


<b>Cover Sheet for Proposals</b> <i>(All sections must be completed)</i>			
<b>Name of Capital Programme:</b>		<b>e-Research : e-Infrastructure</b>	
<b>Name of Lead Institution:</b>		University of York	
<b>Name of Proposed Project:</b>		iRODS Evaluation and Demonstrator (iREAD)	
<b>Name(s) of Project Partner(s):</b>		University of York, White Rose Consortium	
<b>Full Contact Details for Primary Contact:</b> <b>Name:</b> Professor Jim Austin <b>Position:</b> Professor of Neural Computation <b>Email:</b> austin@cs.york.ac.uk <b>Address:</b> Department of Computer Science, University of York, York YO10 5DD  <b>Tel:</b> 01904 432734 <b>Fax:</b> 01904 432767			
<b>Length of Project:</b>	15months		
<b>Project Start Date:</b>	01.01.08	<b>Project End Date:</b>	30.03.09
<b>Total Funding Requested from JISC:</b>		£176,641	
<b>Funding Broken Down over Financial Years (Apr–Mar):</b>			
<b>Apr 07-Mar 08</b>		<b>Apr 08-Mar 09</b>	
£35,328		£141,313	
<b>Total Institutional Contributions:</b>		20%	
<b>Outline Project Description</b> The project provides an in-depth evaluation of the iRODS system to investigate and demonstrate the capabilities of iRODS and to assess its impact and potential benefit to the UK e-Science community. This evaluation will address iRODS compatibility with existing e-Science infrastructure and assess its integration with leading edge security frameworks for VO management. The evaluation will be driven from real world VO requirements and Use-Cases derived from a major UK e-Science grid project, CARMEN. A reference demonstration implementation of the new iRODS system will be built, which will be accessible to the wider e-Science community. The project team have extensive experience of large-scale implementations of SRB systems and of the requirements for VO's within active Grid environments. This knowledge will be brought to bear in building a reference iRODS system on the White Rose Grid.			
<b>I have looked at the example FOI form at Appendix A and included an FOI form in the attached bid (Tick Box)</b>	YES ✓	NO	
<b>I have read the Circular and associated Terms and Conditions of Grant at Appendix B (Tick Box)</b>	YES ✓	NO	

## FOI Withheld Information Form

We would like JISC to consider withholding the following sections or paragraphs from disclosure, should the contents of this proposal be requested under the Freedom of Information Act, or if we are successful in our bid for funding and our project proposal is made available on JISC's website.

We acknowledge that the FOI Withheld Information Form is of indicative value only and that JISC may nevertheless be obliged to disclose this information in accordance with the requirements of the Act. We acknowledge that the final decision on disclosure rests with JISC.

<b>Section / Paragraph No.</b>	<b>Relevant exemption from disclosure under FOI</b>	<b>Justification</b>
Section 5	Not on JISC website	Sensitive institutional finance information

## 1 Introduction

---

This document describes a proposal for an evaluation and demonstrator project within the JISC e-Research e-Infrastructure initiative within the themed area of **Federated Tools and Services**. The project will provide an in-depth evaluation and demonstration of the iRODS system<sup>1</sup>, assessing its capabilities and role within a distributed data management scheme and complex virtual organisation driven by real-world requirements from the CARMEN<sup>2</sup> e-Science project.

## 2 Fit to Programme Objectives and Project Outline

---

The project call identifies that distributed data management issues are still a major challenge in addressing the ubiquity and usability of Grid systems, particularly within the context of complex virtual organisations (VO) where flexible security models and fine-grained role based access are a pre-requisite. The challenges set out in the proposal match with our own extended experience of dealing with distributed data management systems within the context of e-Science collaborative projects and working environments (such as DAME, BROADEN and the White Rose Grid, see section 6). Our research activities have made extensive use of the Storage Request Broker (SRB) software (SRB is the pre-cursor to iRODS), both as an enabling data handling middleware stack and as a platform for integration with our novel distributed data-mining software, the Pattern Match Controller system.

