# Simulation as a Service
## *A Case Study of Provisioning Scientific Simulation Software via a Cloud Service*

Morgan Eldred, Alice Good and Carl Adams
*School of Computing, University of Portsmouth, Portsmouth, U.K.*
*morgan.eldred@myport.ac.uk, {alice.good, carl.adams}@port.ac.uk*

Keywords: Cloud Computing, Simulation as a Service, Software as a Service (SaaS).

Abstract: This paper reports on a case study that was conducted on a large scale cloud service project that moved scientific simulation software to the cloud, one that used sensitive data. The study aimed to explore the challenges and practicalities of initiating and evaluating simulation as a cloud service. Action research was used to examine the nuances throughout the project as the service was moved from on-premise into a public cloud, lasting over one year from start to finish. The study was able to identify some emergent issues affecting initiation, technical security challenges and the evaluation of a significant change in a critical applications provisioning model.

# 1 INTRODUCTION

During the last 20 years there has been a continuing trend towards IT industrialisation. This has resulted in IT services becoming repeatable and usable by a wide range of customers and service providers. This is because of the increasing commoditization of technologies, virtualization and the rise of service-oriented software architectures, along with the dramatic growth in use of the Internet. These factors constitute the basis of a discontinuity that offers opportunities to shape the relationship between those who consume and those who provide IT services. The discontinuity implies that the ability to deliver specialized services in IT can now be paired with the ability to deliver those services in an industrialized and pervasive way. The reality of this implication is that users of IT services can focus on the business capability of what the services provide, rather than how the services are implemented or hosted. Similar in nature to how utility companies sell power on demand to subscribers, IT services can now easily be delivered an provisioned as a contractual service. This is not a new concept, but it does represent a different model from the licensed-based, on-premises models that have traditionally dominated the IT industry.

Cloud services provide a new way of delivering computing resources. Several types of cloud computing platforms exist, of which the main types are public, private and hybrid. Public clouds are normally offered by commercial organisations that provide access for a fee. Private clouds exist within are contained within a specific organisation and typically are not available for outside use. Hybrid clouds are a mixture of private and public clouds with the typical setup being that of a private cloud that has the ability to call upon additional resources from a public cloud (Chang, 2014).

The main advantage of cloud computing is the ability of equilibrating the access to computing resources for all types of businesses, regardless their dimensions and investment capabilities. These advantages include cost efficiency, scalability, concentration, security and accessibility with a further list below.

This paper outlines the overview, key issues and themes that emerged in a study of a large scale project within a mid-sized multinational company that ran a pilot to provision a scientific simulation software package via a public cloud.

# 2 RESEARCH METHODOLOGY

The research was conducted via a case study, taking an action research approach which used an iterative approach to collecting and analysing data. The benefits of an action research approach are that it focuses on generating solutions to practical problems and empowers the researcher to engage with the research and subsequent implementation activities (Mayer, 2000). A typical action research methodology takes a five step approach, as follows:

- Step 1: Identify the Problem
- Step 2: Devise a Plan
- Step 3: Act to Implement a Plan
- Step 4: Observe
- Step 5: Reflect and Share

Using this methodology, the approach starts with identifying the problem, which in this case was to determine if the simulation software was able to run via a cloud service. The second step was then to devise a plan around the migration of the service to the cloud and then test the success criteria. The next step was to execute the plan and implement the service, via a cloud provisioning model. This is the part of the approach where the action research is taking place via an iterative approach. After the plan was implemented, the researcher would observe how the service was or was not working. Once the researcher has had time to observe the situation then the entire process of action research was reflected upon, and at times the whole research approach may start over again (McCallister, 2011).

For this research, the researcher was a participant observer who was present for: top management meetings; from the inception of project to start-up; to designing the service; all the way throughout the whole project, till the end state of deciding if the service would be provisioned via the cloud. This access provided rare insight into what goes on in a multinational organisation during a large scale cloud service project. Along with access to top management meetings, the researcher had access to the critical role of the Head of IT strategy & projects, which was the primary role for orchestrating one of the biggest cloud pilots within the industry. The research itself used an academic approach to a real-world case study.

