

vl·e



virtual laboratory for e·science



WS-VLAM a Grid-enabled workflow management system

Adam Belloum
Institute of Informatics
University of Amsterdam
a.s.z.belloum@uva.nl

GlobusEUROPE 2011, Lyon France , September 19 2011

UvA



UNIVERSITEIT VAN AMSTERDAM

Outline

- VLAMG & VL-e projects
 - Mission & Strategy
 - Infrastructure
 - Workflow an approach to model experiments
 - Successful stories
 - Rollout
 - Follow-up
- VL-e "Information has become the fuel of our knowledge society, and our ability to digest, understand and share it will determine scientific, economic and social progress"*

VLAMG (2000-2004) & VL-e (2004-2009)

Mission

To boost e-Science by creating an e-Science environment and carrying out research on methodologies.

Strategy

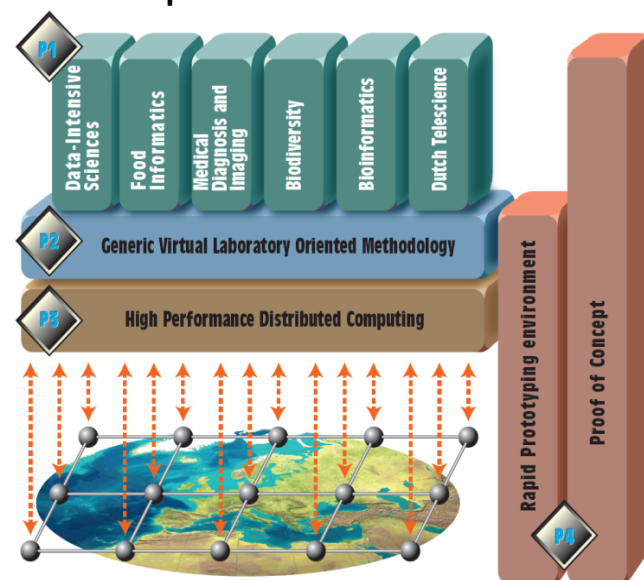
To carry out concerted research along the **complete e-Science technology chain**, ranging from applications to networking, focusing on new methodologies and re-usable components.

vl-e facts

budget 40 M, period 2004-2008
more than 20 consortium partners from industry and academia
director: prof. dr. L.O. Hertzberger
website: <http://www.vl-e.nl>

consortiumpartners

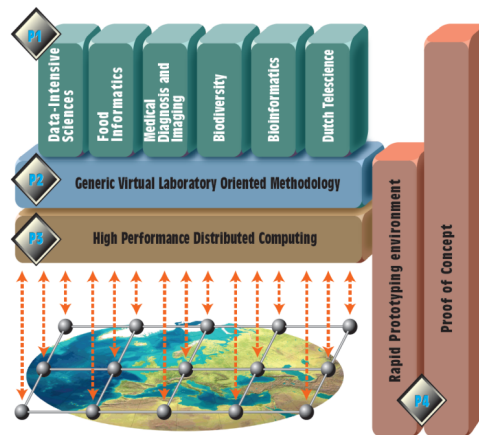
A&F Wageningen, AMC, CWI, DSM, Friesland Foods, FEI, FOM AMOLF, NBIC, Nikhef, IBM, LogicaCMG, Philips Research, Philips Medical, SARA, Top Institute Food and Nutrition, TNO Kwaliteit van Leven, TU Delft, Unilever, UvA-IBED, UvA-IvI, UvA-SILS, VU, VUmc, WTCW



Rapid prototyping Environment

DAS2/DAS3 cluster is composed of 5 clusters located at:

- University of Amsterdam (**28** nodes), Multimedia lab in UvA (**41** nodes),
- Vrije University (**79** nodes),
- Technical University in Delft (**64** nodes),
- Leiden Institute of Advanced Computer Science (**23** nodes)



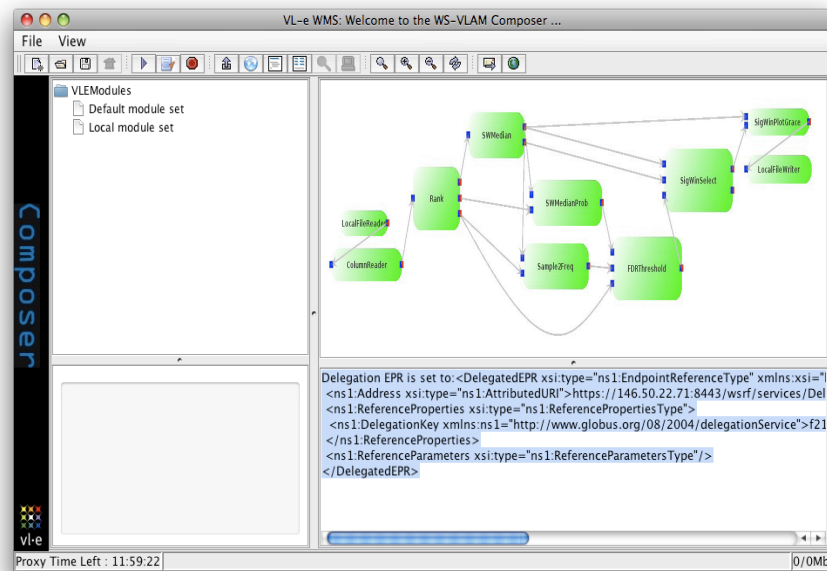


Outline

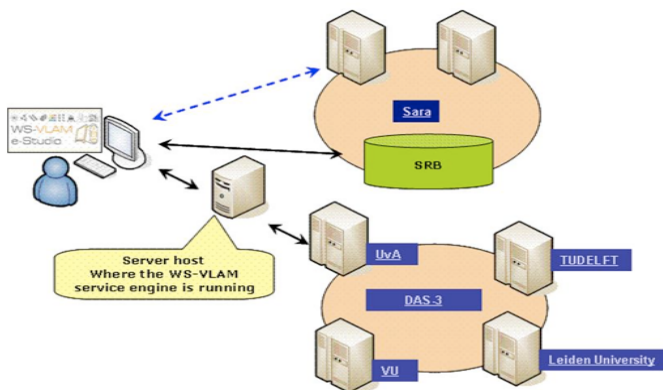
- VL-e project
- Mission & Strategy
- Infrastructure
- Workflow an approach to model experiments
- Successful stories
- Rollout
- Follow-up

Workflow management system

- Workflow management system** is a computer program that manages the execution of a workflow on a set of computing resources.



The user interface of the WS-VLAM a workflow management system developed in the VL-e project to execute application workflow on geographically distributed computing resources



