Secure Access to Service-based Collaborative Workflow Across Organisations

Duncan John Russell

Submitted in accordance with the requirements
for the degree of Doctor of Philosophy

School of Computing

Faculty of Engineering

The University of Leeds

September 2006

Abstract

This thesis addresses the problem of providing secure collaborative workflow across different organisations with an architectural solution. This involves addressing such issues as collaborative workflow and business processing, Grid Computing, service-oriented architecture and role-based access control.

Grid Computing has been developed to provide middleware for collaborative access to distributed processing services and distributed data sources, supporting distributed users that form Virtual Organisations. Some distributed services and data are commercially sensitive, and need to be protected by controlling access to them, ensuring access is only for permitted users in the collaborative team. The collaboration is controlled in a workflow management system and links role-based workflow with role-based access control. Workflow management in Grid Computing provides the capability to integrate and coordinate distributed users, stateful Grid Services, Information systems and Grid Compute Resources.

The research is supported by the UK e-Science Project, DAME (Distributed Aircraft Maintenance Environment), a collaborative project that demonstrates the use Grid Computing for collaborative problem solving across organisations. DAME uses the domain of aircraft engine diagnostics and maintenance in a global context, requiring the support of workflow management to coordinate the sharing of globally distributed users, processing services and data.

This research extends the understanding of access control to Grid Services, by producing an architecture for the definition and control of dynamic access control policies for collaborative service-based workflows. In particular, the research addresses collaborative access to stateful Grid Service instances across organisations.

The proposed solution for secure collaborative service-based workflows is called "Workflow-Team Policy Architecture". An implemented web-based portal and workflow management system controlling Grid Services instances across the White Rose Grid is evaluated using the business example of aircraft engine diagnostics.

Acknowledgements

I would like to thank my supervisors Professor Peter Dew and Dr Karim Djemame for all their assistance and guidance during my research and writing.

I would like to thank the members of the DAME project at the Universities of Leeds, Oxford, Sheffield and York and industrial partners Rolls-Royce and Data Systems & Solutions for all their involvement in the demonstrations and their feedback during meetings. In particular, thanks go to Georges Honoré and Martyn Fletcher for their work on the DAME integration and requirements modelling. Also, thanks go to Charlie Dibsdale and Graham Hesketh for their expert opinions in evaluating the demonstration environment.

I would like to thank my fellow researchers at University of Leeds for their support with the demonstration platform and the White Rose Grid for provision of the Grid Computing platform.

My thanks also go to the funding support from the DAME project under UK Engineering and Physical Sciences Research Council Grant GR/R67668/01.

Finally, I would like to thank my family for their support, especially my wife Sarah for putting up with me.

Declarations

Some parts of the work presented in this thesis have been previously published in the following:

Russell, D., Dew, P. M. and Djemame, K. 2004. Self-Securing Dynamic Virtual Organisations. In: (Eds. Chivers, H. and Martin, A.) *Workshop on Grid Security Practice and Experience*, 8-9 July 2004, Oxford, UK. University of York. pp. II-1 - II-6.

Russell, D., Dew, P. M. and Djemame, K. 2004. Access Control for Dynamic Virtual Organisations. In: (Ed. Cox, S. J.) *UK e-Science All Hands Meeting 2004*, Nottingham, UK. EPSRC. pp. 332-339.

Russell, D., Dew, P. M. and Djemame, K. 2004. DAME Collaborative Workflow & Access Control. In: *DAME Open Day*, 19/03/2004, Leeds, UK. DAME.

Russell, D., Dew, P. M. and Djemame, K. 2005. Service-Based Collaborative Workflow for DAME. In: *Services Computing, 2005. (SCC 2005). Proceedings. 2005 IEEE International Conference on*, 10-15 July 2005, Orlando, Florida. pp. 139-146.

Contents

List of Figures

List of Tables

Glossary

The following glossary is an index of commonly used terms within this work:

Business Process is the sequences of tasks that are initiated by a business event with the intent to achieve a business goal.

Business Requirements are the functional and non-functional requirements to achieve a business goal.

Engine Performance Data is the recorded data from the on-wing system that monitors engine performance parameters, such as fuel consumption and throttle position.

Engine Vibration Data is the recorded data from the on-wing system that monitors the vibration of the aircraft engine, the frequencies of which shows the speeds of the propeller shafts and their harmonics.

Feature is defined in the DAME project to characterise a signal trace found in the engine vibration data.

On-demand is the notion by which Grid Computing resources can be dynamically allocated at the latest time to execute a task.

Grid Service Instance is a stateful service. An example implementation would be Grid Services in Globus Toolkit 3 (GT3, (*The Globus Alliance*, 2005)) that executes a long-term process.

Grid Service Type is the definition of a class of service. In Web Services, this could be the name of the interface defined in WSDL, without being bound to a specific URI (Uniform Resource Identifier), allowing for a loosely coupled implementation.

Virtual Organisation is defined by Foster, et al. in (Foster and Kesselman, 2004). In this thesis, a Virtual Organisation (VO) is defined as a semi-permanent collaboration by different organisations or departments from different administrative bodies to achieve a business goal.

Workflow Definition is the flow of tasks to achieve the business goal. It may be described diagrammatically (UML activity diagram (Aalst et al., 2003)), or in high-level model terms

(BPML (BPMI, 2002), Petri-nets (Aalst and Hee, 2004)) or in an execution language such as BPEL (Andrews et al., 2003).

Workflow Instance is the active workflow executing the description from the *Workflow Definition*. It contains the state of the workflow.

Workflow is the automated execution of a process on computers (or information technology), usually as part of a business process.

Workflow-Team Policy Definition is the access control policy for the *Workflow Definition*. It describes the actions a *Role* can perform on *Service Types* and is formed as a template policy that creates the *Workflow-Team Policy Instance*.

Workflow-Team Policy Instance is the access control policy for the *Workflow Instance*. It executes the policy described in the *Workflow-Team Policy Definition*.

Chapter 1

Introduction

This research is aimed at the need for organisations to respond to globalisation. The Internet has allowed new approaches to collaboration between people and organisations. It is easier for geographically distributed people to collaborate on tasks by sharing the same applications and data resources with the support of distributed computing. The open access of the Internet promotes computer-supported collaborations and processes across organisational borders. However, exposing computing applications to the Internet requires security restrictions to protect the applications being accessed by malicious and unauthorised users, inside or outside the organisation.

This research investigates the problem of secure collaborative use of computer resources in business processing across organisations.

1.1 Research Context

The research presented in this thesis investigates the use of distributed computing and business processing with access control. There is an increasing use of computers to support business processing and business process management (Henderson, 2000, Henderson, 2002, Sayal et al., 2002, Smith and Fingar, 2003, Stanoevska-Slabeva et al., 2001). Business process management is concerned with the definition and enactment of collaborative processes to achieve business objectives. Automation of business processes is achieved through workflow management systems (Aalst and Hee, 2004).

Service oriented architecture makes collaborations possible by dividing the functions within an organisation to form services. A service is a function that is well defined, self-contained and does not depend on the context or state of other services. The service description is an abstraction of the operations provided by the service. The integration of appropriate services by means of executing workflows allows business applications to be built, integrating operations from different organisations. The integration of services by using workflows to meet business requirements is illustrated in Figure 1.1.

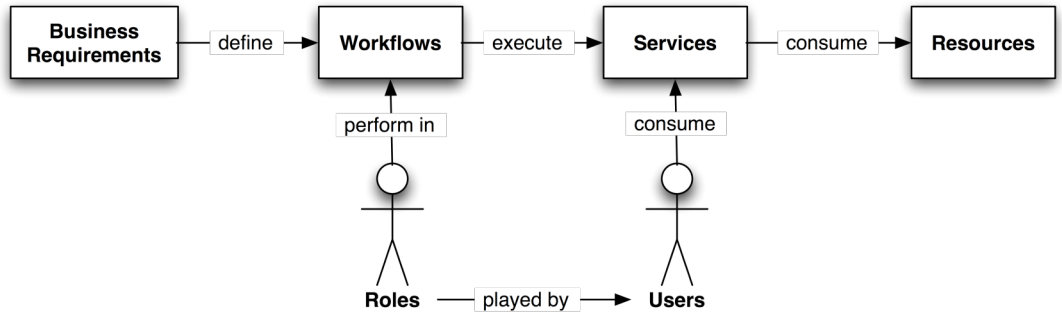

Figure 1.1 Business requirements define workflows and the use of services and resources by users enacting roles

The diagram in Figure 1.1 starts with the business requirements of an organisation, which are used to define the workflows that automate business processing. The workflows call services, which contribute to the tasks in a workflow. The resources consumed by services are required implemented platforms, such as compute processors, databases and data storage. The users consume the services by playing the roles in the workflow.

The creation of a collaborative environment means people can share knowledge while workflow management provides a controlled platform for sharing services and data from different organisations in a consistent manner, reducing errors and tracking progress. Collaborative workflows are executed to support the automation of people sharing a process and the services and data used.

A new form of business has emerged where different organisations can collaborate on achieving business goals by using distributed services, gaining benefit by combining resources, services and data. These distributed services are integrated in collaborative workflows that cross organisational boundaries, where the people, services and data come from different organisations with the support of computing systems that create virtual collaborations, forming Virtual Organisations (VO). To achieve cross-organisational workflow, internal services and data are exposed to external access. At the same time this increases the requirement to protect access to commercially sensitive information and services.

Service oriented computing, an information technology implementation of service-oriented architecture, uses a distributed middleware that supports common standards of inter-operation. Service oriented computing achieves loose coupling between services by using abstract descriptions of services. By abstracting the description of a service from its implementation, integration can be achieved using the service descriptions making is independent of the service implementation. This enables services to be platform and language independent, composing applications by service description, allowing services to

be interchangeable, evolvable and inter-organisational. For business processing, workflows can be composed from service descriptions creating dynamic aggregation of services in workflow engines.

Grid Computing uses abstract descriptions of computing resources, facilitating the virtualisation of compute processing, storage and networking in a loosely coupled manner. Similar to service oriented architecture, compute resources can be discovered and used to fulfil dynamic requirements to processing. Grid Computing introduces the notion of service instances, created when a service implementation requires long term stateful processing. This could be created during a workflow process for collaborative use.

Of significance is the business model of supplying services to create workflows for business processing. Grid Computing enhances the business model by creating service instances that can be consumed in the context of a workflow instance. Service instances link state to a service interface, creating a stateful service. This research investigates the specific area of collaborative business processing when users are from different organisations and services are supplied by different organisations, with the impact of restricting user access to service instances, where groups of individual users requires collaborative access to a service instance. The solution must be aligned with business objectives for protecting competition and administration across many users, consuming many service instances, across organisational boundaries.

For global operations to respond to changes in demand, 24 hours a day, computing resources must be dynamically available. Grid Computing aims to make computing resources (processing, storage and networking) available by virtualisation. Access to resources is made transparent of location, administration domain, architecture and capability. Virtualisation is achieved with a decentralised middleware, providing a standardised interface to the compute resources, allowing for interchange of resources in the same operation. The cross-organisational integration of users, computing services and compute resources to meet business objectives is called a VO. The VO is been made possible by the virtualisation of resources from Grid Computing and by virtualising the functions within an organisation.

The collaboration of people, processing and data requires access management to ensure commercial operating conditions for all parties are protected for unauthorised entry. Current workflow technologies have little support for security requirements, however security work on access control does approach collaboration and workflow (Bussler and Jablonski, 1995, Chandramouli, 2001). Work in the areas of workflow and access control does not currently address how to control collaborative access to stateful services across organisations. This research uses business requirements to capture access requirements to

assets created and shared in collaborative service-based workflows, such that business level access control can specify fine-grained access control to temporary service instances created in distributed systems across organisations. The current grid technologies investigated in this research (Globus Toolkit version 3 and version 4) do not provide dynamic authorisation mechanisms for collaborative access to service instances. With a combination of current technologies in Service Oriented Architecture, Grid Computing, business processing and workflow, and access control there still exists a gap in capabilities to control access to service instances for collaborative use across organisational boundaries.

1.2 Motivation

DAME (Distributed Aircraft Maintenance Environment) was a £3m project funded by EPSRC as part of the UK e-Science program, running from December 2002 to January 2006. It involved the universities of Leeds, Oxford, Sheffield and York and the industrial partners Rolls-Royce, subsidiary Data Systems and Solutions (DS&S) and Cybula Ltd. The investigation aims of the project included the use of Grid Computing in a commercial setting across organisations. The business context for DAME is to support the leasing of aircraft engine to airline companies.

Leasing aircraft engines is a business model used by Rolls-Royce, along with a subsidiary Data Systems & Solutions, to provide the through-life leasing support service. The aircraft engine diagnostics support is to be enhanced with an on-wing engine data recorder, storing vibration and performance data during flight. The data is downloaded and processed by available Grid Computing services and resources. Grid Computing was chosen due to the amount of engine data and processing required. The processed results are available to collaborative teams made up of people at the operating airline, DS&S and Rolls-Royce, who share data and services from different organisations. The collaborative diagnostics team only exists with the support of distributed computing and forms a VO. The services, data and results can be commercially sensitive, especially to competing airlines. Therefore securing the services, data and collaborative processes is essential for the distributed diagnostics environment to be a credible business opportunity.

The DAME environment requires workflow management to maintain accuracy of work and timeliness of diagnoses, and control the processing across distributed services. People in the collaborative team who access the executing workflows in DAME environment require access from global locations. To support this DAME uses a web-based portal into which the VO members login to collaborate securely. This research uses the DAME business requirements for workflow and security on collaboration of aircraft engine diagnostics to investigate securing Grid Service instances in collaborative workflows across organisations.

1.3 Research Objectives

The aim of this research is to investigate the application of workflow, access control and Grid Computing for sharing service instances across organisational boundaries. The following are the research objectives:

- Design and build an experiment to study a workflow management system to execute and control long-running Grid Services across a multi-organisational grid, allowing shared access by users from different organisations.
- Investigate issues of access control in a virtual organisation for collaborative access to services and data, where some services and data are commercially sensitive and could reveal proprietary information to competing members within the virtual organisation.
- Derive a general model for the provision of dynamic fine-grained access control to stateful Grid Service instances used in collaborative teams of users, services and data from different organisations.
- Evaluate the access control for secure service-based collaborative workflows using the business example from the UK e-Science DAME project.

1.4 Major Contributions

The major contributions of this work include:

C1. The Workflow-Team Policy Architecture, an architectural conceptual model for collaborative access to grid services across organisations, where users, services and compute resources form a Virtual Organisation.

C2. The definition of static and dynamic Workflow-Team Policy Architecture components to support C1. The static components define the workflow and policy and are created for simplified administration by linking a role-based process and role-based access control. The dynamic components contain the state of a collaborative team whose members include the users and service instances. The dynamic policy component is linked to the dynamic workflow component in controlling access to the service instances in the context of each workflow.

C3. A general model of secure collaborative workflow to support C2. The model, derived from a business case study, shows that business requirements can be used to generate the static definitions in C2, which in turn create and control the dynamic definitions in the Workflow-Team Policy Architecture.

C4. Implementation of a Workflow Management System and Web-based Portal to illustrate and evaluate the Workflow-Team Policy Architecture. This demonstrates

collaborative access to grid services, by users and service providers across organisations.

C5. Presentation of the DAME business case, detailing the diagnosis process in a Virtual Organisation. The diagnosis process is a collaborative workflow requiring users from different organisations to combine expertise and share access to grid services in the domain of aircraft engine maintenance.

C6. Analysis of the DAME diagnosis process detailing the business requirements for secure access to the grid services consumed during collaboration. This security analysis includes requirements for controlling use and exposure of competitor's services and data. The analysis also includes protection of access to valuable process information, such as preventing discovery of workflow process definitions.

C7. Evaluation of Workflow-Team Policy Architecture and the DAME demonstrator, from C1-4, using semi-structured interviews of industrial experts and analysis against similar and complementary solutions.

1.5 Research Methodology

The research methodology is a cyclic approach to defining the research problem and testing and evaluating a software implementation of the secure collaborative workflow architecture implementation. The research cycle, Figure 1.2, is started with *collect information*, collecting information about the context of the problem and its areas of concern. The collecting activity can include requirements gathered from the motivational problem to ascertain new issues to address. On the technical side, a literature review provides information on the background topics, which can be augmented by attending conferences and reading forums and newsgroups to obtain opinions, and gaining practical experience by experimenting with technologies in the topic area.

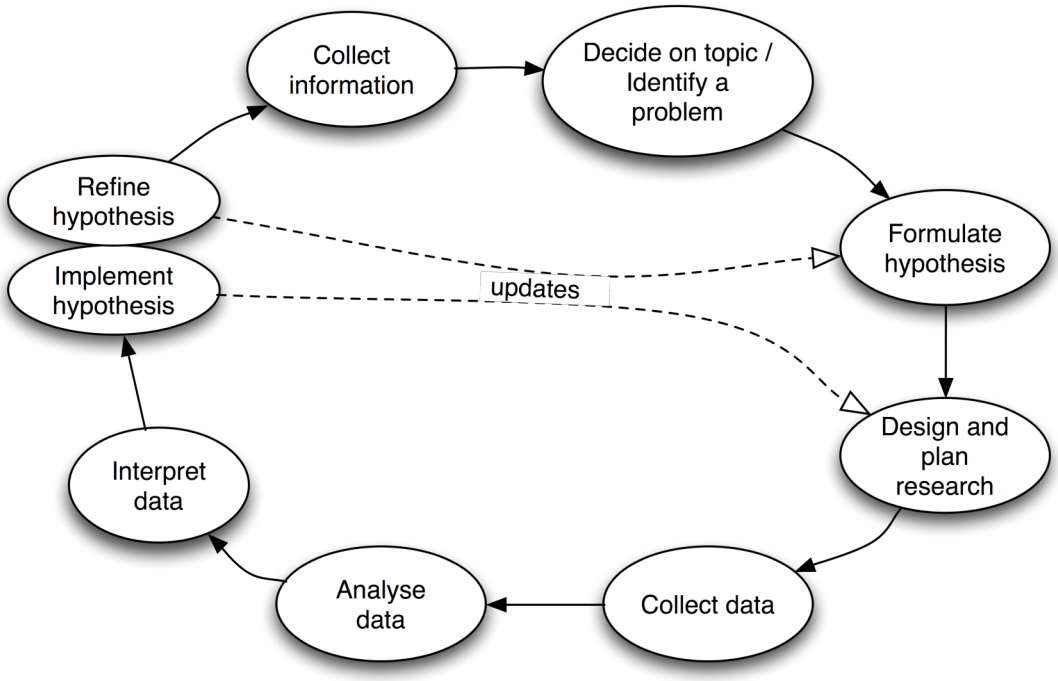

Figure 1.2 A model of the research cycle, (Hutchinson, 2004)

The collected background information and the requirements will show issues not addressed by the current state of the art. From these issues the topic or topics for research are decided. Formulating the hypothesis focuses the topic into manageable research objectives. To support the hypothesis, the *design and plan research* phase becomes a build and test exercise to create a software system or environment. The testing exercise is the experiment under which the hypothesis is tested, and this provides the data to be collected.

The data collected can be quantitative or qualitative. This research concentrated on qualitative collection by gaining expert opinion measured against the motivation scenario. The approach used in this research was the analysis and interpretation of expert opinion, comparing the built system against the outcome required to successfully achieve the motivating scenario. The outcome of the analysis and interpretation is used to refine the hypothesis, which can be tested in another experiment.

The research cycle described in this thesis starts by gathering requirements from the motivational scenario for the collaborative problem solving diagnosis of aircraft engine in the DAME project. This builds a view of the business problem. This information was augmented by a literature review across the areas of business processing, workflow management and Grid Computing. Additionally, works about Grid Computing were substantiated with practical experience by experiments on Grid Computing installations.

Practical experience has been gained by using middleware from Globus Alliance[1], and deploying services across the White Rose Grid[2] (WRG). Other experience with issues of Grid Computing was gained from White Rose Grid user group meetings, and Globus Alliance and GGF[3] email forums.

The hypothesis was formulated by identifying issues within the motivational scenario and resulted in the generic architecture presented previously in Figure 1.1. Details of the issues and identification of the problem can be found in section 4.3. To address the research objective a secure workflow management system was implemented in the DAME demonstrator, an enabled the integration of contributions across the DAME consortium.

The built system was demonstrated, along with presentations of the architecture, during DAME project meetings and at external conferences. Qualitative opinion was collected by semi-structured interview of two experts from the industrial partners at Rolls-Royce and DS&S. Further qualitative data was collected from the discussions at DAME meetings and external conferences. Analysis of the results from the demos and interviews, along with recent developments in the areas for workflow, Grid Computing and access control, was used to feed into updating the hypothesis, refining the model and architecture for secure collaborative access control in workflows. The model was evaluated against published comparative solutions. The results of this analysis are reported along with areas for future research in Chapter 7.

1.6 Outline of Thesis

Chapter 2 provides the background to different the topics areas that are combined in this research. Firstly, it introduces collaboration in business processing and workflow modelling in workflow management. Next introduced is the facilitating concept of service oriented architecture, which promotes loose coupling of components and how this allows composable systems from distributed services to support collaboration, including Web Services, Web Service workflows, and loosely coupled resources using Grid Computing. Chapter 2 completes with an outline of security architecture and solutions to access control, applicable to secure collaborative workflows.

---

[1] Globus Alliance, http://www.globus.org

[2] The White Rose Grid, e-Science Centre of Excellence, is part of the consortium of Universities Leeds, Sheffield and York. See http://www.wrgrid.org.uk

[3] GGF, Global Grid Forum, http://www.ggf.org

Chapter 3 explains the motivating business scenario from the DAME project, with examples of workflows and VOs. Chapter 4 identifies the requirements for collaborative workflow, security and collaborative access to commercially sensitive services from the DAME scenario.

Chapter 5 evaluates the requirements from the DAME scenario, to produce a generalised model for secure collaborative workflow. The considered analysis yields the Workflow-Team Policy Architecture, which is presented as a model for controlling dynamic teams sharing Grid Service instances across organisations.

Chapter 6 presents an experiment to test the Workflow-Team Policy Architecture in the DAME demonstrator. The secure collaborative workflow system is implemented to control Grid Services that are deployed by the DAME partners and execute across the White Rose Grid

Chapter 7 presents the evaluation of the system including expert interview results and compares the Workflow-Team Policy Architecture with developments in security, Grid Computing and workflow, analysing with reference to the experience of building the workflow to execute Grid Services across organisations. Recommendations for future research in this area and complementary work are presented in 7.7.

The thesis and findings are summarised in Chapter 8. After Chapter 8 is the Bibliography and Appendices, which contains the overview model of the constructed DAME workflow management system architecture and detailed interview results and questionnaire.

Chapter 2

Background

In this chapter, the background topics to the research are covered. To address the problem of secure collaborative workflow, several areas need to be addressed. Firstly, the chapter covers collaboration and collaborative working with computer support. This leads into a definition of workflow and workflow management systems, including the definitions of collaborative workflow aimed at computers supporting collaboration across organisations.

Section 2.5 covers the topic of service oriented architectures (SOA), detailing its characteristics, its component architecture for service discovery and how that combines with a business supply chain. To elaborate on SOA, technology solution Web Services is described and linked to workflow solutions that can provide applications by aggregating services across organisations. In section 2.5.5, business processing across organisations is described as the VO and linked to Grid Computing. Grid Computing is defined in section 2.5.6, relating the discovery and use of compute resources to the SOA, Web Service technologies and Grid Computing workflow solutions.

Sections 2.6 and 2.7 provide descriptions of security requirements for computing systems, especially addressing access control for collaborative use of workflow and services, including solutions used in Web Services and for cross-organisational use of Grid Computing resources.

2.1 Collaborative Working

Collaborative working is the task when more than one person is involved in achieving a goal or objective in a piece of work. Collaboration requires communication between people in a collaborative team, such as speech, facial expressions, passing a paper document. Additionally computer supported collaboration involves passing information electronically, such as a document, status of a process or video conferencing. Human communication is information exchange in context and the collaborations can be expressed as a process, capturing the context of the communication.

There are three different styles of collaboration (Jackson, 1999):

- Fixed, where the participants and their actions are predefined;
- Ad-hoc, which is peer-to-peer communication which adjust themselves to the organisation where they exist; and
- Semi-fixed, where the participants have some flexibility in there actions to achieve an objective.

Collaborative models of the processes can be defined using roles. Roles categorise the types of actions a person can perform within an organisation. This allows the process to be defined without specifying people's names.

2.1.1 Collaborative Virtual Working & CSCW

Computer Supported Cooperative Work (CSCW) is characterised by a group of users interacting and collaborating using shared objects towards some common objectives supported by a computing system (Ellis et al., 1991, Papazoglou and Schlageter, 1998, Simon and Marion, 1996). Using a distributed computing system, the collaboration is considered virtual because the users do not need to be face to face to collaborate (Jackson, 1999). CSCW systems are mostly visual interfaces that extend text based collaboration and support different methods for groups to collaborate. These are synchronous communication, such as video conferencing (Churchill et al., 2001, *Access Grid*, 2006), and asynchronous communication, such as email or online forums, and combinations of both. Synchronous and asynchronous communications can be defined within a collaborative process, in the styles fixed, ad-hoc or semi-fixed. The following section concentrates on fixed and semi-fixed processes that are expressed as workflows.

2.2 Workflow

Workflow is based upon office processes where lists of jobs are assigned and carried out. These processes could be issuing an invoice, routing a document or processing an order. With an automatic workflow management system the processes are managed by a computer program that assigns the work, passes it on and tracks its progress. The advantages of reduced errors, not misplacing work and parallel processing can yield a more effective and economic work place (Simon and Marion, 1996).

Workflow is the sequence of a process through which work passes to completion. It is the computerised facilitation or automation of a business process, in whole or part (Hollingsworth, 1995). Business processing can be fully or partly automated, such that workflows may be included within a business process and executed using a workflow

management system (Aalst and Hee, 2004, Smith and Fingar, 2003). Workflow as a technology can be used to support a number of different areas that include:

- Image Processing
- Document Management
- Electronic Mail & Directories
- Groupware Applications
- Transaction-based Applications
- Project Support Software

Business processes describe the core activities of a business, and workflow is the automated component that is managed on an IT system. A business process can be described and modelled, then controlled in a management system to monitor progress. The advantages of capturing a business process and automating workflows include the ability to control and monitor several executing processes in various states, in a repeatable manner, on large scale or scalable systems to accommodate growth or peak activity periods. It also has the ability to define points where work is allocated to appropriate resources, which may be computers or people depending on process definition or task requirements.

Alonso et al. (2004) states that workflow is about defining processes and facilitating the definition and maintenance of business logic. Workflow process models normally have two levels of representation. The first level defines the process of the flow of work through the system. The second level represents the state of execution of a particular process, usually described as an "enactment". As the first level defines the process, we can use the term Process Definition. This representation of a process is usually captured graphically and can be stated using mathematical rules, such as Petri-nets (Aalst and Hee, 2004), pi-calculus (Milner, 1993) and UML activity diagrams (Object Management Group, 2003). In particular, Eriksson and Penker (2000), illustrate the use of UML for business process modelling, showing the use of UML at a higher level of abstraction from traditional software system modelling. The field of business process modelling has specialised these generic process languages (Petri-nets and pi-calculus) with incarnations such as Business Process Modelling Language (BPML) (BPMI, 2002) and Business Process Modelling Notation (BPMN) (Object Management Group (OMG) / Business Process Management Initiative (BPMI), 2006). Business Process Execution Language (BPEL) (Andrews et al., 2003) uses a subset of BPMN and a means to execute a process on implemented systems (BPEL). BPMN contains a rich syntax to describe process flow, participants and collaboration and is independent of implementation on IT platforms. BPEL is exclusively tied to implementations for Web Services containing descriptions of control flows for a

workflow manager and message flows to and from Web Services. BPEL scripts describe processes using interface descriptions of Web Services and can optionally be bound to specific deployed Web Services, by providing URLs (Uniform Resource Locator) in the process script.

Aalst, et al (2002b) categorised the control-flow constructs in workflows. Examples are basic sequence, parallel split and multiple choice split. Each construct has been used to compare different workflow languages by classifying the degree of support each language provides (Wohed et al., 2002). No single workflow language implements all the constructs and opinion is divided on the 'best' workflow language.

Business processing is the execution of business logic in the integration of business components. In the Business /Integration Tier, Figure 2.1, workflow is the aggregation of IT components; often distributed, components can be modelled as services. Web Services and service-oriented architecture are introduced later in this section, 2.5. The second level of representing workflow is the enactment, which contains the state of a particular process, is covered later in this section, in 2.3 Workflow Management.

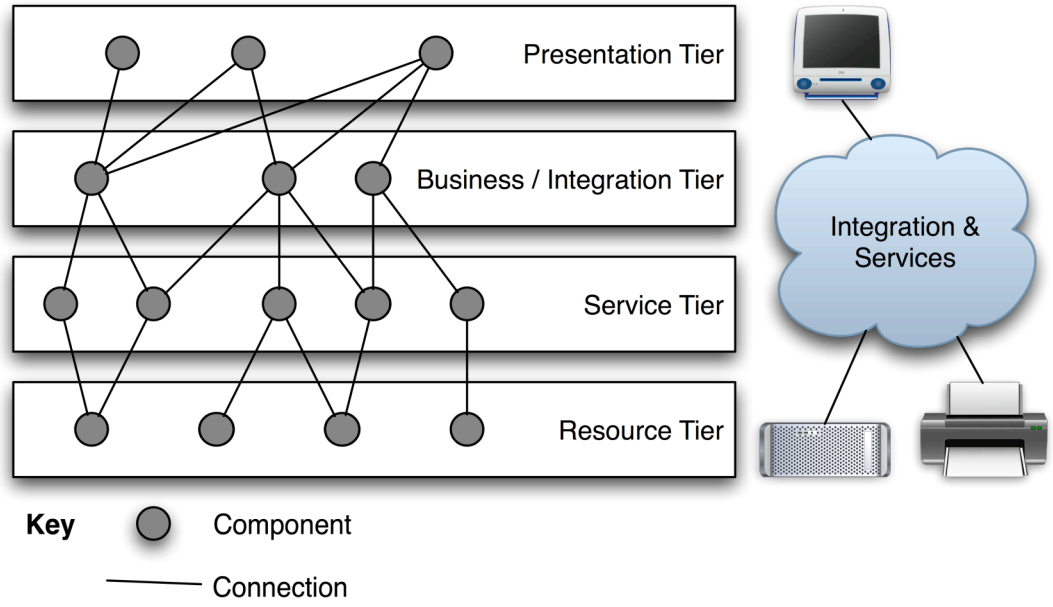

Figure 2.1 Service-oriented architecture tiers

During the workflow, assets may be created for just the duration of the workflow. In a service based workflow it is possible to initiate a service in one task to be accessed later in another task in the workflow, possibly by a different user. In the Grid Computing model this is represented as a Grid Service instance (from Grid Services in the OGSA standard (Foster et al., 2002)). This becomes the temporal business asset for the duration of the workflow, similar to temporal user relationships as described by Chandramouli (2000).

2.3 Workflow Management

A workflow management system is defined by the Workflow Management Coalition (WfMC) as:

> *A system that completely defines, manages and executes "workflows" through the execution of software whose order of execution is driven by a computer representation of the workflow logic, from (Hollingsworth, 1995).*

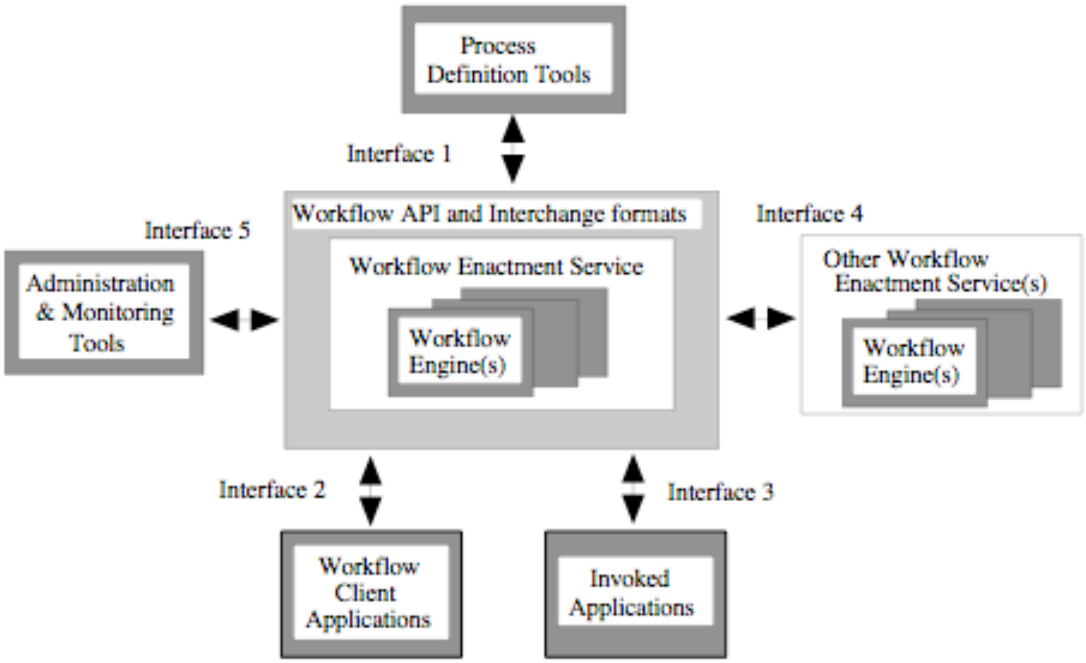

Figure 2.2 Workflow reference model – components & interfaces, (Hollingsworth, 1995)

Figure 2.2 shows the main components of workflow management (Hollingsworth, 1995). The central component is the Workflow Enactment Service. This is the interface point to the execution of workflows by the Workflow Engines. The Enactment Service directs the communication to the Workflow Engines from some of the external interfaces 2, 3 and 4. Each Workflow Engine executes an instance of a defined workflow sequence. The workflow sequences are defined in the Process Definition Tools, which would create workflow scripts in a language such as BPEL, passing them into the Workflow Enactment Service via Interface 1.

The Workflow Client Applications provide user input and output to executing workflows via Interface 2. This is the user point of collaboration in active workflows. The Workflow Engines execute sequences of work instructions, passing invocation and control messages to and from the Invoked Applications, via Interface 3. The Invoked Applications

do the processing work, such as query information from a customer database or perform algorithmic signal processing.

In the Workflow Reference Model, the interface definitions specify methods to connect these components. Interface 4 is the definition on how Workflow Enactment Services can invoke other Workflow Enactment Services. In Web Service workflow engines, this is implemented as another Web Service interface. Interface 5 defines the protocol for workflow management tools to control and monitor Enactment Services. This provides a view of which workflows are executing and methods to manage the Enactment Service component.

A number of interoperability scenarios can be created from this architecture, including the execution and control of services and workflow management systems outside the traditional organisational boundary. By implementing the components as distributed services, a service-oriented architecture (SOA) can provide loose coupling between business process definition and workflow implementation. Distributed services can be replaced or modified without the need to change the business process. This allows the workflow implementation to change more often than the business process definition. It enables services to be outsourced if required. Outsourcing services leads to a business model that includes service suppliers and even commodity computing from compute resource suppliers. The combination of these creates the model for Grid Computing and VOs (Foster and Kesselman, 2004). In a situation where users and services collaborate across organisations, secure methods of access control are important.

2.4 Collaborative Workflow

In a collaborative workflow, people act in roles given in the workflow definition. Since roles group together the job functions a person has within an organisation, a role will have defined permissions within the organisation. People have assignment to roles, which provides them with the permissions of the role (Ferraiolo et al., 2001). In the case of workflows, the role has a responsibility for completing tasks in the business process. As users enact roles in the workflow, they can combine their skills forming a team to bring the workflow to completion. An example of the importance of roles is given in section 2.7.2.

John Yunker (2002) presents a high level view of the ebXML[4] suite that allows one to express a process that is implemented as instances. This definition is concerned with the interfaces between the business that contribute to shared assets and business goals, in Figure 2.3. Yunker explicitly avoids defining a collaborative process. Within the definition the

[4] ebXML Electronic Business using eXtensible Markup Language, http://www.ebxml.org/

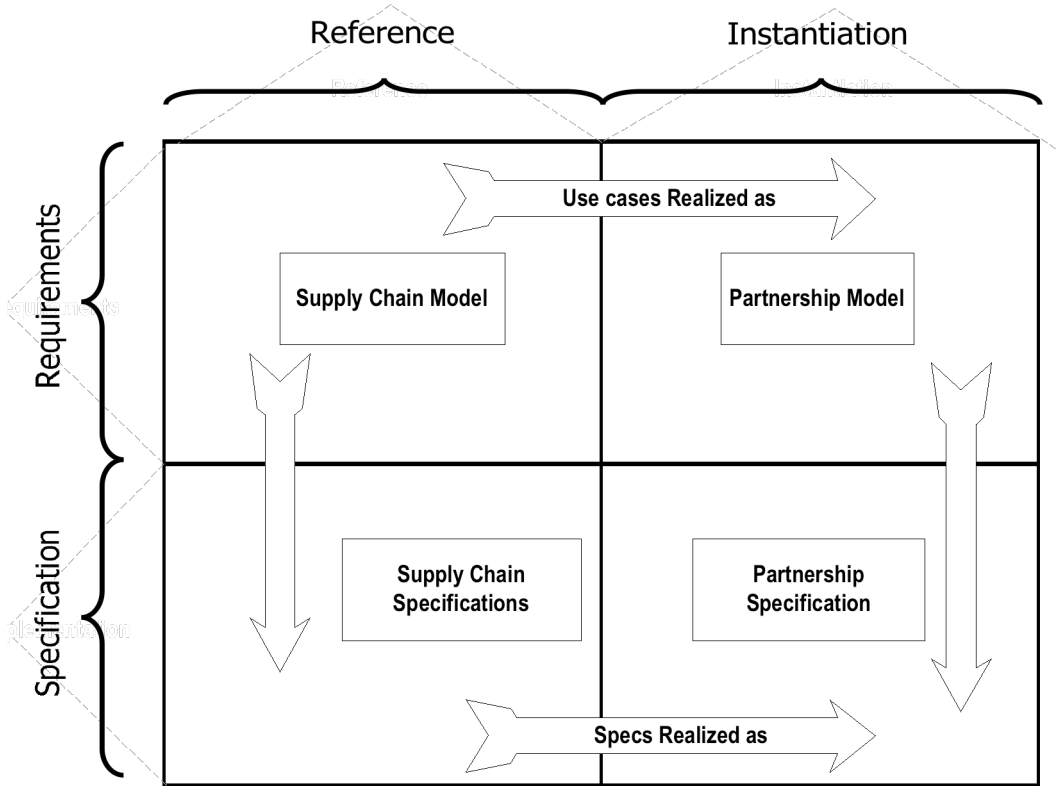

Figure 2.3 Elaboration methodology from business process to collaborative workflow, (Yunker, 2002)

supply chain is modelled, along with the business partners who contribute to a shared partnership specification and allows the definition of types of shared assets, and subsequently what controls are required on shared instances. This shows that during instantiation roles are realised as people and processes have state and can create assets, that may be shared and have a lifecycle that may or may not extend beyond the lifetime of the workflow.

A business process explicitly defines tasks that a role can perform; yet, it does not link process definitions to access permissions for the role, either during process definition or execution. Chandramouli (2000) discusses how role based permission can be linked to process definitions, but does not illustrate the how to map from one to the other. Mendling et al. (2004) does propose a method to extract role based access control from BPEL using XSLT (Extensible Stylesheet Language Transformations). The later section 2.7.2 goes into further detail of role based access control and links to business processing.

2.5 Service Oriented Architecture

Service Oriented Architecture (SOA) offers flexible approaches to distributed systems engineering with quality of service and evolution. A service is a function that is well

defined, self-contained and does not depend on the context or state of other services. The following section describes the characteristics of SOA, without directly linking to technologies. Later sub-sections outline the implementations of Grid Computing and Web Services using SOA. It should be noted that some capabilities of Web Services fall outside the scope of SOA, and not all SOA characteristics are implemented in Web Services (Papazoglou and Dubray, 2004).

2.5.1 Main Characteristics

The main part of SOA is its decentralised middleware (Alonso et al., 2004). By ensuring the middleware is not implemented in a single place, it allows individual services to be independent of other components within the system. A centralised middleware implementation would provide a single point of failure and tend to force vendor specific tie-in. However, a decentralised middleware requires replication of functionality at every location a service is deployed. This overhead is justified due to flexibility provided by the following characteristics:

- Loose Coupling is an architectural property exhibited by services that makes them independent from the state and context of other components in the system. Loose coupling is defined by the following characteristics:
  - Defining services by interface, including data exchange and behaviour (pre/post conditions).
  - Platform/Language independence. By separating the interface from the implementation, this promotes language and platform independence. This also prevents the dependence on a particular vendor.
  - Discovery by abstract descriptions. The interface definition provides the means to locate services, irrespective of deployment environment, therefore independent of implementation platform or language.
  - Evolvable systems independent of implementation/platform. Services can evolve by adding functionality, improving operation or moving deployment environment. Loose coupling and interface definition allows this to happen independent of other system components.
  - Interchangeable by interface definition. An interface definition allows different service providers to offer the same or similar services. Differentiators may be defined by cost, response time, accuracy, security or other quality of service measures.
  - Autonomous services. Loose coupling is achieved by removing dependence across the implemented system. In a distributed system, independence allows services to

be reused in different contexts, without repercussions on the original use, also evolution of individual services without affecting other services or the entire system.

- Reusable services. Loose coupling promotes the reuse of services in new contexts not previously envisaged. Rich definitions of interfaces and behaviour contracts would enhance the validation of processes employing services in new contexts. Contracts of interface and behaviour are part of SOA, but tend to be weakly defined in implementation technologies such as Web Services.
- Inter-organisational, by using a loosely coupled system, an application or process is able to use services developed outside of the organisational bounds. Interface contracts play a major part in this, and where inter-organisational use of services currently exists in Web Services. Business contracts tend to specify acceptable use. Specifying these contracts in the implementation domain and policing them is currently limited, with current research into this area (Padgett et al., 2005).

- Encapsulation. There is an implicit level of granularity for services. Most literature attempts to puts this in terms of previous software technologies, so that objects are the smallest units built, then components are made of objects, finally services are the integration of components. Services should encapsulate a useful business component that can be offered in different contexts.
- Discoverable, mechanisms in SOA allow a service to be discovered. This could be by service name, interface type, context, behaviour, performance/Quality of Service (QoS), or organisation.
- Message Based – platform and language independent. A service uses a message-based asynchronous data exchange (Papazoglou and Dubray, 2004). This is different from synchronous remote method invocation, and ensures independence from the execution environment. Asynchronous messaging is useful when the client requesting a service does not require an immediate response and can continue with other tasks. Examples are sending a purchase order, and waiting for confirmation of goods despatch. New approaches extend this to document exchange, which proposes methods to pass a document in a peer-to-peer manner (Schoder et al., 2005), with mechanisms to protect sections from unauthorised access.
- Well-defined interfaces. The strength of SOA lies in defining services by interface, not by implementation. The interface allows independence of implementation. Interface definition can include data exchange formats, policies of use, service behaviour (pre/post conditions) and QoS.

- Service Level Agreements, as mentioned in interfaces service level agreements (SLA) may be used in the interface definition to specify QoS (Padgett et al., 2005).
- Cross-organisational integration (by processes). Services can be provided for use by other organisations. Usage policies and access control mechanisms are required in most situations to protect the service from misuse.

2.5.2 SOA Components

The main components of SOA (Endrei et al., 2004) are summarised below and illustrated in Figure 2.4 showing the basic communication between the components for locating and invoking service requests:

- **Service**: Logical service execution entities. The contracts are defined by one or more published interfaces.
- **Service provider**: The execution entity that implements a service specification.
- **Service consumer** (or requestor): The software entity that calls a service provider. Traditionally, this is termed a "client". A service consumer can be an end-user application or another service. The act of finding a type of service from a service registry or using a service broker requires that the request be required to "bind" to the implementation. The service consumer can be bound to the service implementation during the design/creation of the service consumer, or using late binding during runtime, when the service consumer and implementation are bound before the service is requested.
- **Service registry**: A specific kind of service provider that acts as a registry and allows for the lookup of service provider interfaces and service locations.
- **Service broker**: A specific kind of service provider that can pass on service requests to one or more additional service providers.

The architecture in Figure 2.4 illustrates how loose coupling is achieved. A service consumer that wants a kind of functionality can first locate the 'service type' that performs the desired function. The service type will describe the properties of the service interface. The service consumer will then need to locate a deployed implementation of the service type. The location address is used to bind the service consumer to the service from the service provider. Discovery of the service type and deployed service can use a form of registry, at design time, system configuration (i.e. by the system operator, before execution) or during runtime (W3C, 2006).

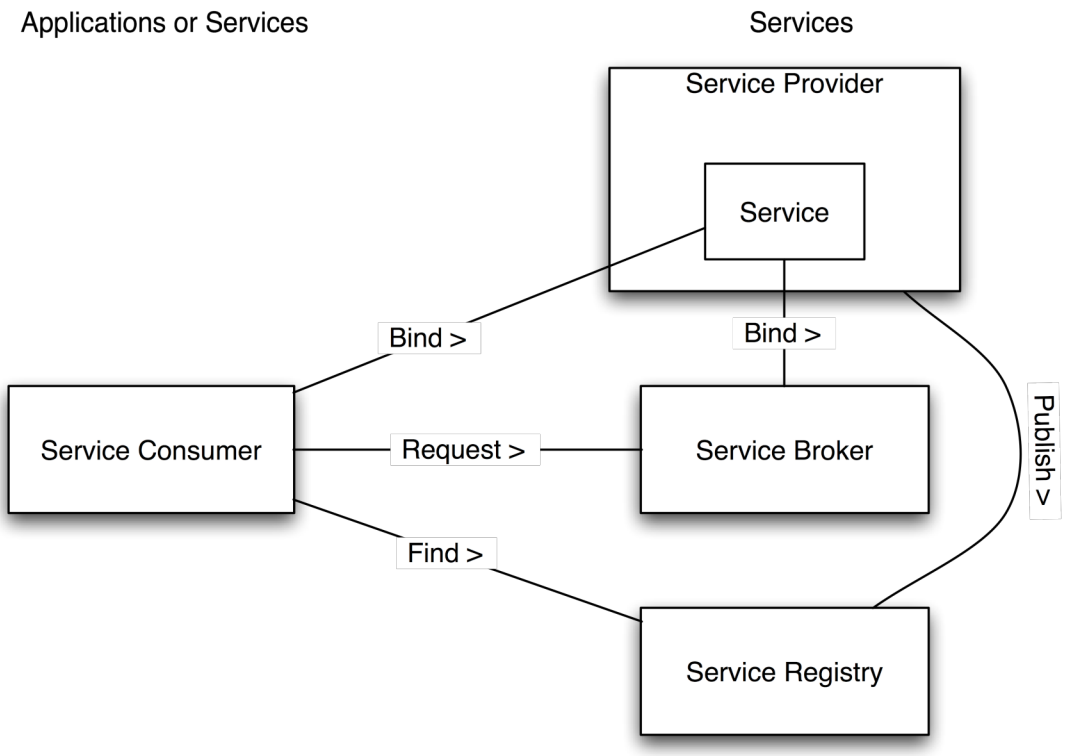

Figure 2.4 Service-oriented architecture main components

Looking back at Figure 2.1, the SOA components are intended to reside in the Service Tier. They provide an interface for Resources, such as computing, printing and processing, to be used within the Integration Tier. Service integration can be performed in a workflow engine or within another service, since the integration or aggregation of services can be consumed as another type of service. Therefore, system components within the Integration Tier can be the implementation of service interfaces. For example, a workflow engine would be both service consumer and service provider, for a given interface to a define workflow sequence.

