Written Testimony of
Gregory R. Ganger
Professor of Electrical & Computer Engineering and Computer Science,
Carnegie Mellon University

United States House of Representatives
Committee on Oversight and Government Reform
Subcommittee on Government Management, Organization, and Procurement
Hearing on
Cloud Computing: Benefits and Risks of Moving Federal IT into the Cloud
July 1, 2010

I thank you for the opportunity to testify about the benefits and risks of using cloud computing for federal IT functions.

About me: My name is Gregory (Greg) R. Ganger. I am a Professor of Electrical & Computer Engineering and Computer Science at Carnegie Mellon University (CMU). For the last ten years, I have also served as Director of CMU's Parallel Data Lab (PDL). The PDL is a world-renowned research center focused on storage and large-scale infrastructures, such as cloud computing and more traditional data centers, regularly working with and annually supported by most of the major developers of technology in these areas. Current industry sponsors include Google, Microsoft, Yahoo!, VMware, HP, IBM, Intel, Oracle, Facebook, APC (of Schneider Electric), EMC, Hitachi, LSI, NEC, NetApp, Seagate, and Symantec.

I have been conducting research on large-scale computing and storage infrastructures (e.g., cloud computing) and their operation/administration for over a decade. Among the cloud computing projects I lead are CMU's Data Center Observatory (DCO) and the CMU portion of OpenCirrus. The DCO was conceived as a consolidated data center and private cloud for research computing/storage needs, but heavily instrumented and forward-looking to enable research into efficiency, and it is being realized with active collaboration from several of the PDL sponsor companies. OpenCirrus (https://opencirrus.org/) is an open cloud computing testbed currently consisting of ten sites worldwide, each of which provides public cloud computing resources via open interfaces and open source software.

Testimony roadmap: I have been asked to testify about the use of cloud computing for federal IT needs, including potential benefits, risks, challenges, and consequences. My

written testimony is organized as follows. First, I provide a brief review of cloud computing generally, highlighting a few forms that it can take, including the highly relevant concept of a so-called "private" (or "internal") cloud. Second, I discuss the large potential benefits of using cloud computing for federal IT functions, which are similar (in many cases) to those for large corporate organizations. I highlight the benefits first because, while I suspect that most questions will focus on the risks and challenges, overall thinking about the concept of using cloud computing resources for federal IT functions should not lose sight of the large potential benefits of this young, maturing technology. Third, I discuss various risks, challenges, and consequences. Some of these (e.g., resistance to change) will require continuing education and strong guidance, possibly including explicit incentives. Some of these (e.g., lock-in and management complexity) will require patient and incremental approach to moving federal IT into the cloud, as advancement in both technology creation and standards bodies address unresolved issues. A few (e.g., security) may require certain IT functions to never migrate fully to a public cloud. None, however, preclude rapid partial migration of federal IT function into the cloud and expanded migration over time.

It is important to keep in mind, while considering pros and cons of moving federal IT into clouds, that it is far from an all-or-nothing decision. For some federal IT functions, it will be the right choice, and for others it may not be. The choice need not be the same for all IT functions, and movement can happen independently for each, allowing incremental movements that each yield benefits.

A. Cloud computing basics

Very broadly, "cloud computing" involves using someone else's computers (and possibly software setups), shared with yet other groups, for some task instead of using your own. There are many technical issues involved, which have delayed the realization of this long-sought notion of computing services as utilities, but the basic concept of outsourcing work is natural in today's service-based economy.

The "cloud" aspect refers to the fact that the computers used are on the network, somewhere, but that the cloud computing customer need not be aware of where they are or

details of how the outsourced work is completed – it is referred to as "in the cloud", because large networks (e.g., the Internet) are often illustrated as clouds in technical diagrams.

The term "cloud computing" has been applied to a broad class of IT outsourcing activities, leading to broad definitions. For example, NIST's definition[1] is more technical than my very brief description above, but it closes with "and is composed of five essential characteristics, three delivery models, and four deployment models." Just the cross-product of the three delivery and four deployment models yields twelve configurations that fit the definition. I will not detail the full breadth here, but I will highlight a couple configuration options in an attempt to help clarify cloud computing and important issues involved.

Raw resources vs. software services: The delivery model axis relates to the form of computing service purchased from a cloud computing provider. One option, called "Infrastructure as a Service (IaaS)" by NIST, is to rent raw computing resources, such as computer time or storage capacity. Which programs a customer runs in their rented computer time,[2] or what data is stored in rented storage capacity, is entirely up to the customer (who must, therefore, configure and maintain the programs themselves). Setting aside technical details, the IaaS concept should be familiar to anyone who has rented a car, exercised in a fitness center, or stayed at a hotel. The other two options, called "Software as a Service (SaaS)" and "Platform as a Service (PaaS)" by NIST, provide complete applications (e.g., email) and/or building blocks (e.g., database systems) for use by customers (and perhaps provided by customers to third parties). Setting aside technical details, these concepts are akin to outsourcing of food services, patent litigation services, or accounting services.