Within our e-Science SRB implementations we have repeatedly hit against the limitations of the Storage Request Broker, particularly in regard to integration with complex VO's. To this end DAME constructed a bespoke role-access management system which integrated SRB and portal based role-access mechanisms built on the UK e-Science X509 certificate infrastructure. However, this bespoke solution, whilst adequate for DAME's demonstration purposes was not a scalable solution. Similarly, many of our ongoing e-Science projects have hit upon limitations with the metadata schemes of SRB, which have proved to be too restrictive and inflexible to support complex meta-data schemes for large-scale e-Science projects. This limitation is particularly pertinent to the new CARMEN e-Science project, which is handling complex neuroscience data for which semantically rich and flexible meta-data labelling schemes are critical.

One potential solution for managing some of these inherent data management challenges is the newly released version of Storage Request Broker, called iRODS. This offers the potential for integration within complex VO's through the addition of rule-based components for data access, which can support integration with role-based access schemes, and offers the capability of more flexible user-defined meta-data schemas.

Hence, we are proposing a 15month, in-depth evaluation project to investigate and demonstrate the capabilities of iRODS, within the context of our ongoing e-Science and White Rose Grid activities for the wider benefit of the e-Science community.

The project addresses the objectives and aims of the call through the definition of five main deliverables:

1. A demonstration implementation of the iRODS system data management system deployed for open access and evaluation via the White Rose Grid and NGS system, which will demonstrate controlled access to distributed data across a collaborative VO environment;
2. An evaluation of the iRODS system for fine-grained role-based access in a real-world Grid based virtual organisation scenario, driven by the use-case requirements from an active national large-scale e-Science project, CARMEN;
3. An assessment of the maturity of the iRODS technology for integration with existing UK e-Science security infrastructures and systems such as Grid-shib and Kerberos;
4. Analysis of the capabilities of iRODS for integration with existing e-Science SOA infrastructure and middleware systems, notably the York PMC middleware stack for distributed data-mining;

---

<sup>1</sup> See [http://irods.sdsc.edu/index.php/Main\\_Page](http://irods.sdsc.edu/index.php/Main_Page)

<sup>2</sup> See <http://www.carmen.org.uk>

5. Wide-scale dissemination and technology transfer of the project results via national initiatives including the National Grid Service, White Rose Grid e-Science centre, the OMII and WUN.

The focus of activities 1-4 is elaborated below; item 5 is addressed in section 4.

## 2.1 Deliverable 1: iRODS Demonstration Implementation

A core element of the project will be the development of a reference demonstration implementation of the new iRODS system, which will be accessible to the wider e-Science community. The project team have extensive experience of large-scale implementations of SRB systems and of the requirements for VO's within active Grid environments, most notably the White Rose Grid (WRG). This knowledge will be brought to bear in building a reference iRODS system on the White Rose Grid.

There are several benefits to implementation within the WRG environment:

1. The White Rose Grid is a fully operational Grid environment, running a mix of leading edge middleware infrastructure with a distributed user-base across the three host University sites (Leeds, Sheffield, York). This environment provides the opportunity to assess the impact of iRODS deployment in a realistic, large-scale operating environment, across a range of diverse security policies and user authentication schemes (such as Grid-shib).
2. The White Rose Grid has an existing and functional SRB infrastructure, which is populated with a diverse range of datasets across numerous data zones. This will provide:
  - a. The opportunity to assess the impact of migration from SRB to iRODS (particularly those features of iRODS which have been developed to support this migration path)
  - b. The opportunity to build upon existing WRG SRB demonstrators and data sets to provide a demonstration system for the wider e-Science community. This re-use will remove the overhead associated with building new demonstration applications and allow the development team to focus on the core functional capability assessment of iRODS
  - c. The WRG is a compute node within the National Grid Service (NGS) and hence is ideally placed to scale the demonstration capability to the NGS and to provide demonstration access to NGS users.

### DELIVERABLES:

- iRODS reference implementation
- iRODS application demonstrator on WRG
- Reports on migration to iRODS infrastructure

## 2.2 Deliverable 2 – iRODS Role-based access in VO Context

An extremely important aspect of the project will be the assessment of iRODS within the context of real-world operating constraints for Grid-based VO's. To this end, iRODS assessment will be extensively supported by an evaluation within the scope of the CARMEN e-Science project.