2.5.3 Supply Chain

SOA supports traditional inter organisation supply chains by enhancing a traditional service business approach, such as Banking (Rust and Kannan, 2003). SOA itself promotes inter-organisational trading, however, current implementation such as Web Services lack the mechanisms for accounting and payment. Work in Grid Computing for economic modelling (Buyya, 2002) is aimed at marketing grid resources for traditional supply chains and VOs. Buyya (2002) addresses different marketing models for trading compute power as a service, that can be dynamically consumed from different suppliers on demand.

2.5.4 Web Services

Web Services provide internet-based, machine-to-machine communication. Not to be confused with web sites, which are human-to-machine communication. Web Services are an implementation of service-oriented architecture. Although it is possible to use Web Services to create a system that is not service oriented, the following describes how Web Services do relate to SOA. According to Papazoglou and Dubray (2004), Web Services follow the model of software-as-a-service provided by ASP (Applications Service Provider) allows the consumer to 'rent' software applications per use from a remote hosting system. Web Services provide connection of components, so the consumer can create their own application or consume a component which itself may be an aggregation of Web Services.

Papazoglou and Dubray (2004) state:

> *Web Services constitute a distributed computer infrastructure made up of many different modules trying to communicate over the network to virtually form a single logical system. Web Services are modular, self-describing, self-contained applications that are accessible over the Internet. (Papazoglou and Dubray, 2004 p.2)*

They enable developers to construct applications across the Internet, and across organisational boundaries using any platform or language required. Once a Web Service is deployed, other applications and Web Services can discover and invoke that service.

2.5.4.1 Web Service Communication

The structure of Web Service communication uses messages of XML (extensible Meta-Language) sent using SOAP (Simple Object Access Protocol) sent over HTTP (Hypertext Transfer Protocol). At the base level of Web Services, HTTP is chosen to allow messages to pass through firewalls, especially corporate firewalls, which were not the case with other distributed object communications, such as CORBA (Common Object Request Broker Architecture). SOAP is used as the message container, providing addressing and message structure formatted in XML. Residing on HTTP its request/response method operates similar to HTTP.

The SOAP structure is briefly described below:

- Envelope, which describes where to send the message and states the format for the rest of the message (header and body) and how to process it;
- Header, which states where the message came from, the policies required to understand the body and possibly digital signature for the message body;
- Body, this is the content of the message defining the method to be invoked and its arguments.

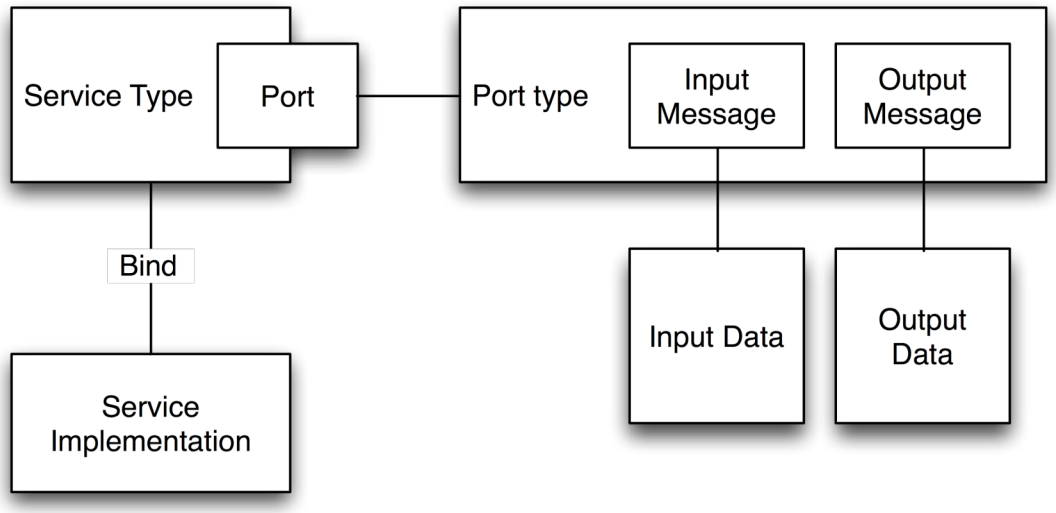

Figure 2.5 WSDL structure overview

WSDL (Web Service Description Language) contains a description of the Web Service, specifying the messages passed in and out and the where to find the service to execute. Service type descriptions may be defined in a separate document from the binding of the service implementation, such that there may be multiple implementations of the same service type, possibly on the same or different servers or from the same or different organisations. Figure 2.5 shows a simplified view of the WSDL structure. The Service is described as a service type that has a port of port type. The port type describes the input messages that must be sent to the service and the output messages that are received. The URL of a deployed service binds the service implementation to the service type.

UDDI (Universal Description, Discovery and Integration) is a registry intended to list businesses similar to Yellow Pages™. In Web Services, it is used to list the Web Services provided by businesses, with a description in human readable text (for example English) and a location to the WSDL document. The WSDL can be retrieved, usually from the remote server hosting the Web Service, and interpreted for selection and binding for use in Web Service workflows.

2.5.4.2 Web Service Workflow

Aggregation of Web Service to form applications can be specified using workflows. That is the orchestration of sequences required to perform actions within an application can be defined using a process language such as BPEL. Other processing languages for Web Services have been proposed, however BPEL is the most commonly used and supported in workflow tools. BPEL is the result of joining two previous workflow languages. One from IBM WSFL (Web Services Flow Language) (Leymann, 2001) which included support for service lifecycles and used constructs relating to WSDL making composition from service

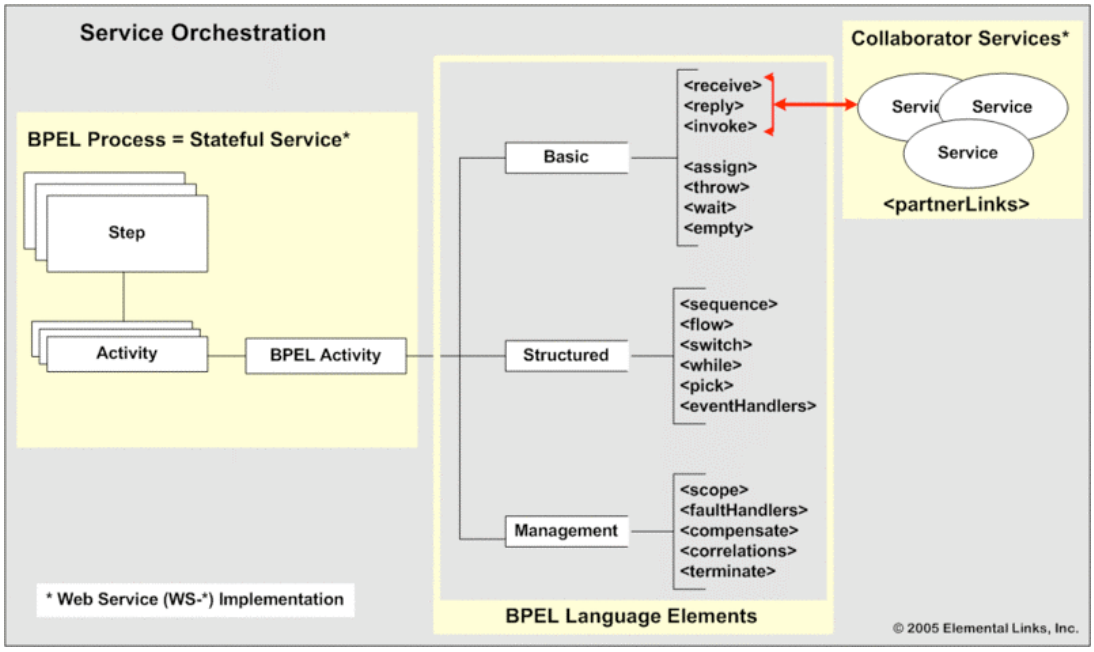

Figure 2.6 BPEL Illustration, from (Michelson, 2005)

descriptions logical and one from Microsoft XLANG that included support for exception handling (Thatte, 2001).

An illustration of the constructs in BPEL from (Michelson, 2005) is shown in Figure 2.6. This illustration shows the basic language elements of BPEL and how it relates to the definition of a process and the execution of services. Two statements back up the diagram:

> *"A BPEL process is a service orchestration, used to describe/execute a business process (or large grained service), which is implemented as a stateful service"*
>
> *"A process is comprised of steps, steps have activities AND activities are BPEL language elements, and the basic activity elements are the ones used to interact with the collaborating services <partnerLinks>", from (Michelson, 2005).*

Other important workflow languages include WSCL (Web Services Conversation Language (Banerji et al., 2002) proposed by Hewlett-Packard is concerned more with extending the description of a Web Service to detail how the service interface is used within a process. Another is WSCI (Web Services Choreography Interface) (W3C, 2002), proposed by Sun Microsystems, Intalio, SAP and BEA, it provides another method of describing workflow. However, it concentrates on controlling the sequence of operations required for a single service to operate. Whilst it does describe workflow, it is intended to sit between BPEL and WSDL in describing "choreographed message exchange" (Aalst et al., 2002a).

Examples of workflow engines from major workflow tools vendors are IBM Websphere, Microsoft BizTalk and BEA Weblogic. There are many other smaller companies providing workflow solutions due to Web Services open standards. Lists of current vendors are available from the BPMG[5] and WARIA[6].

2.5.5 Virtual Organisations in Grid Computing

Grid Computing is a solution to the growing number of IT platforms used within a business or scientific research environment that require connectivity to integrate applications and better utilise unused compute processing power and data storage. As a technical solution, it is a middleware of connectivity standards, specifying data protocols and behaviour across heterogeneous platforms. For the problem domain, it aims to achieve greater utilisation of existing resources and scalability of resources to achieve new high power applications, that are not constrained by geographical location and provide new opportunities for integrating processing across traditional organisational boundaries.

This last point is the basis of the Virtual Organisation (VO), from Foster et al. (2001), who define a VO as:

> *The real and specific problem that under lies the Grid concept is coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations. The sharing that we are concerned with is not primarily file exchange but rather direct access to computers, software, data, and other resources, as is required by a range of collaborative problem-solving and resource-brokering strategies emerging in industry, science, and engineering. This sharing is, necessarily, highly controlled, with resource providers and consumers defining clearly and carefully just what is shared, who is allowed to share, and the conditions under which sharing occurs. A set of individuals and/or institutions defined by such sharing rules form what we call a virtual organisation (VO), from (Foster et al., 2001)*

---

[5] Business Process Management Group, lists current workflow tools at http://www.bpmg.org/chl_bpmg_solution_providers.php

[6] Workflow and Reengineering International Association lists current workflow tools at http://www.waria.com/databases/wfvendors-A-L.htm

The VO is a business enabler to create business teams across organisational boundaries. An illustration of this can be seen in Figure 2.7. This shows business activities that use resources to achieve business goals. Traditional organisations are shown as shaded rectangular boxes that encompass the required resources (squares) and business activities (circles) to achieve the business goals (triangles) for that organisation, within its own administrative bounds. VOs are shown as with the dotted line around an irregular area. The VO is the collaboration between organisations to achieve new business goals, consuming resources and activities across organisational boundaries. The business goals and some activities only exist within the VO. However, the resources have one owning organisation each. Management of shared resources and activities across organisational boundaries is key to defining VOs. The following text summarises the ideas behind Grid Computing and VOs, defined by Foster et al and used across many Grid Computing projects in academia (e-Science) and commercial, such as IBM On-Demand.

2.5.5.1 The need for the Grid

Grid technologies are needed to support the sharing and coordinated use of diverse resources in dynamic VOs. A dynamic VO is one that can change its boundary during the

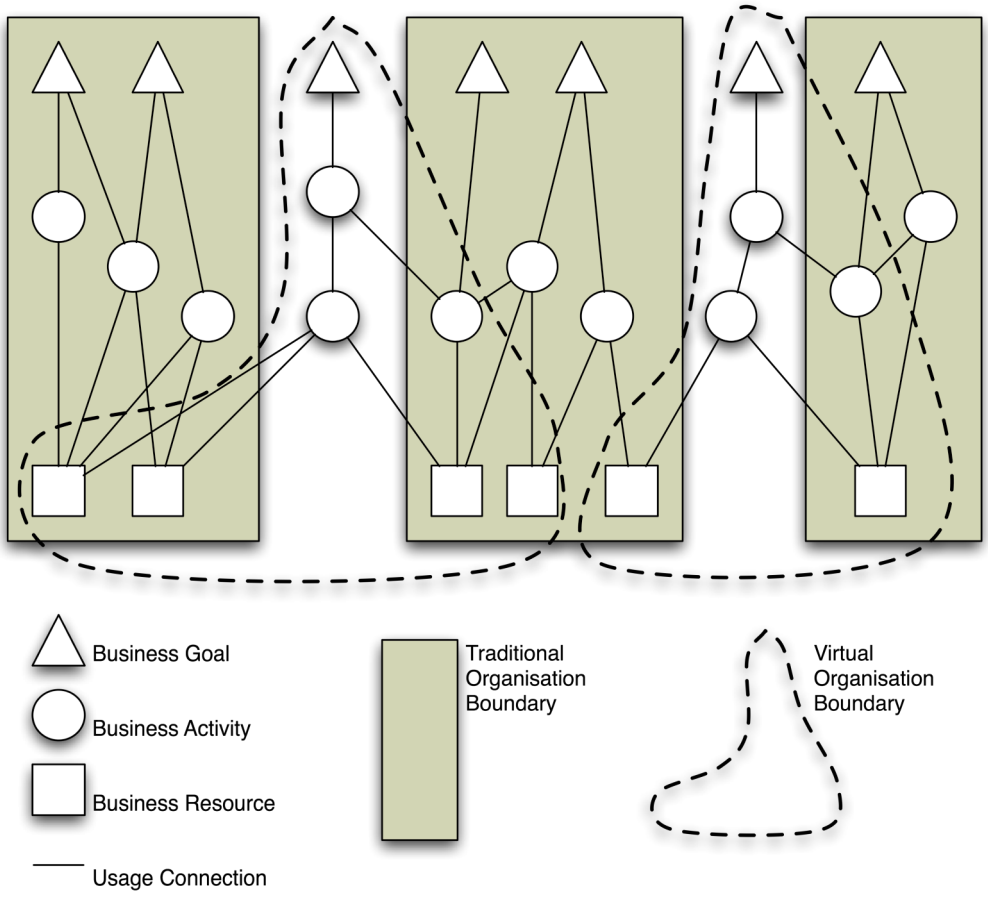

Figure 2.7 Virtual organisations create new business goals across organisations

lifetime of the VO. This allows temporary use of resources or services offered by another organisation in a collaborative manner. As shown in Figure 2.7, this differs from a traditional supply chain, highlighted by new business goals attributed to the VO and sharing of resources between organisations. An example of a VO applied to a business case can be found in (Russell et al., 2005), and is shown in the next chapter.

2.5.6 Grid Computing

During this research, the dominant Grid Computing standard came from The Globus Alliance (*The Globus Alliance*, 2005), with implementations distributed as Globus Toolkit released as version 2, version 3 and the current version 4. The Globus Toolkit 2 (GT2) (The Globus Alliance, 2002) was aimed at sharing computing resources; this is defined as compute processing, data storage and networking. Due to the high cost of large high-performance computing (HPC), Grid Computing was aimed predominantly at providing access to HPCs within and across organisations. GT2 implements a resource management middleware that includes resource discovery, monitoring, secure transport, basic access control and an interface for sending processing tasks to compute resources.

Globus Toolkit 3 (GT3) (Foster and Kesselman, 2004) improved the connectivity of Grid Computing by using Web Service interfaces to access the new concept Grid Services. Grid Services extend the Web Service interface by providing access to the grid management interface, such as resource management and security. It also provides a method to create stateful services, by using the factory pattern (Gamma et al., 1995) to create a service instantiated with its own URL, using a defined service interface. A stateful service is a service in which its internal state is reflected in the interactions with its clients. The creation of a stateful service creates an executing service instance, which has its own address that is returned by the service factory on creation.

The model of connectivity changed in GT4 (Globus, 2004) in favour of WS-Addressing (W3C, 2004), where the location of the service instance, now known as a WS-Resource, is passed in the SOAP envelope using a static URL. This is more compatible with Web Service workflow standards and tools, such as BPEL and IBM Websphere (*IBM Websphere*, 2003).

Another Grid Computing standard is UNICORE, (*UNICORE Forum e.V.*, 2006) which stands for Uniform Interface for Computing Resources. Comparing GT4 with Unicore, Unicore provides a more complete and secure method to access computing resources. However, since making resources commonly accessible is the objective and GT4 uses open Web Service standards, and has been adopted by more organisations, it is therefore, more commonly used.

This section expands on the main characteristics and reasons behind the development of Grid Computing.

2.5.6.1 Grid Resource Management

To achieve resource sharing there needs to be a mediation layer, a resource management layer. In grid technologies, this is manifested as middleware using common interoperability protocols between service to resource, resource to resource and service to service communications. The middleware offers this commonality in both message exchange and behaviour. The main difficulty is the middleware's connection to the resource. The grid solution has found that most middleware services are best implemented at each resource in a peer-to-peer type manner, with very few centralised services. Such central services would be information look-up services, such as identity servers for authentication.

The following sections describe the services provided by the middleware to coordinate resource sharing.

2.5.6.2 Grid Connectivity Security

When connecting to a service or resource these security issues apply:

- Identity assurance – a means of authenticating a user, service or resource as the source of a message/request. The authentication is bi-directional as both parties in the request need to be sure of each other's identity.
- Private, confidential data exchange – a means to protect message contents from being read by anyone who intercepts it. A message usually has three parts, the address (including the senders address for response), the request and the message content. The privacy needs may require any or all of these parts to be kept private. It should be noted that it is usually only possible to protect the message contents.
- Integrity, assurance of content – a means to 'sign' the message to ensure the contents are what the sender intended. Usually achieved by digitally signing with a checksum that will only match the message originally sent. It assures the authenticity of the sender.
- Delegation of access rights – a means to allow the requester's identity to be used by the service to initiate requests to other services on behalf of the user of the service. The user delegates permission to the service and the service can request actions from other services using the user's identity. In this case permission for delegated requests are decided from the rights of the original user.
- Access Policies – a means to specify actions that can be performed. These can be restricted by identity, role, organisation, context, message content, request type,

date/time, etc. Permission can be specified as: 'deny all rights except those allowed in the policy'; or 'permit all rights except those denied in the policy'.

- Local security (resource sharing) – is a means to protect the integrity of the computing system executing a task. It is a means by which actions executed on a resource may not compromise the resource, and that those actions are not compromised by the resource or any other actions/requests executing on that resource.

2.5.6.3 Resource

The Grid Resource Allocation and Management (GRAM) service provides a uniform interface to job scheduling systems on different compute platform implementations for the remote execution and management of jobs. As part of resource management, the resource provides monitor and control services:

- Monitor – this provides state information about a resource, it can also provide static information about a resources configuration/attributes.
- Control – the management of resource, protocols that control a resources management such as enabling, updating, reconfiguration or modifying a policy. This is separate from protocols to use a resource.

2.5.6.4 Coordination

The coordination of resources, services and users are summarised by these actions:

- Discovery – the ability to find a resource or service by searching on descriptions of service properties. Such as, finding a service by name or function, or a resource by capacity or level of security.
- Allocation, scheduling, brokering – managing the time on a resource or service is achieved by these functions in the middleware. These can also be used to discover appropriate resources for execution of tasks before sending information to deploy the task on the resource.
- Usage policies – these policies would typically control access to resources/job queues dependent on how much a resource is used or what functionality is used. Reasons for controlling use can include economic and security. Usage policies can be linked to charging mechanisms (see Grid Economy).
- Access to HPC, Storage – Grid Computing is means to share expensive large-scale resources in different organisations, such as HPCs and very large databases.
- Fault tolerance – the following fault tolerance mechanisms can be applied to improve reliability and sustainability of resources:

- Monitor failures (inc intrusion detection) – awareness of failures in a resource.
- Data replication – services to address data independent of location, allow data to be replicated. Data availability is improved and data replication services transparently provide data from appropriate/available sources.
- Job migration, check pointing – if a resource fails, then check pointing marks useful points of completion, such that the job can be restarted from the last check point, on the same or a different resource.
- Task pool management – similar to job migration, a central service records scheduled jobs that a removed from the job queue on completion. If a resource or job fails it can automatically be restarted.

2.5.6.5 Services

The services in Grid Computing take the form of Web Services. In GT3 (Foster et al., 2002), a service factory creates service instances that provide an interface to the instance of a grid process. In GT4 (Globus, 2004), the service interface uses a static URL, providing access to the resource factory, which creates processing instances called resources. A resource is addressed via the Grid Service URL using WS-Addressing. Both of the GT3 and GT4 methods achieve stateful services. Further detail of services have been given in section 2.5 Service Oriented Architecture.

2.5.6.6 Grid Economy

Grid economy is an area of research that uses economic modelling for the marketing and trading of grid resources in VOs. Buyya (2002) addresses different marketing models for trading compute power as a service, that can be dynamically consumed from different suppliers on demand. The following points are required capabilities of the architecture to enable an economic model for the trade of grid resources as a service:

- Accounting, charging, payment services – a means to record usage of a service/resource and charge for it. Similar services will be required by the consumer to record usage and pay for it.
- Negotiation – a means to negotiate payment for the use of a service/resource. This would be written into a contract including quality of service parameters.
- Service Level Agreements – a contractual agreement on the requested/deliverable quality of service (QoS) that a service/resource must perform to. These qualities may include: response time, security, accuracy and/or availability.

- Economic models – different economic models have been investigated for market based trading of services and resource in the grid community. These include trading models based on auctions and share trading

2.5.6.7 Grid Workflow

Krishnan et al. (Krishnan et al., 2002) proposed GSFL (Grid Services Flow Language) as an extension to WSFL to incorporate the developments of GT3 Grid Services into workflow description. The main proposal was to make provision for peer-to-peer style connections between processing services. In traditional workflow engines, all the control and messaging is centralised in the workflow management system. When the workflow is a pipeline of data processing then large amounts of data are passed back and forth to the workflow engine, which can be costly in a distributed environment. Krishnan et al. propose that data is pipelined between services and that workflow engines control and monitor the process without handling all the data, illustrated in Figure 2.8.

Although there is no evidence that GSFL has been implemented it raised some important issues of workflow in grid systems that still need to be tackled. Such as how to separate control flow from data flow in systems where the data is large (>100MB) and it is not necessary to route the data through the workflow engine.

There are many projects concerned with providing workflow tools and engines for large-scale data processing for scientific grid workflows. These include myGrid (Dept. of Comp.Sci., 2004) which has a Web Services based workflow engine Taverna that uses a proprietary workflow language; Gridbus (*Gridbus Workflow Engine (GWFE)*, 2006), which has a workflow engine using an XML-based workflow language that supports GT4 middleware; and GRACE (*GRACE - Grid Search and Categorization Engine*, 2003), coordinating searches across distributed data.

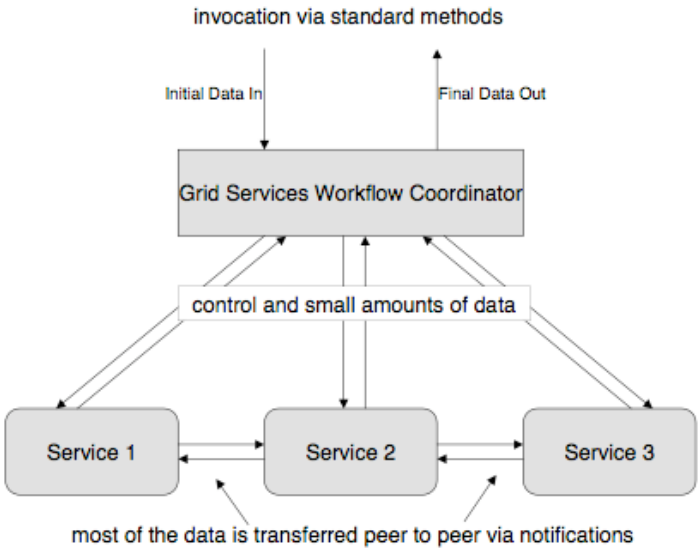

Figure 2.8 Grid Services workflow model, (Krishnan et al., 2002)

2.6 Security

The topic of security becomes increasingly important as more business critical applications and services are exposed to access outside the physical protection of an organisation, such as using the Internet (Agarwal et al., 2002, King et al., 2001, Atkinson et al., 2002). In a distributed environment, each point of entry requires protecting. In this case, the control of access to a service may use a gateway. The gateway would allow legitimate access and deny illegitimate access. This type of protection aims to ensure internal components and information can only be used or seen by those that are trusted to use them fairly. In addition to access control, the identity and type of access can be recorded to audit trails and non-repudiable logging. Access can also be encrypted to prevent others being able to view the messages (privacy), and the access can be digitally signed which prevents others from tampering with the messages (integrity). These are typically associated with 'man-in-the-middle' attacks (Burr et al., 2006).

The technical security solutions are only parts of the overall security solutions related to the respective security policies and must be coordinated with the solutions regarding methods, organisation and competence. Technical solutions are used to fulfil security objectives such as identification and authentication, authorisation and access control, protection against intrusion and attacks, maintaining confidentiality, privacy and integrity of information, non-repudiation, and auditing.

Confidentiality is concerned with keeping the action performed with the service private. This ensures that other parties cannot determine who made the request to a service, the details of the request (the action requested and the contents of any message sent) and the

details of the reply. The content of the messages can be protected by encryption, but full privacy is very difficult since messages contain details of the sender, receiver and the action. Another security concern is denial of service, where the service is attacked, typically by overloading an operation with data, to prevent processing of legitimate accesses.

Integrity is the means to prove that a message is complete and unchanged. It is an important part of any security mechanism to prove that any message has not been tampered with and contains the information originally intended.

Auditing is the recording of actions. This may be used as part of security to track users' actions on a system, both at the system boundary and within the system. Non-repudiation is an important part of auditing; it is the assurance of the integrity of the audit. A non-repudiable log of a user's actions cannot be refuted, and can be used by either the user or supplier of a service or system to prove that an activity took place (Zhou, 1997).

The two main components in security are authentication and authorisation. Authentication is the verification of the identity of a person or a process. For distributed systems, authentication verifies that a message has come from its stated source. The source can be identified as a person, or a process or service, or a computer system or server. The method of verification in computer messages involves a signed message. This message may be as simple as stating the name of the person, and it will be signed by a trusted third party, similar to using a passport to identify someone. Identity certificates such as X.509 (Tuecke et al., 2003) contain the name and organisation of the entity being identified, and will be digitally signed by a certificate authority. Digital signing, using public key cryptography (IEEE, 2000), creates a hash-code from the message data, which is then encrypted using the signer's private key. In addition to verifying identity, the contents of messages can be verified as originating from the sender by digitally signing with their private key, which can be recovered using their public key.

Authorisation is the act of granting permission. In computing, this is granting permission to access files (read, write, delete, etc.) and perform operations on the operating system or remote services.

A policy is an explicit representation of constraints and rules that govern the behaviour of an agent or a system. Policies define the actions that may be performed on the target by a subject. The target or subject can be specified by identity or as an abstraction from identity such as role or organisation. Additionally policies may include other constraints that restrict the action such as context, message content, request type or date/time. Permission can be specified as deny all rights except those allowed in the policy or permit all rights except those denied in the policy. Resolution of policies may result in binary decisions of permit or deny. Additionally some policies may result in decisions of

"don't care" when an action is not critical, or even "don't know" when there is not enough information about the subject, action or target to form a certain decision.

2.7 Access Control

Access control is the limiting of rights or capabilities of a subject to communicate with other subjects, or to use functions or services in a system or network (Department of Defence, 1987). Simple access control definitions are in the form of access control lists, that list subjects authorised for specific access to an object, the specific access is sometimes called an action. This simple list is inflexible in the face of a large number of users or objects, or frequent changes in subject, object or action.

The following section 2.7.1 describes different approaches to access control for computer systems. These different approaches are required to address issues of administration of large numbers of users, or dynamic policies for changing access rights due to properties of the subject, object or context, including environmental changes such as time of day, or access control approaches to cope with changes in objects to be accessed.

2.7.1 Access Control Architecture

To control access to an object there has to be a gateway where two functions are performed. There has to be a decision made on whether access is permitted or not. This is performed at the Policy Decision Point (PDP). On the result of this decision, if it is permitted, then there is the enactment of the action. If it is denied, then the action is blocked. This happens at the Policy Enforcement Point (PEP). The PEP is also responsible for passing the details of the action to the PDP to obtain the resultant decision. This is based on policy framework definitions used in the IETF (Internet Engineering Task Force) (Yavatkar et al., 2000) and illustrated in Figure 2.9.

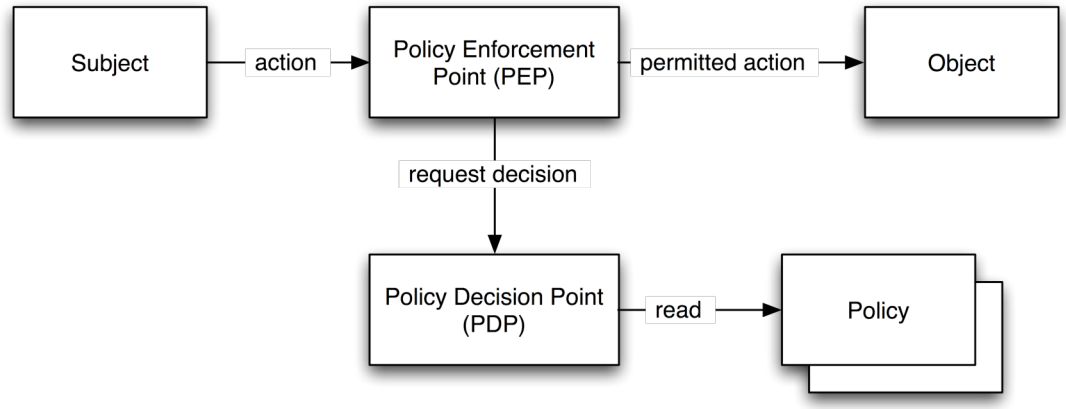

Figure 2.9 Access control architecture

2.7.2 Access Control Solutions

Access control solutions provide methods to specify which users can perform specific actions on specific objects. Fine-grained specification of authorisation is when a policy specifies the actions that each subject can perform on an object, for a policy where the base rule is deny all actions unless specified as permitted. Alternatively, a policy could permit all actions unless denied by constraints in the policy document. A fine-grained policy becomes difficult to maintain as the number of subjects, objects or actions increase or when any of those change frequently (Ferraiolo et al., 2003b), where as access control policies tend to be at a higher level (coarse-grained), specifying types of users and types of resources, such as only personnel staff and managers may view an employees details (King et al., 2001).

Role-based access control (RBAC) (Ferraiolo et al., 2003b) provides a coarse-grained description for access permissions, using an intermediary entity (the role) to separate the access permissions of users to target objects. RBAC allows the subject in an access control rule to be specified as a role. Users are mapped to roles. Roles, in RBAC, can be specified in a hierarchy, where a role in a high hierarchical position adopts the permissions of the roles lower down. In some collaborative processes, roles may not be sufficient in defining the permissions for users. One such case is separation of duty, where a task requires two different people to complete two separate actions. An example would be the authorisation of a purchase order that requires two signatures. The roles required to complete the two actions may be available to one person, or the two actions may be performed by the same role, however the separation of duty rule would prevent one person performing both actions. Another restriction is conflict of interest; this is usually a legal restriction against unfair competition, where a person has gained some knowledge from one action that can be used to gain an unfair advantage in another. An example would be insider dealing in share trading, or a solicitor representing both parties in a divorce.

During a process, a collaboration is formed. This collaboration can involve users, services and resources, that would be defined respectively as roles, service and resource types in a workflow definition. The following background summarises work on access control in the area of users accessing services and resources during a collaborative process.

The NIST standard for RBAC (Ferraiolo et al., 2001) defines a session as the period of interactions during which the user can act in a given role, Figure 2.10. This provides an example of how the user to role mapping can be defined for a process. However, the NIST standard does not explicitly support processes or workflow.

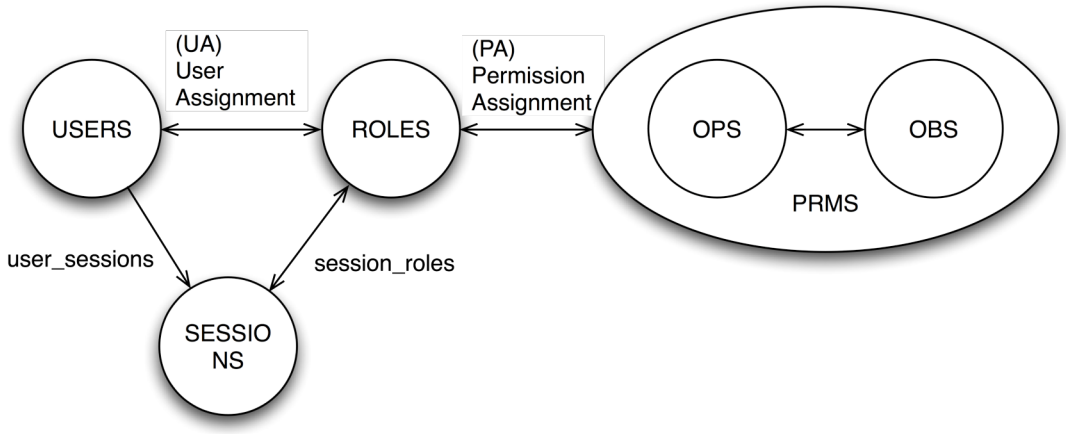

Figure 2.10 Core RBAC, (Ferraiolo et al., 2001)

Chandramouli (2000) describes a method for defining control constraints by analysing business process, identifying roles in the process and assigned target rules for the roles, where the targets are methods of classes. Mendling et al. (2004) shows an approach where RBAC rules are generated from a business process definition in BPEL. As BPEL does not include security aspects, Mendling at al. use the Partner Role definitions in the BPEL script to automate the creation of xoRBAC, XML Role-Based Access Control policies.

Liu and Chen (2004) describe in WS-RBAC (Web Service RBAC) a 'session' that can be used to map users to roles when accessing Web Services across enterprises in a business process described in BPEL. The paper does not however demonstrate a link between the illustrated BPEL script and a WS-RBAC policy, which would be useful in extending the previously mentioned work by Mendling et al..

Other examples relating access control with processes include, the Task-Based Authorisation Controls (TBAC) (Thomas and Sandhu, 1998) which shows how access control can be automated to define fine-grained associations of subject to object. The TBAC model demonstrates how to apply roles (i.e. RBAC) to dynamic policies, but does not define this in the context of a workflow.

Similarly, Team-based Access Control (TMAC) (Thomas, 1997) uses roles in teams for access control in collaborative environments. This presents a link between role-based permissions across object types, and provides fine-grained, identity-based control on individual users to individual object instances. The team here is defined as a project team (Thomas, 1997), with generalised policy for defined users acting in roles for a given context. Georgiadis et al. (2001) extend the TMAC model with contextual information, which ensures team members only access when their context is true. In this example, the context is time and location. This makes for a finer-grained access control than forming teams with user-to-role mapping, but does not show dynamic resources.

In TMAC 2004 (Alotaiby and Chen, 2004), the team definition is an instance of a collaboration of users, the permissions are derived from access control predicates using business requirements. However, the permissions are attached to the context of the collaboration. Therefore, Alotaiby and Chen propose a different approach to provide access control for teams than those that involve workflow enactment.

2.7.3 Access Control Solutions for SOA, Grid Computing

There is notable work from the Web Service and grid community that is aimed at tackling authentication and authorisation problems for distributed SOA. The collaborative workflow described in this thesis requires authentication of users from different organisations. For grid systems, the Globus Toolkit 2.4 and 3.2 (*The Globus Alliance*, 2005) employ X.509 for user identity and additionally use centralised user attribute management systems from the Community Authorisation Service (CAS) (Foster et al., 2003) and the Virtual Organisation Management System (VOMS) (Alfieri et al., 2003). CAS releases fine-grained permissions to users by attaching access assertions to the users grid certificate. The user passes the grid certificate containing their identity and access permissions the requests to the service. The fine-grained access permissions to resources are stated for the duration of the certificate. VOMS, also uses certificate attachments, in this case the attachments specify group membership and other user attributes, such as role. When connecting to the desired service, the VOMS certificate is passed with the request. The service then uses a local access control policy to makes a decision on access permissions based on membership to the group and other user attributes. Both of these schemes issue the user with attributes that are passed to the service, this is a ‘push’ model. Both of these employ the notion of a VO policy, where a user has permissions to use services because they have membership to the VO. This makes it difficult to restrict user permissions to specific collaborative workflows and service instances within the VO, where service instances are dynamically created and shared. Adding and removing members to access a service instance in a workflow would be difficult unless the user requested a new certificate for each service request (also used in previous work by the author (Russell et al., 2004b)).

Policy decision engines such as Akenti (Thompson et al., 2003) and PERMIS (Chadwick and Otenko, 2002) can retrieve policy documents on each service request, and given a dynamic policy these schemes would be useful to interpret a dynamic fine-grained policy attached to a workflow, although neither system would relate access decisions to the workflow context. Typically, Akenti and PERMIS provide policy decisions at the service request, requiring users to be authenticated by every service supplier. The policy engines can use either the push model of sending permissions and user attributes with the request, or the pull model, which locates attributes from remote locations for the user when a request is made.

Park and Hwang (2003) discuss peer-to-peer sharing of Web Service based assets and resources. It uses RBAC to restrict access, with central 'enterprise' polices mapping users to roles and resource owners creating peer level policies defining role access to the resources. The resource model is static, however since the distributed policies are checked on each access it would be possible to link temporal resources to a dynamic local access control policy. The overhead of policy access could be significant. The peer-to-peer model is still restrained by a business model of central administration of users and roles, with users controlling access to assets. In some cases, the access to assets may be specified in the enterprise policy.

Another issue for security in service based computing across organisations is the means of authenticating users. The discussion above concentrates on RBAC to separate user identities from permission specifications, thereby a mechanism is required to attribute users to roles in systems across organisational boundaries. Shibboleth (Cantor, 2004) provides a user authentication system that crosses organisational boundaries and can include attribute assertions, such as role. It would be feasible to integrate Shibboleth into the grid security architecture, to provide role assertions in SAML (OASIS, 2004) for the distributed users, and subsequently be used in workflow access policy decisions.

Agarwal et al. (2002) discuss the issues of authentication across different systems and different administration boundaries. Agarwal et al. propose that X.509 certificates and Kerberos tokens can be used to identify users, extending this to using SAML to transport identity assertions between different systems. The intention is to provide a common interface to link the different methods of identification. It does not, however, address the issues of matching identities between X.509 and Kerberos when a person's name is not identical, for example using first name and surname, when another system provides middle names and a further system only provides initials and surname.

2.8 Summary

In this chapter the background topics were introduced, showing previous work in collaborative working and virtual working by support of computing concentrated on visual sharing and basic communications methods, such as email and video conferencing. To provide control in business processing, workflow management can be used to support collaborations, reducing errors in sharing work by promoting coherence, and recording status of progress. The concepts and technology in Service Oriented Architecture enables business processing, and therefore collaboration, across organisational boundaries.

Service Oriented Architecture, loose coupling, middleware and virtualisation of services leads to virtualisation of the resource tier in Grid Computing. Collaborative use of services and resources across organisational boundaries leads to VOs.

Exposing services and resource to Internet access creates new issues of security about provisions of controlling who can perform what function, when and in what capacity. This requires standardisation of authentication methods to identify users, services and resources. It also requires method to identify attributes of the user and the context of their access. This may include the role of the user, the organisation they belong to, or the identity of the workflow in which they are collaborating.

Different access control solutions have been presented that build on role-base access control (RBAC) (Ferraiolo et al., 2001) and team based access control (TMAC) (Thomas, 1997) to define permissions for roles in collaborative sessions and task based access control (Thomas and Sandhu, 1998) to define permissions for users in workflows. In the grid community, much of the access control work has focused on methods for authentication across organisations, and user attribute based access control in policy decision engines. The user attributes do not include capturing the context of collaborations. Importantly, the collaborations are defined by the combination of the users, the executing workflow (which captures the context) and the services and resources consumed. Grid Computing introduces service instances, which have not been considered in the described access control solutions for collaborative use across organisations.

Chandramouli (2000) defines temporal business associations as temporary assignments of users to roles in collaborative workflows, using rules derived from business process models. Grid Service instances created by Grid Service factories during a collaborative workflow can be considered as temporal business assets in a dynamic context of users, service instances and workflow state. In the next chapter, the DAME project is described, illustrating the motivating scenario that uses Grid Services as temporal assets during collaborative diagnostics of aircraft engines. Chapter 4 describes the requirements for secure collaborative workflow for the DAME scenario. Chapter 5 defines the model for

control of fine-grained access permissions to Grid Service instances executing across organisational boundaries, called Workflow-Team Policy Architecture. Chapter 6 illustrates the implemented secure collaborative workflow management system used in the DAME demonstrator.

Chapter 3

DAME

This chapter presents DAME (Distributed Aircraft Maintenance Environment), the motivational scenario for workflow coordination of services with secure access by teams of collaborating users. DAME is a UK e-Science project to demonstrate the use of distributed services and Grid Computing resources for the support of aircraft engine diagnostics. This chapter outlines the different organisations; the roles they play and the roles the users from the organisations play in the collaborations. The main workflow used by the DAME demonstrations is illustrated showing the business relationships and the collaborative scenario.

3.1 DAME Project Introduction

DAME (Austin and et al., 2001) is an EPSRC-funded project and is supported by industrial partners Rolls-Royce, the aircraft engine manufacturer and Data Systems & Solutions (DS&S), who provide IT support and maintain service contracts for engine leasing. Between the two partners, they manufacture, sell and lease aircraft engines to commercial airlines. DAME includes the academic research partners from the universities of Leeds, Oxford, Sheffield and York.

3.1.1 Project Partners

From the University of Leeds there are two partners in the project consortium. From the Informatics Institute in the School of Computing, Professor Peter Dew leads the expertise on integration of Grid Computing infrastructures. From the Keyworth Institute, School of Mechanical Engineering, Professor Alison McKay provides expertise in data provenance. From the University of Oxford, Professor Lionel Tarrasenko provides the expertise in data acquisition from the aircraft engines and signal processing for initial feature detection. From the University of Sheffield, Professor Peter Fleming provides case based reasoning (CBR) to match new engine cases with past engine diagnoses. Finally, from the University of York, Professor Jim Austin provides fast pattern matching services.

The industrial partners in the project provide the motivating scenario for this research. Rolls-Royce is an aircraft engine manufacturer and supplies gas-turbine aero engines for commercial aircraft. DS&S, a subsidiary of Rolls-Royce, are experts in decision support systems and provide the online environment for airline operators to diagnose aircraft engines. Together they are expert in the supply and support of aircraft engines, with significant experience in online management of engines with human diagnosis support.

The consortium partners form an example of a VO. The universities provide with computing services and integration service to demonstrate the virtual collaborative working in the supply and use of Grid Computing. This VO illustrates how different organisations can contribute to DAME product delivery. This chapter also presents a VO that is formed in the business scenario involving the end user, the Airline.

3.2 DAME Operational Overview

DAME is a collaborative tool to support the leasing of aircraft engine to airline companies. Leasing aircraft engines is a business model from Rolls-Royce, along with a subsidiary Data Systems & Solutions who provides through-life leasing support service. One part of this business model is to be enhanced with an on-wing engine data recorder that stores vibration and performance data during flight. When out of parameter behaviour is detected, the downloaded data is used in collaborative diagnostics workflows that involve people, processing services and data from different organisations. The services, data and results can be commercially sensitive, especially among competing airlines. Therefore, securing the services, data and collaborative processes is essential for the distributed diagnostics environment to be a credible business opportunity.

DAME requires collaborative processing in the task of aircraft engine diagnostics. The DAME environment, illustrated in Figure 3.1, cannot be described as one system, because by its nature it requires independent distributed resources to store, process, manage and communicate the data and results. The data sources (aircraft engines) will be located at airports globally, and so will the maintenance staff dealing with the engines. Data storage and processing resource will also need to be globally distributed to cope with limitations in the transport of large amounts of data in short space of time. Other personnel involved in the diagnosis will also be globally distributed to provide 24-hour support.