During the observation and reflection stages of the action research approach, mixed methods were used to evaluate the success of the project. These included quantitative methods that were used to determine the technical success criteria. These methods looked at indicating whether the data would be consistent before migrating to the cloud and then in determining the run-time performance of the simulations within a cloud environment.

The researcher used a mixed method of both qualitative and quantitative data, in the form of surveys. These were distributed to twenty four employees of both technical and business staff to find insight into trends that occur and organisational challenges. The use of a mixed method helped back up one set of findings from one method of data collection underpinned by one methodology, with another different method underpinned by another methodology. The researcher designed a series of survey questionnaires that included both boolean and open-ended questions, so that the resulting data would be both qualitative and quantitative. Qualitative data was used and analysed in the following approach. Questionnaires with open-ended questions were sent to twenty four pre-selected participants, coming from a wide range of both technical and business staff. The Questionnaires were distributed electronically via an online survey tool, with replies sorted and trends were identified to find commonalities. Upon the initial analysis another set of quantitative questionnaires was distributed to further investigate the findings and commonalities. The decision on the selection of interviewees was determined via a deductive approach to responses.

Examples of the specific questions that were used in the survey are listed below.

- Was the project a success
- Is cloud a viable provision model for scientific applications
- Is cloud scalable to run simulations
- Was the organization ready for cloud
- Does the organization need to introduce new processes for the adoption of cloud

Examples of the specific open-ended questions that were asked are listed below.

- What emerging themes were identified
- What key issues that were identified
- What was the impact of key issues

# 3 CASE STUDY

The case study is based on a midsized international company with a headquarters in Europe with operations in Europe, the Middle-east and Africa. The company has approximately 5,000 employees located over seven countries with revenues consistently averaging between $8-10 billion dollars and has a corporate culture that promotes innovation.

The company was exploring the possibility of migrating scientific simulation software that has significant computing and storage requirements into a cloud based HPC environment service delivery model. The benefits of moving from on-premise to a cloud provisioning model were that it would enable the scientific community within the company to flexibly increase compute via a cost effective, on-demand, pay-per use model (Jackson et al, 2012).

If successful, this new capability would enable the company to compete with much larger companies who had the capital to invest in the development and maintenance of large scale on-premise super-computing environments. A project was conducted to do a formal evaluation of migrating the service to the cloud and to determine if the concept was feasible from a technical and economic perspective, before a decision to invest further into a cloud provisioning service model was decided. The project was a multimillion dollar project, lasting over a year that consisted of a five person project team, with twelve other stakeholders from IT and outside IT whom were involved in the project.

## 3.1 Problem Statement

Scientists within the organisation within this case study were being challenged with a need for superior simulation modelling, as both the supply of information and the sophistication of quantitative techniques increases. The organisation invested heavily in technology that providing a vastly higher resolution of raw data, generating unprecedented volumes of data. All this additional data enables finer-scale simulation, as the geo-cellular models they simulated burst through the 10 and 100 million cell thresholds. As impressive as these advances were, they only represent more granular approaches to traditional modeling methods.

As the models increase in size, the organisation requires significantly more computing resources to run, given the increasing complexity with detailed models needing to run hundreds of times to quantify uncertainties and define the risks. The growth in demand for high performance computing was exceeding the supply from vendors. This results in the organisations' science community needing to limit simulations due to computing capacity, and was the driving force in running a proof of concept to explore new sources of high-performance computing capacity via a cloud provisioning model.

With the organisation investing in the project, it would need to determine real practical questions in relation to simulation as a service such as:

- Will the project be successful
- Is cloud a viable model for scientific applications
- Is cloud scalable for simulations
- Is the organisation ready for cloud
- Will new processes need to be implemented for cloud
- What themes would emerge
- What key issues would be faced

## 3.2 Plan

The project plan was drafted during the development of the business case. The project was broken down into six milestones, with each milestone having an estimated time. The milestones consisted of having: a signed off business case; a contract agreement with vendors; a high and low level design; the implementation; testing and finally an analysis conducted on the results leading to the project findings. Overall the initial estimation of the project plan was that it would take around five months to complete, however the actual time for the project was fifteen months. Significant challenges were faced in almost every step of the project plan.

