Virtual Infrastructure with Database as a Service (VIDaaS) Project

Monday 1st December 2011

James A J Wilson
James.wilson@oucs.ox.ac.uk
VIDaaS objectives

- Deliver costs savings / efficiencies to UK research
- Create production-ready ‘Database as a Service’ (DaaS)
- Create hybrid cloud infrastructure on which to host DaaS
- Enable sharing of infrastructure and services
- Provide appropriate user training and support for DaaS
- Ensure services are sustainable beyond the life-span of the project
Research data infrastructure at Oxford

• Programme begun in 2008 with an internal scoping study

• Eidcsr (JISC funded, 2009-2010)
  – Scoping and piloting institutional data management infrastructure (software, metadata, responsibilities, etc.)

• Sudamih (JISC funded, 2010-2011)
  – Researcher training (organisation, software tools, etc.)
  – Pilot ‘Database as a Service’ (DaaS)

• VIDaaS (JISC & HEFCE funded, 2011-2012)
  – Full production-level DaaS, hosted on cloud infrastructure

• DaMaRO (JISC funded 2011-2013)
  – Integrate existing Oxford data management tools into enhanced infrastructure
Project team

Prof. Paul Jeffreys – Principal Investigator
Dr. Michael Fraser – Co-investigator
Dr. Stuart Lee – Director of OUCS
Dr. James A J Wilson – Project Manager
John Ireland – Lead Architect (DaaS)
Jon Hutchings – Lead Architect (VI)
Peter Jones – Lead Architect (infrastructure)
Asif Akram – Software Developer
Christian Fernau – Software Developer
Adrian Park – VI Specialist
Dr. Meriel Patrick - Analyst

Elena Blanco – Technical Author
Dr. Miko Flohr – Research Representative
Diane West – PA to the Director of IT
What is the DaaS?

- A web-based system that will enable researchers to quickly and intuitively
  - build a relational database from scratch, or
  - Import an existing database in common formats (such as Access)
- Generic data addition, editing, and querying interfaces
  - Research groups may, if desired, develop their own Web front-end interfaces to databases hosted by DaaS
- Databases centrally hosted, maintained, and routinely backed up
- Access controls to determine who can view or edit each database
  - Easy to share data with colleagues or even the general public
- Metadata capture to improve data rediscovery
DaaS Components

1. 
2. 
3. 
4. 
& 5...

Virtual Infrastructure with Database as a Service (ViDaaS)
User requirements

• Intuitive interface
• Better collaboration functions (data editing and sharing; multiple permission levels)
• Straightforward means of publishing datasets
• Automated back-up
• Ability to import and export data in various formats
• Training on principles of database design
• Cheap
• Great visualization tools [mostly out of scope]
VIDaaS enhancements

• Extended functionality beyond the humanities
• Frequent testing with ‘early adopters’ group
  – In exchange for 3 years hosting without charge
• Improved user interface, documentation, and support
• Enabling ‘publication’ of data, or sub-sets of data
  – Both dynamic and static data publishing options
• XML databases as well as relational databases
How the DaaS works

When a user registers a project, they are given a unique URL where they can access their database(s)

Each project = 1 WM

VMs can be moved between cloud infrastructures to cheapest / most appropriate hosting environment

Users need not be kept informed about where their application is being stored behind the scenes

Cloud-hosted online research databases on demand

- Be inspired with an idea for a new research project
- Choose the type of database you want (large or small, relational or XML)
- Cost the database into your funding proposal
- Dial-up the DaaS and start structuring your database or import an existing one

Gather your data

Analyze the data

Share your data with colleagues or the general public

Archive data of long-term value

Virtual Infrastructure with Database as a Service (ViDaaS)
## Delivery dates and dependencies

<table>
<thead>
<tr>
<th>Deliverable</th>
<th>Due Date</th>
<th>Dependency</th>
</tr>
</thead>
<tbody>
<tr>
<td>VMWare vSphere hybrid cloud in place</td>
<td>August 2011</td>
<td>VMware</td>
</tr>
<tr>
<td>Non-humanities user requirements identified</td>
<td>August 2011</td>
<td></td>
</tr>
<tr>
<td>Pilot DaaS on Vmware delivered to Eduserv</td>
<td>December 2011</td>
<td>Eduserv</td>
</tr>
<tr>
<td>Launch of DaaS to early adopters</td>
<td>December 2011</td>
<td></td>
</tr>
<tr>
<td>Launch of DaaS as production service within Oxford Oxford Research Database Service (ORDS)</td>
<td>January 2011</td>
<td></td>
</tr>
<tr>
<td>Launch of DaaS on Eduserv infrastructure</td>
<td>January 2012</td>
<td>Eduserv</td>
</tr>
<tr>
<td>Final ROI &amp; business case</td>
<td>February 2012</td>
<td></td>
</tr>
<tr>
<td>Functioning IAM and monitoring &amp; accounting system in place</td>
<td>February 2012</td>
<td></td>
</tr>
<tr>
<td>All documentation and training materials delivered</td>
<td>March 2012</td>
<td>DCC</td>
</tr>
<tr>
<td>Final report</td>
<td>March 2012</td>
<td></td>
</tr>
</tbody>
</table>
Business models & sustainability

Some basic sustainability models:

1. Each institution hosts DaaS themselves (whether on their own or national cloud infrastructure).
   Institution provides user support.
   An organisation ‘owns’ the service and coordinates software updates.
   A ‘service board’ collectively develops software, documentation, & training.

   - Each institution hosts DaaS themselves
   - Institution provides user support
   - An organisation ‘owns’ the service and coordinates software updates
   - A ‘service board’ collectively develops software, documentation, & training

2. As 1, but ‘owner’ also offers ‘train the trainer’

3. ‘Owner’ provides full user support via central helpdesk.

<table>
<thead>
<tr>
<th>Model</th>
<th>Package details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Packaged software</td>
<td>Simplest model &amp; necessary starting place. Operational burden on customer</td>
</tr>
<tr>
<td>Pre-configured software</td>
<td>Simplifies customer installation</td>
</tr>
<tr>
<td>Appliance</td>
<td>Highly portable solution with reduced burden on customers &amp; greatly increased return on investment</td>
</tr>
<tr>
<td>Cloud SaaS</td>
<td>No technical responsibilities for customers. Requires substantial development of self-service interface and charging mechanisms</td>
</tr>
<tr>
<td>National Service</td>
<td>Maximises economies of scale. Facilitates researcher mobility. Reduces HEI autonomy. Requires establishment of governance and operational organisations</td>
</tr>
</tbody>
</table>
Contacts

http://vidaas.oucs.ox.ac.uk/

http://sudamih.oucs.ox.ac.uk/

vidaas@oucs.ox.ac.uk