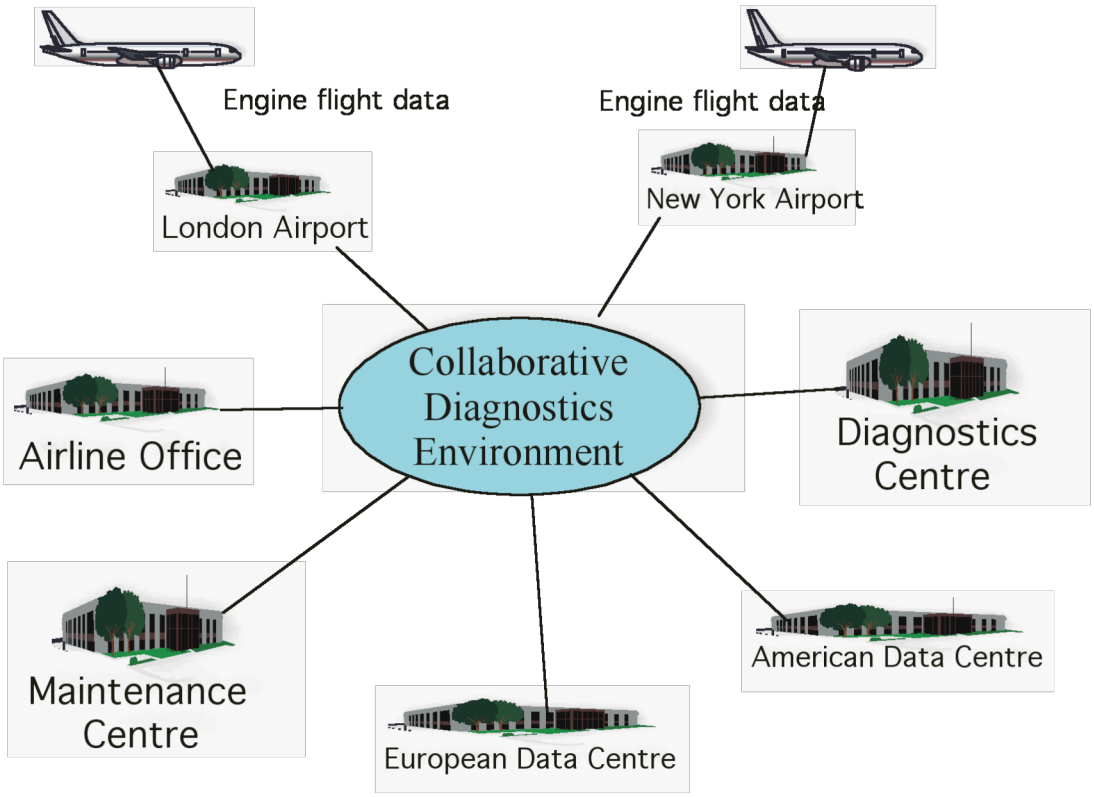

Figure 3.1 DAME operational scenario

3.2.1 DAME Scenario - Aircraft Engine Diagnostics

Rolls-Royce has a requirement for improved engine diagnostics since the change in business model from selling to leasing engines. This current business model used by Rolls-Royce is called Power By The Hour®. This provides the airline operator with a fixed engine maintenance cost over an extended period of time. The business model includes scheduled and unscheduled engine maintenance, replacement parts, exchange of replaceable units and continuous spare parts replacement. Specific programs are tailored according to engine family and operator needs (Rolls-Royce: Services, 2004).

This business model is a move away from product-based supply to a service supply model, by providing the airline operator with 'engine thrust' instead of selling the engine unit outright. The through-life supply is supported by DS&S. To best support the operator, the service supplier needs to understand the operating conditions of the engine and part of the operator's business. Currently DS&S have significant experience in supporting aircraft engines with predictive maintenance cycles using knowledge of flights, typical engine and parts lifecycles and the types of supporting components that operators use, such as the type of oil.

Developed with the help of Oxford University, Rolls-Royce have been using a vibration monitoring system in ground-based testing of aircraft engines. The engines are exercised in the test-bed to test dependability in the design, and for pass-off testing production units before release. This concept has been redesigned as an on-wing embedded system to monitor and record vibration and performance parameters of the aircraft engines whilst in flight.

To improve diagnostics and maintenance scheduling a ground-based system will analyse recorded data to monitor engine behaviour. This is facilitated by downloading the data on landing, then processing it with distributed services, to provide an on-demand diagnosis of the condition of the engine. Automated workflow services execute the diagnostics processing and by employing QoS (Quality of Service) requirements, results can be returned within the turn around time of the aircraft. This will be used to detect wear of components, foreign object damage (e.g. ingestion of birds) and other out of parameter conditions.

Another aim is to provide information on the condition of engines to the maintenance team at the airport for predictive maintenance. The objectives in providing more information and diagnostics through the engine lifecycle are to reduce cost, improve safety, increase availability and improved scheduled maintenance models.

To achieve an improved level of diagnostics the analysis processes historic engine records across all the airline operators. During the diagnostics process that is started on landing, case-based reasoning is used to match with the historic records and provide a likely diagnosis. Historic records also support improved predictive maintenance by performing pattern matching to identify trends of product lifecycles.

The DAME scenario requires enhancement in the IT support systems due to the scale and distribution of the problem. The data sources and users are distributed. Every global airport is a potential target for access to the diagnosis system. Each flight generates large amounts of data, each engine will typically generate >35MB per hour and there can be up to six engines per flight. As already stated the diagnosis process uses the historic records of past flights, this requires access to large amounts of stored data for processing. The volume of data will continually increase, and locations for storage are likely to be distributed in several global data centres, hosted by different organisations, such as Rolls-Royce, DS&S and the Airlines. Along with data hosted in different organisational domains, there is potential to use data processing services for the organisations already identified and other organisations that are specialists in certain processing algorithms, such as fast pattern matching.

The diagnosis processing is required to produce an advisory result on engine condition within a short period of time. It needs to process the information and return a result within the ground staff's available time in the turn-around time of the aircraft. To achieve this, processing resources will need to be dynamically available to respond to varying patterns of demand.

The access to data storage and processing services requires Internet exposure of the supporting IT. This creates a security risk by opening access to commercially sensitive data and processing algorithms. Sensitivity of the data includes the engine and performance data and the derived results from analysis of the raw vibration data. This data can reveal the operating conditions of the airline. This has commercial sensitivity, for example other airline could use knowledge operating conditions to gain competitive advantage, or to damage reputation by releasing information to the journalistic press or business analysts. Commercial sensitivity of the processing services would be to prevent somebody replicating the service, then using it for there own use without paying the service provider, or reselling it as their own.

For the DAME scenario to be realised, the DAME e-Science project is used to investigate the appropriateness of Grid Computing in addressing the problems listed above. Grid Computing supports the VO formed by the collaborating organisations to improve diagnosis and predictive maintenance. Among its properties is high-speed networking, large scale distributed storage, dynamic resourcing (of compute power, storage and network), supported by strong authentication and authorisation mechanisms for distributed access across organisations.

The users in DAME require access to the distributed diagnostics services and data. It was identified during DAME project meetings and requirements analysis that one possible solution is by using a web-based portal and workflow management system. A web-based portal is a web site in the World Wide Web that provides personalised access capabilities (typically with secure access), with pages served containing tailored views and controls to back-end systems, such as databases and workflow management systems.

The next section 3.3, presents how the distributed partners in the DAME business case, shown in Figure 3.1, form a VO, collaborating on the engine diagnosis process

3.3 DAME Business Model

The DAME business model is the VO that involves the people, processes and resources required to support the diagnosis business processes. The partners in the DAME VO are captured in a UML business class diagram in Figure 3.2, using the Rational Rose™ business modelling profile (Johnston, 2004). The diagram shows the structural relationship of the

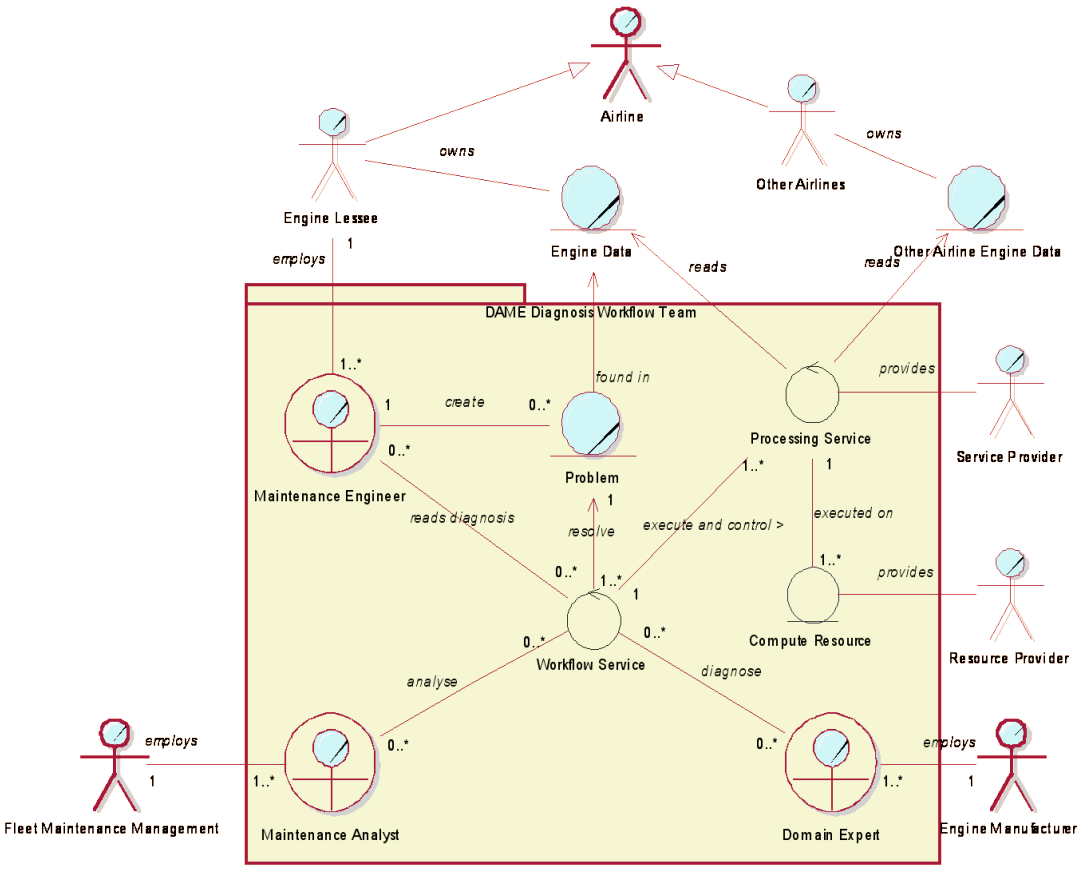

Figure 3.2 DAME Virtual Organisation, showing the diagnosis workflow-team

business partners during the collaboration in the engine diagnosis scenario. The organisations in the VO are shown on the outside of the package containing diagnosis team. The collaborating diagnosis team includes people, processes and compute resources.

The main partners of DAME are shown in Figure 3.2 on the outside of the package DAME Diagnosis Workflow Team and are represented as *Business Actors* are:

- Airline (shown at the top), are the operator of the aircraft engines. Shown, as specialisations are the Engine Lessee and Other Airline. The Engine Lessee is the customer of the diagnostics process, operating the engine and employing the Maintenance Engineers, who start the process when the aircraft lands. The Other Airlines involved in the business model illustrate the pool of historic engine data that is owned by an airline, but used in the diagnosis of other airline's engines. The sharing of data improves the quality of the result in the diagnosis process, but the raw data owned by one airline cannot be accessed by another airline.
- Fleet Maintenance Management (bottom left), in this case represents DS&S. They support the aircraft engine operators (Airline) by providing the diagnostics support and employ Maintenance Analysts. They manage the integration system, executing a portal

and workflow manager to execute diagnostics workflows, consuming processing services.

- Engine Manufacturer (bottom right), in this case represents Rolls-Royce. The Engine Manufacturer supplies engines by service contract, which is managed by the Fleet Maintenance Management organisation. They employ the Engine Designer, who can be consulted during the diagnosis process.

Additionally, the DAME scenario includes two more types of organisation:

- Service Provider (middle right), which represents organisations that specialise in processing algorithms, such as fast-pattern matching. The processing algorithms are offered as services to be used in the diagnosis workflows. The deployment of services may be at a fixed location, or the services may be dynamically deployable on grid resources from a Resource Provider. The Service Provider can be an external organisation, or the organisation may be one of the three types already mentioned.
- Resource Provider (lower middle right) represents organisations that provide the grid resources of data storage and compute processing power, as a service. The key provision for DAME is short-term available compute power, where processing tasks can be dynamically deployed to compute processing resources to meet the demands of workflows.

In the centre of Figure 3.2 is the DAME Diagnosis Workflow Team package. Within the package are the *Business Workers* that represent the following human roles in the diagnostics:

- Maintenance Engineer (ME): carries out inspection, diagnosis and maintenance of aircraft engines; Employed by the Airline and is based at the airport;
- Maintenance Analyst (MA): provides technical advice and coordinates analysis; Employed by the Fleet Maintenance Management organisation and is based at the diagnostics support centre where the airline's aircraft maintenance contracts are managed;
- Domain Expert (DE): acts as a repository of knowledge and will provide expert diagnostic advice on unidentified features; Employed by the Aircraft Engine Manufacturer, and is based at the engine manufacturer's design centre and is an experienced aircraft engine designer.

The other elements inside the Diagnosis Workflow Team package are created during the diagnostics process. When an aircraft lands, the Engine Data is downloaded to a data store. The ME creates the *Business Entity* Problem from the new Engine Data and starts the automatic diagnosis process, shown as Workflow Service. The Workflow Service executes

Processing Services to perform analysis of the Engine Data, which also accesses Other Airline Engine Data. The Processing Services are executed on one or more Compute Resources.

The Diagnosis Workflow Team is dynamically created as new team members join to collaborate on solving the Problem and execute Processing Services, consuming Compute Resources. The multiplicity in Figure 3.2 shows that the Diagnosis Workflow Team can contain many people acting in the roles using many processing services and compute resources. The dynamic model showing how the team evolves over time, consuming services is provided in the next section, 3.4.

3.4 DAME Workflow

The DAME workflow is the diagnostics process that is created whenever an aircraft lands and the vibration and performance data is downloaded. This diagnostics process has been captured from interviews with the business partners and documented in the Use Case Analysis (Fletcher, 2002). Figure 3.3 contains a simplified activity diagram showing the diagnostics process, interpreted from the results of the use case analysis and accepted by consultation with the business partners. The use of activity diagrams for this work is supported by the adoption of UML throughout the DAME project to capture the system structural and behavioural architecture, and has been shown to be useful in specification and communication of workflows (Eshuis and Wieringa, 2002).

The diagnostics process in Figure 3.3 uses swim lanes to denote the activities carried out by each role. The roles are the *Business Workers* illustrated in the Workflow Diagnostics Team from Figure 3.2.

The process starts when new data is downloaded after an aircraft has landed. The process splits into three parallel paths. On the first path, the ME performs a visual inspection of the engine's condition. Concurrently, the diagnosis result of the on-wing system, named Quote Diagnosis, is downloaded. This result is from preliminary signal processing on the vibration data, which is executed during flight and indicates any out of parameter behaviour. The third path is the automated workflow, called Brief Diagnosis/Prognosis and labelled as WF1 in Figure 3.3. This is a workflow of processing services and is explained later in this section. The result of WF1 is joined with the other results and processed in Check Diagnosis. Depending on the status there are three outcomes from Check Diagnosis. The first is when the result is Clear allowing the ME to Release Engine for its next operation. The other two outcomes from Check Diagnosis are categorised as Feature Detected. The Feature Detected is out of parameter behaviour, which can be categorised as either Known or Unknown. A Known result is when WF1 produces a

conclusive diagnosis of the flagged behaviour. From a Known result the required Maintenance Procedure can be executed by the ME and then the engine released. If the Feature Detected is Unknown then the process is escalated to the MA.

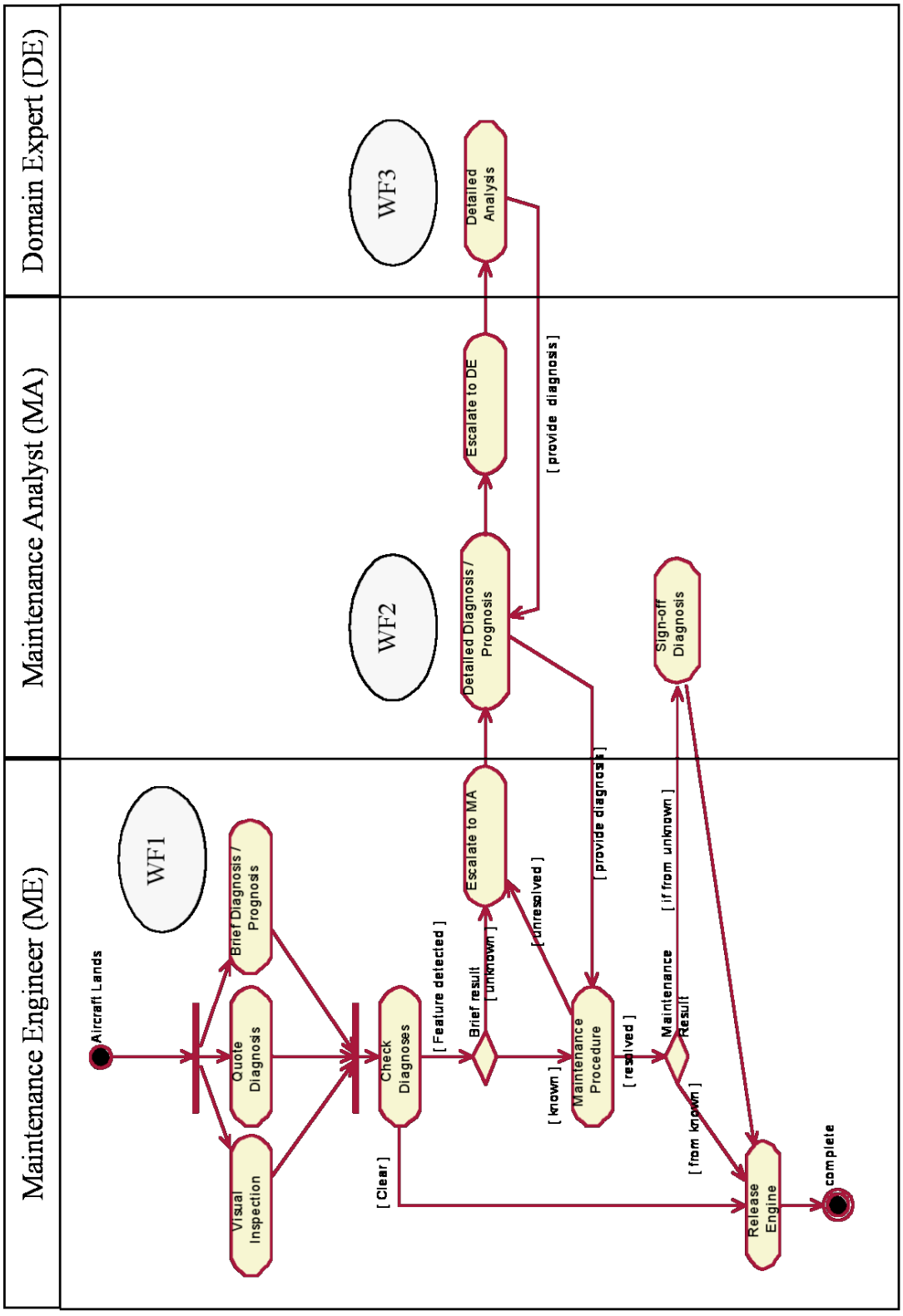

Figure 3.3 Diagnostics business process

The escalation task is a request for assistance from more specialised personnel. This is the mechanism to build the team and control team membership. The ME escalates the problem to the MA when the initial automated process (WF1) cannot recognise out of parameter vibration signals. The ME initiates the escalation and then automated resource allocation selects the MA to join the process. The MA investigates the problem with a range of tools (WF2) and, if needed, escalates the problem to a DE. Escalation is the same as before, but allocating a DE to the process. The DE uses further investigation tools (WF3). At escalation the user initiating the task provides annotation in support of the data, results and processing services being used in that instance of a process. The escalation task is part of defining the members of a collaboration, by adding users to the team. The release of the diagnosis process removes users from the team.

The activities shown in Figure 3.3 are high-level business activities. There are three high-level activities, indicated as WF1, WF2 and WF3, which contain workflows in themselves. The first of these, WF1, is shown in Figure 3.4. It is an automated workflow triggered by the arrival of new engine data from the on-wing system. WF1 uses a chain of data processing tools to produce a most likely prognosis. These processing tools are part of on-going development in diagnosis and in order to support the inclusion of new versions and different implementations each tool is implemented as a service component.

The first task in the WF1 sequence is signal processing on the engine vibration data to extract fragments of the signal that are performing outside the boundaries of normal operation. Of which, some types of behaviour are identified in this process, such as foreign object damage or a bearing failure. All signal fragments are sent to the second process, which looks for matches in historic engine data. The third process takes the matching records then uses case-based reasoning to identify the likely cause of the signal fragment.

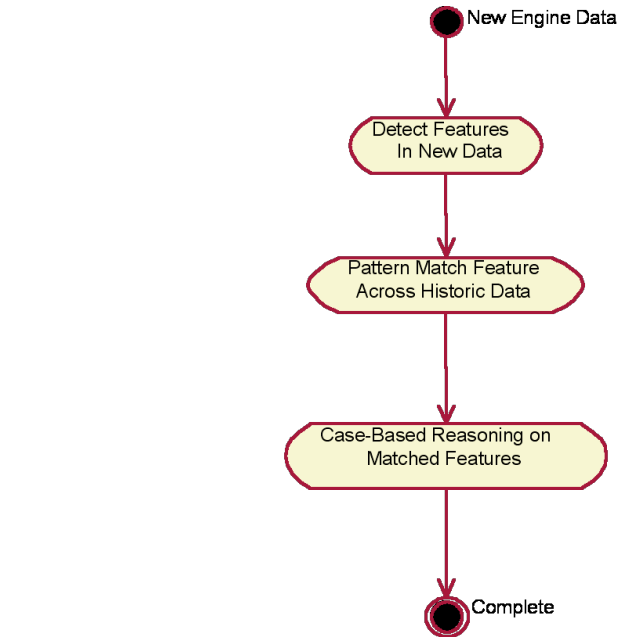

Figure 3.4 WF1 - Brief Diagnosis / Prognosis

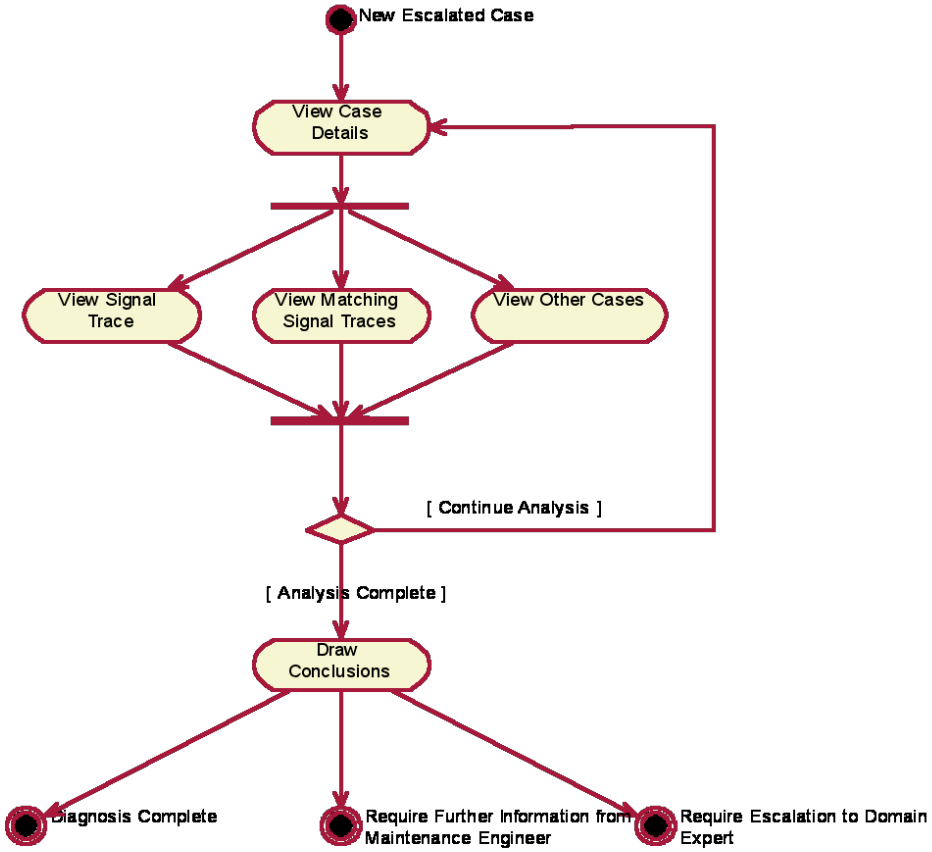

Figure 3.5 WF2 - Detailed Diagnosis / Prognosis

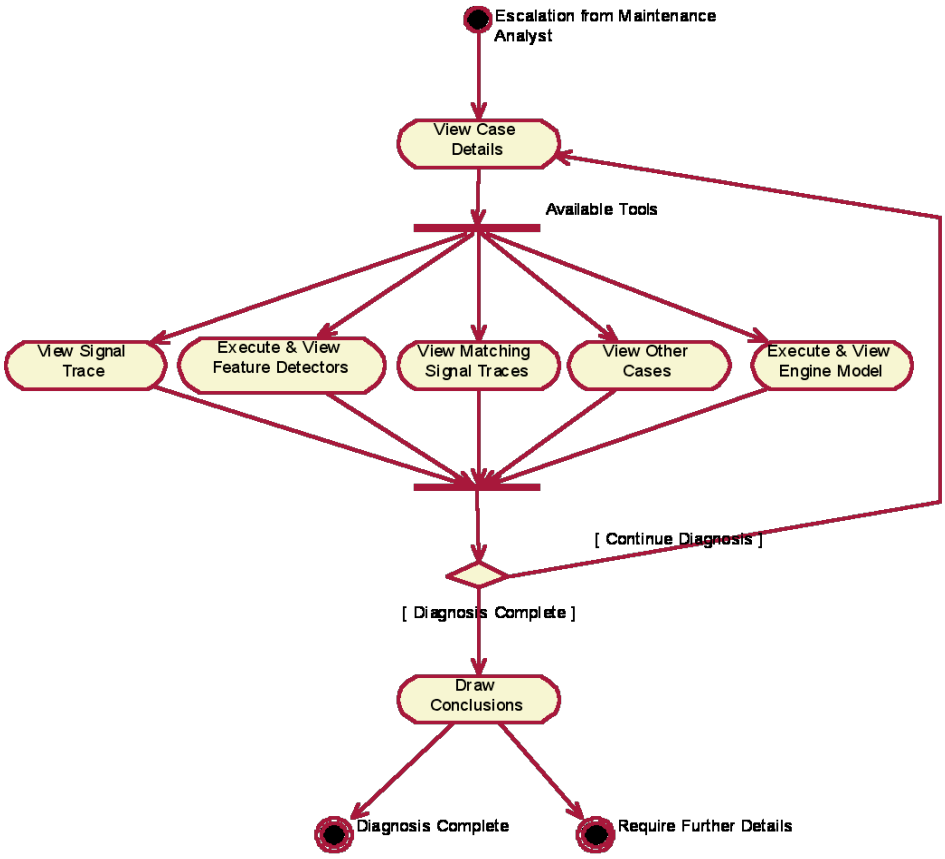

Figure 3.6 WF3 - Detailed Analysis

The workflows WF2 and WF3 are shown in Figure 3.5 and Figure 3.6 respectively. These examples show that the MA and DE roles have more flexibility by choosing which diagnostics tools is used. WF2 and WF3 represent a combination of static and ad-hoc workflows. The tools are defined in the static definition, but the order of execution is ad-hoc and determined by the user at runtime. The static definition means that permissions can be assigned by role to the workflow task. The ad-hoc elements of the workflow allow the user to choose which services are required to fulfil the goal of the workflow.

The MA, who has access to tools that aid in visualisation of the engine data, pattern matching and further search capability across engine case history, executes the WF2 workflow. The DE executes the WF3 workflow, who has access to specific analysis tools produced by the engine manufacturer. This includes more detailed information about engine behaviour including engine simulation models.

3.5 DAME Summary

The chapter has provided an introduction to the DAME UK e-Science project, which has been used as the motivational scenario for this research. The next chapter uses the DAME scenario to outline the requirements to achieve the business implementation and describes the security requirements for collaborative workflow across a globally distributed VO.

Chapter 4

Requirements for Secure Collaborative Workflow

In this chapter the DAME scenario is used to extract system requirements to achieve a globally distributed system for collaborative diagnostics.

4.1 System Requirements

The DAME scenario requires enhanced IT support due to the scale and distribution of the task. The user population can be summarised as:

- Across a large number of sites (at least every international airport, Rolls-Royce and DS&S);
- Across many organisations (Rolls-Royce, DS&S and many Airlines);
- Dynamic; changes will occur to the population of DAME users within organisations, independent of the DAME environment.

The processing is performed by workflows coordinating the different tasks. The tasks are implemented as services in a SOA. Loose coupling mechanisms in SOA means that changes to service deployments can happen without requiring changes to workflows. In particular, abstraction by service type means that workflow definitions can be independent of service implementation changes. Thus, SOA allows the services defined in a workflow to change implementation. The changes could be using services from different providers; or, deploying the services on different platforms, which includes dynamic use of available grid resources; or, evolving the services with extended functionality.

In current SOA implementations, namely Web Services, most services are stateless and interactions are simply transitory. In stateless Web Service implementations, a service response is typically no longer than 30 seconds to complete without causing a timeout. Long-term services in DAME are required to execute for longer periods, this can be addressed using stateful service implementations, such as those offered in Grid Services. These long-term services will also be started by one user and subsequently accessed by

other users, collaborating on steering the process and viewing the results. Examples of the DAME long-term services are:

- Feature Detection from WF1. The test-bed data supplied by Rolls-Royce takes more than 5 minutes to process, a long flight taking as much as 1 hour.
- Engine simulation model and visualisation services. Executed by the MA and DE, these services can be run throughout the workflow (once launched) providing a common view on engine condition and shared between users authorised to see the output.

The data sources for DAME are large scale. At the start of the DAME project, Roll-Royce had over 50 GBytes of engine vibration data from the test-bed system. With each flight, the capacity required for engine data will increase. This also increases the processing requirements when searching across the historic records of data. The processing requirements are dynamic in that diagnostics results are required for maintenance to be completed within the turn-around time of the aircraft.

The project partners are interested in addressing the scalability issues of storage and processing by using Grid Computing. In the future, they hope to use commodity computing resources, as a commercial product, which can be used to cope with dynamic usage requirements.

Therefore, the DAME project is using Grid Computing to address dynamic processing and storage requirements, along with issues of distribution of data, processing, users and organisations.

4.2 Security Requirements

The importance of security for DAME has been noted in section 3.2. This is characterised by the nature of a global distributed system, the exposure of services and data to Internet access. This creates a security risk of access to commercially sensitive data and processing algorithms and can reveal the operating conditions of the organisations. This has commercial sensitivity, for example other airline could use knowledge operating conditions to gain competitive advantage, or to damage reputation by releasing information to the journalistic press or business analysts.

There are many areas of security and this research concentrates on access control to the assets used in DAME. An asset is defined in the DAME Dependability and Security Study, by Fletcher et al. (2004a), as:

> *Asset is a resource of value to an organisation. Assets may include hardware, software, data, people and soft assets such as reputation or intellectual property.*

The types of assets analysed in DAME were data and services used in the diagnostics workflow. The requirement to protect the assets has been classified in terms of system security and dependability goals. These are defined as goals, identified by the stakeholders, where success is necessary for business objectives. Failure to meet a goal will have an adverse impact on the business. The stakeholders are someone or something that has a vested interest in the behaviour of the use cases (Fletcher, 2002). The main system security and dependability goals from the DAME Study are listed below:

- To maintain the Confidentiality of Detailed Engine Design and Performance Data;
- To maintain the Confidentiality of Operational Data;
- To ensure that any Diagnostic advice provided by the system is Reliable;
- To record the provenance of diagnostic decisions and identify individuals' actions in the diagnostic process;
- To provide predictable availability;
- To protect the confidentiality of technical industrial property used in the system's implementation.

The possible threats to the assets are concerned with confidentiality, integrity and availability. The following table summarises these concerns for data and services, from (Fletcher et al., 2004a):

Table 4-1 Notes on possible threats (concerns) to data and service assets

	Data Asset	**Service Asset**
Confidentiality	Lack of individual data confidentiality may include: A. An unauthorised party observing the actual information contained within the data. B. An unauthorised party observes the existence of: the transmission or storage of a particular data asset. Depending on the data asset the impact of this may be that it may divulge: • Proprietary information. • Business process information. • Operational and maintenance data e.g. fault incidence, deduction of general fault incidence from data volumes, etc.	Lack of individual service confidentiality may include: A. Ability of an unauthorised party to access the internal algorithms. B. Unauthorised execution allowing, access to other resources, "chosen data" attacks, etc. Depending on the service asset the impact of this may be that it may divulge: • Proprietary information. • Business process information. • Operational data e.g. fault incidence, deduction of general fault incidence through execution occurrences, etc.
Integrity	Lack of individual service integrity may include: A. The loss or corruption of data. B. Inappropriate modification of data.	Lack of individual service integrity may include: A. Uncontrolled modification of software. B. Critical failure in software.
Availability	Lack of individual data availability may include: A. Delay in the availability of data. B. Loss of data. Depending on the data asset the impact of this may be: • Lack of availability of the automatic workflow WF1 diagnosis. • Impede the ability of the Maintenance Analyst (MA) and Domain Expert (DE) to make a diagnosis. • Lack of availability of other information provided by DAME e.g. information provided for the repair database as part of the completion of the "repair loop".	Lack of individual service availability may include: A. Any regular or prolonged loss. B. Unauthorised access or attempted access leading to denial of service. C. Unauthorised execution leading to denial of service. D. Loss of the availability of a communications path. E. Loss of the availability of a processing node. Depending on the service asset the impact of this may be: • Lack of availability of the automatic workflow WF1. • Lack of availability of other information provided by DAME e.g. information provided for the repair database as part of the completion of the "repair loop".

In the DAME Study there were 33 data assets and 20 service assets identified. The main assets to protect in are:

- Engine vibration data, this is raw data downloaded from the aircraft into an available data store, if compromised there is little threat unless it can be attributed by its metadata;
- Engine records; when engine data is downloaded, a database record is created containing flight details, such as airline, aircraft number, airline, flight number and times, this provides provenance for the vibration data;
- Diagnosis results: this information is very sensitive and could reveal proprietary operating conditions.
- Partial results: at each stage of the workflows, these are the outputs of processing services and annotations from the uses. The DAME Study (Fletcher et al., 2004a) identifies each of these results, and puts a quantitative measure on the impact of unauthorised access. Examples are:
  - Annotations from the escalation process could convey proprietary operational information and may have a medium impact to confidentiality;
  - Engine Simulation Result could divulge proprietary engine information and may have a high impact for new engine types, but medium for older engine type.
- Workflow definitions capture knowledge of the diagnosis process and have commercial value to the operators of the service integration (such as DS&S).
- Workflow records that capture details of the diagnosis case can reveal an airline's operational information, and tampering with live records may inhibit flights in the short-term, or damage business reputation in the long-term.
- Long running service instances. The service instances require access during the workflows by collaborating users. Confidentiality is required for:
  - User access, because this can reveal the operators details;
  - The initial parameters a service is started with, and steering parameters sent during execution, which could also reveal operator and flight details;
  - Results from the process (see Partial results above).
- The proprietary algorithms used in services may have commercial value. Competing service providers may want to replicate part or all of the algorithms, or potential customers may try to gain access without paying.

- The algorithms may contain proprietary knowledge, such as the model to simulate an engine's performance, which would be of interest to competing engine manufacturers.

4.3 Requirements for Secure Collaborative Workflow

The DAME scenario is a collaborative process between the identified roles of Maintenance Engineer, Maintenance Analyst and Domain Expert. This occurs when an aircraft lands at an airport, data from the engines are downloaded and workflow WF1 is automatically started. According to the workflow model (illustrated in Figure 3.3), workflow WF1 is assumed to start under the identity of the ME that is responsible for connecting the engine and downloading the data. If WF1 produces an unclear diagnosis, then escalation causes the other participants to join the diagnostics process. The MA uses tools WF2 and the DE uses the tools WF3. Both of the people in these roles can access data and service results that were executed by the ME.

It has been identified that Grid Computing will be used to address the processing and storage needs, and that a Workflow Management System is required to control and manage workflows of distributed processing and data access. These technologies support an open model of collaboration between distributed users, distributed data sources, distributed service providers extending to distributed grid resource providers on a global scale. The previous section, 4.2, identified the need to protect access to the components of the VO. Therefore, access control must prevent undesirable access and permit the business collaborations.

For each execution of the business process there exists a workflow instance. The workflow instance captures the context of the collaboration. It contains the state of the workflow, such as the data associated with the workflow and the identities of the Grid Service instances. In the example scenario, long-running stateful services can be shared between users. The users act in pre-defined roles in the workflow instances. The workflow instances controls the use of Grid Service instances. Therefore, the context of the workflow instance also describes the collaborative team.

The team has the properties:

- Workflow instance executing the sequence from the workflow definition;
- Users playing roles from the workflow definition;
- Users sharing access to Grid Service instances;
- Grid Service instances executing the service types from the workflow definition;
- Grid Service instances executing commercially sensitive algorithms;

- Grid Service instances accessing commercially sensitive data;
- Grid Service instances returning either anonymised or commercially sensitive results.

The link between role-base workflow and role-based access control has been explored in previous work. The role-based definitions allow the user base to change more often than the business requirements that define the business process and access permissions. However, the DAME VO scenario highlights two specific issues not addressed in the previous work (Chandramouli, 2000, Liu and Chen, 2004, Kang et al., 2001, Lepro, 2003, Koshutanski and Massacci, 2003). The first is that role-based access control is not sufficient for processes that involve users and services across organisations, since services and data can reveal proprietary information about an organisation operating conditions. The second is that the previous work does not show how to handle collaborative access to Grid Service instances with automated dynamic fine-grained access control.

4.4 Problem Summary

The following table outlines the requirements for the DAME architecture to deliver secure collaborative workflows.

Table 4-2 DAME problem summary

The Problem of:
Providing and restricting access to commercially sensitive data and services in distributed systems where system resources, services and users belong to different organisations. Users collaborate in task based problem solving, from geographically distributed locations.
Affects:
Owners of data and authors of services that allow access to parties on a need to know basis avoiding conflict of interest and exposing an organisations operating data to a competitor.
The impact of which is:
• No sensitive data should be read or modified by unauthorised users or services • No service should be instantiated by unauthorised users or services • No service instance should be accessed by unauthorised users or services • The workflow of the collaboration should not be restricted in access to resources/services to resolve problems, in its normal flow of execution • Access to data should be restricted by conflict of interest mechanisms • For information that is sensitive to security level, the access control policy should restrict read and write access • Access policy mechanisms should not impede the workflow time frame when

authorising security assertions.
A successful system would be:
A system that supports single sign-on by users and permits access to all resources/services allowed without detrimental time or execution overheads. It will react to changes in collaboration membership before unauthorised access can occur. It will support the workflow dynamics of a problem solving team using identities, roles and policies created to protect multiple organisations and service instances. The system must be managed from multiple locations, with each organisation able to impose policies on aspects under their own interest. There should be visibility of available services and resources and the impact of policies imposed. Access to services should be logged under fine-grained detail, with security exceptions clearly identified. There should be a means of recording access to services in a manner that is non-repudiable to support auditing. This includes intruder detection whether access has been gained or not, and auditing for economic reasons. The workflow instance should manage its own access control policy to itself and the service instances currently being used. This may or may not be triggered by user intervention.

4.5 Summary

The DAME scenario to support aircraft engine maintenance requires a highly distributed solution, which raises issues of managing user access to data and resources. The diagnostic team model requires the use of data from all participating organisations, some of which compete on a commercial basis. The operating data from that recorded on the engines is commercially sensitive and requires protection. Supporting the diagnosis are algorithms that process the data, which are also commercially sensitive to the service provider. The actions by all parties in the collaboration need to be protected against the release of sensitive data or access to protected algorithms.

Using the scenario presented in Chapter 3, this chapter presented the list of requirements for securing the collaboration of users from different organisations, consuming services and grid resources from suppliers in further different organisations. The next chapter presents the generalised solution to secure collaborative workflow. In Chapter 6, concepts in the generalised model are tested in the secure collaborative workflow implementation, part of the DAME project demonstrator. The evaluation of the implementation and model are presented in Chapter 7.

Chapter 5

Workflow-Team Policy Architecture

This chapter presents a generalised architectural model for automated control of dynamic access permissions in collaborative use of Grid Service instances in workflows across organisations.

The requirements from previous chapters are analysed and compared with the published work to produce the access control model. The proposed solution is the Workflow-Team Policy Architecture, which address the issues raised in the analysis, and fulfils the requirements for secure collaborative Grid Services across organisations. The architectural structure and components are described in section 5.3. To express the scope of the design, the assumptions about the commercial policy, technology and the case study are listed in section 5.4. At the end of the chapter there is an example of how the Workflow-Team Policy Architecture would be implemented in a portal-based, collaborative problem-solving environment.

5.1 Analysis of the Workflow Security in the DAME Demonstrator

The DAME demonstrator was successful in illustrating how Grid Services can be combined for collaborative use of Grid Service instances in a diagnostics environment. The evaluation showed that there are still issues for controlling access and controlling business processing for distributed services. This analysis uses Chapter 3 and Chapter 4 and compares the requirements with published work. Also used in the analysis are the DAME Security and Dependability Study (Fletcher et al., 2004b), feedback and discussions from the DAME project meetings and during the Workshop on Grid Security Practice and Experience, July 2004, Oxford, UK, at which the author presented part of this work (Russell et al., 2004b).

The analysis of the data resulted in some architectural decisions for a scheme to manage the authorisation to use collaborative services instances in a grid environment across organisations. The Workflow-Team Policy Architecture solution is presented in this chapter in section 5.3.

Key issues from the results of the demonstrator are discussed in the following sub-sections:

- Business organisational issues relating to the demonstrator's appropriateness to the DAME business model;
- Business process definition issues relating to how processes are defined and mapped to access control, and how process and access control are coupled in the demonstrator;
- Access control issues, discussing other approaches and complementary solutions; and,
- Architectural issues, regarding scalable solutions for DAME.

5.1.1 Business Organisational Issues

The DAME business model has shown the need for workflow to control the user of distributed services. It requires a portal to provide a location independent collaborative environment with multiple users accessing common services that are required to be Grid Services for reasons of scale and availability. Users interacting with the workflows controls access to the services. Therefore, the point of collaboration is the workflow. This is because:

- It is the point where all users access services for launching, service control and retrieving results;
- It is the point where all data passes to and from services.

The collaborative business process creates service instances that require collaborative access by specific users. However, the business process definition cannot specify specific users and specific service instances ahead of use. For manageability, the business process is defined by role and service type. Therefore, on instantiation of a workflow fine-grained access control policies are required to state which users can access which service instances. This policy needs to be dynamically modified to follow the actions in the workflow, to add access rules for the created service instances and to provide changing access for users.

The model used for the demonstrator is based on a VO of consumers and suppliers of services. The business model reflects the DAME scenario shown in Figure 3.2, from section 3.3. The DAME demonstrator has illustrated a feasible VO by modelling the different roles and organisations using the different partners of the project.

The different partners in the business model relate to the SOA tier shown in Figure 2.1. This shows how a VO relates to a traditional supply chain, which is reflected in the DAME VO model. The supply chain is illustrated in Figure 5.1. The labels are generalised, but also indicate the positions DS&S and Rolls-Royce occupy in the DAME VO. External organisations or any of the VO member organisations could provide the Grid Services and

Grid Resources. Whoever provides the services or resources is subject to contractual responsibilities between supplier and consumer.

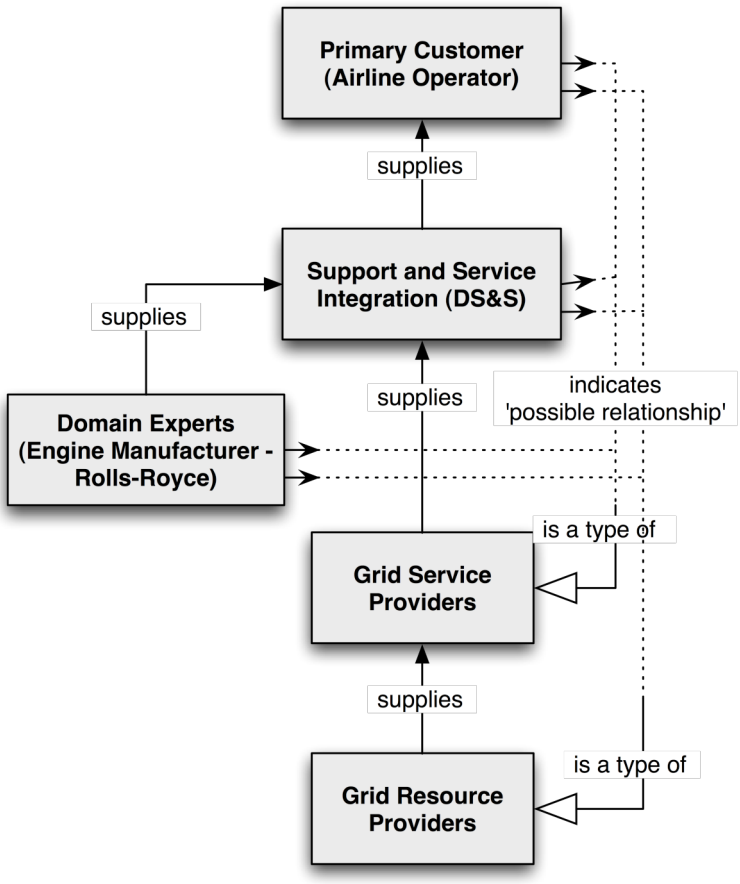

Figure 5.1 Supply chain, related to DAME VO

In a traditional vertical supply model, the contractual responsibility for provision is between the direct consumer and provider. For example, if a service provider dynamically uses resources from the grid resource provider to meet performance requirements of the service consumer, then dynamic selection and payment of the grid resource is the responsibility of the service provider, not the service consumer. Therefore, grid resource and service consumer would not be in direct contact. This is required to maintain business relationships when there are many customers and many suppliers.

In Figure 5.1, the Support and Service Integration organisation executes the workflow management system. The workflows consume the processing services from the Grid Service Providers. The integration of services decouples the consumer of the support (Primary Customer) from the organisations that provide processing services and grid resources. In the example scenario, the Service Integration also decouples the Primary Customer from the Domain Expert.

