



Project no: 511085

STREP

Thematic Priority

## **Provenance: 18 Month Project Report**

Deliverable D1.1.3

“Enabling and Supporting Provenance in Grids for Complex Problems”

Period covered: 01-09-2005 to 28-02-2006

Date of preparation: 27-02-2006

Start date of project: 01-09-2004

Duration: 24 Months

Project coordinator name: John Ibbotson

Project coordinator organisation name: IBM United Kingdom Limited

---

Copyright © 2006 by the PROVENANCE consortium

*The PROVENANCE project receives research funding from the European Community's Sixth Framework Programme*

**Members of the PROVENANCE consortium:**

- IBM United Kingdom Limited United Kingdom (IBM)
- University of Southampton United Kingdom (UoS)
- University of Wales, Cardiff United Kingdom (UWC)
- Deutsches Zentrum für Luft- und Raumfahrt e.V. Germany (DLR)
- Universitat Politècnica de Catalunya Spain (UPC)
- Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutató Intézet Hungary (STA)

## **Foreword**

This document has been edited by John Ibbotson (IBM) based on input from project partners.

## Table of Contents

<b>1</b>	<b>Executive Overview .....</b>	<b>5</b>
1.1	<i>Management Report Overview .....</i>	5
<b>2</b>	<b>General Management .....</b>	<b>6</b>
2.1	<i>Project objectives .....</i>	6
2.2	<i>Objectives for the Period.....</i>	6
2.3	<i>Major Achievements during the Period .....</i>	7
<b>3</b>	<b>Workpackage Progress .....</b>	<b>8</b>
3.1	<i>WP2: Requirements.....</i>	8
3.2	<i>WP3: Architecture .....</i>	9
3.3	<i>WP4: Security .....</i>	13
3.4	<i>WP5: Scalability.....</i>	14
3.5	<i>WP6: Tools and Setup .....</i>	15
3.6	<i>WP7: Application 1 – Aerospace.....</i>	18
3.7	<i>WP8: Application 2 – Organ Transplant Management.....</i>	19
3.8	<i>WP9: Implementation, Integration and Test .....</i>	22
3.9	<i>WP10: Collaboration .....</i>	23
<b>4</b>	<b>Consortium Management .....</b>	<b>26</b>
4.1	<i>Management Activities during Months 13-18.....</i>	26
4.2	<i>Management Issues .....</i>	26
	<b>Appendix 1 - CPL Agreement .....</b>	<b>30</b>

# 1 Executive Overview

This Project Report provides details of the Provenance project for the third 6 month period (September 2005 to February 2006). It describes the workpackages that were active during that period, which milestones, deliverables or other tangible achievements were completed and relates these to the overall project objectives. The report also summarises the activities planned for the next 6 months and provides indications of any potential changes to the workplan. Exploitation opportunities identified by the project are described in a separate revised deliverable D1.3.1, Exploitation Strategy.

At the Provenance review held in November 2005, the reviewers provided a number of comments related to the project. The consortium will provide a written update to address these comments by the end of March 2006.

## 1.1 Management Report Overview

The Project Report is structured with the following sections:

**General Management:** Describes the overall project objectives and which objectives were achieved during this reporting period. It also itemises the major achievements of the project so far.

**Workpackage Progress:** Provides a detailed breakdown of all workpackages active during this reporting period together with milestones achieved and deliverables produced. It also details any deviations from the workpackage plan together with expected activities for the next 12 months and where appropriate, details of any knowledge dissemination activities associated with the workpackage.

**Consortium Management:** Provides a description of the project management for the reporting period. It includes details of any plan deviations together with reasons and corrective actions. It also highlights any issues identified by the project coordinator.

## **2 General Management**

### **2.1 Project objectives**

The overarching aim of the Provenance project is:

*To design, conceive and implement an industrial-strength open provenance architecture for Grid computing, and to deploy and evaluate it in complex grid applications (aerospace engineering and organ transplant management).*

Specifically, the objectives of the project are:

1. To specify the contents of provenance in relation to workflow enactment.
2. To design and implement a scalable and secure distributed co-operation protocol to generate provenance data in workflow enactment.
3. To conceive and implement tools to navigate, harvest and reason over provenance data, also in a scalable and secure manner.
4. To design and engineer a scalable and secure software architecture to support provenance generation and reasoning.
5. To deploy and evaluate the provenance system in two different grid applications, namely aerospace engineering and organ transplant management.
6. To propose a draft provenance specification for input to an open standardisation process thereby contributing to the standardisation efforts in this area within the Grid and Web Services architecture domains.

### **2.2 Objectives for the Period**

According to the project plan provided in the Provenance Technical Annex, objective 1 from section 2.1 above would be achieved at the end of the first six month period. To meet this objective, the Provenance project will have provided:

1. A set of User and Technical Requirements for provenance, for the two selected applications, and also for related applications with provenance needs, so as to ensure the generality of the project effort.
2. A pre-prototype of a provenance system; the pre-prototype will consist of an assemblage of existing software, including an exemplar provenance service and a prototype implementation of a provenance recording protocol of the PASOA project ([www.pasoa.org](http://www.pasoa.org)).

At the end of the first twelve month period, objective 2 would be achieved. To support this objective, the Provenance project will have provided:

1. A project internal deliverable detailing a logical architecture for Provenance.
2. Version 1 of documents providing specifications of the Security and Scalability characteristics of the logical architecture.
3. A specification for the Tools and Setup components of the Provenance architecture.
4. Mappings of the two example applications for Aerospace design and Organ Transplant Management to the Provenance logical architecture.
5. A functional prototype that implements the Provenance recording and query interface functionality. This prototype will be enhanced over the next six month period to include the security and scalability features identified in the associated specifications

At the end of the third six month period, objective 3 would be achieved. To support this objective, the Provenance project will have provided:

1. An initial proof of concept tools prototype was implemented based on the eXo portal framework

---

Copyright © 2006 by the PROVENANCE consortium

*The PROVENANCE project receives research funding from the European Community's Sixth Framework Programme*

2. A workflow visualization portlet was implemented using the framework following close interaction between the tools and application workpackages
3. A specification of the setup protocol has been produced based on the requirements of the application workpackages.

