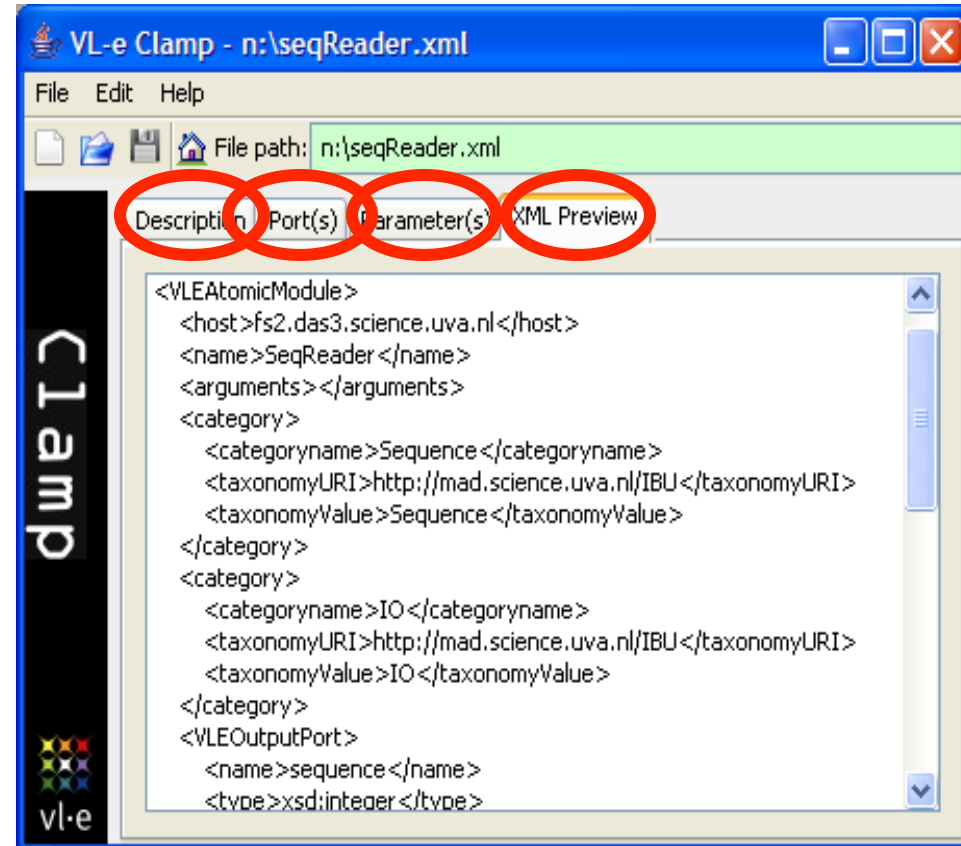
(1) Problem investigation

- 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**-**S**cience **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		SWMedian	rank	rank

