



# **AFCEA Bethesda Chapter High Performance Cloud Computing Symposium**

## **White Papers**

**Wednesday, March 23, 2011**

---

## Dear AFCEA Bethesda Guest,

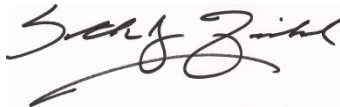
AFCEA Bethesda Chapter is pleased to present **a new and value-added component of our Cloud Computing symposium – namely a compilation of white papers** designed to support and enhance our event discussion. We have coordinated with knowledge and thought leaders across government, academia and industry to offer and capture their opinions and perspectives that span the topics addressed within this symposium, specifically:

- FedRAMP
- High Performance Computing in the Cloud
- Sharing Data in the Cloud

These papers were specifically written for this event. Our authors' intent is to educate, inform and evoke further discussion and contemplation, and as such, follow-on contact information is provided within each. Moreover, these papers are oriented towards the "next-phase" of cloud computing understanding, implementation and adoption, rather than focusing on Cloud 101 topics and subject matter. Our goal is to help take our AFCEA audience to another level in their cloud contemplations.

On behalf of the authors, their respective organizations, and the AFCEA High Performance Cloud Computing Symposium planning committees, we thank-you for the opportunity to share our experience and views, and offer the following white papers for your edification and reading enjoyment.

Warm regards,  
-Seth Finkel

A handwritten signature in black ink that reads "Seth Finkel".

## White Paper Synopsis and Table of Contents

Title	Author(s) / Organization	White Paper Summary	Page
“Magellan: A Testbed for Exploring Cloud Computing for Science”	Kathy Yelick, R. Shane Canon, Lavanya Ramakrishnan, Iwona Sakrejda - <b>NERSC, Lawrence Berkeley National Laboratory</b>	<b>High-Performance Computing in the Cloud</b> - Provides a status summary of Magellan, the DOE-funded program designed to explore how cloud computing can serve the needs of the scientific community. Amongst topics highlighted are applicable workloads and deployment models, initial performance findings and future considerations for adoption.	4
“NOAA’s Implementation and Use of High Performance Computing in the Cloud”	<b>The National Oceanic and Atmospheric Administration (NOAA)</b> , Office of the Chief Information Officer	<b>High-Performance Computing in the Cloud</b> - Provides an introduction to NOAA’s high-performance computing objective, current state and status, and future challenges to be addressed.	5
“DHS Cloud and Impediments to Sharing Data”	<b>Department of Homeland Security (DHS)</b> , Office of the Chief Information Officer, Enterprise Systems Development Office	<b>Data Sharing in the Cloud</b> – Provides an overview of the DHS private cloud and its level of certification and accreditation. Also highlights challenges and DHS perspectives related to sharing data.	6
“High Performance Cloud Computing and its’ Impact Within Academia”	Jimmy Lin – <b>University of Maryland</b>	<b>High-Performance Computing in the Cloud</b> - Provides an overview of the challenges within the academic community to educate the next generation of cloud-computing technologists and what next-generation problem-solving approaches and capabilities are now possible with cloud computing infrastructure.	7
“Agency Case Study per High Performance Cloud Computing”	Read Maloney – <b>Amazon Web Services</b>	<b>High-Performance Computing in the Cloud</b> - Provides a case study of NASA’s JPL Research and Training Studies	9
“An Introduction to FedRAMP”	Phil Horowitz - <b>Apptis</b>	<b>FedRAMP</b> – Provides an introduction to FedRAMP including its premise, value proposition, and proposed benefits to Agency and cloud solution providers alike.	10
“Cloud’s Great Challenge: Data Sovereignty & Sharing”	Phil Horowitz - <b>Apptis</b>	<b>Data Sharing in the Cloud</b> – Provides a summary of issues to include data sovereignty that are being considered and addressed by governments worldwide.	11
“Data Sharing in the Cloud Aided by the Service Request Lifecycle”	Lilac Schoenbeck & Dan Trevino - <b>BMC Software</b>	<b>Data Sharing in the Cloud</b> – Provides a summary of the elements that comprise the service request lifecycle and how effective management can lessen data sharing risks and concerns.	12
“CloudFirst and FedRAMP; “New Wine in Old Wineskins?”	<b>CSC</b> - Based on an essay by Ron Knode posted at <a href="http://www.csc.com/cloud/blog/">http://www.csc.com/cloud/blog/</a>	<b>FedRAMP</b> - Provides explanation of a premise that FedRAMP requires a re-visitation of agency controls and assumptions that are more with cloud computing structure and delivery conditions.	14
“Data Sovereignty and the Cloud”	<b>CSC</b> - <a href="http://www.csc.com/cloud/blog/">http://www.csc.com/cloud/blog/</a>	<b>Data Sharing in the Cloud</b> – Provides a summary of the issues pertaining to data sovereignty and the underlying inconsistency that exists between cloud ubiquity and governments’ requirement for defined boundaries and containment.	16
“High-Performance Computing: a Viable Cloud Workload?”	<b>Quest Software</b> , Public Sector <a href="http://www.quest.com/virtualization">http://www.quest.com/virtualization</a>	<b>High-Performance Computing in the Cloud</b> - Provides an argument for why HPC is an ideal candidate and solution for cloud-based processing and infrastructure models	18
“Data Sovereignty: the Biggest Impediment to Public Cloud?”	<b>Quest Software</b> , Public Sector <a href="http://www.quest.com/virtualization">http://www.quest.com/virtualization</a>	<b>Data Sharing in the Cloud</b> – Provides a proposed progression of how data sovereignty will be solved in cloud-based solutions, and what agencies can do in the meantime to satisfy their cloud-based IT requirements	20

# Magellan: A Testbed for Exploring Cloud Computing for Science

Kathy Yelick, R. Shane Canon, Lavanya Ramakrishnan, Iwona Sakrejda  
NERSC, Lawrence Berkeley National Laboratory

Cloud computing provides a new resource model where multiple virtual servers hosted in data centers are used by individuals or groups, usually through a *pay-as-you-go* model. Cloud computing promises greater economy of scales compared with private data centers since it extends consolidation across organizations. This new paradigm has already made an impact in the way many enterprises and Web 2.0 organizations address their computing needs. The promise of highly elastic, inexpensive computing cycles and customizable software environments has also started to attract the attention of science communities. However, the suitability of cloud computing for various scientific workloads is still poorly understood. The goal of Magellan Project is to inform Department of Energy's Office of Science and other scientific communities by exploring how cloud computing can address the computing needs of scientist.