## **2.3 Major Achievements during the Period**

### **2.3.1 Achievements for Months 13-18**

The following are considered the major achievements and activities during this six month period:

1. A project face to face meeting was held at SZTAKI Budapest during October 26-28 2005
2. A project face to face meeting was held at UWC Cardiff during February 22-24 2006
3. The first annual project review was held in Brussels on November 14<sup>th</sup> 2005
4. Following the project review, the recommendations of the review panel were implemented by the project.
5. The project provided briefings for IT analysts in November 2005 and February 2006.
6. Although the Requirements workpackage WP2 has formally completed, it has instigated a feedback process with the groups that submitted requirements to the project following suggestions made at the annual review.
7. The Architecture workpackage WP3 continued to refine and develop the Provenance Architecture. The draft logical architecture provided as an additional deliverable provided at month 12 was further expanded and reviewed to produce the Final Architecture deliverable D3.1.1 at month 18.
8. The Architecture workpackage WP3 has identified the need for a methodology to support Provenance. This will provide a best practice approach to designing distributed systems that are Provenance aware. This methodology is not a contracted deliverable of the project but will be provided during the next period.
9. The Security workpackage WP4 has continued to develop the specifications resulting in a final draft. The decision to base the reference implementation on the Globus Toolkit (GT4) has allowed the specification to concentrate on the use of the GT4 security framework to meet the requirements collected in WP2.
10. The Scalability workpackage WP5 has developed a further specification draft which refines the architectural and implementation recommendations identified in deliverable D5.1.1 into a form that is closer to an implementation using GT4. This workpackage has identified the need to use the Web Services Resource Framework (WS-RF) which has influenced the Implementation, Integration and Test workpackage WP9.
11. The Aerospace Application workpackage WP7 has completed the description of its application mapping to the Provenance architecture and deployed an early Provenance Store implementation into the TENT environment.
12. The Organ Transplant Management (OTM) workpackage WP8 has completed its analysis of the Provenance requirements for the OTM scenario. It has completed a preliminary OTM/EHCR demonstrator with integration to a Provenance Store.
13. The Implementation, Integration and Test workpackage WP9 has continued to develop the reference implementation for delivery at the end of month 18. To speed the deployment to partners, an existing implementation (PreServ) from UoS has been used which allows early integration using a client API that will also be used for the WP9 implementation. The reference implementation has been ported to GT4 to take advantage of the WS-RF and security support within the toolkit.
14. The project has continued to collaborate with other research projects within Europe and worldwide.

### 3 Workpackage Progress

The following sections describe in more detail the activities in each of the non-management workpackages WP2 to WP10. The project management activities are described in section 3 – Consortium Management.

#### 3.1 WP2: Requirements

##### 3.1.1 Objectives

Formally this workpackage has ended, however the first year review report contained recommendations related to this workpackage. Therefore activities were carried put in this workpackage in accordance with the recommendation of the first year project review: *"There should be a follow-up with those external projects that provided requirements. They should receive feedback on what has happened with their requirements and keep them updated on the progress of the project to attract them as potential end users."*

##### 3.1.2 Progress

Mappings from user requests to accepted software requirements and the corresponding design features for each client project were created and sent to the following client projects:

1. Combechem
2. Diligent
3. DataMiningGrid
4. eDiamond
5. GENSS
6. HealthCare and Life Sciences Framework
7. myGrid
8. OTM application
9. TENT
10. Traffic Management Application (K-WFGrid project)

##### 3.1.3 Deviations

None

##### 3.1.4 Deliverables

No deliverables are planned.

##### 3.1.5 Plans for next 6 months

The evaluation of client feedback and continued contact with them is planned.

##### 3.1.6 Knowledge Dissemination

None

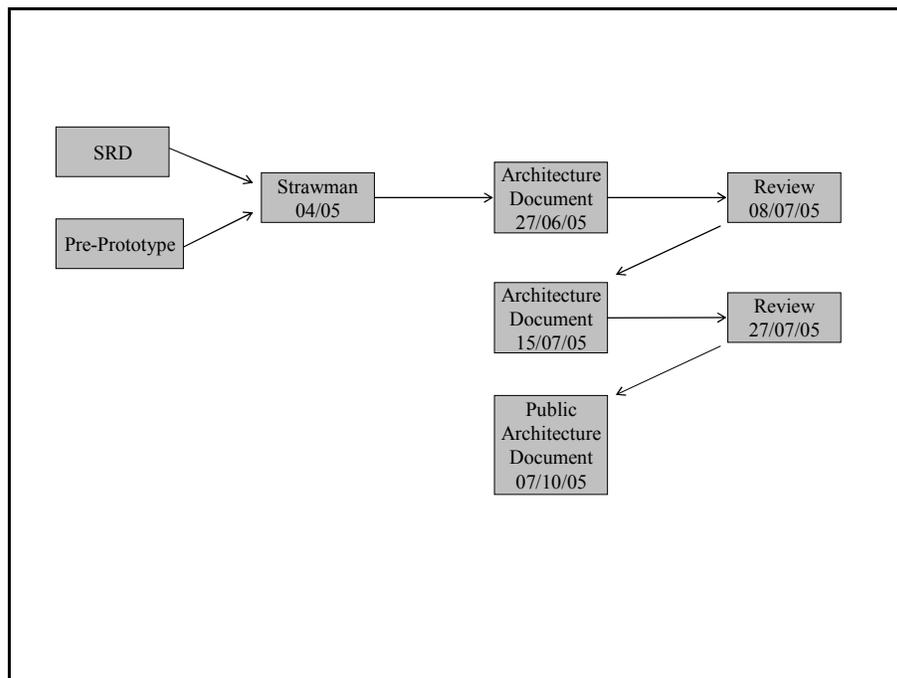
## 3.2 **WP3: Architecture**

### 3.2.1 **Objectives**

The architecture will be described in several ways in order to address the concerns of its multiple stakeholders, end-users, developers and system managers. Different views of the architecture will be adapted to this end: a logical view addressing functional requirements, process and physical architectures taking into account non-functional requirements and physical deployment, and development architecture identifying modules and libraries at the level of software development. We will adopt an iterative design process, deriving a first architecture definition (strawman) from the technical requirements, iterating it into a final architecture definition, using feedback from the different application specific studies, tool design, and security and scalability analyses. The architecture will ultimately be defined into a standardisation proposal.

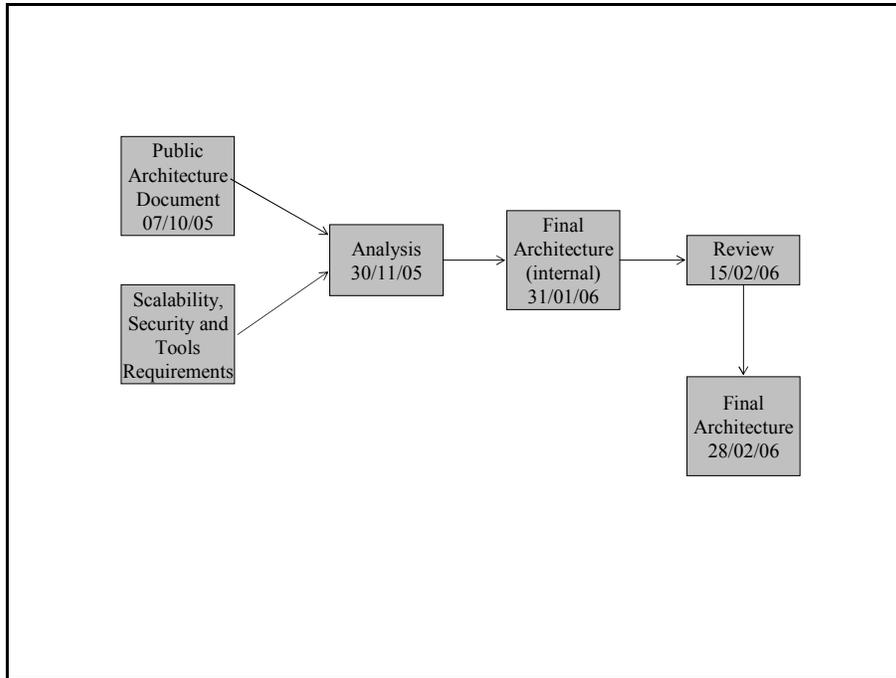
### 3.2.2 **Progress**

We completed the specification of the provenance architecture over the last six month period. The document is being released as deliverable D3.1.1 for this period. Overall, the progress of this document is illustrated in the following figure (presented at the Review meeting): this resulted in a first public release of the architecture document in October.