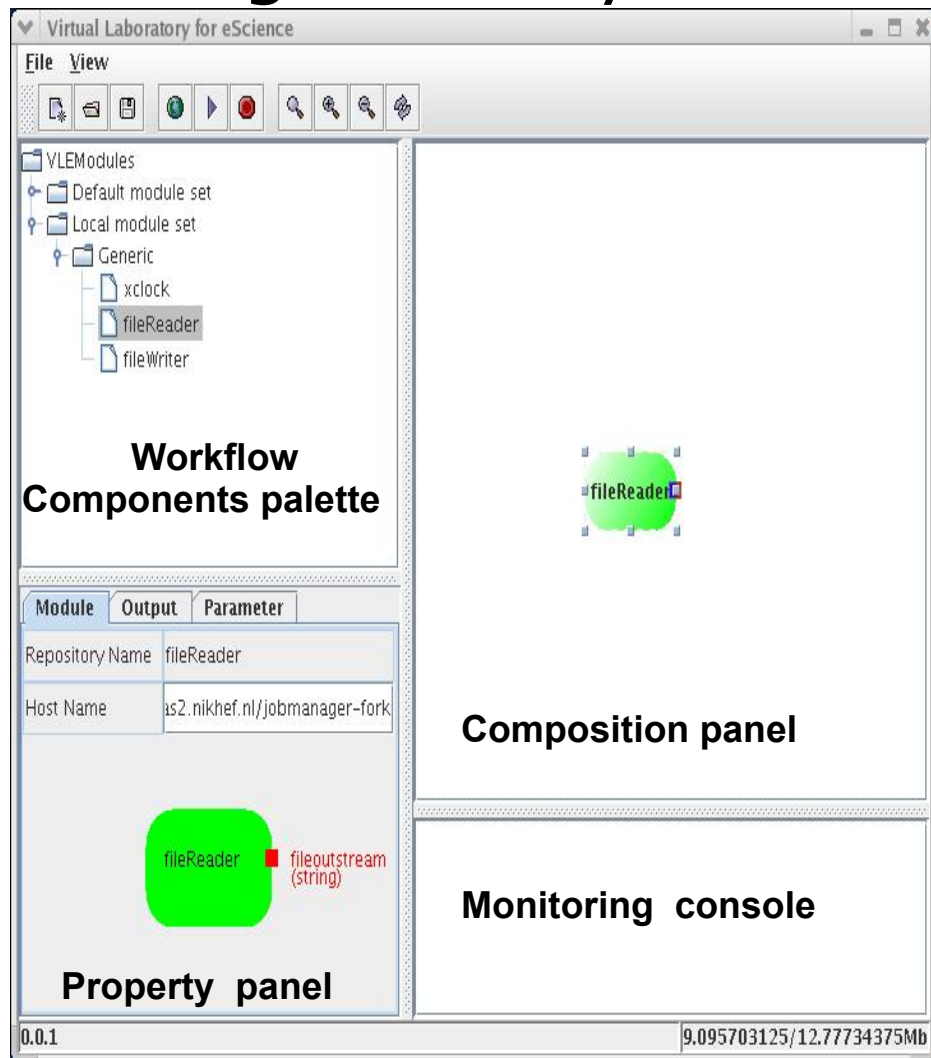
Proxy Time Left : 11:56:20

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 large central workspace where a 'fileReader' module is being dragged and dropped, indicated by a green glow and a dashed border.
- Property panel:** A table at the bottom left showing the configuration for the selected 'fileReader' module.

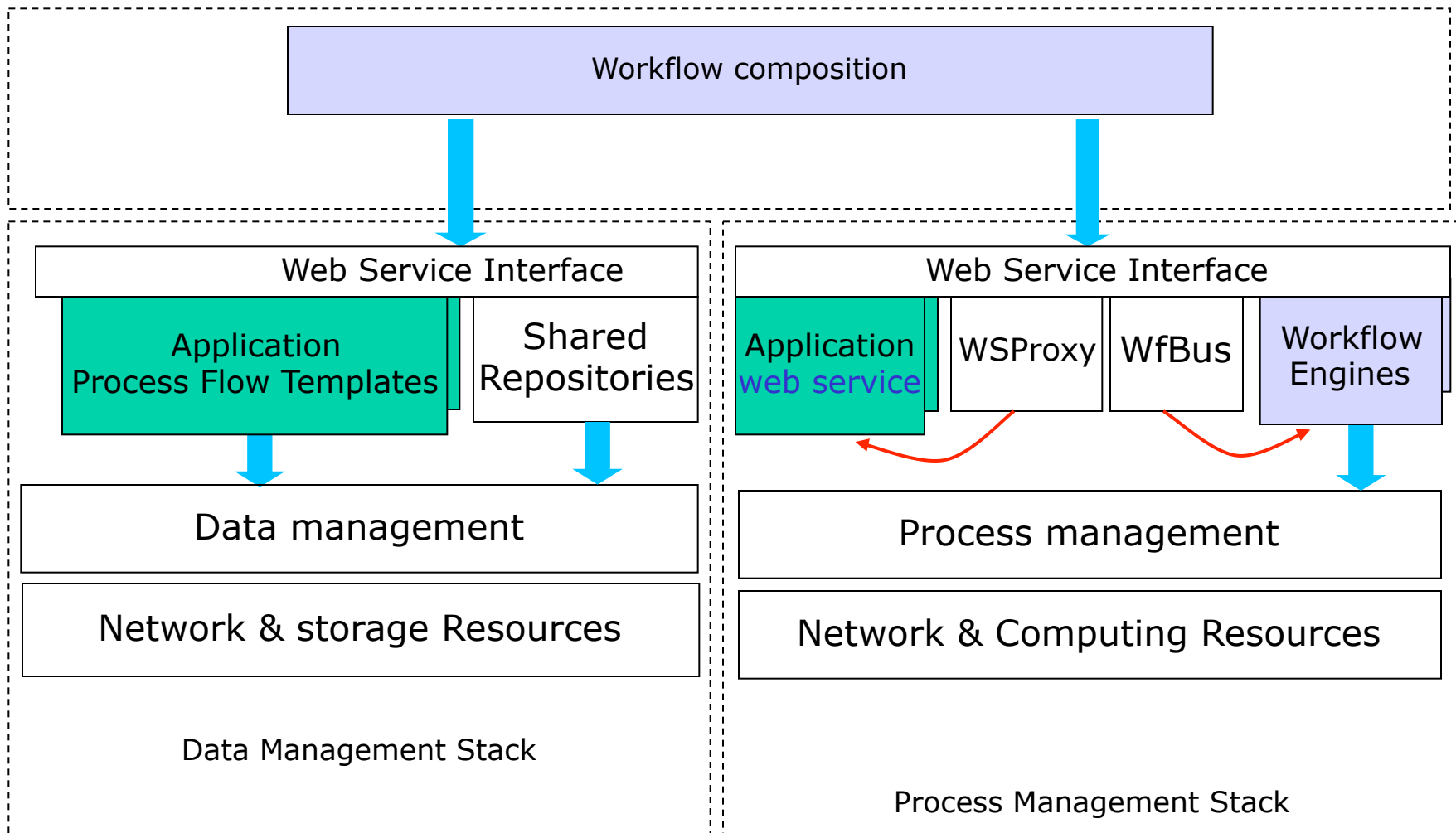
Module	Output	Parameter
Repository Name	fileReader	
Host Name	as2.nikhef.nl/jobmanager-fork	
- Monitoring console:** A panel at the bottom right, currently empty, intended for monitoring the workflow's execution.