CARMEN (Code Analysis Repository and Modelling for e-Neuroscience) is a 4-year, £4million, e-Science Pilot Project funded by the EPSRC which started Oct 2006. The project brings together experimental and theoretical neuroscientists with computer scientists to address the complete lifecycle of neuroscience knowledge, and to develop a collaborative working environment for neuroscience data management and data analysis. The system will rely heavily on distributed data management capability, and is a complex VO with diverse stakeholders and highly sensitive datasets which must be securely managed, including authorised and authenticated role-based access. CARMEN will be a live UK system, used by neuroscientists in the project. It is being aligned with the international neuroscience initiative (INCF), as well as the national INCF node.

CARMEN is already deploying SRB across its Grid nodes, CAIRNS, which is being used to provide data virtualisation. However, issues of role-based access and integration with the CARMEN security infrastructure, GOLD, and sophisticated meta-data management will be pivotal to the success of the project. GOLD is a UK e-Science EPSRC funded project investigating practical aspects of developing and deploying virtual organisations (VOs) and developing proof-

of-concept OGSA/SOA middleware and tools for VO management. The project focuses on highly dynamic VOs across full R&D lifecycles. The developed middleware is generic and applicable to other e-Science sectors. As such it has been adopted as the VO and security framework for CARMEN. Assessment of iRODS will be carried out to investigate the mapping between the GOLD VO mechanisms, and the ability to support rules for role-based access in iRODS.

The York team have already reviewed the features of iRODS and noted that there was considerable potential for enhanced functionality within CARMEN deployment. However, the adoption of the technology is not in the project plan. Furthermore, the technology is an unknown quantity, and the project will not adopt the technology until it is proven. This project will allow us to do this.

This study will permit a parallel evaluation path, whereby the iRODS system can be fully evaluated alongside the current SRB deployment, **with a view to early adoption in CARMEN** if the iRODS functionality is stable. The functional requirements for CARMEN (which are being reviewed in Q4 2007) will produce a set of detailed use-case and scenario documents which will form the basis of the iRODS reference implementation. Having clear e-Science project driven requirements for role-based access and security schemes will permit the development team to approach the iRODS evaluation from an applied, systems integration approach rather than as an abstract academic assessment.

#### DELIVERABLES:

- Use Case scenarios for complex Grid-based VO
- iRODS implementation of role-based access for VO
- Integration with GOLD VO infrastructure
- Documentation and reports on iRODS role-based access mechanisms
- Analysis for CARMEN pilot project adoption

### **2.3 Deliverable 3 – Integration with Grid Security Frameworks**

Distributed security models which are scalable and flexible are a critical requirement for the real-world creation of dynamic VO's in Grid based systems. The UK e-Science programme has funded extensive research in these issues and a number of security frameworks and middleware stacks are emerging as a result. Two emerging security models with wide academic take-up are Kerberos and Shibboleth (and its derivative Grid-Shib). iRODS will need to interact with these two security models if it is to be widely adopted within the UK e-Science community. Hence, the proposal will seek to evaluate and demonstrate the integration of both Shibboleth (Grid-Shib) and Kerberos security models within a Grid deployment. Key aspects of the iRODS technology that will be evaluated are its publicised features for GSI authentication, Shibboleth authentication, and access controls on rules and micro-services. A successful integration of these two technologies will demonstrate how iRODS can underpin a complex VO in a distributed Grid framework.

This work will be aided by existing experience and infrastructure within the WRG, which has instances of both Kerberos and Grid-Shib already deployed.

#### DELIVERABLES:

- iRODS integration with Grid-Shib security framework
- iRODS integration with Kerberos security framework
- Demonstration implementations on the WRG
- User Documentation and reports on integration process

### **2.4 Deliverable 4 – Integration with UK e-Science Infrastructure**

Early UK e-Science adoption of iRODS will be dependent upon its ability to readily integrate with existing e-Science infrastructure. To this end, an example, pilot, integration of iRODS with the Pattern Match Control system from York will be assessed. PMC has been developed by the York proposal team, and provides a highly representative OGSA/SOA compliant, e-Science middleware infrastructure.

Pattern Match Controller (PMC) is a distributed data mining technology for remote data management and analysis. It provides the capability to mine remote, non-relational datasets in-situ, and to abstract away the problems of distribution from the end-user or client. The current system has been deployed on top of several federated SRB data zones. It is a critical component within the CARMEN system for managing distributed pattern matching on neural spike train data.

