RAND CORPORATION

Perspective

Expert insights on a timely policy issue

Cost Considerations in Cloud Computing

Kathryn Connor, Ian P. Cook, Isaac R. Porche III, and Daniel Gonzales

Increasing costs for government technology acquisition programs, coupled with decreasing budgets, have the acquisition community looking to exploit new trends in data storage and processing. The Department of Defense (DoD) Chief Information Officer (CIO) describes that organization's current information technology (IT) state as one that is "duplicative, costly, and complex," a decades-long result of Components developing their own IT architectures to meet their individual needs.[1] One option is to consider shifting enterprise services to cloud-based computing.[2] Cloud computing offers the potential to reduce duplication and cost, especially in government data centers. The U.S. Office of Management and Budget (OMB) has issued guidance to reduce the number of data centers in all parts of the U.S. government.[3] As part of this effort, the U.S. CIO established a cloud computing strategy for the federal agencies to follow.[4] These agencies, including DoD, could save money on hardware, software, and the maintenance needed to keep pace with the technology refresh cycles in the commercial sector by sharing cloud computing resources.[5]

Cloud computing has garnered the attention of virtually all parts of the federal government as data and computer processing needs grow and budgets shrink. Despite this interest, insufficient guidance exists regarding how to estimate the costs—and potential cost savings—related to cloud information storage and processing. Such estimates are needed to identify, prioritize, and justify cloud resource needs, including for such DoD programs as Distributed Common Ground System–Navy (DCGS-N), Military Tactical Command and Control (MTC2), Distributed Common Ground System–Army (DCGS-A), and others.

Until more formal policies and best practices for acquisition of cloud computing–based systems are available, cost estimators and other acquisition analysts who have limited experience with cloud computing alternatives can benefit from this Perspective, which

provides context, background, and common terminology for cloud computing based on lessons learned from a recent estimate. Cost estimates should be coupled with effectiveness measures to quantify the impact of security-related and other risks, and assist decision-makers in selecting the most appropriate storage option.

Background

Recent research allowed us to better understand the cost drivers and important decisions that can affect the costs associated with moving to the cloud, to create the RAND Cloud Cost Model, and to develop a structure for comparing the cloud with other information storage and management alternatives. To support the analysis, we leveraged existing cost estimating structure based on operating and support cost guidance from MIL-STD 881C and Cost Assessment and Program Evaluation (CAPE), formerly the Cost Analysis Improvement group. After we completed our study, Agrawal and Manring proposed a work breakdown structure that would work for some cloud computing studies.[6] Our team was asked to consider moving a current IT-intense government program to cloud architecture. The program houses large amounts of data in multiple file formats and is used by the defense and homeland security–related agencies, as well as other government partners.

The wide range of alternatives allows analysts to assess whether existing requirements, policy, and practices help or hinder cloud adoption.

The necessary sharing of information among partners; the potential savings; and the need to collect, access, and analyze data on a 24-hour basis makes cloud storage an attractive option. However, good acquisition practices suggest that alternatives need to be studied. We compared a number of alternatives, including the existing program, two commercial cloud options, and an external data center with virtualized servers. Our cost analysis considerations were informed by professional literature and by interviews with U.S. government program officers pursuing similar products, private-sector application developers, and data center and cloud computing providers. Government data, commercial data, and tools were used to estimate costs associated with program staff, other personnel costs, software development costs, and commercial web services. Finally, the team used actual costs, based on data availability, to validate estimates.

The wide range of alternatives allows analysts to assess whether existing requirements, policy, and practices help or hinder cloud adoption. One potential hindrance that was important in this project involved information assurance (IA) risks specific to moving to a shared cloud architecture. While some federal agencies have adopted security guidelines for moving to the cloud, DoD has yet to issue definitive cloud security requirements.[7] Many individual systems currently house highly sensitive data, and the security of data stored in the cloud is a major concern. Another RAND study has enumerated some of these issues as well as the divergence of related legal and regulatory frameworks, and has suggested ways to mitigate them.[8]

Here, we present the considerations for cost analysts to explore when federal agencies select a cloud-based alternative. This document covers basic definitions, background, structured considerations for cloud analysis, and findings from applying this prototype

to a federal agency program. The structured considerations consist of seven primary cost areas. We briefly put each consideration in context of a federal acquisition program, describe potential cost drivers, and offer questions for cost-estimating teams to address with technical experts. Questions for each area are followed by a specific cost driver; where appropriate, the anticipated direction of the cost relationship appears in parentheses. We conclude with an example comparison of data management and storage system alternatives and a discussion of findings from the recent project.