The interface also includes 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 unique identifier '9.095703125/12.77734375Mb'.

Tutorials for both developers and End-users are available at

<http://www.science.uva.nl/~gvlam/wsvlam/Documentation>

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

Hammer

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: Next Previous Highlight all Match case

Done

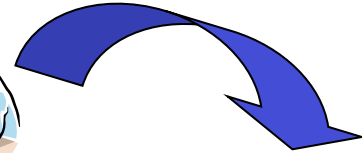
(3) Experiment Execution

Virtual resource Browser

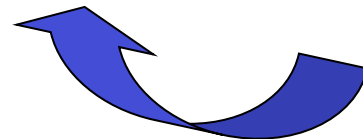
- Tools with **intuitive interface** allowing end users (Scientists):
 - **Find** results: access resources to manipulate data
 - **Re-start** experiment and monitor its execution



*analyzing the
experiment results*



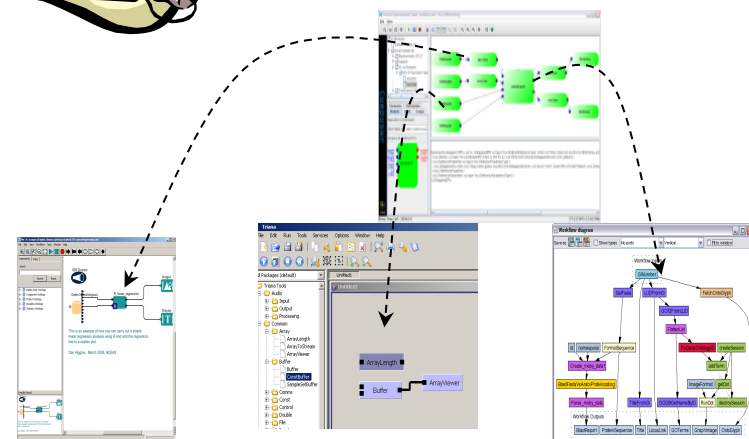
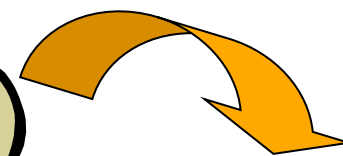
*Perform more
experiments*



(3) Experiment Execution VLE-WFBus

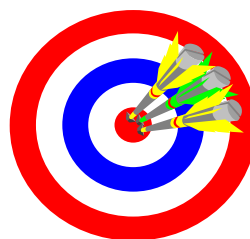
- Tools to **improve interoperability** among various execution platforms.
 - Recognize different workflow descriptions.
 - Coordinate the execution of workflows, and monitor their execution.

*performing
the
experiment*

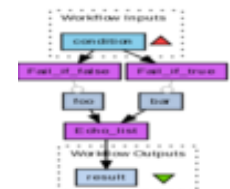


(4) Results Publication

- 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)



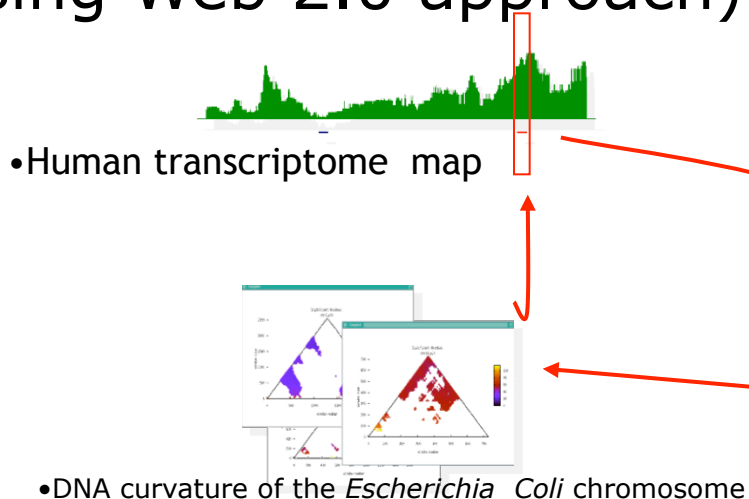
success



publish

(4) Results Publication

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



• myexperiment web site

myexperiment beta About | Mailing List | Publications Log In | Register | Give us Feedback | Invite