Business Case: This stage took four months instead of an expected one month, as the biggest delay was in getting approval and buy-in for the business case, with the critical element revolving around the way that the sensitive data used within the simulations would be protected.

Contract: Similar to attaining business case approval, getting a contract signed with the software vendor took four months instead of the expected one month. This was because neither the organisation nor the software vendor could come to an agreement around intellectual property rights. In the end the situation was resolved as both parties agreed to waive rights.

Design & Implementation: The design took twice as long as expected, due to stringent internal information and data security requirements. This impacted the implementation as the vendor software required using a physical license dongle. As the project took an action research approach, the design and implementation phases took an iterative approach and required reworking of the designs during the implementation phase.

Testing of the service was approximately three

Table I: Milestones.

| Milestone | Timelines | |
|---|---|---|
| | *Estimated* | *Actual* |
| Business Case | 1 Month | 4 Months |
| Contract Agreement | 1 Month | 4 Months |
| Design | 6 Weeks | 3 Month |
| Implementation | 1 Week | 1 Month |
| Testing | 3 Months | 3 Months |
| Findings | 1 Month | 2 Months |

months the same time as initially estimated.

Findings: The findings took two months instead of the estimated one month, as the key themes and issues, led to some insightful findings, which required further investigation into finding commonalities and in conducting in depth interviews to validate the findings.

## 3.3 Design

The design was developed with the guiding principle that the architecture would need to be secure, lean and agile. This is because it was hypothesised that it would drive efficiencies and reliability through an elastic architecture that could dynamically scale up or down compute clusters as needed. The objective of the design was to simulate a real-life corporate network within a cloud scenario, so that the cloud service would be almost identical to the on-premise service, so that data could easily be moved in and out. As such a Virtual Private Cloud-VPC in the Amazon European datacentre was setup to act as the “corporate network”. The next step was to create a VPC in an Amazon US datacentre to act as the “cloud network”. The connections between the installations were facilitated through the use of OpenVPN which was installed on standard Linux Amazon Machine Images.

A major design requirement for running simulations in the cloud involves how to transfer large datasets between the corporate network and the cloud environment. To achieve this, a cloud network attached storage-NAS server was provisioned in the cloud, with a virtual device in the Amazon cloud configured to acts as a NAS front end to Amazon’s object based data cloud, simple storage service-S3. Due to the large storage requirements; there was a need for a common internet file system, which is a standard way that computer users share files across corporate intranets and the internet, with a network file system interface. This design is commonly known as a cloud storage security gateway system and is considered a secure way for encrypting and decrypting data as it is either uploaded or downloaded via Amazon’s S3 by examining the consistency of the contents and preventing data tampering (Wang et al 2013).

The next design step was to create the Simulator software head node in the cloud. This node would be static and would be were the simulation jobs would be submitted to and then run in a dynamically-created compute cluster. The head node rand on a red hat enterprise linux node running on an Amazon C3.xlarge size server. This server was selected as it had the required compute capacity need for the simulations along with having a solid state drive and a 10GB network interface. The design choice for using a solid state drive is that they offer higher performance compared to traditional storage devices and are needed in HPC systems, especially those with a growing demand of supporting big data applications (Chen et al, 2013).

A major security challenge was in the need to connect the simulation software’s physical Universal Serial Bus-USB license dongle to a virtual server. This was resolved via the use of a USB network device server being placed within a de-militarised zone-DMZ within the corporate network. This enabled the mapping of a USB port to a virtual server over the network. The USB port on the device server was then mapped to the simulator license server in the DMZ and was configured with a public IP. The DMZ was also configured to allow traffic between Amazon and the license server.

Figure 1 depicts the conceptual design which indicates the three main networks within the setup, the Amazon cloud, the virtual office and the DMZ and the major components within each of these networks.

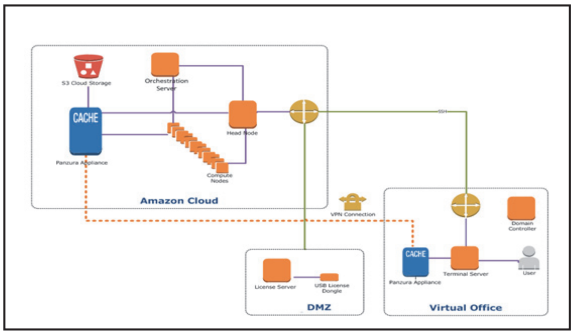


Figure 1: Conceptual Design.