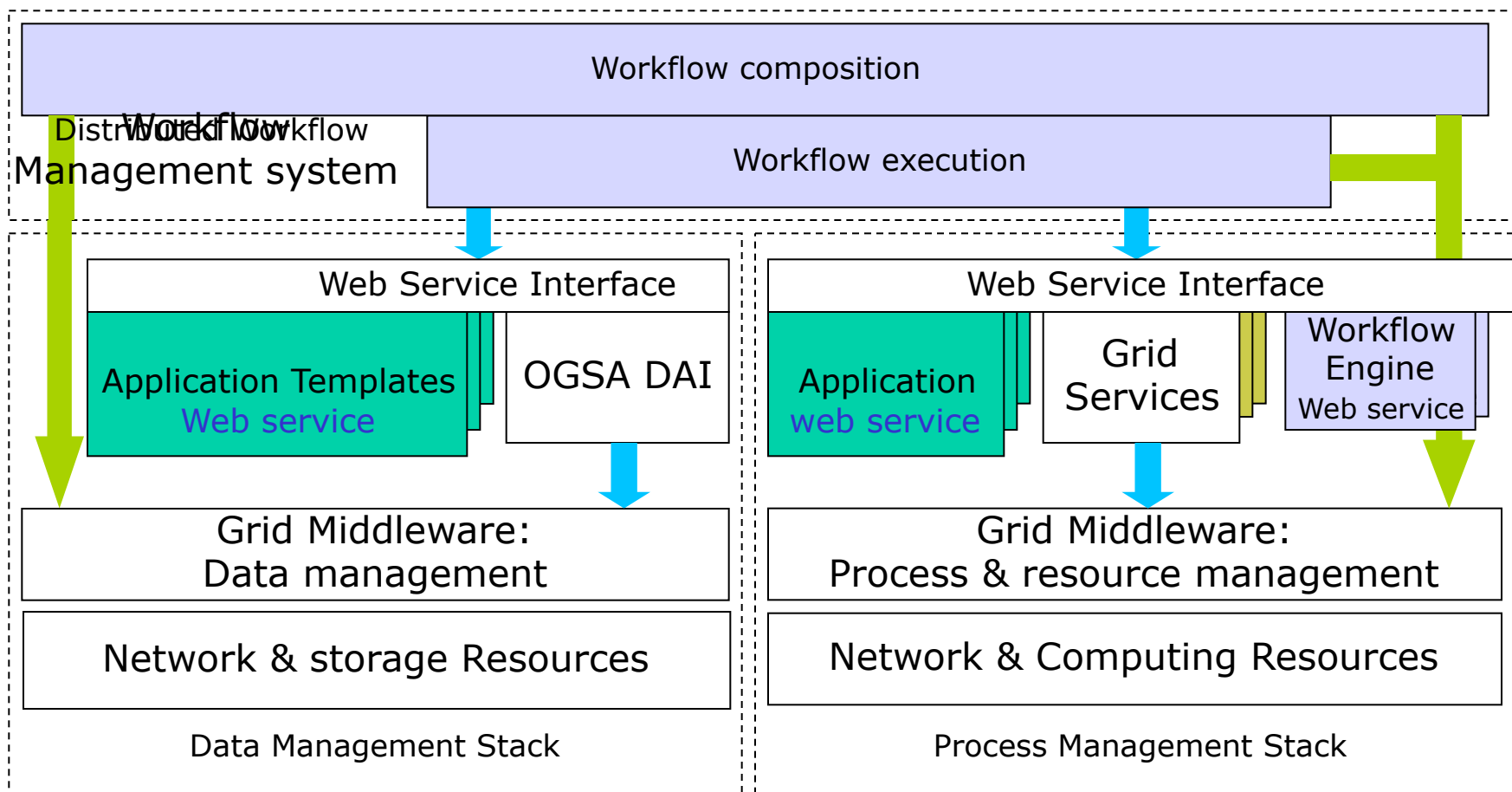
WS-VLAM workflow management system

- Follows the OGSA/WSRF standards.
- WS-VLAM **workflow engine** is implemented as a **WSRF compliant Web service**.
- The engine supports the **monitoring** and **interactive** runtime control of the application workflows. It offers the possibility to disconnect from a long running workflow without stopping it.

WS-VLAM Features enabled by Gt4

- Provide **streaming** facilities between applications executed on resource geographically distributed
- Workflow **provenance** using history-tracing XML-based provenance framework.
- Support the **composition** and the execution of **hierarchical** workflows.
- Support for **remote** graphical output.
- Provides the **detach/attach** capability for long running workflows.
- Provides a **monitoring** facilities based on the WS-notification.
- Provides **workflow farming** possibilities.
- etc.

A WSRF enabled workflow engine

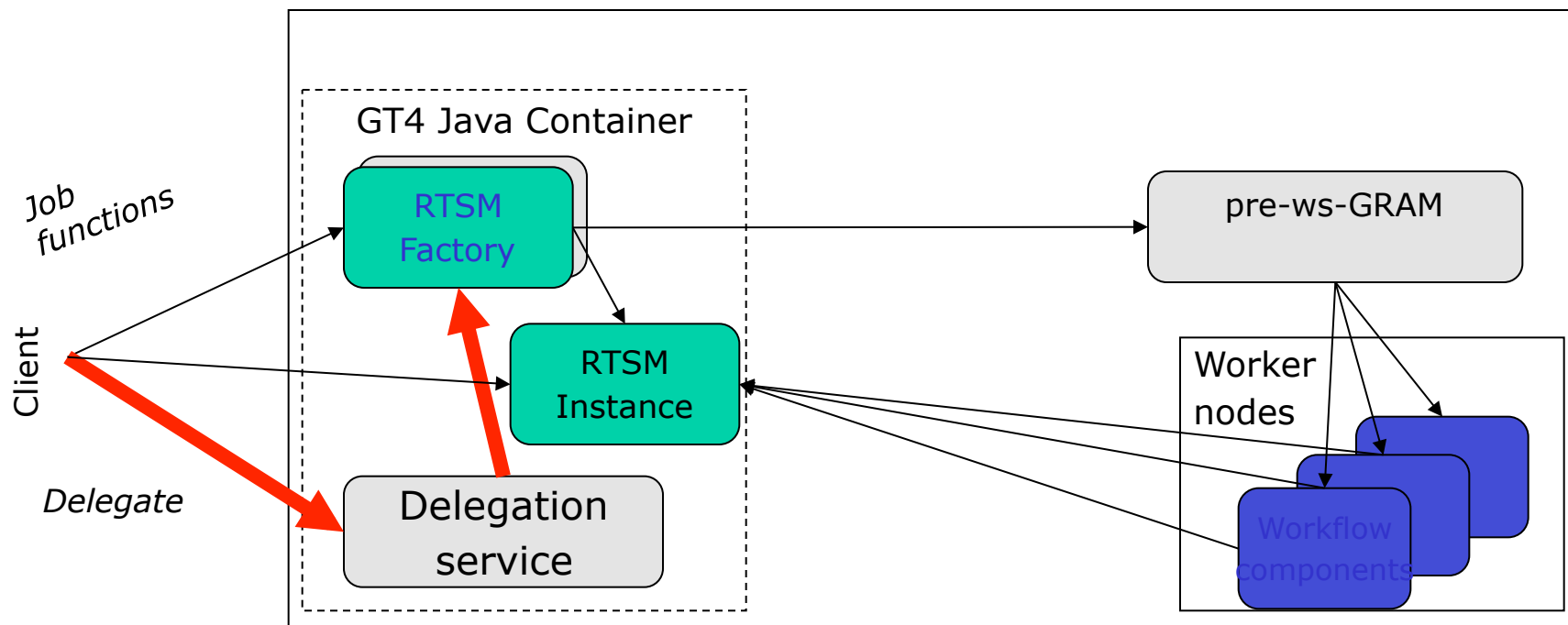


WS-VLAM Engine characteristics

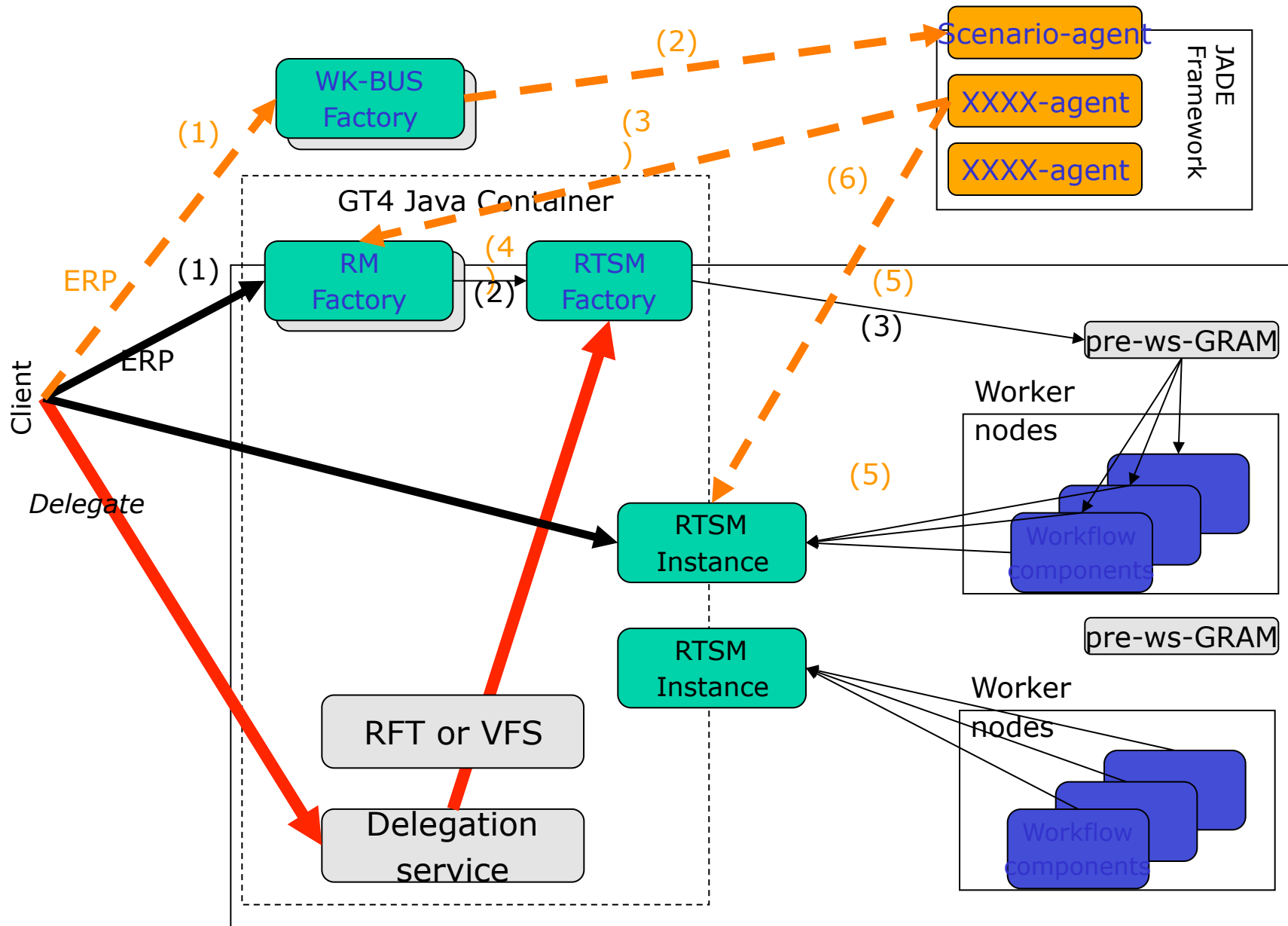
- Implemented as GT4 WSRF service
- Uses [GT4 delegation service](#) for security
- Uses [GSI enabled](#), private VNC for Graphical modules
- Uses [notifications](#) events provides by GT4 for monitoring the execution