Home Users Groups Workflows Files Packs

Home > Workflows > SigWin-detector Config-Basic

Workflow Entry: SigWin-detector Config-Basic

Created at: 30/10/08 @ 14:26:17 Last updated: 24/04/09 @ 09:37:57

License | Credits (1) | Attributors (0) | Tags (1) | Featured in Packs (0) | Ratings (0) | Attributed By (0) | Favourited By (0) | Citations (0) | Version History | Reviews (0) | Comments (2)

Version 1 (of 1)

Version created on: 30/10/08 @ 14:26:17 by: Adambel | Revision comments

Last edited on: 24/04/09 @ 09:37:57 by: Adambel

Title: SigWin-detector Config-Basic

Type: Chemistry Plan

Preview

Description

Detects significant windows in a sequence.

[Launch in wslam]

Input: A space-delimited file with (at least) one column containing the input sequence E1, E2, ..., EN. A two-line header should precede the data. Input file example: "sig-win.txt"

Output:

(1) A file containing the detected significant windows for each label. Each data row represents a stretch of consecutive significant windows. Column 1 gives the window size and columns 2 and 3 give the first and last significant windows in the stretch.

Original Uploader: Adambel

Credits (1)

Attributions (0)

Tags (1)

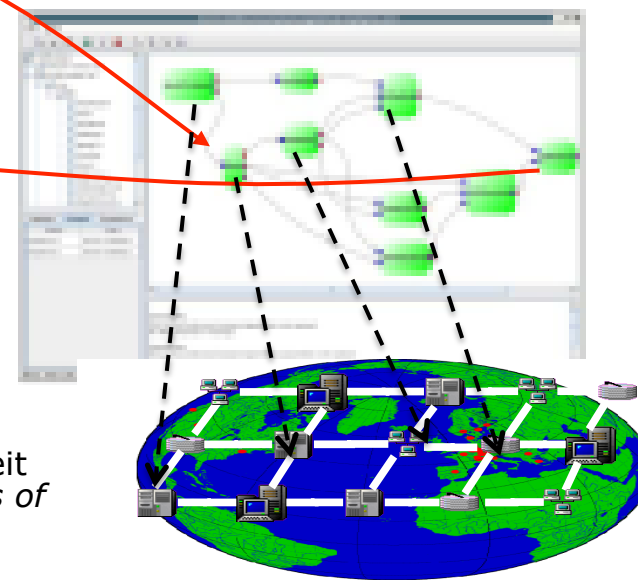
Original Uploader tags: ws-vlam; wslam

Shared with Groups (0)

Featured in Packs (0)

Ratings (0)

