RMAS – Delivering an integrated Research Management & Administration System

Paul Fairley[a], Kevin Ginty[a], Simon Kerridge[b], Simon Foster[c], Stephen Trowell[c], Steve Hustwayte[c], Steve Butcher[d], Josh Brown[e], Roland Major[e]

[a] University of Sunderland, UK; [b] University of Kent, UK;
[c] University of Exeter, UK; [d] HEFCE, UK; [e] JISC, UK

Summary

In 2009, two feasibility studies were carried out to ascertain if there was a need for a 'cradle to grave' research and administration system. Several Higher Education Institutions (HEIs) expressed interest, identifying problems encountered when dealing with the management of research activities. With no supplier offering a comprehensive solution, the pathfinders began to solve the problem through the RMAS project. The Universities of Exeter, Kent and Sunderland, in conjunction with Nexus began the development of a framework which would allow data from the RMAS modules to be passed between systems and institutions via a Central Enterprise Service Bus (ESB).

Keywords: Cloud, Higher Education, integrated system, Interface, Research Management and Administration System [RMAS], CERIF, CRIS, ERA, Communications Bus, EuroCRIS

1 Introduction

During 2009 the Higher Education Funding Council for England (HEFCE) Shared Service Program made funding available for a two-part study to explore the demand and options for a 'cradle to grave' research and administration system. Studies were carried out as a product of a number of universities identifying the need to resolve a generic problem relating to the requirement to manage research activities in a more coherent way through improved use of data across a range of existing IT systems. The studies invited input from all higher education establishments in the United Kingdom.

The studies found that no supplier offered a solution which supported the full functionality required to support the management and administration of a University's research programme, from conception through to research outputs and project closure (Binge 2009a). Research management and administration includes aspects such as project & management and costing & pricing. It was reported that universities were typically using a combination of databases and spreadsheets, working alongside existing cooperate systems such as Finance, HR and Student records to manage these activities.

To address this gap JISC funded the Research Management and Administration System[1] (RMAS) project under the University Modernisation Fund (UMF). This paper presents the scope and design of an integrated modular RMAS framework for UK research organisations which allows

[1] See http://www.exeter.ac.uk/research/rmas/ (accessed on 20th February 2012)

for the management and administration of research projects from concept through to closure. The framework enables best of breed modules to work together to produce an integrated RMAS and to interface with a University's existing corporate systems.

This paper demonstrates how, utilising the overarching RMAS architecture, the three pathfinders institutions, the Universities of Exeter, Kent and Sunderland have:

- Developed and deployed RMAS standards (based around CERIF XML[2])designed to allow the exchange of information between both research systems, and between core corporate systems and research systems;
- Developed adaptors to integrate RMAS modules with existing local research and corporate systems using standard approaches.

This paper will firstly review previous work before defining the characteristics of RMAS, exploring problems that were encountered by the pathfinders, and finally the paper will considering the impact that RMAS can make within the sector.

2 Background

Research into RMAS began by looking into the Electronic Research Administration (ERA) systems [CRIS/RMAS], which showed that there were various complications research systems at that time. The first of which being that there was no 'off the self' full ERA system within the UK Market. thus meaning that different modules must be purchased from different suppliers which can cause interoperability issues. This meant that in order to manage the full research project process solutions had to be procured from a number of suppliers. This in turn often caused interoperability issues.

Two RMAS feasibility studies were carried out in 2009. The first in May 2009, (Binge, A. 2009a) which revealed that there was a strong demand for 'cradle to grave' RMAS, in which 19 HEIs contributed. The second study (Binge, A. 2009b) showed that drastic savings could be made by the development of an RMAS, starting with the estimated 10% to 20% savings that can be made in staff efficiency, which equates to an average saving of £375,000 per annum (based on an institution the size of Exeter). Thus meaning that if all 29 of the HEIs that had expressed initial interest incorporated the RMAS structure there could be a saving of more than £25M over ten years.