WS-VLAM Engine: architecture (1/2)

Service host(s) and compute element(s)

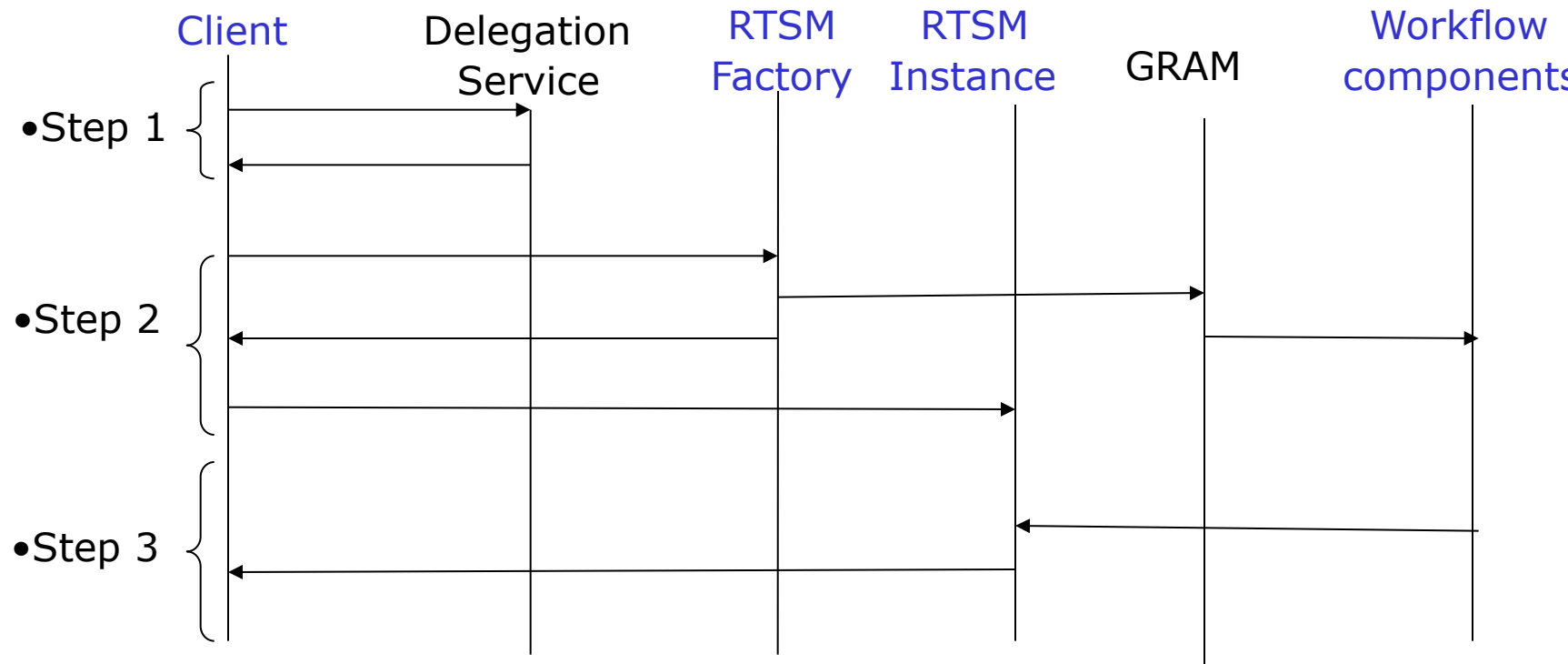


WS-VLAM Engine: architecture (2/2)

