EPSRC-JISC Cloud Workshop 2011

Monday 26th September 2011

All Hands Meeting, Ron Cooke Hub, University of York

1. Introduction

At a workshop hosted by RCUK in July 2010, delegates from across academia explored the opportunities of Cloud computing for research as well as the barriers to its wider exploitation by researchers in the UK (see the report of the workshop on the Research Councils Workshop on Cloud Computing for Research page on the RCUK website).

The barriers identified at the RCUK Cloud Workshop included: the heterogeneity of Cloud platforms and the absence of standards; the lack of benchmarking information about the various commercial Cloud platforms; issues around the licensing of software; issues of privacy, ethics and security particularly in relation to data management and storage; the portability of tools and data to and from different Cloud platforms; uncertainties in prediction of the cost of research activities utilising Cloud; lack of education and training about Cloud computing methods; and, the absence of software tools, methods and platforms for researchers to properly exploit the Cloud.

In response to the outputs of this workshop EPSRC and JISC issued a joint call for pilot projects in October 2010 to explore and develop new Cloud computing technologies for research in engineering and the physical sciences. As a result of this, 11 short exploratory projects were supported to probe Cloud computing as a platform for enabling, facilitating and undertaking research in a range of EPS fields. A list of the pilot projects is given below, more detailed descriptions of each project is available in Annex 1 of the Workshop Annexes document:

- Flood Modelling for Cities Using Cloud Computing
- GATE Cloud Exploratory: Adapting the General Architecture for Text
- Rapport: Robust Application Porting for HPC in the Cloud
- Elastic Virtual Infrastructure for Research Applications
- myTrustedCloud: Towards a virtual private cloud
- CloudBIM: Exploring the Feasibility and Potential for Cloud Research in the Architecture, Engineering and Construction Sector
- My Private Cloud
- Fair Benchmarking for Cloud Computing Systems
- Optimal Scheduling of Scientific Application Workflows for Cloud-augmented Grid Infrastructures
- Clouds in Space
- ECHO: Enabling Cloud Hosted Organisations

2. Aims and Objectives

EPSRC and JISC organised a workshop for the Principal Investigators of the Cloud pilot projects and other interested stakeholders at the All Hand Meeting in York on Monday 26th September 2011 (see Annex 2 of the Workshop Annexes document for a full attendees list). The workshop was an

opportunity for the Cloud pilot projects to share their findings, to engage with the other Cloud pilot projects and to meet other interested stakeholders in Cloud research. EPSRC and JISC wanted to use this opportunity to canvas the opinion of the Cloud pilot project awardees so this could feed into any future strategy for the support of Cloud related research and investment.

The aims and objectives of the EPSRC-JISC Cloud Workshop were to:

- Disseminate the outputs and findings of the Cloud Pilot projects so far.
- Learn from the projects' experiences of using Cloud in a research environment.
- Understand what is required to widen the uptake of Cloud computing in research and what actions are required to enable this.
- Highlight the work being performed as part of the EPSRC-JISC supported Cost Analysis of Cloud Computing study being undertaken by Curtis and Cartwright.
- Encourage participation in the Cost Analysis of Cloud Computing study.

The agenda for the workshop can be found in Annex 3 of the Workshop Annexes document.

3. Introduction and Lightening Talks

The Principal Investigators of the Cloud pilot projects were given the opportunity to showcase the progress and outcomes of their projects to date during a poster session at the All Hands Meeting.

The attendees were given a short introduction from EPSRC and JISC on the aims and objectives of the workshop, and the current and future activities that are relevant to Cloud based research. This included:

1. An update on e-Infrastructure including the development of the community led document on the e-infrastructure eco-system led by Peter Coveney and the Tildesley report and roadmap that was commissioned by BIS.
2. EPSRC's development of a strategy around Software as an Infrastructure.
3. Relevant JISC call and events, for example calls in the Access & Identity Management (AIM) Programme: AIM Embedding and Research Tools Programme (RTP): Exploiting e-Infrastructure for Research and a workshop on curation on the Cloud planned for 2012.

Representatives from the Cloud pilot projects then gave a series of lightening talks. Each of the pilot project representatives was given 3 minutes to speak about the key findings from their project.

4. Lessons Learnt

As the Cloud pilot projects were coming to an end EPSRC and JISC felt that it was important to capture from the attendees the lessons that they had learnt through the research projects, and what advise they would give to future applicants looking to use the Cloud for research.

Attendees were divided into 5 groups and asked to discuss the following questions:

- ***What are the most important lessons that you have learnt as a result of the Cloud Computing Pilots***
- ***What advice would you give to a researcher looking to use the cloud based on your experience***

The groups were then asked to prioritise the top three lessons that they had learned from their experience and the top three pieces of advice and present this back to the workshop.