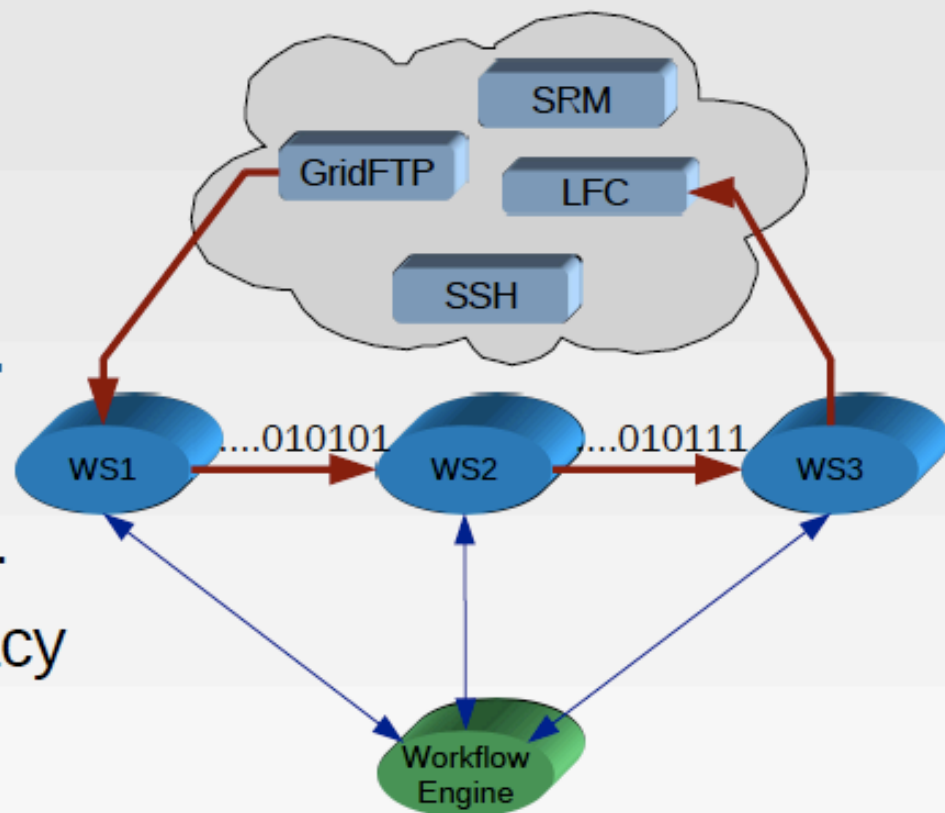
• WS-VLAM composer



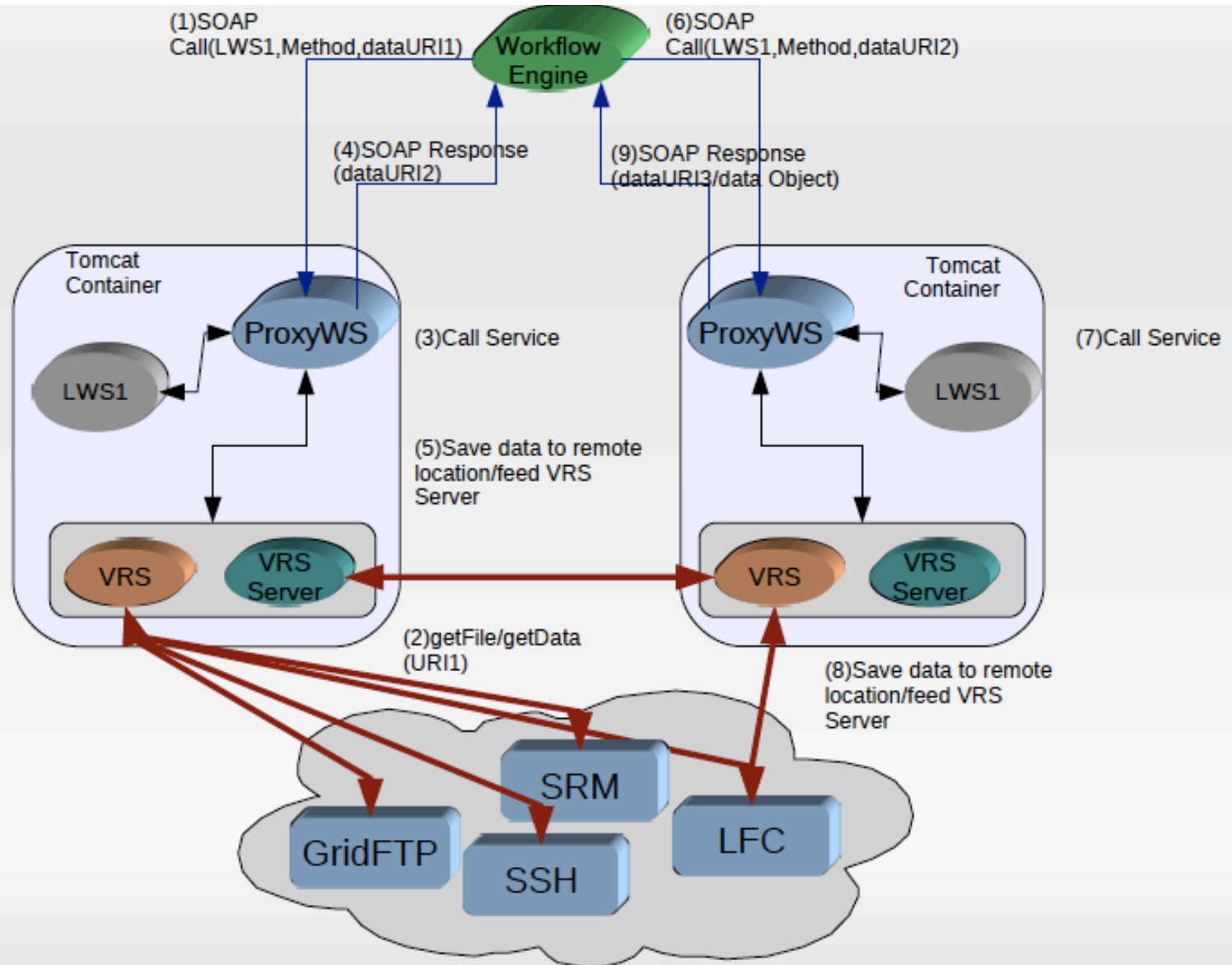
Marcia A Inda, Marinus F van Batenburg, Marco Roos, Adam SZ Belloum, Dmitry Vasunin, Adianto Wibisono, Antoine HC van Kampen and Timo M Breit
 SigWin-detector: a Grid-enabled workflow for discovering enriched windows of genomic features related to DNA sequences, BMC Research Notes 2008,