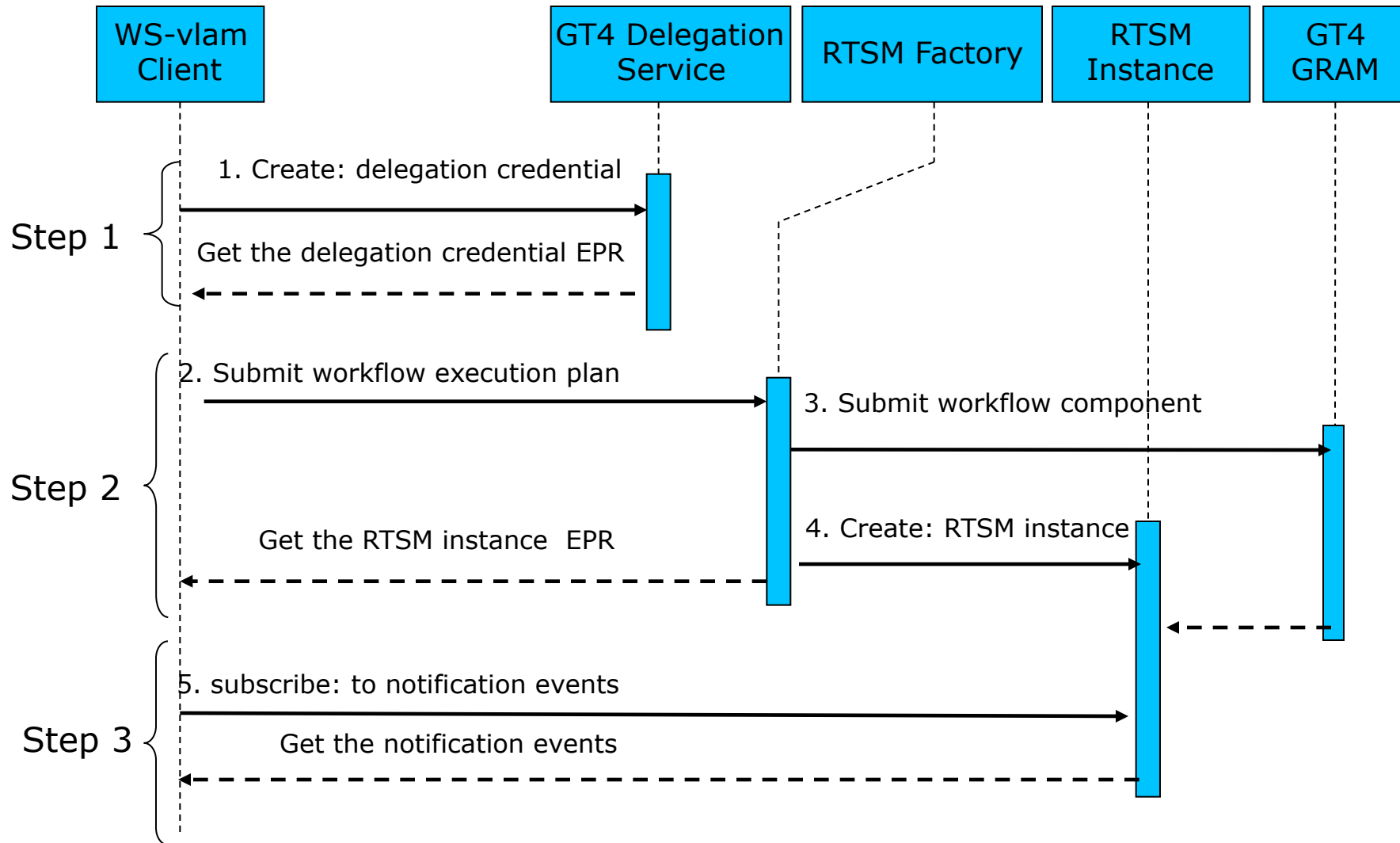


A three step execution process

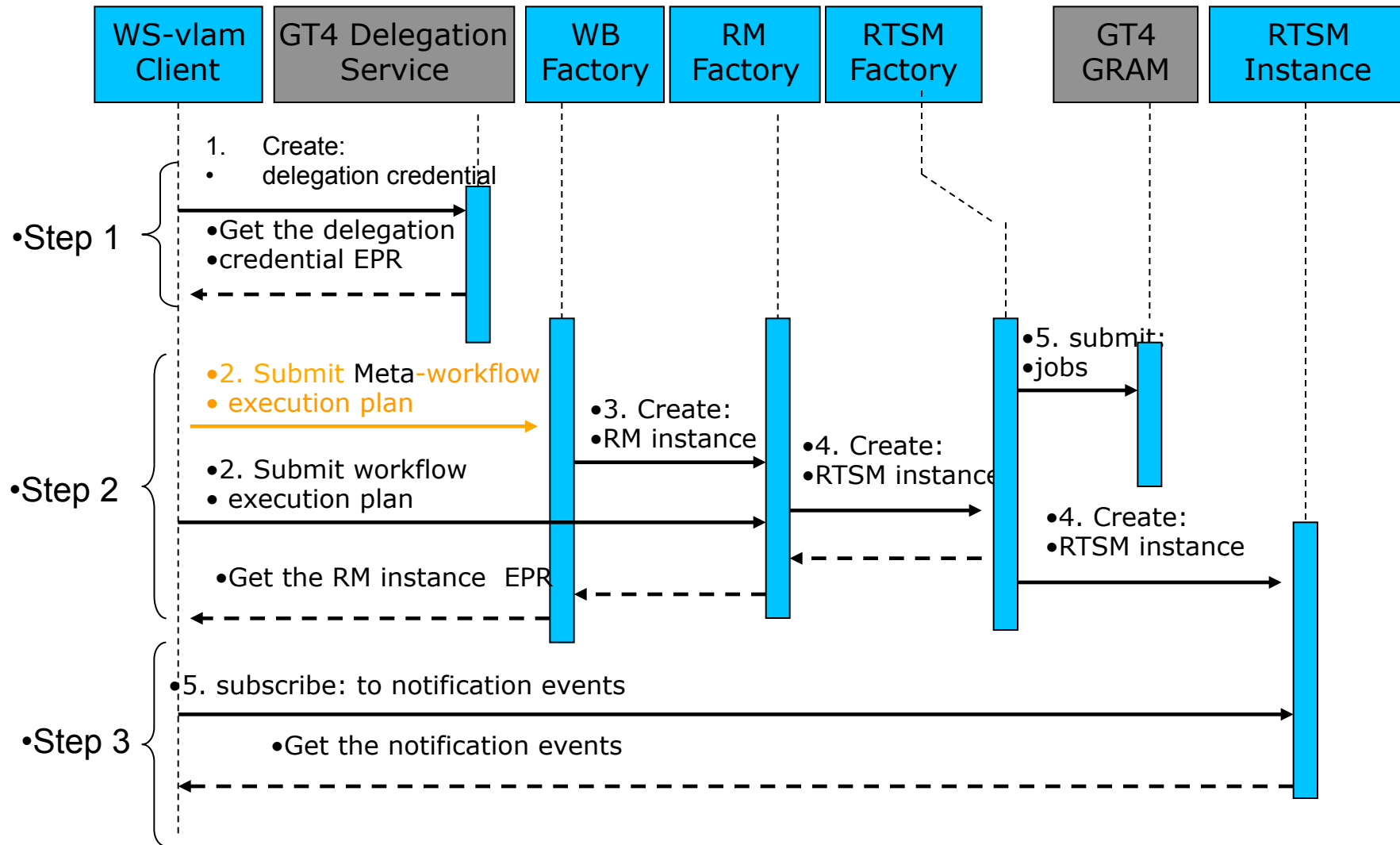
- Step 1: Create Delegated Grid Credentials
- Step 2: Instantiate the workflow components
- Step 3: Monitor the workflow execution



Sequence-diagram



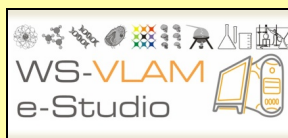
Sequence-diagram



Current depl^o

VLe Studio

- WS-VLAM composer
- VBrowser
- Semantic tools



SAW: Semantic Annotation for Workflow

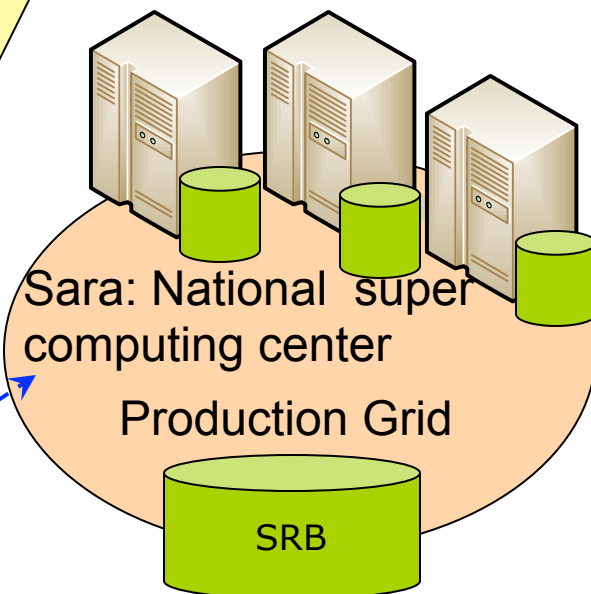
CLAMP: Connecting LANGUAGE for Modules & Programs

HAMMER: Hybrid-bAsed MatchMaker for e-Science

Resources

WSRF Services

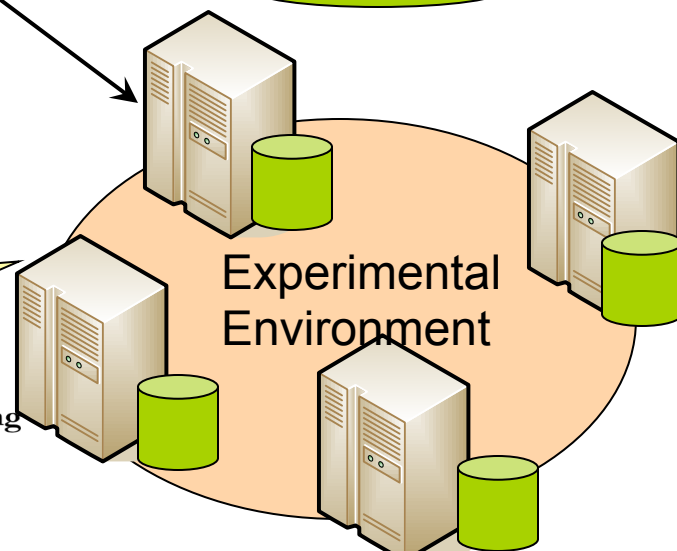
- WS-VLAM engine
- workflow component repository



Server host

Computing Nodes

- Workflow components
- Grid Middleware → GT4

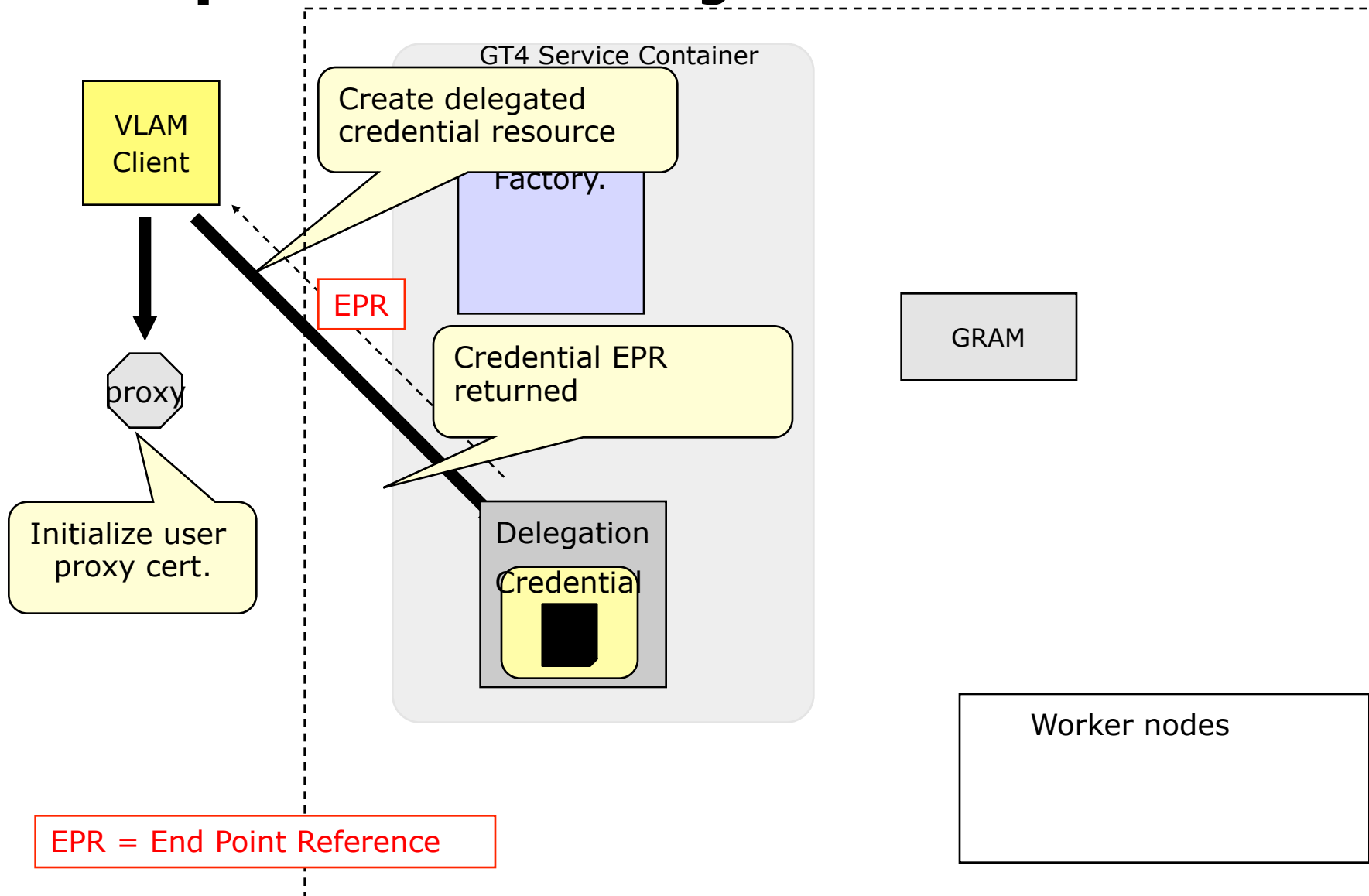


Collaborative e-Science experiments: from scientific workflow to knowledge sharing