It serves as a well-understood, highly functional middleware application that can benefit from integration with iRODS. Key integration issues will include:

- a. Assessment of the ability to transmit semantic search constraints with a request for a PMC driven search, such that the search can be directed or constrained at a remote node.
- b. Assessment of the ability to launch remote OGSA services via iRODS, at a remote data node, to facilitate pre or post-processing of the data.

The results of integration with this SOA middleware system will serve as the basis for 'lessons-learnt' documents, describing the development route for iRODS and its level of compatibility with UK e-Science infrastructure. Successful integration with the PMC technology will also lead to its early adoption within the deployed CARMEN system.

**DELIVERABLES:**

- iRODS integration with middleware SOA system, the PMC
- Demonstration implementations on the WRG
- User Documentation and reports on middleware integration process

### 2.5 E-Science Benefit and Impact

SRB has been a leading middleware infrastructure for the management of distributed data within Grid collaborative environments and VO's. However, it has limitations in regard to fine-grain access control, meta-data schema flexibility and integration with SOA applications. iRODS, as the next generation SRB, is intended to address some of these deficiencies. This project will determine the extent to which it has been successful in each of these strategic areas, and provide the UK e-Science community early insight into the suitability and efficacy of iRODS for adoption within ongoing Grid and e-Science projects. In funding a single experienced UK SRB development team to assess and demonstrate the iRODS system, there is the potential to save considerable re-work amongst the general e-Science community by avoiding replication by individual research groups needing to tackle the early deployment and implementation challenges. By providing a reference implementation of iRODS with security infrastructures such as Grid-Shib and GOLD, there will potentially be a generic framework available for the e-Science community which addresses the end-to-end issues of fine-grained access control for distributed data. This will be made open source via OMII and the WRG and would be available for early adoption within their programmes. The reports and evaluations will provide lessons-learnt documents which will facilitate easy adoption by other project teams. Any shortcomings of iRODS within the context of the UK e-Science infrastructure will be addressed with the SDSC iRODS development team with a view to influencing and enhancing the functional development of the software.

If the implementation is stable and mature, then immediate deployment within the CARMEN project will be ensured, and will provide further development momentum as a result of iRODS serving a pivotal role within a leading UK and global activity for e-Science research in neuroscience.

## 3 Work plan and Project Management

---

### 3.1 Summary Project Plan

Work Task ID	Task Description	Start (month)	End (month)
T1 Project Management	Define detailed project plans and Quality plan, webs-site setup	0	3
T2 Infrastructure	Set-up of baseline iRODS infrastructure on WRG and test environment	2	4
T3 Requirements Capture	Capturing VO stakeholder requirements and development of Use-Cases from CARMEN Project	0	3
T4 Role-based access techniques	Evaluate the mappings between security models and iRODS, inc GOLD and Grid-Shib security policies	4	9

T5 Distributed Data Mining	Integration of PMC distributed data mining application with iRODS to evaluate dynamic process invocation via iRODS rule-base.	6	12
T6 Assessment of CARMEN deployment	Assessment of iRODS maturity and robustness for e-Science project deployment.	12	14
T7 Demonstrator – Role Based Access	Development of the role-based access demonstration system on the WRG	12	15
T8 Demonstrator – Application Integration	Development of the demonstrator to explore iRODS integration within existing UK e-Science middleware infrastructures and the PMC system.	12	15
T9 Reporting	Creation of final evaluation reports and documentation of demonstrators.	13	15
T10 Dissemination Activities	Workshops, Demonstrations, Publications	6	15

### 3.2 Project Management

The University of York will run the project, with close collaboration with the support staff teams from the White Rose Grid and the development teams and user-groups from the CARMEN project.

The project management team will be drawn from the academic leaders, Prof. Austin and Dr Tom Jackson. The project management team will be charged with the tasks of:

- Developing Project Management plans and Quality Plans
- Managing the projects technical progress;
- Reporting and collaboration with JISC Programme Manager;
- Managing coordination between diverse e-Science stakeholders;
- Liaison with the e-science projects CARMEN and the WRG;
- Wider UK dissemination and exploitation;

Project progress will be reviewed against the project's work-plan at regular project team meetings, and in consultation with the JISC project manager, at timescales set to JISC's preferred schedule.

The management team have considerable experience of managing the successful implementation of large-scale e-Science projects (see section 6) and the management formalisms and experience from those projects will be applied.

### 3.3 Intellectual Property

In line with JISC's policy the results of the project evaluation and the demonstration system will be made available to the wide UK HE and FE community. This will be ensured through dissemination of outputs via the OMMI, NeSC and the WRG. The demonstration system will be made available for assessment and interaction via the WRG, and will remain supported as a demonstration system beyond the life of the study.

The PMC software used for one demonstration application is proprietary to Cybula (York spin-out company) and the IP for this will remain with Cybula. Cybula will grant full rights for use during the project and for the life of the demonstrator system

### 3.4 Risk Assessment

Risk is an inherent part of any research project. However, the project management team will take steps to provide ongoing risk analysis of the project and will develop a risk register to help manage project plan exceptions. The following risks have been identified as part of the initial project analysis. This risk list will be reviewed in greater detail on development of a detailed project plan.

Risk	Like- lihood (1-5)	Impact (1-5)	Rank	Risk Mitigation
Instability of iRODS software	3	4	12	Project team will maximise working relationship with SDSC to promote bug fixes and feature enhancement
Learning Curve for iRODS	2	3	6	Experienced SRB developers will be deployed on the project from existing project team
Accessibility to CARMEN stakeholders	1	4	4	Key personnel will be shared between the projects to ensure close collaboration
Incompatibility of iRODS with UK e-Science infrastructure	1	4	4	The purpose of the project is to assess this issue and to influence the iRODS development programme where possible
Recruitment delays	1	4	4	The project staff are already in place
Access to WRG resource	1	3	3	WRG support staff will be on the project team and WRG Exec have committed to support of the project
Project Scope creep	1	3	3	Focused deliverables and demonstrators have been defined.
Impact of other software services/solutions	1	3	3	SRB is leading edge in terms of Grid data virtualisation

## 4 Dissemination and Engagement with the Community

Dissemination of the project results to the wider e-Science community and within the JISC e-Science infrastructure will be an important focus within the project. The project team and the WRG are well positioned to maximise the engagement with the wider e-Science community during the progress of the study, having an established reputation within the UK and international e-Science community, and engagement with a wide range of e-Science initiatives. The project lead members have active participation in the following activities and initiatives, all of which will serve as conduits for dissemination and partnership activities:

- Professor Austin is a member of the OMII advisory steering group;
- Members of the OGF data mining work groups and UK e-Science data mining groups (contributors to NeSC workshop events on this theme);
- Steering Committee members on the Grid Computing Now network<sup>3</sup> funded by the DTI;
- Members of the Steering Committee for the UK EPSRC All Hands Conference
- Board Members of the White Rose Grid Executive;
- Regular invitees and presenters to the Global Grid Forum and Globus World events (invited speakers 2005-2006);
- Lead partners in the WUNGrid initiative, as part of the World Wide University Network.

The working relationship with SDSC and Regan Moore will also be used to enhance the UK opportunity to influence the future strategic developments of iRODS.

The White Rose Grid consortium will provide full support for the project, including the host evaluation environment and, where required, technical support for hosting the demonstration system. The WRG is a node of the NGS system, and hence there will be immediate transparency of the WRG demonstration system through to the NGS community.

A number of specific activities will be created to disseminate the results of the evaluation and study:

1. A **Workshop event** will be hosted within the framework of the WRG e-Science centre. The e-Science centre is ideally placed to host and advertise a workshop event, given its established position within the UK e-Science community and due to its links with the NGS, the ETF and the other regional e-Science centres.
  - A **themed conference session** on distributed data management will be hosted by the project team as part of the Fifth International Conference on Condition Health Monitoring<sup>4</sup>, Edinburgh 15-18<sup>th</sup> July 2008. The ongoing project results will be presented.