**Figure 1: Provenance Architecture review process**

This was then followed by a finalisation phase of the architecture, which concluded this period with the release of D3.1.1



**Figure 2: Deliverable D3.1.1 review process**

The Analysis of Scalability, Security and Tools Requirement is an explicit chapter of the architecture document (chapter 9, justification). It relates requirements to design features in the architectures. Some requirements also lead to explicit recommendations for the implementation of specific components, which we also make explicit in this chapter.

Overall, this is a very substantial document, of over 150 pages, which provides us with solid foundations for the standardisation effort in the web services, space.

### 3.2.3 Deviations

A methodology is not a contractual deliverable of EU Provenance. However, with project PASOA, we felt it presented a good opportunity to convert our know-how into an explicit methodology, which we can apply, refine, and test, and ultimately release in the future. We proposed that such a document formalising the methodology would be produced in the following period.

After the review meeting, following feedback from reviewers, we propose to produce a standardisation white paper, which would be used as a basis for standardisation discussions with other third parties. This document is in preparation, with a first draft, intended to be produced by end of February for internal review, and final public release by end of March.

### 3.2.4 Deliverables

List of deliverables, including due date and actual/foreseen submission date

**Table 1 UoS Architecture Deliverables**

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
D3.1.1	Final Architecture	WP3	28/02/2006	28/02/2006	UoS

Copyright © 2006 by the PROVENANCE consortium

*The PROVENANCE project receives research funding from the European Community's Sixth Framework Programme*

### 3.2.5 Milestones

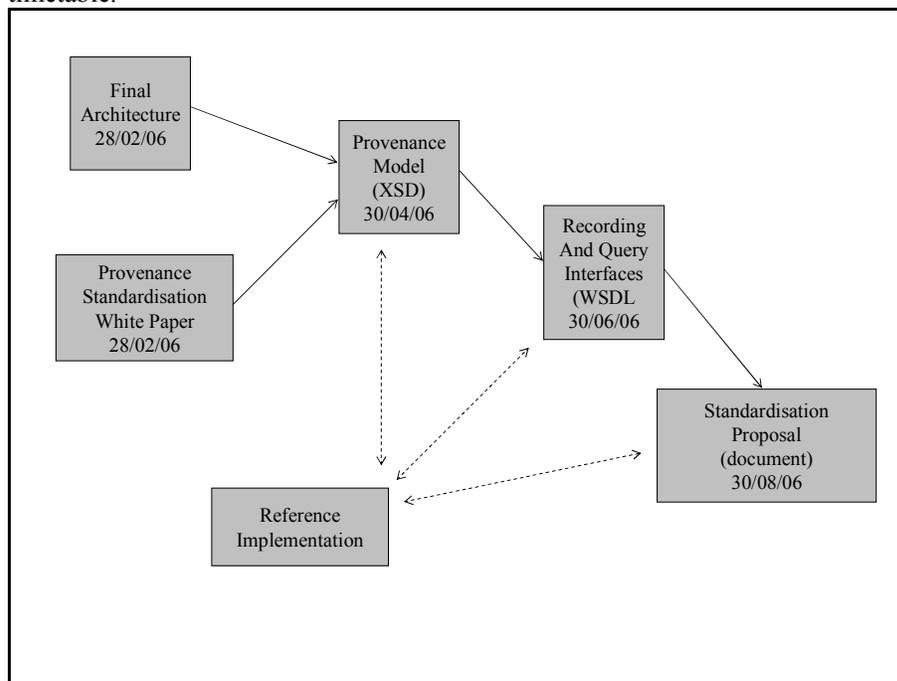
List of milestones, including due date and actual/foreseen achievement date

**Table 2 UoS Architecture Milestones**

Milestone Number	Milestone Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
M3.2.1	Analysis of tools, security, scalability specifications	WP3	30/11/2005	30/11/2005	UoS
M3.3.1	Standardisation white paper	WP3	28/02/2006 (draft)	30/03/2006 (final)	UoS

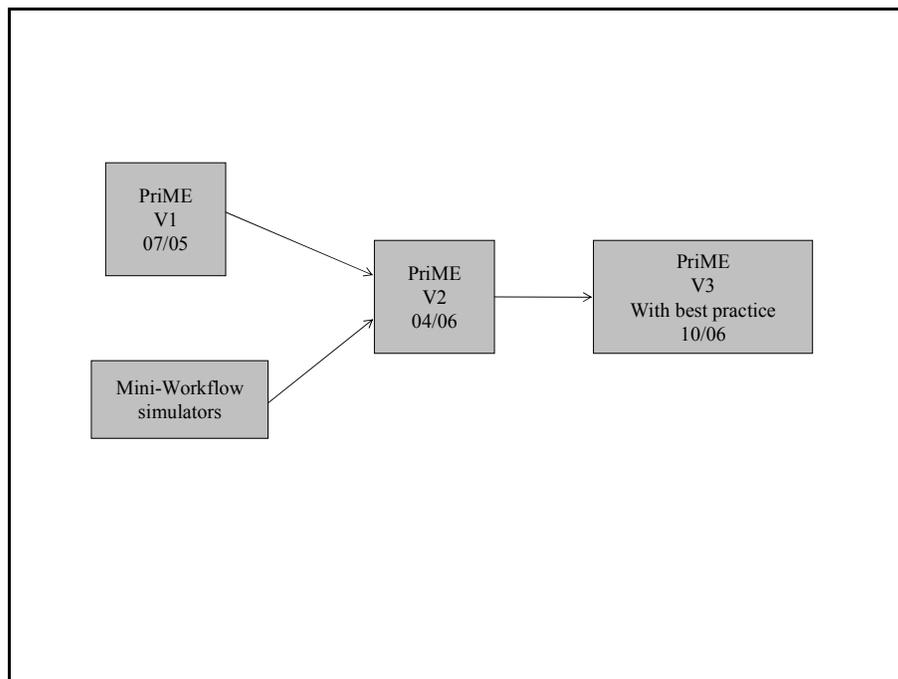
### 3.2.6 Plans for next 6 months

The plan for the next 9 months is to produce a standardisation proposal, based on the following timetable.