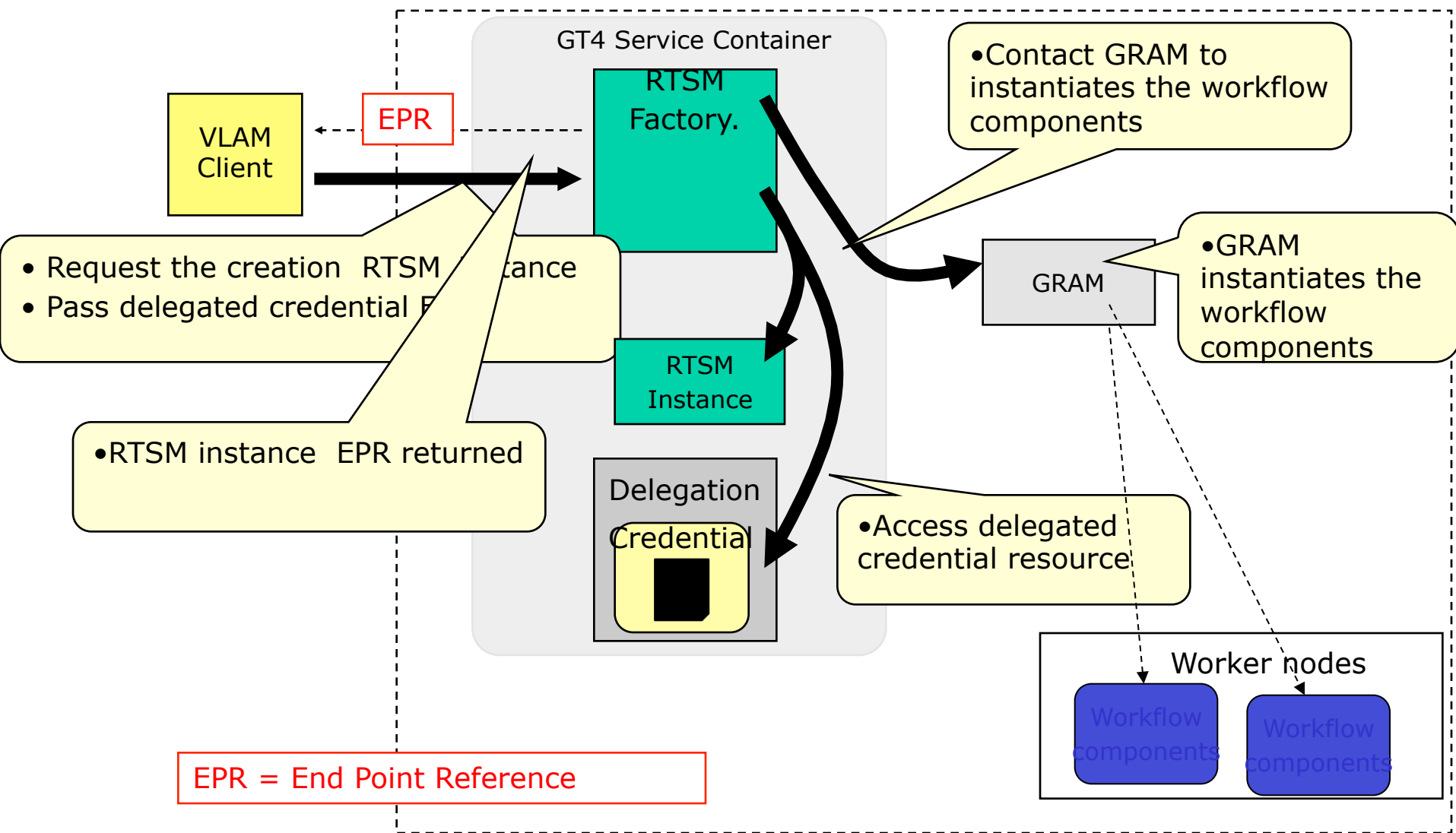
A.S.Z. Belloum, Vladimir Korkhov, Spiros Koulouzis, Marcia A Inda, and Marian Bubak