A key component in the design was the NAS secure storage gateway (which also acted as a cache). Figure 2 depicts the dataflow from the simulation cloud which was within the amazon cloud. The data was de/encrypted as it passed through the cloud NAS and resided in the amazon storage, until it was to the NAS in the virtual office cloud, where it was then de/encrypted and passed into the main data source.

Along with the data flow there was a need to have a connection into the corporate DMZ as the simulation license server needing to be on the same subnet as the USB device server, thus not enabling the license server to be placed in the Amazon Cloud. The detailed design for this is depicted in Figure 3.

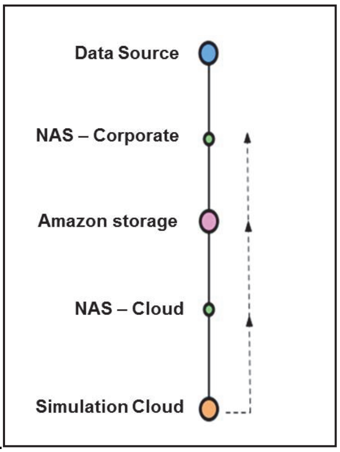


Figure 2: Data Flow.

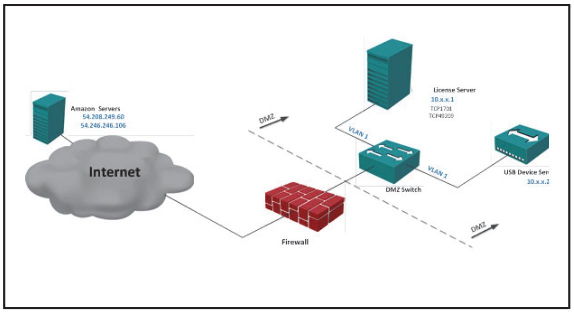


Figure 3: DMZ Design.

## 3.4 Technical Validation

Before the simulation cloud service was provisioned, an on-premise validation test was needed to ensure that simulations would run with the same accuracy regardless of the technology provisioning model. The approach taken for testing was to test out the characteristics with different stakeholders, to ensure requirements from users and those supporting the simulation software were properly gathered. The guidance provided was that the validation would need to ensure that the business characteristics, such as the need to ensure consistency when running simulation on the scalable dataset. The technical characteristics were that cpu performance scalability would need to be performed.

To achieve this, four cases were run on a workstation and then moved to the companies on-premise cluster which had a maximum of 8 core cpu's. The test ran the cases on the on-premise cluster using one, four and 8 cpu's to ensure consistency. The results from this indicated that moving the simulation jobs, did not have an impact on the jobs and once this was successful the simulation jobs could be migrated to the cloud service. Figure 4 indicates that the four cases demonstrated identical results for production rate and cumulative production for the duration of the field history as expected:

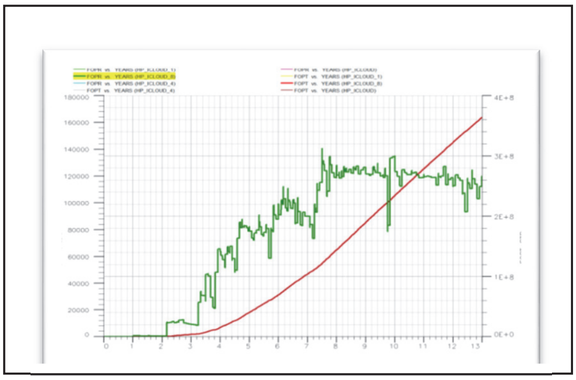

Figure 4: Simulation Validation.

The case with a single CPU on a workstation was completed in 20.9 hours, while the case on the on-premise cluster was completed in 19.6 hrs. The run time for the on-premise cluster with four and eight 8 CPUs ran at 7.7 and 5.7 hours respectively. The in-house cluster setup at the time of this study did not allow job executions on more than 8 CPUs. Figure 5 demonstrates the relationship between number of CPUs and wall clock time.