4.1 Discussion

Below is a summary of the lessons learnt and advice that the attendees highlighted as being most important in the future use of Cloud. The full, unedited outputs from this session can be found in Annex 4 of the Workshop Annexes document.

4.1.1 *What are the most important lessons that you have learnt as a result of the Cloud Computing Pilots*

- A clear understanding of the research problem is needed before Cloud is used to ensure that the Cloud has the features your application needs before you commit.
- Application performance on the Cloud needs to be benchmarked and monitored.
- Currently the science being performed on the Cloud may be limited due to the lack of understanding of the apparatus – however ambition should grow with greater understanding.
- There are genuine cases coming from the Cloud pilots where the use of the Cloud has been cost effective and useful.
- Costing models for Cloud are still complex and not well understood so can be difficult to justify, for example there is variability in charges, variability in performance and a lack of information about physical attributes and data and storage transfer charges. There is a need for dashboard metering where accumulated cost is easy to monitor.
- There can be issues associated with charging and paying for Cloud services because of institutional issues with credit card payments.
- Engagement with users is vital to ensure that the correct user interfaces are chosen and the application requirements and costs are well understood. There may be a need for domain specific services that can mediate between non-expert users and cloud providers.
- There are still a number of outstanding issues on data governance, privacy and security. There needs to be clarification of these issues and new technology (for example access controls) introduced before trust can be built between Cloud service providers and researchers.

4.1.2 *"What advice would you give to a researcher looking to use the cloud based on your experience"*

- Cloud is not simple and is not the easy option. Other computational resource options should be considered and a risk / benefit analysis performed before committing to using the Cloud.
- Researchers must understand their application before they start in terms of what they need and why. This should help users understand what the Cloud offers, its capabilities and its costs, for example:
    - CPU time and data volumes.
    - How much of the researchers requirements can be met by local *vs* external resource.
- Researchers need to talk to and work with knowledgeable and experienced people, for example:
    - Software consultants and developers.
    - Researchers with data and compute expertise, e.g. those experienced in the use of HPC.
- Researchers should spend time understanding the cost models and try and find out all the cost implications before committing larger resources.
- Researchers should work together with other researchers to apply pressure for uniform interfaces and application programme interfaces.

- There is a need to keep up to date on regular changes to features, pricing models and available services.

5. Cost Analysis of Cloud Computing

Rob Hawtin from Curtis and Cartwright introduced the Cost Analysis of Cloud Computing study that EPSRC and JISC commissioned in the summer of 2011.

The purpose of this study is to analyse the costs of Cloud computing for the types of research task within the remit of EPSRC and beyond, specifically in comparison to the costs of other approaches such as institutional facilities or distributed infrastructures such as grids. The study is looking to provide guidance to three distinct stakeholder groups:

- Researchers considering the appropriate computing approach to their own research problem.
- Institutions making investment decisions for computing infrastructure for research in engineering and physical sciences.
- EPSRC and funders making grant allocation decisions regarding cloud and other computing infrastructures.

Curtis and Cartwright were keen to involve the participants of the Cloud pilot projects and asked for information and feedback on the following points:

- What are the key things you've learned about the *cost* of doing research on the Cloud?
- Can you help us make comparisons with local provision?
- What advice do you think researchers, institutions and research councils need regarding the cost of Cloud computing?

The Cost Analysis of Cloud Computing study is due to report in early 2012.

6. What Next?

EPSRC and JISC wanted to draw on the collective experience of the Cloud pilot projects to get information on how the uptake of Cloud computing could be encouraged and supported in the research community (where appropriate) and who should take ownership of the actions that would help facilitate this increased uptake. A steer on how EPSRC and JISC involvement could add value was requested to help inform any future strategy for the support of Cloud.

Attendees were divided into 5 groups and asked to discuss the following two questions:

- ***From your experience what conditions do you think need to be met for a wider uptake of cloud computing in research in terms of:***
  1. ***The technical conditions to enable wider uptake***
  2. ***The conditions for institutional uptake***
  3. ***The conditions for individual researchers / research groups uptake***

- ***What specific actions are required to support this?***
  - ***Funder actions e.g. what do EPSRC, RCUK and JISC need to do?***
  - ***Institutional actions***
  - ***Researcher and community actions***
  - ***Industrial actions***

Attendees were then asked to prioritise the top 4 actions and the action owners and present these back to the workshop audience.

6.1 Discussion

Below is a summary of the discussion that the attendees had on how the wider uptake of Cloud computing could be supported and the roles that funders, institutions, researchers and industry should take in this. The full outputs from this session can be found in Annex 5 of the Workshop Annexes document.