JULY/AUGUST, IEEE Internet Computing, 2011

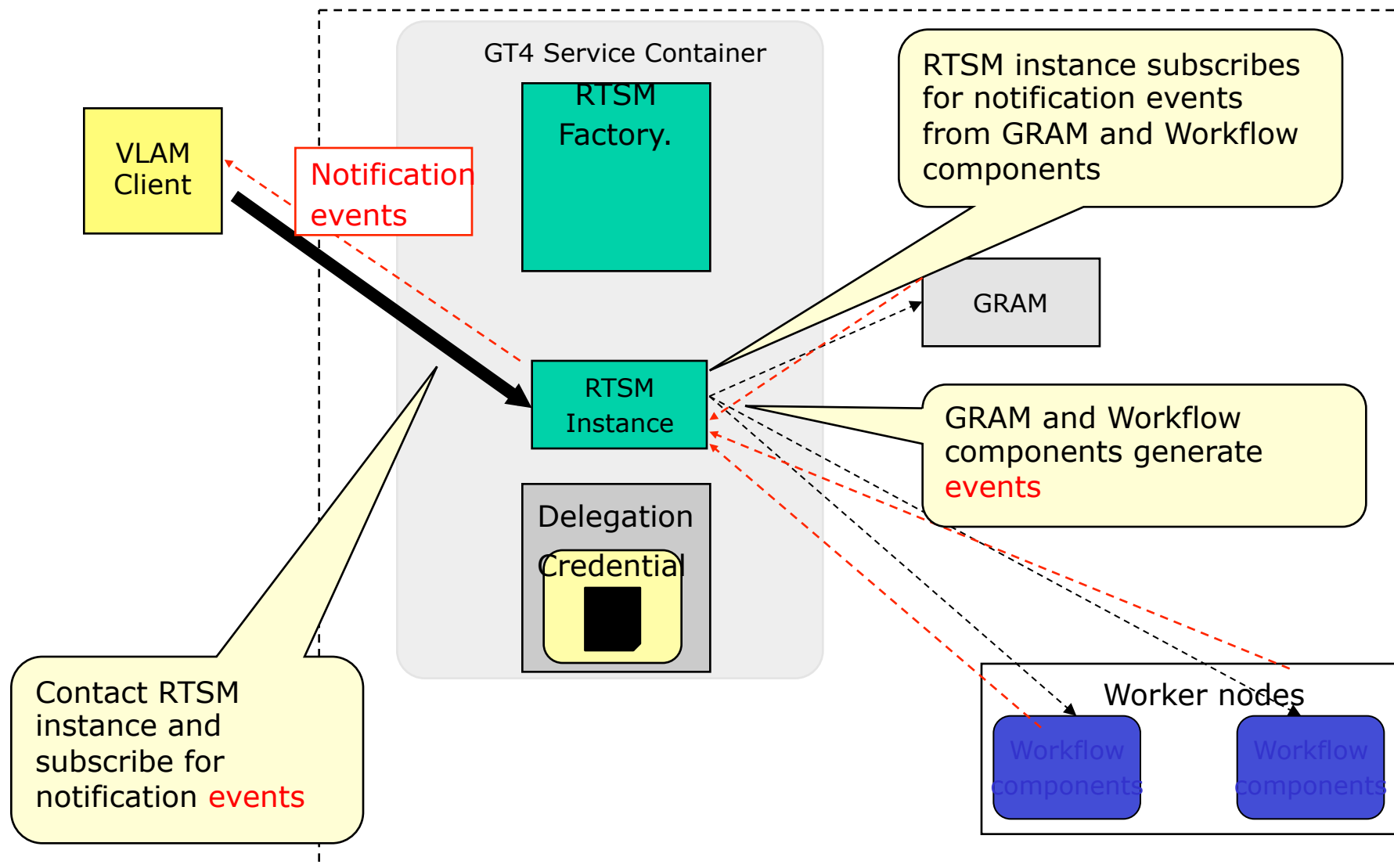
Step 1: Create Delegated Credential



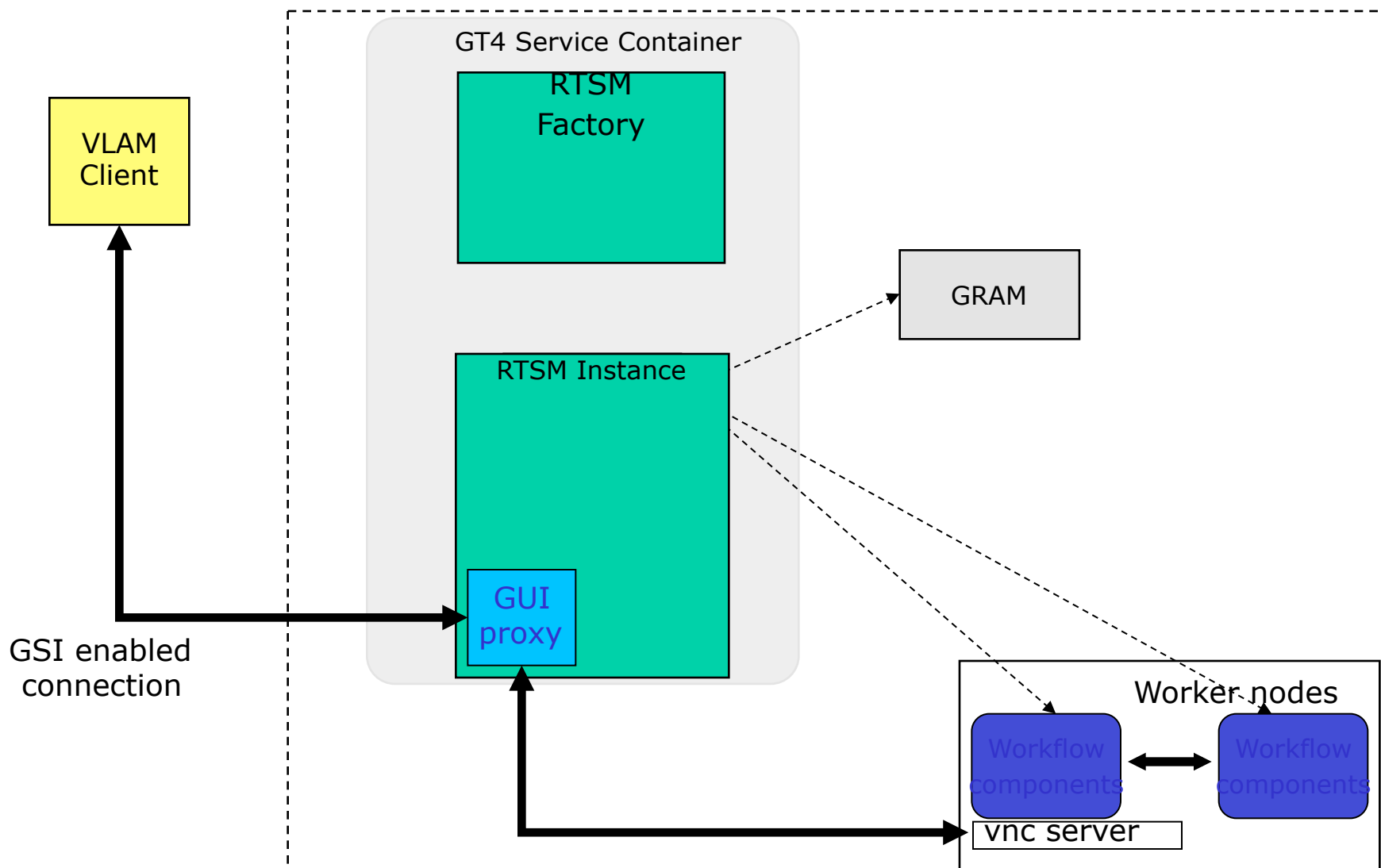
Step 2: instantiates the workflow components and Create a RTSM instance



Step3: monitors the application workflow



Workflow components with graphical output



WS-VLAM communication library

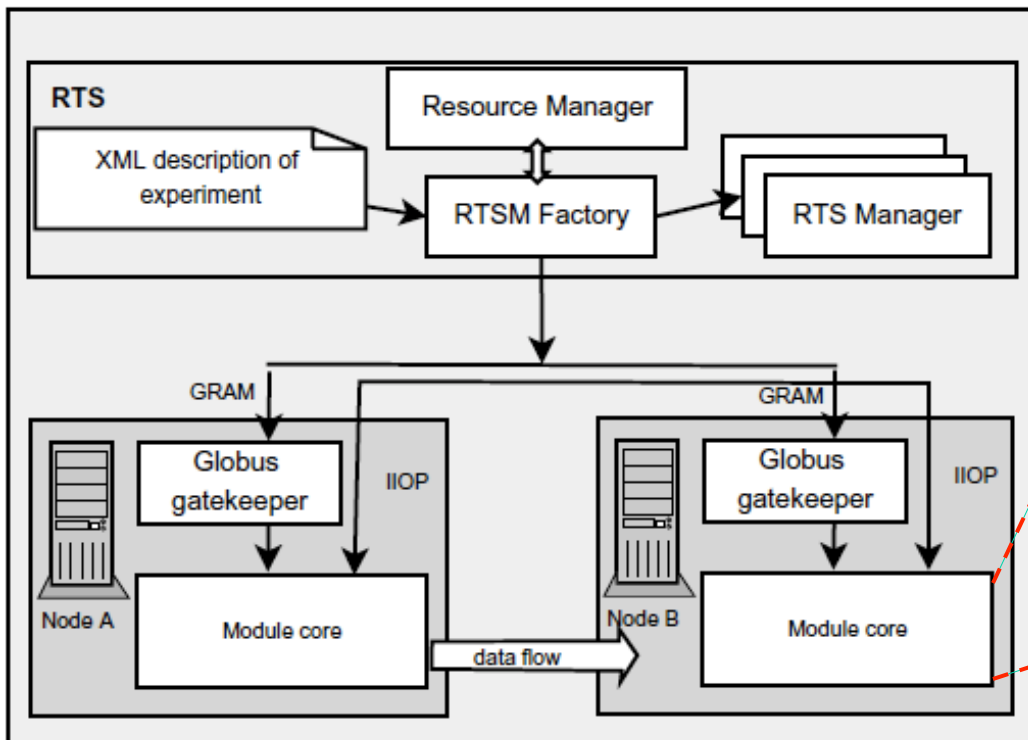


Fig. 1. Run-Time System Architecture.

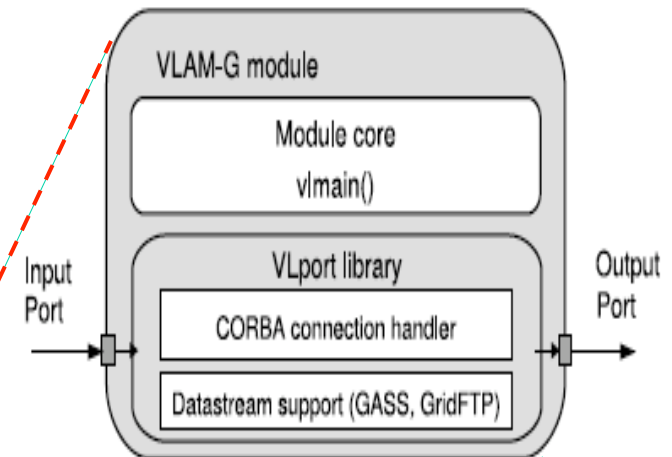


Fig. 4. Port library component architecture.

WS-VLAM Wrappers

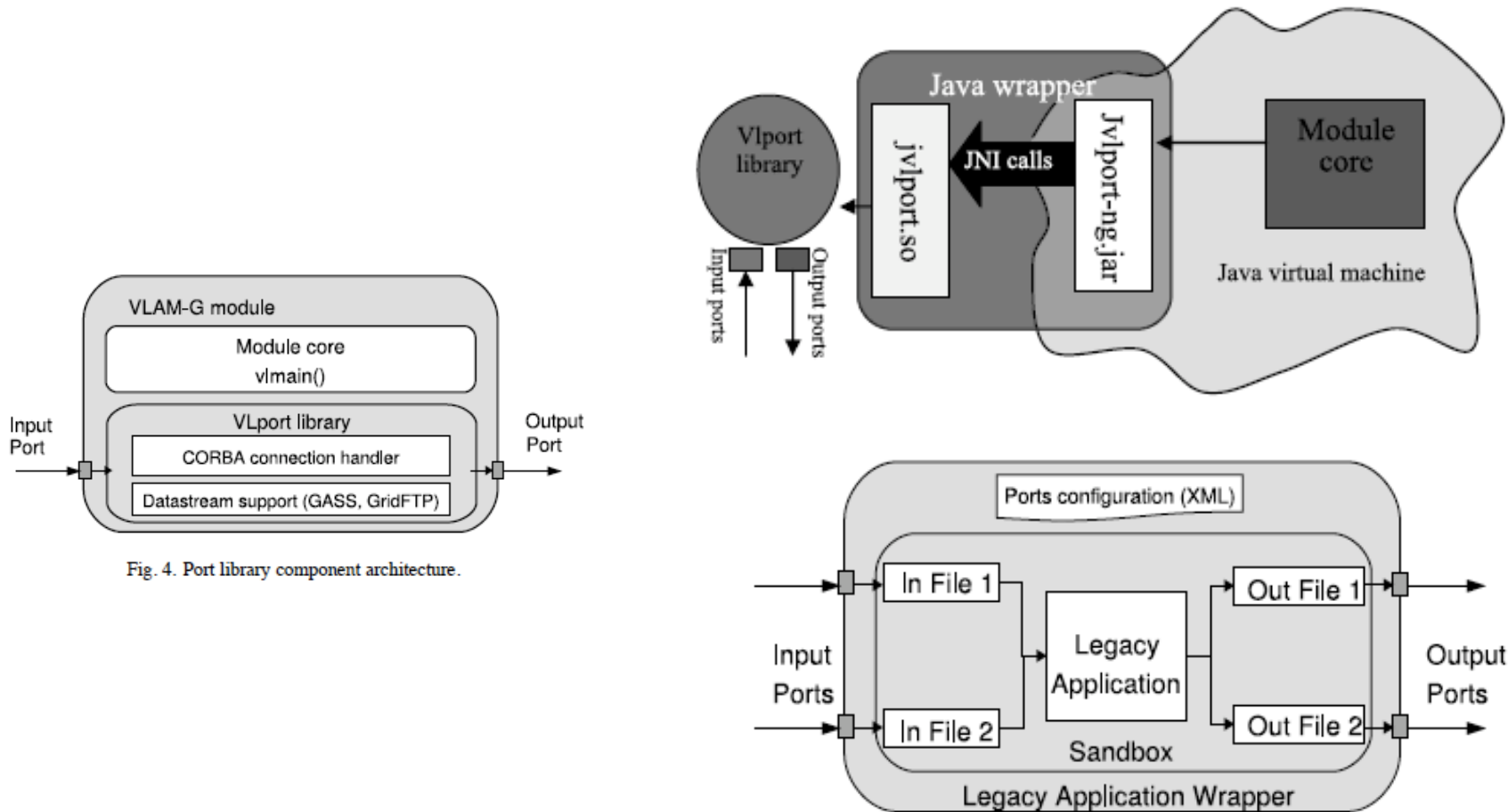


Fig. 4. Port library component architecture.

WS-VLAM communication library

- Data transfer rate as a function of the data block size (average of 10 measurements per each data-block)
- with the deviation not exceeding 5 percent)

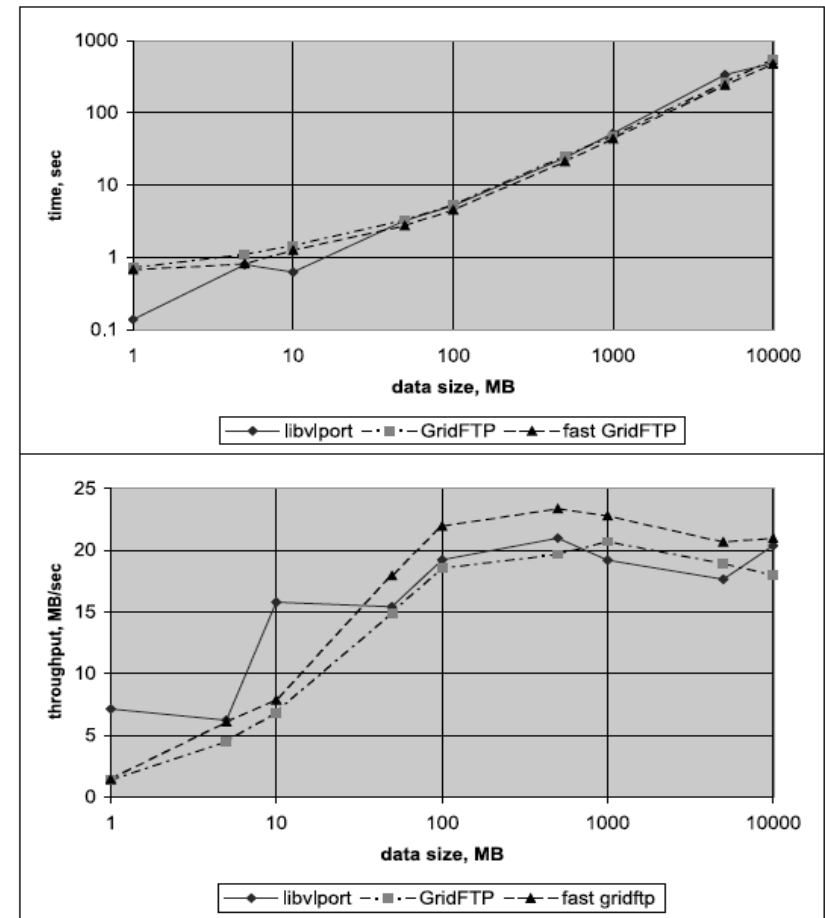
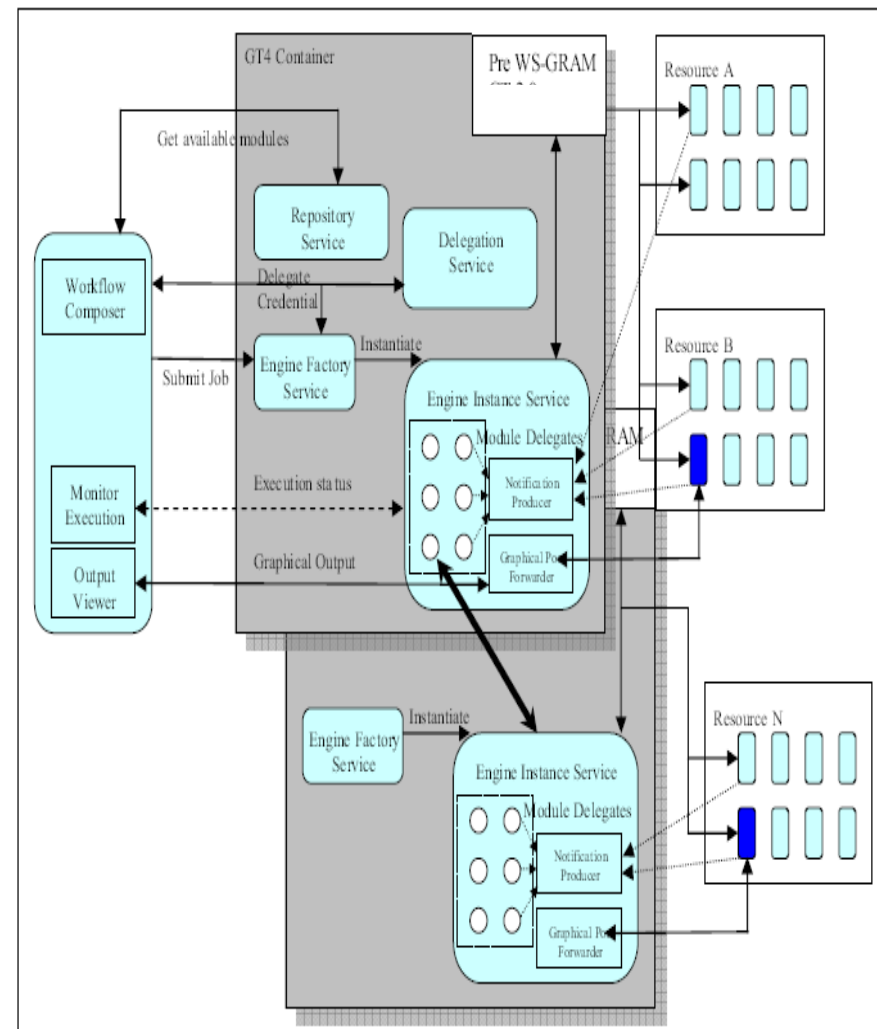