---

[1] The full NIST definition is two pages long, but the primary paragraph states "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. This cloud model promotes availability and is composed of five essential **characteristics,** three **service models**, and four **deployment models**." Most of the remainder details the five, three, and four. The latest version (v15) can be found at http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc .

[2] Rented raw computer time in most cloud offerings is used to execute software encapsulated in a so-called "virtual machine", which appears to the customer as a physical machine. Indeed, all cloud resources are "virtualized" in the sense that details of how they are provided are hidden from customers and may not match the appearance given to the customer – such virtualization enables improved efficiency and is fine for customers, so long as the behavior promised to the customer is realized.

Public cloud vs. private cloud: The deployment model axis focuses on who shares the cloud. One option, termed a "public cloud" by NIST, is made available to the general public by a provider selling cloud computing services. This is the option usually in mind when people first think about cloud computing, since it matches the general accessibility of the Internet. But, it is not the only option. Another option, termed a "private cloud" by NIST, is operated solely by one organization and shared by its various sub-parts. For particularly large organizations, such as the federal government or a large Internet service company (e.g., Google or Microsoft), many of the benefits of cloud computing can be realized with a private cloud model – for such organizations, the economies of scale and aggregation are sufficiently present without sharing externally, because of their many sizable sub-organizations.[3] Of course, an organization can use more than one cloud, including of different types, and can also used both cloud and non-cloud (i.e., their own) computer resources.

B. Potential benefits of moving federal IT functions into the cloud.

Cloud computing has the potential to provide large efficiency improvements for federal IT functions. As with outsourcing in non-IT domains, such as rental cars and food services, the efficiency arises from having multiple customers (organizations) share the provider's offering instead of each providing for itself. Efficiency improvements come from multiple fundamental sources, including: (1) increased utilization of resources, since sharing allows the portions unused by one customer to be sold to (used by) another, while each customer pays for just what they use; (2) economies of scale, since operational costs usually do not scale down linearly with resource size – for example, one cannot use a part of a car, and cooking for two takes nearly as long as cooking for five; (3) increased specialization, since experts working for the provider can focus on the one offering rather than being "jacks of all trades"; (4) low entry cost (in terms of time, effort, and dollars) for new customers, since the resource is already set up by the provider and ready for use. These benefits can all be present for cloud computing, with large potential reductions in IT costs (both capital and personnel), energy demands (due to the need for fewer total computers), and time to establish new IT functions.

[3] As one example, the National Business Center (NBC) of the Department of the Interior now provides some private cloud capabilities (http://cloud.nbc.gov/).

Although concision precludes full analysis here, two examples can help illustrate potential infrastructure efficiency benefits of even just one or two of these sources:

- Although an imperfect example, because of artifacts of CMU's smaller size and relative resource-poorness, our experiences making a case for using cloud computing for research computing at CMU provide some insight. In surveying the separate infrastructures used by research groups on campus, we found average utilizations around 25% -- that is, ¾ of the work potential of the computers went unused, over time, even in a University research environment that struggles to find funds to purchase equipment.[4] A private IaaS cloud computing approach with 75% utilization would reduce the number of servers needed by 66% or allow three times the work to be completed during heavily active times, which has induced us to aggressively pursue deployment of such a private cloud at CMU. Such numbers are normal, even laudable for the traditional "every group for themselves" approach, not a sign of misbehavior. Indeed, a GSA presentation[5] indicated "Average Server Utilization" values of 7-15%, offering even more room for improvement.

- HP's recent data center consolidation effort provides a second example. In 2006, HP identified their "many separate data centers" deployment (85 data centers across 29 countries) as a significant source of inefficiency. They noted plans to consolidate into six large data centers, estimating $1B/year savings in IT expenses and significant energy savings as a result.[6] Recently, HP's CIO Randy Mott shared some outcomes of this successful consolidation effort, including 60% reduction in overall data center costs.[7] Despite ever-growing demands for computing, HP reduced their number of server computers by 40%, which would combine with their improved cooling approaches to yield significant energy savings.