**Figure 3: D3.3.1 Standardisation Proposal (end of month 23)**

We will also produce a methodology document identifying a methodology by which developers can make their applications provenance aware:



**Figure 4: D3.4.1 Methodology (end month 23)**

### 3.2.7 Knowledge Dissemination

List of activities to promote and disseminate knowledge from the project.

**Table 3 UoS Exploitable Knowledge and its Use**

Exploitable Knowledge	Exploitable Product(s) or Measure(s)	Sector(s)	Timetable for Commercial Use	Patents or other IPR Protection	Owner and other partners involved
Final Architecture	A reference implementation of the architecture				UoS (owner) all other partners involved
	Web site		International		All partners
	Publications		International		UoS

**List of presentations (all material available from project Web Site):**

**2006**

1. Luc Moreau. **Provenance: Progress Update**. Presentation at CT1 Technical Collaboration Meeting, Heathrow 06, February 2006. [[WWW](#) ]

**2005**

---

Copyright © 2006 by the PROVENANCE consortium

*The PROVENANCE project receives research funding from the European Community's Sixth Framework Programme*

1. Simon Miles. **A Proof of Concept: Provenance in a Service Oriented Architecture**. Paper presentation at All Hands Meeting (AHM'05), Nottingham, UK, July 2005. [[WWW](#)]
2. Luc Moreau. **Provenance-based reasoning in e-Science**. Invited Talk at Automated Reasoning Workshop ARW'05, Edinburgh, UK, July 2005. [[WWW](#)]
3. Luc Moreau. **Provenance: an open approach to experiment validation in e-Science**. Intelligence Agents Multimedia Group, University of Southampton, Group Seminar, Southampton, UK, December 2005. Note: [full mp4](#), [pod cast](#). [[WWW](#)]
4. Luc Moreau. **Provenance: an open approach to experiment validation in e-Science**. University of Birmingham, Departmental Seminar, Birmingham, UK, November 2005. [[WWW](#)]
5. Luc Moreau. **Provenance: concepts, architecture and envisioned tools**. Invited Talk at Grid@Work'05, Sophia-Antipolis, France, October 2005. [[WWW](#)]

**Press Releases (available from project web site):**

1. **1/2006**: ECS News: [Building trust and validation into distributed computer networks](#)
2. **1/2006**: IST Results: [In Grids we trust](#)
3. **12/2005**: ECS News: [Data provenance project endorsed by industry](#)

**Organisation of the IPAW Workshop**

**11/2005**: EU Provenance project co organises the third International Provenance and Annotation Workshop (IPAW'06). Call for papers can be found on the [IPAW'06 page](#).



**International Provenance and Annotation Workshop**

**Publishable Results**

1. Pre-Prototype architecture (published at AHM'05) and submitted to journal (under review)
2. Architecture
3. Methodology (not a formal project deliverable)

**3.3 WP4: Security**

**3.3.1 Objectives**

The previous six month period of the project led to deliverable D4.1.1 containing the first version of the project Security Specification. This specification focused making the logical Provenance architecture secure. It addressed the security of a single Provenance Store together with a set of Provenance Stores in a federated security domain. This period of the project enhanced the specification by including details which were more closely focused on the implementation of the security architecture including the mapping of the architecture to the Globus Toolkit GT4 which is the framework chosen by the project for the reference implementation. A first implementation of a secure Provenance Store was also an objective for this period.

**3.3.2 Progress**

Work towards these objectives was carried out in tasks listed below. The main objectives of the period were achieved on schedule and without deviation from plan. Activities carried out included:

---

Copyright © 2006 by the PROVENANCE consortium

*The PROVENANCE project receives research funding from the European Community's Sixth Framework Programme*

1. Revision of deliverable D4.1.1 to include security components based on the GT4 framework. This led to deliverable D4.2.1, Security Specification version 2
2. Implementation of Provenance Store access control mechanism as part of the final security deliverable D4.3.1 due at the end of month 24.

As a result the deliverables listed in the table below were completed on time.

### 3.3.3 Deviations

There were no deviations from the plan. Future implementation of the security infrastructure will be planned as part of the Implementation, Integration and Test workpackage WP9.

### 3.3.4 Deliverables

Table 4 IBM Security Deliverables

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
D4.2.1	Security Specification version 2	WP4	28/02/2006	28/02/2006	IBM

### 3.3.5 Milestones

Plan milestones were used to monitor the document editing and review processes.

### 3.3.6 Plans for next 6 months

In the next period, the objectives for this workpackage are:

1. Implementation of a federated security infrastructure to meet the requirements of the Organ Transplant Management application (WP8)
2. Testing and review of the security implementation by the Aerospace and OTM application partners (WP7 and WP8).

### 3.3.7 Knowledge Dissemination

No knowledge dissemination activities have taken place for this workpackage in the last 6 months.

## 3.4 WP5: Scalability

### 3.4.1 Objectives

The objectives for this six month period build on the deliverable D5.1.1 Scalability Specification version 1. This document defined a set of recommendations for implementing a scalable provenance architecture. In this period, these recommendations were extended to include an implementation strategy based on the Globus Toolkit GT4 which is the toolkit for implementing the project reference implementation in workpackage 9.

### 3.4.2 Progress

Work towards these objectives was carried out in tasks listed below. The main objectives of the period were achieved on schedule and without deviation from plan. Activities carried out included:

1. Revision of deliverable D5.1.1 to include scalability components based on the GT4 framework. This led to deliverable D5.2.1, Scalability Specification version 2

2. Re-basing the Provenance reference implementation on the Globus Toolkit GT4 to provide support for the Web Services Resource Framework (WS-RF) allowing a scalable implementation to be written.

As a result the deliverables listed in the table below were completed on time.

### 3.4.3 Deviations

There were no deviations from the plan. Future implementation of the scalability infrastructure will be planned as part of the Implementation, Integration and Test workpackage WP9.

### 3.4.4 Deliverables

**Table 5 IBM Scalability Deliverables**

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
D5.2.1	Scalability Specification version 2	WP5	28/02/2006	28/02/2006	IBM

### 3.4.5 Milestones

Plan milestones were used to monitor the document editing and review processes.

### 3.4.6 Plans for next 6 months

In the next period, the objectives for this workpackage are:

1. Implement support for recommendations contained in deliverable D5.2.1 including:
  - a. DAIS patterns for scalable query management
  - b. Distributed Provenance Store support
  - c. Clustering of Provenance Stores
2. Testing and review of the scalability implementation by the Aerospace and OTM application partners (WP7 and WP8).

### 3.4.7 Knowledge Dissemination

No knowledge dissemination activities have taken place for this workpackage in the last 6 months.

## 3.5 *WP6: Tools and Setup*

### 3.5.1 Objectives

The aim of the “Tools and Setup” WP is to produce a software suite containing a collection of independent modules that can support the navigation, accessing and reasoning over provenance data placed in one or more Provenance Stores. Such tools are intended to be “generic”, i.e. Application independent, and would interact with the PS using a management, query and submission interface. Each module may be considered as an independent tool that is accessible by an application or a user, and tools may also make use of services provided by other tools. The “setup” protocol involves choosing suitable Provenance Stores, obtaining access to Provenance Stores, and identifying schemas that can be used for publishing data into a PS, and for querying a PS.