Previous studies have suggested that a conceptual VO policy can be used to define access to services and resources in a VO (Russell et al., 2004b, Bertino et al., 2004). The conceptual VO policy, from (Russell et al., 2004b), defines vertically the authorisation of all types of components (user roles, services and data) from the top-level consumer to the lowest level provider.

However, the DAME example illustrates the contractual relationships in a vertical supply chain. The contract would be enforced by policies between the consumer and provider, and does not require a policy managed across the whole supply chain. To produce a single policy for user types and service types for the VO would require collaboration on its contents, possibly lengthy negotiations, and would be less flexible for the inclusion of new service types and new service providers. Most importantly, it reduces the flexibility in defining access control for workflows.

Another issue in the VO policy is privacy. The workflow is used as the context for a collaboration and it has been stated that the roles in role-based workflow and RBAC can be linked by the business requirements. Therefore, if the VO policy is defined for the workflow across the VO, then partial knowledge about the workflow can be gained by interpreting the policy. The workflow definitions are assets that belong to the Support and Service Integration organisation. Therefore, using a VO policy could reveal proprietary information.

Before the access control policy can be used to determine authorisation, the users, services and resources need to be authenticated. For a VO policy to make decisions based on each user, they would need to understand the identities of every user in the VO. The example highlights problems with identifying users at each resource. The example has a large user base. It can be assumed that a large user base will be dynamic due to changes in staff.

Another issue with user identities is the workflow management system. For example, in the DAME workflow the process is started by the ME results and access to the services can then be escalated to the MA and DE. The workflow process permits this. However, if the workflow did not understand access permissions, then the results from a service accessed on behalf of one user could be passed to another user. This would occur, for example, if the MA retrieved results a service and escalated the process to the DE, then the DE viewed the results via the workflow, and not accessing the service directly. The access control policy at the service would have restricted the action by the DE, however the MA had already performed the permitted action. This is because the workflow is the point of collaborative access to the service instances.

5.1.2 Business Process Definition Issues

The aim of workflow systems is to provide a flexible approach to defining and managing business processes. This is achieved by defining the process at an abstract level, by describing the workflow in terms of the tasks performed, rather than defining it by the implemented resources. This creates loose coupling between the definition and implementation, which is echoed in the philosophy behind SOA, as discussed in section 2.5.1. SOA promotes decoupling of components by abstract description. Referring back to the Introduction, Figure 1.1 illustrates the high level creation of workflow from business requirements. The abstraction method of defining by types (service, method, data) that is used in SOA is required to form workflow definitions that match business process requirements. The abstraction removes them from rapidly changing implementation. Such that, well formed business processes remain unchanged whilst the implementation of services can change without redefining workflow descriptions. The Workflow-Team Policy Architecture illustrates this in section 5.2, along with exploiting the link between role-based process definition and role-based access permissions.

5.1.3 Access Control Issues

Analysis of the business requirements from the DAME scenario has been used to define access control rules. The VO model illustrated in Figure 3.2 is accurate for the limited requirements of the demonstration scenario, but does not show enough detail to determine access control across the organisations. This statement is vindicated in the evaluation in Chapter 7 and in Fletcher et al. (2004b). The airlines share sensitive data within the DAME system and this needs to be protected. Therefore, the organisation is an important attribute in the control of access to the workflow instance and its service instances. This extends RBAC (Sandhu et al., 1996) and has been recognised that additional attributes are needed for RBAC, for example in (Simon and Zurko, 1997, Thomas, 1997, Ferraiolo et al., 2003a).

The team-based approach modelled in the demonstrator has been accepted in the evaluation. In the demonstrator, the portal provides some of the task-based control over the actions a role can perform. In the general model, the team-based access control implemented by the workflow control point must enforce the task control by role. This compares with team-based access control, TMAC (Thomas, 1997) and TMAC 2004 (Alotaiby and Chen, 2004) reported in section 2.7.2. Both of these schemes control fine-grained permissions by team membership, which includes user identities and object instances. TMAC 2004 extended the TMAC framework with collaboration instances. The approach used in the demonstrator uses a similar fine-grained policy by defining the workflow as the collaboration. However, this research has combined the team based

approach with the task-based ideas presented in TBAC (Thomas and Sandhu, 1998) for task based creation of a roles permissions.

The DAME scenario requires that fine-grained access permissions be defined from high-level tasks and role permissions. The high level definitions form a template of the team-based access control, which is used in collaborative workflow instances. The Workflow-Team Policy Architecture provides the framework to define policy templates from role-based collaborative workflows that are used for fine-grained control when executing workflow instances.

5.1.4 Architectural Issues

The architecture of the business model has been illustrated in Figure 3.1, Figure 3.2 and Figure 5.1. In a distributed system, such as a SOA using grid-computing resources, there are issues about where access control components are located, and how the components are defined. The following are issues about the position of access control components in the business architecture.

- Point of access control to collaborative service instances, PEP as shown in Figure 2.9;
- Location of user identities and attributes;
- Location of access control policies;
- Management of dynamic access control policies;
- Scalability of the architecture, for multiple users and organisations in many collaborative teams, each with permissions and service instances;
- Privacy of permitted actions and instances used in any collaboration; and
- Loose coupling of SOA components.

As already stated, the workflow establishes a context for the collaborative access to services. Therefore, a solution would be the concept of a single access control policy for each workflow containing fine-grained access constraints for the service instances consumed. Dividing the policy into components at each service, as reported in (Russell et al., 2004b), would require each service provider and resource to understand the user identities and attributes.

If the policy and its control are distributed, then managing the collaboration introduces problems of synchronisation, ensuring all service instances have the correct users added to their policy on changes in the users in the team. Dependant on the situation, it may be important to ensure the propagation of a change in policy rapidly throughout the system. From discussions, it has been noted that ensuring permission to team members in a collaboration is more important than a denial of access.

Distributing policy rules to each service instance also has implications of privacy. The policy rules are fine-grained, specifying the users identity, role and permitted actions on the service. In a vertical supply chain, it would not be desirable for the Integration organisation to reveal details of their customers to the service providers.

The user identity and attribute servers would need trust to be assured in the contractual relationship. The employer of the user would be trusted to assert the user's identity and their organisation name. However, the employer may only suggest the role. The role relates to a definition in the workflow, therefore the Support and Service Integration organisation would assert the role on the user. One method would be for the external organisations using the workflow to assert the user's qualifications. A policy would contain a template of the qualifications a role requires that would be checked at the point of access to the workflow management system or web-based portal.

Management of the policies has implications on scalability. In the VO policy approach (Russell et al., 2004b), a single, dynamic policy associated with the collaboration instance would require remote access on each request of a service instance. Alternatively, the VO policy approach could distribute policy rules to each service instance. Therefore reduce the number of remote accesses and requirement for distributed policy management, along with reducing issues of synchronisation and privacy.

It would be preferable to use a single access control policy per workflow instance, which is controlled in one place. Ideally, this would be the point of collaborative access to the workflow, and restrict users access to the services based on role-based constraints for the workflow.

5.1.5 Summary of Objectives

The requirements for secure collaborative workflow can be summarised in by the following rules:

1. The architecture must cope with a dynamic user base.
2. The architecture must cope with dynamic target objects.
3. The architecture should not expose participants to each other through the vertical supply chain. This includes user and service identities and the organisations of customers and suppliers.
4. The access control mechanisms and workflow management system should not expose details of the business process.
5. Access control should relate to the collaborative process using contextual access control.

6. The architecture must scalable to cope distributed users, data and services.

From the analysis of the DAME requirements and the requirements for secure collaborative workflow, the following objectives for the general model are stated:

- Authentication and authorisation information flow should correspond to contractual links in Business-to-Business relationships;
- The definition of business process and access policy should remain unchanged whilst allowing changes in users and implementation of services;
- The collaborative access policy is controlled and managed in one place;
- The authentication and authorisation architecture should be scalable. That is, easy to add users and services, especially if they fit into existing roles and service types.
- The architecture should be heterogeneous to exist in a heterogeneous grid-computing platform.

5.2 General Model of Secure Collaborative Workflow

In this research, SOA abstractions have been used to promote decoupling of workflow definition from implementation. As Figure 1.1 from the introduction illustrates, the workflow definition is created from the business requirements. The business requirements also define the access permissions within the workflows. Figure 5.2 expands on Figure 1.1 by presenting the definition of the workflow and access control policy from the business requirements, and illustrating the instantiation to workflow instance and policy instance, discussed in Chandramouli (2000). Mendling et al., (2004), also discusses this issue, in which they investigate automating the definition of RBAC rules by interpreting BPEL workflow scripts.

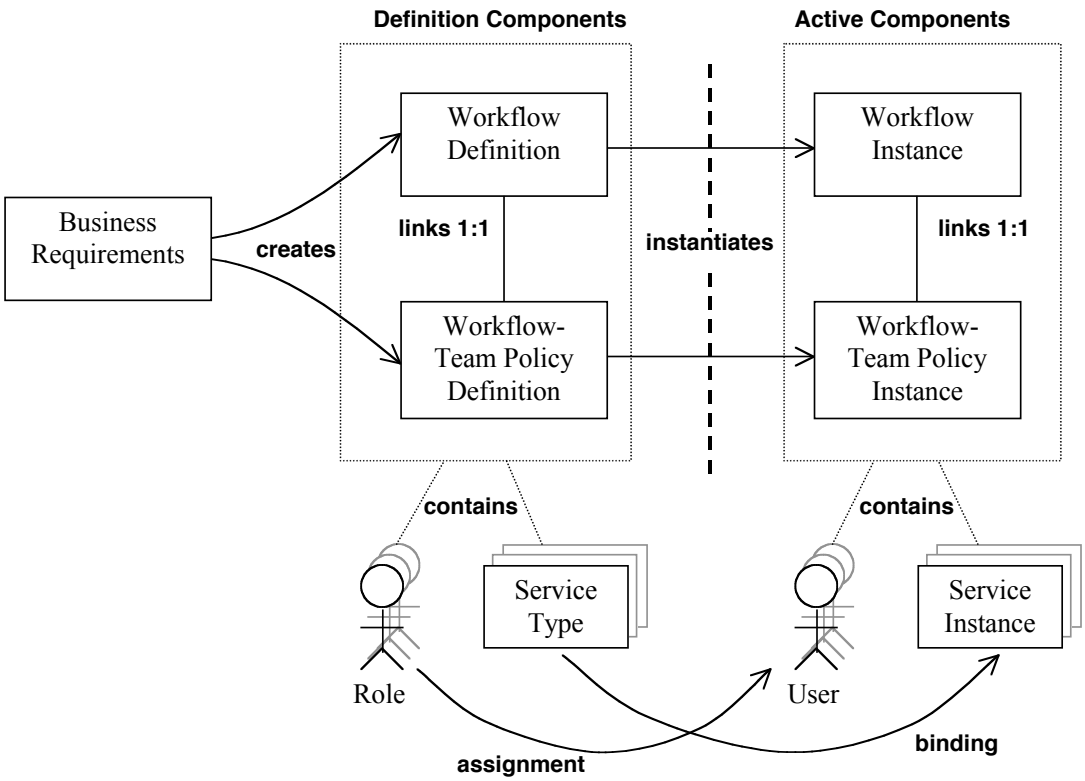

Figure 5.2 Business requirements creates workflow and access policy used in instances of workflow enactment

The access control policy template in Figure 5.2 is called the Workflow-Team Policy Definition. It is defined using the roles from the workflow definition, and is instantiated as the Workflow-Team Policy Instance when the workflow is instantiated. The new scheme retains the Role to User assignment from Figure 1.1, but extends it with the same concept for services. In the definitions of workflow and policy, the service type is given. When instantiated the service instances are recorded in the workflow instance, and specified in the fine-grained access control policy instance with user access.

The Definition Components in Figure 5.2 are the static, structural descriptions of the workflow and policy. On enactment, they become Active Components. The workflow instance and policy instance capture the dynamic changes of time, recording the users in role instances and service instances. Chandramouli (2000) defines temporal business associations, that define fine-grained permissions for users, based on roles from process definitions. In the Workflow-Team Policy Architecture, there are temporal business assets created for consumption during the process of the workflow. The Workflow Instance and Workflow-Team Policy Instance capture the workflow context. This includes the state of

the workflow, the users assigned to the roles and the service instances created to execute the service types.

Since a workflow management system has the ability to handle the data received from a service executed in the workflow, it is possible to pass the data to any user with access to the workflow. Regardless of the access control on a service, it is still essential for the workflow to ensure service results are presented only to the permitted user.

5.3 Workflow-Team Policy Architecture

The collaborative workflow tested in the DAME demonstrator is refined by the analysis to form a general model for secure collaborative workflow, using teams of users and service instances. As discussed in section 2.7.3, grid security mechanisms, such as CAS & VOMS, can provide control for predetermined teams using static services and data in the VO. However, a workflow has a subset of users from the VO and includes service instances as temporal business assets created for the workflow. The proposed Workflow-Team Policy Architecture is used to control collaborative access to service instances from definitions of role-based workflows. The Workflow-Team Policy Architecture is designed to control the collaborative access to services without service providers implementing collaborative inter-enterprise access control mechanisms.

The workflow creates the context for the use of services; it is also the point of collaborative access to the services and service instances. Therefore, the workflow instance has an access policy, which is called the team policy. There is one team policy per workflow instance. It uses the template policy, which defines role permissions to access service types. The team policy linked to the workflow instance resolves user to role mapping and defines fine-grained access to service instances.

The Workflow-Team policy is concerned with the authorisation of users' access to workflow instances and the corresponding services instances that are managed and deployed by different organisations. One important fact is that the Workflow Instance is trusted, such that it guarantees to carry out team policy. Therefore, service instance access control only needs to validate the workflow context. The next section illustrates the components of the Workflow-Team Policy Architecture.

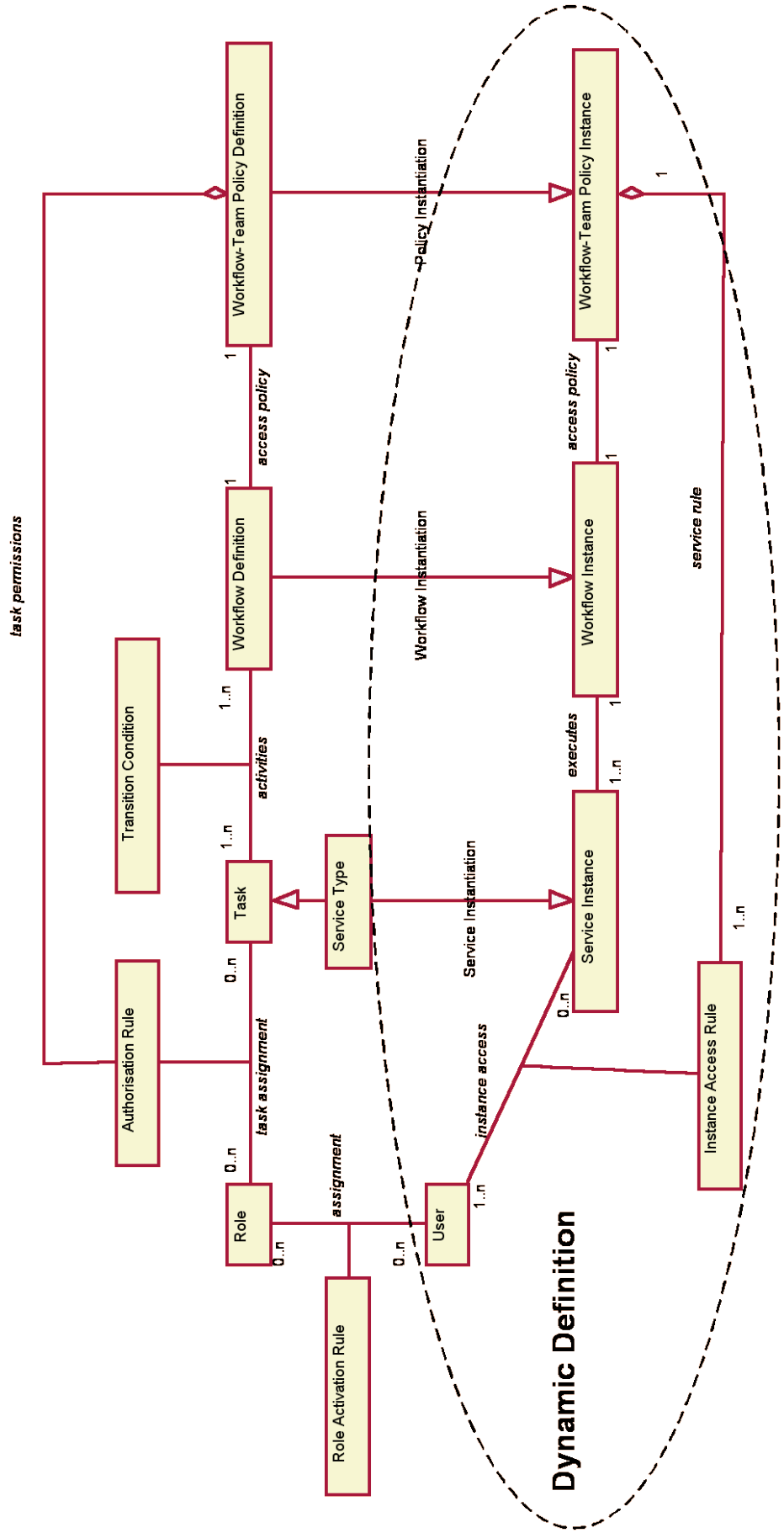

Figure 5.3 Workflow-Team policy architecture

5.3.1 Architecture Description and Static Workflow-Team Classes

Figure 5.3 shows the Workflow-Team Policy Architecture in its structural context, using analysis style syntax for UML class diagrams. This diagram shows the template definitions used to derive the active parts of the policy that are highlighted by the Dynamic Definition. The parts outside of the Dynamic Definition are considered static definitions because they are created from the slow changing business requirements. The *Workflow Definition* is made of *Tasks* that are linked by *Transition Conditions*, as in Aalst and Hee (2004 p.49). The classes in the static definition in Figure 5.3 are described below.

5.3.1.1 Role Activation Rule

When conditions are true, this rule in the business policy permits a *User* to assume the identity and permissions of a *Role*. For example, in DAME the rule is governed by the attributes a user is asserted by the employing organisation and that match a template of the *Role* requirements. There is also a list of organisations that can be permitted to use a *Role*. The organisation is another attribute of the user, asserted by the organisation. The organisation assertion is also checked that the organisation attribute matches the organisation providing the assertion. The *Role activation rules* are maintained in a business access control policy, because this applies across workflows. However, *Role activation rules* may be placed in the *Workflow-Team Policy Definition* if the rule is only ever applied to this workflow context. For example, if only one user can enact this role at a time in this type of workflow. Workflow specific controls over role usage in a workflow are contained in the *Authorisation Rule*.

5.3.1.2 Role

A template for users, a *Role* is usually constrained by the qualifications of users. Users are mapped to the roles that can enact in collaborative workflows, reducing the administration overhead of mapping users to permissions to enact workflows. In this architecture, the users are mapped to roles dynamically at the start of each workflow instantiation. The user to role mapping is validated against the *Role activation rules* and is recorded in the *Workflow-Team Policy Instance*.

5.3.1.3 Authorisation Rule

The *Authorisation Rule* provides conditions for the link between *Roles* and *Tasks*. For example, a user in the role Maintenance Engineer must belong to the organisation of the Maintenance Engineer that started the workflow.

5.3.1.4 Task

The *Task* is an activity in a *Workflow Definition*. The *Task* details a single service that executes and its parameters, as with Activities in BPMN (Object Management Group

(OMG) / Business Process Management Initiative (BPMI)) and in UML Activity Diagrams and <invoke> in BPEL (Andrews et al., 2003). Each task is linked to authorisation rules that map to the required role or roles that execute it and access it.

5.3.1.5 Service Type

The *Service Type* is a type of *Task*. Specifically it is a Grid Service type definition that creates service instances. The WSDL of the Grid Service is captured in the *Workflow Definition* document. Depending on the WSDL contents the service may be bound to a service implementation by providing the URL of the service factory that creates the service instances.

5.3.1.6 Transition Condition

This defines transition between each *Task* in a *Workflow Definition*. In the UML Activity Diagram it is the link between activities, this is a Flow, Fork or Join and can constraints of Condition, Decision or Merge. The transition types include Sequence, OR-split, AND-join (Aalst and Hee, 2004) from Petri-Net style definitions and relate directly to the Sequence and Gateway blocks in BPMN which include the Fork/Join and Inclusive Decision/Merge. When mapping to BPEL these notions become <sequence>, <flow>, <switch>, <while>, <pick> and <eventHandlers>.

5.3.1.7 Workflow Definition

The *Workflow Definition* is the document that captures the tasks, transition conditions, authorisation rules and roles. One entity exists for a workflow definition that is a static capture of a business process, such as a BPEL script. Each *Workflow Definition* entity links to one *Workflow-Team Policy Definition*.

5.3.1.8 Workflow-Team Policy Definition

This static description is linked one-to-one to a *Workflow Definition*. The document captures how *Roles* enact the *Tasks* by containing the *Authorisation Rules* for *Roles* to enact *Tasks* in the *Workflow Definition*. This is a template policy for the active dynamic workflow. It captures the rules by type, i.e. roles rather than users, that create the instance rules in the *Workflow-Team Policy Instance*.

5.3.2 Dynamic Workflow-Team Classes

When the workflow is enacted, the definition of the workflow is instantiated, becoming a *Workflow Instance*. The *Workflow-Team Policy Definition* becomes a *Workflow-Team Policy Instance* linked to the *Workflow Instance*. *Tasks* from the *Workflow Definition* that enact *Service Types* become *Service Instances*, and *Roles* that were permitted to access those *Services* via the *Authorisation Rule* are entered in the *Workflow-Team Policy Instance*

as *Users*, by their corresponding *Instance Access Rule*. In Figure 5.3, the *Dynamic Definition* is the architecture used in the enactment of the *Workflow Instance*. These types are defined in the following.

5.3.2.1 User

The *User* is a person whole can execute tasks in a workflow by enacting roles. As a minimum requirement the *User* has a name and may include additional attributes such as qualifications, experience and location. A *User* can be linked to many roles and can enact different roles at different times. The *Role* played by a *User* for a particular workflow is recorded in the *Workflow-Team Policy Instance*.

5.3.2.2 Instance Access Rule

The *Instance Access Rule* is derived from the *Role Activation Rule* and the *Authorisation Rule*. The *Instance Access Rule* defines which *User* can execute the *Service Instance,* stating the *Role* of the *User* and retaining constraints from the static rules.

5.3.2.3 Service Instance

A *Service Instance* is a stateful Grid Service instance created by a service factory. The *Service Instance* is of a type defined in the static *Service Type* derived from the *Task* in the *Workflow Definition*.

5.3.2.4 Workflow Instance

The *Workflow Instance* is the runtime state associated with the *Workflow Definition*. It records to progress of execution for a workflow, position in the sequence and acts as the document to contain information about the *Service Instances*. It replicates the static structure by linking one-to-one *Workflow Definition* to a *Workflow-Team Policy Instance*.

5.3.2.5 Workflow-Team Policy Instance

This is the policy instance created from the *Workflow-Team Policy Definition*. It is the dynamic document that contains the constraints on the executing *Workflow Instance,* providing fine-grained control as to *Users* permissions for actions on *Service Instances*.

This dynamic policy architecture can be used across a VO, so long as VO partners trust the workflow engine. However, the architecture can be easily implemented to provide team based access control within a single organisation, and prevent unauthorised access within an organisation.

5.4 Design Assumptions

The following lists the assumptions made on the supporting policies, business relationships and technologies used to complement the structure and behaviour of the Workflow-Team Policy Architecture.

5.4.1 Commercial Policy Assumptions

As mentioned in section 5.1.1 the business relationships form part of the structure for protecting information, services and the business processes. The following are assumptions derived from the DAME case study about the commercial relationships.

1. The Workflow-Team Policy Architecture must protect the business process definition and fine-grained instantiation from access by the consumers and service providers. Valuable business relationships could be revealed, and once the process is exposed consumers could go directly to providers.

The business process definition and the names of service providers are likely to be discovered at some point. The next two points address this:

2. A number of service providers and late-binding of services during the process can be used to select the service providers depending on cost or other quality of service metrics (such as speed, reliability).
3. Constant innovation will be used to evolve services, improving delivered results and operating conditions, such as speed to result and dependability of service. Another target for innovation is improved efficiency, resulting in lower operating costs.

Access control to the services is dependant on business relationships. The following are assumptions about the business relationships between the DAME stakeholders:

4. The workflow engine is trusted to execute the access control policy.
5. The definitions of access control and workflow understand the level of secure access that is required to protect the data returned by a service. This needs to be agreed as part of the service interface definition and implemented by the service provider.

5.4.2 Technology Assumptions

The following points relate to complementary technologies that support the Workflow-Team Policy Architecture.

6. An auditing system will record the actions of the workflow manager, to be stored by a third party. Access to the data is append only and only from the workflow management system. It will also include access control decisions (permit and deny) that are linked to execution of the workflow.

7. A policy decision engine such as PERMIS can be used to provide decisions on access determined by the workflow action (Task) and the Workflow-Team Policy Instance and linked policies.
8. All communications that need to be secure are assumed to be secure by the use of appropriate encryption.
9. User identity is asserted using X.509 certificates. The issue and management of certificates is the responsibility of the user's employer, and a trusted certificate authority signs certificates, which could be the employer.
10. The employer asserts user role attributes and the assertion could be an extension to the X.509 user certificate or retrieved via LDAP. The employer digitally signs the role assertions.
11. Performance of the access control mechanism is improved by reducing the number of policy documents required to form a decision, therefore a single document is used for the workflow rules.

5.4.3 Case Study Assumptions

The following point relates to the DAME case study presented in Chapters 3 and 4.

12. The captured workflow includes all the interactions required to complete the business objective. It is likely that the workflow used is incomplete in the real business situation since the case study is speculative and cannot be tested in operation.

5.5 Example policy system

The example in Figure 5.4 illustrates a system using the Workflow-Team Policy Architecture. The example illustration is based on the established DAME case study, however the Workflow-Team Policy Architecture is intended to apply to general collaborative workflow using service instances.

A user connects to the *Portal* using a standard web browser and logs in using their identity and password. The *Portal* retrieves the user's attributes, such as role and organisation, from an external location, such as the user's organisation. This method is scalable by using Shibboleth (Cantor, 2004), where the user authenticates to their own organisation's portal to retrieve SAML (OASIS, 2004) assertions about the user's attributes. The login authentication can use more than 'username and passphrase'. Further forms of identity authentication in combination with username and passphrase, are such as hardware key or fingerprint.

The *Portal* customises the users' views according to the role and organisation. It presents the work available from the *Workflow Management System* (*WFMS*) by a connection that passes the user's identity, role and organisation. The *Portal* displays status and results from available workflows. Stages in the workflow change the team membership or create new service instances, which are recorded in the Workflow-*Team Policy Instance*, by the *Policy Controller*. The *Portal* and *WFMS* form the collaborative environment.

The users and service instances in a *Workflow Instance* are specified in the *Workflow-Team Policy Instance*. Access to the *Workflow Instance* is controlled by its own Workflow-*Team Policy Instance*, via the *Policy Enforcement Point* and *Policy Decision Point*. This extends the architecture presented by Bertino et al. (1999) which also uses workflow instance policies, that are generated for the fine-grained user access control on rules that can only be determined at runtime. The *Workflow Instance* creates and controls *Service Instances* according to the *Workflow Definition*. The *Policy Controller* records the *Service Instances* in the *Workflow-Team Policy Instance*. The changes made by the *Policy Controller* are subject to the *Workflow-Team Policy Definition,* which defines the dynamic rules. Since, the user access has already been managed, it is not necessary to pass the user's identity to *the Service Instance*. Instead, each *Workflow Instance* is given a unique identifier (automatically generated) that is digitally signed by the *WMFS*. The *Service Instances* and *Compute Resources* are contacted only by using this unique identifier. This presents advantages when considering a business model, such as DAME; some of which are:

- Services and Resources do not need to understand collaborative access conditions; therefore, they do not need to interpret roles, organisations, or understand the context in which the users access the service/resource;
- Services and Resources do not need collaborative access mechanisms, making it easier to create new services and bring in new service suppliers; this also supports ease of creating a loosely-coupled service architecture;

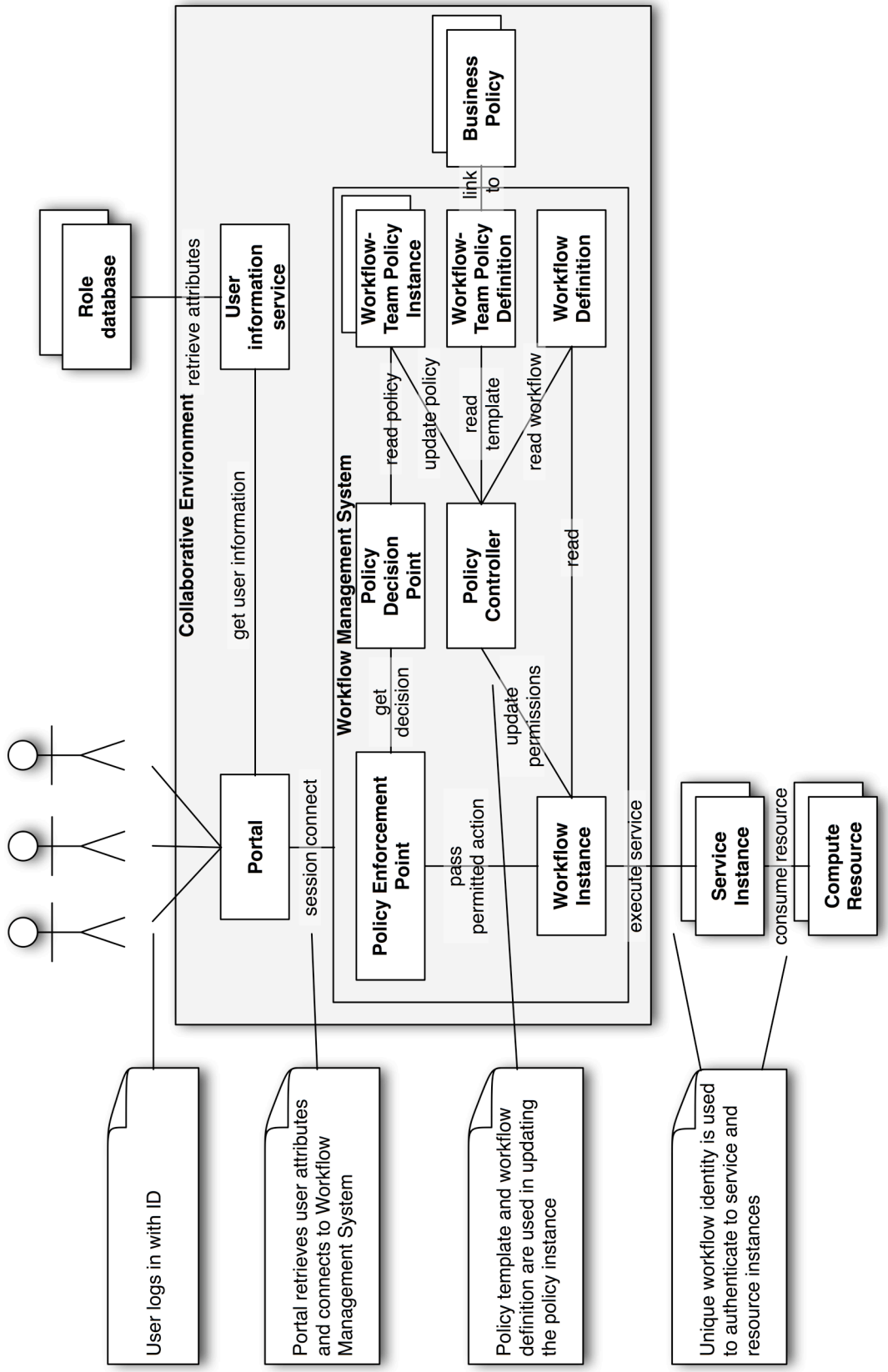

Figure 5.4 Workflow-Team implementation architecture

- Services and Resources do not need knowledge of every user in the system. In the DAME demonstrator, suppliers of processing services and compute resources are employed by the company operating the *WFMS* and not by the DAME customers, therefore services suppliers only need to deal with the workflow system, and rely on its mechanisms to control collaborative user access.

5.6 Summary

This chapter presented the Workflow-Team Policy Architecture from analysis on the DAME requirements and experience with the DAME project. The analysis addressed issues on VOs, linking business process definitions with access control rules for roles in teams and maintaining loose coupling in a distributed scalable architecture.

The result illustrates the management of a Workflow-Team Policy Instance for each Workflow Instance that links control of collaborative access with the execution of the workflow in the workflow management system. The main advantages are:

- No need for many remote Grid Services to retrieve user information from several sources;
- The remote Grid Service do not need to implement collaborative security controls;
- The management of policy components is restricted to one place;
- The architectural model for policy usage relates to the business-to-business contractual arrangements for service provision in a VO.

The next chapter described the implementation of secure workflow for the DAME demonstrator, which is used to test the Workflow-Team Policy Architecture.

Chapter 6

Secure Collaborative Workflow for DAME

This chapter describes the experiment to test the concepts of the Workflow-Team Policy Architecture by implementing a secure collaborative workflow in the DAME demonstrator. As part of the DAME project, a collaborative demonstrator was implemented to investigate the use of Grid Computing for the business case of distributed aircraft engine diagnostics. The implementation described in this chapter provides the integration of the Grid Service provided by the academic project partners by using workflow management and access using a secure portal.

The next section, 6.1, defines the objectives of the experiment. Section 6.2 describes the architecture of the workflow management system in the context of the DAME demonstrator. In section 6.2.1, the implementation and deployment of the workflow management system is described and related to the secure web-based portal and the deployed Grid Services. Section 6.2.4 lists how the Workflow-Team Policy Architecture is implemented in demonstration environment. Finally in this chapter, section 6.2.5 lists the assumptions about the implementation system in representing the case study requirements and Workflow-Team Policy Architecture.

6.1 Secure Collaborative Workflow Experiment

The experiment demonstrates the VO and diagnostic workflow illustrated in Chapter 3. The experimental setup used a portal and workflow manager to control the execution of remote Grid Services. Its main purpose was to evaluate the method of dynamically controlling collaborative access to Grid Service instances. This demonstrated the constructs of the Workflow-Team Policy Architecture.

6.1.1 Demonstrator Process

The experiment executes the business process from Figure 3.3 in section 3.4 DAME Workflow, by walking through paths represented in the DAME Use Cases (Fletcher, 2002).

The process simulates the arrival of new flight data that has been loaded into an engine data store and automates the launch of WF1 in Figure 3.4. The process uses real engine vibration data obtained from the Rolls-Royce engine test-bed. By choosing data known to cause a feature to be detected, the WF1 result was 'Unknown Feature'. The portal is collaborative, supporting simultaneous users, in different roles from different organisations. The ME could view the results of WF1 and escalate the process to a MA assigned automatically. Since the workflow instance is attributed to a particular case, it would only appear on the worklist of the correct user assigned to the ME role. The escalated process would appear on the worklist of the assigned MA, who could view the results and services already associated with the case. The MA could execute new processes, as per WF2 in Figure 3.5, and escalate to a DE. The automatically assigned DE would access results and services permitted to them from the escalated case and run their own services from WF3, in Figure 3.6.

The workflows use test data from Rolls-Royce, simulated case history for case-based reasoning services and the deployed engine model executed proprietary algorithms from Rolls-Royce engine modelling. The process was completed when the DE provided a diagnosis annotation about the cause of the detected feature and release the case to the MA. The DE no longer has access to the workflow. The MA can also provide annotation, and then must release the case to the ME. Again, this modifies the access policy and restricts the MA. The ME can then complete the case by releasing the engine for flight. All services instances and temporary data are destroyed.

The test procedure was not limited to the one case. The chosen input data was varied to produce different results from WF1 and subsequently the escalation requirements were different (i.e. WF1 results was 'Clear' and no escalation was required). Different users in the ME role could run cases and the case would be escalated to different MA users, who escalate to different DE users. Members of the consortium were assigned different roles and different simulated organisations to test the team dynamics. The Grid Services were hosted and executed at different organisations. One service employed a Grid resource broker to dynamically assign and transfer the processing task to different Grid Computing machines.

The test was repeated on several occasions, with different stages of implementation during the life of the DAME project. Changes occurred in services deployments, with changed location, functionality and service interface. Changes also occurred in the security used in the Portal and Workflow Management System, improving methods of login, by using myProxy server and grid X.509 certificates for all users and methods for defining policy templates, which define how, the dynamic team is controlled (subsequently known as the Workflow-Team Policy Definition).

6.1.2 Experiment Setup

The secure collaborative workflow was tested in the DAME demonstrator. The demonstrator components consist of:

- A web-based portal, with secure login, myProxy identity server and supporting databases;
- Workflow management system;
- Grid Services, running on GT3 (v3.02) middleware and provided by the project partners at the universities of Leeds, Oxford, Sheffield and York;
- White Rose Grid compute servers and storage, running GT3 (v3.02) middleware, hosted by the Universities of Leeds, Sheffield and York.

The implementation of the demonstrator uses GT3 Grid Services, and GT2 resource management; as GT4 was only fully released in 2005, part way through this research. The analysis of the results in Chapter 7 includes a discussion relating the research to GT4 implementations.

The architecture of the demonstrator is shown later in this chapter in Figure 6.1, and the implementation architecture in Figure 6.2. Further implementation details are contained in Appendix A.

6.1.3 Secure Collaborative Workflow Implementation Objectives

The objectives of the implementation of the secure collaborative workflow is to demonstrate and evaluate the following:

- Secure login of users from distributed locations and different organisations;
- Secure access to stateful services and grid resources;
- Role-based access control to commercially sensitive service and data, in particular the protection of competitors operating conditions;
- Collaborative business process that executes the diagnosis process identified in the Use Case Analysis (Fletcher, 2002);
- Dynamic use of grid resources, for dynamic process deployment across organisations;
- Single sign-on, for automated authentication to services and grid resources without the user logging in to each service or resource on each access request;
- Automation of collaborative team, to demonstrate users joining the team on escalation of diagnosis problem and leaving team on release of the problem;

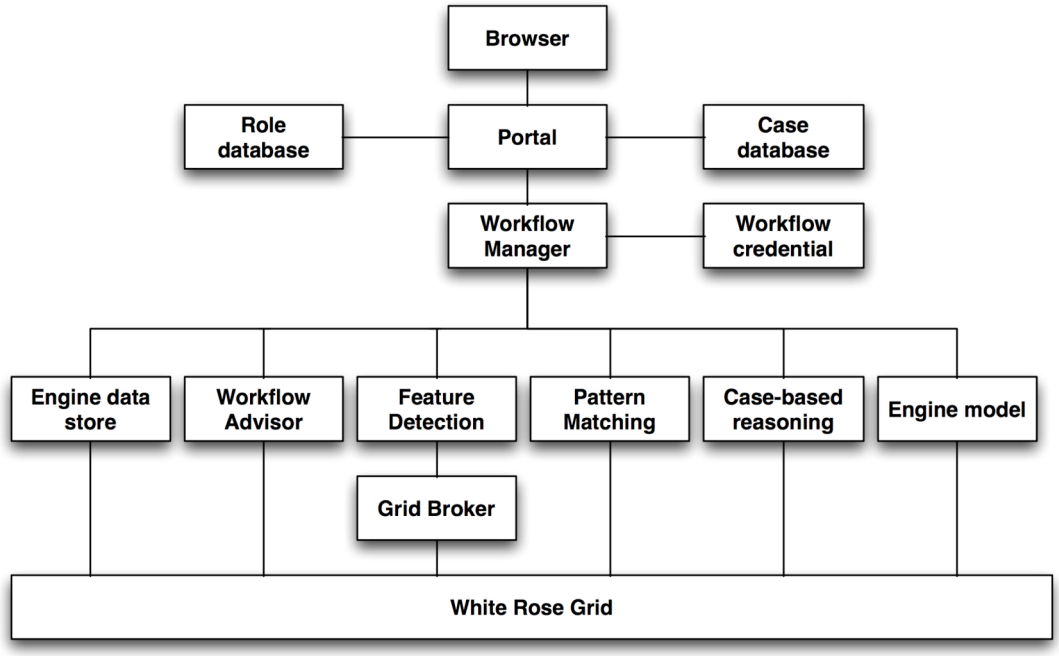

Figure 6.1 DAME demonstrator architecture

- Automation of access control permissions for collaborative team members to access long running stateful services.

Also under investigation are the following objectives:

- Determine the point of collaboration, so that access control can clearly identify the user requesting access;

These objectives have been used to focus the development of the workflow management system in the DAME demonstrator. The implementation has been developed along with the Workflow—Team Policy Architecture and used to prove the structure presented in Chapter 5.

6.2 DAME Demonstrator

The architecture for the DAME demonstrator is shown in Figure 6.1. The collaborative user accesses the distributed system using a web browser client, connecting to a web-based portal that executes on a White Rose Grid server at Leeds. The user logs into the portal matching the user's grid certificate that is installed in the Role database. The Role database contains user information, such as user identity, role and organisation. It also contains the policy template. The Case database contains a record of the users involved in the workflow and the workflow identity.

The portal connects to the workflow manager for users to launch, control and monitor workflows. The connection between the portal and workflow manager creates a session containing the user's identity and role. The session exists whilst the user is logged into the portal, and can be used to update the user on workflow events.

The workflow manager executes workflows, which in turn call GT3 based Grid Services. It was recognised that the workflow manager is the point of collaboration between different users, and the point where multiple services are called from different service providers. Using the users role and organisation information access to the services is control in the workflow manager and a single X.509 grid certificate, the workflow credential, is used to identify the workflow manager to the Grid Services.

The Grid Services are deployed on White Rose Grid compute nodes. The Feature Detection is deployed as a GT3 service on the grid and dynamically uses any compute node on the White Rose Grid selected using a broker service. The Feature Detection executes grid jobs in GT 2.4, and requires an identity and account space on the machine to which the job is submitted. This is achieved using the grid extension to X.509 that allows delegation of rights. This means the Feature Detection service can use the identity provided by the workflow manager to access the White Rose Grid.

The execution of workflow needs to be accessible by the distributed users. A web-based solution was chosen for its capability to reach users within corporate networks without modifying firewalls. This assumes the use of standard web server ports[7] and the organisation in the business case do not block the Integration Service Supplier's web site.

The analysis of the requirements captured for DAME, result in the first build and test phase of the project in the form of the DAME demonstrator, with the workflow management implemented as the integration point for the services to support engine diagnosis.

[7] Web server ports are port 80 for standard unsecured traffic over http and port 443 for secured traffic over https.

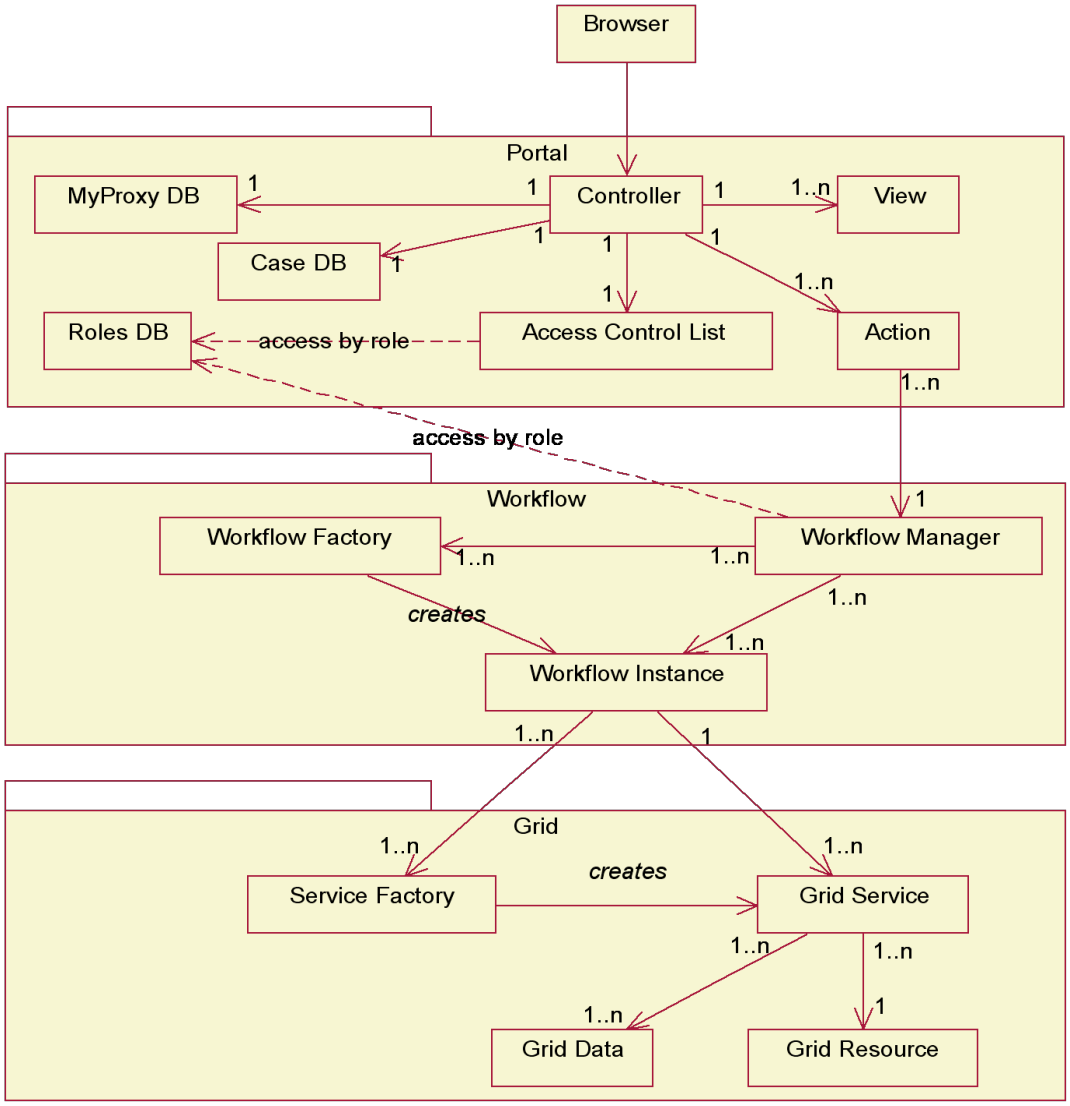

Figure 6.2 DAME implementation architecture

6.2.1 Workflow Management System

The implemented architecture for the DAME demonstrator is illustrated in Figure 6.2 as a concept class diagram. The diagram shows three packages Portal, Workflow and Grid. The Portal is the user's interface to view and control the workflows that are executed in the Workflow package. The Workflow package provides service integration by executing the Grid Services in the Grid package.