Technical Definitions

In this document we use some technology vocabulary, such as *data centers*, *cloud*, and several "**-as-a-service*" acronyms, which we will introduce briefly. A *data center* is a facility with space, power, cooling, and security that houses, operates, and manages computer systems such as servers and associated telecommunications. *Cross-domain solutions* (CDS) are technologies that allow information to be transferred between classified and unclassified networks.[9]

Virtualization enables hardware separation from software and can provide substantial benefits by enabling server consolidation and live virtual-machine (VM) migration. Live migration is an important tool for moving VMs across physical servers in data centers and clusters, which facilitates load balancing, fault management, and maintenance. Consolidating VMs on physical hardware (e.g., resource pooling) can reduce energy consumption and data center operations costs.[10] The Navy's Consolidated Afloat Networks and Enterprise program of record is an example. By making each computer or server more efficient, fewer are needed on a ship.

Cloud elasticity enables computing and storage resources to be elastically provisioned and released, so cloud tenants can scale up resources rapidly to meet demand and then release them so they can be used by other tenants.[11]

According to the National Institute of Standards and Technology (NIST) definition, a computing cloud system must provide resource pooling, rapid elasticity, on-demand self service, and broad network access and measured service. Table 1 provides a brief definition of each of the three basic service models.[12]

Structured Cost Considerations

1. Software Development and Maintenance

An examination of the existing software code, software licenses, and future requirements for a specific program of interest indicated that the existing applications were not optimized for a cloud environment and could not be supported by VMs. While these applications relied on common business operating systems, the requirements for speed and the desire to add applications from additional vendors made it clear that Linux might support faster response and a wider variety of the necessary proprietary applications. In addition, existing software licenses were tied to specific pieces of hardware. This made virtualization impossible without further software development. Programs considering cloud storage should discuss:

A computing cloud system must provide resource pooling, rapid elasticity, on-demand self service, and broad network access and measured service.

Table 1. Overview of Cloud Service Models

Models	Description	Examples
Software as a service	Provider hosts applications online that users reach via browser	Gmail, Microsoft Office 365, Cisco GoToMeeting, DropBox
Platform as a service	Provider has software platform in cloud	AWS Elastic Beanstalk, Heroku, Force.com
Infrastructure as a service	Provider gives access to computing resources like virtual machines, servers, storage, and load balancers	HP Cloud, Windows Azure, Rackspace Openstack, Amazon EC2, Softlayer

- What operating system is preferred (or required) if any? Does this allow for more competition? (increased competition, decreased costs)
- How much (if any) retrofitting needs to be done to bring current programs up to speed with the cloud provider? (increased amount of retrofitting, increased costs)
- What licenses can the cloud provider include? For instance, some providers have Linux and Microsoft Windows, while others support only one. (increased provided licenses, decreased costs)

Cost estimators should question whether the virtualization software will interact appropriately with other aspects of the development. Also, the upgrade path for the cloud provider may dictate updates (and thus spending) for the government side. Questions that should be discussed with the cloud provider include:

- How often has the cloud provider upgraded systems or licenses? (frequent upgrades, increased cost)
- What is the potential for vendor lock in? What additional steps are necessary to protect the program from this risk?
- What is the future feature development for the cloud provider?
- Do currently owned/operated program systems support these, or will there have to be significant rework as the cloud provider modernizes? (increased rework, increased cost)

The frequency of upgrades in hardware and software licenses combined will inform the software refresh cycle that is most appropriate to the program. Program software would need regular refreshes every 2–3 years to keep up with changes in the software; the path of hardware improvements in server storage and speed indicates a useful lifespan of four years. The cycle of technological development must be coordinated with the spending profile for the program, and the rapid cycle suggested by software upgrades and server improvement may conflict with the budgeting preferences of an agency that would rather plan for a longer lifespan for its technology investments.