<sup>3</sup> See [www.gridcomputingnow.org](http://www.gridcomputingnow.org)

- A **demonstration and presentation** will be made at the EPSRC All Hands Meeting as part of the WRG e-Science centre activities at the AHM.
- A **Webinar presentation** will be made via the Grid Computing Now's on-line, live webinar broadcast system.
- A **demonstration** will be provided across the WUNGrid infrastructure to present the results to the wider WUN membership (SDSC are a partner member);
- A **presentation** to the **OMII** will be made on the project outcomes and wider implications of iRODS use within other OMII initiatives such as OGSA-DAI.
- Participation in JISC e-infrastructure events will be fully supported, including participation in:
  - JISC bi-annual programme meetings
  - Cluster meetings and topic-related workshops, where appropriate
  - JISC CETIS special interest groups

A **website** will be developed to support the ongoing activities of the project and to disseminate the final project results. This will be hosted within the current WRG website portal<sup>5</sup>, to ensure wide exposure of the project activities and to maximise longevity of the website, as part of the WRG's ongoing portfolio of projects and community reports.

---

<sup>4</sup> See [www.cm2008-mfpt2008.org](http://www.cm2008-mfpt2008.org)

<sup>5</sup> See <http://www.wrgrid.org.uk/>



## 5.2 Institutional Support

The level of support being provided by the host institution is 20%. This has been set because the project has only relatively little direct institutional benefit, and the demonstrators and evaluation are being carried out predominantly as an exercise to advise and inform the UK e-Science community. The host institution will provide direct financial support through use of key WRG personnel to support the development and deployment of the demonstrators. The host institution will also meet the cost of the WRG compute resource time and facilities for hosting the demonstrators (including beyond the immediate life of the project). In addition, the costs of the dissemination activities will be met through the WRG e-Science centre, as a part of its national UK position within the e-Science programme.

## 6 Previous Experience of the Project Team

---

The project members are ideally situated to lead an analysis of iRODS; having a long-standing, working knowledge of the SRB system and the requirements for distributed data management, and an on-going and active research relationship with SDSC (Reagan Moore, SRB team leader) and Nirvana (commercial realisation of SRB software).

The core eScience team at York (Austin, Jackson, Jessop, Fletcher) has been responsible for several SRB implementations and the group was one of the first UK large-scale adopters of the SRB technology. The team have not been passive end-users, but have pushed SRB deployment extensively through a number of leading edge e-Science projects. These include:

**DAME**<sup>6</sup> (Distributed Aircraft Maintenance Environment) – An EPSRC funded e-Science pilot project which ran from 2001 to 2005, with a project budget of £3.5M. DAME was one of the first e-Science applications for SRB. SRB was deployed as the core data management tool for handling the virtualisation of distributed data assets for remote asset monitoring applications. The system was developed and operated across the White Rose Grid, as a fully functioning e-Science demonstrator for a Grid based condition-monitoring system. Issues of role access and data management were investigated in detail and a bespoke model for role-based access was developed and implemented on top of SRB. The York Pattern Match Control Technology was also integrated with the SRB system, as an active service that facilitated remote data-mining of the data assets managed via SRB.

**BROADEN**<sup>7</sup> (Business Resource Optimisation for Aftermarket & Design on Engineering Networks) – A DTI Technology Programme Funded project within the Inter-Enterprise Computing theme, running from 2005 to 2008, with a budget of £4.2M. BROADEN is a commercial follow-on to the successful DAME project, with Rolls-Royce and EDS taking lead role in a project to roll out the DAME system on an industrial scale. The York team have provided the expertise on the data management and data mining architecture, including the development of the Nirvana SRB implementation across the Rolls-Royce Internal grid.

**WUNGrid**<sup>8</sup> - York University are members of the World Wide University network (WUN) and have been a lead partner in the development of WUNGrid (with Reagan Moore of SDSC). WUNGrid is an initiative which brings together the international research expertise at the WUN member institutions to work on a series of innovative collaborative projects which are made possible by the emerging Grid technology. SRB has been a core component of this system and York have played the lead role in developing a distributed data-mining demonstrator across WUNGrid and SRB to show the capability for shared data management in academic VO's.

**WUN Medieval Studies** network<sup>9</sup> – also as part of the WUN activities York has been involved in providing technical support for the deployment of SRB as a data management stack for this medieval data repository project.