3 Architecture

The RMAS platform has the overall purpose of connecting different systems for managing research. In order to plan the architecture of the RMAS, the pathfinders initially looked into the different modules that would be used within a HEI. From this investigation, the pathfinders came to an agreement on which modules they would use to send data between each other.

In order to focus development of this concept, the three pathfinders selected a subset of potential modules from the proposed RMAS framework. The Costing & Pricing Management module was selected as two already had such a module from the same supplier and the third was preparing to procure such a module. This selection provided a focus for the initial work on the RMAS messaging.

[2] See euroCRIS CERIFXML www.euroCRIS.org.uk (accessed on 26th February 2012)

The pathfinders then examined how the RMAS installations could be integrated in RMAS; this connecting of systems happens on two levels. At the institutional level, a communications bus facilitates data connectivity and/or message exchange between applications within the institution. Figure 1 shows a simplified example of how messages are routed within the local RMAS systems. Data connectivity via database integration was also demonstrated and documented.

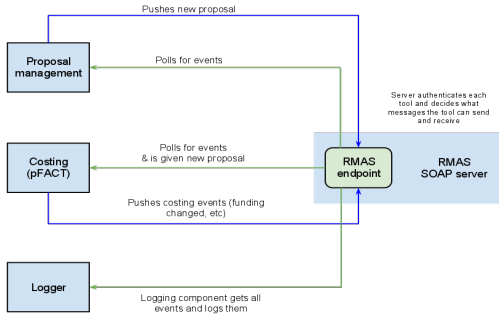

Figure 1: Local Architecture

The overall purpose is to improve efficiency, reducing costs, and facilitating communication.

At a national level a central ESB, developed by JISC Advance as a workstream complimentary to RMAS, will support data exchange and messaging between institutions and/or between externally hosted RMAS modules and institutions. Figure 2 shows how messages from institutes can be sent out to each of the other pathfinders, and how the national cloud infrastructure could host research system solutions in the long-term. Message exchange between pathfinders via the central ESB has been demonstrated

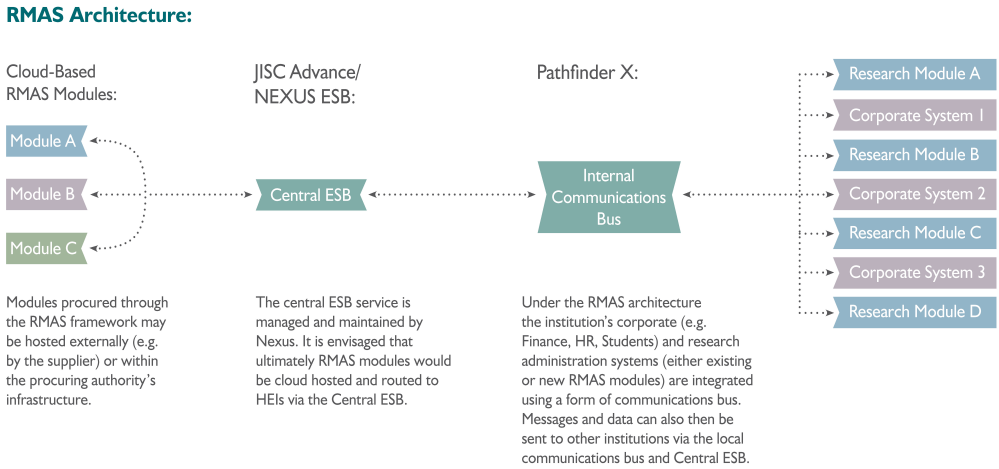

Figure 2: communications bus connecting to Central ESB

4 Extracting Data

The universities of Kent and Sunderland initially investigated messaging out of their respective Costing & Pricing Management tools since both Institutions operated systems from the same supplier. However, this is where the first difficulty arose for the pathfinders. It soon became apparent that although the two the pathfinders had the same "database" the structures were completely different, which meant that the converter that had been developed by Kent was of little use to the other pathfinders. This meant that additional converters had to be developed by Sunderland, highlighting the important issue that although systems may be from a common supplier, integrations solutions are often specific to each installation.