### 3.5.2 Progress

WP6 has been working closely with WP7 and WP8 to elicit requirements for the use of tools (for instance, how rules should be defined for the workflow scenarios being defined in the applications), and the requirements of the setup process. There has also been assessment of the impact of a GT4 implementation – such as the particular security capability that is made available via the GT4 container, and the mechanism used for credential checking using this system.

#### December/January 2006

4. The operations needed to be supported in the setup protocol were specified, based on the logical architecture document (from WP3) and the application scenarios defined in WP7 and WP8.
5. Members of WP6 visited DLR to better understand how tools could be used in the TENT application, and how best to integrate the tool suite with the application scenario defined in TENT.

There continues to be close collaboration between the personnel involved in the Provenance project at Cardiff, and the UK EPSRC-funded “PASOA” project. There is also now a stronger collaboration with the Triana group at Cardiff. This was not possible previously since no Provenance specific tools were available.

### 3.5.3 Deviations

No deviations from the plan and, consequently no corrective action deemed necessary.

### 3.5.4 Deliverables

Work in the period has produced the following documents. Addition input was provided to deliverable D3.1.1.

**Table 6: UWC Tools Deliverables**

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
I.6.3	Issues arising from the installation, deployment and use of the eXo Portal Framework.	WP6	N/A	19/01/2006	UWC
I.6.4	Integrating ProActive with the Provenance tool suite	WP6	N/A	24/01/2006	UWC
D.6.1.1	Update to Deliverable: Tools for Using Provenance	WP6	31/12/2005	16/12/2005	UWC
D.6.1.2	Formal Deliverable: Setup Protocol Implementation	WP6	28/02/2006	28/02/2006	UWC

### 3.5.5 Milestones

### 3.5.6 Plans for next 6 months

Plans for the next period to the end of project are centered on delivery of the main workpackage outcomes (deliverables below) with a number of intermediate milestones to ensure sound progress along the way. Major targets are as follows:

1. Month 18 [Deliverable D.6.2.1]: this is currently being undertaken and on-target for delivery.
2. Month 19 [Deliverable D.6.3.1]: involves specification of configuration utility for tools that have so far been implemented.
3. Month 23: Delivery and demonstration of the final tool suite that can be integrated with the application examples from WP7 and WP8.

To achieve these targets, the following intermediate tasks will be implemented:

1. Updates to the implementation of a tool suite;
2. Usage of the tool suite for particular application scenarios;
3. Updates to the setup protocol,
4. Design of the configuration utility/tool
5. Integration of configuration utility with the tool suite and setup protocol.
6. The setup protocol is currently being implemented in the context of the WP7 TENT application.
7. Implement workflow visualization portlet. As part of this effort, it will also be necessary to better understand how an application specific workflow should be described, and subsequently how it should be reconstructed.
8. Implement example of a relationship plug-in. This will be undertaken in collaboration with WP7.
9. Integrate tools with OTM (WP8) and DLR (WP7) application mini-workflow examples.
10. Investigate any constraints on use of Portal Framework.
11. Design of configuration tool.
12. Integrate Tools with OTM (WP8) and DLR (WP7) application
13. Integrate Tools with Security Framework (for setup protocol) – based on the current implementation from WP4.

### 3.5.7 Knowledge Dissemination

Interaction has taken place with the Triana group at Cardiff and the ProActive group at INRIA for integration between the tool suite and these systems. A two page proposal about how an active object in ProActive can submit and retrieve data from a Provenance Store has been discussed between Arnaud Contes, Omer Rana (Cardiff) and Francoise Baude (INRIA). Funding is currently being sought to take this idea to fruition.

**Table 7: UWC Knowledge Dissemination**

<b>Dissemination of Knowledge</b>					
<b>Planned/Actual Dates</b>	<b>Type</b>	<b>Type of Audience</b>	<b>Countries Addressed</b>	<b>Size of Audience</b>	<b>Partner Responsible or Involved</b>
12/12/05	“Self-Organisation and Multi-Agent Systems”,	Artificial Intelligence researchers	various (International)	120	UWC

Copyright © 2006 by the PROVENANCE consortium

*The PROVENANCE project receives research funding from the European Community's Sixth Framework Programme*

<b>Dissemination of Knowledge</b>					
	University of Paisley, UK				
14/12/05	“Scientific Data Management” workshop, at National eScience Centre, Edinburgh, UK	Grid Computing and Astrophysics community	various (International)	30	UWC
18/1/06	“Virtual Research Environment and Portals” workshop in Portsmouth, UK	Grid Computing and Computational Science	Primarily European	25	UWC

### **3.6 WP7: Application 1 – Aerospace**

#### **3.6.1 Objectives**

The main workpackage objectives during the period were as follows:

1. Integration of client side library into the aerospace application.
2. Input and requirements for WP 6 (tools and setup).
3. Deployment of Provenance store in DLR.

#### **3.6.2 Progress**

Work towards these objectives was carried out in tasks listed below. Both of the main objectives of the period were achieved on schedule and without deviation from plan. It is expected that the results from the first period will continue to be revised in the next period since they are interim results. Activities carried out included:

1. Description of the application mapping to provenance. This includes the methodical and detailed identification of interaction, actor states and relationship p-assertions.
2. Identification of concrete queries.
3. Integration of the PreServ client side library into TENT. This allows the first recording of p-assertions and p-assertion retrieval from the store.
4. Installation/deployment of a PreServ Provenance store on a DLR server.
5. Input for WP6: Discussion and documentation of requirements for tools development. This includes interaction of application and Provenance tools, usage of query and management tools, tools usage for p-recording. Also the possible usage of ProActive for the application has been discussed.

#### **3.6.3 Deviations**

No deviations from the plan and, consequently no corrective action deemed necessary.

### 3.6.4 Deliverables

There were no official deliverables during the last 6 months.

### 3.6.5 Milestones

While no formal milestones were set in the period, the workpackage did carry its own internal project plan.

### 3.6.6 Plans for next 6 months

Plans for the next period to the end of project are centered on delivery of the main workpackage outcomes (deliverables below) with a number of intermediate milestones to ensure sound progress along the way. Major targets are as follows:

1. Month 20 [Deliverable D.7.2.1]: Delivery and demonstration of preliminary TENT demonstration system. Showing limited functionality, however already integrating with provenance components. Deployed on a local site.
2. Month 23 [Deliverable D.7.3.1]: Delivery of the evaluation report for the work-package describing the systems developed, evaluating their use of the provenance architecture / components and other elements of the project.
3. Month 23 [Deliverable D.7.3.2]: Delivery and demonstration of the final TENT/SikMa demonstration system. This system will show the TENT application running in conjunction with deployed provenance components in distributed sites.