**Project Goals and Testbed.** Magellan has been funded by the DOE Advanced Scientific Computing Research Program to investigate how cloud computing models can be used to serve the needs of midrange computing and future data-intensive computing workloads for the Office of Science. These workloads are typically not served by DOE data center facilities today. Some key questions the project will address are: What are the unique needs of typical scientific workloads and can these workloads run efficiently in today's cloud offerings? Are there new service models and programming models popularized in the cloud that can benefit scientific communities? What are the implications on security and other operational aspects of providing and using Cloud resources? What is the cost and energy efficiency of clouds?

To support this exploration, a distributed testbed infrastructure has been deployed at the Argonne Leadership Computing Facility (ALCF) and the National Energy Research Scientific Computing Facility (NERSC). At NERSC, the testbed consists of 1,440 Intel Nehalem quad-core processors (5,760 cores total) connected by a High-Performance InfiniBand interconnect, nearly 1 PB of storage, and 8 TB of high-performance flash storage.

**Early Results.** Early results of the project demonstrate some of the benefits and challenges in employing clouds for scientific computations. Scientific pipelines with complex dependencies can benefit from the flexibility offered by virtualization. The Magellan testbed has been utilized by scientists from a variety of disciplines. These users are exploring various models of cloud computing including MapReduce using Hadoop and virtual instances using Eucalyptus. High-energy physics and bioinformatics projects have been heavily engaged given the natural fit of their workloads to the underlying models. However, even in these cases, a significant amount of effort has been required in order to port applications to these framework.

Typical commercial cloud offerings built on virtualized environments connected by gigabit ethernet results in a signif-

icant performance degradation for even small scale applications [1]. Performance decreases of more than 40x were measured for some scientific applications. These performance differences also impact the cost analysis as commercial providers charge per CPU hour. Our analysis shows that much of the slow down comes from the choice of interconnect typically employed in these systems. This is consistent with the improvements we have measured in application performance from cloud offering that are better optimized for HPC requirements. Cloud systems intended for scientific workloads may benefit from a non-virtualized offering in order to achieve higher performance.

Similarly, large scale data parallel scientific applications could benefit from the MapReduce implementations such as Hadoop. However, scientific applications often exhibit data access methods and patterns that are not well addressed in the framework. Also scientific workloads typically have unbounded resource needs that are not accounted for in current day cloud scheduling. Thus, there are significant challenges in exploiting cloud resources and in many cases the current service models do not easily map to the needs of most scientific workloads. [3, 2].

**Future Work and Conclusion** While the Magellan Project has made progress on many of the questions, it stills has many questions left to explore. Now that the testbed is fully operational, more scientists are using the resources to explore how their applications behave in these environments and what changes are required to fully leverage these systems. While cloud computing has attractive features that can benefit scientific workloads, current cloud offerings do not provide out-of-the-box solutions for these applications.

The momentum behind cloud computing is reshaping the entire computing landscape. The Magellan project aims to help scientists navigate it.

**Acknowledgements.** The Magellan Project is funded by the Department of Energy from The American Recovery and Reinvestment Act of 2009 under contract number DE-AC02-05CH11231.

## REFERENCES

- [1] K. Jackson et al. Performance Analysis of High Performance Computing Applications on the Amazon Web Services Cloud. In *2nd IEEE International Conference on Cloud Computing Technology and Science*, 2010.
- [2] L. Ramakrishnan, S. Campbell, S. Canon, T. Declerck, I. Sakrejda, S. Coghlan, N. Desai, R. Bradshaw, P. T. Zbiegiel, and A. Liu. Magellan: Experiences from a Science Cloud. In *Science Cloud Workshop*, June 2011.
- [3] L. Ramakrishnan et al. Defining Future Platform Requirements for e-Science Cloud (Position paper). In *ACM Symposium on Cloud Computing 2010*, June 2010.

## **WHITE PAPER – NOAA’s Implementation and Use of High-Performance Computing in the Cloud**

**By The National Oceanic and Atmospheric Administration (NOAA), Office of the Chief Information Officer**



Cloud computing offers opportunities to meet increased demand, but it introduces new challenges for optimal use of high performance computing capabilities. NOAA is partnering to optimize computing resources and we’re addressing issues such as software management, allocation of computing time, and data transfer to ensure efficient use of available high performance computing.

The American Recovery and Reinvestment Act enabled NOAA to accelerate the implementation of its High Performance Computing Strategic Plan. Under this plan, NOAA is consolidating its research and development (R&D) high performance computing resources and building out more robust infrastructure to support the use of computing remote from its scientists.

NOAA has located its high performance computing on-site or in close proximity to the scientists lending to an organizationally based allocation and utilization of each system. These systems were configured to suit the needs of each organization; however, no one system had the capacity to support the extensive experimentation with large, more complex, higher resolution, future generation models. Transitioning to consolidated R&D high performance computing resources, NOAA is implementing higher capacity systems enabling the allocation of resources in a mission focus manner with more agility.

NOAA has and continues to utilize high performance computing resources external to the agency in remote locations from its scientists. Recently, NOAA was able to run experimental versions of hurricane models in near real-time on a supercomputer at the Texas Advanced Computing Center. NOAA continues to utilize computing resources with the Department of Energy for experimentation with climate models. These experiences have informed the thinking of NOAA’s strategy and have illuminated the challenges associated with creating a cloud of consolidated R&D high performance computing within NOAA.

This strategy poses the challenge of creating a common production environment across these consolidated systems with transparent access to the data which are in data centers remote from the scientists. To address these challenges, NOAA is implementing a common workload manager to enable intelligent scheduling of jobs across the enterprise. NOAA has also invested in building out a high bandwidth, low latency national wide area network to enable staging and automated movement of data between data centers and the scientists.

In addition to building out the computing infrastructure and common environment, the challenge of optimizing the software to easily port between each of the systems in the cloud and effectively scale on higher capacity systems remains.

## DHS Cloud and Impediments to Sharing Data



**By: Department of Homeland Security, Office of the Chief Information Officer, Enterprise Systems Development Office**

The Department of Homeland Security's (DHS) private cloud enables on-demand network access to a shared pool of configurable computing resources. Cloud Computing also reduces costs, allows rapid service provision, and provides secure data sharing, while only requiring minimal management and nominal service provider interaction. The DHS Common Operating Environment (COE) provides infrastructure dedicated solely to serving DHS needs.

### PRIMARY SERVICES

The DHS Cloud environment now spans the two new, geographically diverse Enterprise data centers located in Mississippi and Virginia. The private Cloud deploys more than 90 virtual machines supporting applications as part of the Office of the Chief Information Officer's (OCIO) four primary services being stood up by the OCIO Enterprise System Development Offices (ESDO):

- Infrastructure as a Service (IaaS)
- Platform as a Service (PaaS)
- Software as a Service (SaaS)
- Consulting as a Service (CaaS)

### CERTIFICATION AND ACCREDITATION

The DHS Chief Information Security Officer (CISO) certifies and accredits that the infrastructure within the COE meets the criteria of the National Institute of Standards (NIST) Special Publication (SP) 800-37, *Guide to Security Certification and Accreditation of Federal Information Systems*, SP 800-53rev3 *Recommended Security Controls for Federal Information Systems and Organizations*, and current DHS IT Policy directives and guidelines.

The DHS Cloud is inventoried and accredited as a General Support System (GSS), with a FIPS 199 security baseline of M/M/M for confidentiality, availability, and integrity. This certifies that the infrastructure follows NIST and DHS security guidance and policy. Individual production applications added to the COE must obtain an Authorization to Operate by performing an independent certification and accreditation, have an assigned Information Systems Security Officer (ISSO), and provide any additional tailored application-level security controls protecting the data within the system. The certification and accreditation of individual applications is highly accelerated due to the large number of inheritable controls from both the data center and the COE environment. The continuous monitoring and Operations and Management of the environment is provided by the data center providers.

The cloud observes traditional data stewardship and governance principles. All shared data have stewards that govern access, ensure quality, and manage the asset. As part of the security process to host an application in the DHS cloud, the system sponsor must complete a privacy impact assessment and system of record notice, unless an existing agreement or addendum covers the system.

### IMPEDIMENTS TO SHARING DATA

The most significant impediments to sharing data in the cloud environment are lack of trust and understanding of the data. "One DHS" and Federal Information Sharing Environment (ISE) policy currently govern information sharing and access between agency clouds. Providing standardized audit services is critical to establishing trust and compliance mechanisms. As information is provided through cloud services, data is classified data security markings, but applying data classification algorithms and roll-up logic to dynamically classify data continues to be a challenge. Additionally, a private cloud can only provide data on behalf of stewards if predefined security, privacy, and business controls can be realized.

## Cloud Computing and its' Impact Within Academia

By: Jimmy Lin – University of Maryland (interviewed by Seth Finkel)



UNIVERSITY OF  
MARYLAND

### INTRODUCTION/SUMMARY

The influence of cloud computing on the world of academia is wide-ranging, and impacts not only how universities are leveraging cloud computing infrastructures within their engineering and science curricula, but also how these institutions embrace previously unachievable contemplation and analysis of socio and cultural behaviors thru initiatives within emerging areas of study.

### EDUCATING THE CLOUD-ENABLED GENERATION

Many within the academic community believe that we are not currently training the people in the right way to deal with our emerging and changing, large data world. Conventional methods of teaching and learning need to keep pace with technological change. Currently, despite the high percentage of unemployment in the United States, there are countless jobs in this large data space, and not enough people to fill them. Among the questions universities must ask and answer are, “How do we train the next generation of data scientists to have the necessary skill sets to derive insights and value from these extremely large data sets?”

Many universities are exploring new partnerships with cloud infrastructure providers. The University of Maryland, for example, is working with with IBM, Google, and Amazon and leveraging their IT infrastructures within the engineering and computer science students' curriculum. Students use cloud resources provided by these companies for scientific analysis and learning.

Universities are also examining how to enhance their curricula, specifically with the new and emerging career areas that will likely result from our ability to perform with dramatically increased compute and storage resources. Certainly, “core” computer science skills will still be needed such as developing expertise in distributed architectures, algorithms, computations and data structures. But in addition, there are new and emerging competency areas such as data visualization, probability and statistics and machine learning will likely be keys to 21st century careers. As Hal Varian, Chief Economist of Google has said, “in 10 years, a statistician will become a very sexy job.... Now we have these systems that collect everything, and the bottleneck is analyzing. There simply aren't enough people to do it.”

Large data will always be large data. This is a problem that will never be solved. Today's “large data” is defined in terms of petabytes and exabytes, while tomorrow's data will be defined in exabytes and yotabytes. We will always be generating data faster than we can process it. Universities must step-up their efforts to train the next generation of experts who will address this burgeoning demand for effective management of information.

### EMBRACING NEW SCIENTIFIC VISION AND OUTCOMES

Another important aspect of cloud computing solutions that academic institutions are working to understand is what new problems can be solved – or old problems solved in a new way - with the advent of massively available (and affordable) IT resources such as those made available in cloud computing-based environments.

For example, suppose a scientist runs an experiment, presents a result on this data set and writes a paper on it. Before cloud, if a colleague wanted to use a copy of the data set to run their own experiment/analysis, both parties would need to figure out a way to provide the data set – and if the data set was too large, this simply couldn't be done. With a centralized cloud store, a single data set could be accessed by countless

scientists – each running their own analysis – resulting in a potentially compounded/beneficial result.

Whether it is a climatologist sharing climate models to help predict natural disasters, or high-energy physicists sharing the output of their super-collision machines – the cloud encourages reproducibility – which is at the bedrock of all science. Benefits of this work model include improved validation and accuracy of findings and the complementary building upon each other's work. Simply put, cloud establishes the shared infrastructure to help scientists collaborate within and across the academic community.

(Of course, this scenario presents its' own set of challenges amongst scientists – to include origination of ideas and evidence of discovery. With a shared or compounded result, who takes credit for what? Perhaps, this is better left to the folks at Pulitzer to decide.)

Another fascinating scenario to contemplate, and one that I (Jimmy Lin) am directly involved, is the enhanced processing and analysis of user-generated content and the social media revolution. It's an emerging area of study known today as *computational social science*. The idea, at a high level, is that for the first time in recorded human history we have details over time of the activities of literally hundreds of millions of people. What they are doing and thinking – and why. These inputs go beyond the fun and games to include issues on a global scale – the Egyptian revolution tumbling Mubarak, revolution in Tunisia, impact and relief efforts for massive natural disasters such as earthquakes and tsunamis.

Imagine this... We have these records in a computational form. Can we use this data to better understand ourselves and solve questions that have challenged sociologists for decades or even centuries? How do rumors spread? How do people become influential? How do groups self-organize? We now have the raw data to begin to answer these questions. And if we can answer them, we can perhaps provoke better civic life and desired outcomes for our society.

## WHITE PAPER – Agency Case Study per High Performance Cloud Computing

### Case Study: NASA Jet Propulsion Lab's Desert Research and Training Studies – Read Maloney <readmal@amazon.com>



*Summary: NASA's Jet Propulsion Laboratory (JPL) uses Amazon's cluster compute environment to process high resolution satellite images that provide guidance and situational awareness to its robots. To streamline processing, JPL relies on Amazon Cluster Compute Cloud Instances and Amazon Simple Queue Service (Amazon SQS) to deploy massive computations with less effort.*

NASA's Jet Propulsion Laboratory (JPL) has developed the All-Terrain Hex-Limbed Extra-Terrestrial Explorer (ATHLETE) robot. As a multi-purpose vehicle, each of the ATHLETE's six limbs is attached to a wheel, enabling the vehicle to travel across various types of terrain—ranging from smooth surfaces to rolling hills to ruggedly steep terrain. However, the wheels can also be locked to transform the limbs into general purpose legs that can be used as feet. The ATHLETE robot can also be used for loading, unloading, and transporting cargo for long distances.

As part of the [Desert Research and Training Studies](#) (D-RATS), JPL performs annual field tests on the ATHLETE robot in conjunction with robots from other NASA centers. While driving the robots, operators depend on high-resolution satellite images for guidance, positioning, and situational awareness. To streamline the processing of the satellite images, JPL engineers developed an application that takes advantage of the parallel nature of the workflow. JPL relies on Amazon Web Services (AWS) for this effort.

The application is built on [Polyphony](#), which is a modular workflow orchestration framework designed to streamline the process of leveraging hundreds of nodes on Amazon Elastic Compute Cloud (Amazon EC2). By accommodating excess capacity on local machines and spare resources in the supercomputing center, Polyphony meshes perfectly with Amazon's cloud computing. Most important, Polyphony enables the resources to work together to achieve a common goal. By using Amazon Simple Queue Service (Amazon SQS), JPL developers can deploy massive computations on Amazon EC2 by writing as little as a single class.

JPL had previously used Polyphony to validate the utility of cloud computing for processing hundreds of thousands of small images in an Amazon EC2 environment. However, JPL has adopted the cluster compute environment for processing huge images and recently processed a 3.2 giga-pixel image to support the ATHLETE robot operations in its 2010 D-RATS field test. Khawaja Shams, Senior Solution Architect, reports that "AWS's resources completed the work in less than two hours on a cluster of 30 Cluster Compute Instances. This demonstrates a significant improvement over previous implementations."

In addition to its support for the ATHLETE robot, Polyphony has been delivered to the [Mars Science Laboratory](#) to serve as one of the primary data processing and delivery pipelines that process data downloaded from Mars. Khawaja Shams, Senior Solution Architect, explains that the application "allowed us to process nearly 200,000 [Cassini](#) images within a few hours under \$200 on AWS." Due to the lack of elasticity available internally before switching to AWS, Khawaja explains that "we were only able to use a single machine locally and spent more than 15 days on the same task." The efficiency and cost-savings offered by AWS has proven invaluable.

## WHITE PAPER SUPPLEMENT – An Introduction to FedRAMP

By: Apptis | 4800 Westfields Blvd, Chantilly, VA 20151 | 703-745-6016 |  
[www.Apptis.com](http://www.Apptis.com) | [www.FedCloud.com](http://www.FedCloud.com)



Although many security issues are introduced by cloud computing, there are also advantages to having applications operate in a cloud environment. Moving public data to a public cloud reduces the exposure of internal, sensitive data to the Internet. The homogeneity of the cloud environment also simplifies security auditing and testing. An example of leveraging these features is the Federal Risk and Authorization Management Program (FedRAMP), a cross federal government initiative, led by GSA, which creates agreed upon security requirements among federal departments and agencies for certifying and accrediting a cloud-based system. Using FedRAMP eliminates the duplication of effort and cost for each agency working to address security requirements for cloud applications.

FedRAMP recognizes that the individual agency risk management of shared systems leads to duplicative management efforts, incompatible vendor requirements, a slowdown in system acquisition and fielding because of lengthy compliance processes, and has the potential to result in the inconsistent application of federal security requirements. FedRAMP minimizes these by creating a government-wide initiative to provide joint authorization services, continuous monitoring, and unified government-wide risk management. Agencies can leverage FedRAMP accreditations for Assessment and Authorizations while retaining their responsibility and authority.

FedRAMP provides centralized security authorizations of cloud computing systems (both commercial and government) to be used government-wide. Centralizing authorizations allows multiple federal agencies to leverage a single security authorization. The FedRAMP Joint Authorization Board (JAB) is comprised of three permanent members – the Department of Defense (DoD), Department of Homeland Security, and the GSA, and the FedRAMP Common Security Requirements (CSRs) are for moderate and low impact systems. CSRs are heightened security requirements from the National Institute of Standards and Technology (NIST) Special Publication 800-53 baseline that includes additional controls and control enhancements for low systems and additional controls and control enhancements for moderate systems. These additional controls address issues of multi-tenancy, shared resource pooling, lack of trust, visibility, and control of the service provider's infrastructure.

FedRAMP makes public a set of CSRs and process documents which the FedRAMP Office and the JAB uses to assess and authorize systems up to FIPS Publication 199 Security Category Moderate and DoD Mission Assurance Category II (MAC II). The DoD will begin transition to FedRAMP in 2011 and will complete the transition by 2014.

Once FedRAMP authorized, it remains the responsibility of each agency to review FedRAMP authorization packages prior to making a decision to accept risk, determine the suitability of the FedRAMP authorization to the agency's mission and risk posture, determine if additional security work is needed, and perform agency specific security activities.

**Contact:**  
**Phil Horvitz, Apptis CTO**  
**(703) 579-0721**  
**[Phil.Horvitz@Apptis.com](mailto:Phil.Horvitz@Apptis.com)**

The results are risk management cost savings and increased effectiveness, an interagency vetted approach, rapid acquisition through consolidated risk management, and consistent application of federal security requirements.

## WHITE PAPER – Cloud’s Great Challenge: Data Sovereignty & Sharing

By: Apptis | 4800 Westfields Blvd, Chantilly, VA 20151 | 703-745-6016 |  
[www.Apptis.com](http://www.Apptis.com) | [www.FedCloud.com](http://www.FedCloud.com)



Security continues to dominate the discussion of the adoption of cloud computing and other light technologies yet it is not the most complex or intractable challenge agencies will face when optimizing their resources and complying with OMB Budget Guidance, the Federal Data Center Consolidation Initiative, the *25 Point Implementation Plan to Reform Federal Information Technology Management*, and the *Federal Cloud Computing Strategy*.

Consider the alert issued by the New Zealand Inland Revenue Department in December 2010: “It is the Commissioner’s view that only business records stored in data centres physically located in New Zealand will comply with the record keeping obligations in the Inland Revenue Acts. Taxpayers are responsible for ensuring they comply with their record keeping obligations. Therefore, taxpayers using a **cloud computing** service will need to be satisfied that all their **business records will be stored in data centres located in New Zealand.**”

If New Zealand’s position seems extreme, NIST SP 800-53 Security Maintenance Control MA-5 states: “The intent of this control enhancement is to **deny individuals** who lack appropriate security clearances (i.e., individuals who do not possess security clearances or possess security clearances at a lower level than required) or **who are not U.S. citizens, visual and electronic access to any** classified information, Controlled Unclassified Information (CUI), **or any other sensitive information** contained on the information system.” NIST SP 800-53 defines “sensitive information” as “Information ... to which individuals are entitled under 5 U.S.C. Section 552a (the Privacy Act)...”; the clear instruction being that Personally Identifiable Information (PII) maintained by an agency may only be viewed or accessed by U.S. citizens.

The U.S.-EU & Swiss Safe Harbor Framework maintained by U.S. Department of Commerce in consultation with the European Commission has been bridging the gap between the European Union (EU) “adequacy” standard for EU and U.S. approaches to privacy protection, but recent legislation in Germany requires additional diligence when transferring data to Safe Harbor-certified U.S. entities and undermines the European Commission’s decision that Safe Harbor certification is sufficient to demonstrate an adequate level of privacy protection.

Where security concerns are addressable through people, process, and technology, data sharing and sovereignty are legal and policy issues that influence an agency’s choice for short and long term that must guide adoption of light technologies. This creates a conundrum for developing governments policy and implementations of cloud computing. If traditional approaches to Data Sovereignty and Governance, where laws, regulations, and policies that control access to data continue to fail to accommodate the evolution of housing and maintaining data outside of conventional boundaries, countries will be forced to develop national cloud infrastructures that are restricted to geographically local solutions. These “country-based” cloud infrastructures will limit choice, increase cost, and diminish the benefits of the continuity and economies of scale that can be achieved when multinational clouds seamlessly interoperate and port data to each other. As we look to past security into our next series of cloud challenges, we must prioritize the development of policy and governance that permits the secure, free flow of data across nations.

**Contact:**  
**Phil Horvitz, Apptis CTO**  
**(703) 579-0721**  
**Phil.Horvitz@Apptis.com**

## WHITE PAPER – Data Sharing in the Cloud Aided by the Service Request Lifecycle



By BMC Software - [www.bmc.com/cloud](http://www.bmc.com/cloud) - Susan Lynch Susan\_Lynch\_CW@bmc.com

The federal government is increasingly pressured to reduce the footprint of IT across the country. At the same time, it must deliver the IT services required to support government agencies. That’s why federal IT organizations are turning to cloud computing as a model for a shared infrastructure, attaining economies of scale while serving the needs of the country.

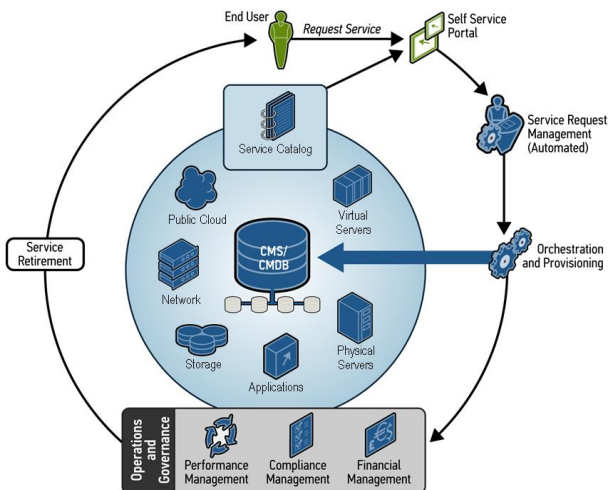
To achieve success with a cloud computing environment, federal IT must offer a set of configurable cloud services to their users, while at the same time, maintaining control of the shared infrastructure to ensure its integrity and efficiency. With robust service levels, this “dial-tone” approach to computing has the power to transform IT into a reliable, effective and ubiquitous asset to the federal government.

Federal IT organizations are currently developing their visions and plans for cloud computing. They are also exploring the role of external cloud services in their long-term IT plans. Often, they are engaging vendors to co-develop an initial solution and roadmap — or a complete cloud environment — to leverage existing IT management investments and supporting a broad range of use cases.

Cloud environments require self-service request and management of IT services, automated service provisioning, and ongoing service management. Contextualizing cloud computing within existing IT environments is an evolutionary approach that requires heterogeneous support and integrated management — from the CMDB, to automation and orchestration, to capacity management, performance management, compliance, and IT business management.

In the next year, federal IT organizations will establish initial cloud environments, with plans to grow those environments in scale and scope to meet user needs. Further, as implementations mature, organizations will leverage operational and governance solutions to infuse Business Service Management best practices throughout the cloud environment.

### The Cloud Service Request Lifecycle



### The Cloud Service Request Lifecycle

The cloud service request lifecycle is driven by automation and integration to deliver cloud services to the user. The major steps include:

- **Self-Service Request and Management Portal:** Enable users to configure and request new IT services and manage existing cloud services
- **Approval and Change Workflows:** Ensure all requests are appropriately passed through customizable, automated or manual, change management workflows
- **Automated Full-Stack Provisioning:** Deliver user-configured, multi-tier cloud services — from the resources through the operating system and the applications — efficiently using automation

- Ongoing Management of Cloud Services: Once provisioned, manage the cloud services appropriately throughout their lifetime, including:
  - Financial Management: Apply metering and accounting for the resources utilized by each department, organization, or user for chargeback or showback purposes
  - Compliance Management: Ensure the ongoing compliance of cloud services and the broader cloud environment
  - Performance and Capacity Management: Maintain the performance of cloud services and ensure the capacity requirements of each is met throughout its lifetime
  - Automated Service Retirement: Retire a cloud service and re-use its resources once they are no longer needed

### **BSM Solutions for Cloud computing**

Cloud computing can bring tremendous value to the federal government, aiding in achieving the efficiency gains and consolidation goals of the federal IT leadership. With some careful planning and innovative, integrated solutions, organizations can build cloud environments that meet the needs of the government agencies, while at the same time, optimizing the resources of IT.

**Contacts:**

**Lilac Schoenbeck** is Senior Product Marketing  
Manager for Cloud Computing

**Dan Trevino** is a Senior Product Marketing  
Manager for Cloud Computing

## WHITE PAPER – CloudFirst and FedRAMP; “New Wine in Old Wineskins?”



By: CSC; Based on an essay by Ron Knode posted at  
<http://www.csc.com/cloud/blog/>

The [Mitre-hosted cloud blog](#) presents new questions each month for invited blog responses. For the month of January 2011, the question dealt with the U.S. government’s recently announced “cloud first” policy for all agencies and departments.

The “cloud first” policy declaration in OMB’s 25-point plan of 9 December 2010 is aggressive thinking and terrific branding. The triple play promises of economy, flexibility, and speed are precisely the kind of IT payoffs that any enterprise would want government or commercial.

However, these promises are themselves based on another promise in the same plan, i.e., the promise of a cloud strategy that can deliver safe and secure cloud adoption across the U.S. government. While there is much to like about the ambitious vision and the no-nonsense “let’s get going now” message for cloud processing in the plan, real success hinges on making the underlying promise of a practical cloud strategy come true. That promise is the more difficult one. It must respond not only to the needs and realities expressed by (government) cloud consumers, but also to the needs and realities of cloud service providers who can actually deliver these payoffs. Only when both constituencies are accommodated in strategy and mechanics can we move from a hit or miss “Ready, Fire, Aim” process to a reliable “Ready, Aim, Fire” process for cloud adoption and payoff.

And, there’s the rub. According to OMB plan, the promise for a practical cloud strategy is rooted in the development of standards for cloud service security, interoperability, and portability. The initial public draft of the [Proposed Security Assessment & Authorization for U.S. Government Cloud Computing](#) (aka “FedRAMP”) took a healthy first swing at such standards, but does not yet tend to the needs of all the constituencies involved. Continuing ambiguity about overall risk governance and accountability, a monitoring framework that excludes the cloud consumer, and a complicated scheme for trying to shape Spec Pub 800-53 for cloud services all present high hurdles to overcome. The process being followed invites industry comments, and we would expect subsequent versions to improve on such circumstances.

Yet, one cannot but wonder if the biblical admonition against “pouring new wine into old wineskins”<sup>1</sup> must be observed here. The cloud processing phenomenon is an evolution of technology but a revolution in the consumption model. The payoffs of resource pooling and rapid elasticity come from having a process that brings cloud services to the consumer in “cloud time”. Trying to bend the conventional machinery for government certification and accreditation (C&A) into a community process for assessment and authorization (A&A) without clarifying who is accountable for risk acceptance in cloud services only slows cloud adoption. The attempt to fashion existing Spec Pub 800-53 controls into a set of requirements suitable for cloud processing is laudable, but does not suit the consumption model of the cloud. In other words, the old wineskins of traditional C&A models and Spec Pub 800-53 cannot yet handle the new wine of cloud processing.

We will see if the multitude of comments submitted against the November 2010 draft of FedRAMP can create a new wineskin of assessment and practice for government cloud computing. Recently, however, NIST

---

<sup>1</sup> See Mark 2:22 or Luke 5:37-38

issued [Draft Special Publication 800-144, Guidelines on Security and Privacy in Public Cloud Computing](#). While this draft does not offer requirements or controls, it does offer a set of good practices on how to risk manage the adoption of (public) cloud computing for a government enterprise. In so doing, the draft suggests some prudent steps for the cloud consumer to take without prejudicing the techniques for security delivery by the cloud provider. The Guideline is an example of (almost) a “new wineskin” process for the adoption of the “new wine” benefits of cloud computing.

Until we fulfill the promises made in the OMB plan, we will be constrained within the government to applications that satisfy the compensating techniques first introduced in “[Digital Trust in the Cloud](#)” and subsequently amplified in many other places. We can gain some benefit from “safe” applications like non-sensitive email, development and test, backup and restore, and even a bit of collaboration and social networking. But, such applications do not deliver the kinds of payoff we need and expect from cloud processing.

Government enterprises are not unique in their attempt to find a way to grab the benefits of cloud computing without unnecessarily raising information risk thresholds. Industry enterprises share this need. In both cases there is not only the need to clear a set of qualifying hurdles with regard to the security characteristics of the cloud service being offered, but more importantly to include a [mechanism for reclaiming visibility](#) into how the cloud is operating right now on our behalf. This need for transparency in cloud processing is unrelenting regardless of the cloud provider or the deployment model being used. And, satisfying such a requirement has the biggest impact on trust generation and payoff with cloud services.

In his earlier blog on this matter Chris Hoff declared “[we’re gonna need a bigger boat](#)” Simply enlarging the vessel may not be enough. The biblical warning declares that “both the wine and the skins will be ruined”<sup>1</sup> if we try to pour new wine into old wineskins. The new wine of cloud processing may well need completely new wineskins (standards and practices) for us to enjoy the rich bouquet of enterprise payoffs.

## WHITE PAPER – Data Sovereignty and the Cloud



By: CSC; <http://www.csc.com/cloud/blog/>

Sharing data is a natural extension or desired outcome of residing on the same infrastructure which comprises the cloud. In the past and even today, we share information over a network and use servers as data repositories to which we all can connect. Now, with the cloud, not only can sharing be simpler but the added abstraction of where the data physically sits should raise the concern that the data being shared may have a potential for exposure – intentional or unintentional. This issue describes the security of sensitive information as it is passed from one user to another and is known as data sovereignty.

Traditionally, sovereignty has been an idea usually connected with the political context: control of a nation's borders. It may also be applied in political, economic and social systems in the form of self reliance or independence. Those concepts do not immediately come to mind when one thinks of privacy and data protection.

Sovereignty centers on the concept of control. It requires a specific level of freedom before it can operate and have meaning. It flourishes in environments where there is reciprocal respect, trust and transparency – and usually languishes elsewhere. Interestingly enough, the environments used in cloud do not foster these traits as it depends and relies on the individual having control of the actions on data.

The concept of sovereignty today may be outdated but the issue is as important as ever. Government and businesses require more from us and in turn we often surrender our personal details upon request and in many arenas.

For governments and business seeing the promise of what widespread cloud computing adoption can deliver, this is going to be one of the biggest challenges for them as they need assurances as to where the data is stored. Today, most countries have rules and regulations about where personal data can be stored. For example, the European Union has the Data Protection Directive which regulates the processing and the free movement of personal data within the European Union. Today, a company covered by this directive and using any cloud platform located in the U.S. to store personal data of their customers could be in clear violation of this directive – depending on who you ask. The fact that data residing in the cloud could effectively be anywhere geographically is going to be an issue.

Data sovereignty has long been a topic of keen interest and discussion in the industry because there is no easy answer. Of all the barriers articulated as reasons not to move to the cloud, data sovereignty may be the only one with any degree of legitimacy. Organizations like the Cloud Security Alliance have recognized this need and have formed teams specifically focused on addressing, maintaining and mitigating the risk of data sovereignty.

It is not a question which will be solved quickly as the answer is not uniquely technical, legal or policy driven – it is a composition of all three. Coupled with the fact that service providers vie for business based on price (assuming the feature sets are comparable to their competitors), it

adds another wrinkle as their costs have a geography element too – two identical data centers have very different costs if one is in Delhi and the other is in Manhattan.

The problem is not impossible to solve but has a few more pieces to consider and tightly integrate.

Regulated industries such as banking who deal with significant privacy and security issues overcame similar issues years ago. Today, cash can be withdrawn from ATMs and credit cards can be used in stores from almost anywhere in the world. Additionally, the call centers who handle their customer calls are often located in different countries but routed to facilities in specific geographies when handling personal data to be in accordance with laws by which they perceive to be governed.

As the concept of the cloud, especially the data cloud, grows, the physical location of the storage is supposed to be unimportant. The whole notion of the cloud is to “trust your data to us, and don’t worry about the details.” However, laws around data retention and storage, which vary widely from country to country (and at even lower levels such as states and municipalities), are at odds with this notion.

While the issue of data sovereignty is probably never going away and will need to be addressed as technology and laws change, it will most likely raise a series of new questions with every change. One which stands out today relates to when an entity owns data, but that data is located outside of the country, what laws are applicable to that data – the laws of the country where the data resides, the laws of the country where the company who holds the data is located, or the laws of the country identified as owning the data in question? This is where a sectoral approach (combining of legislation, regulation, and self-regulation rather than government regulation alone) can have overlaps and gaps which lead to confusion in comparison with clear cut government regulated directives such as what the European Union has established.

This one question has no easy answer nor will any answer received today will stand the test of time because laws are subject to change without notice. Cloud providers in an attempt to respond to this issue have offered to let users choose what data centers to keep their data in. This breaks the value and, even more, the more abstract premise of the cloud.

In summary, maintaining and ensuring data sovereignty will be critical to widespread cloud adoption regardless of the level of sensitivity of the data which is stored in it. CSC understands these issues well and has been working on solutions using trust models and protocols which will balance the mandates placed on our customers to maintain data sovereignty in the cloud.

## High-Performance Computing: a Viable Cloud Workload?

### By Quest Software, Public Sector



Although many argue that high performance computing (HPC) is not a suitable workload for the cloud – we think it is ideal. One of the issues often mentioned by those considering high performance cloud computing (HPCC) is that HPC workloads require a different infrastructure than typical “general purpose” workloads. But of course they do! But that doesn’t mean that you can’t put HPC in the cloud. Why does there have to be only one cloud?

To start, let’s take a quick look at the NIST (National Institute of Standards and Technology) definition for cloud computing:

*“Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.” [NIST, v15]*

This definition never mentions the types of workloads that can be processed; It simply addresses the three tenets of the computing model: the access model (*convenient, on-demand network access*), the resources (*shared, configurable*) and their management (*rapidly provisioned and released with minimal ... interaction*). The first two tenets, the access model and shared, configurable resources, are fairly self-evident and directly relevant to the Cloud computing model.

The third tenet of the NIST definition, management, is where HPC can really benefit from Cloud technologies. Let’s review that last part of the definition:

*“...that can be rapidly provisioned and released with minimal management effort or service provider interaction.”*

This tenet, the ability of the end user to provision and de-provision assets with minimal management effort, is what separates Cloud computing from simple time sharing (a concept that HPC has been using for years). The ability to effectively schedule computer resources enables multiple organizations to share the same HPC infrastructure (whether a single super computer or a cluster of general purpose systems). This spreads out the cost and gives each organization access to greater capacity than they could afford on their own.

This shared access model offered by cloud computing is effective as long as no individual organization’s resource requirements consume too much of the overall resource pool for an extended period of time. Effectively, the cloud must have enough capacity to meet a large percentage of the utilization requests by all of its consumers.

From a practical perspective, the resource that is most likely to constrain a cloud’s ability to deliver is storage. Most HPC applications consume enormous datasets during their execution. These datasets must be provisioned, managed and de-provisioned – including protection from unauthorized access or disclosure – with adequate performance to meet the throughput needs of the HPC application.

So, in answer to the initial question “Is HPC a viable workload for the cloud?”, we believe the answer to be a resounding “Yes!” However, concerns over data sovereignty may, for the foreseeable future, push most

users to a private or hybrid/community cloud where they will have control and security over their own data. Until there is sufficient intelligence to adequately protect the data, large-scale adoption of public clouds by public sector users, whether they are HPC or “regular” users, will remain pretty much a pipe dream.

---

*Our integrated portfolio of solutions simplifies the management of your physical, virtual and cloud infrastructures to reduce costs and improve efficiency. Quest is the market leader in **virtualization and cloud management**, with solutions spanning server virtualization, desktop virtualization, and cloud automation that support multiple hypervisor platforms.*

*Quest is simplicity at work – choose your hypervisor platforms, then choose Quest.*

*<http://www.quest.com/virtualization> / or call 703-820-8400*

## Data Sovereignty: the Biggest Impediment to Public Cloud?



By: **Quest Software, Public Sector**

Cloud computing offers many significant advantages and a public cloud provides them on a large scale. Having the ability to spread the cost of a vast infrastructure across an equally large number of consumers results in an amazing low price per unit, whether the unit be a gigaflop of CPU power, megabit of network bandwidth, or gigabyte of data storage. Couple the low cost of service with the ease of provisioning, payment, and access and you have a compelling use case for a public cloud!

So, what's the "gotcha"? If the lining of the public cloud is all silver, why shouldn't I use it? Well, the darker side of cloud computing lies in data sovereignty. When you hand over control of your infrastructure to a public cloud service provider, you're essentially saying "Here, take my data and do with it as you please." But wait! Don't you have regulatory requirements to ensure that your data doesn't cross country boundaries? How do you know that your sensitive information isn't sitting on the same infrastructure as that of an international drug smuggler? What happens when the Department of Homeland Security (DHS) walks in with a seizure order and takes possession of all systems used to house and process the smuggler's information? Your information is on those same systems, and it walks out the door with the DHS marshal – there's nothing you can do. What happens if the cloud service provider goes out of business and turns off their computers, shuts and locks the doors, and everything winds up in bankruptcy court? How do you *know* that some other user of the cloud system isn't loading up your dataset and casually browsing through it – just for fun, or worse?

You don't. And that's the problem with public clouds. Until data is appropriately labeled and applications and operating systems are intelligent enough to reliably honor those labels, there is no way – other than procedural – to enforce data sovereignty. What does this mean? A couple of things:

- There will be small, incremental improvements in data sovereignty protection over time
- The first steps will be procedural in nature (i.e. the cloud provider will implement policies and procedures to ensure, through operational means, that your data is appropriately protected)
- Data tagging techniques currently used in highly-secure environments (multi-level security systems) will be adapted for general purpose data uses
- Critical middleware applications (think databases, message queuing systems, etc.) will be updated to honor the tags (metadata) associated with the data being processed
- General purpose operating systems (Windows, Linux, OS/X, etc.) will be updated to honor data tags, much like applications will

And even if all these things come to pass, it will only apply to new documents and data. It's just not feasible to go back and tag all of the existing documents and data for many reasons; the most obvious being cost and capability. It is simply too expensive to do this manually, but automated classifying is not feasible (e.g. What if a top secret document is misclassified and revealed publicly?) In addition, older document and data formats don't support today's tagging and labeling methods and converting them to current formats adds cost and risks lost data.

Now that you know about the dark side of the public cloud – what can you do until data sovereignty can be protected? You don't want to miss out on all of that cloud goodness, so what are your options? If data sovereignty is important to you and you still want the benefits of cloud computing, begin by deploying a private cloud. With a private cloud, you house the cloud technologies in an environment that you control – in

your data centers. Plus, the private cloud can continue to live on into perpetuity to house those older data sets and documents once public and hybrid options become safer and more secure. This way, you gain much of the flexibility and cost savings associated with cloud computing, but you avoid the issues caused by allowing someone else to control your data.

---

*Our integrated portfolio of solutions simplifies the management of your physical, virtual and cloud infrastructures to reduce costs and improve efficiency. Quest is the market leader in **virtualization and cloud management**, with solutions spanning server virtualization, desktop virtualization, and cloud automation that support multiple hypervisor platforms.*

*Quest is simplicity at work – choose your hypervisor platforms, then choose Quest.*

*<http://www.quest.com/virtualization> / or call 703-820-8400*