Trust Challenges in a High Performance Cloud Computing Project

The emergence of the political cloud: An exploration into fault lines of organisational dynamics into the perception of information and data security risks of big projects when moving to the cloud

Morgan Eldred
School of Computing
University of Portsmouth
Portsmouth, United Kingdom
Morgan.Eldred@myport.ac.uk

Dr. Carl Adams
School of Computing
University of Portsmouth
Portsmouth, United Kingdom
Carl.Adams@port.ac.uk

Dr. Alice Good
School of Computing
University of Portsmouth
Portsmouth, United Kingdom
Alice.Good@port.ac.uk

Abstract— *The literature on cloud computing has been dominated by examples and issues from small to mid-size companies. This paper reports on a large scale cloud pilot project within the petrochemical industry, perhaps one of the biggest examples of cloud computing pushing the boundaries of what you can do with sensitive data. The study aimed to explore the challenges and practicalities of initiating and evaluating cloud projects. Action research is used to examine the nuances throughout a million dollar cloud pilot which lasted over one year from start to finish. The study was able to identify some emergent issues affecting initiation, implementation, technical security challenges and evaluation of a significant change to the security provision of critical data within a large international company affecting many stakeholder groups. One emergent theme was that of the 'political cloud' which was represented by a clash between organisational behavior, the perception of security and internal politics within an organisation; much like the fault lines between tectonic plates. The paper hopes to make contribution by modeling some of the complexities of cloud computing data security and trust challenge by providing insights on the trust and protection of data within companies, particularly large international companies.*

Keywords; cloud computing; political cloud; cloud information security; data security; cloud trust

I. INTRODUCTION

Cloud computing is not viewed as a new technology, but more as 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 Private clouds 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 [1]. High performance cloud computing environments are computer clusters located in the cloud which address complex computational requirements, support applications with significant processing time requirements, or require processing of significant amounts of data [2].

Security has been identified as a concern for some organisations in the adoption of cloud, with privacy and data ownership amongst the key factors for organisations deciding not to move to the cloud [3], with the common reason for security being the skeptical question "who would trust their essential data out there somewhere?" [4].

This paper outlines the overview, key issues and themes that emerged in a study of a large scale project within a mid-sized multinational petrochemical company that ran a pilot of a standard upstream simulation software package that requires significant computation power and storage requirements in a public high performance cloud computing environment running an industry standard upstream simulation.

Certain companies and countries involved within the petrochemical industry have strict rules about the movement of seismic, well and reservoir data outside of their jurisdiction, as data around their natural hydrocarbon resources are used to assure the growth, govern-ability, control, independence and sovereignty of the industry. From a legislation perspective this data is known as the national data repository [5] as it is accepted as being a national asset. This paper focuses on the insights of trust and security of data as there is significant fear and uncertainty about what data can be kept where within the petrochemical industry.

II. RESEARCH

A. Methodology

The research method taken was an action research methodology which used an iterative approach to collecting and analysing data. The benefit of this approach is that it focuses on generating solutions to practical problems and the ability to empower researchers to engage with research and the subsequent implementation activities [6]. The researcher was a participant observer who was present for top management meetings, from the initiation throughout the whole project. This provided a very rare window into what goes on in a multinational organisation and had access to the critical role of Head IT strategy & projects who was orchestrating setting up one of the biggest cloud pilots within the industry. The research itself used an academic approach to a real-world case study.

B. Design

The research design used a mix of deductive, quantitative, qualitative and inductive approach as it was hypothesised that the use of multiple approaches would help in overcoming the shortcoming of using a single methodology. The pilot used quantitative methods to determine the technical success of the project and interview and questions where conducted among a sample frame of 24 members of the organisation to understand and to identify key themes and issues that arose. The list of interviewee's where those within the pilot team, those whom where stakeholders in the pilot and others whom where interested in the results of the pilot. A large portion of the data collected was on the perceptions around key themes and issues and focused deeper into the perception of technical risk, and trust in the cloud.

III. CASE STUDY

The case study was on a midsized international Oil & Gas company with a headquarters in Europe, Operations in Europe, the Middle-east and Africa. The company has some 5,000 employees located over 7 countries with revenues over 10$ Billion and has a corporate culture around innovation.

They were exploring the possibility of cloud as a competitive advantage in their ability to more accurately determine the characteristics of the oilfield and extraction techniques for hydrocarbons by moving to a probabilistic scientific approach that requires a significant amount of computational power and storage capabilities. High Performance Cloud Computing (HPCC) was identifying as possibly being able to conduct this, as the scientific computing community has shown increasing interest in exploring cloud computing due to the on-demand, pay-as-you-go model which creates a flexible and cost-effective means to access compute resources [7].

If successful this new capability would allow the company to compete with much larger companies who had the capital to investment in the development and maintenance of large scale in-house computing environments. The industry as a whole is changing and Oil was becoming more expensive to explore and extract due to depleting reserves of easy oil and the emergence of unconventional oil which requires higher investment in the research and development of technology to reduce the cost of extraction techniques and distribution [8].

A pilot was conducted to do a formal evaluation and determine if the concept was feasible from a technical and economic perspective before a decision to invest further into this method was decided. Due to the sensitivity of the data, a publically available data set from the North Sea was used.

A. Details of Pilot

The pilot used Amazon Web Services (AWS) as a public HPCC, with an industry standard upstream simulation application being used. Multiple simulation cases where launched to the cloud while transferring large amounts of data in the terabytes between simulated offices. The pilot involved designing and developing a secure lean agile technology model, which was software vendor agnostic, as it was hypothesised that this would drive efficiencies and reliability by being able to dynamically scale up or down computing clusters depending on needs.

For security purposes four sites where configured as a virtual private cloud with a virtual private network linking them to the field office of the remote cluster, a cloud storage controller was used to encrypt and convert the AWS Simple Storage Service (S3) object storage into a Network Appliance Storage (NAS) server sitting on a local network with a Common Internet File Storage interface. The cloud storage controller also acted as a cache for S3 storage, which sped up the uploading and downloading of files. The head node in the virtual upstream simulator cluster was a C3.xlarge instance on AWS, which had a 10 Gb/sec Ethernet controller, two 40 gigabyte solid state drives, and four virtual Central Processing Units (CPUs). The compute nodes had eight virtual CPUs, and the test ran on eight nodes for a total of 64 cores.

The overall cost of the software licenses was valued at over $1.15 million, while the cloud costs of three months being just over $10,000. However the software licenses where provided on a trial basis by the vendor, whom was interested to understand if it was possible to run its scientific application via the public cloud.

TABLE I. PILOT COSTS

Project Costs		
Software	***Quantity***	***Cost***
Core Simulator Licenses	10	$563,190
Parallel Licenses	64	$609,777
Cloud Server Infrastructure	3 Months	$5,000
Cluster Servers	2,000 hours	$1,000

A major security challenge was how to connect the physical Universal Serial Bus (USB) license dongle to a virtual server. This was resolved via the use of a USB network device server placed within the de-militarised zone (DMZ). This enables the mapping of the USB port to a virtual server over the network. The USB port on the device server was mapped to "Upstream Simulator License Server" in DMZ and configured with a public IP. It was also configured in the firewall to allow traffic between Amazon and the license server in the DMZ. One limitation that was encountered was that the virtual upstream simulation license server needed to be on the same subnet as the USB device server, so it was not possible to place the license server in the Amazon Cloud.

The requirements for the license server to be placed in a DMZ was due to the organisation having recently introduced new information security policies that had a zero tolerance area where it was strictly prohibited to connect non company equipment to the internal network and that this must happen in an isolated DMZ specifically build for that purpose.

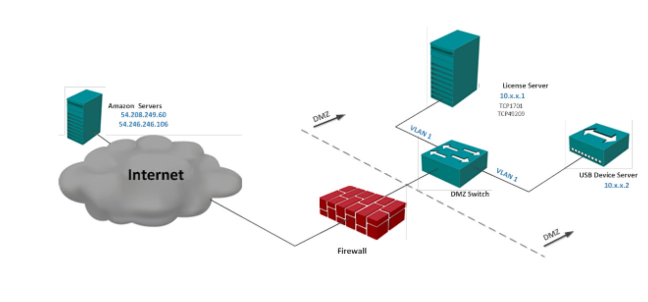

Figure 1. Security Design

B. *Pilot Team*

The pilot team consisted of the following:

- Project Manager
- Solution Architect
- Sr. Reservoir Engineer
- Reservoir Engineer
- Technical Application Engineer

Other stake holders included:

- Head of Research
- Reservoir Research Team Lead
- Legal Associate
- Head of IT
- Head of IT Strategy & Projects
- Head of Information Security
- Head of Infrastructure
- Head of Technical Applications
- Head of Data Management
- Head of Networks
- Head of Servers
- Network & Server Engineers

C. *Pilot Plan*

The initial expectations of the pilot where that it would take a total of 5 months to complete however the total time taken was around 15 months. The biggest delays where around getting buy-in on the business case, with the critical element being assurance that data would be protected along with the signing of contracts with the vendors; which required agreement on the intellectual property rights. Other significant delays where around the high and low level designs which took much longer than expected due to the stringent information and data security requirements and the challenges around using physical license dongles.

TABLE II. PILOT MILESTONES

Phases	**Details**		
	Description	***Expected Time***	***Actual Time***
1	Business Case	1 Month	4 Months
2	Contact Setup	1 Month	4 Months
3	High Level Design	2 Weeks	1 Month
4	Low Level Design	4 Weeks	2 Months
5	Implementation	1 Week	1 Month
6	Testing	3 Months	3 Months
7	Results Analysis	1 Month	2 Months

D. Pilot Methodology

The methodology that was developed by this pilot was an iterative one and this fits in with the research design of an iterative action research project with a clear connection to the action research that was conducted.

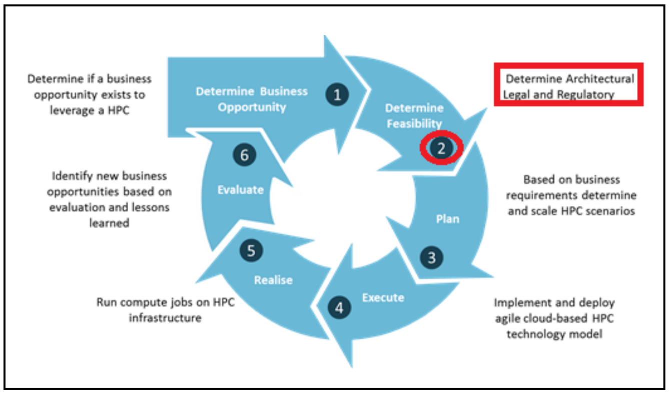
Determine if a business opportunity exists to leverage a HPC
Determine Business Opportunity
1
Determine Feasibility
2
Determine Architectural Legal and Regulatory
Based on business requirements determine and scale HPC scenarios
Plan
3
Execute
4
Implement and deploy agile cloud-based HPC technology model
Realise
5
Run compute jobs on HPC infrastructure
Evaluate
6
Identify new business opportunities based on evaluation and lessons learned

Figure 2. Pilot Methodology

TABLE III. METHODOLOGY STEPS

Steps	Details
1	Determine Business Opp
2	Determine Architecture, Legal & Regulatory Feasbility
3	Plan
4	Execute
5	Realise
6	Evaluate

A major contribution to the methodology was the inclusion of the feasibility step, which is the second step that occurs right after determining the business opportunity. This step had a governance aspect specifically looking at the architecture, including information security by determining what impact the business opportunity had on the current architecture in terms of information security and the technical landscape and if the opportunity was in compliance with information security guidelines. This would also include what the license restrictions where put forward by vendors and if they could be incorporated into the organisations security architecture guidelines without violation. The feasibility step also looked at the legal and regulatory feasibility by determining the security requirements for the data used and setup of the cloud according to region. If data privacy or national data repository requirements specify that data may not move between geographical borders, then this is a no-go for the solution.

E. Pilot Results

The findings of the project where that High Performance Cloud Computing Environment provide an opportunity to leverage technology enhancements and that while it is technically possible to securely run upstream simulations in the cloud, challenges exist with certain national data repository regulations on the movement of upstream data and with the licensing models of existing upstream simulation software, which can be overcome but do require non-ergonomic designs to ensure that stringent information security guidelines are applied.

The framework which was developed was a secure agile method that determines if the business opportunity is ready for the technical architecture, information security and data protection requirements for cloud computing.

IV. EMERGING THEMES & KEY ISSUES

Interviewees where asked what where the three major themes which emerged from this project. Surprisingly 62.5% indicated that politics where a prevalent theme, with 38% of respondents indicating it was the first theme while 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 followed by 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, the lack of skills, internal processes, business value and change.

TABLE IV. EMERGING THEMES IDENTIFIED

Theme	First	Second	Third	Total	Total %
Business Value	1	2	0	3	12.5%
Change	0	2	1	3	12.5%
Innovation	6	2	3	11	45.8%
Intellectual Property Rights	2	2	1	5	20.8%
Politics	9	6	0	15	62.5%
Processes	0	2	1	3	12.5%
Security	3	4	1	8	33.3%
Skills	0	1	2	3	12.5%
Vendor Solutions	3	1	1	5	20.8%

Interviewees where asked what where the three major issues that occurred during the pilot. 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. Project Management was second with 37.5%, 21% for the first theme and 16.5% for the second. Contracts and Processes where 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 issue, while 4.15% of respondents 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 V. KEY ISSUES IDENTIFIED

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%

TABLE VI. EXPECTATIONS, FEELINGS & LEARNINGS

Theme	Low (1)	Medium (2)	High (3)	Average %
Expectations	13	8	3	1.58
Feeling at Start	10	4	10	2.0
Feeling at End	5	10	9	2.17
Organisation Politics	3	2	19	2.67
Impact of People on Politics	8	6	10	2.08
Impact of Trust in Security on Politics	5	9	10	2.21
Impact of Technical Risk on Politics	15	7	2	1.46

Interviewees where then asked to rank on a scale of low, medium, high their expectations, how they felt at the start and at the end of the pilot, the rating of organisational politics, and the impact of people, trust in security and technology risk measured against politics. The main outputs from consolidating this data was how organisational politics where the highest and that the biggest impact on this was down to the trust in security which was high, followed by people which was medium and lastly the technology risk having a low score.

V. POLITICAL CLOUD

The pilot was initially resisted by internal members of the IT department that where responsible for supporting the upstream simulation software. This delayed the approval of the business case and finally sign-off was given by executive management. Constant issues arouse due to specialists in the software not being available to work on the project; this halted progress until a mandate was given to give full support to the project. During the installation a firewall change request was raised and approved to allow access from the cloud into the corporate network and issues where raised about the motivations and ethics of the approver, however new information security policies had just been introduced and had not properly been communicated to all approvers.

Respondents when asked did not indicate information security to be an issue as the upstream data was publically available and the technical risk was deemed to be very low. This along with the input from the interviewees which indicated politics as being the number one theme and key issue provides insight into the fault lines that exists with how organisation behavior and the perception of trust in security pose a real threat to the adoption of cloud.

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 [9]. This aligns with the insights provided by the study as respondents clearly indicated that the biggest impact on organisational politics in regards to people, trust in security and technology risk with the number one reason was the trust in security which was high even though they thought the technical risks where very low.

VI. CONCLUSION

High Performance Cloud Computing is attracting more attention in the literature, in big business and in governments. This paper has reported on research exploring the practicalities of conducting a significant pilot HPCC project within a large company. The pilot project consisting of multimillion dollar resources, lasting 15 months and pushing boundaries of applying cloud computing to big processing needs within a commercial environment.

The research involved an iterative methodology based upon the action research and covered all the stages of the HPCC pilot from creation to evaluation. The pilot project and research focused on evaluating the possibility of using cloud computing to address the high performance needs of the multinational company using a range of criteria, including technical capability and wider business case. The company is an innovator within its industry and open to technological change particularly that could add value to the company. The research captured many practical aspects and issues of applying a significant change to computing provision within a multinational company, such as information security, legal and regulatory compliance for data protection along with providing measures for better acquiring insights into the perception of trust in the cloud.

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; the security, legal and privacy issues that cloud computing raises [10]. This paper further explores the practicalities and contexts the issues of applying cloud to larger processing needs such as HPCC projects. 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 where the realisation that the change from one modal of working to another different modal has a significant impact on the success of a project. Though technically feasible several emergent themes impacted the adoption and perception of the pilot. One significant emergent theme from the research was the concept of the 'political cloud' from different stakeholder groups which impacted the pilot throughout most of the project. A particular aspect of the political cloud was the challenges around the perception of trust in security even if the technology is determined as being a low risk. The political cloud in the pilot acted as fault-lines of issues and consequently resistance to areas of change around a HPCC within this multinational company. The research shows that the evaluation and adoption of HPCC projects, with their considerable change to business practices, will likely involve more than technical performance and business improvements: It will also need to consider the wider political cloud and fault-lines of issues that would impact the acceptance from various stakeholders. Developers and project managers need to have an understanding of the potential political cloud within considerably sized projects, particularly with HPCC that are likely to involve multiple and diverse stakeholder groups and deal with sensitive data.

I. Future Work

This study provides a foundation from which further insights could be captured from future work on the success of cloud projects, by using the measures that captured the fault-lines on the perception of trust in security, i.e. 'the political cloud' and by incorporating the pilot methodology framework could produce a framework which would further advance the success of large scale cloud projects that use sensitive data in the cloud, by mitigating delays in project schedules or by ensuring that cloud projects do not get completely derailed by the emergence of the political cloud.

References

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

[2] 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

[3] 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.

[4] 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.

[5] Energistics, *National Data Repository*, http://www.energistics.org/regulatory/national-data-repository-ndr-work-group: Energistics, 2014.

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

[7] 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

[8] M. Noel, *The Promise of Unconventional Oil*, Edgeworth Economics, 2011.

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

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