The following major milestone is in the following period:

1. Month 17 [Milestone M7.1.1]: Application-specific provenance generation implemented.

### 3.6.7 Knowledge Dissemination

Dissemination of Knowledge					
Planned/Actual Dates	Type	Type of Audience	Countries Addressed	Size of Audience	Partner Responsible or Involved
January 10 <sup>th</sup> , 2006	Talk	Scientists from Cologne/Bonn area	Germany	30	DLR

## 3.7 WP8: Application 2 – Organ Transplant Management

### 3.7.1 Objectives

The main workpackage objectives during the period were as follows:

1. Create a detailed description of the provenance concepts applied to a subset of the workflow in the OTM scenario, to be the basis of intense technical discussions about integration of the OTM/EHCR systems with the Provenance Client API and the Provenance Tools during the Budapest F2F meeting.
2. Design and implement the communication between the OTM and the EHCR systems using the European ENV 13606 pre-standard as a basis.
3. Build a first, reduced version of the OTM / EHCR demonstration system showing integration with provenance components and showing connectivity between the OTM modules and the

EHCR system. The demonstrator will be the basis of intense technical discussions about the next steps on integration of the OTM/EHCR systems with the Provenance Client API and the Provenance Tools during the Cardiff F2F meeting.

**3.7.2 Progress**

Work towards these objectives was carried out in tasks listed below. All the main objectives of the period were achieved within the period. It is expected that the results from this period will continue to be revised in the next period since they are interim results.

1. Detailed description of the provenance concepts applied to a subset of the workflow in the OTM scenario. This mini-workflow scenario was presented in the Budapest F2F meeting and was used as a basis for technical discussion with the other partners.
2. The Budapest F2F meeting, held in October 2005, was an important technical milestone at the middle of the reporting period. Technical discussions with other partners helped to increase common understanding and to decide next steps for the OTM and the EHCR applications, and its integration with the Provenance components.
3. Analysis of the European ENV 13606 pre-standard. Adaptation of the Catalan Healthcare Record to the pre-standard.
4. Development of the communication protocol between the OTM modules and the EHCR system.
5. Implementation of preliminary OTM / EHCR demonstration system, implementing a sub-set of the workflow, which was presented in the Budapest F2F meeting. The demonstrator shows limited functionality; however it already integrates with provenance components and shows connectivity between the OTM modules and the EHCR system.

An internal technical meeting, between SZTAKI and UPC, was held in Barcelona from January 16 to 19, 2006, in order to fully connect the EHCR and OTM systems, and to plan the work from that moment until the Cardiff F2F meeting in Month 18.

**3.7.3 Deviations**

No deviations from the plan and, consequently no corrective action deemed necessary.

**3.7.4 Deliverables**

There were no official deliverables during the last 6 months.

**3.7.5 Milestones**

While no formal milestones were set in the period, the workpackage did manage its own internal project plan. Significant steps in this were as shown in the following table.

**Table 8 UPC Application Milestones**

Milestone Number	Milestone Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
N/A	Demonstration specification document, specifying provenance issues in a mini-workflow	WP8	30/11/2005	30/10/2005	UPC

Milestone Number	Milestone Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
	scenario.				
N/A	Adaptation of the Catalan EHCR to the ENV 13606 pre-standard	WP8	20/12/2005	20/12/2005	SZTAKI
N/A	Connection between OTM and EHCR system. Connection between OTM and provenance stores	WP8	25/01/2006	05/01/2006	UPC
N/A	Revision of demonstration specification and D8.1.1 contents to ensure it is in line with final application demonstration scenarios.	WP8	31/01/2006	20/02/2006	UPC

### 3.7.6 Plans for next 6 months

Plans for the next six month period are centered around the implementation of the OTM/EHCR demonstrator with a number of intermediate milestones to ensure sound progress along the way. Major targets are as follows:

- Month 18 [Milestone M8.1.1] Work towards D.8.2.2.: OTM / EHCR demonstration system.(preliminary)
- Month 18: Cardiff F2F meeting, to be held at the end of month 18, will be an important milestone to prepare the work for the next period and to coordinate work with other partners.
- Month 22: Köln F2F meeting, to be held at the beginning of month 22, will be an important milestone to coordinate the work with other partners.
- Month 23 [First draft version Deliverable D.8.2.2]: Delivery and demonstration of a first draft of the OTM/EHCR demonstration system. This system will show the OTM / EHCR application running in conjunction with deployed provenance components. Delivery of the final version is planned by the end of month 25.

### 3.7.7 Knowledge Dissemination

**Table 9 UPC Dissemination of Knowledge**

Planned/Actual Dates	Type	Type of Audience	Countries Addressed	Size of Audience	Partner Responsible or Involved
31/01/2006	Submission of technical paper to the First International Workshop on Privacy and Security in Agent-based Collaborative Environments (PSACE2006),	Experts on Privacy issues in distributed applications	International	Potentially 30 persons (to be held in May 2006)	SZTAKI, UPC

### 3.8 WP9: Implementation, Integration and Test

#### 3.8.1 Objectives

The main workpackage objectives during the period were as follows:

1. To develop a scalable, secure reference implementation of the Provenance architecture for release outside the consortium by the end of month 18
2. To provide design and installation documentation to assist the infrastructure deployment
3. To base the reference implementation on the Globus GT4 toolkit

#### 3.8.2 Progress

Work towards these objectives was carried out in tasks listed below

1. Responsibility for the software components in WP9 were divided between IBM and UoS with IBM being responsible for the server side and UoS for the Client Application Programming Interface (CAPI)
2. The reference implementation was re-based on the GT4 toolkit to support the use of the Web Services Resource Framework (WS-RF) and Grid Security. This brings the implementation into line with the most widely used Grid infrastructure.
3. To allow WP7 and WP8 partners' early access to a Provenance Store implementation, the PreServ code developed by UoS was distributed. The CAPI being developed will be common to both PreServ and this project's reference implementation.
4. The WSDL and XML Schemas defined by UoS as part of WP3 were periodically released as a result of design changes. The reference implementation was updated to reflect these changes.
5. The Provenance Store access control developed as part of WP4 was merged into the reference implementation.
6. A full design document for the deliverable is being prepared

#### 3.8.3 Deviations

No deviations from the plan and, consequently no corrective action deemed necessary.

### 3.8.4 Deliverables

**Table 10 IBM Implementation Deliverables**

<b>Deliverable Number</b>	<b>Deliverable Name</b>	<b>Workpackage Number</b>	<b>Date Due</b>	<b>Actual/Forecast Delivery Date</b>	<b>Lead Contractor</b>
D9.3.2	Functional Prototype (Early release)	WP9	28/02/2006	28/02/2006	IBM

### 3.8.5 Milestones

While no formal milestones were set in the period, the workpackage did manage its own internal project plan.

### 3.8.6 Plans for next 6 months

In the next period, WP9 plans to:

1. Concentrate on adding functionality to the first release of the Functional Prototype. This will include items from the Security and Scalability workpackages. This will support the requirements of the application workpackages WP7 and WP8.
2. An interim code release containing this additional functionality is planned at the end of the first three month period.
3. This will be followed by the final code release at the end of August 2006.