The internal architecture of the workflow package was based on the WfMC reference model, described in section 2.3. Further detail on the implementation architecture of the WFMS can be found in Appendix A. The key points of the workflow management system (WFMS) and the demonstrator architecture are summarised below.

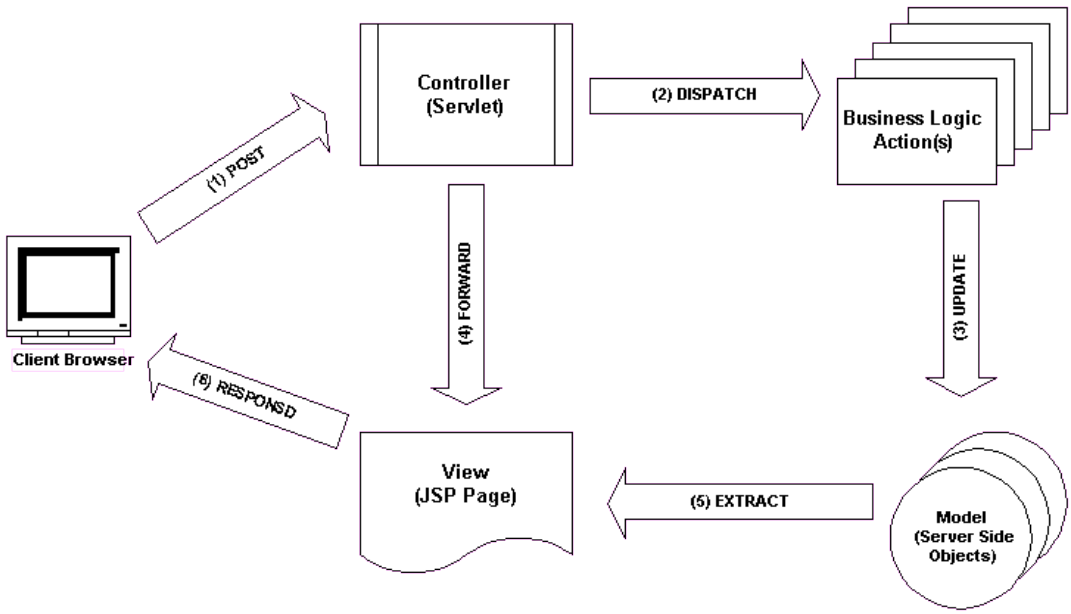

Figure 6.3 Apache Struts Architecture (JSP Model 2), from (The Apache Jakarta Project, 2004b)

The user logs into DAME using a web browser with secure login, illustrated at the top of Figure 6.2. The portal uses Apache Struts (The Apache Jakarta Project, 2004b) to control user access by role. The *Controller* provides the *Views* and *Actions* based on the configuration for that user's role. The flow of the portal actions is illustrated in Figure 6.3. On login, the portal displays a worklist of workflows available to the user that the portal retrieves from the workflow manager. Example worklists are shown in Figure 6.4.

The portal connects to the workflow manager using a session for each user logged in. The session is a bidirectional connection that is made by using the identity, role and organisation of the user. The session exists whilst the user is logged into the portal. The portal is trusted to handle user authentication and provides the user's grid certificate. The bidirectional connection is used to inform the user of workflow events, so a user can be notified on workflow events, relating to the *observer* and *event*[8] patterns (Gamma et al., 1995). The WFMS contains the worklists of workflow instances associated with the users. This means that workflow events can be sent to collaborating users assigned to a workflow instance.

The workflow manager executes workflows, which in turn call GT3 based Grid Services. In GT3, a service factory creates a service instance and assigns it a unique URL

[8] Observer and event patterns originated from the composite MVC (Model-View-Controller) pattern in Smalltalk

for http communication. As mentioned in Chapter 2, workflow management systems execute workflow defined in scripting languages, such as BPEL. However, at the time of implementation, no existing workflow language had support for the dynamic URL assignment used by GT3 Grid Services. Another issue raised in section 2.7.2, is that BPEL and other workflow languages do not support any security mechanisms. Therefore, to demonstrate secure collaborative use of GT3 services, the workflows were modelled as sequences in UML and coded in Java.

The Java code included the Grid Service client code generated from the WSDL for the services. Service discovery was a manual process. The generation of client code was also started manually by executing scripts to generate code stubs by retrieving remotely located WSDL.

During development of the demonstrator, interface changes to services required regeneration of the client code, using the updated WSDL. Depending on the level of interface change, this could require changes in the workflow code, typically when data exchange formats changed.

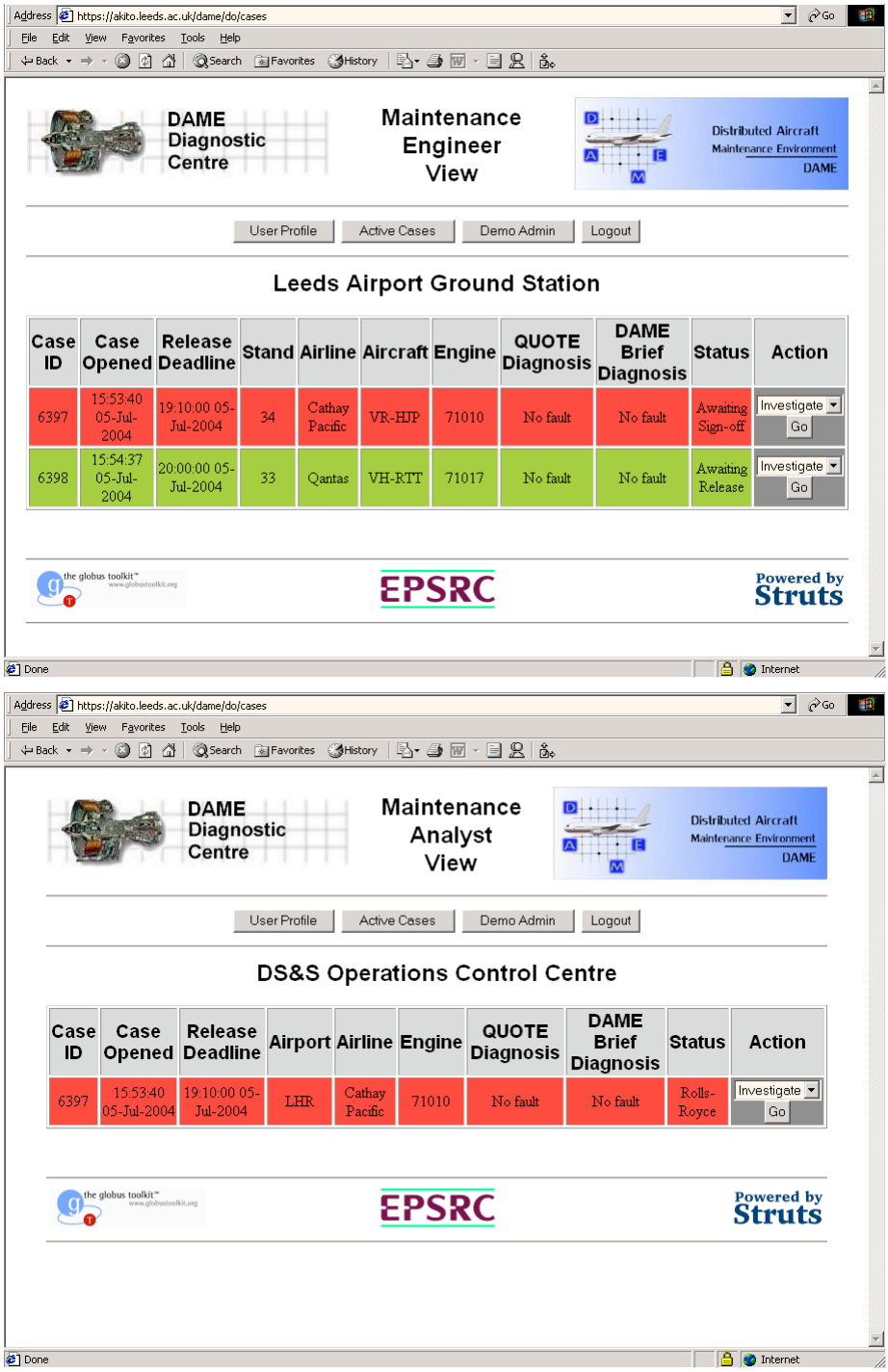

Figure 6.4 Portal view - worklist view for ME and MA role

By coding the workflows, the access to service results, initiation of service execution and event notification rules could be handled in the java class. Whilst not being very flexible, this allows the complete access control solution to be tested in the workflow manager. Some of the access control rules were duplicated in the configuration of the portal. Both WFMS and the portal access the Role database. The Role database contains user attributes such as user identity, role, organisation and the escalation route. The escalation

route details to which user the diagnostics workflow can be escalated, forming a template for the collaboration team.

Even though the user is connected to the WFMS by passing in their identity, the execution of Grid Services uses the identity of the WFMS. This identity is a X.509 grid certificate that identifies the WFMS to the services. The WFMS identity is recognised as 'DAME' and trusted by the services to provide adequate access control for the collaborating users.

The portal could initiate the execution of workflow by proxy, without the user needing to login, because it contains the identities, roles and organisations of the users. This is to simulate a user launching the automated workflow WF1 by retrieving data from an engine on an aircraft that had just landed. The workflow is launched and attached to the worklist of the ME responsible for that engine at the airport. The workflows achieve single sign-on by using the workflow certificate, and the users' worklists. Users can log on and off, creating and destroying session connections, to monitor and control the independently autonomous workflows.

6.2.2 Portal

The portal uses Apache Struts (The Apache Jakarta Project, 2004b), running on a Tomcat web server (The Apache Jakarta Project, 2004a). The Role database contains user information, such as user identity, role and organisation. It also contains the escalation path for users in the diagnostics workflow, a template of access control policy. The Case database contains a record of the users involved in the workflow and the workflow identity. The Case database also holds records of the access control policy instances, specifying the users involved in a workflow instance.

The management of users' identities uses the myProxy service that contains grid proxy certificates. The Tomcat web server contained an applet for users to install their grid proxy certificate. Once the grid proxy certificate is installed, the user logs into the portal by providing their username and passphrase that matches their proxy certificate, which is retrieved and used in the session connection to the WFMS.

The portal connects to the workflow manager for users to launch, control and monitor workflows. The workflow manager is also executed in the same web server as the portal, but could be executed on a separate server.

The Apache Struts portal uses an MVC pattern based architecture that is configured by 'role' to serve views and respond to controls. This reinforces the access control in the WFMS by allowing the users to only see the types of views and controls their role is permitted. The WFMS provides the fine-grained access control by ensuring that only

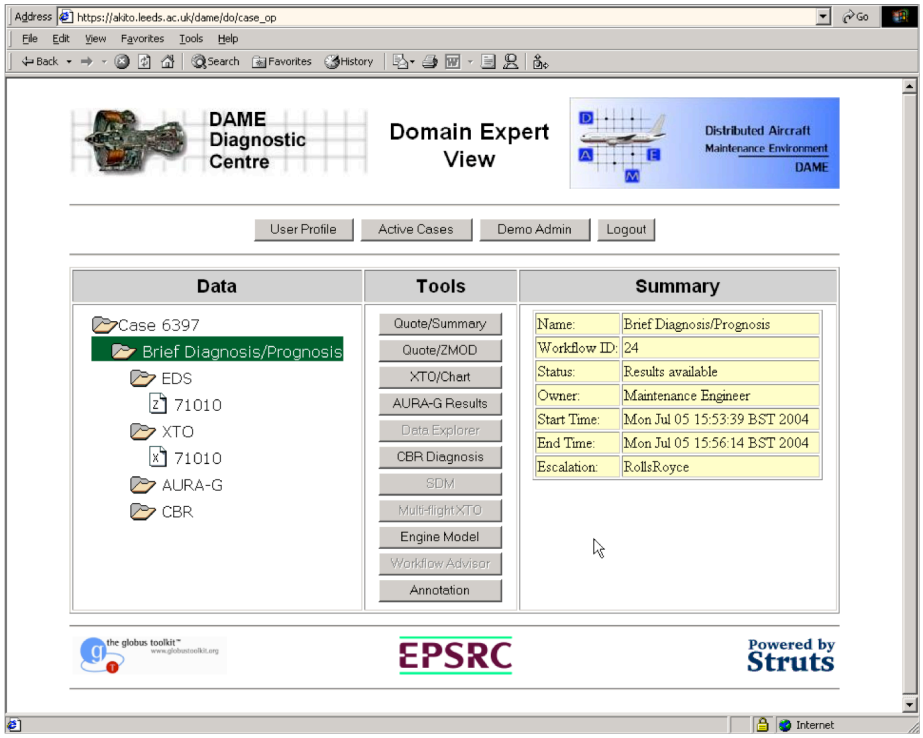

Figure 6.5 Portal view - diagnosis tools view Domain Expert

particular users can access services instances, and destroys service instances on destruction of the workflow instance. The example view in Figure 6.5 shows the tools the DE can view for a particular case, some buttons are greyed because the service is either unavailable at this point in the workflow, or the service is currently executing and results are not available. The same view can be seen by the ME and MA, however the range of tools is different, showing only those they are permitted to use.

The portal can support multiple users, each with individual sessions connected to the WFMS. It can even support multiple logins from the same user, which would share the WFMS connection session. The MVC pattern in Apache Struts allows the users to get updated views on updates events from the WFMS.

The original DAME specifications indicated that the roles ME to MA to DE are hierarchical with the MA inheriting the permission of the ME and DE inheriting permissions of the MA (and therefore the ME). On implementation, it was found not to be the case. The tools used by the ME and MA may not all be available to the DE. The DE is an expert in engine design and not directly part of the through-life delivery of engines, whereas, the MA is the business provider for the through-life support of aircraft engines.

6.2.3 Deployment

The DAME demonstrator deployment involved a collaborative VO. It involved services supplied from each of the four universities. The users that tested the portal were entered in

the Role database and given simulated identities for the DAME business model. Since the business model involved many MEs, most users were given the ME role and each one on a different airline. Other users were given the MA role, with organisation DS&S, and less users were given the role DE at Rolls-Royce.

The WFMS and portal were deployed on the White Rose Grid machine 'chablis'. The Role database, the Case database and myProxy server were deployed on White Rose Grid machine 'alba'. The Feature Detection GT3 service was also deployed on 'alba', with the XTO processing service dynamically deployable using GT2.4. The other GT3 services were deployed on the White Rose Grid compute nodes 'maxima', 'snowdon', 'pascali' and 'tatania'.

The White Rose Grid machines create a heterogeneous system with two multi-processor architectures: either Sun SPARC or Intel-based. The operating system on these machines is either Red Hat or SUSE Linux. Sun Grid Engine is used for queue management and Globus Toolkit 2.4 is used to manage the heterogeneous Grid Computing access to all machines. The Globus Toolkit 3 services are deployed in a GT3 container that connects to the installed GT2.4.

The White Rose Grid is a consortium, but has separate administration domains at the three universities. Users wishing to access White Rose Grid machines can obtain a White Rose Grid X.509 grid certificate and apply for accounts on each machine they wish to use. Access to the grid machines used the DAME identity via the WFMS. This only required DAME to be recognised and mapped to a user account on each White Rose Grid machine.

6.2.4 Implementation of Workflow-Team Policy Architecture

This section describes the relationship between the DAME Demonstrator and the Workflow-Team Policy Architecture. The following table describe the implementation of the components from the Workflow-Team Policy Architecture:

Table 6-1 Workflow-Team Policy Architecture Implementation

Workflow-Team Component	**Implementation in DAME Demonstrator**
Role Activation Rule	The activation rules are coded into the relationship in the Role database, relating the list of users to the roles. This is a limited implementation of the user to role mapping. Extensions would enable organisations to specify the user's role and the rule would list organisations against roles and include attributes the user must have in order to enact that role.

Role	The list of roles is in Roles database. A full specification would include role attributes, such as qualifications required.
Authorisation Rule	The activities a role can perform are captured in the rules in the Struts configuration. The Struts files link the roles to permissions in the MVC model relating directly to the workflow. The limitation of the implementation is that the rules are not be contained within the workflow document. However the XML configuration of Struts would allow the individual rules to be directly mapped between the Workflow Definition and the Struts configuration.
Task	The tasks are listed in the java coded workflow definition as service calls, from generated GT3 service clients.
Service Type	The service types are captured as GT service definitions, which is contained within the WSDL file and also implemented as the service client within the workflow manager. Extensions to the implementation would be to capture the service type as the type described in the WSDL, or semantic description of the function and parameters. Both extensions would require an activity within the workflow execution to locate and bind a service implementation, possibly generating service clients at bind time.
Transition Condition	The demonstration scenario required simple transitions between tasks. The sequence and decisions were encoded within the java workflow definition. The decisions were simple loops based on completion of the workflow as stated by the user.
Workflow Definition	The workflow definition captures the tasks and transition conditions by coding in a java class. The class reflects the workflow as captured in the UML activity diagram. Each coded workflow conforms to a defined interface in the workflow management system, so that workflows can be discovered by name and dynamically included in the system.
Workflow-Team Policy Definition	The policy definition is the collection of authorisation rules relating to the workflow. These are captured in the configuration of the struts portal, coded a rules in XML that

	map role permissions to actions and viewable results for a workflow type. This includes checking attributes of the role, so that organisation of the Maintenance Engineer is validated against the organisation of the launching Maintenance Engineer.
User	The list of users and their attributes is contained in the Role database. The limitation is such that user information is contained centrally and not obtained externally. However, the implementation was structured such that user details were retrieved on workflow instantiation, simulating the effect of retrieving them remotely. The retrieval code could be simply replaced to access the relevant external information system.
Instance Access Rule	The rules that capture which users are operating on which workflows and in which role is captured in the Case database. This is a dynamic list of workflow, user, role and task for each task in the workflow. This is the main part of the dynamic definition of the Workflow-Team Policy Instance.
Service Instance	The GT3 service instance is created by remote service factory and exists on a remote server. The workflow instance retains the reference to the remote service instance via the create service client instance.
Workflow Instance	The Workflow Instance is created by the workflow manager, from the given workflow name. It creates a instance a java class by locating it using the given name and adds it to the worklist with a unique identifier. The Workflow Instance then autonomously executes the Workflow Definition, coded in java, passing information to and from the user via the workflow manager.
Workflow-Team Policy Instance	The Case database captures the attributes of the executing workflows. In addition to the Instance Access Rule, the database links to the Workflow Instance and list other attributes: the organisation of the Maintenance Engineer that launched the workflow, start time and details of the aircraft engine.

6.2.5 Implementation Assumptions

The design of the DAME Demonstrator to test the Workflow-Team Policy Architecture required some assumptions about the implementation of complementary systems. The assumptions are listed in the following statements:

1. The implementation of the workflow security is assumed to provide the correct access to users, with respect to their role, organisation and permitted level of participation as defined in the business process derived from the DAME Use Cases, as described in Chapters 3 and 4.

2. The use of GT3 Grid Services does not restrict the architecture to GT3 implementation. The Grid Service instances are created in a service factory. They are referenced using a unique URL. This reference provides access to the long running process that is maintaining its own state. This is much like accessing the workflow process containing state of the workflow. The general model of the Workflow-Team Policy Architecture assumes that techniques other than a URL referencing the long running process can also be used by modifying the implementation, but not the architectural model. An example of a different reference would be WS-Addressing (W3C, 2004), using in GT4 that contains a reference in the SOAP envelope.

3. Portal access uses X.509 certificates to identify users. In the implementation these are contained in a MySQL database accessed by the Portal. The assumption is that secure login to the portal would include providing X.509 identity and these would be managed and provided by external companies. This was considered outside of the design scope considering there are already techniques for identity management (*Liberty Alliance Project*, 2006, Yao, 2003, Becker, 2002, Mont et al., 2002).

4. The case study has a requirement for a diagnosis result to be returned within the turn around time of the aircraft. The design of the workflow management system uses Grid Services. The implementation assumes that workflow processing uses available compute power to achieve adequate results in the time frame. In the DAME Demonstrator the Grid Broker, shown in Figure 6.1, fulfils this task.

5. The implementation assumes that the security protocols in GT3 Grid Services provide sufficient protection against attacks such as snooping, man-in-the-middle or taking control. Since the implementation uses the security standards in the Globus distribution, when security protocols are updated then the system will inherit the new protocol. It is expected that changes in security standards would require little or no change to the Demonstrator implementation and would be require no change to the architecture.

6. The implementation assumes that access to sensitive data is via secure Grid Services and that the remote data is stored securely.

6.3 Summary

In this chapter, the DAME requirements have been used to define the requirement for secure collaborative workflow. The implementation of which has been integrated into the DAME demonstrator. This provides the integration of services into a diagnosis process, supplying part of the aircraft engine support application. It also creates the collaboration point in the distributed DAME architecture. This is a good candidate for secure control of access to the Grid Service instances temporarily created during the workflow execution. The implemented workflow management system uses fine-grained access control for users' access to service instances in workflow instances based on template collaborations defined in the user identity attributes.

The next chapter presents the evaluation of the demonstrator using the results of interviews with the industrial experts and the results of the DAME Dependability and Security Study (Fletcher et al., 2004b). The evaluation analyses the implemented workflow management system against the Workflow-Team Policy Architecture and compares it with recent developments in the field.

Chapter 7

Analysis of Workflow-Team Policy Architecture

This chapter present the results from the qualitative evaluation of the DAME demonstrator. The primary results are from interviews with industrial experts, providing valuable feedback on the secure collaborative workflow concept and implementation. The industrial experts were used to provide opinion on the validity of the implementation and the business model in the case study. Other results have been obtained from feedback on DAME project meetings and conference presentations. Section 7.3 presents a comparison of the Workflow-Team Policy Architecture with recent work in several areas, including workflow and business processing. In section 7.4 the Workflow-Team Policy Architecture is compare with developments in access control, including Grid Computing middleware GT4. Section 7.5 reviews the architecture of the solution. Section 7.5 discusses the limitations of the research and section 7.7 looks at the next steps in Future Work.

7.1 Evaluation of Collaborative Workflow Security in the DAME Demonstrator

Feedback about the DAME demonstrator was obtained during live demonstrations of the portal, showing multiple users collaborating of the diagnosis and engine data. The demonstrations were based on the identified use cases (Fletcher, 2002) and specifically they presented walkthroughs of the scenario presented in Chapter 3. To gain in-depth knowledge, an industrial representative from DS&S and one from Rolls-Royce were interviewed using a semi-structured interview. Full transcripts and a list of questions can be found in Appendix B. Further feedback was obtained by presenting the architectural models to the consortium partners during DAME project meetings and presentations at external conferences (Russell et al., 2004b, Russell et al., 2004a, Russell et al., 2005). Additionally, this analysis draws on the DAME Security and Dependability Study from DAME project partners (Fletcher et al., 2004b).

7.1.1 Interview Process and Scope of the Evaluation

The aim of the interviews is to assess the research into secure control of collaborative workflows, in the context of the DAME case study. This provides a concrete example of how the research has targeted a real problem. The interview candidates were representatives of the business concerns of the case study and the interviews were aimed at testing the following topic areas:

- The validity of the DAME case study against real business objectives.
- The implementation of the DAME Demonstrator against the case study.
- The collaborative use of processing and the extent of the secure protection of the collaborative workflow and access to commercially sensitive information.

The scope of the above topic areas was restricted using the following:

- Collaborative use of workflows and grid processing tools in the diagnostics procedures.
- Designation of users and roles in inter-organisation processes.
- The protected access to competitors operating data using business processes.
- Access control to data, workflows and services across organisational boundaries.
- The architectural management of the link between workflow control and access control.
- The architectural link between the secure collaborative workflow and the DAME business requirements.
- Relevance of the demonstrator collaborative workflows to DAME business requirements.

During the interview the interview candidates were guided as to whether the question applied to the case study as known at the start of the project, or if it should be compared against recent understanding of the DAME requirements with future business objectives in mind. The interview exercise also benefited from more exploratory questioning, where the interview candidate was free to provide extended answers.

The industrial interviewees were chosen because they have broad experience of the business requirements for predictive maintenance service contracts and understand the technical challenges for employing SOA and Grid Computing across organisations. The interview was designed to gain the opinion of the experts by using open questions targeting the collaborative workflow concept implemented in the DAME demonstrator.

The interview candidate from DS&S was Charlie Dibsdale, who has experience of the service supply business supporting Rolls-Royce products. DS&S support business for Rolls-Royce includes decision support systems and services for aircraft engines, marine

engines and gas turbines for power generation. Dibsdale's interests lie particularly in the secure collaboration and control of decision support services, with the opportunity to use available compute power and storage via Grid Computing.

The interview candidate from Rolls-Royce was Graham Hesketh, who is leading research groups in agent based software, and is experienced in supporting aircraft engine design, manufacture and through-life issues with IT based solutions. Hesketh's interests in DAME are based on the diagnosis support of engine data, and maintaining the privacy of data, results and services employed by Rolls-Royce.

The interview questionnaire addresses the collaborative workflow, the collaboration procedure, the DAME roles and the access control scheme. The questionnaire was first tested on DAME project members at Leeds that have been involved with the portal and business modelling. Based on the test answers the questionnaire was modified to clarify question wording and change the order to a more logical progression. The interview questionnaire and full results from both expert interviewees are in Appendix B. Quotations from Dibsdale are referenced from Appendix B.2. Quotations from Hesketh are referenced from Appendix B.3.

7.2 Evaluation of Interview Results

The interview results show that the DAME system successfully demonstrated the secure role-based collaborations and proves the feasibility of integrating grid-services from different organisations. The demonstrator was used as a proof of concept. It proved the feasibility of GT3 based grid-services in a pseudo environment of different service suppliers and compute resource suppliers, modelled using the White Rose Grid.

The demonstrator captured the collaboration procedure in the business process definition, using the escalation task to add users and the release task to remove users, for the limited use cases investigated. However, future development of a DAME system needs to refine the business process and access control definitions. Dibsdale expressed that there are other similar opportunities for collaborative diagnosis with other types of operator, such as power generation. Hesketh expressed that the access control and workflow should not be to inhibit other collaborations that are acceptable to the business of solving problems:

> *Not telling this fleet operator that somebody else's fleet has a problem with its engines, they are simply providing a recommendation based on their own data. But it is an analysis which is inspired, if not the conclusions drawn from somebody else's data … What we wouldn't want to do is construct a system that would prevent that from happening* (Hesketh).

The demonstrator was useful in illustrating that role alone is not enough to restrict access to the workflow. A user's organisation attribute would be used to ensure that a user in a similar role in one Airline could not access data owned by other Airlines. This was reinforced in the Dependability and Security study (Fletcher et al., 2004b), but had not been identified in the original use case requirements (Fletcher, 2002). An example from Dibsdale states that at the Maintenance Analyst level, sharing diagnoses provides benefits to the customers (albeit a benefit to one customer gained from experiencing problems from another), which is typical in a diagnostics environment.

> *The specific data relating to a specific operator would not be divulged to other operators, but it is going through the same people* (Dibsdale).

Therefore, a Maintenance Analyst can view data across airlines, but should not be able to divulge specifics to another airline. However, Dibsdale did illustrate ways of anonymising data by aggregation so that it can be shown to competing operators:

> *However, there are ways and means of anonymising data, because a lot of operators are very interesting in bench marking themselves against the whole industry averages and therefore can plot their performance with some derived key performance indicators. Against the fleet average is a very powerful output that most operators are happy that their data contributes to and that they get a comparative view of their performance and others get the same* (Dibsdale).

Dibsdale went on to relate another example where access control by role is not enough and organisation needs to be taken into account. The example discusses that support operators at Rolls-Royce and DS&S enact the same role. Both roles can access costing data from Rolls-Royce. Users at DS&S are restricted to certain fields whereas Rolls-Royce employees can see all the data.

Specific data privacy issues still remain in the DAME demonstrator, such as controlling the data returned from the services that can be seen by the airline MEs. The workflow and access control definitions assume the level of information disclosure returned from a service remains constant. Specifically, the case when deriving information from the combination of results from different services is difficult to address. Dibsdale stated:

> *An intelligent user of the system could fire up a whole series of them [workflows] and reverse engineer who it came from. You have to be sure that the instances of the conglomerated results are not attributable to specific customers. It is going to be very interesting in the CBR [Case Based Reasoning] approach*, (Dibsdale).

Using dynamic resource allocation in Grid Computing for dynamic deployment of a service can create another issue of privacy. Schemes to protect algorithms and data in remotely deployed services are in a developing field (Yang and Xu, 2005), but essential for applications like DAME. Dibsdale expressed concerns of privacy when using brokered services across available grid resources:

> *The other aspect is in general grid usage that the security model needs to be extended to actually protecting applications as they are running on external processing devices, not only the application but also the data*, (Dibsdale).

The portal and workflow manager have been demonstrated with user authentication using single sign-on across all organisations, which is required for integration of commercial services. DS&S have expressed that organisations are responsible for asserting username to role assignments. DAME would provide a template that details the requirements of a user to assert the role. The requirements would be similar to a job description that may include qualification levels and be passed as user attributes to validate the role assignment.

Another point to arise from the evaluation was that this system provides the opportunity to separate the business critical workflow definition from the service implementation by using SOA and grid-services. Dibsdale expressed concerns about adapting the business process by managing changes to services and the level of services offered:

> *"…you have to constantly re-invent the service level and adapt it to higher and higher levels as you are going along. Spotting those up sells and deciding to delegate some of those services down to a freebee is an art, a science in its own right"*, (Dibsdale).

Both Hesketh and Dibsdale stated that the business process and security environment must be adaptable to new users and services. From Hesketh, concerns were expressed about the accuracy of the role definitions; in particular, the hierarchy of roles defined early in the DAME project do not reflect access permissions:

> *"The security environment must be adaptable to new users and services*", (Hesketh).

These points reinforce the need to provide abstractions from users and service implementations by using roles and service definitions, which need to be managed along with managing changes in business processes (Smith and Fingar, 2003).

The architecture of the DAME demonstrator received differing comments. Both Rolls-Royce and DS&S agreed that a centralised approach to managing the portal and workflow system meets the business requirements, in particular control of workflow

definitions and access policies. Airline user attributes would be provided by the airlines, such that assertions would state the users' qualifications. These qualifications would establish that the user could fulfil a particular role. Dibsdale stated that the role would equate to a job description with qualification requirements. The Dependability and Security Study stated that the centralised workflow provenance records could cause issues of trust. Both Dibsdale and Hesketh stated that the workflow provenance records are valuable assets that can be used to improve the diagnostics business process. However, a trusted third party should hold the workflow records, accessible by the customer (the Airline) and management company (Fletcher and DAME Architecture Working Group, 2003). The workflow records system, called the Provenance System, was still under development at the time of the study and is outside the scope of this research.

One comment on the architecture, from Dibsdale, conveyed that the workflow management system and workflow instances should be accessible by external components, for integration into other organisations' processes. In the DAME demonstrator the workflows are accessible only via the user interface of the portal. In the demonstrator the workflow management system was executed in the same web container as the portal, using a simple applet connection to link the two. It would have been straightforward to expose the workflow manager to remote access. However, the implementation did not model requirements for authentication by external components.

7.3 Comparison with Workflow and Business Processing Concepts

Discussed in the high-level model for the Workflow-Team Policy Architecture in section 5.2. The transformation of business process requirements into the Workflow Definition and Workflow-Team Policy Definition was similar to techniques linking business process definitions to workflow definitions and collaboration control are stated by Aalst and van Hee (2004:258) and Smith and Fingar (2003:250). Aalst and van Hee discuss mapping Petri-nets to workflow engines. They consider issues of how to map dynamic workflow allocation rules, including separation of function (separation of duty) and authorisation requirements. Smith and Fingar discuss using a pi-calculus based approach to flexible business process management, extending the early techniques of process modelling and re-engineering by adding monitoring and continuous process improvement. Smith and Fingar present a tiered approach, shown in Figure 7.1. The business process executed in the automation tier creates a 'Business Firewall' to protect the internal assets of the company in the integration tier. The components in the automation tier control who can perform which functions in the business. Similar to the Workflow-Team Policy Architecture, this requires external access to services to be captured in the business logic and controlled by a workflow management system.

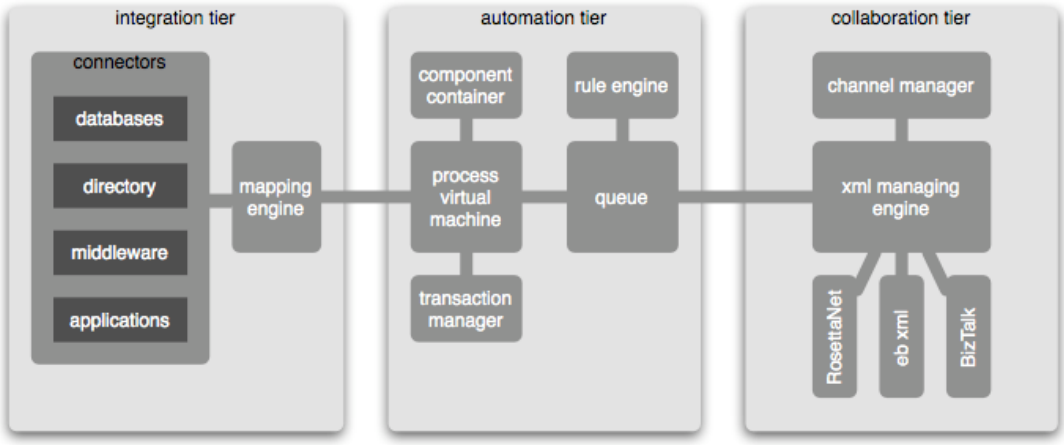

Figure 7.1 3-tier, BPMS architecture (Simplified), (Smith and Fingar, 2003)

The works from Aalst and van Hee, and Smith and Fingar highlight the need to automate the generation of workflow and access policy definitions from business requirements. Bussler and Jablonski (1995) propose a user interface to create access policies from workflows. Chandramouli (2000) and Mendling et al. (2004) describe methods to derive access control policies from process descriptions. Smith and Fingar describe a *Process Integration Environment* for business process developers to specify the mapping of high-level business process modelling, deployment and management to workflows. None of the examples address collaborative access to temporal assets. They also do not cite any example methods for mapping business requirement to workflow. In practice, this is current research that requires rules to link the business process domain; collaborative workflow integration and service oriented computing. Such mapping rules would be best supported by semantic ontologies with inference processing rules to automate mapping and validation between domains (Shadbolt et al., 2006, Beco et al., 2005, Chen-Burger et al., 2004).

The concept of the Workflow-Team Policy Architecture has been based on secure sharing of stateful service instances deployed in GT3. The more recent development of GT4 changed from the stateful service model to the Web Services Resource Framework (WS-RF). The concept describes state as a property of the resource that is exposed by a Web Service interface. The main impact on the Workflow-Team Policy Architecture is that all stateful services created from a deployed service factory in WS-RF use the same URL, unlike in GT3 where each Grid Service instance has a unique URL. This change means GT4 services can be integrated using the workflow language BPEL. The 'instance' resource is now located using WS-Addressing. Therefore the addressing mechanism is the only change and the instance would be recorded in the Workflow-Team Policy Architecture as a shared temporal business asset.

One of the original concepts for using Grid Computing in the DAME scenario described a requirement for VO support for delegation of rights to execute processes across grid compute resources. The DAME business process shows that when the ME from the airline launches the diagnosis process a service can operate on the user's behalf but the user does not have permission to view the output result, because it may reveal sensitive data. Another example is revealed in the interview with Dibsdale is when benchmarking data is released to the airlines. This is obtained by aggregating data across all fleet operators but without exposing a particular operators data. An additional workflow would allow airlines to access the results of an aggregation process but not see any intermediate data. This highlights the need to protect access to a service instance within workflows, supporting the 'business firewall' model proposed by Smith and Fingar and is also highlighted in grid security models by Norman (2006).

7.3.1 Comparison with Workflow Management Systems

The following table is a list of features used in workflow management systems and indicates which features are implemented in major commercial and research-based workflow management systems. As can be seen, some service based workflow systems, such as Triana and MyGrid Taverna did not implement security in the workflow engines. The most complete is the commercial package from IBM Websphere that includes many add-on tools for modelling and monitoring. However the Workflow-Team Policy Architecture is the only system to provide dynamic collaborative access control, which is essential in offering service integration in a competitive global market.

Table 7-1 Comparison of Workflow Management Implementations

Feature\Workflow System	Workflow-Team Policy Architecture	IBM Websphere	Triana	MyGrid Taverna	Gridbus (GWFE)	IBHIS
Cross-organisational security	Yes	Yes	No	No	Yes	Yes
Dynamic automated access control	Yes	Dynamic, not automated	No	No	No	Partly, fixed policies reference semantic data descriptions
Single-sign on	Yes	Yes	No service security	No service security	N/A	Yes
Role-based access with additional attributes	Yes	Yes	No service security	No service security	No	Yes
Collaborative secure access (Team-based)	Yes	No	No	No	No	No
Large user base	Yes	Yes	N/A	Yes	Yes	N/A
Collaborative access to service instances	Yes	N/A	No	No	No	No
Large number of collaborative workflows	Yes, but limited testing	Yes	No	No	No	No
Secure document parts	No	Yes	No	No	No	Filtering of data results
User interaction / steering	Yes, limited to interaction with service instances	Yes	Yes	Limited	Yes	No
Process modelling and exchange	Java class can be dynamically imported to another DAME deployment	Yes (BPMN & BPEL)	Yes (export & import BPEL)	Yes, between Taverna using proprietary SCFL	Yes, using proprietary xWFL	Implements a single process to execute submitted queries

7.4 Comparison with Access Control Developments and standards

This section compares the Workflow-Team Policy Architecture with areas of authentication, authorisation, specific development for security in GT4 and methods for delegation of rights. Comparisons are made to complementary systems that could be used in the implementation of the Workflow-Team Policy Architecture. These complementary systems include the Akenti policy decision engine (Thompson, 2003), the PERMIS policy decision engine (*Modular PERMIS Project*, 2006) and the Shibboleth identity service (Internet2, 2006).

7.4.1 Authentication

The Workflow-Team Policy Architecture uses X.509 certificates for identity assertions, for users and services, but could be implemented independent of the authentication format. The Akenti system is limited to using X.509 for identity. However, the PERMIS policy engine is identity format agnostic and relies on the application to validate identity, such as verifying the Certificate Authority (CA) signature in X.509. Shibboleth can be used to issue identity assertions and uses SAML to provide the identity assertions in different formats including X.509 and can include attributes attached to the identity. In Shibboleth, the release of user attributes and identity can be controlled by the context of the request to improve privacy of a user's identity and attributes.

Shibboleth could support the assertions for identity of users connecting to a workflow management system executing the Workflow-Team Policy Architecture (as illustrated in Figure 5.4). Specifically in the DAME example, the assertions for user identities requires the additional attributes of the user's organisation and role identity, which in turn requires assertion that the user conforms to the required qualifications. The issuer of the user's identity (the user's employer) could assert the role on trust, alternatively, the issuer could be required to provide assertions of the qualifications along with the role assertion. The policy decision point would assess the claim for the user's role based on a policy that states required qualifications and a list of organisations that can claim that role.

Authentication across a large VO can cause concerns of privacy, as shown in the DAME example. Shibboleth offers methods of anonymity in the use of remote resources. The method of authenticating with a trusted external server could issue the user with an anonymous assertion about their attributes. This has been developed so that universities can issue assertion that a user is a member of the university and is permitted to use the remote resource, without the resource owner needing to know who the person is. The trust model in this scenario works in the commercial sector, since the use of a resource would involve some cost to the asserting organisation. However, as with the DAME scenario, there are requirements for non-repudiation of a user's actions.

Shibboleth can provide another method of authentication to provide traceability of anonymous actions. The user is issued with a pseudononymous identity. This would be a randomly generated unique identity that is recorded by the issuer, but the name in the identity has no value to an external organisation. Except that the external organisation can record the actions that can be traceable to a person if needed. In DAME, the organisation and role is probably the most important identifier, rather than user identity. The privacy requirements for DAME would need to be explored further to consider future work on identity management.

7.4.2 Authorisation

In RBAC (Ferraiolo et al., 2003a) the role assignment is considered temporal by using a session to control the user to role assignment. The role assignment tends to be linked to a higher-level agreement, such as job specification and therefore employment. TBAC concentrates on using RBAC in secure workflow management to create just-in-time permissions, thereby reducing vulnerabilities. The Workflow-Team Policy Architecture addresses further to these by managing dynamic policies for a workflow. This is runtime management of changes as users acting in roles join or leave the workflow and their permissions to the assets that are created or destroyed during the workflow.

By moving the policy decision to the collaborative point of access (the workflow), services need not authenticate individual users. The service only needs to authenticate the identity of the workflow engine. The DAME demonstrator illustrates this point by showing where the users' identities are interpreted, and then the workflow accesses the services with its own certificate.

It would be possible to use Akenti or PERMIS in the Workflow-Team Policy Architecture to interpret the dynamic policies, by configuring them to retrieve the policies on each request and provide policy decisions based on attributes. The proposed solution would need to manage the dynamic policies to record role assignments, service instances as temporal assets and create subject-action-object predicates according to the workflow constraints.

7.4.3 GT4 Grid Security Infrastructure

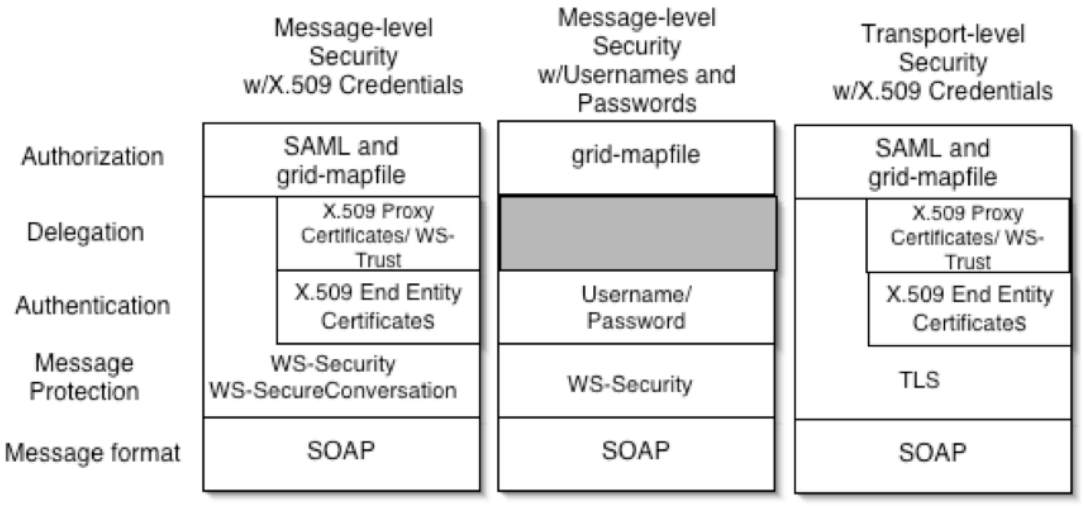

Figure 7.2 Overview of the GT4 Grid Security Infrastructure, (The Globus Security Team, 2005)

Developments of new grid technologies that occurred during this research included the release of GT4. The Grid Security Infrastructure (GSI) in GT4, illustrated in, Figure 7.2 contains the following areas:

- Communication security, using transport-level and message-level security. Transport-level security is proprietary to Globus Toolkits and is to be deprecated. Whereas, message-level security complies with WS-SecureConversation (Della-Libera et al., 2002) of secure message exchange;
- Authentication, using X.509 certificate or username and password;
- Delegation, which supports single-sign on by using X.509 proxy certificate, when authenticating with X.509 certificates. This conforms to WS-Trust;
- Authorisation uses either SAML authorisation assertions from an external policy decision engine such as CAS or the GSI uses the GT3 model of grid-map file, which maps user identities to local user accounts, and inherits the user's rights on that resource.

Additional methods of authentication and authorisation not part of the GT4 distribution can be included in a customised installation. Code hooks that allow security extensions have been placed in the Grid Service container. The container is the host environment for Grid Services. However, this would make authorisation mechanisms non-standard and a possible barrier to integration of external service providers. External service providers would be required to implement a security mechanism unique to the customer's requirement, a barrier to reselling a service. However, the security requirement could also

be a business enabler by defining a standard for suppliers, stated by Dibsdale (Appendix B.2).

7.4.4 Delegation

This section briefly considers the work in PERMIS on delegation. The development of PERMIS policies definition and management has included a method for specifying a delegation policy for the Delegation Issuing Service (Chadwick, 2005). The policy defines the delegation permissions a user has in order to delegate their rights to another user, specifying which actions on objects can be delegated to which users. The policy can be written with coarse-grained constraints, so that users and objects can be defined by attributes, such as role for a user or type for an object. For example, using the DAME diagnostics process, the role ME can delegate the right to read the results of WF1 to the MA. The delegation of rights must be fine-grained, specifying users, actions and objects. The delegated rights can be constrained by properties such as a time period.

The delegation policy performs a similar function to the Workflow-Team Policy Definition. However, examples of the Delegation Issuing Service in operation show manual methods of delegation, by using a browser interface to demonstrate temporarily granting access to documents in a structured VO (such as pre-defined roles, permissions, object types). The DAME scenario requires the changes in access permission to automatically update on changes in the team membership, including users and service instances. Therefore, the Workflow-Team Policy Architecture is designed to automate dynamic changes in fine-grained permissions and would be complementary to provide contextual process control of PERMIS delegation rules.

