



Title: **Provenance Deliverable PC3**
Work Package 10

Authors: Omer F. Rana and Luc Moreau

Editor:

Reviewers:

Type: Public

Version: 1.0

Date: November 2006

Status: Final

Class: Confidential

Summary

This document covers the liaison and co-operation activities with the other IST projects under Strategic Objective “Grid-based systems for solving complex problems”. Contribution made by each partner across seven different tasks is highlighted.

PROVENANCE

Enabling and Supporting Provenance in Grids for Complex Problems

Contract Number:

511085

Members of the PROVENANCE Consortium

IBM United Kingdom Limited

University of Southampton

University of Wales, Cardiff

Deutsches Zentrum für Luft- und Raumfahrt e.V.

Universitat Politecnica de Catalunya

Magyar Tudományos Akadémia Számítástechnikai

és Automatizálási Kutató Intézet

United Kingdom

United Kingdom

United Kingdom

Germany

Spain

Hungary

Foreword

This collaboration workpackage covers the liaison and co-operation activities with the other IST projects under Strategic Objective “Grid-based systems for solving complex problems”. Each partner contributed 1 person-month to the work package, except UoS contributing 2 person-months and undertaking the work package leadership. According to the guidelines, the work package is structured around seven tasks, for each of which a contact person was identified. For each task, we revisit our plan, discuss our progress, and review concrete activities.

Contents

1 Task 1: Exploitation of synergies / technical concertation	5
2 Task 2: Joint fora for exchange and dissemination	6
3 Task 3: Co-ordination of standardisation efforts	7
4 Task 4: Repository of reference implementations and Grid middleware	8
5 Task 5: Collaboration on research inventory and roadmaps	8
6 Task 6: Indicators and impact assessment	9
7 Task 7: Training activities	9

1 Task 1: Exploitation of synergies / technical concertation

(Contact: Steven Willmott, steve@lsi.upc.es)

The Provenance project commits to supporting concertation activities proposed between work groups. It will also support the promotion of synergies between participating projects. Specifically, the project is involved in TG6 (Security) and TG7 (Semantic Grid), and also follows the activities of TG1 (Architecture).

For TG6, the goal was to identify how the conception of provenance, provenance architecture and data model to be developed by EU-Provenance might be useful in addressing certain project requirements of the SIMDAT project. Several generic provenance use cases for SIMDAT were developed, which could be addressed by the provenance architecture developed by EU Provenance. This included tracking replacement of parts in the automotive sector, components that developed faults during manufacturing process and having provenance information being stored in the same trust domain. Specific activities undertaken included:

- Luc Moreau. Provenance: a progress update was provided at a presentation to CT1 Technical Collaboration Meeting, Heathrow, February 2006.
- Victor Tan, gave a presentation at TG6 Collaboration meeting in Sankt Augustin, in June 2006.
- Jan 2005: Various discussions (through UPC) were undertaken with Hospital St. Pau (HSP), Barcelona, Spain – such as:
 - Joint work on organ transplant management application.
 - Regular inter project meetings with HSP lead Carrel FIS project (Spanish-funded HealthCare Project).
- Oct 2006: Ongoing collaboration (through UPC) with the EU CONTRACT project (<http://www.ist-contract.org/>)
 - The provenance architecture is being evaluated in the first phase of the project to determine if it can be used for the electronic contract monitoring mechanisms. This evaluation is expected to continue past the end of the Provenance project.
- Oct 2006: Ongoing collaboration (through SZTAKI) with the EU K4CARE project (<http://www.k4care.net/>)
 - The EHCR subsystem of the OTM application together with the provenance architecture is being evaluated in the first phase of the project, to determine if it can be used to support home care and check adherence to guidelines. This evaluation is expected to continue past the end of the Provenance project.

- Cardiff: Omer Rana and Tom Goodale were involved in collaboration with participants from the Cactus Toolkit (<http://www.cactuscode.org/>) – primarily to better understand how the Tool Suite and the Client Side Library could be used with data generated from the Cactus toolkit. This collaboration will continue after the project.
- Cardiff: Arnaud Contes and Omer Rana were involved in collaboration with the ProActive project (<http://www-sop.inria.fr/oasis/ProActive/>) – to better understand how to make ProActive provenance-aware. Two ideas were investigated: (i) how the ProActive container could be extended to support provenance queries; (ii) how the Tool Suite and Client Side Library could be used alongside ProActive. This collaboration will continue after the project.
- Cardiff: Vikas Deora, Arnaud Contes, Ian Wootten and Omer Rana were involved using the Tool Suite alongside the Triana workflow engine (developed as part of FP5 “GridLab” project). The Triana workflow engine is also being used in the FP6 “Data Mining Grid” project.
- On-going collaboration with “Scheduling Institute” (led by Ramin Yahyapour, Dortmund University) of CoreGRID in using Provenance information to establish trust relationships between parties involved in a Service Level Agreement. Omer Rana participated at the CoreGRID meeting in Sophia Antopolis, November 27, 2006.

2 Task 2: Joint fora for exchange and dissemination

(Contact: Steven Willmott, steve@lsi.upc.es)

The Provenance project commits to supporting and contributing to any joint fora and dissemination activities.

Specifically, EU Provenance contributed to the projects portal created for the IST projects under Strategic Objective “Grid-based systems for solving complex problems”.

Luc Moreau and his team at Southampton were involved in organizing the third workshop on provenance in May 2006. The first workshop, entitled “Data Derivation and Provenance”, took place in Chicago in October 2002 (<http://www-fp.mcs.anl.gov/foster/provenance/>). The second workshop, entitled “Data Provenance and Annotation”, was held in Edinburgh, in December 2003 (<http://www.nesc.ac.uk/esi/events/304/>).

Specific activities:

- IPAW’06 took place in Chicago, in May 2006
<http://www.sigmod.org/sigmod/record/issues/0609/sigmod-record.september2006.pdf>

- Proceedings of IPAW '06 published by Springer Verlag
<http://www.springer.com/uk/home/generic/search/results?SGWID=3-40109-22-173681711-0>
- First Provenance Challenge – <http://twiki.ipaw.info> – a special issue of CCPE Journal from the event is in preparation.
- John Ibbotson from IBM Hursley Laboratory represented the EU Provenance project at the Grid Concertation Meeting (part of the European Grid Technology Days) in Brussels, September 20-21, 2006.
- Omer Rana (Cardiff) represented the EU Provenance project at the “Knowledge Grid” workshop at the IST 2006 event in Helsinki, November 23, 2006.

3 Task 3: Co-ordination of standardisation efforts

(Contact: Luc Moreau, L.Moreau@ecs.soton.ac.uk)

The Provenance project commits to collaborate with other working groups and initiatives to promote and coordinate standardisation efforts.

A standardisation proposal for a Provenance architecture has been produced – as part of work being undertaken in WP3. The Provenance project has a representative (Luc Moreau) on the proposed “European GRID Standardisation Strategy Group”, and has contributed to the definition of the group’s charter, and filled COPRAS questionnaire.

In this context, we suggested that the group could help promote the use of standards that various projects propose, which could result in early adoption of standards proposals and associated reference implementations. Such adoption and resulting deployments could lead to the write up of best practice documents.

In order to be able to make standards reusable, we need to ensure that such standards address users’ needs. To that end, we have proactively interacted with the SIMDAT project, disseminating the ideas of provenance, an activity shared with TG6 (security). Close collaboration has also been undertaken with the GridOneD, Cactus and ProActive projects to promote the use of provenance within these projects. Additionally, the EU Provenance project has approached other EU projects to capture their requirements.

Specific activities:

- Contribution to the OGSA Data scenarios document.
- April 06: Luc Moreau gave a talk at IRISA, Rennes, to CoreGRID project.
- November 2006: Omer Rana gave a talk at ETSI, Sophia Antopolis to the CoreGRID project.

4 Task 4: Repository of reference implementations and Grid middleware

(Contact: John Ibbotson, john.ibbotson@uk.ibm.com)

The Provenance project commits to contribute to the building of repositories containing reference implementations. Its workpackage WP9, led by John Ibbotson, is in charge of building releases according to the technical annex's schedule. We will make such releases available in the adopted repository.

Specific activities:

- Public release of the Provenance Store software has been made available for public download. The source code is also available.
- The Client Side Library software is available for public download. The source code is also available.
- The Tools Suite is available for public download. The source code is also available.

5 Task 5: Collaboration on research inventory and roadmaps

(Contact : O. F. Rana, O.F.Rana@cs.cf.ac.uk)

The Provenance project commits to collaborate with other working groups on the creation and management of any research inventories and roadmaps. Specifically, the project has contributed to the AgentLink roadmap (coordinated by Michael Luck from Southampton).

Specific activities:

- The project has contributed (through UoS and UPC) to the AgentLink roadmap (coordinated by Michael Luck from Southampton) – <http://www.agentlink.org/roadmap/index.html> – M. Luck, P. McBurney, and O. Shehory and S. Willmott, “Agent Technology: Computing as Interaction (A Roadmap for Agent Based Computing)”, AgentLink, 2005. ISBN 085432 845 9.
- The project has contributed to the CoreGRID network of excellence and the roadmap activity being undertaken by Wolfgang Ziegler (Fraunhofer Institute).

6 Task 6: Indicators and impact assessment

(Contact: John Ibbotson, john.ibbotson@uk.ibm.com)

The Provenance project commits to providing a working document describing and managing the market and technology indicators appropriate for the Provenance project. To this end, John Ibbotson is member of the committee working on these issues. Guidelines issued by the committee will be instantiated for Provenance.

Work relating to this task has been subsumed under other activities – primarily tasks 1 and 5.

7 Task 7: Training activities

(Contact: Luc Moreau, L.Moreau@ecs.soton.ac.uk)

The Provenance project commits to providing training activities associated with project deliverables.

Specifically, material from the Provenance project has been included in the “Strategic Information Systems” MSc programme at Cardiff – with the “Distributed Multi-Agent Systems” (CMP629) course.

Preliminary discussions with Thierry Priol (CoreGRID) indicated that our dissemination and training activities could take place in the context of CoreGRID events. Despite volunteering provenance as a topic for the CoreGRID summer school, the topic was unfortunately not shortlisted by the organisers. Instead, a tutorial was organized at Southampton and via the Access Grid. Furthermore, Provenance-related material to the CoreGRID newsletter was provided [1].

Specific activities:

- Zheng Chen participated in the CoreGRID Summer School.
- Luc Moreau presented Provenance at “Grid At Work’05” organised by CoreGRID.
- 30-31/03 Southampton ran a ProvenanceTutorial with attendees from OntoGrid, myGrid, NeSc, NextGrid, DataMiningGrid
- August 2005 – Cardiff included material about Provenance in the “Distributed Workflow Management” tutorial delivered at the EuroPar conference in Lisbon, Portugal. 15 people attended the tutorial. The tutorial is now also archived at the <http://www.gridworkflow.org/> portal.
- May 2006 – an extended version of the workflow tutorial from EuroPar was delivered at the IEEE CCGrid/GridAsia conference. 25 people attended the tutorial – from the US, Europe and the Asia Pacific regions.

Other relevant collaboration activities

Specific activities:

- Use case development focusing on OntoGrid (UoS), myGrid (UoS), GridOneD (UWC), BDWORLD (UWC), Electronic Health Care Records (SZTAKI and UPC) and business compliance (IBM).

References

- [1] CoreGRID Newsletter, “Provenance for Grid Applications”. Available at: <http://www.cetic.be/article452.html>, No. 6, June 2006.