### 3.8.7 Knowledge Dissemination

No knowledge dissemination activities have taken place for this workpackage in the last 12 months.

## 3.9 *WP10: Collaboration*

### 3.9.1 Task 1: Exploitation of synergies / technical concertation

The Provenance consortium was represented by Luc Moreau at the TC steering committee meeting at Heathrow in Feb 2006. Presentation is available from <http://twiki.gridprovenance.org/pub/Provenance/ProjectPublications/heathrow06.ppt>. The presentation resulted in some new interaction with the DataMiningGrid, and the possibility of DataMiningGrid making use of the provenance software. There has been no further progress on the collaboration with Simdat, but the agenda for a future meeting is being drafted.

### 3.9.2 Task 2: Joint fora for exchange and dissemination

We had planned to launch a workshop on Provenance and this has taken place. A review of our work is as follows:

- 17/11/2005:
  - Venue availability confirmed
  - Finalisation of the Call for Papers
    - [cfp.txt](#): IPAW'06 Call for Papers (draft)
  - Program Committee members invitation
- 21/11/2005
  - First call sent
  - Mailing Lists used to announce Call for Papers

- [sem-grid@gridforum.org](mailto:sem-grid@gridforum.org), [agents@cs.umbc.edu](mailto:agents@cs.umbc.edu), [ogsa-wg@gridforum.org](mailto:ogsa-wg@gridforum.org), [semantic-web@w3.org](mailto:semantic-web@w3.org)
  - [service-orientated-architecture@yahoogroups.com](mailto:service-orientated-architecture@yahoogroups.com)
  - [ict\\_governance@yahoogroups.com](mailto:ict_governance@yahoogroups.com)
  - [web-services@yahoogroups.com](mailto:web-services@yahoogroups.com)
  - [enterprise-information-integration@yahoogroups.com](mailto:enterprise-information-integration@yahoogroups.com)
  - [business-process-management@yahoogroups.com](mailto:business-process-management@yahoogroups.com)
  - [e-business@jiscmail.ac.uk](mailto:e-business@jiscmail.ac.uk)
  - myGrid and ontoGrid
  - ACM calendar of events
  - Notification to Harvard e-Science team
  - [agoodman@cfa.harvard.edu](mailto:agoodman@cfa.harvard.edu)
  - PASS team at Harvard
  - Core Grid mailing list and web site
- 29/11/2005
    - Submission of workshop proposal to LNCS
  - 30/11/2005
    - Extension of Program Committee
  - 23/01/2006
    - Workshop website [www.ipaw.info](http://www.ipaw.info) went live
  - 01/2006
    - Approval by Springer Incs for post-proceedings
  - 23/01/2006
    - Final Call for Papers
  - 27/01/2006
    - Confirmation of hotel and conference venue
  - 27/01/2006
    - Launch of new web site
  - 30/01/2006
    - Setting up of web submission system
  - 30/01/2006
    - Direct notifications to selected researchers
  - 06/02/2006
    - Deadline extension to 19/02/2006
  - 06/02/2006
    - Discussion with Microsoft of the nature of their sponsorship

### **3.9.3 Task 3: Co-ordination of standardisation efforts**

No activity

### **3.9.4 Task 4: Repository of reference implementations and Grid middleware**

No activity

### **3.9.5 Task 5: Collaboration on research inventory and roadmaps**

No activity

### **3.9.6 Task 6: Indicators and impact assessment**

No activity

### **3.9.7 Task 7: Training activities**

- 06/12/2005
  - First tutorial session on Provenance given over the Access Grid. While the session was initially aimed at the GriPhyN team in Chicago, it was extended to OntoGrid.
- 02/2006
  - Planning for a half a day tutorial on Provenance, aimed in priority to OntoGrid, but to be extended to other projects.

## 4 Consortium Management

This section describes the management of the consortium during the reporting period. It identifies the major activities during the third 6 month period followed by issues identified during the reporting period and their resolution.

### 4.1 Management Activities during Months 13-18

The following management activities were completed during this reporting period:

1. The project completed the delivery of deliverables due at month 12 of the project including financial and management information.
2. Two face to face meetings were held during this reporting period hosted by partners in Budapest and Cardiff. The management issues discussed at these meetings were:
  - a. Face to Face meeting held at SZTAKI Budapest during October 26-28 2005
    - i. Preparation for project annual review
      1. Finalisation of agenda and content of presentations
      2. Organisation of logistics and preparatory meeting at DLR prior to review
      3. Discussion of aims and objectives of review
    - ii. Discussion and decision to request project extension for three months
      1. Project Officer to be notified prior to review
    - iii. Planning for tasks for remaining period until month 18
    - iv. Planning for remainder of project
      1. Timetable for application scenario demonstrations
  - b. Face to Face meeting held at UWC Cardiff during February 22-24 2006
    - i. Details to be added following meeting
3. The University of Southampton advised the coordinator that a major fire occurred at the University on October 30<sup>th</sup> 2005. The fire caused disruption to the department working on the Provenance project. Initially, the University informed the project that a force majeure event had taken place that affected the work on the project. This information was received in a letter dated 11/11/2005. Fortunately, the disruption was not significant to the project and the University informed the coordinator by letter on 05/01/2006 to that effect.
4. The following activities are discussed in more detail in the following sections:
  - a. The first annual project review was held in Brussels on November 14<sup>th</sup> 2005
  - b. A plan for the remainder of the project was developed and provided to the Project Officer
  - c. Revised financial information was submitted to the Project Officer
  - d. The project coordinator generated a clarification to the Consortium Agreement in preparation for the release of the reference implementation using CPL

### 4.2 Management Issues

Throughout the reporting period, the following management issues were highlighted and solutions provided.

#### 4.2.1 Provenance Annual Review

The first Provenance review took place in Brussels on November 14<sup>th</sup> 2005. The Project Officer in his preliminary feedback stated that:

*The project has developed good work, has a good consortium and is addressing a very interesting topic with a lot of potential. This generates high expectations on this project and the review team would like to see excellent results at the end of the project including the dissemination and use of the results.*

Actions highlighted by the review panel and addressed by the Provenance consortium were:

1. A revised Management Report was provided that corrected errors found by the reviewers
2. The Tools workpackage deliverable D6.1.1 was resubmitted to include:
  - a. Sections on conclusions and further work
  - b. A definition of expected tools users
3. The Exploitation Plan was revised with actions included that could be tracked by the project
4. A draft standardisation plan was prepared that formed part of the revised Exploitation Plan
5. The project plans described in the next section were submitted
6. Implementation issues highlighted will be addressed in the next deliverable

The material was provided before the end of 2005 as requested by the review panel.