(*) Web Service and Data centric workflows

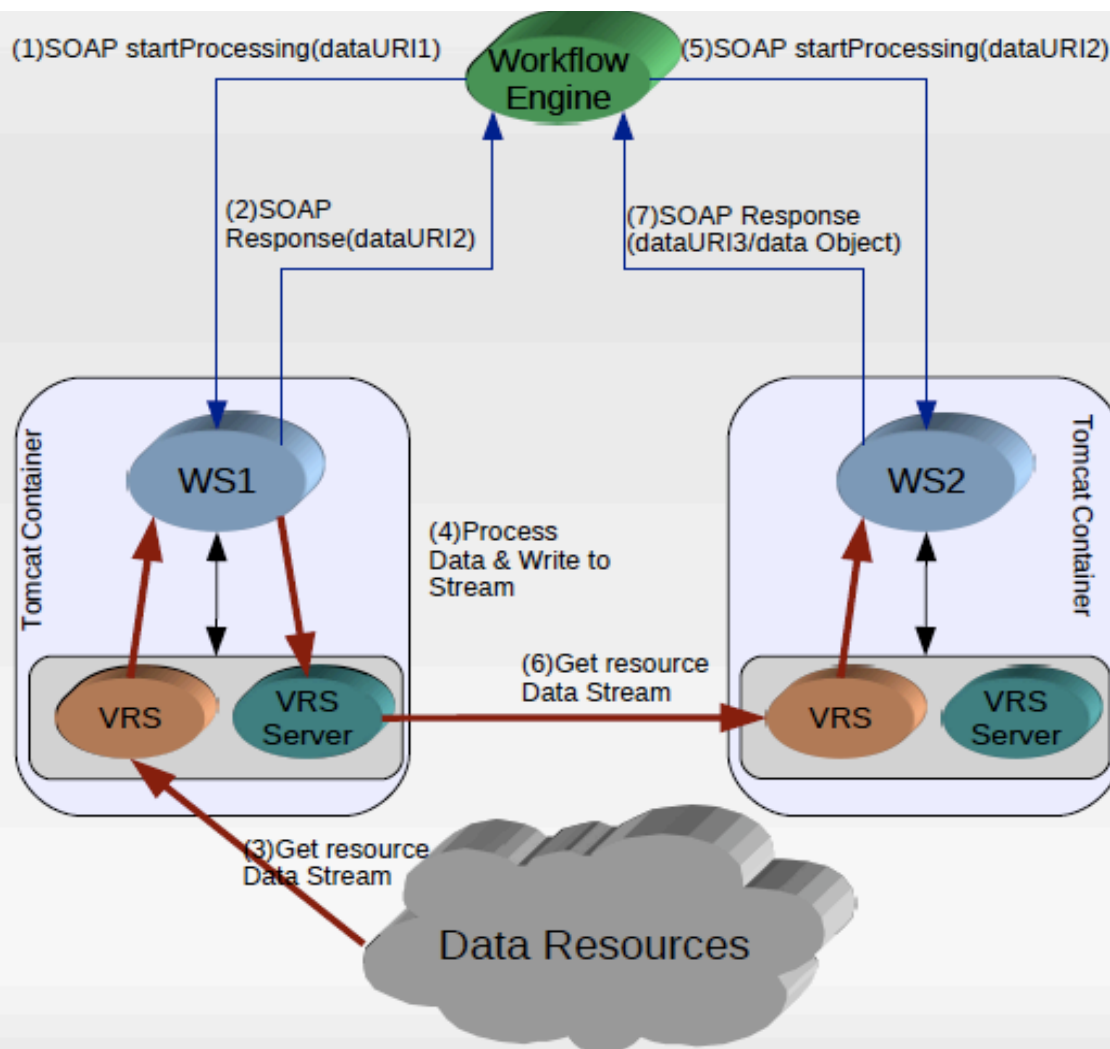
- Introduce a new execution model for data driven workflows.
- It will enable e-Science applications to meet increasing data demands.
 - Provide scalable data transports for web services.
 - Bring data to and from legacy web services
 - Develop an API that web services can use.



(*) Web Service and Data centric workflows



(*) Web Service and Data centric workflows



Example of Complex eScience

VLe Studio

- WS-VLAM composer
- VBrowser
- Semantic tools



SAW: Semantic Annotation for Workflow

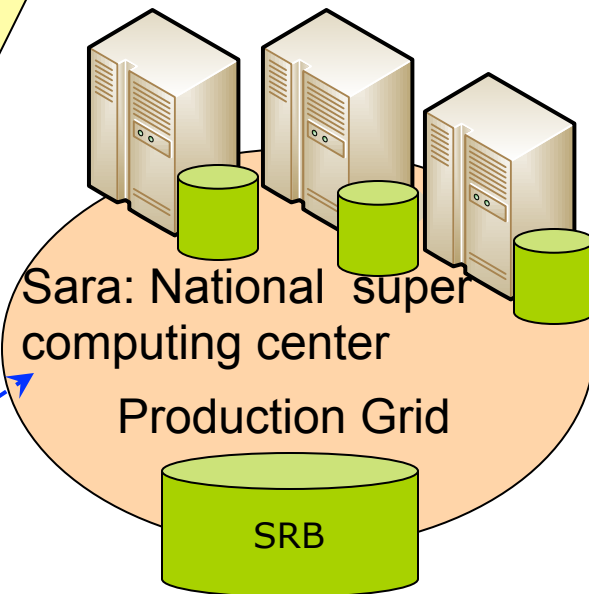
CLAMP: Connecting Language for Modules & Programs

HAMMER: Hybrid-based MatchMaker for e-Science

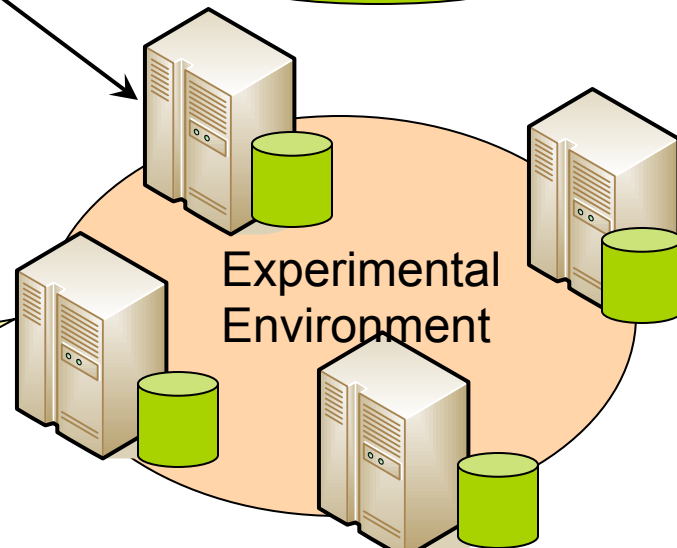
Resources

WSRF Services

- WS-VLAM engine
- workflow component repository

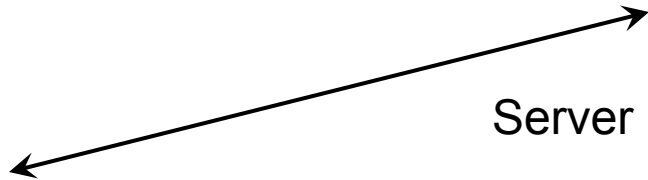
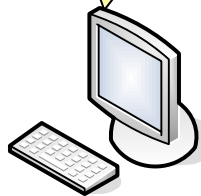
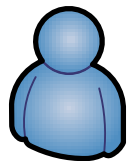


Server host

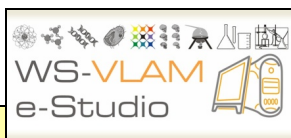


Computing Nodes

- Workflow components
- Grid Middleware → GT4

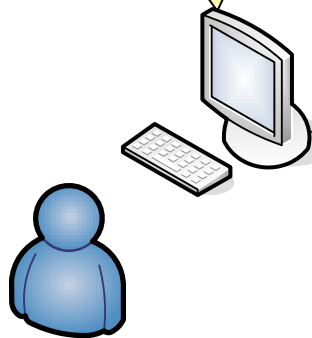


Example of usage Scenario ...

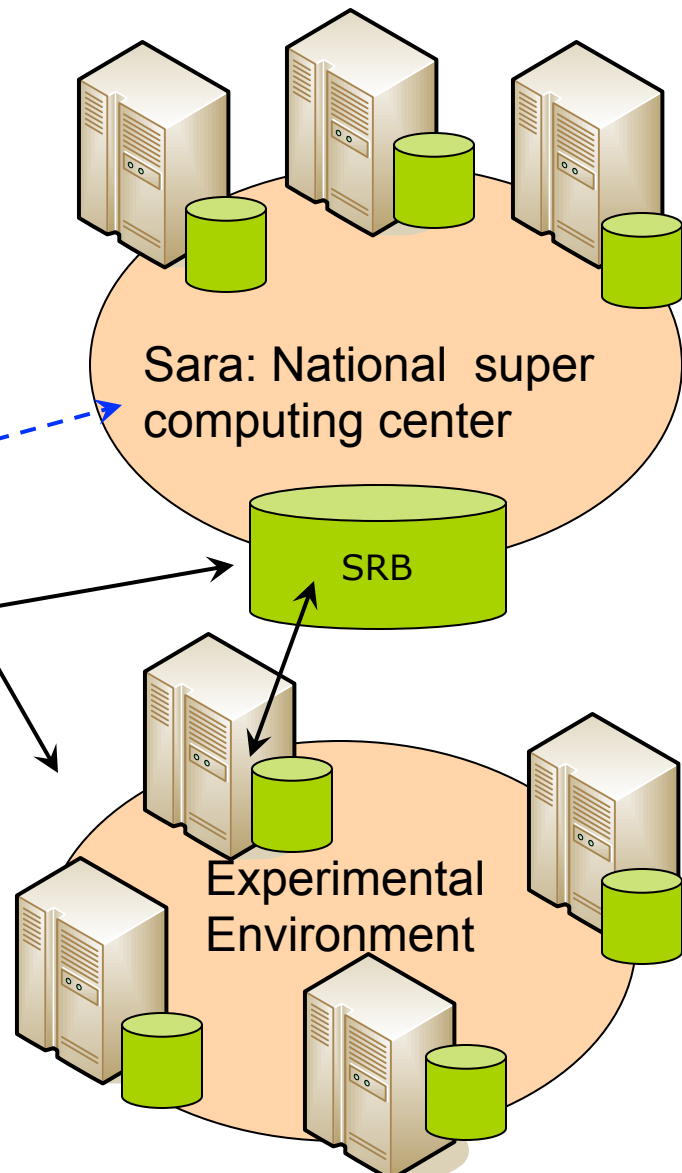


User start VL-e Studio

- 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



Sara: National super computing center

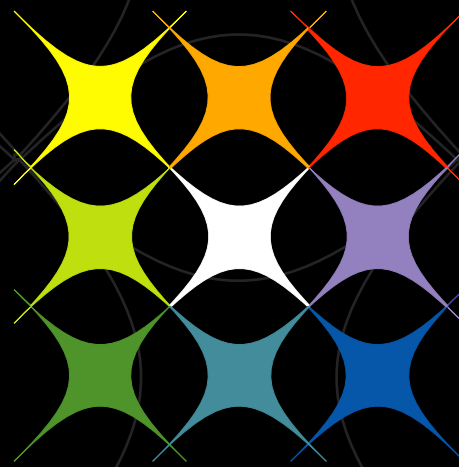
SRB

Experimental Environment



References

1. **A.S.Z Belloum**, Z. Zhao, and M. Bubak, *Workflow systems and applications* Future Generation Computer System. 25 (5): 525-527 (2009), doi:10.1016/j.future.2008.10.011.
2. M.A. Inda, M.F van Batenburg, M. Roos, **A.S.Z Belloum**, D. Vasunin, A. Wibisono, A. HC van Kampen, and T.M. Breit, *SigWin-detector: a Grid-enabled workflow for discovering enriched windows of genomic features related to DNA sequences*. BMC Research Notes 2008, 1:63 doi:10.1186/1756-0500-1-63.
3. Z. Zhao, **A.S.Z. Belloum**, and M. Bubak. *Support for Cooperative Experiments in VL-e: From Scientific Workflow to Knowledge Sharing*, In Proceedings of the IEEE 4th International Conference on e-science and Grid Computing (eScience 2008), pp. 329-330, 7-12 Dec. 2008, doi:10.1109/eScience.2008.120.
4. A. Wibisono, Z. Zhao, **A.S.Z. Belloum**, and M. Bubak, *A Framework for Interactive Parameter Sweep Applications*, In Proceedings of the 8th International Conference on Computational Science, Part III, pp. 481-490, 2008, doi:10.1007/978-3-540-69389-5_55.
5. Z. Zhao, **A.S.Z. Belloum**, C. De Laat, P. Adriaans, and L.O. Hertzberger, *Distributed execution of aggregated multi domain workflows using an agent framework*, In Proceedings of the *IEEE Congress on Services*, pp. 183-190, 9-13 July 2007, doi:10.1109/SERVICES.2007.30.
6. V. Korkhov, A. Wibisono, D. Vasyunin, **A.S.Z. Belloum**, *Interactive Dataflow Driven workflow Engine*, Scientific Programming, vol.15 (3), pp. 173-188 (2007), ISSN 1058-9244.
7. V. Korkhov, A. Wibisono, D. Vasyunin, **A.S.Z. Belloum**, C. De Laat, P. Adriaans, L.O. Hertzberger, *WS-VLAM: Towards a Scalable Workflow System on the Grid*, In Proceedings of the 2nd workshop on Workflows in support of large-scale science, in conjunction with HPDC 2007, June 25 2007, pp. 63-68, doi: 10.1145/1273360.1273372
8. A. Wibisono, D. Vasyunin, V. Korkhov, **AS.Z. Belloum**, *WS-VLAM: a GT4 based workflow management system*, In Proceedings of the International Conference on Computational Science (ICCS 2007), Lecture Notes in Computer Science, vol. 4489-2007, pp. 191-198, doi:10.1007/978-3-540-72588-6_34.
9. Ketan C. Maheshwari , Sílvia D. Olabarriaga , Charl P. Botha , Jeroen G. Snel , Johan Alkemade , **Adam S.Z. Belloum**, *Problem solving environment for medical image analysis*, in CBMS 2007, In Proceedings of the 20th *IEEE International Symposium on Computer-Based Medical Systems (CBMS '07)*, pp.165-170, 20-22 June 2007, doi:10.1109/CBMS.2007.87



vl-e

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