On closer inspection the Costing & Pricing Management database had no resemblance to the graphical User Interface (GUI) that is used by the respective HEIs to input data into the system. Investigating the Costing & Pricing Management system showed that there were various empty tables that had been imported when the system was installed and never removed. Furthermore, the database also has various tables with staff details, however this became problematic when attempting to extract data, this was due to the tables which linked members of staff had multiple Staff_ID's, which meant that data had to be investigated further to find out which of the ID's were linked to each project, and which were link to other tables within the database. The next complication that was encountered was that various aspects of the Costing & Pricing Management system that had no link to other tables within the system, an example of this was the "Project Status" field which is a numerical value. However, once the data was linked the pathfinders could run their converters, to output data as an Extensible Markup Language (CERIF-XML) file.

At Exeter, development work focused around the use of Microsoft SQL Server and a component called Linked Server to present integrated systems data. Microsoft who supply SQL Server have a multitude of drivers freely available on the internet to connect to various database sources.SQL Server has the ability to use these drivers and create named connections that allow the database structure of non SQL Server databases to become visible. This allows queries to be run in a single SQL Server environment against different database formats.

Connections to the underlying databases for HR, Finance, Project and Student records, were made using these linked servers. From these connections, another part of the SQL Server toolset was utilised: SQL Server Integration Services – SSIS. Using this tool, data could be extracted, transformed and loaded into, either XML format or directly into another database table for reporting. Often combining disparate datasets for cross system reporting.

The design of SSIS workflows, also allowed data checks across systems to alert administrators by email of exceptions or to perform joined up exception reporting.

5 EuroCRIS and CERIF in the RMAS Project

First developed in 1991, CERIF had a single-entry, as well as a simple record focus. Now in 2012 as it progresses into version 1.4,, CERIF is increasingly focusing on the evaluation of research outputs. Increasingly organisations are incorporating the CERIF standard.

RMAS is engaging with EuroCRIS on two fronts:

- Facilitating the exchange of information between different tools by ensuring each RMAS-compliant tool agrees on an identical taxonomy for research entities. To achieve this pathfinders have formulated a list of requirements that EuroCRIS are engaged to deliver.

- RMAS contracted the services of EuroCRIS experts in order to provide mappings of data sets common to research projects. This will facilitate the exchange of information from core systems for example HR, Publications, Students etc.

6 Sending Messages

When the pathfinders had their XML files with the data extracted from their research system, the next stage was to ensure that they could communicate with each other via the global communications bus. This was achieved through close work with JISC Advance – Nexus, previously known as Shared Services Procurement Support (SSPS), who developed a communications bus for the pathfinders to connect to. The first step was to create a Web Services Description Language (WSDL) file. This XML format describes network services as a set of 'endpoints' which act upon either document or procedure orientated information. The WSDL files also contained a Simple Object Access Protocol (SOAP) header containing information such as the message destination and source HEI.

Creating the WSDL file allowed Nexus to design the endpoint for the pathfinders to pass messages onto the global communications bus, and then through service routing the message gets pushed to the correct HEI, by way of reading the 'HEI code' which is contained within the SOAP Header, the codes that were used are the HEI codes from HESA (H-0122 - The University of Kent, H-0078 - University of Sunderland, H-0119 - University of Exeter).

Following this, each of the pathfinders needed to define endpoints, which will allow the routing service to have an address to which the messages will be sent.

7 Pushing/Polling the RMAS Bus

It is important for the local communications Bus to be able to both push and poll the local RMAS bus to ensure that messages can be both sent and received by the HEIs, this is achieved via SOAP. The way that RMAS achieves this is via connectors, which are used to call web services, via a service WSDL.

7.1 Pushing Event Messages to the RMAS Bus

When a RMAS event occurs in the RMAS supplier tool, an RMAS message (in the EuroCRIS defined RMAS-CERIF XML) should be sent to the local communications bus. These events are the same as the message types (which are currently being defined with EuroCRIS). Examples for proposal management include 'proposal-created', 'proposal-updated' and 'proposal-deleted'. When any of these actions occur the RMAS supplier system must trigger the sending of an RMAS message.

The message itself must be a well-formed, valid CERIF XML string which contains the message type (e.g. 'proposal-created') and the content (in this case, information about the proposal). The content must also include the entity's RMAS ID to allow other tools to recognise it.

The RMAS ID is proposed to be URN of the form:

urn:rmas:0122:pfact:510f5015-c5bd-4619-8c11-676b4a13fc36:12456:7891011

Where: 'rmas' indicates the RMAS namespace

'0122' is the institution ID that the message is sent from. E.g. H-0119, H-0122, H-0078. This is also contained within the SOAP header for the HE Cloud relaying service but is also useful at the receiving end.

'pfact' is the tool identifier

A UUID that distinguishes this message from others.

The unique record id out of the sending tool. In this case the identifier would be the unique record id in the senders pFact system.

Currently, the SOAP's *pushEvent* method (Englander 2002) will respond with a boolean *true* if the message is accepted, and *false* if it is was rejected.

```
// where to send the message
string uri = "http://www.heftv.ws/rmasSendService";

// create a request
HttpWebRequest request = (HttpWebRequest)
WebRequest.Create(uri); request.KeepAlive = false;
request.ProtocolVersion = HttpVersion.Version10;
request.Method = "POST";

// turn our request string into a byte stream
byte[] postBytes = Encoding.ASCII.GetBytes(msg_to_send);

// specify the message type
request.ContentType = "text/xml; charset=utf-8";
request.ContentLength = postBytes.Length;
Stream requestStream = request.GetRequestStream();

// now send it
requestStream.Write(postBytes, 0, postBytes.Length);
requestStream.Close();

// get the response handle
HttpWebResponse WebResp = (HttpWebResponse)request.GetResponse();
Stream resStream = WebResp.GetResponseStream();

// information about the response
Console.WriteLine(WebResp.StatusCode);
```

Figure 3 - C# code - pushing message

7.2 Polling the RMAS Bus

RMAS supplier tools also need to receive messages from other tools. They do this by polling the Central ESB which responds with any relevant messages. It was decided that connectors should poll, rather than have information pushed to them, to reduce complexity on the supplier side (as it is simpler to be a SOAP client than a SOAP server). This will also benefit institutions as applications are often hosted across different servers and firewalls.

To poll the SOAP endpoint, the client needs to send a *getEvents* message with a timestamp. The endpoint will respond with a list of RMAS messages (CERIF XML) that have been generated since the provided timestamp, and are allowed (by the bus's routing strategy) to be delivered to the requesting connector. If there are no applicable messages for a given poll, the endpoint will respond with an empty list. The timestamp must conform to the ISO8601 specification and must have both a date and time. It must also be UTC (not local time) and must therefore end with a 'Z', e.g. 2011-10-12T07:15:23Z.

It is the responsibility of the connector to remember the last time it polled, to ensure it does not miss any data, or get duplicate messages. The connector can also decide how frequently to poll. Shorter polling lengths will cause a minor increase in load, but will update more frequently. Longer polls mean the application will spend more time in an 'out of date' state.

Fig. 4 shows the PHP code for polling the local HEI Communications Bus for messages.

```
$client = new SoapClient('http://localhost:6980/EventService?wsdl');
$messages = $client->getEvents( array('timestamp' => '2012-01-05T07:00:00Z') );
print $messages
```

Figure 4 – PHP code - polling for message

8 The RMAS 'Toolkit'

Recognising that all UK HEI's have different infrastructures, processes, workflow and business drivers, it was essential that RMAS provided a range of tools to accommodate the requirements of the whole sector. In addition to the messaging technology described above, the University of Exeter focussed on the development of system integration based on readily available SQL tools. Benefits analysis shows how these are delivering efficiency savings in excess of £100k Per annum and have facilitated a step-change in data quality in source systems.