**CARMEN** - (Code Analysis Repository and Modelling for e-Neuroscience) is a 4-year, £4million, e-Science Pilot Project sponsored by the Engineering and Physical Sciences Research Council (UK). The project brings together experimental and theoretical neuroscientists with computer scientists to address the complete lifecycle of neuroscience knowledge, and to develop a

---

<sup>6</sup> See <http://www.cs.york.ac.uk/dame>

<sup>7</sup> See <http://www.dti.gov.uk/files/file23443.pdf>

<sup>8</sup> See <http://www.wungrid.ac.uk>

<sup>9</sup> See <http://cbers.shef.ac.uk/themes.htm>

collaborative working environment for neuroscience data management and data analysis. The system will rely heavily on distributed data management capability, and is a complex VO with diverse stakeholders and highly sensitive datasets which must be securely managed, including authorised and authenticated role-based access, some of which is managed by virtualised workflows.

In addition to the above academic record, the group has a spin-out company, Cybula Ltd, now selling SDE (from DAME), thus is aware of the commercial market and drivers in Grid technologies.

## 7 Supporting Letter(s) and Key Personnel

---

There are three supporting letters for the proposal:

- Institutional support – Professor John McDermid – Head of Department
- White Rose Consortium support – Dr Julian White – Chief Exec WRG
- SDSC iRODS team support – Dr Reagan Moore – Director, SDSC

### 7.1 Prof. Jim Austin, Department of Computer Science, University of York

Jim Austin directs the Advanced Computer Architectures group, one of the largest groups in Computer Science at York which covers the areas of Computer Vision, Neural Networks and Advanced Architectures. The Department has a 6\* REA rating and an Excellence in teaching. The Department has strong links with industry including the British Aerospace Dependable Systems Centre, a Rolls Royce Technology centre, labs funded by Nortel Networks as well as close collaborations with industries in most sectors of the computer industry.

Jim Austin has been running the group at York for over 20 years; it now contains over 30 members. Jim has managed over 22 industrially funded or co-funded grants and contracts since 1986, including large multi-site projects. He has served on many EPSRC and DTI grant reviewing panels as an active member of the EPSRC IT&CS College, as well as acting as panellist for the EPSRC strategic review of grants.

In 2000 he founded a successful commercial spinout, Cybula Ltd.

### 7.2 Dr Tom Jackson - Project Manager, Department of Computing Science, University of York)

Tom Jackson has a PhD in neural networks and a first degree in Electronics. He has extensive experience in academic and industrial project management. He was project manager for the 3-year e-Science project, DAME, which developed the SDE and SRB system to be applied within CARMEN. This involved coordinating the research work of over 16 RAs, four academic sites and two companies. He is now the project manager for the four-year DTI funded BROADEN project, which is taking the DAME technology into use by Rolls Royce. He has a B.Eng in Electronics (University of Salford University) and a PhD in Neural Networks (University of York University). Prior to coming to York he worked for 5 years at the European Joint Research Centre in Italy, managing technology research in support of EU framework programme policy and strategy. His PhD was based on work that has become the basis for many projects in the Advanced Computer Architectures Group at York, involving the use of neural associative memories for intelligent systems.

### 7.3 Mr Mark Jessop – Senior RA, Department of Computing Science, University of York)

Mark Jessop is a senior Research Assistant within the Advanced Computer Architecture Group, Department of Computer Science. He has extensive experience of e-Science software development and large-scale collaborative projects. He was responsible for the successful deployment and management of several leading edge SRB installations on e-Science demonstrator Grid projects, including DAME, GridWeld, BROADEN and now also within CARMEN. He has a comprehensive grounding in Grid middleware standards and developments, and is active within the SRB developer forums and OGF.



Our Ref: JAM/DJH

27 September 2007

JISC  
Northavon House  
Coldharbour Lane  
**BRISTOL**  
BS16 1Q

John A McDermid, FEng  
Professor of Software Engineering  
Head of Department

Dear Sirs,

## **JISC Circular 02/07: Federated Tools & Services Call**

I am writing to confirm the Computer Science Department's support for the University of York's proposal under the above call for work on the evaluation of the SDCS iRODS system.

Grid-related research within the department has gained considerable momentum during the course of the UK e-Science programme, and the Department has been a leading partner in several large-scale funded projects, proving effective both in basic science and in technology transfer, e.g. through the DAME and BROADEN projects. The issues identified within the current proposal, if funded, would both support our fundamental research on Grid computing and would contribute to our ongoing initiatives to maintain the White Rose Grid as a leading edge compute facility for both the White Rose Consortium and the national UK e-Science programme.