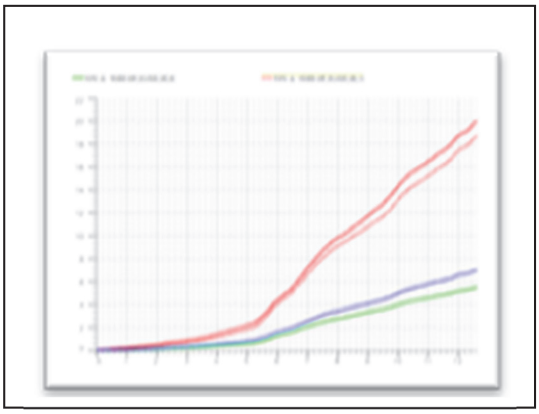

Figure 5: CPU Performance Validation.

With the run simulations being validated and with the performance of the on-premise service being measured, the next step was to compare this against the results of the performance within the cloud provisioned serviced.

Figure 6 shows the results, and indicates that the wall clock time decreases as more CPUs were added, for both calculations for the on-premise and cloud service. It was observed that on-premise calculations stagnated at more than 4 CPUs, resulting in sublinear scaling. This is contrary to the performance via the cloud service, which observed close to linear scaling. Assuming that this linear scaling persists when adding more than 8 CPUs, extrapolating from this observation, it was hypothesised that for larger jobs, the performance of the cloud provisioned service would significantly increase compared to what can be achieved on-premise. This analysis is not exhaustive, but was severely limited by the size of the on-premise service having only 8 CPUs.

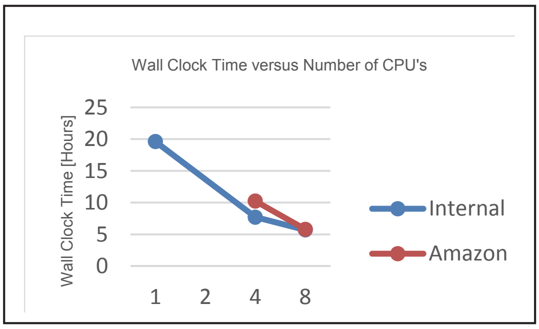


Figure 6: On-premise & Cloud Performance Validation.

## 3.5 Data Collection & Analysis

Once the system was implemented and had passed the technical validation aspect, twenty four individuals involved in the project completed a questionnaire, with 15 coming from IT and 9 from outside.

The 24 interviewees were asked: whether they thought the project was a success; whether cloud was a viable service model for scientific applications and if cloud was a scalable to run simulations; if the organisation was ready for cloud and if the organisation needed to introduce new processes for cloud. For data analysis purposes yes equating to a score of 1, while a no equated to a score of 0. The response and standard deviation were calculated as indicated in Table 1. Not all questions had an input, as respondents preferred not to say and the breakdown is as follows:

- Question 2: 4 participants declined
- Question 3: 1 participant declined
- Question 4: 1 participant declined

The following key insights were indicated

- 80% indicated that the project was a success.
- 71% indicating that cloud was a viable service model for scientific applications.
- 79% indicating that cloud is a scalable.
- Surprisingly only 25% indicated that the organisation was ready for cloud.
- 71% indicating that the organisation needed to implement new processes for cloud.

Table 1: Success Criteria.

| Question | Response | | Mean | Standard Deviation |
|---|---|---|---|---|
| | Yes & (Total %) | No | | |
| 1. Was the project a success | 20 (80%) | 4 | **0.833** | **0.381** |
| 2. Is cloud viable for scientific applications | 17 (71%) | 3 | **0.850** | **0.366** |
| 3. Is cloud scalable | 19 (79%) | 4 | **0.826** | **0.388** |
| 4. Is the organisation ready for cloud | 6 (25%) | 17 | **0.261** | **0.449** |
| 5. The organisation needs new processes for cloud adoption | 17 (71%) | 7 | **0.708** | **0.464** |

The same interviewees were then asked their opinion of the three major themes which emerged during the life of the project.