Noting that RMAS solutions may be enhanced through the use of CERIF, the pathfinders have worked collaboratively with euroCRIS to extend the CERIF standard and vocabulary. This workstream encompassed the full range of attributes relating to HR, finance, project, outputs and publications datasets that would be required for RMAS solutions incorporating CERIF XML data exchange formats to be adopted across the sector. All of the outputs from RMAS will be made available for the sector via a centrally maintained repository of open source software tools, user guides, case studies and technical documentation.

9 How does RMAS progress?

In the RMAS project and previous feasibility studies it has been found that many HEIs would benefit from a RMAS framework, and it is clear that when implemented substantial savings can result. However the project encountered various difficulties during development. The biggest problem was inconsistent and/or redundant data. The most appropriate way to address this issue is to configure the databases when a RMAS system is installed, ensuring that all systems are identical. This will allow messages to be sent to the Central ESB, and then pushed to other HEI systems with relative ease.

Working closely with euroCRIS, will enable the pathfinders to implement the standards of CERIF, which provides a canonical reference at data level, as well as metadata level (Jörg ***et al*** 2012). Further to this, CERIF will allow the formalisation of research entities and their relationships, this will allow for the use of a current research information system (CRIS). The CRIS will enable the pathfinders to define a data exchange format, which will enable the creation of a common data warehouse.

10 Conclusion

This paper was intended to investigate the feasibility of integrating RMAS systems, and to conclude what this integration would mean to UK HEIs and to the funding bodies HEFCE and JISC. The investigation showed that the development of an integrated RMAS would be beneficial to all through the evidenced potential savings of RMAS. Furthermore there would be a greater chance of information being passed between HEIs housing valid data, more often. This would be possible through the adoption of the CERIF standard for sharing of data.

In 12 months RMAS has delivered an outline framework for a suite of proven solutions in the HE sector to communicate in order have the capability to transform many of the research administration processes within HEIs, leading to substantial efficiency savings and significant data quality improvements.

References

Binge, A. (2009-a) *Research Management and Administration System – feasibility study part 1, University of Exeter*. Available online (http://www.exeter.ac.uk/research/rmas/about/keydocuments/) (accessed on 21st February 2012)

Binge, A. (2009-b) *Research Management and Administration System – feasibility study part 2, University of Exeter.* Available online (http://www.exeter.ac.uk/research/rmas/about/keydocuments/)(accessed on 21st February 2012)

Common European Research Information Format, euroCRIS 2012, version 1.3 (www.eurocris.org) available online (http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.3/Specifications/CERIF1.3_FDM.pdf) (accessed on 21st February 2012)

Englander, R. (2002) – *Java and SOAP*, O'Reilly publishing.

Jörg, B., Dvořák, J. Vestdam, T., van Grootel, G., Jeffery, K. and Clements, A. (2012) CERIF 1.3 *XML - Data Exchange Format Specification*. Available online (http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.3/Specifications/CERIF1.3_XML.pdf) (accessed on 1st March 2012)

Contact Information

Paul Fairley
paul.fairley@sunderland.ac.uk
Kevin Ginty
kevin.ginty@sunderland.ac.uk
University of Sunderland
Centre for Internet Technologies
Digital Media Building
St Peter's Campus
St. Peter's Way
North Sands
UK

Simon Kerridge
S.R.Kerridge@kent.ac.uk
University of Kent
Research Services
Canterbury
UK

Simon Foster
S.Foster@exeter.ac.uk
Stephen Trowell
S.Trowell@exeter.ac.uk
Steve Hustwayte
S.Hustwayte@exeter.ac.uk
University of Exeter
Research and Knowledge Transfer
the Innovation Centre
Rennes Drive
Exeter
Devon
UK

Steve Butcher
s.butcher@hefce.ac.uk
HEFCE
Northavon House
Coldharbour Lane
Bristol
UK

Josh Brown
j.brown@jisc.ac.uk
Roland Major
roland.major@jiscadvance.ac.uk
JISC
Lower Ground Floor
Brettenham House
5 Lancaster Place
London
UK