2. Database Options

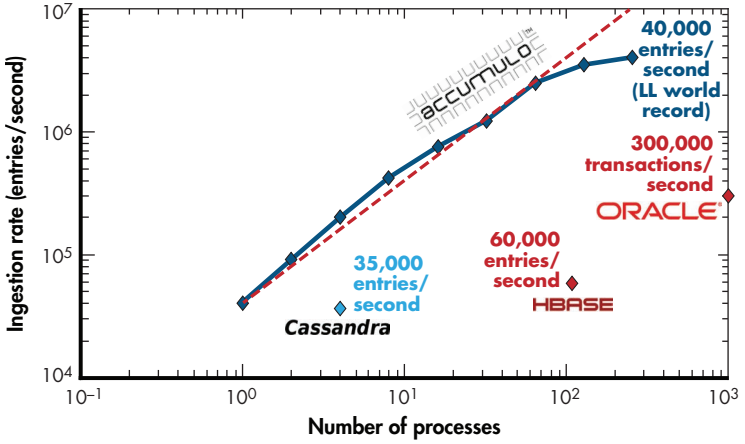
The potential promise that "big data" analytics holds for many enterprise mission areas makes relevant the question of the database types and data stores used. The costs associated with moving an enterprise application to the cloud depends upon the types of data

and databases involved. Many cloud implementations involve the transition of legacy data and databases from enterprise networks. Traditional enterprise systems that must meet high transaction rate requirements have traditionally been implemented using Structured Query Language (SQL) relational database management systems (RDBMSs). SQL databases are based on a strict tabular structure and use fixed data formats. The market leader in SQL RDBMS is Oracle, which produces the fastest RDBMS and custom high-performance RDBMS hardware. However, this proprietary software comes with significant licensing costs. Oracle's custom RDBMS hardware is also relatively expensive compared with commodity servers. In addition, Oracle RDBMS licensing costs typically increase with increasing database size. On the other hand, their products are well-known and it is relatively easy for contractors to hire company software specialists.

Oracle SQL databases typically run on single servers and consequently have size and scalability limitations. These limitations, although they only affect very large systems, have led to the development of a range of new distributed file systems and databases that have better scalability properties than traditional SQL databases. Hadoop is a widely used example of an open-source distributed file system that includes an algorithm for parallel processing of extremely large sets of data. Many systems exist that extend or supplement Hadoop—such as Apache Accumulo, which provides a highly granular mechanism for managing security and access control within a distributed file system.

Many of these so-called "NoSQL" databases have advantages over traditional SQL databases for large-scale applications. They can be used with unstructured data, including raw documents and "untagged" data. Many are open source and do not require licenses.

Figure 1. Relative Accumulo Database Performance



SOURCE: Jeremy Kepner, Christian Anderson, et al., *D4M 2.0 Schema: A General Purpose High-Performance Schema for the Accumulo Database*,IEEE High Performance Extreme Computing Conference, MIT Lincoln Laboratory, Lexington, Mass., 2013.

Furthermore, these distributed file systems and NoSQL databases do not require specialized hardware and work well on commodity servers found in cloud computing systems. A potential drawback to some of these software code bases is that contractors must be able to hire software experts familiar with these open-source code bases. Such experts are now in high demand.

Figure 1 compares the performance of a number of open-source databases against Oracle for a "graph-like" data set (which can be represented as a sparse matrix of connected vertices and arcs). One can see that the Accumulo database, when implemented on Hadoop, has a data ingestion rate significantly higher than that provided by Oracle. However, it should be noted that these performance results are specific to a graph-like datastore. RDBMSs perform better when the data used is densely packed and naturally

Some data center providers suggest not virtualizing the database portion of the new system to avoid this redesign, but splitting functions across a government-owned enterprise network and a cloud-based system could involve additional communications and security costs.

falls into tabular forms, and when the database is queried at a high transaction rate. These results indicate that the cost implications of moving to a cloud depend critically on two factors: the type of data that is to be processed in the cloud and how it will be processed (i.e., the database and file system used to support this processing).

Many government and commercial entities are considering designs that utilize Hadoop and other distributed file system technologies.[13] One key program consideration should be to weigh the cost of continuing with current database providers against the cost and time that will be required to modify existing data structures to effectively use new distributed file systems and distributed databases offered by open-source programs.