6.1.1 *From your experience what conditions do you think need to be met for a wider uptake of cloud computing in research in terms of?*

1. ***The technical conditions to enable wider uptake***
2. ***The conditions for institutional uptake***
3. ***The conditions for individual researchers / research groups uptake***

6.1.1.1 Technical conditions to enable wider uptake:

- Improve the range of resources and services available to researchers.
- Development of simple standards and application programme interfaces that link with popular environments.
- Improved bandwidth for external data transfers.
- Further research and development into the use of multi-cloud which is still difficult to do.
- Improved billing systems, which allow central accounting and usage caps.

6.1.1.2 The conditions for institutional uptake:

- The business case for the use of Cloud needs to be made i.e. how does it compare with the true cost of local support and provision.
- There needs to be interoperability between local resources and Cloud, and between different Cloud bursting services.
- Institutions and users need to be able to trust the service provider. In order to promote this there needs to be greater predictability and confidence in the billing, accounting, authorising and authorising caps.
- Better arrangements for charging institutions i.e. credit card issue.
- Better 'sandboxing' or testing environments for learning and teaching about the Cloud in institutions.
- Better understanding by University management of the potential benefits.
- Brokerage is needed, for example, the use of campus champions, community experts or support from the Software Sustainability Institute or through the running of institutional "how to" events or road shows.
- Monitoring of the service.

6.1.1.3 The conditions for individual researchers / research groups uptake:

- There needs to be a better awareness that Cloud is an option and that funders accept the use of Cloud computing resource on research projects.
- Ensuring that there is the right level of expertise/skills or available training for researchers and groups looking to use the Cloud.
- Researchers need to understand whether using the Cloud will be worth it, through the production of good and bad exemplars.

- Available funding to support initial community efforts on the Cloud.
- Access to faster/more powerful machines.
- More transparent billing charges and costs.
- Greater level of industry and academic engagement leading to the production of clear comparable business models.

6.1.2 What specific actions are required to support this?

- ***Funder actions e.g. what do EPSRC, RCUK and JISC need to do?***
- ***Institutional actions***
- ***Researcher and community actions***
- ***Industrial actions***

6.1.2.1 Funder actions e.g. what do EPSRC, RCUK and JISC need to do?

- Guidance on the type of support that can be applied for, for example pay-per-use computing as oppose to capital expenditure, should be provided to applicants by funders. Awareness needs to be raised that Cloud is an alternative or an option.
- Funders could help researchers understand whether it is worth using Cloud by commissioning cost / performance studies that show both good and bad exemplars of Cloud usage. These should be disseminated to the community and at an institutional level.
- In order to change researcher culture funders should consider what incentives, penalties or training could be used to encourage uptake where appropriate.
- Funders should provide support for application based work using Cloud experts in research domains and for the user informed development of tools and techniques on Cloud.
- Funders should look at extending support for the Cloud pilot projects, or the groups that have resulted from this call.
- Funders such as JISC (possibly through the University Modernisation Fund) could support the building of a better sand box / test environments for research communities to use to familiarise themselves with the Cloud and learn how to use it.

6.1.2.2 Institutional actions

- Institutions need to understand and buy-in to the technology, from the top and on the ground (potentially running institutional “how to” road shows), promoting the use of Cloud services where appropriate. To do this institutions need to undertake a proper risk analysis and look at the results of studies on the cost and performance of Cloud to understand the cost implications and how Cloud can add value to existing local IT provision.
- A clear internal model needs to be developed for institutional use and access of Infrastructure as a Service.

6.1.2.3 Researcher and community actions

- The research community can work to provide reference models for Cloud use, publishing raw data and algorithms.
- Researchers should understand their application requirements and try out the Cloud before making a commitment.
- Researchers can look to develop community action at EU level *c.f.* EGI/PRACE.

6.1.2.4 Industrial actions

- Cloud providers and research user communities should develop a better and stronger engagement which could lead to:
    - Improved services, such as provision of better Cloud middleware.
    - Provision and agreement of suitable Service Level Agreements.
    - More flexible / Institutional friendly billing practices, including greater price transparency.
    - Clear and comparable business models.
    - Academic discounts.

7 Next Steps

EPSRC and JISC agreed to write up the workshop outputs and circulate to attendees.

The Cloud pilot project Principal Investigators were encouraged to get involved with the in the Cost Analysis of Cloud Computing study by contacting Rob Hawtin: rob.hawtin@curtiscartwright.co.uk.

A fuller write up of the outcomes of the Cloud pilot projects will be issued by JISC and EPSRC in early 2012.