York is developing a new campus, and Computer Science will be one of the first Departments to move to the new campus, known as Heslington East. This gives us the opportunity to provide first rate facilities to support the White Rose Grid which was one of the major elements of the University's bid to HEFCE for support for the development. We plan to appoint new academic and support staff (probably in 2009) to support these activities. Thus there will be a major Institutional investment to complement the JISC-funded work, if it is approved.

Please let me know if you require any further information, e.g. on our plans for Heslington East.

Yours sincerely

John A McDermid



SAN DIEGO SUPERCOMPUTER CENTER

9500 GILMAN DRIVE, MC0505  
LA JOLLA, CALIFORNIA 92093-0505  
TELEPHONE: (858) 534 5073  
FAX: (858) 534 5077

Dr. Tom Jackson  
Department of Computer Science  
University of York  
York YO10 5DD

### Letter of Support for the Proposal on Federated Data and Services by the DAME/BROADEN Team

Dear Tom:

I write in my capacity as Director of the Data and Knowledge Systems group of the San Diego Supercomputer Center (SDSC), and as principal investigator for development of the integrated Rule-Oriented Data Grid (iRODS) technology. I base my response on SDSC collaborations with the DAME/BROADEN team including the Worldwide Universities Network data grid, which federates multiple SRB data grids to encourage collaborative research across international boundaries; the DAME distributed decision support system which has been integrated with the SRB data grid technology, and the CARMEN, Gridweld, and White Rose Grid projects that also utilize the SRB data grid. Our interactions with these projects have been highly successful, and have enabled the use of data grid technology to broaden access to shared collections and promote collaborative research.

The current research and development activities at SDSC are focused on the execution of management policies that are applied to the distributed shared collections. The iRODS system automates the application of rules that control the execution of remote micro-services. Rules can be defined that implement each management policy, whether related to time-dependent access controls, trustworthiness assessment criteria validation, integrity validation, data distribution, minimization of the risk of data loss, application of Institutional Research Board approval criteria or patient confidentiality mechanisms, and application of data analysis processes. The management processes are expressed as sets of micro-services applied at each remote storage location. This approach ensures that the management policies are enforced no matter which client is used to access the distributed shared collection.

The success of the iRODS technology is highly dependent upon the application of the technology in production data grids, the identification of those management functions required by each application, and the creation of generic infrastructure that enables the specific policies of each community to be implemented. Towards that end, SDSC seeks collaborations with projects that implement data sharing environments, data publication libraries, data preservation archives, and data analysis environments. We welcome the opportunity to collaborate with the DAME/BROADEN team on a large-scale implementation of the iRODS data grid technology.

Yours,

A handwritten signature in blue ink that reads 'Reagan W. Moore'.

Reagan W. Moore  
Director, Data and Knowledge Systems group  
San Diego Supercomputer Center

JISC,  
Northavon House,  
Coldharbour Lane,  
Bristol, BS16 1Q

24 September 2007

Dear JISC

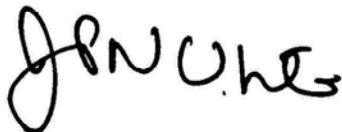
**Letter of Support Regarding: JISC Circular 02/07: Federated Tools & Services Call**

In regard to the above call, and York University's submission of a proposal relating to the evaluation of the SDCS iRODS system, I would like to indicate our full support for this activity within the context of the White Rose Grid e-Science activities.

The SRB system has been a major component within our project developments across the White Rose Grid, given its pre-eminent position as a middleware service for distributed data management. It has played a significant part in e-Science projects hosted and developed within the White Rose Grid, such as the EPSRC funded DAME and CARMEN activities. We anticipate that iRODS will also play a major role in the future, and this project proposal provides an ideal opportunity for the White Rose Grid to continue in its position as a leading edge adopter and developer of this software.

The White Rose Grid continues to be a leading edge facility within the White Rose Grid consortium and indeed within the larger UK e-Science infrastructure, and we welcome this opportunity to build on this position to host the proposed iRODS reference implementation and demonstrator.

Yours sincerely



Dr Julian White  
Chief Executive