Some data center providers suggest not virtualizing the database portion of the new system to avoid this redesign, but splitting functions across a government-owned enterprise network and a cloud-based system could involve additional communications and security costs. The price structure dictated by a service provider can provide insight into whether splitting the database from the cloud environment is feasible. The price can vary based on speed, size of storage, number of uploads, number of downloads, number of regions, messaging between VMs, and a variety of other factors. Valuable questions to ask during cost estimating include:

- How much of the whole system can/should be moved to the cloud?
- What portions of the software can be bought as a service? (increased dependence on commercial services, decreased procurement costs, potential increase in sustainment costs)
- What parts of this system, if any, need to be hosted by the government? (government host, increase in fixed costs)
- What types of users are there for the system? (increase in user types, increased cost)
- Does this data management and storage style work well for this type of user?
- Are the license costs associated with the software sustainable for the program?
- What queries and reports will need to be redesigned? (increased queries, increased cost)

3. Hardware and Communications

Cost and size estimates of hardware are very different for traditional data centers, government data centers, and commercial cloud provider environments.

When a government program runs its own data center, it controls the type and refresh rate of hardware. When using another facility, program staff must translate requirements into generic units of VMs that may be based on a generic server specification, or may require detailing a number of servers, number of CPUs/server, gigabytes of RAM, gigabytes of disk space, and other data storage requirements. As each data center and cloud provider does this

differently, cost estimators will have to reach out to providers to understand the best way to provide computing requirements data for a rough order-of-magnitude cost. On top of these basic options, there are often options to lock in lower costs with longer contracts or opportunities to acquire different levels of service quality. Other hardware issues we considered include update frequency, continuity of operations (COOP) requirements, and communications costs.

In most situations where cloud computing is being considered, it is desirable because it allows for resource pooling. The government can benefit from not having to size its data center to peak loads. Because of data sensitivity and consistent utilization, we explored dedicated servers from commercial providers for this report. Dedicated hardware would be cost-prohibitive when utilization is inconsistent (i.e., many peaks and valleys in demand), but if the program is expected to use a large percentage of the estimated processing power at all times, then the servers can still be affordable in the cloud context.

Alternatives that take advantage of the cloud may still require the government to keep abreast of hardware trends. As cloud providers upgrade their hardware, this may require program staff to update any government-furnished equipment (GFE) that interacts directly with the service provider's equipment. The oldest, least-powerful part of the system frequently restricts the top speed of transfer and processing. The burden of upkeep, then, resides with the government, even when the preponderance of hardware and software maintenance is outsourced. Proper planning and budgeting for an upgrade path is crucial. This is particularly important if the aforementioned database is not virtualized.

Another program driver was the second site to provide COOP capability. While this requirement is common in the traditional data center world, it is less frequently used in cloud environments because the cloud can easily shift from one set of servers to another. The way that data are stored and managed is different, and often programs decide that they do not need full COOP capability because the cloud allows for sufficient redundancy. The requirement for a COOP location nearly doubles the amount of hardware needed by a cloud provider when compared to the program maintaining its own data center. Estimators should identify the strictness of requirements for continuity in their programs.

Finally, each alternative had different combinations of communications support. Cost estimators need to work with software engineers to understand the bandwidth necessary to support the system. When the program manages its own data center, it needs to consider both unclassified and classified connectivity at its main and COOP location, as well as between the two sites. In contrast, alternatives at the government data centers meant that communications costs were included and classified connectivity was already provided. For a commercial cloud provider, the basic communications lines

Alternatives that take advantage of the cloud may still require the government to keep abreast of hardware trends. As cloud providers upgrade their hardware, this may require program staff to update any government-furnished equipment that interacts directly with the service provider's equipment.

The security of a cloud is potentially challenging to ensure because data center systems may not be under the physical control of the government agency.

would be included, but the government would have to arrange for classified connection to its CDS and unclassified connectivity from that location to the private cloud provider. While the Cost Analysis Requirements Description (CARD) outlines many of these factors for a traditional data center, it may not be sufficiently detailed to support a virtualized or cloud computing estimate where requirements are shared among the government and contractors.

Cost estimators will need to find the answers to the following hardware questions:

- What is the best way to characterize the number of VMs required by the program?
- What level of utilization does the program expect? What will peaks and valleys in usage look like? What will be typical utilization?
- If computing resource requirements grow, then when do they become unsustainable (by provider, in terms of cost, software limitations)? What other scalability issues may affect the program?
- How will government furnish equipment interfaces with commercially provided computing capability? Will updates by cloud providers affect updates in GFE? (increased updates, increased GFE obsolescence potential)
- How will moving to the cloud affect the hardware license costs?
- Will the program need COOP capability in all alternatives or just noncloud data center options? (required COOP, increased costs)
- What communications lines will the program be responsible for, as opposed to the service provider? (increased communications by provider, increased costs)

4. Security and Privacy

The security of a cloud is potentially challenging to ensure because data center systems may not be under the physical control of the government agency. The challenge for DoD acquisition programs in particular is that DoD policy for clouds is still in development. The DoD CIO has committed the DoD to leverage FedRAMP. In addition, DoD CIO is updating and aligning DOD IA policies, IA controls, and processes with those used across the federal government.[14] DoD is taking a cautious approach as it works to fully understand the challenges and establish the appropriate risk mitigations.

Our discussions indicated that a clear guide for commercial cloud compliance with the DoD Information Assurance Certification and Accreditation Process (DIACAP) does not yet exist and therefore it may be more of a challenge than with data hosted in-house. Therefore, cost estimators should incorporate significant uncertainty around supporting the DIACAP process and other IA and testing requirements.

- What data can be allowed outside the government walls?
- What are the risks for the program if it is unable to access the data due to provider problems?
- Consider data access permissions: What has to be developed to allow different government users variable access to data that is stored in the cloud? How complex will this be for the software developer? (more permission types, increased costs)

- Are there security concerns about data being mixed on servers with nongovernmental data? (secured data, increased costs)
- Is there a reason to investigate buying dedicated servers with the commercial provider because resource pooling is undesirable for some processes or data? (dedicated servers, increased costs)
- If sharing computing space/power with another program, can data be fenced off or managed by U.S. citizens only? (U.S. citizens only, increased costs)

5. Data Compatibility and Migration

While not inherently cloud-focused, data compatibility and migration is a major cost driver for programs.[15] In our research, one option involved combining government clouds to include multiple programs, so that operation, support, and (to some extent) costs were centralized and redundant services minimized. The concern is the kinds, size, and quality of data. Requirements for data completeness will either default to the program with more lax standards (a.k.a. "dirty" data, where fields contain errors, are empty, or may reference multiple categories) among the contributors to the database, or compliance with highest requirement will force extra work on all but the referent program. Being robust to different data quality and standards is a partial solution, but is highly dependent on the ability of other tools to manipulate such data. Support tools may have to be dramatically recoded, making this an issue that extends beyond simple choice of storage formats. The issue is compounded when cloud providers host those tools and are meant to support multiple programs. To understand the implications of data compatibility and migration, the cost estimator should interview system users and program staff on:

- What data standards exist for the program and how often are they updated? (frequent updates, increased costs)
- How many lines of code or hours of labor were required to update the existing code to new data standards in the past? (increased effort, increased costs)
- What costs are associated with making one system support multiple standards?
- How will the program prioritize data from different sources?
- Will more personnel be required to deal with data compatibility in some alternatives more than others?
- What does the data migration plan entail? What risks could affect the resource requirements?

6. Classified Computing and Cross-Domain Solutions

Classified communication networks can drive significant cost in any government IT system, including the cloud. While there are a growing number of commercial providers that are supporting the intelligence and military communities with cloud services, a key question for a program is whether they want the cloud provider to support classified communications or data. The data are unclassified,

While there are a growing number of commercial providers that are supporting the intelligence and military communities with cloud services, a key question for a program is whether they want the cloud provider to support classified communications or data.

The kinds of program personnel required for a cloud program are very similar to those required of any other information system.

but often users in the field have better access to classified networks. One solution was to keep a CDS at a government facility to step information up and down from the classified networks. This would prevent having to pay for classified communication lines in the service-level agreement and remove the requirement for cleared personnel to work with the system. CDSs have posed significant risks to past programs because of known challenges associated with spillages. In some cases, it is less costly to create a new network than to shift between classified and unclassified data, especially if there is a large amount of unstructured data.[16] Cost estimators will need to find the answers to the following hardware questions:

- Will any of the data be classified? (classified, increased costs)
- Will the system need to transfer data between unclassified and classified networks? (transfer, increase costs)
- What are possible locations for hosting a CDS?
- Does an existing CDS meet the needs of the program?
- Is the data being exchanged structured, unstructured, or both? (unstructured, increased costs)

7. Personnel

The kinds of program personnel required for a cloud program are very similar to those required of any other information system. During the cost estimating process, we found there were some potential personnel savings. First, a program could significantly reduce—though, crucially, not eliminate—the number of system administrators associated with the operation of servers and networking devices, since those would be provided under the service-level agreement. Second, programs may be able to reallocate staff time previously used to operate and maintain the program's assets to realize further savings.

Other personnel impacts are worth consideration. Programs often have a diverse user base, including trained specialists, other government agencies, and military personnel in the field. The diversity of this group indicated that the program needed staff to support queries of the system by these various user groups, regardless of the type of data center or cloud arrangement. At the same time, the staff needs to provide technical oversight and be informed purchasers of cloud products. While the cloud provider is managing facilities, server maintenance, some licenses, and hardware disposal, the government needs to have strong contract oversight to ensure that the cloud system is meeting the needs of the program. The previously mentioned COOP capability also affects the number of people who need to be on staff to manage a second round-the-clock operation. Cost estimators who are working with programs considering a cloud alternative should ask:

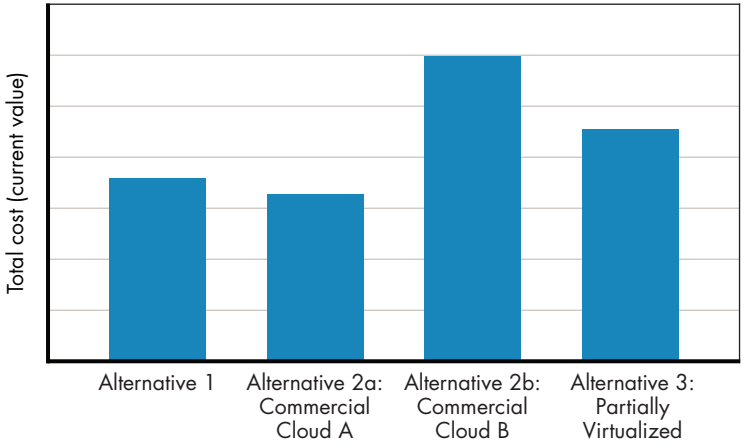
- How will the quantity of system administrators change across alternatives? (increased number of sysadmin, increased costs)
- What are the appropriate size, technical expertise, and experience level of the contracting staff? (larger contracts staff, increased cost)
- How will moving disposals to the service provider affect the tasks handled by the logistics staff? (government disposal responsibility, increased cost)

- Will technical requirements for COOP change with a cloud solution? Will they affect staff for round-the-clock operations at a secondary site?

Findings

This study allowed us to compare the costs of the existing program that is based on government-owned and -operated hardware to the costs of several alternatives, including incremental improvements to current assets (alternative 1), shifting to one of two commercial vendors for a hybrid cloud option where the majority of the system is in the cloud except the CDS (alternatives 2a and 2b), or selecting a government data center where the majority of the servers would be virtualized, except the database (alternative 3). We can compare the relative costs of the various alternatives in Figure 2, which is based on results from the RAND Cloud Cost Model.

Figure 2. Cloud Systems Are Not Universally Higher- or Lower-Cost Alternatives



SOURCE: RAND analysis. Proprietary data prevent inclusion of magnitude of costs.

Commercial Clouds Do Not Always Create Savings

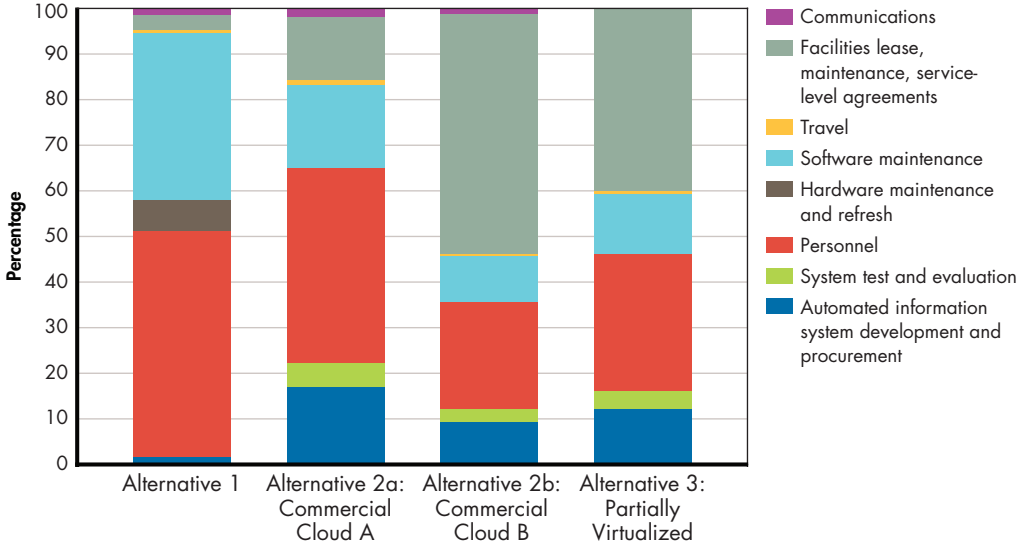
Figure 2 demonstrates the differences that can arise even when the overarching framework (in this case, utilizing a commercial cloud vendor) is the same. Commercial cloud vendor A was able to support the same size of data center for a little more than half of the cost of vendor B. Most importantly for future cost-estimating practices, there is no evidence to guarantee that a cloud solution (Alternatives 2a and 2b) will be cheaper than a more traditional data center in Alternative 1, or a partially virtualized data center (Alternative 3). Such a claim is frequently made about moving to a cloud-based system, but cannot be taken at face value as it depends heavily on program requirements and the provider's pricing structure. Figure 3 demonstrates in more detail the constituent costs of the alternatives, and is based on results from the RAND Cloud Cost Model.

Costs Shift from Hardware to the Service-Level Agreement

While inclusion of proprietary data does not enable us to share the magnitude of results, cost estimators may benefit from observing how individual cost elements change as a percentage of the total program cost across the alternatives. Three cost elements—site activation, training, and disposal—are of such small magnitude that they are not visible in the chart.

In Figure 3, the cost of facilities, maintenance, and service-level agreements change dramatically across alternatives and have a profound effect on total cost. At a high level, software development costs are very similar regardless of the environment chosen (except for Alternative 1, where no new development is included). We found the software costs for our reference program were more dependent on the selection of commercial off-the-shelf software

Figure 3. Variation in Percentage by Cost Element for Alternatives: Program-Level Choices Can Mitigate Claimed Cost Savings of Cloud-Based Systems



SOURCE: RAND analysis. Proprietary data prevent inclusion of magnitude of costs.

selected as the backbone of the system (in Figure 3, we show results assuming a primarily custom effort for the commercial cloud alternatives and Alternative 3).

The majority of the difference in software maintenance costs stems from licenses for software, rather than development. Facilities, leases, and service-level agreement costs rise when using an outsourced data center provider. Hardware disposal costs decrease for the program when the data center is outsourced (Alternatives 2a, 2b, and 3), because the service provider covers disposal through their service-level agreement, rather than the program using standard military disposal through the Defense Logistics Agency. Personnel costs decrease from Alternative 1, because fewer system administrators are on the program staff. These personnel become the responsibility of the cloud or government data center provider.

Discussion

Cloud computing is often presented as an all-or-nothing alternative to traditional ownership of massive amounts of hardware. Even when "*-as-a-service" products are considered, they are actually minor pieces of a larger program position, that may not be sitting close to one pole (purely contracted cloud service) or the other (physical ownership of all mechanicals and software). Sophisticated programs may not be able, or even need, to face such a binary

choice, and cost estimates cannot be developed as if they do. We found that some, but not all, existing commercial off-the-shelf products allowed for virtualization if versions are updated without new license costs. This may not be the case for all programs, but virtualization should become more common, and therefore more affordable. Major portions of an IT program could be virtualized, while the storage of sensitive information, idiosyncratic application processing, or particular digital products of a process could be kept under strict control of an agency if required for security. When virtualization of databases or applications is very costly because of license costs or code development requirements, the government may choose to select a combined path. Migration from current database providers may lower license costs, but can incur significant upfront data migration and programming costs.

As the reliance on massive amounts of data increases in government functions, the need to consider cloud and in-house hardware solutions will only grow. Rigorous, defensible estimates require identifying the associated drivers and risks. Further research on cloud-specific cost estimating structure elements would be valuable to support cloud cost analysis policy development and help ensure analysis is of sufficient rigor. Current DoD cost modeling lacks good examples that consider the range of options for a program in the cloud, combined with the future costs of expanding and maintaining program-owned hardware and software. The points we have raised are an important, though by no means exclusive, set of prominent concerns when considering cloud, traditional, and potential partial cloud solutions.