The savings in these examples do not even account for the much improved IT staff efficiency (#3 above) or the faster pace of deployed IT improvements (a consequence of #3). With consolidated infrastructures, IT staff specializing in particular aspects can focus on those aspects – because of the large scale, such specialization does not lead to excessively sized IT staffs. Since the particular aspects (e.g., network management or storage management) are handled by the provider, none of the customers need to employ staff focused on those aspects – one set of staff handles them for all, eliminating redundancy across customers and allowing customer IT staff's to focus on the customer's missions instead. Also, because specialized

[4] But, during active times, they tend to be overburdened.
[5] "GSA Presentation on the Federal Cloud Computing Initiative" by Michael Goodrich (Project Manager, FedRAMP and Apps.gov, General Services Administration) on Software & Information Industry Association panel. See slide 22. Available at http://www.siia.net/blog/index.php/2010/06/gsa-presentation-on-the-federal-cloud-computing-initiative/
[6] http://news.cnet.com/HP-plans-data-center-consolidation/2110-1011_3-6073187.html
[7] http://www.enterprisenetworkingplanet.com/news/article.php/3878966

staff have fewer aspects to manage, they can focus more attention on improving their specific aspects, leading to more rapid adoption of new technologies and best practices from which all customers immediately benefit.

In addition to significantly increasing efficiency across a set of current customer IT functions, cloud computing can greatly improve the situation for new IT functions (#4 above). Traditionally, a lengthy start-up process is often involved with establishing a new IT function, including procuring new computers (and sometimes building machine room space to power and cool them), installing and configuring the computers, and only then finally starting to set up the IT function in question. With cloud computing, one can rent pre-setup computer resources as soon as one has budget to do so, leading to much quicker progress on new directions. Moreover, one does not have the danger of incorrectly guessing how many computers are needed (which can lead to waste or delays), since the cloud provider allows rapid incremental scale-up (charging only for what is used) as long as the customer is willing to pay for what they use. Among other things, therefore, cloud computing could significantly accelerate deployment of e-government applications.

Overall, the potential benefits from cloud computing are huge, both for global efficiency (total equipment and energy used) and for each customer (dollars and mission focus).

C. Risks, challenges, and consequences

Cloud computing is very different from the traditional approach of each organization (e.g., agency) creating and maintaining their own computing resources, from top to bottom. Naturally, there are many challenges to be faced in making the significant transition to outsourcing aspects to external providers, particularly given the relative youth and rapid evolution of cloud computing. Of course, there are security concerns when an external provider is made part of an agency function. There are also "lock-in" concerns caused by lack of standardization and (in some cases) the difficulty of moving large data sets. Another significant source of challenges is the massive IT culture change inherent in a transition to cloud computing, which will require overcoming resistance to change and retooling IT staff skill sets.

Security concerns: Security is an issue for all networked computer activities. It is natural to imagine that security might be weakened by involving an external provider, particularly when confidential data are involved. But, it is not necessarily the case in all, or even most, circumstances. As in the real world, computer security is about risk management, not absolutes – most of us feel relatively secure in our homes, for example, despite glass windows on the ground floor.

Having federal agencies maintaining infrastructure does not guarantee their security, both because humans are imperfect and because no perfect computer security technologies exist. Public cloud providers are capable of employing the same best practices and technology as government agencies and potentially upgrading more rapidly to new advances (because of #3 above). The question is whether or not they can be trusted to do so. To establish that trust, there will need to be certification of the degree of trust that can be placed in a given provider, using established (e.g., FISMA) and perhaps new mechanisms – standardized approaches to doing this is an area of necessary, and ongoing, effort in technology working groups. Movement of IT functions to providers must be limited to those pairings with acceptable risk. Certain functions, and certain data, will perhaps never be appropriate for public clouds – highly classified intelligence activities, for example. But, for many federal computing activities, security needs are likely to be consistent with those of corporate customers of public clouds.

It is worth noting that private clouds, maintained by the government, can be used for IT functions that may require security efforts beyond those that public cloud providers are willing to employ (e.g., because they go beyond what corporate customers require).

Lock-in concerns: Currently, cloud computing offerings are diverse – one can choose among several to which to migrate a function, and then go thru the effort to migrate, but often there is no easy way to switch from one provider to another. Today, such a switch can involve time-consuming extraction of one's data, reprogramming of one's application to fit the new provider's interfaces, and uploading of one's data to the new provider. Each step can be onerous.

One big part of the problem is standardization or, rather, lack thereof. Although various working groups are now focused on standardization, it is still early in the process.

Indeed, cloud computing is sufficiently new that there is some danger in standardizing so quickly, with such a short window of experience from which to draw. Nonetheless, standardization is an important part of promoting compatibility and competition among cloud computing providers.

A technical issue, for IT functions that involve very large data sets, is the time required to upload or download the data. For example, at commonly available wide-area networking (WAN) rates, transferring multi-terabyte datasets to or form a public cloud could require multiple weeks, which would make the concept of migrating a high hurdle. This is a challenge that federal customers share with corporate customers, and technical solutions will undoubtedly be developed.

Resistance to change: Some of the trickier challenges faced when efficiency-seeking leaders push their IT staff to move some functions to a cloud are non-technical, relating to human nature. Some (not all!) IT staff resist changes to currently working practices that they control and understand. I suspect that, where it exists, this resistance will be stronger in consistently-funded government IT settings, where business-style pressures and incentives (e.g., bonuses) for innovative steps leading to tangible savings are not present. Simply demanding an IT change rarely yields desired outcomes, as unhappy IT staff can become inefficient in a variety of ways. A mixture of push (e.g., requests and insistent education) and pull (e.g., incentives) may be needed to effect rapid and positive adoption.

Perhaps the most common form for such resistance to take is aggressive arguing against the change in question, on technical grounds and by overstating the effort required to enact the change. The awkward aspects of such arguments are usually twofold: the IT staff raising them generally know more than anyone else in the organization about the technical issues in question, and the arguments raised generally are at least partially correct. A mixture of education (for the IT staff and their managers) and a technical mindshare (for both to utilize) may be needed to separate the legitimate concerns from those based primarily on a desire to avoid change.

The technical mindshare should also provide for sharing of effort on issues like certification/accreditation (e.g., for security issues discussed above), verifying continued good practices, negotiating Terms of Service (ToS), and procurement (e.g., multiple bids obtained

and okayed periodically). Forcing every agency to independently deal with such issues truly could become a significant barrier, but a shared clearinghouse is a natural way to eliminate redundant effort for common needs. Note that none of my discussion is meant to imply that actions, including those that I mention, are not already being pursued in the context of the Federal Cloud Computing Initiative; indeed, some are (e.g., see apps.gov).

IT culture changes: A consequence of moving to cloud computing is major change for IT staff. Note that even full transition to cloud computing would not mean elimination of all IT staff – not by a long shot. Expert IT personnel will be needed to assist with planning, to provision, and to manage IT functions outsourced to the cloud. But, the expertise that they will need is going to be different. Rather than expertise in managing the aspects now outsourced (e.g., physical computers, networks, and building-block applications), for example, IT staff and managers will need new expertise in working with cloud-based activities, projecting usage costs rather than capital costs, and there may be reduced separation between application engineers and IT staff. Continued education for IT personnel, and perhaps a new breed of staff, will be an important part of such transition.

Not only will new IT expertise be needed to manage functions outsourced to the cloud, but a hybrid IT model is most likely for quite some time – some functions will be moved to one or more clouds, while others remain "in house". Thus, the IT staff will need to manage a set of functions spread across multiple environments, using new integrated management tools. Creation of such tools can be expected, as particular cloud interfaces become very popular and/or standardized.

D. Concluding remarks

Cloud computing is an exciting realization of a long-sought concept: computing as a utility. Pursuing judicious use for federal IT functions is important, given the large potential benefits. Patience, perseverance, incremental adoption, and continued investment in research, education, and standardization related to cloud computing will be needed in realizing that potential. Some specific recommendations for consideration that follow from my observations include:

- First, cloud computing is a big change, and realizing its large potential will require significant formal technical and change management training for IT staff and managers. This need may warrant expansion or adaptation of programs like "scholarship for service" as well as targeted executive education initiatives.
- Second, standardization is important to address lock-in concerns, but continued experimentation (including research, testbeds, and case studies) and innovation are also crucial given the relative youth of cloud computing and the presence of unresolved technical questions (e.g., in security, data transfer, and management). The natural tension between these two needs may warrant focused programs for each in order to avoid lack of progress on either.
- Third, information and effort sharing across federal agencies considering cloud computing will be an important aspect of overcoming resistance to change. Explicit support should exist for shared technical mindshare, provider tracking/clearing, and case study reporting.

It is my hope that my testimony has helped to clarify some of the major technical matters and logistics associated with the idea of using cloud computing for federal IT. For non-technical practitioners, I recognize that digesting the concepts and evaluating the merits of cloud computing is no easy feat. Yet, I understand how important it is for members of the Committee to have trust and confidence in the IT directions taken by federal agencies, given the expense and mission importance of IT. As leaders in the realm of technology and innovation, please know that we at Carnegie Mellon University stand ready to assist you in dealing with technical questions as they relate to your efforts to craft sound public policy and oversee federal IT activities. We applaud your diligence in reviewing this specific matter.

Again, thank you for the opportunity to testify. I will be happy to answer any questions the Committee might have.