## **4.2.2 Relationship between Grid Provenance and PASOA**

At the annual project review, the reviewers requested the clarification of the relationship between the EU Provenance project and the UK Provenance Aware Service Oriented Architecture (PASOA) project [www.pasoa.org](http://www.pasoa.org). The following statements clarify this relationship:

1. PASOA is an academic project that is investigating the broad area of provenance; particularly in the e-Science disciplines.
2. The EU Provenance project is focusing on taking provenance and bringing it into the application and enterprise space based on the academic work of PASOA. Specifically, EU Provenance focuses on security and scalability of the design and implementation, which provide an industrial-strength solution. This is beyond the scope of PASOA.
3. PreServ is an implementation of the PASOA work and the EU Provenance project has permission to use the implementation as an early release of a Provenance Store for use by project partners. The EU Provenance Project implementation is based on the Globus Toolkit GT4 and is WS-ResourceFramework (WS-RF) compliant. The PASOA PreServ implementation is not.
4. By developing a common client side programming library, users can choose between different implementations of the Provenance Store interfaces and from a standardisation point of view it will test the robustness of the interface specifications since both implementations will support a common set of interfaces. The EU Provenance project implementation will support additional interfaces that take advantage of WS-RF.

### 4.2.3 Plans for Next Period

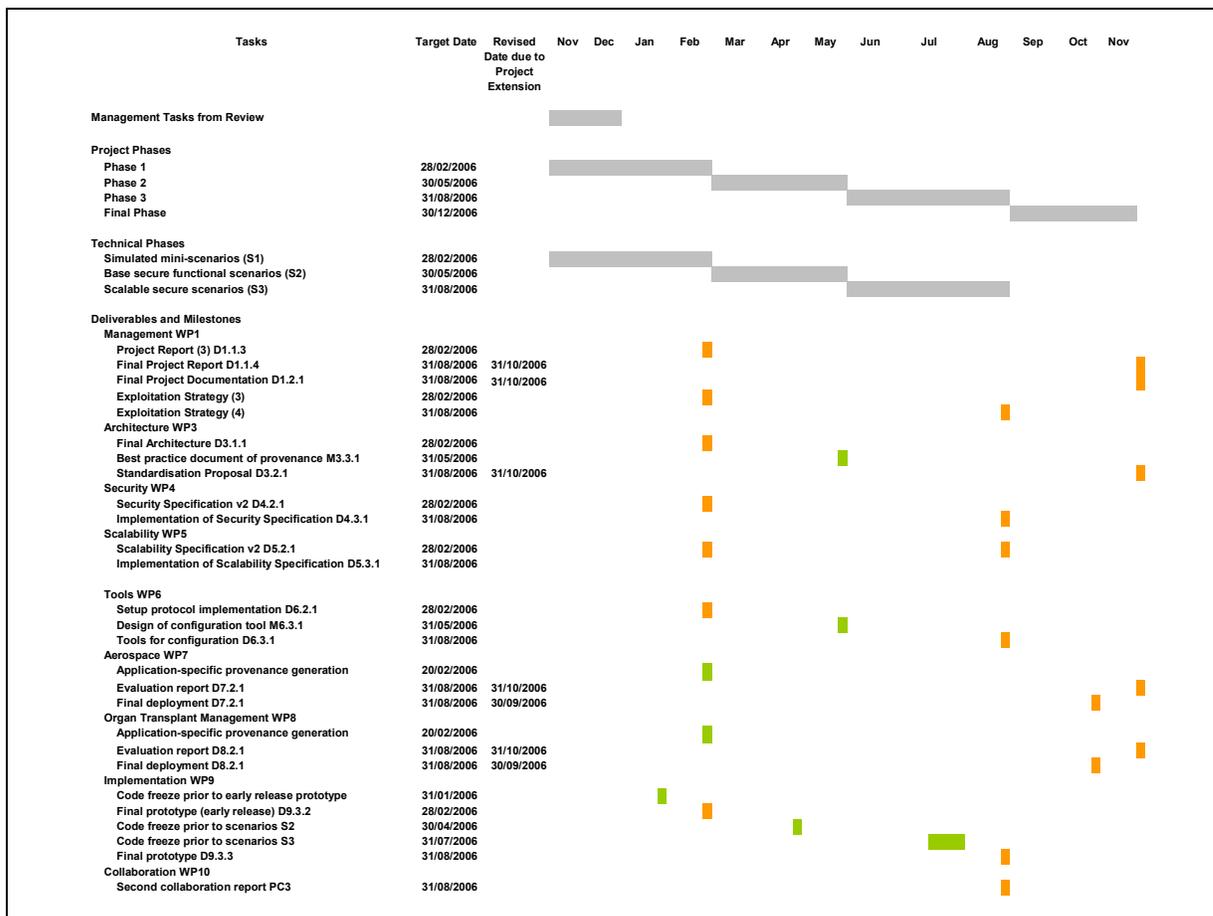


Figure 5: Provenance Project Plan

Following the annual review, the Provenance consortium submitted an outline project plan for the remaining period of the project. This is illustrated above. Key points of the plan are:

1. The project is divided into a set of three moth phases including the requested three month extension period.
2. Following the contracted deliveries at the end of February 2006, there will be two technical phases resulting in subsets of the final demonstration scenarios which will be available at the end of August 2006 with an interim set May 2006.
3. Further external releases of the reference implementation will be provided. An interim release is expected in May 2006 with the final version August 2006.
4. For WP7 and WP8, the final deployment and reports will be provided at the end of September and October 2006 respectively. This allows deployment to take place following the final code release at the end of August 2006.
5. Final management reports will be provided at the end of October 2006 with the expectation that the final project review will take place before the end of 2006.

#### **4.2.4 Revised Financial Reporting**

A set of financial reports were provided to the Project Officer at the end of month 12 as required. Following review by the Commission, it was discovered that the total claim was less than 75% of the previous pre-financing. This meant that the Commission could not make the next pre-financing payment. Following consultation with the Project Officer, the coordinator submitted another claim which met the payment threshold requirements.

#### **4.2.5 Clarification of Agreement to use CPL for Code Release**

The coordinator's legal department has a formal approvals process for the release of any source code under an open source license. The Provenance consortium had agreed in a minuted meeting to use the Common Public License version 1. To confirm this, the coordinator issued an agreement for all partners to sign. A copy of this agreement is provided in [Appendix 1](#). All partners signed and returned the agreement allowing the coordinator's internal approvals to be completed.

## **Appendix 1 - CPL Agreement**

### **Clarification relating to the release of EU Provenance v.1 on Open Source license terms**

Section 6.6, 'Knowledge Management and Intellectual Property Provisions', of Annex I to the EU Contract dated 19th May 2004 provides that the Provenance Consortium is committed to make interfaces available, and to release a reference implementation, on open source terms.

It is understood that, at a meeting on 6th April 2005, you agreed on behalf of your organisation that the Common Public License v.1 ('CPL') should be the relevant open source license that should apply to any release of such items. The CPL is viewable at URL:

<http://www.opensource.org/licenses/cpl1.0.php> .

Please would you now confirm your organisation's agreement (by signing below and returning a copy of this Clarification) to the releasing of 'EU Provenance v.1' under the provisions of the CPL license.

Thank you.

(Name)  
Title  
date

AGREED FOR AND ON BEHALF OF  
*IBM United Kingdom Limited*

---

Name:  
Title:  
Date:

