SIMDAT

Data Grids for Process and Product Development using Numerical Simulation and Knowledge Discovery

Project no.: 511438

Grid-based Systems for solving complex problems – IST Call 2

Integrated project

SIMDAT

Deliverable - Draft

D.11.3.1 Updated Aerospace Activity Scenario and Requirements including analysis of data items and process workflows.

&

D.11.3.2 Updated Prototype System Description and Capability

Start date of project: 1 September 2004 Duration: 48 months

Due date of deliverable: 2007/09/01

Actual submission date: 2007/10/08

Lead contractor for this deliverable: BAE Systems

Revision: 1.0

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination level		
PU	Public	
PP	Restricted to other programme participant (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	X

Revision history

Date	Version	Author	Modification
20/01/07	0.1	M Turner	Initial Version
24/09/07	0.2	W Song	Workflow advisor update
24/09/07	0.3	M Turner	Updates to data manager
25/09/07	0.4	G Alleon	Addition of Section 6
26/09/07	0.5	M Turner	Addition of Section 9
02/10/07	0.6	M Turner	Update following reviewer comments, additional of sections 9,10 and 12.
08/10/07	0.7	M Turner	Minor Corrections
08/10/07	1.0	M Turner	Final Version

Copyright

Table of contents

1 Definitions, Acronyms and Abbreviations

1.1 Acronyms

PKI	Public Key Infrastructure
PSE	Problem Solving Environment
QoS	Quality of Service
CA	Certificate Authority
TLS	Transport Layer Security
RSM	Response Surface Modelling
SLA	Service Level Agreements
VO	Virtual Organisation
URI	Uniform Resource Identifier

2 Executive Summary

This document describes the Aerospace prototype for the Knowledge phase of the SIMDAT project. The aim of this prototype is to make better use of the product and process knowledge within the system. This document describes the architecture of the prototype based on the 5.1 release of the GRIA middleware and outlines the use of a new data management layer. It describes how surrogate data models are being used to decrease the total simulation time in the system. A framework for process support using a workflow advisor system is described. A testing and monitoring framework to give feedback about the current and historic state of the system is also described.

3 Introduction

This document is a report that describes the aerospace prototype implementation for the interoperability phase of the project as specified in SIMDAT Annex 1- “Description of Work” [1].

The plans for the knowledge phase of the project for the aerospace activity involve making better use of the knowledge being generated in the previous interoperability phase prototype. Two types of knowledge were identified they were product knowledge and process knowledge. Product knowledge is the knowledge at the analysis services level and is a combination of specialist knowledge and automatic data extraction. Process knowledge is the knowledge used to construct the workflows used to tie the analysis services together. The aim of the aerospace prototype is to show how SIMDAT technologies can be used to make better use of both forms of knowledge to reduce the total time of simulation and demonstrate process knowledge reuse.

The basic scenario remains the same as the interoperability phase use case described in SIMDAT deliverable D11.2.3 [2] and involves the multidisciplinary optimisation of a high lift system to try and reduce the generated noise whilst maintaining its flight characteristics. The analysis services are provided by the partners at distributed sites and the optimisation is driven by the University of Southampton. In the scenario a VO is constructed with BAE Systems playing the part of Prime Contractor subcontracting the optimisation to the University of Southampton who then enact the analysis services on behalf of BAE Systems.

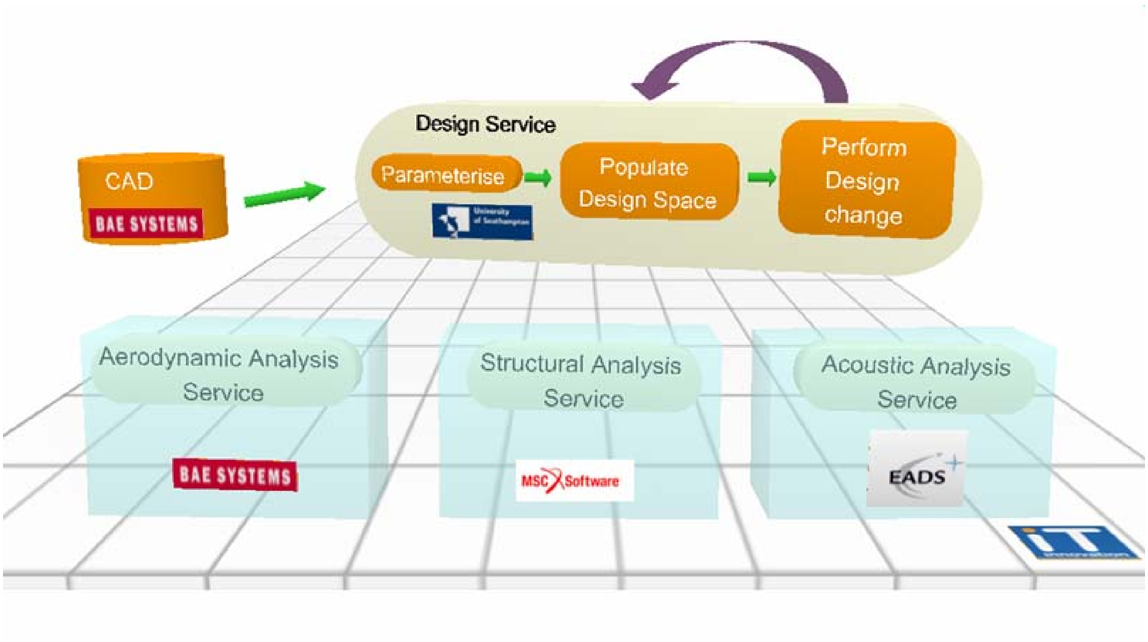


Figure 1 Scenario Deployment Overview

4 Architecture

The knowledge phase prototype will be built using GRIA 5.1. All partners will be moving to this release of GRIA from the previously used GRIA 5.0.1. The upgraded version brings additional services that the aerospace partners wish to use to manage the scenario VO. Figure 2 shows an overview of the software architecture of the aerospace prototype.

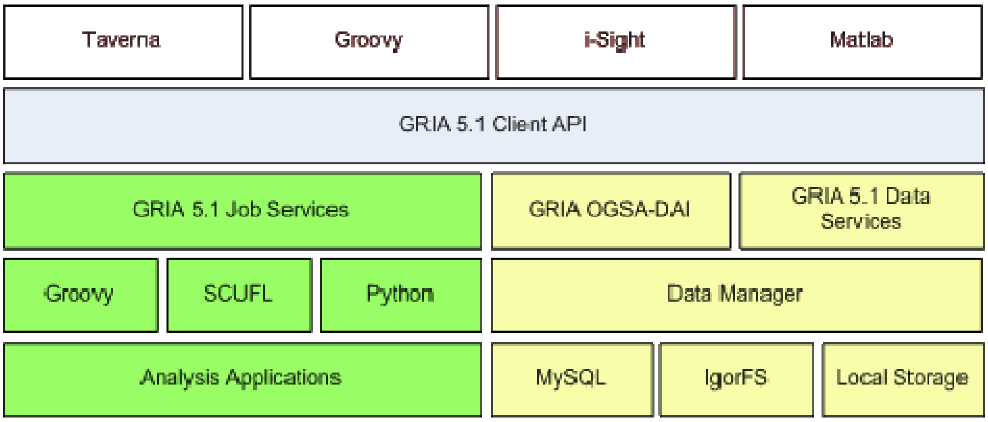


Figure 2 Aerospace Architecture overview

It shows how we are still using a variety of PSE’s and scripting languages to access the job and data services. The architecture diagram also shows the new data manager layer accessible via the GRIA OGSA-DAI interface of the GRIA data services. The data manager and how the analysis services interact with it are described in section 5.1.

As well as the new prototype architecture there have been changes to the surrounding infrastructure. The certificate authority has been changed from OpenCA [3] to a more light weight alternative.

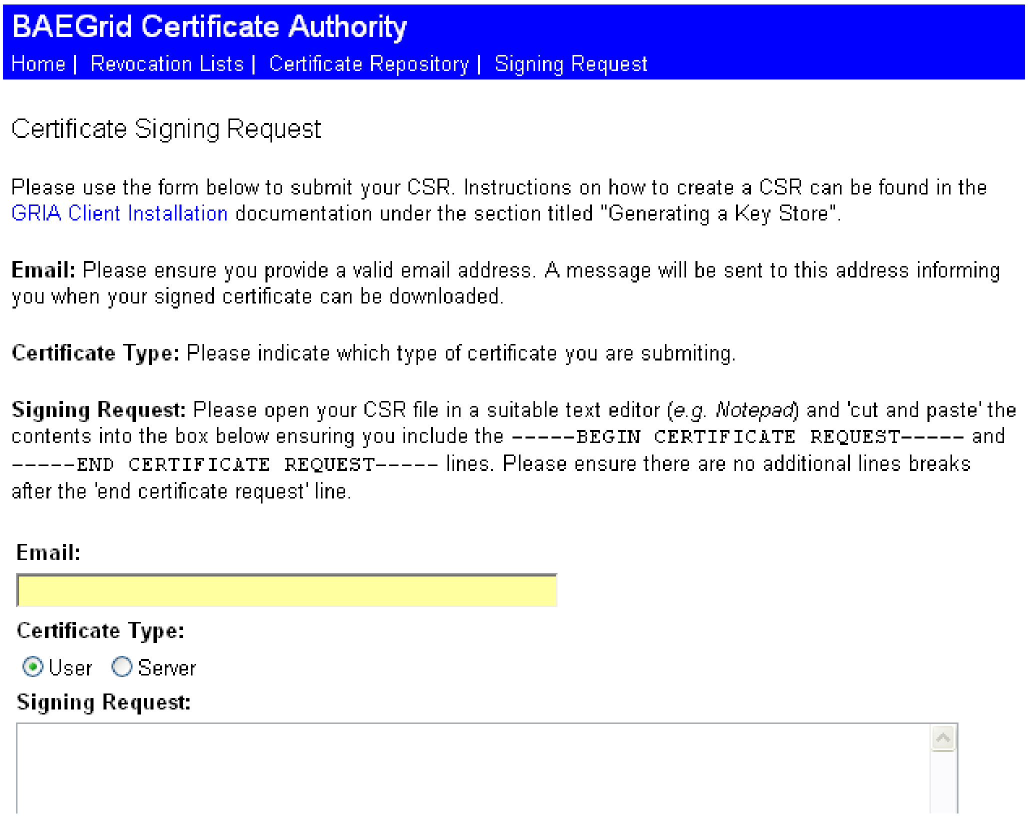


Figure 3 Screenshot of new Aerospace Certificate Authority

Currently only one CA is deployed for all of the partners but in the interest of making the scenario more realistic the aerospace partners will look at deploying a CA per organisation. This will reflect reality were each organisation forming part of the VO will maintain their own identity management schemes.

4.1 GRIA 5.1 VO

The VO model in the SIMDAT aerospace scenario is that of a prime contractor subcontracting aspects of a design out to sub contractors. In the aerospace case BAE is acting as the prime contractor and subcontracting the University of Southampton to perform the wing optimisation. The university does this by using the resources assembled by BAE for the specific VO. The collaboration is an example of the Business Cooperative model, for more information see SIMDAT deliverables D4.2.2 & D4.2.3 [4] [5].

In the last phase of the project the VO was constructed using the private account service provided by GRIA 5.01, with the release of GRIA 5.1 the new Membership service and Client management registry service became available. This gives the aerospace sector a more flexible way of forming the VO. The new architecture involves each partner in the VO managing their own Membership services with the prime contractor also running a SIMDAT VO membership service that aggregates the membership services of the partners. This means that the individual partners can manage users that should have access to the VO and the prime contractor does not need to manage individual users. The prime contractor also runs a registry service that holds resources that are available to the VO. These resources can be any GRIA resource – SLAs, job services, data Services, individual jobs and data stagers. The prime contractor then grants members of the SIMDAT group access to the registry for them to discover and user resources, and for some members of the partner organisations to add new resources to the registry.

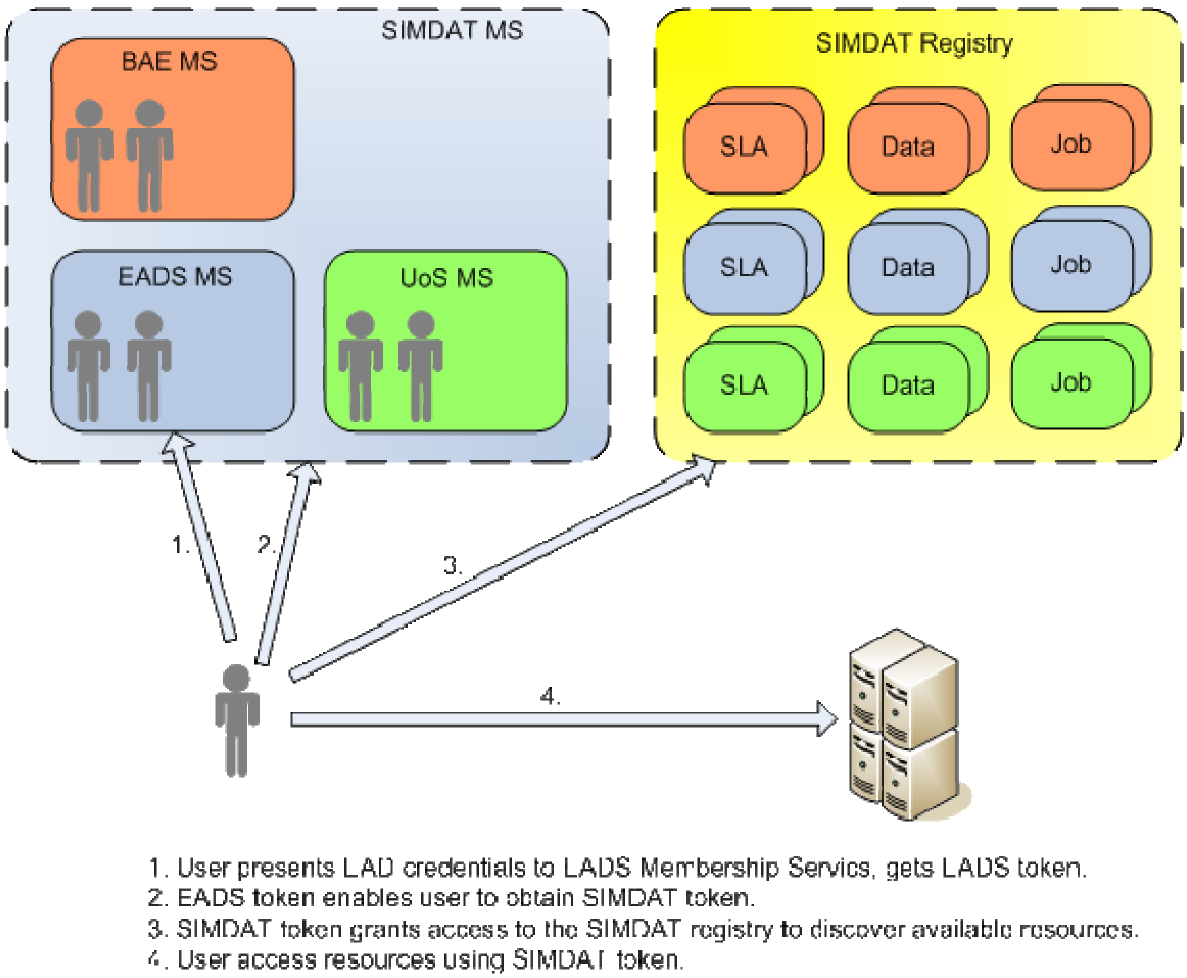


Figure 4 Operation of Aero VO

Figure 4 shows how the new VO operates for an EADS user wanting to discover and access resources within the aerospace VO.

5 Knowledge enhanced workflows

The aim of the knowledge phase of the project is to enhance the prototype to make better use of the knowledge that exists within the system. As part of this various ways of enhancing our workflows will be looked at. The enhancement will be approached in two ways, the extended use of surrogate data models and the management of data produced at the analysis service level.

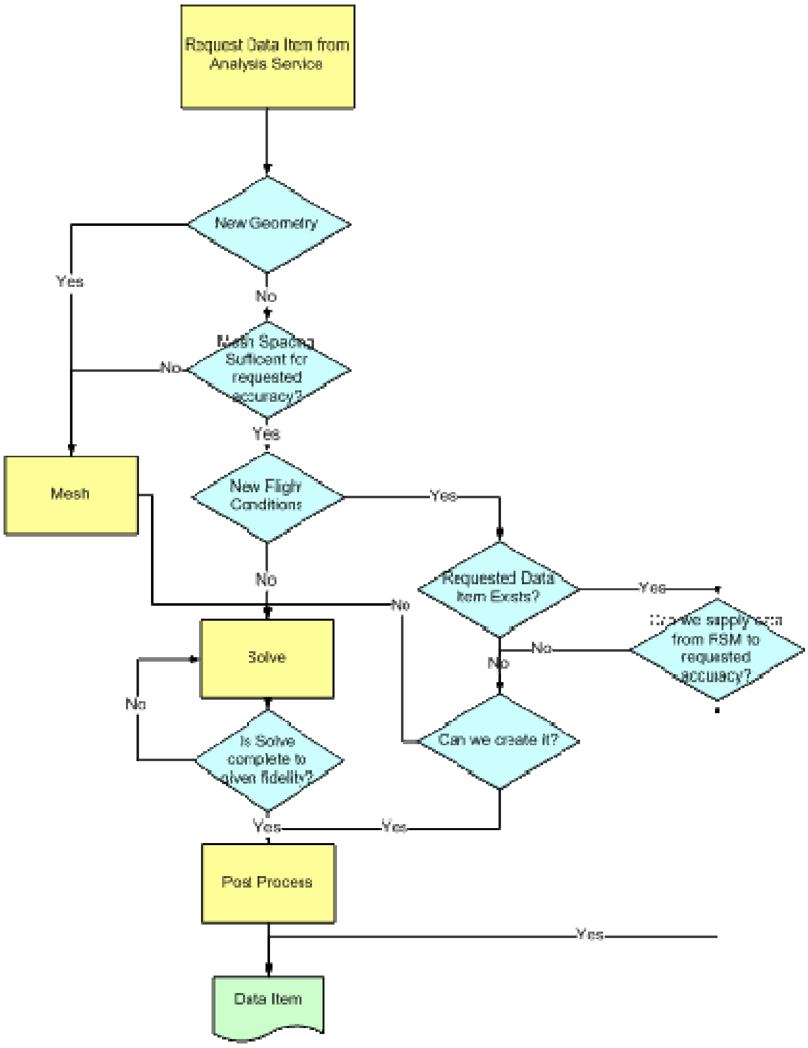


Figure 5 Enhanced workflow example

Figure 5 shows a conceptual workflow and how the enhanced workflow components would sit around the standard CAE process of mesh – solve - post process. The aim of these enhancements is to reduce to total time that the simulation. This is the time that the total optimisation takes not necessary the time taken for an individual simulation run. To do this knowledge gathered from previous runs will be stored and managed in a data manager layer. Then when a request is made of the analysis service the knowledge stored in the data manager can be used to see if various stages of the analysis are needed or whether data can be reused.

5.1 Data Manager

The data manager will be responsible for managing simulation data at the analysis services level. The reasoning behind the data manager is that in some cases data is thrown away between analysis runs that could be reused to speed up subsequent runs. For example in the interoperability phase prototype a CFD mesh would be produced for a particular geometry and then be used to get solution for one particular flight condition, and then discarded. If another request for a solution on the geometry, but at different flight conditions, was made the mesh would be regenerated. With the data manager the mesh could be stored and then reused where applicable. This is simple enough for a single discipline at a single site but the scenario is distributed and so the data management layer will have to work across the aerospace grid.

The top level requirements for the data management layer are:

- Capable of storing files, results and associated metadata
- Capable of retrieving files (or file references), results and associated metadata
- Queryable, verify existence of data items and results.
- Can be tailored for each analysis discipline.
- Accessible to members of the VO, with controlled access
- Distributed, each analysis site able to contribute.
- Flexible ways to access files, Local files, GRIA data stagers, IGOR-FS
- Management of bulk data lifetimes, automatic clearing of data based on data and ease of regeneration.

Currently data within the analysis survives is either stored as local files on a file system or as GRIA data stagers. Sharing of this data between the partners is done by passing references to GRIA data stagers between the workflows at the end of execution. The data manager must be able to deal with GRIA data stager handles as well as references to local files. The aerospace partners are also interested in trying out the IGOR-FS distributed P2P file system and so the data manager must be able to be extended for this.

The meta data associated with each of the bulk files managed by the data manager will be used to give the files some context (i.e. which geometry is this mesh created from, what type of mesh is it etc) but also so allow the data manager to perform some cleanup operations. Each partner has a finite storage capacity and so the data manager must make decisions on what to keep and what to remove. The policy adopted is that storage limits will be built into the system, the data manager will then store everything it can until the volume of data approaches that limit. It must then decide which files to delete to free space. To do this the oldest files could be deleted but the age of the file does not necessarily give an indication of its usefulness. A ranking must be assigned to each file stored and that used to manage the lifetimes.

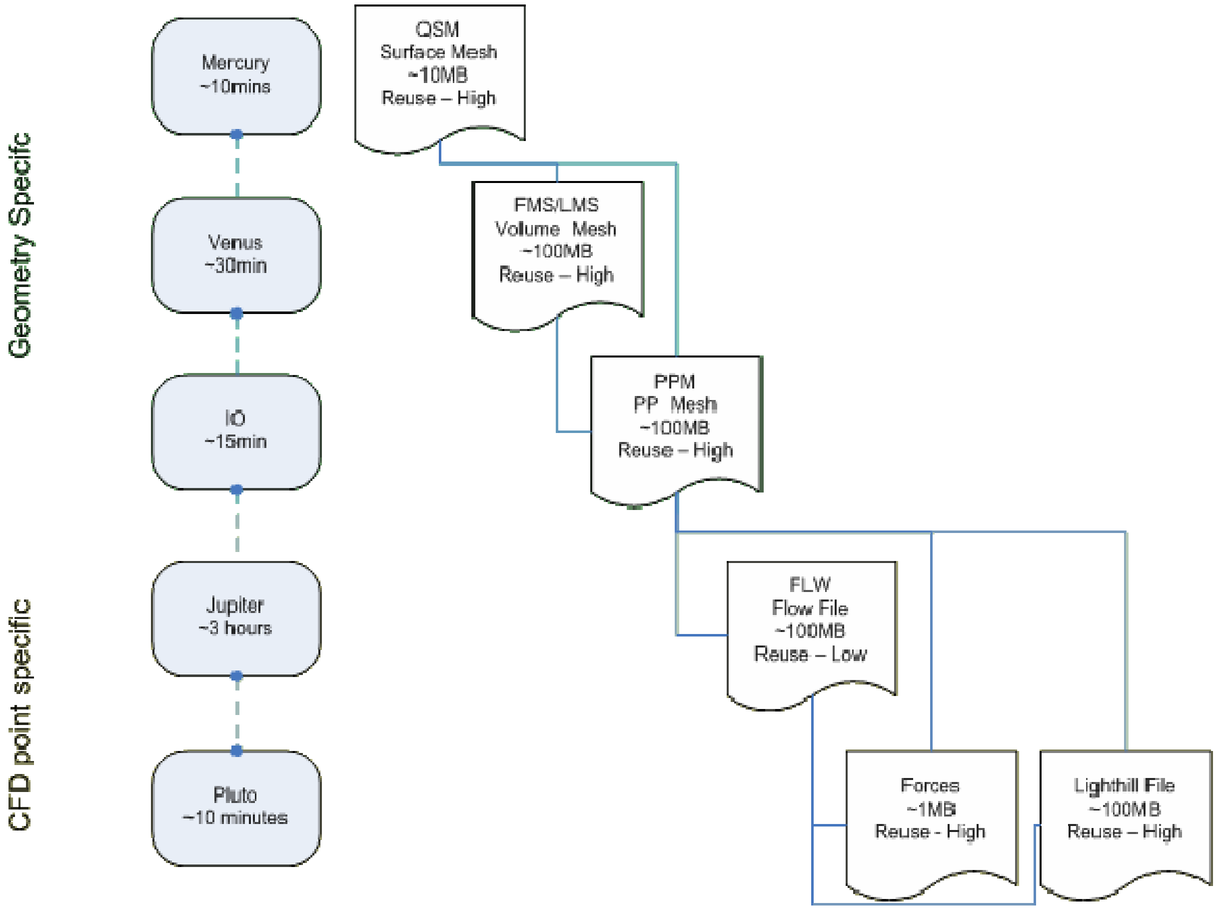


Figure 6 CFD Generated Data

Figure 6 shows the BAE CFD process with the associated generated data files, the approximate time they take to generate, their approximate size and the dependencies between the data items. A link between the data items indicates that to produce the lower item you need the linked item. For example to create the PPM file, both the FMS/LMS and the QSM files are needed. The reuse rankings give an indication of the reuse potential of the file, for example the PPM file may be used to run simulations at multiple flight conditions. A combination of the reuse factor and the time taken to generate can be used to decide what data to remove first. For example if there was a choice between the QSM, FMS/LMS and PPM then the QSM and FMS/LMS files may be removed before the PPM as it takes the longest time to generate or the PPM file might be deleted knowing that it can be regenerated in a reasonable amount of time from the FMS/LMS and QSM. The FLW file takes the longest to generate but is a very large file, once it's been post processed it might not be used again so is a candidate for removal. It is these types of decisions that the data manager must support.

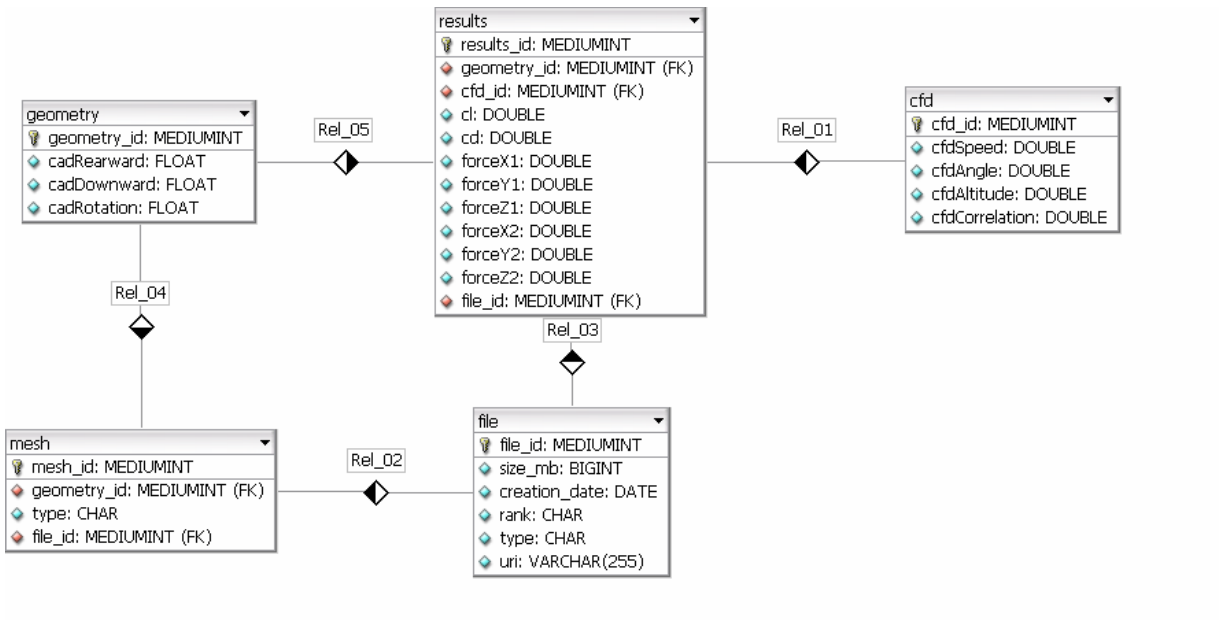


Figure 7 Initial Data Manager Schema

Figure 7 shows an initial schema for the data manager. It is the CFD specific schema but the geometry, mesh and file tables are common amongst the disciplines. The schema shows how the objects are related. For example a single *geometry* will have multiple *mesh* objects associated with it (i.e. Surface, Volume etc) and a result references a specific geometry and specific flight conditions (the *cfd* object). The *results* object is specific to the CFD simulation, there is an analogous results object for the other analysis disciplines. The *file* object represents a reference to a physical file, and includes its size, creation date and unique reference *uri*. The *rank* entry is the priority that is assigned to the file when it is committed to the data manager, the lower rank files will be removed from the data manager sooner that the higher ranked files. The *type* entry allows the object to describe different types of files; this might be Local, GRIA or IGOR-FS. The combination of the *type* and the *uri* should give the user enough information to then access the physical file (assuming they have the correct access permissions).

An example use case for the schema is the committing of a new mesh to the data manager. The user generates the mesh based on a specific geometry, when the mesh is generated the user presents the data manager with a reference to the local file location of the mesh and a reference to the associated geometry. The data manager then copies the local file to a secure, managed partition and creates a new *file* object based on that location. A new *mesh* object is also then created referencing the *file* object just created and the supplied *geometry* reference. The *mesh* is now available to other users of the data manager, they can now query on a geometry and find all of the associated meshes.

In the case of the aerospace prototypes the analysis workflows will need to interact with the data manager, a Java API and a series of Taverna local workers will be developed. The API will also allow the data manager interactions to be scripted using the Groovy language.

5.2 Surrogate Data Models

The response surface modelling modules (RSM creation and RSM evaluation) that are exposed as GRIA services introduce knowledge into the workflow at two levels: at one level it uses knowledge derived from existing data in order to construct a surrogate model of some computer simulation service and at another level it uses error predictions associated with the response surface model in order to advise the user on whether to accept the user on whether to invoke a full analysis or accept the surrogate modelling analysis.

The RSM creation service simply takes some information on a simulation that is available in the data management layer (a set of results of running the simulation) and uses this data to construct a surrogate model or metamodel (a model of a model) that is cheap to run and can be used in lieu of the original simulation model where appropriate. The RSM creation service also outputs a metric which gives a global insight into the accuracy of the RSM fit to the original data.

The RSM evaluation service uses the response surface model created using the RSM creation service in order to predict aerodynamic performance. This service is much quicker to run than the aerodynamics service that uses a CFD model. The service outputs both a prediction of the aerodynamic performance and a local measure of its accuracy.

The global insight into the accuracy provided by the RSM creation service and the local measure of accuracy provided by the RSM evaluation service are used to decide whether the RSM is trusted to give a similar result to the aerodynamics service it approximates. This insight then provides a user with sufficient information on whether to accept the RSM prediction or resort to full aerodynamic analysis.

The RSM creation and evaluation services are implemented as GRIA services and Taverna local workers and embedded into the overall aerodynamics workflow. This use of knowledge extraction intelligently allows users of the aerodynamics workflow to obtain results more economically as users can make use of response surface approximations when appropriate.

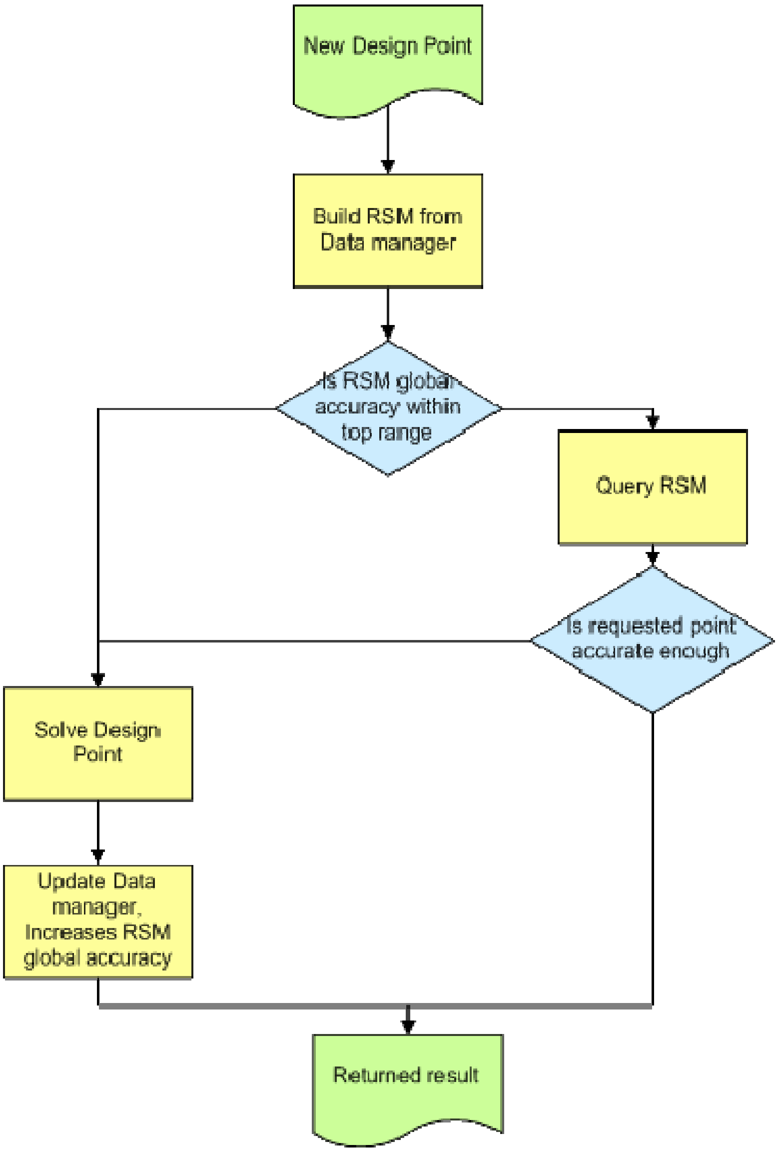


Figure 8 Use of RSMs in analysis services

As more simulation data becomes available in the data management layer the accuracy of the data models increases and so more responses may be given based on the RSM's than by simulation. The embedding of the services within the aerospace workflows means that this decision is made automatically based on the requested accuracy and the system dynamically chooses RSM or simulation responses.

6 Testing and Monitoring Framework

The aerospace partners have set-up a grid infrastructure in order to perform their experiment. At the time of writing, this infrastructure spans 3 countries with 4 sites being deployed. By the end of the project, 4 countries and 5 sites are foreseen. Obviously, such an infrastructure is difficult to operate as each site is independent and have its own schedule for maintenance, deployment for new releases and new services, operations. Moreover, it would be very difficult to have a per service monitoring as it would request special management rights. In order to have a minimum monitoring system, the partners of the aerospace scenario have decided to add beside their managed services a set of free services that can be used for testing and monitoring purpose.

6.1 Basic requirement for the monitoring framework

In order to have an overview of the global infrastructure, one needs at least to monitor the grid connectivity and have a status of the computing resources sitting on each site. In order to achieve this, the partners commonly agreed to deploy the 2 GRIA basic services: DataService and JobService. In the JobService, a basic image processing application has been deployed. This image processing application is based on ImageMagick [6] which is available for most of the operating systems (Windows, Unix & MacOS X). This application is very easy to install and non intrusive since it requests very few CPU cycles per image transformation. Therefore, it is a good candidate for monitoring purpose. Obviously, monitoring the connectivity between sites and an image processing application does not provide a full overview of the status of the applications deployed on a site, but it is very likely that if those very basic services does not operate properly, the business one useful for our overall workflow will not operate at all.

In conclusion, each site willing to participate to the monitoring framework should publish the following services:

- Data Service
- Paint application hosted by the Job Service

All these services have to be made available for free (or be accessible to the SIMDAT group).

6.2 Monitoring framework architecture

The basic bricks of the monitoring infrastructure being presented, let us focus on the overall architecture. The monitoring application is obviously distributed since some services should be available on each site. Nevertheless, thanks to the GRIA API, the service will be run from a single site and the data will be store on that particular site.

Figure 9 presents the way the application is distributed over the partners.

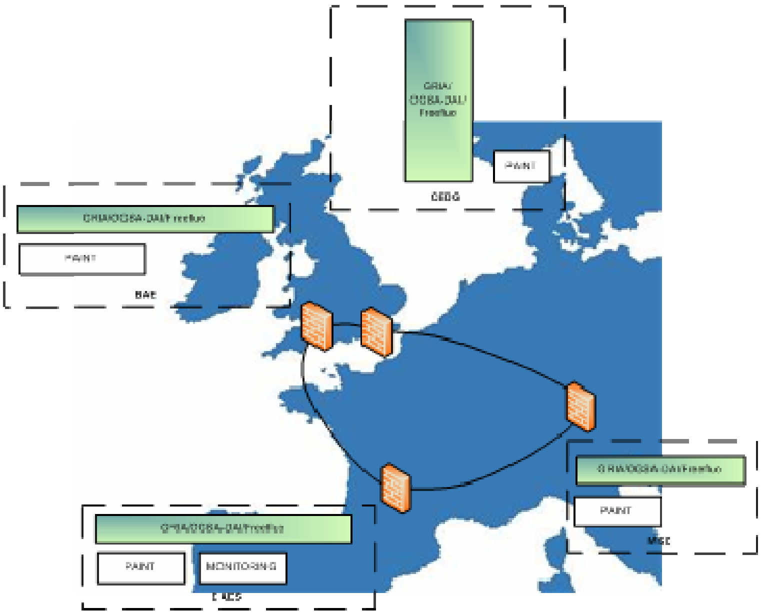
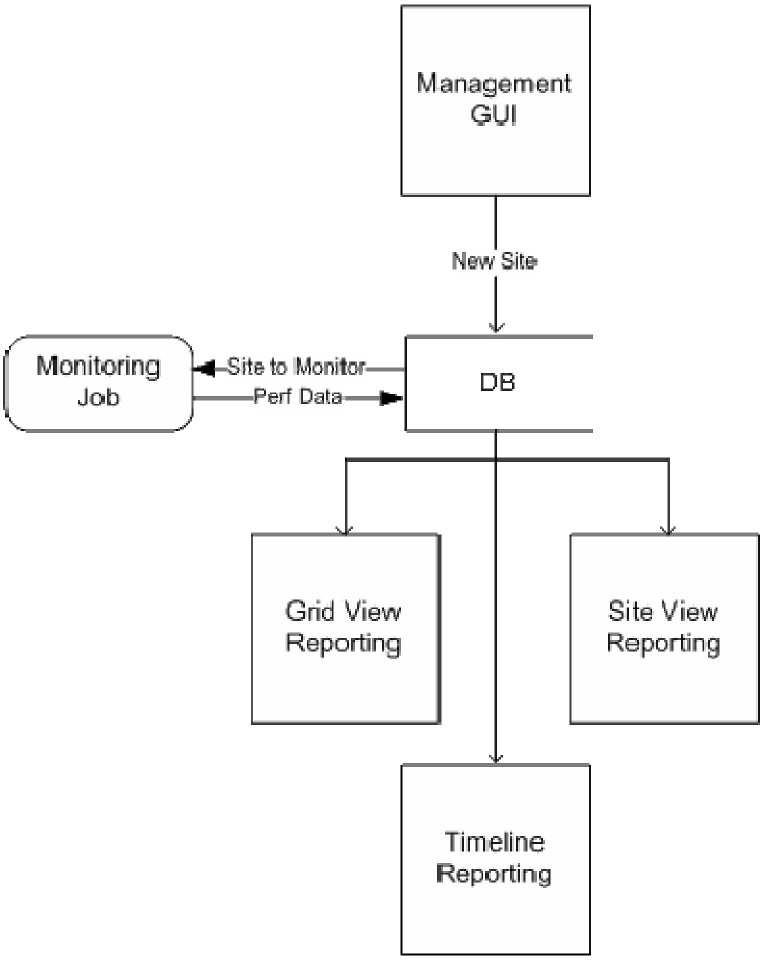


Figure 9: Distribution of the monitoring application

The internal architecture of the monitoring framework is the following:



From this sketch, it is made clear that central to the monitoring infrastructure is a database hosting both the sites and links to be monitored as well as the performance data. These data are obtained from a service scheduled at specified time. A set of graphical user interface pages are used to interact with the application. One is used to enter new sites to be monitored. A site is being defined by its geographical position and its URL. Three other pages are reporting the status of the infrastructure:

- One is providing a timeline featuring all the measures available,
- An other one is providing a map with the performance of the different links between the sites
- The last one is providing the performance on a per site basis.

7 Optimisation Workflow advisor

In the aerospace scenario, workflows are constructed in various PSEs using services provided by partners to perform design optimization tasks. Various optimization methods and strategies have been used in these workflows. These workflows often embed extensive expert knowledge regarding the usage patterns of methods and strategies, which can be useful for building new workflows. One of the forms of these workflows is MATLAB scripts which are of particular interests for Southampton researchers in the knowledge phase of the project. Case-based reasoning and rule-based reasoning are the two technology areas that have been studied in support this work. The architecture adopted is illustrated in figure 10.

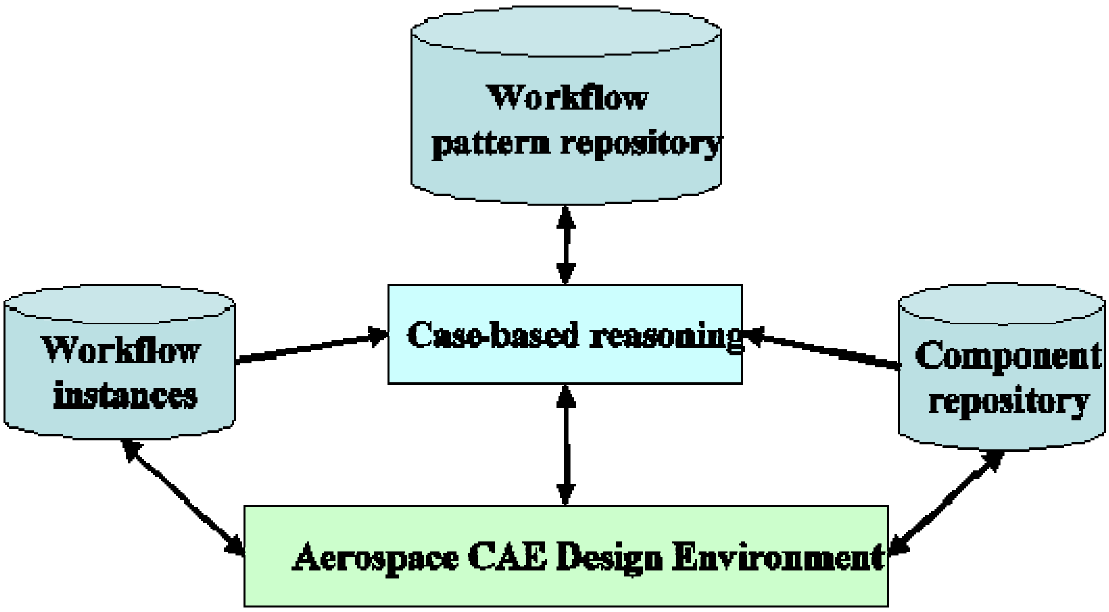


Figure 10 Optimisation workflow advisor architecture using case-based reasoning

Existing MATLAB scripts stored as files on the file system are first parsed to extract use patterns of interested components. These patterns are represented as graphs and stored in XML files, along with references to the original scripts. One distinct feature of this approach is that no manual annotation from users is required, therefore reducing the burden of annotation and possible mismatches.

MATLAB scripts are first parsed, and converted into graphs representing the workflow patterns which describe what optimization components are used, and in what order. Both loop and conditional structures can be identified by the parser. The extracted workflow patterns are then stored in an archive, currently in the format of a XML file. Given a new or not-yet archived workflow, user will be able to find similar workflows: similarities between workflows can be defined in a number of ways:

Two workflows shared a common components;

All the components in one workflow can be found in another workflow;

Same set of components can be found in two workflows in the same order;

The first components of two workflows are the same;

- The last components of two workflows are the same;

These similarity criteria can be used to assist users to search for interested workflows when it comes to building new workflows.

Heuristics are often used in deciding optimization methods to use and relevant control parameter settings. For example, evolutionary computation methods have proved to be effective in finding global or near global optimal solutions; therefore they are always the preferred methods to adopt. However, they tend to converge much slower than gradient based methods, thus hybrid approaches are suggested. It is in this context that rule-based reasoning approach is investigated.

Ontology support to the reasoning process will further extend the scope of the reasoning process, allowing users to find ontologically similar workflows, as well as exactly similar workflows. Since optimization workflows are of main interests in aerospace scenarios, a simple optimization ontology is built to support the classification of optimization methods and similarity match in workflow search. This ontology support, combined with analysis of workflow patterns, provides users with the ability to locate workflows using the same type of methods.

When it comes to building new workflows, it is common practice to start from existing similar ones and to make modifications according to the current application scenario. Rule-based reasoning approach is used to support this way of working. Complete new workflows can still be constructed, and added to the archive by expert users to further enrich the knowledge base.

8 Status and Plans

Figure 11 shows the aerospace plans and milestones up to PM45.

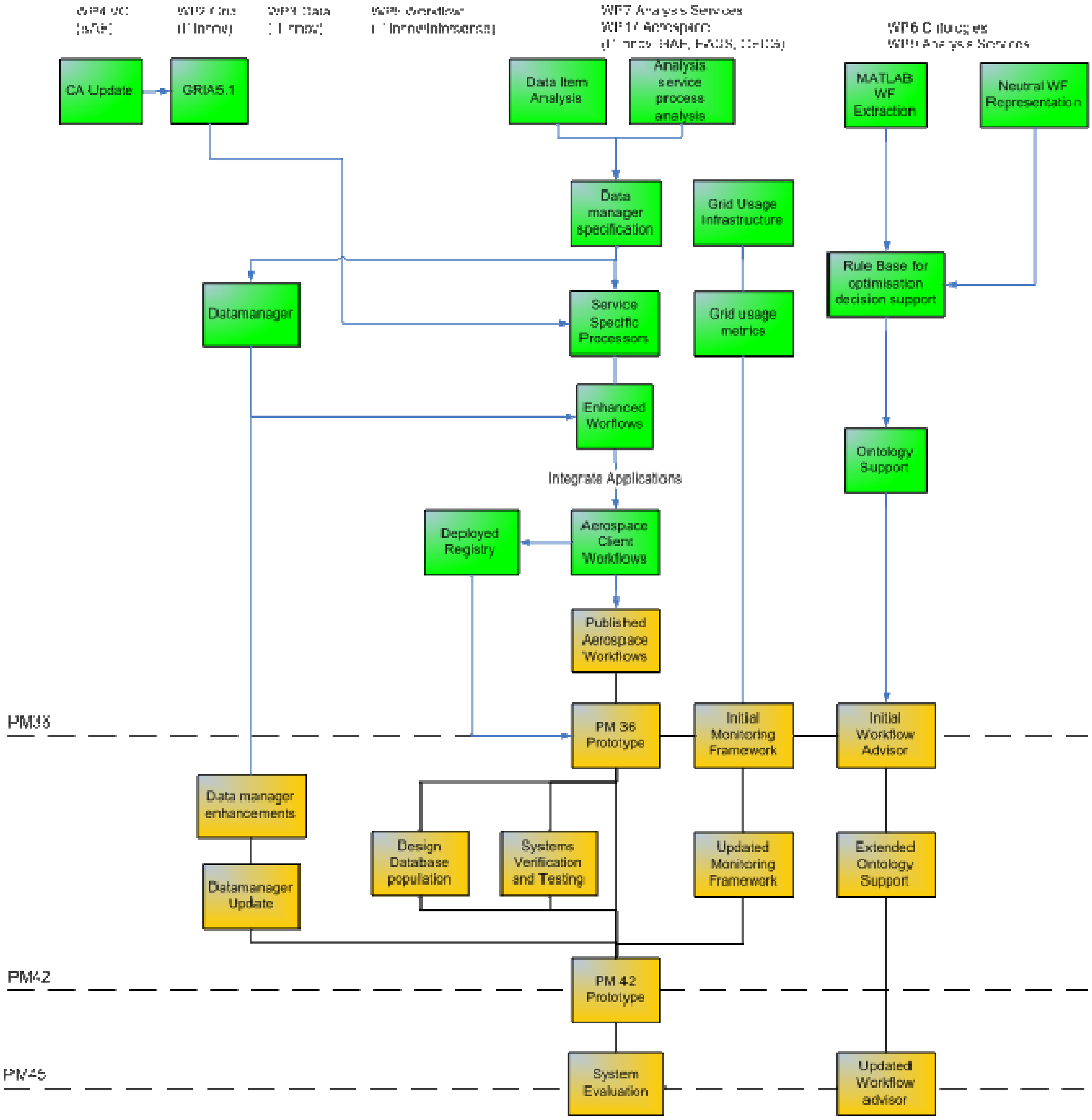


Figure 11 Aerospace Plans to PM45

As the figure shows by PM36 the initial version of the aerospace prototype will be in place, utilising the first version of the data manager and workflow enhancements. The first version of the testing framework will be in place and an initial workflow advisor prototype will have been developed. Between PM36 and 42 the data manager will be refined and the services updated. The PM36 prototype will be populated with data to allow the system to be tested. The second version of the monitoring framework will be deployed by all the partners. The workflow advisor tasks will continue to be developed through to PM45 with the integration of a workflow repository and extended use of the optimisation ontology for reasoning. Between PM42 and 45 the partners will also evaluate the PM42 prototype and continue populating the system with data.

9 PM36 Requirements status

Below are the outstanding requirements from the last phase of the project and their current status, the requirements listed are those not fully fulfilled or that have changed since PM30; for the complete list please see SIMDAT deliverable D11.2.3.

9.1 Grid Infrastructure

Requirement	Description	Priority	Implemented	Comment
Authentication service	Ability to authenticate users (X509 based) Authenticated transactions Ability to handle federated identity for exploitation phase	High	Partial	PM42 release of GRIA will include updates for identity scheme integration
Authorisation Service	Policy driven access control to resources Dynamic policy management for exploitation phase	High	Partial	NEC's DAC technology addresses dynamic authorisation.
Single interface to compute service	Ability to access compute services with different scheduling requirements through single interface Reservation of compute resources Sandboxing runtime in compute service including ability to specify sandbox environment	Medium	Partial	PM42 GRIA release will support multiple schedulers.
Resource discovery	Ability to discover alternate services due to service failure or unavailability	Low	Complete	Registry released as part of GRIA 5.01

9.2 Administration of Virtual Organisation

Some of the requirements for the administration of the Aerospace Activity virtual organisation are also covered by in the Grid Infrastructure requirements section

Requirement	Description	Priority	Implemented	Comment
Authentication service	Ability to authenticate users Authenticated transactions Ability to communicate securely Ability to handle federated identity for exploitation phase	High	Partial	PM42 release of GRIA will include updates for identity scheme integration
Authorisation Service	Policy driven access control to resources Dynamic policy management for exploitation phase	High	Partial	NEC's DAC technology addresses dynamic authorisation.
Resource discovery	Ability to discover alternate services due to service failure or unavailability	Low	Complete	Registry released as part of GRIA 5.01
Auditing	Execution and resource audit for QA purposes	High	Partial	Provided but not validated against Organisational policy
Accounting	Ability to charge for service provisioning	Medium/Low	Partial	Not linked to payment systems

9.3 Workflow

Requirement	Description	Priority	Implemented	Comment
Workflow execution optimisation	Control of workflow execution with constraints of service availability	Medium/Low	×	
Workflow mining	Ability to search workflow database using meta-data	Low	Partial	See section 7
Workflow language interoperability	Ability to exploit workflows in other standards based workflow enactors	Medium	Complete	Runtime interoperability between several workflow systems achieved.

Abstract workflow construction	Ability to construct workflows using abstract service descriptions	Medium	Partial	See section 7
Abstract workflow execution	Ability to execute abstract workflows with runtime binding to services via a semantic service registry	Medium	×	
Workflow composition advisor	Ability for the workflow authoring tool to advise users on next steps based on prior knowledge in workflow repository	Medium	Partial	See section 7

9.4 Analysis Services

Requirement	Description	Priority	Implemented	Comment
Auditing	Execution and resource audit for QA purposes	High	Partial	Log data provided but this has not been verified against organisational policies
Accounting	Ability to charge for service provisioning	Medium /Low	Partial	Provided by GRIA but not linked to payment system
Results lifetime	Ability to specify lifetime of results	Medium	Partial	Limited capability provided by GRIA, being extended with PM42 Data manager
Semantic description of analysis service	Ability to semantically describe analysis service (inputs/outputs/behaviour) to facilitate deployment and to enable the exploitation of semantic service registries	High	Partial	Registry deployed as part of GRIA 5.01 but no semantic information attached.

9.5 Ontologies

Requirement	Description	Priority	Implemented	Comment
Ontology driven data access	Ability of knowledge service to access ontologically described data	High	×	

9.6 Knowledge Services

The integration of the knowledge services in the cost modelling scenario was not achieved due to staffing issues but an architecture was defined. For more information see SIMDAT deliverable D8.2.3.

Requirement	Description	Priority	Implemented	Comment
Cost Prediction	Ability to estimate future costs based on historic cost data.	High	NA	Requirement Depreciated
Multiple source data integration	Ability to search across a variety of data sources i.e. OGSA-DAI, flat file, excel spreadsheets.	Medium	NA	Requirement Depreciated
Distributed knowledge aggregation	Ability for a member of the VO to access aggregated knowledge about the whole problem using ontology described data to generate aggregated knowledge that is not otherwise visible to a single team, organisation or specialism	Medium/Low	Partial	Ongoing with PM2 Data Manager

9.7 Summary

The below table shows a summary of all of the PM 30 requirements and shows that 81% of them are either completely or partial fulfilled.

	Yes	Partial	No
Grid Infrastructure	5	3	
Virtual Organisations	2	4	
Workflow	7	0	5
Analysis Services	5	3	1

Ontologies	3		1
Knowledge Services	2		1
Total	24	10	8
	57%	24%	19%

10 PM45 Requirements status

Name	**Metadata for services and workflows**		
Req. Id	AERO-001		
Application Activity	Aerospace		
Prototype(s)	Aerospace prototype for PM42		
Date Created	2007-03-27	Priority	High
Created By	Southampton CEDG	Technology component	Workflow
Status	Ongoing		
First Implementation Date	End of August 2007	SIMDAT module targeted	WP 11.3
Description			
Relation to prototype			
• Aero PM42			
Requested functionality			
• Adding metadata for published services and workflows • ontology support for metadata • API access to the metadata from workflow authoring environments			
Validation			
• Ability of users of workflow composition tools to add ontology enhanced metadata to services and workflows			
Assumptions			
•			

Name	**Ontology Support for Workflow Composition**
Req. Id	AERO-002
Application Activity	Aerospace

Prototype(s)	Aerospace prototype for PM42		
Date Created	2007-03-27	Priority	High
Created By	Southampton CEDG	Technology component	Workflow, Ontology
Status	Complete		
First Implementation Date	End of August 2007	SIMDAT module targeted	WP 11.3
Description			
Relation to prototype			
• Aero PM42			
Requested functionality			
• Definition of workflow ontology • API access to the ontology from workflow authoring environments			
Validation			
• Availability of ontology to workflow authoring environments			
Assumptions			
•			

Name	**Workflow Repository**		
Req. Id	AERO-003		
Application Activity	Aerospace		
Prototype(s)	Aerospace prototype for PM42		
Date Created	2007-03-27	Priority	High
Created By	Southampton CEDG	Technology component	Workflow, Grid Infrastructure
Status	Ongoing		
First Implementation Date	End of August 2007	SIMDAT module targeted	WP 11.3
Description			

Relation to prototype			
• Aero PM42			
Requested functionality			
• Workflow repository conforming to standards • Archival of workflows into repository from workflow authoring environments • Retrieval of workflows from repository from workflow authoring environments • API accesses to above functionalities			
Validation			
• Ability of users of workflow composition tools to query, search workflow repository, in addition to the ability to add new workflows into repository from workflow authoring environments			
Assumptions			
•			

Name	**Product Knowledge Workflow Components**		
Req. Id	AERO-004		
Application Activity	Aerospace		
Prototype(s)	Aerospace prototype for PM42		
Date Created	2007-03-27	Priority	High
Created By	BAES	Technology component	Grid Infrastructure, Workflow.
Status	Complete		
First Implementation Date	End of August 2007	SIMDAT module targeted	WP 11.3
Description			
Ability to add local components into workflows to make decisions based on product knowledge.			
Relation to prototype			
• Aero PM42			
Requested functionality			

• Each partner able to author they're own custom workflow decision nodes. • These nodes must be deployable in published workflows.			
Validation			
• Ability of users to create nodes and demonstrate path through workflows being affected by product knowledge decisions.			
Assumptions			
•			

Name	**Product Knowledge Workflow Monitoring Components**		
Req. Id	AERO-005		
Application Activity	Aerospace		
Prototype(s)	Aerospace prototype for PM42		
Date Created	2007-03-27	Priority	High
Created By	BAES	Technology component	Grid Infrastructure, Workflow.
Status	Complete		
First Implementation Date	End of August 2007	SIMDAT module targeted	WP 11.3
Description			
Ability to add local components into workflows to make decisions based on the output of currently running jobs and have the ability to stop that job based on product knowledge.			
Relation to prototype			
• Aero PM42			
Requested functionality			
• Each partner able to author they're own custom workflow decision nodes. • These nodes must be deployable in published workflows. • A node that is able to monitor the output of a GRIA job and change the state of that job before it finishes/fails.			
Validation			
• Ability of users to create nodes and demonstrate a GRIA job being stopped or restarted with different inputs based on a decision made in a node monitoring its output.			

Assumptions			
•			

Name	**Data Lifetime Management**		
Req. Id	AERO-005		
Application Activity	Aerospace		
Prototype(s)	Aerospace prototype for PM42		
Date Created	2007-03-27	Priority	High
Created By	BAES	Technology component	Grid Infrastructure, DDRA.
Status	Complete		
First Implementation Date	End of August 2007	SIMDAT module targeted	WP 11.3
Description			
Ability to associate data life times to data items.			
Relation to prototype			
• Aero PM42			
Requested functionality			
• Each item of data within the system to have an associated lifetime, once this lifetime is reached the data is automatically removed.			
Validation			
• Removal of data beyond its lifetime.			
Assumptions			

Name	**Data Management**
Req. Id	AERO-006
Application Activity	Aerospace

Prototype(s)	Aerospace prototype for PM42		
Date Created	2007-03-27	Priority	High
Created By	BAES	Technology component	Grid Infrastructure, DDRA.
Status	Complete		
First Implementation Date	End of August 2007	SIMDAT module targeted	WP 11.3
Description			
Ability to store file references and associated metadata in a data management system.			
Relation to prototype			
• Aero PM42			
Requested functionality			
• Each analysis service is able to store required data items and related metadata and query that metadata to retrieve the data items at a later date.			
Validation			
• Storage and retrieval of data at the analysis service level.			
Assumptions			

Name	**Distributed file access**		
Req. Id	AERO-007		
Application Activity	Aerospace		
Prototype(s)	Aerospace prototype for PM42		
Date Created	2007-03-27	Priority	High
Created By	BAES	Technology component	DDRA
Status	IGOR-FS Delayed		
First Implementation Date	End of August 2007	SIMDAT module targeted	WP 11.3

Description			
Ability to share files between analysis services.			
Relation to prototype			
• Aero PM42			
Requested functionality			
• Ability to use IGORFS with GRIA to share analysis data between analysis service nodes.			
Validation			
• Access of Lighthill sources from BAE by EADS over IGORFS			
Assumptions			

Name	**Surrogate Data Models**		
Req. Id	AERO-008		
Application Activity	Aerospace		
Prototype(s)	Aerospace prototype for PM42		
Date Created	2007-03-27	Priority	High
Created By	BAES	Technology component	Knowledge
Status	Complete		
First Implementation Date	End of August 2007	SIMDAT module targeted	WP 11.3
Description			
Extended use of surrogate data models in analysis services.			
Relation to prototype			
• Aero PM42			
Requested functionality			

• Each analysis service to be able to use RSM's to give analysis solutions based on previous data.			
Validation			
• Demonstrated extended use of RSMs in scenario			
Assumptions			
•			

Name	**Authentication Service**		
Req. Id	AERO-009		
Application Activity	Aerospace		
Prototype(s)	Aerospace prototype for PM42		
Date Created	2007-03-27	Priority	Medium
Created By	BAES, EADS	Technology component	Grid Infrastructure, VO
Status	Existing Requirement		
First Implementation Date	End of August 2007	SIMDAT module targeted	WP 11.3
Description			
Ability to integrate GRIA authentication services with existing enterprise schemes.			
Relation to prototype			
• Aero PM42			
Requested functionality			
• Ability to authenticate to GRIA using enterprise identity scheme.			
Validation			
• Ability to authenticate to GRIA using enterprise identity scheme (i.e. LDAP or MS Active Directory)			
Assumptions			
•			

Name	**Semantic Registry**		
Req. Id	AERO-010		
Application Activity	Aerospace		
Prototype(s)	Aerospace prototype for PM42		
Date Created	2007-03-27	Priority	High
Created By	BAES, EADS	Technology component	Grid Infrastructure
Status	Complete		
First Implementation Date	End of August 2007	SIMDAT module targeted	WP 11.3
Description			
Semantic Registry for Job Services			
Relation to prototype			
• Aero PM42			
Requested functionality			
Need for semantic registry to locate and describe deployed job services.			
Validation			
• A deployed registry in the aerospace scenario			
Assumptions			

11 Software evaluation

Within the aerospace community the Technology Readiness Levels (TRL) [1] defined by NASA, the US DOD and the UK MOD are a widely recognised way of assigning a metric of maturity to a technology. The TRLs run from TRL 1 – Basic principles identified up to TRL 9 – Deployed and proven technology. By assigning a TRL to the SIMDAT technology used within the aerospace activity we will be able to convey the state of the technologies to aerospace organisations external to the SIMDAT project.

We judge the latest GRIA release (5.1) to be at TRL 5/6. The tables below give the UK MOD definitions for TRL 5 and 6 for Software [2].

TRL	Description	Supporting Evidence
6	Level at which the engineering feasibility of a software technology is demonstrated. This level extends to laboratory prototype implementations on full-scale realistic problems in which the software technology is partially integrated with existing hardware/ software systems.	Results from laboratory testing of a prototype package that is near the desired configuration in terms of performance, including physical, logical, data, and security interfaces. Comparisons between tested environment and operational environment analytically understood. Analysis and test measurements quantifying contribution to system-wide requirements such as throughput, scalability, and reliability. Integration of basic components.
5	Level at which software technology is ready to start integration with existing systems. The prototype implementations conform to target environment/interfaces. Experiments with realistic problems. Simulated interfaces to existing systems. System software architecture established. Algorithms run on a processor(s) with characteristics expected in the operational environment.	System architecture diagram around technology element with critical performance requirements defined. Processor selection analysis, Simulation/ Stimulation (Sim/Stim) Laboratory build-up plan. Software placed under configuration management. COTS in the system software architecture is identified. Integration plan.

We judge that GRIA has been demonstrated on a full-scale and realistic problem within the aerospace prototype and the prototype has fulfilled the desired configuration for the defined aerospace scenario. The software has become easy to install and configure and the modular approach allows the system to be configured in a very flexible way. The software is now also compliant with a number of recognised standards. The differences between the test environment and the operational environment have been compared as part of the WP4.2 work package and the feedback used to guide the next GRIA releases.
Below is the description for TRL 7.

7	Level at which the program feasibility of a software technology is demonstrated. This level extends to operational environment prototype implementations where critical technical risk functionality is available for demonstration and a test in which the software technology is well integrated with operational hardware/software systems.	Critical technological properties are measured against requirements in a simulated operational environment. Full integration.

OGSA-DAI is being used within the aerospace scenario to allow partners access to the distributed data manager and results databases. OGSA-DAI has been integrated with GRIA and so benefits from the security and account mechanisms provided by GRIA. We judge this technology, the combination of GRIA and OGSA-DAI to be at TRL5. We have integrated it into the prototype system but have not yet tested it against all desired configurations.
The Aerospace partners have also been waiting to evaluate the IGOR-FS technology within the scenario but, to date, have been unable to. Based on demonstrations of the existing technology IGOR-FS is judged to be at TRL3, at a R&D stage where its feasibility has been proven in laboratory conditions. To move to TRL4 the system will have to be demonstrated as a stand alone demonstrator with a plan for integration into the larger scenario.

12 Technology Summary

12.1 Grid Infrastructure

The release of GRIA 5.1 has solved a number of issues for the aerospace partners and has been seen as a step forwards in terms of usability and manageability. It has been adopted by all the aerospace partners. The new membership and registry services have been employed in the prototype.

12.2 DDRA

Work on the aerospace data manager continues with support from DDRA. The aerospace partners are also waiting for the release of the IGOR-FS and are eager to evaluate it's use within the prototype.

12.3 VO

The GRIA 5.1 membership and registry services have given the aerospace partners a new way to form the prime contractor VO. The new configuration eases the burden of administration from the prime contractor and allows the more seamless discovery and sharing of resources.

12.4 Workflow

The workflow advisor architecture agreed on between the aerospace and workflow work packages has begun to be implemented. The approaches have shown to be complementary and the work continues to be supported by the workflow work package.

12.5 Ontology

The optimisation ontology is being used to enhance the reasoning used with the aerospace workflow advisor.

12.6 Analysis Services

The aerospace scenario continues to make use of analysis services and uses the reference service provided by MSC.

12.7 Knowledge

Knowledge services have been introduced at the analysis services level in the shape of the surrogate data models. These models use the knowledge generated by previous analysis runs to reduce the need for full simulation in later analysis requests.

13 Conclusion

The architecture for the aerospace knowledge phase prototype has been defined based on GRIA 5.1. It has been shown how surrogate data models offer a way to potentially use existing results to reduce the total number of simulation runs that have to take place whilst still maintaining enough accuracy for the larger optimisation. The use of a data manager has been investigated and requirements have been described and several use cases defined. The first version of the testing and monitoring framework has been implemented and the development of the second version is underway. The workflow advisor architecture has been defined and case and rule based reasoning investigated.

14 References

1. SIMDAT Annex I - “Description of Work”
2. SIMDAT Deliverable D11.2.3 Aerospace prototype for Interoperability with updated validation results
3. OpenCA www.openca.org
4. SIMDAT Deliverable D4.2.2 Collaboration Patterns.
5. SIMDAT Deliverable D.4.2.3: Policy Analysis report relating per Application Sector Industrial policies to Grid Technologies including gap analysis
6. http://www.imagemagick.org/script/index.php
7. TRL http://www.ams.mod.uk/content/docs/techman/content/trl_scpe.htm
8. http://www.ams.mod.uk/content/docs/techman/content/trlann/trlanna.pdf

End of Document