7.5 Policy Architecture Considerations

One of the targets for the Workflow-Team Policy Architecture is scalability across many users, many service providers and many concurrent collaborations. Part of the solution addressed the architecture of policy distribution. VO policy solutions implement a distributed policy architecture that interprets multiple policies on each resource access. Some issues when using multiple policies in access decisions are:

- Communications overhead (taking time and bandwidth) when reading policies from multiple locations on every service request, to ensure that interpretation is up-to-date and includes changes and the evaluation of temporal rules;
- Out-of-date decisions may be caused by caching policies to speed decision making by avoiding communication overheads;

- Trust relationships have to be established between every consumer and supplier in a truly vertical VO, if user credentials are used to authorise access to services;
- Privacy of permitted actions, or resources in use may be compromised in common access to VO policies, a concern in DAME since it may reveal process information and operating procedures;
- A remote policy that is included in a VO policy decision may be capable of hijacking the decision and denying particular users or organisations (Chadwick and Otenko, 2003).

The Workflow-Team Policy Architecture addresses this by separating the end user from the service supplier, with the WFMS hosted by the integration organisation controlling access with a single policy. Services are free to restrict access with their own policy, but the Workflow-Team Policy Architecture reduces the requirement to access multiple policies at the point of service access.

The Workflow-Team Policy Architecture requires the management of dynamic access rules. PERMIS uses distributed management of attributes and roles, but requires tools to manage distributed access conditions. Akenti does support distributed policies where each stakeholder can manage its own dynamic policies. However Akenti does not control who can state which constraints, therefore any one of the stakeholders can deny permission to others.

The Workflow-Team Policy Architecture integrates access control and workflow management, but retains the separation of concerns of the two mechanisms. Periorellis et al. (2006) proposes a coordination layer distributed in the middleware to control access and police electronic contracts, combing access constraints, quality of service requirements and task-context. This provides a coordination layer on service access that also records a non-repudiable log of activities across a VO. However, it does not state how the peer-to-peer style contracts are managed for dynamic fine-grained access control, and how privacy of actions can be protected in a commercially sensitive environment.

7.5.1 Experience

During the building and testing of the demonstrator, the workflow management system integrated different services developed by the different partners. During this time, differences in deployment platforms had to be taken into account in the development of the workflows. Despite the use of common middleware, some components were not common on all platforms. The effect of these differences caused complications in the building the workflow management system to cope with different standards of interface.

The use of collaborative access control policies at each deployment platform would cause further compatibility issues. If the business model uses loose coupling of services, each service provider would need to implement the collaborative policies and dynamic policy management control interfaces. This would reduce the overall flexibility to incorporate any service provider. However, Dibsdale remarked that requirements such as security implementations could be part of qualifying standards for services to be DAME providers.

7.6 Limitations of approach

One of the risks in this research is that the DAME scenario is limited in producing and evaluating a generalised solution. There are significant parallels in access control requirements in collaborative use of distributed systems with other publications this work has been compared to, despite the use of different domains and scenarios. Similar scenarios in collaborative processing have used the medical domain (Chandramouli, 2001); others have used typical business processes such as purchasing (Papazoglou and Dubray, 2004), or academic processes such as student admissions (Liu and Chen, 2004). The Workflow-Team Policy Architecture could be applied to these other domains, in particular the collaborative problem solving in the medical domain.

The workflow model used in the research is limited to the example diagnostic collaboration from DAME. This provides a rich example of cross-organisational use of user expertise, with automated decision support, use of sensitive data across organisations and a requirement for Grid Computing. It does however only show a limited view of workflow problems. The workflow scenario is relatively simple and exercises few of the workflow control patterns (Aalst, 2006). An interesting extension to the access control requirements for collaboration would be to incorporate conflict of interest, separation of duty and role hierarchy, such as reported in Ferraiolo et al. (2003b).

This research in this thesis addresses the case for centralised management of workflows and access control. It has been shown that the business model supports the control at the point of integration, however there is work on modelling and validating decentralised coordination and mediation in workflows (Wombacher et al., 2005) and VOs (Periorellis et al., 2006).

Integration of workflow in the portal did not explore all complexities in providing access to external components. Such as, methods of authentication, and subsequently authorisation by this type of access were not explored. However, the connection point should be the same as the portal session-based connection, providing the identity of the user and their role (and qualifications to authenticate role assert). This would define the access

control in terms of the 'business firewall' discussed in section 7.3, from Smith and Fingar (2003).

Since the Workflow-Team Policy Architecture has concentrated on automation of permissions based on workflows, it requires that all activity take place within workflows. Whilst this is a requirement for most business procedures, it does constrain collaborations to fixed or semi-fixed types. As stated by Hesketh, the access control mechanisms should not constrain collaborations that would benefit the business. The Workflow-Team Policy Architecture does not support ad-hoc workflows, since they would not link to a Workflow-Team Policy Definition. A general access control policy could be used to enforce access in ad-hoc workflows.

Another limitation of the Workflow-Team Policy Architecture that it does not manage services instances after the workflow finishes. The architecture assumes that workflows have been defined so as not to leave hanging instances or that a clean-up mechanism is implemented to ensure remote service instances are destroyed at the end of the workflow. In support of ad-hoc workflows, it would be necessary to extend the architecture to the control of access to Grid Service instances beyond the lifetime of the workflow in further collaborations. Methods to achieve this would require extracts from the policy that define access for a particular service and can be attached to that service. This is similar to methods of flow control discussed in a later point.

A limitation on the technology meant the experiment uses GT3. This is the result of the required collaboration between the DAME project partners and the need to use a common grid platform running on the White Rose Grid. Current grid technology from the Globus Alliance is GT4. The differences between the two versions has been described (Harmer and McCabe, 2005, Silva, 2005, Harmer et al., 2005). Required changes for the Workflow-Team Policy Architecture to use GT4 appear to be minimal. The interface protocol for the management of Grid Service instance has changed, however the concepts have not. Therefore, the impact is likely to be mostly syntactical changes to the implementation and would not affect the proposed architecture.

In this work, it has been assumed that all access to information and services are performed using a web-based portal, and that data retrieval is 'eyes only'. In practice, the web output may be printed, or subsequent implementations may allow the user to save results or raw data to a remote system. Access control to confidential data outside the side boundary would require complex techniques such as flow control. Flow control tracks and regulates the dissemination of sensitive data, and thus guarantees only authorised information flow (Denning, 1976).

> *Access controls are insufficient to regulate the propagation of information after it has been released for processing by a program. Similarly, cryptography provides strong confidentiality guarantees in open, possibly hostile environments like the Internet, but it is prohibitively expensive to perform nontrivial computations with encrypted data. Neither access control nor encryption provides complete solutions for protecting confidentiality. (Zdancewic, 2004)*

This work considered access control within the boundary of the business system by ensuring data could only be viewed from recognised applications, identified using the authentication mechanisms with the Workflow-Team Architecture.

7.7 Future Work

This section discusses some areas in which further research would improve the Workflow-Team Policy Architecture contribution.

One area of research, highlighted as incomplete is the automated generation of workflow and access policy definitions from business process requirements. This would extend pervious work (Mendling et al., 2004, Chandramouli, 2000) and assist in agile business process management (Smith and Fingar, 2003). The third wave of business process management (Smith and Fingar, 2003) introduces agile methods of running business processes by ensuring the integration of process definition, monitoring and change control. In the Workflow-Team Policy Architecture this would require automatically reflecting changes to the Workflow Definition in the Workflow-Team Policy Definition.

An issue raised in the interview with Dibsdale related to the changing level of service delivery. Different levels of service delivery (and cost) provide different levels of performance from the services. Dibsdale suggested coarse levels for the marketing of services, such as gold, silver and bronze. These quality of service levels can include time to produce a result, accuracy of results and functionality. As new services are developed these are charged at the premium level and the previous premium service may be demoted to the standard package. As stated changes in the service used in workflows need to be reflected in the access policies. However, appropriate descriptions of services may provide another solution.

New description methods could lead to further work in workflow descriptions. Using abstract descriptions of services, loose coupling between service implementations and workflow definitions can be maintained. If services are described using abstract descriptions of functionality and quality of service levels, this can reduce the cost of maintenance due to service evolution and lifecycles. If a workflow is defined as the premium ‘gold’ level, then

premium level services can be assigned to 'gold' as an attribute in its description. When this is replaced by better functioning services, it loses the gold status and may be automatically offered in the lower level workflows. Loose coupling between workflow definition and service implementation is achieved. However, research is required to understand appropriate descriptions of quality of service and to define methods to monitor that the required level has been fulfilled.

An alternative method of describing services in workflows would be to discover service implementations by semantic definition. Semantic descriptions of the requirements of a workflow could be used to locate the services at runtime. The semantic description could specify the types of data to be processed and the type of process, along with quality of service requirements, such as processing time, accuracy and security. Semantic services is currently an expanding area, as can be seen in the activities on the Semantic Web (Shadbolt et al., 2006) and the Semantic Grid (*The Semantic Grid*, 2003).

The Delegation Issuing Service (Chadwick, 2005) from PERMIS is a possible target for future work to the Workflow-Team Policy Architecture. An investigation would be required into automating the delegation process and how to automatically include new service instances into fine-grained access. The investigation would need to include how to link the delegation policy to the workflow definition and how to restrict delegations based on the workflow context.

The Delegation Issuing Service could also be used to expand on the limitation that the Workflow-Team Policy Architecture does not support ad-hoc workflows. The PERMIS decision engine could use the Workflow-Team Policy Instance and combine it with a general access control policy, the definition of which could support delegation of rights. An interesting exercise would be to see if the Delegation Issuing Services could be used to automate permissions in ad-hoc collaborations, sharing temporal business assets, such as long-running Grid Service instances.

Ongoing work to support secure collaborative workflow for aircraft engine diagnostics will continue in the DAME follow-on project BROADEN (Business Resource Optimisation for After market and Design On Engineering Networks) (ComputingLeeds, 2005).

7.7.1 Extensions using complementary approaches

The identity management in this work could benefit from a restricted release policy on attributes during workflow. The Workflow-Team Policy Architecture uses a role template, where a user must conform to all the requirements to have permission to enact the role. This typically includes their level of qualification. If certain tasks require only partial conformance to the role template and an identity management system contained all of a

users attributes then a policy decision could be made by retrieving only relevant attributes. Koshutanski and Massacci describe this method as the abduction of credentials (2003). This is mainly designed for "mobile" processes that execute across different workflow engines. However, the DAME example could benefit from this by having a more flexible approach to attribute management and the requirement of not revealing the details of the process, due to the process definition being a valuable asset to the service integration organisation (DS&S).

If the Workflow Management System were refactored and redeployed in the DAME demonstrator it would be an opportunity to incorporate standards that have been defined during the building of the original demonstrator. These standards would be:

- BPEL, for the definition of the workflow. To use this would require all the DAME services to be re-implemented using GT4.
- SAML, for identity assertions from remote systems to connect to the workflow management system. The connections would be either from the portal, or from an interface application run by an external organisation.
- XACML, for the definition of policies, although this would need to be compared with the capabilities of using PERMIS policies.

As shown in the previous section, the Workflow-Team policy can be applied to the DAME demonstrator. In order to satisfy the refined objectives, the next stage of work involves an implementation to modify the DAME Portal and Workflow Management System to use the Workflow-Team policy. This also presents the opportunity to utilise some of the rising standards along side the dynamic policy, such as:

- Replacing Globus Toolkit 3 with Globus Toolkit 4 and therefore use the resource model in WS-RF (OASIS, 2005) instead of grid-services;
- By using WS-RF resource model:
    - Service Instance references in the Workflow-Team policy become WS-Addresses (W3C, 2004), which is a change in format but not concept;
    - BPEL, which wasn't suitable to integrate GT3 services, can be used to establish Workflow Definitions.

An initial feasibility study into moving from Globus Toolkit 3 to 4 reveals this to be a logical progression, with minimal architectural changes.

The re-implementation provides the opportunity to investigate distributed identity and user attribute management by using Shibboleth, with SAML assertions. Shibboleth supports the distributed authentication of users required to make DAME scalable across many

customers. SAML assertions are passed from the authenticating organisation stating the users ID, the roles they can assume in DAME and the organisation they belong to. An additional policy would validate that the organisation can assign users that role, it would also be used to add attributes such as the level of service that organisation has paid for. This policy could be expressed using XACML (OASIS, 2003).

In the Workflow-Team Policy Architecture SAML could be used to communicate user attributes. In the WFMS the policies could be expressed in XACML. Communication of access requests and decision responses could use either SAML or XACML. This cross over of the technology standards is subject to further investigation. Also, when implementing the architecture using policies expressed XACML it would be possible to evaluate the use of Akenti or PERMIS policy decision engines and how they manage dynamic policies. Another point of interest would be to compare how the team constraints are dynamically recorded in Workflow-Team Policy Instance with the management of static and dynamic constraints at runtime, as in Bertino et al. (1999).

Another important component of the system is the management of workflow definitions and the generation of the Workflow-Team Policy Definitions. This is a requirement to support changes in the business requirements, and subsequent changes in workflows and roles. One approach would be to extend the work in Mendling et al. (2004) to automate the production of XACML role based policies from BPEL workflow definitions. Like Mendling et al., this could be performed using a XLST script to transform from BPEL to XACML, mapping from the activities assigned Partner and PartnerRole in BPEL to service activation constraints for roles in the Workflow-Team Policy Definition.

7.8 Summary

This chapter presents the results of qualitative evaluation of the Workflow-Team Policy Architecture and the implemented DAME demonstrator. The interview results of the industrial experts are presented in section 7.1, which provides useful feedback on how the business model corresponds to the VO model in the DAME implementation and how the implemented method for controlling dynamic access control for teams matches the business requirements for DAME. Sections 7.3, 7.4 and 7.5 compare recent developments of similar and complementary work against the proposed Workflow-Team Policy Architecture and the implemented demonstrator. Section 7.3 addresses work in workflow and business processing. Section 7.4 addresses the access control areas of identity management across organisation, other authorisation system, the security infrastructure in GT4 and the Delegation Issuing Service for PERMIS. In section 7.5, considers aspects of scalability and dynamic control in the Workflow-Team Policy Architecture. Section 7.6 discusses the limitations of the research, providing a critique of the scenario based approach and the

proposed solution. Finally, section 7.7 proposes how continuing research can extend the work presented here.

Chapter 8

Conclusions

The content of this thesis is summarised in this final chapter. Section 8.1 provides a summary and conclusions of the main findings. In section 8.2, the salient points of the thesis chapters are presented. Section 8.3 lists the contributions of the research, conveying the implications to business processing and e-Science.

8.1 Main Conclusions

This thesis has presented the Workflow-Team Policy Architecture as a solution to address the problem of providing secure collaborative workflow across different organisations. In pursuit of this solution, the research combines the topic areas of collaborative workflow and business processing, Grid Computing, service-oriented architecture and role-based access control.

8.1.1 Summary of Workflow-Team Policy Architecture

The Workflow-Team Policy Architecture has been derived from analysis of previous related work and the requirements from the motivating business scenario from the DAME project. The example scenario shows teams of users in a collaborative workflow sharing Grid Service instances, which make use of commercially sensitive services and data. The security requirements illustrate the need for fine-grained access control for the collaborative team. The team is defined as the users, Grid Service instances and data that exist for the duration of the workflow.

The Workflow-Team Policy Architecture allows users to share temporary instances in secure collaborations within Virtual Organisations. This cannot be achieved in standard approaches to access control. Many variants of access control lists and security policies provide the ability to specify a number of users access to the same asset. The asset may be a service, data file or record in a database. Roles in RBAC allow users to be changed without affecting policy definition and access control rules can be defined in terms of business rules.

Previous work has identified conditions where RBAC is insufficient in defining access rules, requiring addition attributes such as context, or membership to a team. This thesis identified that role-based business processes can be used across organisational boundaries, and derived access control from such processes require extension to protect access by competing organisations.

The architecture describes the static components of the architecture that are linked to the business rules. The static components are the workflow definition, the policy definition, the roles and associated rules for users to enact roles and execute tasks in the workflow.

The architectural description shows the dynamic components of the architecture that have temporal existence during the lifetime of the workflow instance. The dynamic components are the instances of workflow, dynamic policy and service instances. Using the static definitions the dynamic parts are controlled by the workflow engine. The dynamic policy contains the temporal rules for team members to collaboratively access the temporal assets.

The Workflow-Team Policy Architecture uses a centralised policy instance to dynamically manage and control access to the temporal grid services created and shared by users from different organisations during the workflow. This is created and controlled from the Workflow-Team Policy Definition which contains the static rules. The policy definition is linked to a workflow definition. The executing workflow is dynamically managed in the workflow engine using a Workflow Instance, which is linked to and manages the Policy Instance. This linkage is described in section 5.3.

The architectural solution uses a centralised policy instead of a distributed on to reflect the business concerns of revealing valuable process information. Such information could be used by competitors to discover operating practice, and customer-supplier relationships. A distributed policy across the members of the VO was rejected since it opens visibility of the complete supply chain. A VO policy for fine-grained access control would reveal end customers and end suppliers. For DS&S, the most valuable parts of their business are the process information, customer confidentiality and protection of operating conditions. In the DAME case study, the workflow management system is the point of collaboration between users and the point of integration of grid services. Therefore, the Workflow-Team Policy Architecture is designed to control multiple users access to long running grid services with the framework of a workflow management system.

This research has extended previous work on access control and role-based business processing by defining an architectural approach to controlling dynamic policies for teams of users to share temporal assets created and consumed within a workflow. The Workflow-Team Policy Architecture responds to the business requirement of DS&S to protect

customer data. The research illustrates some of the issues with cross-organisational collaborations and methods to share competitive services and data without exposing valuable operating conditions.

The Workflow-Team Policy Architecture has been evaluated, in the DAME project demonstrator, by implementing a role-based portal and workflow management system controlling GT3-based Grid Services instances across the White Rose Grid. The analysis of the architecture has illustrated that the architecture can be independent of implementation technology and is capable of managing dynamic access control to GT4 services.

8.1.2 Summary of Limitations

The business model in the case study was limited to protecting access to services and processes, but did not extend to the commercial sensitivity of the returned data. In the Workflow-Team Policy Architecture the sensitivity of the data returned by the service is derived in the capture of the business process and business rules. An extension to the Policy Definition and Policy Instance is required to dynamically process security levels of return data using meta-data provided by the services, such as in IBHIS (Kotsiopoulos et al., 2003).

The business case study scoped the research to investigating a limited range of business processes across organisational boundaries. Due to the speculative nature of the case study, the business processes exercised a few of the control-flow patterns from van der Aalst (Aalst, 2006). The case study would need to be expanded beyond the diagnosis workflow to test the Workflow-Team Policy Architecture in managing the different workflow patterns, assuming the expanded case study uses other control-flow patterns.

8.1.3 Summary of Future Enhancements

Adaptive processes and security play an important part in coping with change by innovation and different customer usage. The Workflow-Team Policy Architecture adapts using a well-defined process and role-based access policy derived from the business requirements and rules. For the implementation of the demonstrator, the defined process and access control policy were manually created from the analysis of the business requirements. Future work would provide methods to automate the link between the business requirements and the defined process and policy. The automated link may generate the process and access control from business rules, such as in Mendling et al. (Mendling et al., 2004), or there could be a more dynamic linkage as in Smith and Fingar (Smith and Fingar, 2003).

An emerging method to achieve automation uses semantic descriptions, from Semantic Web and Semantic Grid (Shadbolt et al., 2006, *The Semantic Grid*, 2003). Semantic descriptions of services and return data can be used to decouple process descriptions from service descriptions. Services can be replaced easily by discovering

descriptions of the service function. Security protection of the returned data can be automated using semantic meta-data describing the security level of the returned data.

Since the Workflow-Team Policy Architecture is a model of workflow and security, it can be implemented using different technologies. Enhancement of the DAME Demonstrator would be to employ recent developments in technology with the framework defined by the architecture. Current appropriate technologies would be:

- PERMIS policy decision engine and delegation issuing service.
- Shibboleth authentication service.
- BPEL process language.
- Secure grid services Globus GT4 using WS-ResourceFramework.

8.2 Summary of Thesis

In this thesis, the research approach gathered information, identified a problem, designed and built a solution to address the problem and evaluated the solution. The major contributions are enumerated in 1.4 and the research methodology is presented in 1.5.

Chapter 2 contains the background information of the topic areas in this research. Section 2.1 to 2.4 introduces the concepts of collaborative working and workflow management in business processing. To support collaborative work across organisations, section 2.5 introduces SOA discussing the properties of loose coupling. This illustrated that workflow definitions, linked to business requirements, can remain static, whilst deployed services can change, supporting dynamic deployment across grid resources, integration of services from different suppliers and evolutionary change in services to respond to customers needs.

Grid Computing, in section 2.5.6, provides dynamic resources and virtualises the management of available compute processing and data storage in a common middleware. By supporting globally distributed collaborations of users, services and resources, Grid Computing is instrumental in forming VOs.

Security is introduced in section 2.6 and essential in maintaining properties such as privacy, confidentiality and availability. One aspect of security is Access Control is presented in section 2.7. This section describes some of the solutions to defining and controlling access control policies in distributed computing, illustrating the use of Roles in access control (RBAC), and how it relates to roles in processes and workflows. Extensions to RBAC use tasks and teams to create a context for controlling access.

Chapter 3 introduces the DAME project and illustrates the business case study example, collaborative problem solving for aircraft engine diagnostics. In a competitive environment, competitors share engine operating data to improve the accuracy of engine diagnosis. Section 3.3 presents the VO model for the collaboration and the diagnosis workflow in section 3.4. However, only privileged roles can access the competitor's data, that is distributed, large in volume and consumed by distributed processing services supplied by further third parties. The business requirements for security and scalability are given in Chapter 4.

Chapter 5 presents the architecture of the proposed solution, called the Workflow-Team Policy Architecture. The issues raised in Chapter 4 are analysed and compared with the background in securing collaborative workflows. The resultant solution uses role-based access control with role-based workflow management to control collaborative access to remote Grid Services, by defining the workflow as the point of collaboration, forming a team of users, services and data. The Workflow-Team Policy Architecture is presented in two forms. The first, is a high-level model about how the workflow definition and access control policy definition are linked, leading to a coarse-grained access control policy that is used to control a fine-grained policy instance or each workflow instance. The second form shows the architectural links between the static parts, Workflow Definition and Workflow-Team Policy Definition, and the dynamic parts, the Workflow Instance and the Workflow-Team Policy Instance. The scope of the architectural solution is bounded by the assumptions list in section 5.4.

Chapter 6 presents the implementation of the Workflow-Team Policy Architecture in the DAME Demonstrator, a web-based portal to access a workflow management system that executes the secure collaborative workflow and consumes GT3 grid service instances. Section 6.1 provides an overview of the test cases executed in the implementation. In section 6.2, the implemented demonstrator is described, including a secure web-based portal, the workflow management system and an outline of the deployment. This was evaluated by the analysis of interviews with two industrial experts. The results are presented in Chapter 7.

Section 7.2 summarises the interview results in which the feedback on the DAME implementation confirms that the implemented method for controlling dynamic access control for teams supports the DAME scenario test case. Sections 7.3, 7.4 and 7.5 compare the proposed Workflow-Team Policy Architecture and the implemented system with other work in the field of workflow, business processing and access control. Section 7.6 discusses the limitations of the research, providing a critique of the scenario based approach and the proposed solution. Section 7.7 proposes how continuing research can extend the contribution of this research.

8.3 Summary of Contribution

8.3.1 Summary of Research Objectives

The research objectives are been summarised in the following points and referenced against the contributions C1 to C7 listed in section 1.4:

- Contribution C4, the design and build of an experiment to study a workflow management system to execute and control long-running Grid Services across a multi-organisational grid, allowing shared access by users from different organisations. The research tested the constructs of the Workflow-Team Policy Architecture by implementing a workflow management system (WFMS) and workflows that execute GT3 Grid Services. The WFMS integrates with the portal to provide role-based access control to the workflow and Grid Services that is dynamically managed to specify fine-grained access of users to Grid Service instances created and deployed by external organisations.
- The investigation of issues of access control in a VO for collaborative access to services and data, where some services and data are commercially sensitive and could reveal proprietary information to competing members within the VO. The DAME VO business model (contribution C5) illustrates a case where organisations share services and data to achieve business goals, but require protection of the service and data assets. The security requirements illustrated have been incorporated into the model for Workflow-Team Policy Architecture, tested in the DAME demonstrator and evaluated by qualitative interviews with industrial experts.
- Contribution C3 is the general model for the provision of dynamic fine-grained access control to stateful Grid Service instances used in collaborative teams of users, services and data from different organisations. Contributions C1 and C2 fulfilled in the Workflow-Team Policy Architecture. This solution is the result of contribution C6, the analysis of the business case, comparison with other state of the art solutions and experience building the workflow management system with dynamic access control. The research has presented the workflow as the collaborative point of access to Grid Service instances. In order to manage collaborative access control in a scalable solution, across organisational boundaries, requires management of many users and dynamic collaborative teams using coarse-grained access control policies. Therefore, to achieve fine-grained access control to Grid Service instances across organisations, the workflow instance is used to coordinate execution and manage a dynamic policy instance. The role-based Workflow Definition is linked to a Workflow-Team Policy Definition for high-level definition Of policy, which is executed as fine-grained policies to protect access to service instances by permitted members of the collaborative team.

- Contribution C7 is fulfilled by the evaluation of the access control for secure service-based collaborative workflows using the business example from the UK e-Science DAME project. The business VO and collaborative workflow was defined by working with representative from Rolls-Royce and DS&S, and in collaboration with the definition of the DAME use cases (Fletcher, 2002). The business case study is study is presented in chapters 3 and 4, fulfilling contribution C5. The DAME project provided the opportunity to present the concepts and implementations for access control to the Grid Computing environment, gaining useful feedback in developing the resultant Workflow-Team Policy Architecture.

8.3.2 Implications to Business Processing

The Workflow-Team Policy Architecture is an important step in the use of business workflows across organisations. The solution addresses the requirements for the a business to supply a diagnostics support service to many competing customers, using secure access mechanisms to consume Grid Computing services from third party suppliers and benefit from collaborations whilst protecting commercial interest. To quote Dibsdale:

> *It is absolutely not allowed for one airline to see the actual performance of a rivals operating fleet. So that would be the fastest thing that would shut us down as a service business, that you could imagine. A lot of our effort is to ensure that we don't actually do that*. From Dibsdale, Appendix B.2.

8.3.3 Implications to e-Science

There is increasing awareness in the growing scientific use of HPC and large-scale clusters in Grid Computing. The use of Grid Computing supports not only large scale computing, but also large-scale collaboration, similar to the collaboration discussed in this research. Therefore, e-Science use of Grid Computing demands controlled access to HPC and executing services and presents situations where collaborative use of executing services is required. For example, in collaborative visualisation in the gViz project uses dynamic allocation of grid processing to render images that can be steered and viewed by different users simultaneously (Brodlie et al., 2004). The Workflow-Team Policy Architecture can provide the means to control collaborative access to grid services across the VO, by using workflow to mediate collaboration and automate team membership.

UK e-Science projects, such as myGrid and Triana (*UK e-Science Programme*, 2005), show that service and workflow definitions can be shared between users. This research extends the capability by illustrating the collaborative use of active workflow instances. Recent projects have also extended methods of workflow and security in collaborations. For example, controls on service usage in VOs by automated contract negotiation and monitoring in the Gold project (Periorellis et al., 2006). The Workflow-Team Policy

Architecture can be used in these e-Science projects to provide a complementary solution to access control, by using workflow as the context for using Grid Services and linking this to control access of Grid Services.

Bibliography

Aalst, W. M. P. v. d. 2006. *Workflow Patterns* [Online]. [Accessed 01/06/2006]. Available from: <www.workflowpatterns.com>

Aalst, W. M. P. v. d., Dumas, M., ter Hofstede, A. H. M. and Wohed, P. 2002a. *Pattern-Based Analysis of BPML (and WSCI)*. Queensland University of Technology, Brisbane. Report: QUT Technical Report FIT-TR-2002-05.

Aalst, W. M. P. v. d. and Hee, K. v. 2004. *Workflow Management: Models, Methods, and Systems (Cooperative Information Systems S.)*. The MIT Press

Aalst, W. M. P. v. d., Kumar, A. and Verbeek, H. M. W. 2003. Organizational Modeling in UML and XML in the context of Workflow Systems. In: *SAC 2003*, Melbourne, Florida, USA. ACM. pp. 603 - 608.

Aalst, W. M. P. v. d., ter Hofstede, A. H. M., Kiepuszewski, B. and Barros, A. P. 2002b. *Workflow Patterns* [Online]. [Accessed 10/04/2003]. Available from: <http://tmitwww.tm.tue.nl/research/patterns/download/wfs-pat-2002.pdf>

Access Grid 2006. [Online]. [Accessed 01/08/2006]. Available from: <http://www.accessgrid.org/>

Agarwal, D., Jackson, K. and Thompson, M. 2002. Securing Collaborative Environments. *WACE 2002: Second Workshop on Advanced Collaborative Environments*.

Alfieri, R., Cecchini, R., Ciaschini, V., dell'Agnello, L., Frohner, A., Gianoli, A., Lorentey, K. and Spataro, F. 2003. *VOMS, an Authorization System for Virtual Organizations* [Online]. [Accessed 15/01/2004]. Available from: <http://grid-auth.infn.it/docs/VOMS-Santiago.pdf>

Alonso, G., Casati, F., Kuno, H. and Machiraju, V. 2004. *Web Services - Concepts, Architecture and Applications*. Springer: London

Alotaiby, F. T. and Chen, J. X. 2004. A model for team-based access control (TMAC 2004). In: *Proc. ITCC 2004. Int. Conf. on IT: Coding and Computing, 2004.*, IEEE. pp. 450 - 454.

Andrews, T., Curbera, F., Dholakia, H., Goland, Y., Klein, J., Leymann, F., Liu, K., Roller, D., Smith, D., Trickovic, I. and Weerawarana, S. 2003. *BPEL4WS v1.1* [Online]. [Accessed 11/08/2003]. Available from: <http://www-106.ibm.com/developerworks/library/ws-bpel/>

Atkinson, B., Della-Libera, G., Hada, S., Hondo, M., Hallam-Baker, P., Klein, J., LaMacchia, B., Leach, P., Manferdelli, J., Maruyama, H., Nadalin, A., Nagaratnam, N., Prafullchandra, H., Shewchuk, J. and Simon, D. 2002. *Specification: Web Services Security (WS-Security)* [Online]. [Accessed 30/01/2004]. Available from: <http://www-106.ibm.com/developerworks/webservices/library/ws-secure>

Austin, J. and et al. 2001. *Distributed Aircraft Maintenance Environment DAME: A GRID e-Science Full Proposal*. DAME Project.

Banerji, A., Bartolini, C., Beringer, D., Chopella, V., Govindarajan, K., Karp, A., Kuno, H., Lemon, M., Pogossiants, G., Sharma, S. and Williams, S. 2002. *Web Services Conversation Language (WSCL)* [Online]. [Accessed 05/08/2003]. Available from: <http://www.w3.org/TR/wscl10/>

Becker, P. 2002. *Shibboleth: Identity the Internet Way* [Online]. [Accessed 26/01/2004]. Available from: <http://www.digitalidworld.com/modules.php?op=modload&name=News&file=article&sid=90&mode=chrono&order=0>

Beco, S., Cantalupo, B., Giammarino, L., Matskanis, N. and Surridge, M. 2005. OWL-WS: a workflow ontology for dynamic grid service composition. In: *e-Science and Grid Computing, 2005. First International Conference on*, Rome, Italy.

Bertino, E., Ferrari, E. and Atluri, V. 1999. The Specification and Enforcement of Authorization Constraints in Workflow Management Systems. *ACM Transactions on Information and System Security*. **2(1),** pp. 65–104.

Bertino, E., Mazzoleni, P., Crispo, B. and Sivasubramanian, S. 2004. Towards supporting fine-grained access control for Grid resources. In: *Distributed Computing Systems,*

2004. FTDCS 2004. Proceedings. 10th IEEE International Workshop on Future Trends of, 26-28 May 2004, pp. 59-65.

BPMI 2002. *Business Process Modeling Language (BPML)* [Online]. [Accessed 05/08/2003]. Available from: <http://www.bpmi.org>

Brodlie, K., Duce, J., Gallop, J., Sagar, M., Walton, J. and Wood, J. 2004. Visualization in grid computing environments. In: (Eds., Rushmeier, H., Turk, G. and Wijk, J. J. v.) *IEEE Visualization, 2004*, 10-15 Oct 2004, Austin, Texas, USA. IEEE Press. pp. 155-162

Burr, W. E., Dodson, D. F. and Polk, W. T. 2006. *NIST Special Publication 800-63, Electronic Authentication Guideline: Recommendations of the National Institute of Standards and Technology* [Online]. [Accessed 1/07/2006]. Available from: <http://csrc.nist.gov/publications/nistpubs/800-63/SP800-63V1_0_2.pdf>

Bussler, C. and Jablonski, S. 1995. Policy resolution for workflow management systems. In: *Proceedings of the 28th Hawaii International Conference on System Sciences*, IEEE Computer Society.

Buyya, R. 2002. *Economic-based Distributed Resource Management and Scheduling for Grid Computing*, PhD thesis, Monash University, Melbourne, Australia

Cantor, S. 2004. *Shibboleth Architecture* [Online]. [Accessed 10/02/2005]. Available from: <http://shibboleth.internet2.edu/docs/draft-mace-shibboleth-arch-protocols-05.pdf>

Chadwick, D. 2005. Delegation Issuing Service. In: *NIST 4th Annual PKI Workshop*, Gaithersberg, USA. pp. 62-73

Chadwick, D. and Otenko, O. 2003. A Comparison of the Akenti and PERMIS Authorization Infrastructures in Ensuring Security in IT Infrastructures. In: (Ed. El-Hadidi, M. T.) *Proceedings of the ITI First International Conference on Information and Communications Technology (ICICT 2003)*, Cairo University. pp. 5-26

Chadwick, D. W. and Otenko, O. 2002. The PERMIS X.509 Role Based Privilege Management Infrastructure. In: *Proc.7th ACM symp. on Access control models and technologies*, Monterey, California. ACM Press. pp. 135-140.

Chandramouli, R. 2000. Business Process Driven Framework for defining an Access Control Service based on Roles and Rules. In: *23rd National Information Systems Security Conference*, Baltimore, MD. NIST.

Chandramouli, R. 2001. A framework for multiple authorization types in a healthcare application system. In: *Computer Security Applications Conference, ACSAC 2001. Proceeding 17th Annual*, 10-14 Dec 2001, New Orleans, Louisiana, USA. IEEE Computer Society. pp. 137 - 148.

Chen-Burger, Y. H., Hui, K. Y., Preece, A. D., Gray, P. M. D. and Tate, A. 2004. Supporting Collaboration through Semantic-based Workflow and Constraint Solving. *Proceedings of the 14th International Conference on Knowledge Engineering and Knowledge Management (EKAW).*

Churchill, E. F., Snowdon, D. N. and Munro, A. J. 2001. *Collaborative virtual environments : digital places and spaces for interaction*. Springer: London ; New York

ComputingLeeds 2005. *DAME to BROADEN: realising aircraft maintenance on the Rolls Royce Grid* [Online]. [Accessed 01/09/2006]. Available from: <http://www.computingleeds.ac.uk/back/issue14/current/BROADEN.htm>

Della-Libera, G., Dixon, B., Garg, P., Hada, S., Hallam-Baker, P., Hondo, M., Kaler, C., Maruyama, H., Nadalin, A., Nagaratnam, N., Nash, A., Philpott, R., Prafullchandra, H., Shewchuk, J., Simon, D., Waingold, E. and Zolfonoon, R. 2002. *Specification: Web Services Secure Conversation (WS-SecureConversation)* [Online]. [Accessed 02/02/2004]. Available from: <http://www-106.ibm.com/developerworks/library/ws-secon/>

Denning, D. E. 1976. A lattice model of secure information flow. *Communications of the ACM*. **19(5),** pp. 236-243.

Department of Defence 1987. *Trusted Network Interpretation of the Trusted Computer System Evaluation Criteria (TNI), NCSC-TG-005*. National Computer Security Center, Ft. Meade, MD 20755.

Dept. of Comp.Sci. 2004. *myGrid* [Online]. [Accessed 11/08/2003]. Available from: <http://www.mygrid.org.uk/>

Ellis, C. A., Gibbs, S. J. and Rein, G. 1991. Groupware: some issues and experiences. *Commun. ACM*. **34(1),** pp. 39-58.

Endrei, M., Ang, J., Arsanjani, A., Chua, S., Comte, P., Krogdahl, P., Luo, D. M. and Newling, T. 2004. *Patterns: Service-Oriented Architecture and Web Services* [Online]. Available from: <http://publib-b.boulder.ibm.com/Redbooks.nsf/RedpieceAbstracts/sg246303.html?Open>

Eriksson, H.-E. and Penker, M. 2000. *Business modeling with UML - business patterns at work*, John Wiley & Sons, New York.

Eshuis, R. and Wieringa, R. 2002. Verification support for workflow design with UML activity graphs. In: *Proceedings of the 24th International Conference on Software Engineering*, Orlando, Florida. ACM Press.

Ferraiolo, D., Kuhn, R., Chandramouli, R. and Barkley, J. 2003a. *Role Based Access Control* [Online]. [Accessed 26/01/2004]. Available from: <http://csrc.nist.gov/rbac/>

Ferraiolo, D. F., Kuhn, D. R. and Chandramouli, R. 2003b. Access Control Policy, Models, and Mechanisms—Concepts and Examples In *Role-Based Access Control* Artech House, Norwood, MA, pp. 27-49.

Ferraiolo, D. F., Sandhu, R., Gavrila, S., Kuhn, D. R. and Chandramouli, R. 2001. Proposed NIST Standard for Role-Based Access Control. *ACM Transactions on Information and System Security*. **4(3),** pp. 224–274.

Fletcher, M. 2002. *DAME Requirements: Use Cases*. DAME Project. Report: DAME/York/TR/02.001.

Fletcher, M., Chivers, H. and Conmy, P. 2004a. *DAME Dependability and Security Study: Asset Analysis*. DAME Project, York, UK. Report: DAME/York/TR/04.002.

Fletcher, M., Chivers, H. and Conmy, P. 2004b. *DAME Dependability and Security Study: Final Report*. DAME Project, York, UK. Report: DAME/York/TR/04.007.

Fletcher, M. and DAME Architecture Working Group 2003. *DAME Service Definitions and Descriptions*. DAME Project. Report: DAME/York/TR/02.005.

Foster, I. and Kesselman, C. 2004. *The grid 2 : blueprint for a new computing infrastructure*. Morgan Kaufmann: San Francisco, Calif.

Foster, I., Kesselman, C., Nick, J. M. and Tuecke, S. 2002. *The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration* [Online].

[Accessed 21/01/2003]. Available from: <http://www.globus.org/research/papers/ogsa.pdf>

Foster, I., Kesselman, C., Pearlman, L., Tuecke, S. and Welch, V. 2003. The Community Authorization Service: Status and future. In: *CHEP 03*, La Jolla, California.

Foster, I., Kesselman, C. and Tuecke, S. 2001. *The Anatomy of the Grid: Enabling Scalable Virtual Organizations* [Online]. [Accessed 16/11/2002]. Available from: <http://www.globus.org/research/papers/anatomy.pdf>

Gamma, E., Helm, R., Johnson, R. and Vlissides, J. 1995. *Design patterns : elements of reusable object-oriented software*. Addison-Wesley: Reading, Mass.

Georgiadis, C. K., Mavridis, I., Pangalos, G. and Thomas, R. K. 2001. Flexible team-based access control using contexts In *Proceedings of the sixth ACM symposium on Access control models and technologies* ACM Press, Chantilly, Virginia, United States, pp. 21-27.

Globus 2004. *The WS-Resource Framework* [Online]. [Accessed 02/02/2004]. Available from: <http://www.globus.org/wsrf/>

GRACE - Grid Search and Categorization Engine 2003. [Online]. [Accessed 1/8/2006]. Available from: <http://www.ub.uni-stuttgart.de/grace/>

Gridbus Workflow Engine (GWFE) 2006. [Online]. Available from: <http://www.gridbus.org/workflow/>

Harmer, T. and McCabe, J. 2005. *GT3.2 to GT4 Migration: A First HOWTO* [Online]. [Accessed 01/09/2005]. Available from: <http://www.qub.ac.uk/escience/howtos/GT3%20to%20GT4%20Version%200.3.htm>

Harmer, T., Stell, A. and McBride, D. 2005. *UK Engineering Task Force - Globus Toolkit Version 4 Middleware Evaluation* [Online]. [Accessed 01/09/2005]. Available from: <http://www.nesc.ac.uk/technical_papers/UKeS-2005-03.pdf>

Henderson, P. (Ed.) 2000. *Systems Engineering for Business Process Change: collected papers from the EPSRC research programme*. Springer-Verlag, London.

Henderson, P. (Ed.) 2002. *Systems Engineering for Business Process Change: new directions: collected papers from the EPSRC research programme*. Springer-Verlag, London.

Hollingsworth, D. 1995. *The Workflow Reference Model* [Online]. [Accessed 11th April 2003]. Available from: <http://www.wfmc.org/standards/docs/tc003v11.pdf>

Hutchinson, S. 2004. *Scientific Research Methodology - Putting Theory into Practice, SDDU Course*. Leeds University, Leeds.

IBM Websphere 2003. [Online]. [Accessed 11/09/2003]. Available from: <http://www.ibm.com/websphere>

IEEE 2000. *IEEE Std 1363-2000: Standard Specifications For Public Key Cryptography*. IEEE.

Internet2 2006. *Shibboleth* [Online]. [Accessed 01/09/2006]. Available from: <http://shibboleth.internet2.edu/>

Jackson, P., J. 1999. *Virtual working : social and organisational dynamics*. Routledge 1999: London

Johnston, S. 2004. *Rational UML Profile for business modeling* [Online]. [Accessed 1/9/2006]. Available from: <http://www-128.ibm.com/developerworks/rational/library/5167.html>

Kang, M. H., Park, J. S. and Froscher, J. N. 2001. Access Control Mechanisms for Inter-Organizational Workflow. In: *SACMAT'01*, May 3-4, 2001, Chantilly, Virginia, USA. ACM Press. pp. 66 - 74.

King, C. M., Dalton, C. E. and Osmanoglu, T. E. 2001. *Security architecture : design, deployment and operations*. Osborne/McGraw-Hill: New York ; London

Koshutanski, H. and Massacci, F. 2003. An access control framework for business processes for web services. In: *2003 ACM workshop on XML security*, 31/01/2003, Fairfax, Virginia. ACM Press, New York, NY, USA. pp. 15-24.

Kotsiopoulos, I., Keane, J., Turner, M., Layzell, P. and Zhu, F. 2003. IBHIS: Integration Broker for Heterogeneous Information Sources. In: *Proceedings of the 27th Annual International Conference on Computer Software and Applications*, IEEE Computer Society. pp. 378.

Krishnan, S., Wagstrom, P. and von Laszewski, G. 2002. *GSFL: A Workflow Framework for Grid Services*. Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, 1L 60439, U.S.A. Report: Preprint ANL/MCS-P980-0802.

Lepro, R. 2003. *Cardea: Dynamic Access Control in Distributed Systems* [Online]. [Accessed 15/03/2004]. Available from: <http://www.nas.nasa.gov/Research/Reports/Techreports/2003/PDF/nas-03-020.pdf>

Leymann, F. 2001. *Web Services Flow Language (WSFL 1.0)* [Online]. [Accessed 05/08/2003]. Available from: <http://www-3.ibm.com/software/solutions/webservices/pdf/WSFL.pdf>

Liberty Alliance Project 2006. [Online]. [Accessed 01/03/2006]. Available from: <http://www.projectliberty.org/>

Liu, P. and Chen, Z. 2004. An Access Control Model for Web Services in Business Process. In: *Web Intelligence, 2004, IEEE/WIC/ACM International Conference on*, 20-24 Sept. 2004, pp. 292-298.

Mendling, J., Strembeck, M., Stermsek, G. and Neumann, G. 2004. An Approach to Extract RBAC Models from BPEL4WS Processes. In: *Proc. of the 2nd International Workshop on Distributed and Mobile Collaboration (DMC 2004), 13th IEEE Int. Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises (WETICE 2004)*, June 2004, Modena, Italy. pp. 81-86.

Michelson, B. 2005. *elemental links: BPEL Primer* [Online]. [Accessed 1/12/2006]. Available from: <http://elementallinks.typepad.com/bmichelson/2005/09/view_bpel_proce.html>

Milner, R. 1993. The Polyadic pi-Calculus: a Tutorial In *Logic and Algebra of Specification*. Vol. 2005 (Eds. Bauer, F. L., Brauer, W. and Schwichtenberg, H.) University of Edinburgh, pp. 203-246.

Modular PERMIS Project 2006. [Online]. [Accessed 01/09/2006]. Available from: <http://sec.isi.salford.ac.uk/permis/>

Mont, M. C., Bramhall, P., Gittler, M., Pato, J. and Rees, O. 2002. *Identity Management: a Key e-Business Enabler* [Online]. [Accessed 30/09/2003]. Available from: <http://www.hpl.hp.com/techreports/2002/HPL-2002-164.pdf>

Norman, M. 2006. Types of grid users and the Customer-Service Provider relationship: a future picture of grid use. In: (Ed. Cox, S. J.) *UK e-Science All Hands Meeting 2006*, Nottingham, UK. National e-Science Centre.

OASIS 2003. *eXtensible Access Control Markup Language (XACML) Version 1.1* [Online]. [Accessed 02/02/2004]. Available from: <http://www.oasis-open.org/committees/xacml/repository/cs-xacml-specification-1.1.pdf>

OASIS 2004. *OASIS Security Services (SAML) TC* [Online]. [Accessed 02/02/2004]. Available from: <http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=security>

OASIS 2005. *OASIS Web Services Resource Framework (WSRF) TC* [Online]. [Accessed 12/07/2005]. Available from: <http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsrf>

Object Management Group 2003. *UML* [Online]. [Accessed 10/09/2003]. Available from: <http://www.omg.org/uml/>

Object Management Group (OMG) / Business Process Management Initiative (BPMI) 2006. *Business Process Management Notation (BPMN)* [Online]. [Accessed 1/8/2006]. Available from: <http://www.bpmn.org/>

Padgett, J., Haji, M. and Djemame, K. 2005. SLA management in a service oriented architecture. In: (Ed. Gervasi, O.) *Computational Science and its Applications - ICCSA'2005*, 2005, Springer-Verlag. pp. 1182-1191.

Papazoglou, M. and Schlageter, G. 1998. *Cooperative information systems : trends and directions*. Academic: San Diego ; London

Papazoglou, M. P. and Dubray, J.-j. 2004. *A Survey of Web Service Technologies* [Online]. [Accessed 1/10/2005]. Available from: <http://eprints.biblio.unitn.it/archive/00000586/>

Park, J. S. and Hwang, J. 2003. Role-based access control for collaborative enterprise in peer-to-peer computing environments In *Proceedings of the eighth ACM symposium on Access control models and technologies* ACM Press, Como, Italy, pp. 93-99.

Periorellis, P., Cook, N., Hiden, H., Conlin, A., Hamilton, M. D., Wu, J., Bryans, J., Gong, X., Zhu, F. and Wright, A. 2006. GOLD Infrastructure for Virtual Organisations. In:

(Ed. Cox, S. J.) *UK e-Science All Hands Meeting 2006*, Nottingham, UK. National e-Science Centre.

Rolls-Royce: Services 2004. [Online]. [Accessed 20/12/2005]. Available from: <http://www.rolls-royce.com/service/civil/helicopters/fha.jsp>

Russell, D., Dew, P. M. and Djemame, K. 2004a. Access Control for Dynamic Virtual Organisations. In: (Ed. Cox, S. J.) *UK e-Science All Hands Meeting 2004*, Nottingham, UK. EPSRC. pp. 332-339.

Russell, D., Dew, P. M. and Djemame, K. 2004b. Self Securing Dynamic Virtual Organisations. In: (Eds. Chivers, H. and Martin, A.) *Workshop on Grid Security Practice and Experience*, 8-9 July 2004, Oxford, UK. University of York. pp. II-1 - II-6.

Russell, D., Dew, P. M. and Djemame, K. 2005. Service-Based Collaborative Workflow for DAME. In: *Services Computing, 2005. (SCC 2005). Proceedings. 2005 IEEE International Conference on*, 10-15 July 2005, Orlando, Florida. pp. 139-146.

Rust, R. T. and Kannan, P. K. 2003. E-service: a new paradigm for business in the electronic environment. *Communications of the ACM (SPECIAL ISSUE: E-services)*. **46(6),** pp. 36-42.

Sandhu, R. S., Coyne, E. J., Feinstein, H. L. and Youman, C. E. 1996. Role-based access control models. *IEEE Computer*. **29(2),** pp. 38-47.

Sayal, M., Casati, F., Dayal, U. and Shan, M.-C. 2002. Integrating workflow management systems with business-to-business interaction standards. In: *Data Engineering, 2002. Proceedings. 18th International Conference on*, pp. 287-296.

Schoder, D., Fischbach, K. and Schmitt, C. 2005. Core Concepts in Peer-to-Peer (P2P) Networking In *P2P Computing: The Evolution of a Disruptive Technology* (Eds. Subramanian, R. and Goodman, B.) Idea Group Inc., Hershey.

Shadbolt, N., Hall, W. and Berners-Lee, T. 2006. The Semantic Web Revisited. *IEEE Intelligent Systems*. **21(3),** pp. 96-101.

Silva, V. 2005. *Globus Toolkit 4 Early Access: WSRF* [Online]. [Accessed 01/03/2005]. Available from: <http://www-128.ibm.com/developerworks/grid/library/gr-gt4early/>

Simon, A. R. and Marion, W. 1996. *Workgroup computing : workflow, groupware, and messaging*. McGraw-Hill: New York ; London

Simon, R. and Zurko, M. 1997. Separation of duty in role-based environments. In: *10th IEEE Computer Security Foundations Workshop*, June 1997, Rockport, Mass. pp. 183-194.

Smith, H. and Fingar, P. 2003. *Business Process Management: The Third Wave*. Meghan-Kiffer Press: Tampa, FL.

Stanoevska-Slabeva, K., Schmid, B., Tschammer, V., Ifip Conference on E-Commerce E. Business E. Government and International Federation for Information Processing 2001. *Towards the E-Society : e-commerce, e-business, and e-government : the first IFIP Conference on E-commerce, E-business, E-government (I3E 2001), October 3-5, 2001, Zurich, Switzerland*. Kluwer Academic Publishers: Boston ; London

Thatte, S. 2001. *XLANG - Web Services for Business Process Design* [Online]. [Accessed 05/08/2003]. Available from: <http://www.gotdotnet.com/team/xml_wsspecs/xlang-c/>

The Apache Jakarta Project 2004a. *Apache Tomcat* [Online]. [Accessed 23/04/2004]. Available from: <http://jakarta.apache.org/tomcat/>

The Apache Jakarta Project 2004b. *The Apache Struts Web Application Framework* [Online]. [Accessed 23/04/2004]. Available from: <http://jakarta.apache.org/struts/>

The Globus Alliance 2005. [Online]. [Accessed 12/08/2005]. Available from: <http://www.globus.org>

The Globus Alliance 2002. *Globus Toolkit 2.4 Overview* [Online]. [Accessed 01/01/2003]. Available from: <http://www.globus.org/toolkit/docs/2.4/overview.html>

The Globus Security Team 2005. *Globus Toolkit Version 4 Grid Security Infrastructure: A Standards Perspective* [Online]. [Accessed 1/9/2006]. Available from: <http://www.globus.org/toolkit/docs/4.0/security/GT4-GSI-Overview.pdf>

The Semantic Grid 2003. [Online]. [Accessed 05/08/2003]. Available from: <http://www.semanticgrid.org>

Thomas, R. K. 1997. Team-based access control (TMAC): a primitive for applying role-based access controls in collaborative environments. In: *Proc. 2nd ACM workshop on Role-based access control*, Fairfax, Virginia. ACM Press. pp. 13 - 19.

Thomas, R. K. and Sandhu, R. S. 1998. Task-Based Authorization Controls (TBAC): A Family of Models for Active and Enterprise-Oriented Authorization Management. In: *Proc. 11th Int. Conf. on Database Security XI: Status and Prospects*, Chapman & Hall, Ltd. pp. 166-181.

Thompson, M. 2003. *Akenti: Distributed Access Control* [Online]. [Accessed 02/02/2004]. Available from: <http://www-itg.lbl.gov/Akenti/>

Thompson, M., Essiari, A. and Mudumbai, S. 2003. Certificate-based Authorization Policy in a PKI Environment. *ACM Transactions on Information and System Security*. **6(4),** pp. 566 - 588.

Tuecke, S., Engert, D., Foster, I., Welch, V., Thompson, M., Pearlman, L. and Kesselman, C. 2003. *Internet X.509 Public Key Infrastructure Proxy Certificate Profile* [Online]. [Accessed 12/01/2004]. Available from: <http://www.globus.org/security/standards/draft-ietf-pkix-proxy-06.pdf>

UK e-Science Programme 2005. [Online]. Available from: <http://www.rcuk.ac.uk/escience/>

UNICORE Forum e.V. 2006. [Online]. [Accessed 1/8/2006]. Available from: <http://www.unicore.org>

W3C 2002. *Web Service Choreography Interface (WSCI) 1.0* [Online]. [Accessed 1/1/2004]. Available from: <http://www.w3.org/TR/wsci/>

W3C 2004. *Web Services Addressing (WS-Addressing)* [Online]. [Accessed 01/12/2004]. Available from: <http://www.w3.org/Submission/2004/SUBM-ws-addressing-20040810/>

W3C 2006. *Web Services Architecture* [Online]. [Accessed 1/7/2006]. Available from: <http://www.w3.org/TR/ws-arch/>

Wohed, P., Aalst, W. M. P. v. d., Dumas, M. and Hofstede, A. H. M. t. 2002. *Pattern Based Analysis of BPEL4WS* [Online]. [Accessed 16/03/2004]. Available from: <http://xml.coverpages.org/AalstBPEL4WS.pdf>

Wombacher, A., Fankhauser, P. and Aberer, K. 2005. *Overview on Decentralized Establishment of Consistent Multi-lateral Collaborations based on Asynchronous Communication* [Online]. [Accessed 17/05/2005]. Available from: <http://wwwhome.cs.utwente.nl/~wombachera/papers/P2005-03.pdf>

Yang, E. and Xu, J. 2005. Integrating an Attack Tolerant Information Service with Taverna. In: (Ed. Cox, S. J.) *Proceedings of 4th U.K. e-Science All-Hands Meeting*, 19/09/2005, Nottingham, UK.

Yao, W. 2003. *Trust Management for Widely Distributed Systems*, thesis, University of Cambridge, Cambridge

Yavatkar, R., Pendarakis, D. and Guerin, R. 2000. RFC2753: A Framework for Policy-based Admission Control. *IETF Informational Standard*.

Yunker, J. 2002. *ebXML and the e-Business Process Stack* [Online]. [Accessed 05/05/2005]. Available from: <http://www.ebxml.org/presentations/john_yunker.ppt>

Zdancewic, S. 2004. Challenges for Information-flow Security. In: *Proceedings of the 1st International Workshop on the Programming Language Interference and Dependence (PLID'04)*, 25th August 2004, Verona, Italy.

Zhou, J. 1997. *Non-repudiation*, PhD thesis, University of London, UK,

Appendix A

DAME Workflow Architecture

A.1 Workflow Manager Components

The Workflow Manager is shown below using a collaboration diagram. The workflow manager executes autonomously by monitoring the state of individual workflows. Each workflow executes independently in a separate thread. Asynchronous connection to the system is achieved using the Portal base class, via the Portal Connection. This allows workflows to be started, monitored and the results to be retrieved. An additional connection to the workflow manager, the Event Agent, is used by the Ground Support System to launch the automatic workflow. The automatic workflow "Workflow Brief Diagnosis" is shown in the diagram, with multiple instances. However, all workflows share the same interface and there can be other workflow types executing simultaneously.

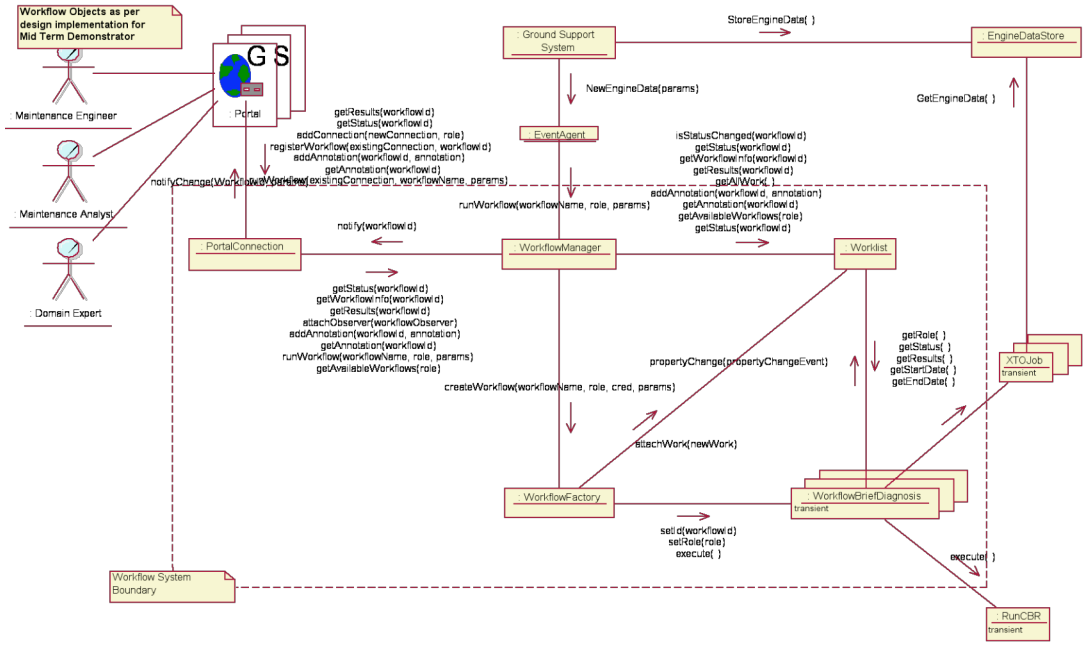

Figure A.1 DAME Workflow Management System Architecture

The text following describes each component shown in Figure A.1.

Event Agent

The Event Agent is a specific connection for the Ground Support System to the Workflow Manager, with the purpose of launching the automatic workflow when new engine data arrives. It provides a simple interface to the Ground Support System, hiding the more complex interface to the workflow manager.

Engine Data Store

The Engine Data Store represents the repository for engine vibration data files. New data is loaded via the Ground Support System when an aircraft lands. Other system components, such as XTO, retrieve vibration data sets by querying the Engine Data Store with flight details.

Ground Support System

The Ground Support System represents the station used to download vibration data from the on wing QUOTE system and upload it to the Engine Data Store. Then, on the event of new data, launch the automatic workflow using the Event Agent.

Portal

This is the abstract base class that provides the external connection to the workflow system. Specific implementation of the Portal can be provided for connection into different systems

such as Struts based MVC style web server, WAP interface for mobile devices or a simple pager interface for alerts on workflow status changes.

Portal Connection

The Portal Connection follows the observer pattern (Gamma et al., 1995) used in the composite MVC pattern. Many Portal classes can connect to the workflow system, using the Portal Connection. Each Portal registers interest in workflow instances, and the Portal Connection notifies the relevant Portals on change of status.

Run CBR

The RunCBR Class represents the client component to execute the remote CBR process.

Workflow Brief Diagnosis

The Workflow Brief Diagnosis is an implementation of the Workflow class. It represents the automatic workflow launched on the event of new engine data arriving. It executes a sequence involving most tools in DAME.

Workflow Factory

The Workflow Factory creates workflow instances. The name of the workflow is passed to the factory and an instance of the relevant workflow type is created and initialised with the parameters passed into the create function.

Workflow Manager

This is the central enactment engine. It is a permanently executing thread that controls the creation of workflow and monitors their execution. Changes in workflow instance status are forwarded to the Portal Connection. Requests from the Portal Connection are forwarded to the worklist.

Worklist

The Worklist contains the workflow instances. Each workflow is accessible by using its unique ID.

XTO Job

The XTO Job represents the Grid Service executing the XTO data analysis process.

Appendix B

Evaluation Interviews

B.1 Interview Questionnaire

Collaborative Workflow Interview

Issue 2D. **Russell**

27th October 2004

Purpose and Scope

The aim is to assess the research into secure control of collaborative workflows, in the context of the DAME case study. This provides a concrete example of how the research has targeted a real problem. By interviewing representative of the business concerns for the case study the data collected will be used to validate the current state of the demonstrator of the research and provide validation for the vision of how the system would ideally be implemented.

Interview Candidates

There are two interview candidates, one from each of the target companies involved in DAME.

Graham Hesketh from Rolls-Royce

Charlie Dibsdale from DS&S

As there are only two respondents it will not be possible to keep the interviews anonymous. However, there is a need to protect commercially sensitive information. Therefore, it may be necessary to collect data that may be useful in validating the demonstrator and vision, but will not be included in the results or publications. This may make it difficult to publicly validate the research.

Contacting the candidates

The candidates will be contacted by email to arrange possible times for the interview. This will take place at a convenient venue to the interview candidate. Typically in Derby for Graham Hesketh and either Bristol or Derby for Charlie Dibsdale.

A confirmation of the arrangement will be sent by email, along with a guide as to the topic of the interview and an agenda for the meeting.

To motivate participation, the interview will be made convenient to the candidate, and they will be provided with the results.

Interview Design and initiation

The interview should be structured in the question format so that on repetition the proposed questions will be similar. The limited number of candidates for the interview makes it important that the interview process is kept similar so as to improve consistency of results.

The structured interview is aimed at directing the questioning in the context of collaborative workflows. However, the exercise will also benefit from more exploratory questioning, where the interview candidate is free to provide extended answers. Some of the questions are intended to prompt the candidate to answer more in-depth, possibly revealing information not thought about when compiling the questionnaire. The interviewer will have sufficient knowledge on the topic and would be able to ask more probing questions if time allows.

Situation

The interview should take place at candidates place of work, for their convenience and to make them feel more comfortable to talk about the problems from their perspective.

Recording information primarily by note taking, both during and directly afterwards. By requesting to candidates it may be possible to use tape recording equipment (e.g. Dictaphone).

It is intended that the interview duration is 1.5 hours maximum. This includes a brief demonstration of the aspects to be evaluated. It is understood that the candidates are familiar with the DAME demonstrator, and the demonstration at the interview will be used as a refresher to the demo and to focus the discussion on the topics in the questions (namely the implementation and control of the collaborative workflows).

Interviewer will dress smartly to appear professional, conscientious and to match the formal environments of the candidate's work place.

Testing the Interview Design

The interview questions will be trial on personnel at Leeds to ascertain:

Relevance to the target subject

Ease of understanding of the questions

Length of time taken to execute the interview process

As a result of the trial, the questions may be rephrased to improve their understanding, or relevance and the number of questions may be reduced. There may also be a list of guidelines to for the interview to use a prompts to gain relevant answers from the candidates.

Recording Information

Other than the answers to the questions recorded information from the interview should include:

Whether all questions were answered

Duration of interview

Opinions on respondent

Perception of interaction

Results of interview test

The interview was tested on 27th October 2004 from 10:20am to 11:30am with Georges Honore. During the 70 minutes the procedure involved an Introduction for 3 minutes followed by structured interview questions and a 2 minutes summary. There was also a 5 minute interruption and some comments on the suitability of the questions.

The answers were drawn from Georges' knowledge of the current DAME demonstrator and the use case study documents (which were used as requirements input to the design of the demonstrator). However, this test revealed some repetition of the questions and some questions that were irrelevant. The results meant that some repeated questions were removed and some were modified. The questions addressed aspects of both the demonstrator and future requirements of a production DAME system. Therefore, where there was repetition the questions needed rephrasing to ensure they addressed either current or future requirements.

Interview Format

Item	Title	Duration	Scope
1	Open Interview	3 mins	Open the interview
2	Introduction	5 mins	Introduction and Purpose of interview, including aims to orient the respondent.

			Reaffirm the confidentiality of the interview?
3	Demonstration	10 mins	Demonstration of collaborative workflow in DAME demonstrator
4	Interview questions	50 mins	Structured interview questions
5	Comments	15 mins	Open Comments
6	Interview Summary	5 mins	Interview Summary
7	Close	2 mins	Close

The total time should be no more than 1.5 hours, including the introduction and demonstration and concluding with a short summary of the interview results.

In general, the evaluation will attempt to elicit answers around the following:

Collaborative use of workflows and tools in the diagnostics procedures,

Designation of users and roles in inter-organisation processes,

Access control to data, workflows and services across organisational boundaries,

Relevance of the demonstrator collaborative workflows to DAME business requirements.

Interview Follow-up

The documented results of the interview with a candidate will be sent to the respondent by email and say 'thank-you'. This will be done before it is used or published to give the candidate the opportunity to correct any mistakes of omissions. These changes should be recorded as such in view of provenance to the data.

Also send any additional material that may have been promised or realised in the interview that the candidate would like to have.

Structured Interview Questions

The questions will be shown on a form that should be filled out with answers during the interview. These will need to be thorough if tape recording is not permitted, and will probably need to be checked and completed after the interview. This would allow the interviewer to concentrate on the rapport with the candidate rather than spending time note taking.

Collaborative Workflow in the DAME Demonstrator

1. Does the workflow in the Demonstrator reflect the expected business process?
2. Does the secure access in the demonstrator match the expected business process?

3. Does the demonstrator represent all the organisations and roles in the DAME diagnostics process?

Collaboration Procedure

4. What happens in the escalation of diagnosis problems between:

4.1 The Maintenance Engineer and the Maintenance Analyst?

4.2 The Maintenance Analyst and the Domain Expert?

5. What information is passed during escalation?

6. In what way would the roles use DAME to collaborate?

7. What information do users in the same role share and are there restrictions on sharing?

DAME Roles

8. Are the three defined roles sufficient to define access control for collaborations?

9. What restrictions are there between:

9.1 Maintenance Engineers sharing a process?

9.2 Maintenance Analysts sharing a process?

9.3 Domain Experts sharing a process?

Workflows

10. How should the processing resources be managed securely?

11. What restrictions would there be on executing workflows from the different roles?

12. What restrictions would there be on sharing workflow definitions?

13. What restrictions would there be on sharing/collaborating in active workflows?

14. Who can define new workflows for DAME?

Access Control

15. How should users login to the DAME portal?

16. How should organisations provide user identities?

17. Who is responsible for assigning roles to the users?

18. How should the roles be defined?

19. What restrictions are there on sharing data and services across organisations?

B.2 Interview Results DS&S

Interview Results from 4th November 2004

Issue: 1 Author: D. Russell Issue date: 17th November 2004

Purpose and Scope

The aim document contains a summary and transcript of the research interview with Charlie Dibsdale that took place on 4th November 2004. This interview conducted by Duncan Russell addressed the author's PhD subject area of 'Secure Collaborative Workflow'.

Summary

The following is a summary of the points discussed separated into categories for Collaboration, Workflow, Access Control, Roles and General Security.

Collaboration

The current demonstrator accurately represents the desired business process for the diagnostics case, but is limited to the identified use cases to be achievable in the project timescale.

The collaboration between roles will require increased sources of data, such as performance data and engine configuration (currently DAME only uses vibration data). DAME would not require much dialogue between the roles, however other examples of diagnostics would. Example being, in power station diagnosis the operator's experience would need to be obtained and captured within the system, usually in longer term diagnostics.

The current operations room has a multi-disciplinary team that would collaborate on solving a problem. This would require shared access to active workflows and possibly collaborative services. These Maintenance Analysts would be performing most of the problem solving.

Workflow

Workflow definitions intrinsically capture knowledge and therefore have business value.

A workflow administrator role was identified to control:

Access to launch workflow, especially with workflows (and their capabilities) and restricted by the level of service (package) the customer pays for. The general example is bronze, silver or gold packages for service contracts.

Who can define/add a workflow into the system. And within that definition what users are able to access which parts.

More complex is in access to workflows automatically defined from captured knowledge (usage history) in the system (e.g. Workflow Advisor). This service could improve one company's performance by using the practices of another.

Access control to services & data

The main assets to protect are data for competitors. This is characterised as data that is identifiable by airline. Methods of anonymising (by aggregation) allow sharing of data. This is mainly concentrated in accessing persistent data, and not diagnostics service results, although the data aggregation may be a service operating on the stored data.

Service access is still restricted mainly by the information that can be obtained. There could be restrictions on launching processing on resources to ensure availability of processing power for other users (DAME processes or internal users, such as design simulation). This is based on resource usage/allocation and one of the decisions used in access to resources will be cost (financial).

Roles

In addition to the roles already defined for DAME, other roles that use the system would have different access rights. For example, regulators may require access for assurance of operation. This may be used to view summary data, showing averages and trends. It is not assumed that access for these roles includes diagnostics tools.

Users will be attributed with role and organisation to define what data, services/tools and screen they can see. Additionally a working context might be a constraint on access to services or data, e.g. the workflow would define the context.

A central administration for DAME would provide the organisations with user templates, detailing role name and access rights. An organisation is then responsible for naming the users in that role and informing the administration when access should be revoked.

The access rights in a role template would be like a job description, detailing how the user fits in the overall process. Attributes in the role template can also be used to check if the user is qualified to that level of access.

General Security

Single Logon is required across organisations. Other companies' portals with access to their own systems could also make use of DAME tools. This maybe access to the services or sections of the portal. It would not be desirable for a user to enter username and password for each system.

DAME would require central administration for control access to workflow/services/data controlled by the DAME contract. This would include user identity mappings to roles and

access policies. A 'central' administration could be a distributed effort, but is essentially within one organisation/trust domain.

Secure applications (obfuscated operation and data) would be required for execution on third-party grids.

Access control policies could also 'judge' suitability of a service by requiring it to correspond to certain standards, such as z-mod for data exchange and the MIMOSA operating standard.

The security framework must be adaptable to changes in users and possibly organisations access by package. Also changes to the packages to respond to a services life-cycle.

Interview Transcript

The following is a transcript from the taped interview between the two people:

CD = Charlie Dibsdale

DR = Duncan Russell

Some details of the speech have been omitted in the transcript such as 'Err' and 'OK' and 'Yes'.

Interview started at 1:05pm 4th November 2004, in Charlie Dibsdale's office, DS&S, Bristol and finished at 2:15pm.

Introduction

DR: OK, so a brief introduction so we know where I am with this interview

CD: Yes

DR: It's part of my PhD research so it's part of my data that I am gathering. As you know, my research is concentrating on the secure access and the collaborative access to the tools within DAME.

CD: Yes

DR: So that's why this is different to the previous interview which you did with Richard and Ian which concentrating on the usability of the portal. I am still going to be referring to the portal in the current demonstrator. And we are looking to see what your views are on how that suites the need and what you see the future need is.

CD: OK

Collaborative Workflow in the DAME Demonstrator

DR: So I have split this up into a few categories and I am going to start with the demonstrator. OK so the first question, I'll just make a few notes. Does the workflow in the Demonstrator reflect the expected business process?

CD: OK. What I would say first off is that the demonstrator project decided to choose a set of use cases based around a very sort of narrow scope of process. Because we had to bound the project and obviously make it achievable. Whilst at the same time demonstrating some of the features of using grid, distributed and all that sort of stuff. First off the demonstrator itself as far as I am concerned only has a small scope of its potential value. And its usage I believe, in an operational environment, would have a far wider scope and set of deliverables, etc. That would substantially increase its expected business value. Is that a fair answer, or…

DR: No, that's good.

CD: The other thing I would say in terms of what the project set out to achieve and did it achieve it in the workflows, then, yes it did.

DR: Right.

CD: Which is the other side of the question I think you are asking.

DR: Yes, OK. Does the secure access in the demonstrator match the expected business process?

CD: I think that we have a secure logon site and probably if we went in Rolls-Royces requirements we would require two-factor authentication. So as it stands, if it went operational, probably not. But I don't see that there is a problem to implement that extra layer of functionality. The other aspect is in general grid usage that the security model needs to be extended to actually protecting applications as they are running on external processing devices, not only the application but also the data. So if we are brokering with somebody else that we want to be protected against is that they can take a look in as things are being processed and not be able to interpret from that point of view. And I am not sure that grid technology goes that far nowadays.

DR: I can say it doesn't quite tackle that, but I know somebody in Leeds who

CD: is looking at it?

DR: yes, looking at privacy in the application itself.

CD: That is absolutely the issue. Now whether that is a big issue or not is debatable because, we may choose to implement a grid either internally or with a trusted third party compute supplier. And it may be years before we get out to the general grid community. Where that level of security will be required.

DR: I agree, also the market model for that is a bit muddled at the moment.

CD: Yes, it is not mature.

DR: I don't know if you heard, that at the moment Hewlett Packard Data Centre. Because they were building a warehouse.

CD: Oh I see

DR: In fact, they did they built a warehouse of compute processors. They were looking at about five thousand processors in this one warehouse. That's what they were aiming towards. But they have closed it now.

CD: Have they really?

DR: Already closed, they didn't get enough take up.

CD: Well it's too immature at the present time.

DR: OK, Does the demonstrator represent all the organisations and roles in the DAME diagnostics process?

CD: Within the bounds of the scope, yes it does.

DR: OK. The interest there is the accuracy of how the airlines are represented, because so far we have used yourselves at DS&S to explain the role of the airline, because they are not partners in the project.

CD: but by the other token what we are moving to, as a business model are product services based for Rolls Royce. As far as diagnosis on the application is concerned its all within Rolls Royce's interest because the are responsible for the availability of the whole unit. The Airline in that business model has got a very much reduced interest in the ins and outs of what is wrong with the engine. So they don't need the diagnostic output. What they need is the assurance that they can fly to their schedules.

DR: Yes, that's good.

CD: So, I am not concerned with that at all.

Collaboration Procedure

DR: So. This is specifically about how you see the process evolving. In particular one area that wasn't captured very well is the escalation process. And really the collaboration between the Maintenance Engineer and the Analyst, and then the Analyst to the Domain Expert. So how would you see that process actually happening?

CD: I think that there will be occasions that some failure modes, and it is a rare occasion, that are not, have not ever been experienced before. And in that sense I think the escalation

will be fairly quick up into the domain expert. Who will need a whole series of tools, capabilities and visibility of data. I think recently I think some of guys who are trying to disentangle the root causes are probably not explicit domain experts but they are a part of the crew who are trying to run the operations of engine support. Therefore they are trying get to a root cause by looking at event data and also configuration data is important. Perhaps unfortunately with aeroplane engines, they are very complex assets with lots of modifications. And a failure occurring in one node, one part of an engine, there might be several candidate equipment items that are fitted there all with different modification states. So it's very important that the failure is sentenced against what modification state of the equipment. Maybe that's one area that we need to look at and exploit, is bringing in these other sources of data. That does not need the Dennis King type who is looking at raw vibration data, that's more at the maintenance level.

DR: Yes. OK.

CD: The output of DAME is more aimed at going to somebody like Dennis King. The real output and the real value exists in the people who are trying to support the asset out in the field. That is through the Rolls Royce operations room.

DR: Right, yes.

CD: But that is where the value to Rolls Royce exists compared with, sending information back to engineering and the reliability groups within Rolls.

DR: OK. When they are looking at the data, obviously you said about the configuration data and things which on the whole you would expect to be able to pull from databases.

CD: Yes

DR: Is there going to be much dialogue between the different layers. Are they going to back to the ground staff, are they going to try to ask questions.

CD: I think that a dialogue with the ground staff in reality is going to hardly exist. That's a consequence of choosing this very narrow use case. So we have this hypothetical question that we have a severe time delay because we want to dispatch aircraft in time to reduce delays. In the real world it doesn't happen like that. The majority of these things are probably longer term investigations. Otherwise we are already able to detect the reoccurring failures.

DR: Right. That has kind of covered my next question which is the information that is passed during the escalation. Do you see any other ways in which the roles collaborate?

CD: I see massive potential for a complete explosion of use cases.

DR: Right

CD: It is that big. I think we have been very narrow focused. The diagnostic capability is not limited to aeroplane engines either. So there will be slightly different use cases because, in different market sectors it is all just as applicable. So collaborating with the actual operator, say in a power station situation we would be talking the operator because we would want perhaps anecdotal information, or trying to talk to them about the experience of how failures, but what you would hardly not envisage in the DAME situation is talking to the pilot. There are many more examples. My belief is that this whole environment has a whole set of other uses and use cases.

DR: That's good. One of the issues with my research is that I am really focused with the DAME use cases, and that's fine if I was in development. But for research in want to address a wider audience. So that's good information.

CD: I am not criticising DAME, because it had to demonstrate that, and it had to be controlled for the project deliverables. So going back on the decision, would we do the same? Yes, absolutely.

DR: I think along with the collaboration, probably applicable to DAME with the fact that you have the different airlines. Would there be, or what sort of restrictions would there be on sharing information.

CD: OK, we have massive experience of this within DS&S, because we are doing on-line condition monitoring as we speak. It is absolutely not allowed for one airline to see the actual performance of a rivals operating fleet. So that would be the fastest thing that would shut us down as a service business, that you could imagine. A lot of our effort is to ensure that we don't actually do that. However, there are ways and means of anonymising data, because a lot of operators are very interesting in bench marking themselves against the whole industry averages. And therefore can plot their performance with some derived key performance indicators. Against the fleet average is a very powerful output that most operators are happy that their data contributes to and that they get a comparative view of their performance and others get the same. So I think in that respect, as long as you anonymise it and you aggregate it then not a problem. So it is benchmarking is a big value string.

DAME Roles

DR: OK, again back to our diagnostics process, covering the roles. I think you have probably already answered this, but are the three roles sufficient to define access control for the collaborations.

CD: For the scope of work that we have, yes. But, not for a wider scope.

DR: And you have mentioned the operations extending the use cases.

CD: That's the Rolls Royce operations room. What they have is a bunch of people to execute the total care contracts to make sure that engines are available to fly. So that set of operations people a group of multi-discipline people. There may be some logisticians, there may be some maintenance planners, there may be a repair and overhaul guy. There are all kinds of different people who are all interested in understanding the condition of certain components of the engine. So we have to take a look at the operations staff and the output they want to see. And in different operating contexts or the market, the operator of the asset may be a person that you would also be speaking to, such as the operator of the power station.

Now I can envisage that regulator may be a valid customer. Because they are ensuring themselves of compliance to health and safety. You could demonstrate to other stakeholders, lessors and financiers of the assets. They are interested in understanding what is the condition, what is the history of an asset. 'Am I safe to lease this thing or not?'. So, I could spout off for quite a long time to widen the potential customers.

DR: That sort of fits in with access control mostly to persistent data.

CD: Yes. What you wouldn't expect is for some of these people being able to dig into the diagnostics tools or that raw data. They want to have a summary status of condition. Things like that, that is derived from... Very much a subset of what we are looking at.

DR: Good. I think we are now looking at specifics of the DAME roles. This is something that we haven't modelled, would be any of the roles sharing between the same roles. So Maintenance Engineers sharing the same process, or Maintenance Analysts or Domain Experts involved collaboratively through the DAME portal.

CD: You talk about the escalation, I don't think it's 'I have finished my bit now I need to delegate up to you'. I think its 'OK, guys I need the Domain Expert to become involved' and I think the collaboration be the Domain Expert will give advice in looking at it. 'Yes, where I think you need to concentrate on is in this area' and the Maintenance Analyst, DS&S personnel would do a bit more of the leg work.

Workflows

DR: This is now going to the diagnostics tools and, how should process resources themselves be managed securely on the grid?

CD: There is going to be some concerns about cost, initially. This could be regarded as being a bit miserly by the commercial guys. If you are able to set off a load of processing that is going to either grab a lot of the available processing power then there are cost consequences of that, if going out to a general environment. There are also priority constraints. For instance, you have to look at the wider use of an internal grid. There might

be times that the design community might want to run an optimisation and take up 90% of the internal processing. Yet, we need to do a fundamental Aura search that is also going to grab 90% of the processing power and come back with the service level. Because the brokering and the service level need to be integral to everybody's use of the shared resource. Who can be authorised to set off the jobs? I don't know the answer. I think it is partly a process thing and partly an authorisation or brokerage problem.

DR: What about executing the workflow. What restrictions would there be on executing the workflows from different roles?

CD: Let's take one side, because there is an automated section that the system itself is authorised to conduct workflows in certain scenarios. I think that is obviously easy controllable and easy enough to think about whether that would automatically happen according to some criteria or rule. So that is probably embedded within the workflow system, classified. And, what workflows? Again, there maybe workflows that the maintenance guys use that maybe subsets of bigger workflows that the domain expert might want to use. How do you mix and match that? Has there got to be the concept of a workflow administrator I wonder?

DR: I think in the general model of grid, it is quite difficult to pin that down. But in this specific case, there is the concept of control administration, so it is easier to place somebody in charge.

CD: That would be the short term answer. But I was thinking that maybe it becomes another role in the use case scenarios.

DR: I think that goes onto the next question, on sharing workflow definitions. Once we get to the state where the Sheffield workflow advisor would be able to create workflows, I suppose it would be the same sort of thing.

CD: There is also another issue to do with workflows, because what you are doing is explicitly capturing knowledge. And the knowledge embedded within the workflows is explicit and it has also got high value and also got a level of intellectual property. Which has intrinsic business value. Who has access to and who can reuse is a debate for the business people to decide. If perhaps they want external people to make use of that as a service in maybe a more traditional business, then it becomes a revenue stream.

DR: OK

CD: You can also think of a case where you can segment service provision from decision support, for a third party operating assets. Because you can say 'Do you want the bronze, silver or gold contract?'. And the gold may involve the full automation of the workflow management system. And by the way, that becomes an issue because if we are allowing a

third party to use the diagnostics capability, we are actually capturing their knowledge within the workflow system. So you can see the kind of issues. As the service provider if you have three or four of these customers and you are looking at similar assets, the fact that you are capturing knowledge explicitly by yourself you can make that reusable and use one persons 'best practice' and improve somebody else performance.

DR: Yes

CD: That is a knotty business problem, that will be an issue.

DR: I think that comes down to the same as the benchmarks, that if you can pool it, and everybody benefits, then you can agree on sharing workflows

CD: Yes

DR: I think we have kind of cover restrictions on sharing active workflows.

CD: I would say restrictions because the workflows have value. You want to restrict access in order that you can exploit that value.

DR: Yes, and another we have kind of covered, but just to clarify. Who can define new workflows for DAME?

CD: Well, if you have got this system that is monitoring usage of workflows, then surely it defines its own new ones. In a way that, perhaps is not a straight question because the system itself is define new workflows. The other aspect is that if the DE or the MA comes up with a novel way of diagnosing a particular classification of problem then of course we want to capture that as best practice. Then both of those parties should be able to define a new workflow.

Access Control

DR: OK Last set of questions. Access Control, again you have kind of answered this: How should users login to the DAME portal?

CD: I'm talking as an administrator, so administrating the DAME system. User access control is very important. I'll give you a scenario: You may have a Maintenance Engineer, who is employed. Gets sacked for some reason and we omit to take the off the access control. Next day, they log in to the system, extract a lot of data and sell it to The Sun. Because they feel aggrieved that they have been unfairly dismissed. But of course we have got to be extremely tight in access control and indeed with the current commercial systems that we have now it is something that we take very seriously. And make sure that the user access profiles are done virtually on an on-demand basis. We also have a whole area of our portal by which we can administer the user access, friendly and easily. It becomes a big overhead when you have potentially thousands of users who are able to log on.

DR: Would that be explicit areas on a per user basis, or does map it to roles and groups?

CD: No, I think you have got to map to roles. There is a whole area of debate here. On some occasions you may require to go down to data field level. I'll give you a for instance. We run forecasting and services for Rolls Royce that are linked into lifecycle cost. Our analysts are interested in the forecast, per say. Rolls Royce doesn't want us to see the cost data associated with it. When we are able to log on, we can see all the machinations of how to do forecasting with all the data transformations that takes. What we don't see are all the cost elements that are associated with that. So, we may be simulating a failure and recovery of an asset, where you need labour and material costs for that. We can see the labour, the people who are required and qualified to do the fix. And we can see the material, but we can't see the cost rates. In that case, the security access for a person designated as a DS&S person, is allowed to see everything apart from cost data. There are possibly the instances where you have to go down to the data level to restrict access.

DR: And effectively that restriction is done by attributing the user to the organisation?

CD: Yes. Possibly you might not be able to foresee. But you do need to be able to build an architecture that could limit you at any level within the architecture to a subset of attributes that you might possibly want to define. So it could be work page, it could be data element.

DR: OK, this here is an interesting one. This is to do with the administering of the system. How should organisations provide user identities?

CD: There is obviously the person. Are you trying to get at the fact that you are trying to authenticate the person or the machine that they are logging in on.

DR: The person, at this point.

CD: I think you have to have named person/people with a defined role, within a defined organisation. I think those attributes are the tree things that drive the security. There maybe more but I can't think at the present time.

DR: So, who is responsible for assigning the roles to the users?

CD: I envisage an administrative function with the portal and there is also a commercial aspect, again. It would put this in an operational environment . The part of the process doing the sale, or whatever it is you are doing to you customer set, will be the definition of what the customer can log on for, what purpose. And what can they use and see. There may several different classifications of different types of customer, because they have different job roles and work with different customer processes. If you essentially say, Person-Role-Organisation, I think you can define that.

DR: Again we are talking about that being centrally administrated.

CD: Yes. Centrally of course, if you think a little laterally, the grid itself allows you to do a distributed effort. And therefore you might have administrators who are scattered, depending on where their market is. You might have a US one, you might have a UK one. What I would guess there is that there are potential conflicts, because the interpretation for a role in the US might be completely different from what it is in the UK.

DR: Yes, so you might want context for that role. I think what we are saying is essential here is because it is in the same trust domain. We are not trying to cross boundaries. I think the question is addressing the issue of whether the airlines could provide identities and attributes to those identities.

CD: I think we give a template definition, and say these are the roles, these are the groupings. They have to supply the names of the people who are actually fulfilling those and when let a contract for say accessibility to health monitoring today. We contractually put the onus of the customer to make sure that they keep us informed of who is an authority user and who is not. We cover ourselves commercially by saying, if you don't tell us, we are not liable to the kind of accident I described earlier. We also have a service level around that, that when we have a new user or somebody who needs to be deleted the system or revoked. There is a timescale and a response time under which we must do that.

DR: OK, That's good. I think we might have addressed this one, but I shall ask it anyway. How should the roles be defined?

CD: Almost like a job description, is what I feel. Give some context of that role in how it fits into the overall process. It's got an expected set of responsibilities. Have a set of processes that you are linked to. How those processes are triggered. In fact, you could probably extract that from the use cases themselves. There maybe other things that the use case doesn't include. Say, what qualifications do we expect a person to have and what attributes, a sort of job description.

DR: That's good, because that fits in with what I am looking at with the policy. So you might be able to define some of those to be used with the policy. So it maps it when it looks up the role, and the role might have a policy. So you can explicitly and quickly see what a role can do, without have to look up the rest of the system.

CD: Yes, and there maybe a qualification that a person gets labelled as a DAME maintenance advisor. It may be training, and a who set of other attributes that you must have in order to fulfil that role.

DR: That makes sense. And again I think we have answered this one, but the restrictions on sharing data and services across organisations? I think what would be interesting is to expand on that question, was the notion where the airline would possibly be providing their

own services to access the data and diagnose their own fleets and look across their own fleets.

CD: Yes

DR: Or you could possibly provide them with some access to other organisations data.

CD: Lets differentiate the difference between data and enablers. With enablers what you could possibly do is open up a restricted set of services such that: Somebody like Lufthansa is probably one of the main airlines in the world that likes to keep a lot things in house. I could see value in that one of the diagnostics services that we have is so good they might want to make use of it, but from their environment. So they are passing data to it, and taking the results back in house as raw data. Not actually accessing via the portal user interface. So, I guess what we are exposing there is some kind of Web Service for a particular enabler, and they are just calling it on an as required basis. Now I believe that is a service up-sell opportunity, because you can charge Lufthansa for the availability of that. We don't have to constrain them to use the user interface that we have.

As far as data is concerned, I think anything that gives (or is seen as giving) competitive advantages to a set of customers is the quickest way you can shut down a business. But, with benchmarking if you can get the complete data, then there is great benefit in customers measuring themselves against the average.

DR: OK, so you might on one level be providing access to data that is aggregated and anonymised.

CD: Now, there is another, If we have access to the full fleet historical data. Let's say that Lufthansa wanted to do a search because they have just seen an anomaly through the whole database. I think that is possible. As long as the results are anonymised, going back. Because it has to be non-attributable to where it has been experienced before. Now that's going to create problems when you look at the use cases. Because, you are extracting event data based on the index of finding similar anomalies with your research, then you are going to have to make damn sure that that event data is anonymised. I don't know how you are going to manager that.

DR: I can see it would be extremely difficult to maintain it without some way of cross referencing to tell where that data had come from.

CD: Yes, an intelligent user of the system could fire up a whole series of them and reverse engineer who it came from. Because you could get clues about who is operating in an operating context. Yes, could be an issue.

DR: So that might mean you would have to provide, I assume that there would be a level of that within DAME anyway. So, that the say, aura-g search result would be processed by

CBR. And then CBR would say this is the likely result because this matches up with the case. If that's the sort of thing you could provide to Lufthansa, that would be anonymised because that would provide a diagnostic result.

CD: You have to be sure that the instances of the conglomerated results are not attributable to specific customers. It is going to be very interesting in the CBR approach, because I can see all kinds of searchable attributes that we haven't delved down into in order to isolate. Because operating context and environment are very big drivers of reliability. So how do you classify those. And certain operators will be in predominately one context i.e. the different between short haul and long haul.

DR: Yes, I remember from some of the discussions with Dennis King that he looks at which airline it is, so he knows what the operating conditions, what altitude they run at.

CD: That's precisely my point.

DR: That is all the question I have. Do have any other comments you would like to add.

CD: I think I made the comment earlier on the general use and not being able to reverse engineer. What we didn't do very much within this project is look at the metadata.

Getting more out into the grid, if you want general purpose providers, i.e. they have got an algorithm. You not going to be able to be a general provider of those services for anybodies z-data. A silly example, you have a really novel method of doing fast Fourier transformations, and you want to offer that out to anybody. You are going to maximise your market if you can get data into a defined standard. We haven't really talked about those defined standards of metadata within DAME. So that we can link up several different diagnostic capabilities much more conveniently. I think the metadata standard that has been intrinsically in there is the z-mod file format. For vibration spectrum data. That I think helps. This also plays into the security game, because you could think about the concept of a certified service provider for a certain algorithm. One of the things in your brokering, you could say what metadata standard are you using for vibration data.

There are standards that are coming on. The MIMOSA is the condition monitoring standard in the US. And of course PLCS is more associated with the event and maintenance data from ISO.

DR: The sort of thing that I am trying to move on with is combining the workflow with secure access and then trying to provide templates. And the template really works directly with the workflow and says, here are the roles and when you generate the instance of the workflow, then it starts mapping users to role instances. And then you can define the users who can use this workflow.

CD: That's a very knotty problem, because there may be some small workflows that say maintenance guy could not access, and yet you could have a super workflow where the one that you are not allowed to run is right in the middle. Can you run that and not see the input and output results? It becomes a black box that you wouldn't see the workflow. It could become a horrible rats nest.

The kind of instance that you might not want, is that you pull out a report that says this is the relative safety this year, this month compared with this month last year. And the industry is saying we have improved safety performance by ten percent and in fact it has fallen by twenty percent. That's dynamite data.

DR: So as a summary, I have certainly got some good information, on expanding the DAME use case with other examples and I have a good idea of the surface of those. So I can apply them to the general model, so that's very useful. I also have a clearer idea of the DAME problem, in aspects which had been touched on but not quite addressed. In particular, where you can see the grid is going to go in the short term and the administration of the system is quite important and with cross organisational systems, the admin and the access control is important.

CD: Absolutely vital. If you haven't got that you haven't got a business.

DR: I think that you have explained with good example (and I don't see any reason that DAME would be any different) I think there would need to be some kind of central control of users. In particular, using the idea follow the money, Rolls Royce is being paid for the contract, but they have to protect the interest of everybody. So you can see why the central administration would work in this instance.

CD: You are right because administration goes further. You have to consider Rolls Royce's fundamental culture and business is going to change to a service-based culture. Now where you have self adapting system that are capturing knowledge and the administration of the site, you have got to be fairly astute because you are always going to be looking for up sells of services. So if you have discovered something new out of your extraction for knowledge it becomes a potential revenue stream as a new service. Not only that you could get into the point of view that the service offerings themselves have got a finite life. Therefore today, the whole of the DAME ability to do say the automatic aura-g search might be charged for at a certain rate. In five years time if all the other rivals have cloned that, operators may expect that as a free service. So you have to constantly re-invent the service level and adapt it to higher and higher levels as you are going along. Spotting those up sells and deciding to delegate some of those services down to a freebee is an art, a science in it own right.

DR: I hadn't quite thought of that. I suppose costing that initially must be difficult if you can't really see the lifecycle.

CD: The lifecycle of a service is obviously much shorter than the lifecycle of a product. You are dealing with quite complex assets. So One, having a handle on the lifecycle of a service, in terms of 'can I charge for it'. The second trick is in trying to understand on what the intrinsic value of that service is. Now, if I am going to give the diagnosis of a failure mode, then internally I can repeat that again and again. How much value is it to that customer? How do I get in a position where I am not charging for the cost base for the service I am providing, but I am charging them for the value. You have be able to have a mechanism to understand the value of what you are delivering as well. That's not easy.

DR: As an aside I think it is quite interesting that Rolls Royce has been able to change to a service model.

CD: Sixty percent of its revenue in civil airspace is on a service base. It is completely different product. It is incredible.

DR: It must have been quite difficult.

CD: Take a look at their share price. Take a look and see what it has done in the last couple of years, then you will see the story.

Some discussion followed about the roles and how their requirements affect the design of user interface. Also about data quality, as fitness for purpose, therefore how this is presented to the user. Including issues of style and control (data manipulation).

Repeated user and ad hoc user access, such that entry into the system would be different and alert requirements.

CD: Another thing about accessibility is single logon. There maybe a requirement that certain users access via another portal. We go back to the Lufthansa point of view, (conjecture). That the Lufthansa user might want to log onto their own intranet and have accessibility to the DAME user interface via their own intranet. What you don't want to burden those customers with is that they have to logon to the intranet, and then have to separately have to logon to the DAME system as well. You want single logon.

B.3 Interview Results Rolls-Royce

Interview Results from 12th November 2004

Issue: 2 Author: D. Russell Issue date: 7th December 2004

Purpose and Scope

The aim document contains a summary and transcript of the tape recorded research interview with Graham Hesketh that took place on 12th November 2004. This interview

conducted by Duncan Russell addressed the author's PhD subject area of 'Secure Collaborative Workflow'. Additionally Martyn Fletcher was in attendance to the interview to offer suggestions and opinion based on previous security work on the DAME project.

Summary

The following is a summary of the points discussed separated into categories for Collaboration, Workflow, Access Control, Roles and General Security.

Collaboration

Collaborations may take place by telephone, but would still involve DAME to determine context of collaboration (such as engine number and previous prognosis results). Using the system to identify the task/problem reduces risk of ambiguity. There would need to be a record of decisions made during the telephone conversation.

Security in the portal should not be restricted to current business process, but should allow for collaborations between people. Especially between different Maintenance Analysts, or between Domain Experts.

Collaborations between users of the same role has not been addressed, but needs to be managed within the portal. Access control should not prevent this from happening within the same organisation (at least). Simply put, this would prevent Maintenance Engineer from different organisations sharing data, but allow Maintenance Analysts all within DS&S to share diagnosis of a problem.

The collaborations between users of the same role, although not currently implemented, will benefit by using DAME in reducing ambiguity, by sharing a workflow instance and the pertaining contextual information.

Workflow

DAME central workflow accurately represents a simplified possible business of engine diagnosis.

Workflow definitions, considered as IPR, would require restricted access. Predominately the Domain Expert analyses vibration data and the Maintenance Analyst analyses performance data. The Maintenance Analyst can use the Domain Expert's results automatically if the Domain Expert's workflow can be captured and automated. Also, this aids training other Domain Experts.

Access control to services & data

Access control to assets is predominately viewed as restricting access to data.

Authorisation mechanisms would need to restrict user by Role and Organisation to view data, or service output or workflow output. Not as strong a need to restrict access to launch workflows or services, compared to retrieving results. Therefore users would be identified by the user id, but access restrictions would use attributes such as role and organisation, possibly extended by task or context (which may be derived from the workflow type/stage).

Restricting access to services, data and algorithms is currently viewed as feasible by using an in-house grid behind a firewall, possibly extended to DS&S. This could be an internally-shared processing/storage resource and would need policies on how the sharing takes place. If methods of obfuscating services and data are available, then execution of parallel tasks could be delegated to external grids. However, no business case for large parallel execution is currently foreseen in Rolls-Royce, except possibly MEROS at DS&S.

Roles

There are more roles in the business than those defined for DAME.

The three basic roles are a simplification on the possible user community, but provides a basis to incorporate other stakeholders with appropriate access privileges.

At the Maintenance Analyst level, sharing diagnoses provides benefits to the customers (albeit a benefit to one customer gained from experiencing problems from another), which is typical in a diagnostic environment. Therefore a Maintenance Analyst can view data across airlines, but should not be able to divulge specifics to another airline.

Maintenance Analysts and Domain Experts typically do not have restrictions on seeing customer data, but would not reveal specifics to other customers, either explicitly or implicitly.

General Security

Portal security is good by using x.509, but not in external tools with separate or no authentication/authorisation.

DAME would need a central control to administer: users and their rights, issuing certificates (or defining who can issue certificates), workflow definitions.

Grid raises the issue of confidential data and services potentially being rapidly available. Protection mechanisms are seen a piecemeal rather than a generic solution that applies to all parts of the system.

Interview Transcript

The following is a transcript from the taped interview between the three people:

GH = Graham Hesketh

MF = Martyn Fletcher

DR = Duncan Russell

Some details of the speech have been omitted in the transcript such as 'Err' and 'OK' and 'Yes'.

Interview started at 10:00am 12th November 2004, in a meeting room, Rolls-Royce, Derby and finished at 11:25am.

Introduction

DR: As introduction, the previous interview with Dennis King was more focused on the usability on the DAME portal. This is more focused on my PhD research, which is still referring to the DAME demonstrator, but is more focused on secure collaborative workflow within the portal and how you see that for a future deployed system.

Collaborative Workflow in the DAME Demonstrator

DR: First question, and this is aimed at the current demonstrator, Does the workflow in the Demonstrator reflect the expected business process?

GH: Now, the workflow that you are talking about is the automated response?

DR: Yes

GH: Which effectively kicks off the search process, the analysis process for XTO. Right, yes it does. That is essentially what we want to do. And I believe that came out of the initial requirements. We spent a lot of time coming to terms with what the business requirements were and constructed a process that would reflect that. And I think that is what the current workflow in the portal does.

MF: Do you mean the actual workflow that is executed for that automatic thing? Because the view was that it could be changed depending on what people thought.

DR: It is kind of encompassing not just the initial stage of it but the whole business process.

GH: You mean the escalation process and everything? Yes, it seems quite natural. It seems to work quite well. I have been through the portal on several occasions and followed through the scenarios. And yes, as far as I can tell for what we would like to do, although I am not an end users in as far as Dennis King is, it matches my understanding of the processes we would like to conduct.

MF: The only comment I have is that, to me it is not exactly clear the difference between the roles of different Maintenance Analysts and the Domain Expert. I am not clear in my mind the exact demarcation of those jobs.

GH: Does that come out in any of the later questions or should we tackle that now?

DR: There is an opportunity to tackle that later. So, with the current demonstrator does the secure access match the expected business process?

GH: That is a bit more difficult to confirm. We are currently using username password access, which is layered on top of the X.509 certification underneath. So you don't get a username and password until you have those underlying certificates created. That's an additional level of security which is good. But having got the password access, what I can't say is whether or not Rolls-Royce would be happy to deploy assets, as in datasets and algorithms which are only secured by that method alone. That I can't confirm.

MF: Also, the dependability study has said that the separate things that pop, such as the engine model, are a big no-no as far as security is concerned. Because you download executables onto your machine and have no access control from that point. There are criticisms of the current demonstrator bore out in this dependability and security study which are fairly valid to this stuff.

GH: Having said that Rolls-Royce is intending to have a corporate portal. Which will allow employees access to some aspects of the business systems. Particularly email, but also corporate intranet and that implies some of the company data. Now if they have a corporate portal over the internet, that will have to be protected again via probably just logging in through username and password connections. It won't even have X.509 certifications. So if they are going to be happy to do that, again it depends on the extent of the information that will be allowed to be transmitted over that portal. Probably not to the same extent as we are talking about with engine datasets, potentially customer information and engine models, algorithms and so on. That is unlikely to be, accessible through the corporate portal. So we are talking about requiring more level of security than that. And I believe that the demonstrator does offer that, but whether it reaches the high standards or levels that are required is another matter. I think what it has done is opened an avenue of opportunity that wasn't previously available and so not enough thinking has been done to decide what is an adequate level of security, given that opportunity.

DR: Good, Does the demonstrator represent all the organisations and roles in those organisation in the DAME diagnostics process?

GH: We have currently settled on just the three roles, although we did identify more stakeholders than that. I think for the purposes of demonstrating the usability and the functions that are available, the workflows that are available and the integration. I think those roles were sufficient. Because you can make the mental leap from there to incorporating other stakeholders with appropriate access privileges.

MF: Such as MRO people?

GH: Yes. So the airline customers themselves could be integrated into that model. Although they may have a different access point and different requirements, of course. But nevertheless, the idea of the virtual organisation that will allow different stakeholders to take part in the process, with appropriate levels of access, I think, is adequately demonstrated within this package that we have (within the demonstrator). And it's a straightforward extension to include other parties.

DR: Great, OK.

Collaboration Procedure

DR: Moving onto what you perceive outside of the demonstrator. Concentrating on the collaboration procedure, can you say what happens in the escalation of diagnosis problems between the Maintenance Engineer and the Maintenance Analyst? And then the Analyst and the Domain Expert?

GH: In what respect?

DR: What information is passed, how they interact, how you expect the collaboration to be?

GH: Within the demonstrator and to an extent possibly in real life as well, they would exchange messages to request additional assistance or to clarify issues, problems and so on. But in real life they may well pick up the phone and talk to people. It depends on the timescale of the problem they are trying to get a resolution from. If it needs to be urgent, then having a physical presence is generally a good idea. You can get somebody from just down the road to come on along and help out, that's one thing. Having electronic access to the same data means you don't have to do that. So I think what you have to do is consider it in the context of ways of working in other domains that use electronic means rather than physical relocation. So people collaborating on software projects or solving problems in modifying spreadsheets, for example. Might be more analogous to the way that we are suggesting that people might work in solving these diagnostics problems. Because previously they wouldn't have had access to the same data. Now they can, it opens up the opportunity for, as I say, picking up the phone, calling somebody as well as going through an automated system of messaging where you can put specific requests in and also point them at the right data without them having to fumble about trying to find it. By having a system where they can log in and you are referring to the same entities by the same labels and it is quite clear what it is you are dealing helps to focus the problem and the resolution much more efficiently. So, I think the mechanism that is already in the demonstrator, crude though it is. I think is effective at channelling the resources in the right direction to solve the right kind of problems.

MF: If you are using the phone, presumably you still want any decisions made through that phone call recorded for traceability?

GH: Again, it comes back to what current processes are. If you have got somebody on the ground at an airport that has an indication, from whatever source, that causes them concern and they want some back up confirmation. If they call a maintenance engineer at DS&S, or whatever, the first problem they have is identifying what assets they are talking about. Once you have done that you then got to identify what the data you want them to specifically look at. And what the nature of the problem is that you want them to give you an opinion on. So, all of these are potential areas of ambiguity in the current system. Plus, how do they access that data. It may be that you are asking for historical information. 'It appear to have this problem with this engine now, was there anything that you saw in the past that might give a clue as to what may have gone wrong?'. Again that is a fair enough question, so long as you get all the labels right, and so on. Once you have done that, for audit purposes you need to have some kind of physical record that the conversation has taken place or you've had a particular recommendation from a particular person. That, I don't know. Though that's current working practice that I can't really comment on. What I can say is that the demonstrator, the way that it is implemented, does lend itself to solving all those issues. So what it is you are talking about, the ambiguity is removed. And what has actually occurred is logged. So, that particular issue is removed, as well.

MF: I think, at one stage we got some criticisms for DS&S about the annotations being inside the DAME system and not part of the SDM. Would that lead to difficulties later on?

GH: We have to bear in mind, this is a demonstrator. It's not meant to be a final deployed commercial solution. The commercial solution, as far as DS&S are concerned, revolves around SDM. That's a business decision that's been made between DS&S and Rolls-Royce. And that's what we have to live with. What that implies, is that any future solution that we are suggesting, based around an integrated environment (we won't call it a demonstrator because it is a fully deployed commercial solution) will have to integrate with the solutions that are already agreed and in place. So it will have to be SDM based. What we are talking about here is an addition to the already proposed SDM2 tools. Which are the analysis suite of tools that sit on the infrastructure that has already been put in place. So, yes the databases that we are talking about should already exist and there should be facilities for incorporating the kinds of data that we need. Or at least relational links to additional databases that contain QUICK data and other engine data. Which may not have a place in the current databases but necessarily you could have links to those databases.

It should be possible to, in fact it is necessary to put all those things in place, so that the additional integrated tools that you layer in through a portal will be able to access

commercially available systems and will not therefore be at variance with the business processes that DS&S are already saying that are mandatory. It is a question of integration, and we couldn't do that with the current demonstrator in its current form. We didn't have the time or we didn't have the resources to do that. And it was also unnecessary. For an EPSRC funded project which is doing basic research with a steer towards an industry problems, I think it would have been inappropriate to go down the route of full integration with a commercial system like SDM.

DR: It would have been difficult to release the software, as well.

GH: Correct.

DR: OK. I think we have covered a few questions. There is only one left on collaboration and this is more to do with not so much the escalation, but the users who are operating in the same role in the system. What sort of information would they share between each other and what restrictions would there be on sharing?

GH: Do you mean multiple, say, maintenance engineers?

DR: Yes.

GH: At the moment the escalation process restricts the information to a chain of command. So , a particular maintenance engineer would escalate to a particular analyst. That's an arbitrary restriction that is placed on the demonstrator. It is a particular implementation of an approach that is modifiable, scalable and so on, such that you can have multiple maintenance engineers all allocated in the same escalation path. I haven't experimented with that. I don't know how easy it would be to implement. I am assuming that it is fairly trivial to implement. And I don't not know what the working practice impact of that would be. For example, if an engineer escalated that to one of a pair or a multiple number of maintenance analysts and said 'Can any of you guys help me out?'. That's a more robust solution, because it overcomes the possibility that the particular engineer that you are escalating it to isn't available. So obviously you have to take that into consideration. But it then opens up the question of what happens when multiple people all want to access the same data and all getting involved in the same diagnostic process. Would there potentially be any conflict there? Well I presume that there wouldn't, because what we are talking about here is read only access to data sources. So you may have the effect of bandwidth limitations, or if people are trying to access the same resources at the same time and they are already in use. So the duplication of access to potentially limited resources is an issue.

So, I don't think at the moment that the portal has anything built into it that helps to manage the process of multiple similarly capable people, or authorised people to access the resources at the same time. And it doesn't help to manage those in a coordinated fashion.

But, I think that is really an additional level of complexity which would have been intended to be addressed in follow-on work rather than at this stage. I think, at this stage, identifying that you have got this escalation path that allows you to have limited access by individuals to limited problems, is fine. How you then deal with how you duplicate this up is a separate issue. Which I think is an important one, because may well be that if you let loose a whole bunch of people all at the same access level, all at the same data, that has a potential for creating problems. Which you would hope that the infrastructure, the portal, the workflows and so on, would be able to give you handle on and help you with. And at the moment we haven't addressed that issue, the portal doesn't address that issue.

MF: After we did the use cases, we talked to you more about security. During the course of that, the view that I came up with and I think Howard [Chivers] came up with, is that at the maintenance engineer level, maintenance engineers work for particular airlines. They would be denied access to other airlines information. So there would be walls between maintenance engineers. At a slightly higher level, you would have maintenance analysts who looks over a few jobs, and he would be able to see information from across airlines. He may not be able to see specific things, but he may be able to see results of diagnoses. He may be working on contracts across airlines, so there might be a few of those. And domain experts would be able to see more. .

GH: That's fine and I think that process works. And it comes down to the attributes that each individual user that logs into the system has got and there access privileges. And that is feasible for the portal control and I could see how that could be implemented and work. What I don't know and I think because we haven't tried it is what is the impact of having people with access to multiple, or having multiple people with access to the same resource. As apposed to an individual having access to multiple resources. The issue that you are talking about there is where one maintenance analyst can see multiple engines from multiple airlines.

DR: So, they are controlling the conflict of interest?

GH: Yes. Well in fact for them it is not really…Well there is a potential for conflict of interest but as an analyst they are entitled to see that data. And they are expected to operate on it. For them it becomes an issue of ambiguity. Which particular engine are they talking about, which one are they dealing with. Because they have multiple engines that they can see and refer to.

I think the other issue. Which is where you have got multiple analysts all having access to the same engine data and none of them knowing which one is expected to operate on it. So when a request comes through from the maintenance engineer at the airport, an escalation

comes through we don't know which of the maintenance analysts are going to respond, and they don't know which is expected to.

MF: If you have a set of maintenance analysts who work on the same group of airlines and you give them a token. You say that is your job at the moment. If he is on holiday, or something, then it goes to somebody else.

GH: There are ways of working it through. My point is simply that the current implementation of the demonstrator, the portal, doesn't address that issue.

DR: Is there scope in the future to, not so much that the job is allocated to one person, but several maintenance analysts may be working together to solve a particular problem.

GH: They might do. Again, it depends on there level of expertise. Domain experts themselves may be few and far between, thin on the ground. So getting multiple domain experts together to collaborate on a problem might be an issue. But analyst , yes there could be multiple analysts that get involved. Again, the portal at the moment doesn't have the facility within it, that one analyst for example could escalate it another analyst. You have an escalation to domain expert but not really calling the services of another analyst.

DAME Roles

DR: So here's the opportunity to expand on the roles. Are the three defined roles sufficient to define access control for collaborations?

GH: You mean for a full implemented commercial solution or just for the purposes of the demonstrator?

DR: For the full system.

GH: Probably not. Again, following from Howard and Martyn's analysis, there is a much more detailed investigation of the stakeholders and potential threats to the system. Which will necessitate a more detailed and structured approach to annotation of access privileges and protection of individual assets. So I suspect that the breakdown that we have currently got is not detailed enough in order to provide all of the necessary levels of access restriction, that would be necessary in a commercial solution.

MF: An example would be a Maintenance Engineer that works for a particular airline. So the role of that person is particularly important, the actual airline he works for (the organisation he works for) is important. Because that determines what can do and what he can't do. At the Maintenance Engineer level there's other attributes as well.

GH: My feeling is that, the particular structure that we have adopted in the demonstrator is intended to be the final structure for a commercial system. It is intended to demonstrate the capabilities of a distributed environment with appropriate access restrictions and a level of

collaboration between the different parties. It demonstrated the principle of it. And I think it works in that respect, it shows what you can do. It then points to the fact that a more detailed analysis is required in order to fully specify what we need for collaborative working in a proper commercial system. The outcome of that will enable somebody to specify what you need for a full system, including all the roles and access rights and levels of collaboration, and so on. Which is a full mixture of the functionality that's required and the level of security that's required at each level. That I think is feasible. It's not done currently within the demonstrator. The demonstrator merely points the way to how it could be implemented.

DR: This is going back to the last point we were talking about. It is more concentrating on access control between roles of the same type again, sharing a process along with sharing data. What restrictions are there be on, say Maintenance Engineers sharing. Martyn already said you would put walls between the airlines. And going up to the next level what restrictions are there on Maintenance Analysts sharing a process?

GH: Sharing a process? I don't think so. In that the processes are deployed appropriately. For example, within DS&S, they may well be servicing multiple airlines, with multiple different engine types. And they might not all be Rolls-Royce engines, there may be engines form different manufacturers, but DS&S nevertheless have the contract on providing the diagnostics support for. All the processes that they operate, they would expect be used by whoever requires to do so within DS&S. So they are not going to distinguish that a process is only for a particular airline and not to be used across airline, or by different analysts.

MF: Did you mean process? Or did you mean the information about a particular diagnosis?

DR: Kind of both really.

MF: I appreciate that you only have one process that you use, probably for Maintenance Analyst,Maintenance Engineer, the process or the working process that they using is the same. Not across levels, but on a level. Do you mean the actual data, say for instance a Maintenance Analyst looking a Singapore Airline and another one looking at… Is that what you mean?

DR: If we go back, I think what you are describing is the process definition.

GH: To give you an example, if an analyst within DS&S is examining fleet data for a particular engine type from an airline. And recognising that there is a potential problem, that is occurring and the get to the root, or believe that get to the root of the problem. And identifies a danger, a hazard. Then you expect them to translate that into all of the other fleets, other airlines that are using that same engine. And again, to follow that up with an analysis, a risk analysis, potentially provide advise back to other fleet operators as well, that

they may be remedial action required. That is taking information from one specific airline and translating it into recommendations for another. They do that because it is the same people involved, whether or not the is a process that says you need to do this. Because it is the same people involved they will naturally do that. Now , whether you want to automate that process so that whenever analyses come up with conclusion that are applicable to an engine type then translate that into workflows and recommendations for other airline customers that use the same engine type. That is a different matter, but the specific data relating to a specific operator would not be divulged to other operators, but it is going through the same people. An analyst is aware that there is a problem on a particular operators fleet, then extracting what is common information from that, which is it is something to do with the HP compressor, and then translating that into analysis of somebody else's airline data, to see whether or not there is an HP compressor problem. And if so, informing them of what needs to be done. There is an aliasing there of the information. So not telling this fleet operator that somebody else's fleet has a problem with its engines, they are simply providing a recommendation based on there own data. But it is an analysis which is inspired, if not the conclusions drawn from somebody else's data.

That I believe is a legitimate business process that is conducted behind the scene. What we wouldn't want to do is construct a system that would prevent that from happening. Whether or not you would want to construct a system that would help to make that, to automate that process is another matter. But you wouldn't want to prevent it from happening.

DR: So at the analysts level, if an analyst is working on a particular problem and would want another analyst to help with that problem, are there any restrictions in the business.

GH: Not that I am aware of. I think within DS&S they do not have that kind of Chinese wall, that says, these people only deal with these problems, but I may be second guessing here. So you can't take it as verbatim that it is true. But, to the best of my knowledge that isn't done. What they would do is have Chinese walls between customers. So that they would be communicating customer information from one customer to another.

MF: So, essentially DS&S probably have, as far as you know, DS&S have a bunch of maintenance analysts who can look across any engine that DS&S might have data for.

GH: They have a predictive services business and it's their job to do the backroom analysis of health monitoring data. And those people may expertise in particular engine types, they may have generic expertise across all engines types. And may well be able to personalise it where necessary. But typically you would expect an individual that has expertise of 535-E4 engines, the RB211's, to concentrate on doing that. Rather than expect them to do Trent 500's or Trent 800's, or whatever. But nevertheless they can be called in where necessary. So it may well be that there is internal demarcation…

MF: To match actual domain knowledge.

GH: And it may also be done by a fleet operator as well. So particular people have responsibility for particular fleets. But I don't think there is anything that explicitly prevents them from calling-in assistance from other analysts, as and when necessary, or the opportunity arises.

DR: I assume that's a similar practice at Rolls-Royce?

GH: Yes.

DR: You would be able to look across anybody's data, but you are not going to reveal specifics.

GH: That's right. And they do. Dennis looks across pretty much all engine types. Lessons that he learns on one he will implicitly transfer, if not explicitly, across to the others. And hopefully a system like the one that we have been developing within DAME would provide the repository for that kind of knowledge to be encapsulated to be reused more automatically. And also by other people. So, Dennis has learnt something and incorporates it into the system in such a way that the next Domain Expert that comes along that's expecting to address a different problem for a different customer. They can still make use of the knowledge that's been gained by an exercise that Dennis has done. And it's really the lesson that's been learnt, the knowledge that been incorporated which is the important aspect. The fact that it's for a particular customer may well be completely aliased out. There may be still records that that is were it came from, but that isn't necessarily made available, it doesn't have to be made available for the next domain expert that wants to use it.

Workflows

DR: Now we are looking at processing resources and the workflows that are using the resources. So, How should the processing resources be managed securely?

GH: The simple answer is to do it all in house. We have enough resource, or you make sure you have enough resource available to meet peak demand. But because it is all in house, or within trusted suppliers (potentially within DS&S) then you don't have the security considerations of where is my data going, where is my algorithm going. I think communication over the internet or over a grid can be made secure through encryption. But it has to be decrypted somewhere and where it is decrypted and run you have got a potential threat to security. Whether or not in the short term we are going to find ourselves in a situation where that becomes an unacceptable risk. That's a different question. I don't know I am not the person that would answer that. But, for the shorter term one resolution to the problem is to provide a distributed set of resources with are inside the Rolls-Royce firewall, or inside a trusted parties firewall. Now, what that does mean, it still gives you the

opportunity to use those resources for other things. So, you don't have to put in a dedicated resources for doing engine diagnostics. What you can do is put in a multi-purpose resource which is capable of doing that and other functions within the company as well. And that's really the short term expectation that we will be able to provide addition computed resources and storage resources which are intended to be more general purpose. And shared among a variety of users. So, that we still get optimal usage, but still have the capacity to be able to respond to short term demand.

MF: I don't want to keep going on about this dependability and security study. But, one of the conclusions out of that, for things like the engine model you would want to keep it in house any way. Because you want to keep the algorithms secure. But things like Aura you might want to put out onto the grid because the data when it is out there is not really understandable.

GH: The purpose of the security analysis amongst other things is to compartmentalise the risks and identify those bits which are amenable to being taken outside to a less secure framework. Having done that it gives you the opportunity for saying if you want to have a full diagnostics system then these are the parts that have to be maintained securely, these are the parts which are less secure. And therefore, it helps you to map out the distribution of those resources. The easiest thing is to keep it all within a framework that you have control of. But, with the eye to the future of identifying which bits can be farmed outside as and where necessary.

DR: OK, the other issue is obfuscating the operation itself, so you don't know what the algorithm is doing. So that you would be able to farm that out.

GH: That is a possibility that we raised in one of the meetings at York when talking with Howard. That you have the processor is executing instructions which it doesn't really understand, it doesn't know what they are doing. The net result of the operation appears opaque at the outside. So everything is effectively encrypted, from the point of view, the data comes in, the algorithm comes in. It is executed on a processor, the results are turned round and sent back. And it's only when it is decrypted at the other end that you can tell what it has done. Whether that is feasible, or not I don't know. I think I had some reservations about whether that can actually be done or not.

MF: I know there is some work at Leeds on that.

DR: Yes, Erica is working on that. And I am not sure to what level it is secure but the operation itself is feasible.

GH: That will open up a range of possibilities.

MF: I think if you had a situation where you have and engine model and you want to run hundreds of them. Then you may run out of resources inside. If you get into that situation, that's when you want to start farming things out. But then you have to be really sure that the security is okay.

GH: I think the issue of whether or not you run hundreds and hundreds of engine models comes down to opportunity in that historically we've come up with the approach that if you don't have that resource, you don't rely on it. You don't run hundreds of engine models simply for the sake of it because typically that resource isn't available. So you build your processes around not being able to do that. If it becomes an opportunity that could run hundreds of engine models, you have then got to consider, well what is the benefit of doing so. And what is the cost and risk of doing so. You then look at the cost of farming out those models and the risk of doing so on secured machines on the Internet for example. And look at the benefit you got from running all those models. And given that the way that the business has evolved and we have had to do without that opportunity the question then arises 'are our business models tuned to not needing hundreds of engine models to have been run such that the benefit of having that is now much reduced?' So it would take time before you could say, We evolved the process along these lines now that we know this opportunity arise what kind would we be able to get. So I can't say for certain that it is something that would become immediately appealing.

MF: Some people suggested running lots of engine models. I am not sure whether it came from a business case, or a desire to have a massive computational problem.

GH: Well, if you look at something like DS&S's business model and you look at some of the other processes, in particular MEAROS, which is looking at engine [unknown from tape?], or modules being required for shop visit replacement. That is something that parallelises well, because you have got a probabilistic framework that you need to sample. And you need to come up with statistical averages panning out into the future in order to determine whether the particular parameters that you are suggesting are the best. That does lend itself naturally to be able to find, garner compute resources out there to run each one of these models. Again, it comes down to the sensitivity of the data that is in that model, whether or not you would want to be running that on somebody else's machine. But clearly there, if you can crack that problem, then this is a good opportunity, because this is the kind of thing that DS&S do want to do. They do need to run multiple MEAROS models. They do need to have it encapsulated within an optimisation framework.

DR: OK. This is probably fairly obvious, I just need confirmation answers. What restrictions would there be on executing workflows from the different roles? This is a

workflow that is available to run and just making sure that the obvious fact that you would restrict a particular role.

GH: You would restrict it. But what would be the down side to either deliberate or accidental access to it. There probably isn't that much. If you run a workflow, it generates data and that data sits the system. Now data itself is still access restricted. So, even if somebody either deliberately or accidentally ran a workflow that they weren't entitled to that still wouldn't mean that they would be entitled to look at the results. Because they would be secured by other access permissions. It may have an implication in terms of using resources, compute resources that you wouldn't want to happen. So you might consider the situation that the system, inadvertently or incorrectly kicked off a workflow that it shouldn't have done because it wasn't necessary it generates results that really aren't useful. That's a waste of resources. So that's one issue. But the second one would be if an individual in the system tried to kick off a workflow that they should have had access to, what would be the implication of that? Well I said one thing would be a waste of resources, but the second one would be if the data that is generated from that is still access controlled then that individual still wouldn't get access to that even though they had access to the workflow. Although it would depend on how they've managed to get access to the workflow. Whether it's a breakdown in the system somewhere where a privilege has been set or unset incorrectly. Or that person has managed to change their own access restrictions so that they can now see that workflow, and kick it off. If they have done that, then may also be able to see the results. But that I think is a more generic security consideration. That's an individual that has managed to acquire access privileges that they weren't entitled to.

DR: Here is one the extends to the workflow advisor and CBR. So, what restrictions would there be on sharing workflow definitions?

GH: Definitions? Can you clarify what you mean by a workflow definition?

DR: Currently we have the automatic workflow, but you could envisage the results of either captured workflows and using workflow advisor and produced a definition. Or somebody who has organised the tools in a certain way and produced a definition. So what are the restrictions on sharing that definition and possibly how you go about defining access for roles to both execute and access.

GH: I think that there's an issue there of IPR and knowledge capture. That the workflow itself encapsulates the new knowledge. An individual has recognised a new way of doing something or there analysis of particular data that is pertinent to that situation, leads them to believe that a particular workflow is the right response to that and it turns out to be the case. Then they have created new knowledge and that is valuable. And the organisation that has done that would want to protect that for their own purposes. So, if that was done within

DS&S they would probably expect that other DS&S analysts should have access to that IPR and that new workflow. But not other maintenance analysts that might be collaborating through the portal from other organisations. Similarly within the Domain Experts. If you have one Domain Expert that has created a valuable workflow, then you would expect that to be accessible by other Domain Expert from the same organisation. Whether or not you would want to allow Maintenance Analyst from a different organisation to also have access to that work. Now, that's a tricky situation. DS&S is not owned by Rolls-Royce, it is only part owned. They are a joint venture organisation, a separate body. So we have commercial relationships between us. So, it's not everything that we do that have to get access to and vice versa. So, in that sense they may well want to retain IPR on workflows that they create. And not make them directly accessible to the domain expert, without some kind of collaboration agreement. Because at the moment we have the Domain Experts that have access to pretty much everything. So, naturally I think the way that the system is implemented at the moment, a workflow that is accessible to an analyst will also be accessible to an expert, a domain expert. And if new knowledge is created within the analyst domain, they may not necessarily want that to be directly available in an expert without some other considerations. You are right, it is not a given that it should be an open process along the escalation path, which it currently is at the moment. It may well be that you will want to be able to label and protect individual workflows that are not available off the shelf within the implementation of the system. So, that if it is something that is created after the system has been implemented, you may want to label those with access privileges.

MF: There would probably be an owner to it, an owner attribute.

GH: Yes.

DR: Or maybe projecting into a business agreement that it would be part of a package, and that being an attribute.

GH: Well that's right, yes. And if the controls were in place, then they could release individual assets that they have within the system that are labelled as such, to other trusted parties on an access control basis. And that would be subject to external agreement. But, nevertheless, that would be implemented within the system.

MF: If a maintenance Analyst comes up with a workflow that does a particular thing, would it be aimed at deciding when maintenance is required or aimed at finding out what the diagnosis was? What I am trying to say is, at the Maintenance Analyst level are they more interested in the actual good management of maintenance activities and the Domain Expert more interested in the diagnosis of unknown faults. Therefore the results of a workflow at the Maintenance Analyst level would not be directly of interest to the Domain Expert, is what I am trying to say.

GH: there is a bit of overlap. The Maintenance Analysts are people with domain knowledge as well. There are not just people who turn a handle. There are people who have fairly intimate knowledge, in a lot of cases of the gas turbine itself, performance characteristics, gas path analysis.

MF: So, they would be into producing workflows to actually produce a diagnosis.

GH: Yes. What they would is they would be predominately concerned with capturing repeated actions and automating them, automating the analysis. So, that it just speeds up the process for the future. If they have got a particular pattern of data they are not sure what it implies, then they will go through an investigation process that may or may not involve a Domain Expert. They may well be able to do it all in house. But, once they have done that and they have verified that there is an approach that they have adopted that works. They will then want to capture that an reuse it. And if that can be automated within the system, then the next time that it occurs, they should get a much quicker resolution, involving less of their time. Now, primarily the distinction between the Maintenance Analyst and the Domain Expert is crudely the distinction between performance analysis and vibration analysis. That's really where the distinction comes, in QUICK and within DAME. In that, the Domain Experts are the ones that know in detail about the vibration information and what its ramifications are. Whereas the Maintenance Analysts tend to be more concerned with performance analysis and that's the one that gives them the long term information. Now, they will be getting some indications from the vibration analysis that QUICK is currently capable of delivering, that's over and above what they have been used to in the past. And that will help point them to particular resolutions, and they will want to use that. But, they want have the in depth knowledge of where that information came from. That's the Domain Expert area.

So, when comes to creating workflows and automating a particular analysis and resolution they may well be making use of summary information, features, from the vibration data. So, it's primarily their knowledge of dealing with performance issues that will drive the workflows that they are creating. The Domain Expert is much more concerned with in-depth analysis in the fine detail of the vibration data. And, they, as you say, are not so concerned so much about the workflows. They are hands-on. They want to get in at the very fine detail level, where it is much more interactive rather than. Well this analysis is done off-line, these are the results in summary. The time may come when the systems are intelligent enough that they can actually replicate what the Domain Expert is doing in fine detail. And they can do it well enough that it isn't necessary enough to call upon their services. That time hasn't come yet, but you can envisage that it may do in the future. And therefore being able to have documented reusable workflows that duplicate what the domain expert has done. Will then help other Domain Expert to quickly go through that process without too much manual

intervention. So, I think the two workflows are quite different and distinct and there may not be a necessity for one to kick off the other. You can imagine that there is still this escalation process. You can effectively imagine the situation where the Maintenance Analyst has already had the benefit of the automated workflows that they have been creating and still wants help from the domain expert, so they still haven't got the final answer. In which case, when the call upon the Domain Expert, the domain expert will be able to see the results the benefit of the previous automated workflow, without necessarily knowing what it entailed. So they might not have had access to it in any detail, but the see the results and they understand why they have been called in to provide their advice. It would probably mean that they wouldn't need to initiate the workflow themselves, because it has already been done.

MF: Is the Domain Expert not too concerned about performance?

GH: The Domain Expert does know about performance, but all of the knowledge required to draw conclusions about performance issues, probably resides in the Maintenance Analyst at DS&S.

MF: So, as far as running the engine model is concerned, who of those two roles would actually be interested in running the engine model.

GH: That's a good question.

MF: I've always, I thought, that the Maintenance Analyst would maybe run the engine model but not in a very interactive way. And the Domain Expert would run it in a very interactive way because he understands everything about it. But, that's probably a wrong assumption from what we have just said.

GH: Again, it comes down to the levels of distinction. In that, I think there are gradations with the Maintenance Analyst domain, they are not all the same. Some Maintenance Analysts will deal with high level data, high level performance data and not go into too much detail. Other Maintenance Analysts may have an inclination to get more in-depth and those are the kind of people that might want to run the engine models, to look for specific indication in the performance data. Because currently the engine model doesn't provide vibration data. It is just performance analysis. So the people primarily doing performance analysis calculations are the Maintenance Analyst. And they are probably the customers for the particular service. That isn't to say that the Domain Expert wouldn't want todo that or need to do that. They may well do, but they will be doing that probably in conjunction with other analyses. In particular, looking at QUICK data. So, I that there is a range of requirements.

MF: Apart from the fact that we have allocated the Domain Expert to Rolls-Royce and the Maintenance Analyst to DS&S, around those two there is probably a fuzzy boundary. There people in DS&S coming up to Domain Expert.

GH: DS&S may want to get more experts who are capable of doing in-depth analysis of QUICK data. Or they may just want to outsource that back to Rolls-Royce, and that's the business model that they want to stick with. Similarly within Rolls-Royce we've got plenty of performance experts who could do the role that DS&S do, but don't at the moment because that is a function that DS&S get involved in. But, nevertheless any particular problem that comes along that gets escalated to a Domain Expert in Rolls-Royce. They may well call upon resources within the company that look at performance issues rather than vibration issues. And that will always be the case. So it isn't that all the performance analysis is done with DS&S and all vibration analysis is done within Rolls-Royce. It's just a natural distinction at the moment because that is where the distinction lies.

DR: That kind of brings up an interesting model. When you said outsourcing the expert you could almost outsource the Maintenance Analyst back to DS&S if you have people who can fulfil the role.

GH: That is where the role is currently conducted. Previously before DS&S were formed Rolls-Royce were responsible for providing that Maintenance Analyst role to the airlines. When we set up the joint venture company the staff that went into DS&S came from Rolls-Royce. It was those people that had that function within Rolls-Royce that went over to setup the new business. But previously, when they were in Rolls-Royce, they still called up the services of other people within Rolls-Royce. Now that they are within DS&S, they can still do that to an extent. Because we have very good cross linkages between the two organisations. But now it is a cross company link rather than an intra-company link.

DR: This one we have kind of covered, but I will ask it anyway, just in case you can think of something else. What restrictions would there be on sharing or collaborating in active workflows? Which we have kind of discussed with the exchange of roles.

GH: Although I said that the workflow itself might contain valuable IPR. And that therefore the individual organisations might want to protect it. The likelihood is that if we are collaborating at this level of using an integrated tool for virtual organisations. Then, we probably may well find ourselves in the situation where we don't want to restrict access to the workflows, that we don't want to consider them as protected IPR that our partners shouldn't have access to. It is probably the case that it is going to be an open collaborative venture. That the workflows may well evolve out of the combined efforts of the partners, rather than any one individual. So, yes, if what you are doing is a process that involves multiple partners, and the net result of all of that is a captured workflow. Then it is really

joint ownership, at that point. That may well be the way that it is seen. And subsequently if it is only going to be utilised within that environment then it will be for the benefit of both parties anyway.

MF: It is probably for the overall benefit of both organisations in that respect. You might have some open-source workflows within the organisation document.

GH: It is an interesting question though because the whole idea, the whole opportunity of having virtual organisations collaborating together. Sharing data. Sharing development, and so on. Does raise the issue of IPR as who owns it. Who created it, where does it reside.

DR: That actually leads nicely into the next question. I shall throw it in here, not to stop you at this point. Who can define new workflows for DAME?

GH: There has to be an authority maintaining what DAME is. And that will reside with a nominated authority. That authority will be the one tasked with maintaining the system integrity and augmenting it in terms of functionality. However, who could suggest what might be a new appropriate workflow is probably unlimited. That any of the users of the system that are currently involved in deriving benefit, using day-to-day or whatever, would have the opportunity of interacting with it and creating potentially new workflows that could be of value. You still need to evaluate whether it would be useful, and that is a separate issue, as well. Somebody has to take the decision, yes, this workflow has been used umpteen times and its providing value. Therefore we would like to capture it, and put it in our library and even kick it off automatically under the following circumstances, etc. Somebody needs to do that evaluation and it may well be the role of the Maintenance Analyst primarily that would come to that conclusion. But it is whoever is responsible who is the nominated authority for maintaining the integrity of the system that would have to decide on whether it gets incorporated and implemented. I think anybody can suggest it.

MF: But, the automatic workflows that people would run, presumably that would be fairly well controlled. Strictly controlled, because you don't want some half baked stuff getting in.

GH: Clearly not.

MF: So that would have to be really well controlled.

GH: Yes it would.

MF: I am not saying that the other stuff doesn't have to be, but…

GH: Any user of the system can effectively produce a workflow by initiating a set of actions that they are entitled to do. But whether there is any net value in it that is worthwhile capturing is another matter. If they do that over and over again and every time they do it and every time they believe they are getting benefit out of it. Then capturing what is that they

have done, documenting it and putting it in as a suggestion to the nominated authority that says 'This is a workflow that I would like to see stored within the system. I don't have to keep on doing it myself I would just like to call upon it. That would be a useful function to have. Somebody at some point might take the decision that it needs to be an automated workflow because its value is such that we would really want it to be automatically executed. And again, that would not be the decision of an individual user. That would have to be from whoever is responsible for maintaining all the system.

Access Control

DR: Good. OK, How users should login to the portal, I think we have already discussed that one. So, how should organisations provide user identities?

GH: (long pause). Well that I am not sure about. I got a certificate to access the system because I knew the person who was capable of providing those certificates. I suppose the issue becomes if, say, somebody in DS&S said we need to get a certificate for an individual to allow them access to the system. What is the process we should undergo to grant that? That I don't know. I think it comes down to whether we are talking about a commercial system or whether we are talking about a demonstrator.

DR: Really, the commercial system. So, this would encompass all users, those from the airline as well.

GH: Well, they would have to be, as well as a certificating authority within whatever the grid system that we are using, there will have to be an authorising mechanism, a nominated person, a body, that would have to deal with who is entitled have access to the system and how they would verify that. That would I think depend who is considered to be the owner of the system. So whoever is the owner, deploying the asset. They would have responsibility for delegating access privileges and verifying that it is the right people that are getting them. I mean, a process for doing that would have to be drawn up with whoever the security representatives are in those organisations.

DR: Good. Who is responsible for assigning roles to the users?

GH: Again, I think it would be the same body. The one that would be expected to validate a request for access, would also be the one that would validate the role. It would be one and the same body. You would expect it to be. You cannot have somebody saying yes we will allow another person access to the system, but we will leave it to somebody else to decide on what their access privileges are, their role is. It would have to be one and the same body.

DR: OK, fine. I think this one poses quite a wide scope. How should the roles be defined?

GH: Within the current portal demonstrator, we have had a stab at it. In that, we have come up with three representative roles that demonstrate that a role based approach is feasible and

works. In a fully deployed commercial solution. I think there would have to be two things. One I think there would have to be a detailed analysis conducted to identify the set of roles which is appropriate for the solution. And also you would need a mechanism by which that can be changed and augmented. So I think its two things that a required there. One is an initial analysis and a specification and an implementation. And on top of that a mechanism for changing that. I Because it, the chances are that the role based analysis that you initially come up with isn't the one that is ultimately going to be the most efficient, or the one that you need when you put the system into practice. So will you also want a mechanism for being able to author it. I think if you have got a mechanism for authoring it, it becomes less problematic and less of a concern that you get the initial analysis right. So, if you start, if you bootstrap the system with a set roles in there that will allow initial users to access and use the system. Then as it evolves and you identified that role set, sorry that set of roles is not the most efficient or doesn't in fact meet all the requirements, then you just modify if. And that may mean changing the actual access privileges of the previously identified roles or simply adding new roles with different access privileges and functions. Then, so long as you have got that mechanism that you can change the roles that are already there, or you can augment them with new roles. Then it becomes I think a reasonably straightforward process to implement one.

DR: Good. This is quite a wide question, but I don't expect too wide an answer. What restrictions are there on sharing data and services across organisations? Or in fact, to put it the other way round, what scope is they for sharing data and services across the organisations?

GH: Well there are big concerns, obviously. Commercial concerns where we have got customer data. As well as, our performance data, the algorithms all of those things are address by the security analysis that Martyn and Howard have done. We know the kind of scope of the problem that we are dealing with. We also know to what extend we have measures for dealing with that, currently in place. Physical security of items, non-disclosure agreements with people, electronic access, encryption and so on. All of those things are currently physically in place. The problem of the grid is that it provides rapid access to an awful lot of resources and confidential data, potentially confidential data. That was previously not available. It raises a whole new set of problems. I think that it will, the issues will be slowly addressed. There won't be an immediate, this is it, we now know how to do it, and therefore we are going to implement this, this method across all of the access routes. I don't think that will to happen. What will probably happen is that there will be a piecemeal approach. In that, access to individual resources, data, information will be made available under specific circumstances and under specific mechanisms and they will be tended to be done on a one by one basis. So where there is a requirement a necessity for access to that

data and information and there is a business case for it, and we have a driver. Then, an appropriate mechanism will be found and deployed. The lessons that you learn from that may well make it easier to deploy the next dataset, information, resources, etc. But I don't that what they are going to do is withhold all of them until they've got a universal solution that is going to work for anything. I think that what they will do is on a case by case business basis they will find a mechanism for deploying access and learn from that for the next resource and so on. Until eventually, evolving out of that will become a generic solution, that will then be available for all. I think it will happen that way round. I mean, we have already seen during the lifetime of the project that we have delivered datasets on CDs, we have provided modified engine models. We have provided access over the internet and so on. Each one of these is just an evolving set, they not a universal solution. But they are a pragmatic response to an individual problem and I believe that's what we will see more of.

DR: Excellent. OK, that was my last question. I think that to summarise. I have got a very good overview, and from when I interviewed Charlie last week, and this has given me a view from Rolls-Royce perspective. It is extremely useful in validating my model. So, my next step is using the information…

Any further comments you wish to add?

GH: I suspect that quite a few of the issues that we have raised are going to be addressed within Broaden. Because we will to be talking about deploying a system along the lines of an internal, or close collaboration with us and DS&S, type model. Nevertheless there will be a range of roles and accesses to be required to be defined within that. And so, some of the questions that have been raised and the possible solutions that we have discussed we possibly see the light of day in the next few years.

DR: Thank-you very much, it has been very useful.