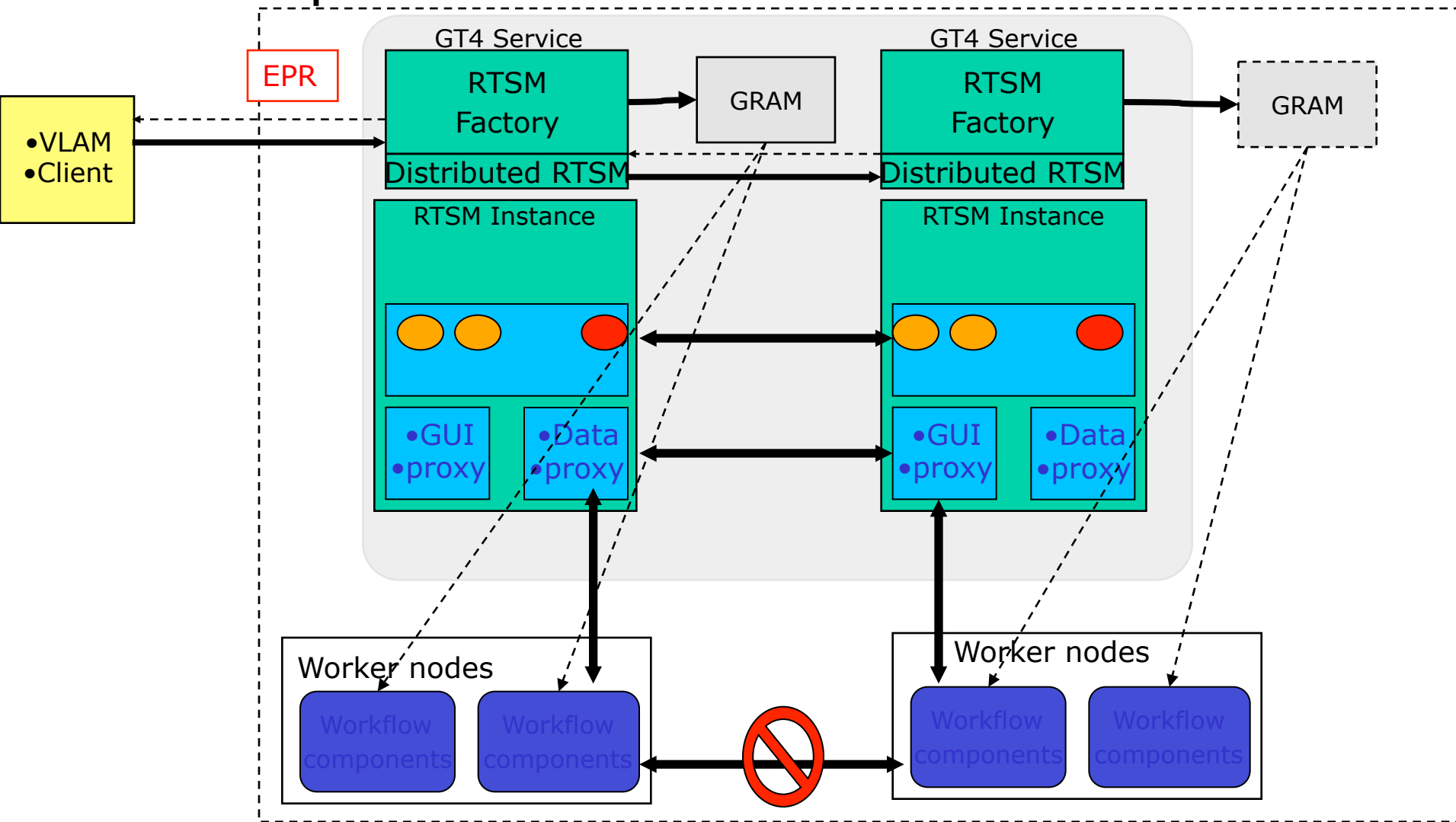
Fig. 7. Average performance of the RTS library on WAN compared with standard Globus data transfer tool.

Support hierarchical workflows ...

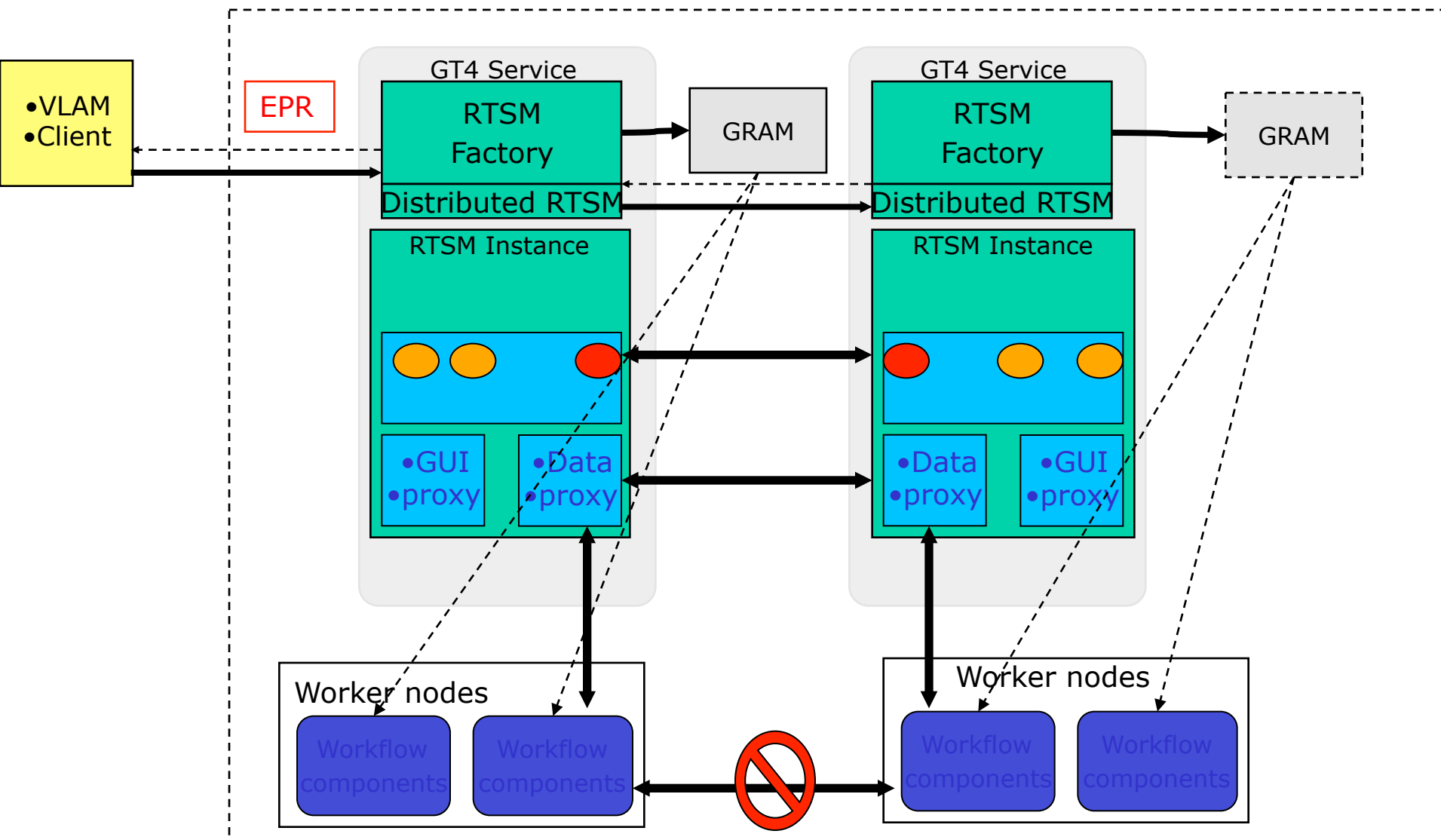
- workflow components can be **composed** of a **collection** of other **components**
- The advantage is now both the **composition** & the **execution** of complex workflows become **easier**



Step2: instantiates the distributed workflow components & Create a RTSM instances



Step2: instantiates the distributed workflow components & Create a RTSM



Outline

- VLAMG & VL-e projects
- Mission & Strategy
- Infrastructure
- Workflow an approach to model experiments
- **Successful stories**
- **Rollout**
- **Follow-up**

List of applications developed using WS-VLAM

- **sigWin detector** *[Dr. T. Briet Micro-Array Dept-UvA]*
- *Affymetrix Permutation* *[Dr. T. Briet Micro-Array Dept-UvA]*
- *Omnimatch* *[T.P van der Krif UU/Leiden]*
- wave propagation *[Dr. F.N van de Vosse , TUE]*
- Blast *[Dr. S. Olabariga, AMC]*
- gut microbiota *[Dr. F.J. Bruggeman, CWI]*
- Smart Infrastructure *[Dr. C. Delaat SNE-UvA]*
- Dynamic network control *[Dr. C. Delaat SNE-UvA]*
- GridSFEA, *[Dr. E. Else TU Munchen]*

More applications www.science.uva.nl/~gvlam/wsvlam/Applications



Outline

- VLAMG & VL-e projects
- Mission & Strategy
- Infrastructure
- Workflow an approach to model experiments
- Successful stories
- **Rollout**
- **Follow-up**

Other results of the VL-e project



More than one organisation



National Grid Initiatives & European Grid Initiative

- At the national level a grid infrastructure is offered to national and international users by the NGIs. BiG Grid is (de facto) the Dutch NGI.
- The 'European Grid Initiative' coordinates the efforts of the different NGIs and ensures interoperability
- Circa 40 European NGIs, with links to South America and Taiwan
- Headquarter of EGI is at the Science Park in Amsterdam

Conclusions

- GT4 offers useful **basic services** to develop a truly grid enabled workflow engine:
 - Delegation service & ws-notification mechanisms proved to be very useful
- **Separating** the workflow **composition** from **execution** suits better the Grid environments

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. Ilkay Altintas, Manish Kumar Anand, Daniel Crawl, Shawn Bowers, Adam Belloum, Paolo Missier, Bertram Ludascher, Carole A. Goble, Peter M.A. Sloot, Understanding Collaborative Studies Through Interoperable Workflow Provenance, *IPAW2010*, Troy, NY, USA
3. A. Belloum, Z. Zhao, and M. Bubak Workflow systems and applications , *Future Generation Comp. Syst.* 25 (5): 525-527 (2009)
4. Z. Zhao, A.S.Z. Belloum, et al., Distributed execution of aggregated multi domain workflows using an agent framework *The 1st IEEE International Workshop on Scientific Workflows*, Salt Lake City, U.SA, 2007
5. Zhiming Zhao, Adam Belloum, Cees De Laat, Pieter Adriaans, Bob Hertzberger Using Jade agent framework to prototype an e-Science workflow bus Authors *Cluster Computing and the Grid*, 2007. CCGRID 2007

References

1. 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.
2. 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
3. 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.
4. 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/>

<http://www.science.n/~gvlam/wsvlam/>