The following key insights were indicated

- 62.5% indicated that politics were a prevalent theme, 38% indicated that politics was the first theme, 25% indicated it was the second theme.
- Innovation was second with 45.8%, 25% for the first theme, 8% for the second and 13% for the third.
- Security at 33.3%, 13% for the first, 17% for the second and 4% for the third.

Other key themes included vendor solutions, intellectual property rights, a lack of the required skills, internal processes, business value and change.

Interviewees were then asked their opinion of the three major issues which emerged during the life of the project.

The following key insights were indicated

- Politics was again the highest at 66.7% which was aligned with the responses from the emerging themes, with 42% of respondents indicating it was the first issues while 25% indicated it was the second issues and similar to themes zero respondents indicated that it was the third issue.

Table 2: Emerging Themes.

| Theme | First | Second | Third | Inclusion Total % |
|---|---|---|---|---|
| Business Value | 1 | 2 | 0 | 12.5% |
| Change | 0 | 2 | 1 | 12.5% |
| Innovation | 6 | 2 | 3 | 45.8% |
| Intellectual Property Rights | 2 | 2 | 1 | 20.8% |
| Politics | 9 | 6 | 0 | 62.5% |
| Processes | 0 | 2 | 1 | 12.5% |
| Security | 3 | 4 | 1 | 33.3% |
| Skills | 0 | 1 | 2 | 12.5% |
| Vendor Solutions | 3 | 1 | 1 | 20.8% |

- Project Management was second with 37.5%, 21% for the first theme and 16.5% for the second.
- Contracts and Processes were tied for the third issue both with a response of 20.8% with contracts had a response of 12.5% for being the first issue and 4.15% for being the second and third issues.
- 4.15% identified processes as being the first issue and 16.65% for being the third issue.

Other key issues included capability of staff, lack of clear KPI's to measure and information Security.

Table 3: Key Issues.

| Key Issue | First | Second | Third | Total | Total % |
|---|---|---|---|---|---|
| Capability | 2 | 2 | 0 | 4 | 16.7% |
| Contracts | 3 | 1 | 1 | 5 | 20.8% |
| KPI's | 0 | 1 | 1 | 2 | 8.3% |
| Politics | 10 | 6 | 0 | 16 | 66.7% |
| Processes | 1 | 0 | 4 | 5 | 20.8% |
| Project Management | 5 | 4 | 0 | 9 | 37.5% |
| Security | 0 | 0 | 1 | 1 | 4.2% |

# 4 DISCUSSION

The pilot was initially resisted by internal members of the IT department that were responsible for supporting the simulation software. This delayed the approval of the business case. Design challenges arose during the design and implementation stages due to stringent internal security requirements.

This was a surprising finding as the respondents did not indicate information security to be an issue, this along with the data captured from the surveys indicated that politics was the pervasive theme and key issue of moving to the cloud. Considering that the project was determined a success, but that the organisation was not ready and would need to introduce new processes to support cloud. This leads to a finding of how organisation behavior and the perception of trust in security pose a real threat to the adoption of cloud. This is indicated by 2009 Gartner survey with indicates that politics is a challenge of cloud service adoption (Gartner, 2009).

Resistance to change is a normal human response as employees seek to translate the change to a personal context, which can be greatly magnified by fear of the unknown (Berube, 2012). If internal policies and security concerns are a significant challenge in cloud service adoption, then building a maturity assessment at the start of the project to understand the organisations culture, internal processes would clearly assist in migration to a cloud service, and the delivery of a cloud service project could then plan accordingly to further train and develop staff on the impact of cloud, before any implementation would occur. This activity was not included in the project, but the insights received after the fact, could help in any future cloud project by being able to measure and mitigate the risks of the cloud.

The data which was collected using an action research approach indicates that a lot is still unknown about dealing with challenges during the initiation stages of a cloud project were the realization that the change from one modal of working to another different modal has a significant impact on the success of a project. Though the project was validated as being a success, several emergent themes impacted the adoption. One significant emergent theme from the research was that the organisation did not have the appropriate internal policies.

The research shows that the evaluation and adoption of simulation as a service project, which is a considerable change to business practices, will likely involve more than technical performance and business improvements: It will also need to consider the wider political fault-lines of issues that would impact the acceptance from various stakeholders.

# 5 CONCLUSIONS

Cloud computing is maturing, but there is still a lot that remains uncertain for its adoption within enterprises, such as the organizational changes brought about by cloud computing. Cloud services that support simulation via a HPC environment are attracting more attention in literature, in big business and in governments.

This paper has reported on research exploring the practicalities of conducting a significant simulation as a service project within a large company. This paper further explores the practicalities and contexts the issues of applying cloud to larger compute processing needs

This is one of the few works that covers simulation as a service in a real life project.

The research involved an iterative methodology based upon an action research methodology and covered all the stages of the project from creation to evaluation. The pilot project and research focused on evaluating the possibility of running simulation as a service which leverage a cloud infrastructure to address the HPC needs of the multinational company using a range of criteria, including technical capability and wider business case.

It was a successful project and the insights taken from this work can further be used to make informed decisions about moving simulations to the cloud. Lessons learned from this would be that doing a proof of concept is a good method.

The data which was collected using an action research approach indicates that a lot is still unknown about dealing with challenges during the initiation stages of a cloud project were the realization that the change from one modal of working to another different modal has a significant impact on the success of a project. Though the project was validated as being a success, several emergent themes impacted the adoption. One significant emergent theme from the research was that the organisation did not have the appropriate internal policies

The research shows that the evaluation and adoption of simulation as a service project, which is a considerable change to business practices, will likely involve more than technical performance and business improvements: It will also need to consider the wider political fault-lines of issues that would impact the acceptance from various stakeholders.

Developers and project managers can take practical guidelines from this paper that can be used to make informed decisions about moving simulations to the cloud. These examples are in the form of design, validation steps but more importantly the need to get feedback from different stakeholders before starting a project and the need to have an understanding of the potential political impact may occur similar to this project in terms of project delays and in design requirements. Key contributions to knowledge are that even if the project is successful, the organisation may not be ready for cloud and that new processes would need to be developed to operate via a cloud provisioning model. For considerably sized projects of this type the recommendation is to run a pilot first and to plan and execute the development of internal processes that are required to enable the organisation to be cloud ready.

# REFERENCES

H. F. Wang, L. J. Wang, Pingjian Institute of Information Engineering, Beijing, China*, Chinese Academy of Sciences*, 2013.

N. Agrawal, V. Prabhakaran, T. Wobber, J.D. Davis, M. Manasse, R. Panigrahy, *Design Tradeoffs for SSD Performance*, Microsoft Research, Silicon Valley.

J. Chen, Using pattern-models to guide SSD deployment for Big Data applications in HPC systems, *2013 IEEE International Conference on Big Data,* 2013.

V. Chang, The Business Intelligence As a Service in the Cloud, 37, 512-534 ed. , *Future Generation Computer Systems*, 2014.

C. Vecchiola, S. Pandey, R. Buyya, High-Performance Cloud Computing: A View of Scientific Applications, Pervasive Systems, Algorithms, and Networks, *10th International Symposium on Pervasive Systems, and Networks*, 2009.

V. Chang C S. Li, D. De Roure, G. Wills, R. Walters, C. Chee, The Financial Clouds Review, 1 (2). pp. 41-63. ISSN 2156-1834, eISSN 2156-1826. ed. , *International Journal of Cloud Applications and Computing*, 2011.

M. Armbrust, A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, M. Zaharia, *A View of Cloud Computing*, Vol. 53 No. 4 ed. , ACM, 2010.

J. Meyer, Using qualitative methods in health related action research, 320: 178-181 ed. , *British Medical Journal,* 2000.

J. McCallister, *Contemporary Social Work Issues*, MSW, 2011.

K. R. Jackson, L. Ramakrishnan, K. Muriki, S. Canon, Performance Analysis of High Performance Computing Applications on the Amazon Web Services Cloud, *IEEE Second International Conference on Cloud Computing Technology and Science*, 2010.

Gartner, Managing and Deploying Clouds, *Gartner 2009 Datacenter Conference,* 2009.

D. Berube, *Resistance to Change is a Good Thing*, Life Cycle Engineering, 2012.

A. Khajeh-Hosseini, I. Sommerville, I. Sriram, Research Challenges for Enterprise Cloud Computing, *1st ACM Symposium on Cloud Computing,* 2010.