Notes

[1] DoD CIO, *DoD Cloud Computing Strategy*, June 2012.

[2] Nicole Blake Johnson, "Is the Cloud Overhyped? Predicted Savings Hard to Verify," *Federal Times*, October 8, 2012. As of July 3, 2013: http://www.federaltimes.com/apps/pbcs.dll/article?AID=2012310080001

[3] Vivek Kundra, *25-Point Implementation Plan to Reform Federal Information Technology Management*, Washington, D.C.: The White House, December 9, 2010. As of September 4, 2013: http://www.dhs.gov/sites/default/files/publications/digital-strategy/25-point-implementation-plan-to-reform-federal-it.pdf

[4] Vivek Kundra, *Federal Cloud Computing Strategy*, Washington, D.C.: The White House, February 8, 2011. As of July 3, 2013: http://www.whitehouse.gov/sites/default/files/omb/assets/egov_docs/federal-cloud-computing-strategy.pdf

[5] DoD, 2012.

[6] Raj Agrawal and Jennifer Manring, "Cloud Computing Starter Kit: Cost and Business Case Considerations," briefing delivered at the 2013 International Cost Estimating and Analysis Association Professional Development and Training Workshop, New Orleans, La., June 2013.

[7] Federal Risk and Authorization Management Program (FedRAMP). FedRAMP defines requirements for cloud computing security controls, including vulnerability scanning, and incident monitoring, logging, and reporting. FedRAMP has been adopted by the Department of Homeland Security. Accessible at http://cloud.cio.gov/fedramp

[8] Neil Robinson, Lorenzo Valeri, Jonathan Cave, Tony G. Thompson-Starkey, Hans Graux, Sadie Creese, and Paul Hopkins, *The Cloud: Understanding the Security, Privacy and Trust Challenges*, Santa Monica, Calif.: RAND Corporation, TR-933-EC, 2011. As of August 16, 2013: http://www.rand.org/pubs/technical_reports/TR933.html

[9] The Intelligence Community is working on a similar capability called Cloud Security Gateway.

[10] Divya Kapil, Emmanuel S. Pilli, and Ramesh C. Joshi, "Live Virtual Machine Migration Techniques: Survey and Research Challenges," in Advance Computing Conference (IACC), 2013 IEEE 3rd International, 2013, pp. 963–969. As of January 28, 2014: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6514357

[11] Robinson et.al., 2011.

[12] More detailed definitions can be found in Peter Mell and Timothy Grance, *The NIST Definition of Cloud Computing: Recommendations of the National Institute of Standards and Technology*, Gaithersburg, Md.: National Institute of Standards and Technology, NIST Special Publication 800-145, 2011. As of July 12, 2013: http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

[13] Merv Adrian, *BYOH—Hadoop's a Platform. Get Used to It*, October 9, 2013. As of December 16, 2013: http://blogs.gartner.com/merv-adrian/

[14] DoD, 2012.

[15] Agrawal and Manring, 2013.

[16] Chad C. Serena, Isaac R. Porche III, Joel B. Predd, Jan Osburg, and Brad Lossing, *Lessons Learned from the Afghan Mission Network: Developing a Coalition Contingency Network*, Santa Monica, Calif: RAND Corporation, 2014, RR-302-A. As of May 19, 2014: http://www.rand.org/pubs/research_reports/RR302.html

About This Perspective

This Perspective is intended to help guide DoD and other federal government agencies in estimating the costs associated with acquiring cloud-computing capacity. It was commissioned by RAND Arroyo Center's Force Development and Technology program and is based on lessons learned from a U.S. Army sponsored analysis. The Force Development and Technology Program assesses technological advances and new operational concepts to improve Army mission performance by seeking opportunities in new technologies, evaluating alternative force structures, and identifying efficiencies from acquisition reform.

The authors would like to thank John Davis and Kevin Cincotta for their careful and very thoughtful reviews of the manuscript.

About the Authors

Kathryn Connor is a cost analyst who joined the RAND Corporation in 2009. Since then, she has worked in RAND Arroyo Center, the RAND National Defense Research Institute, and RAND Health. Her research has included the development of cost models for analyzing cost effectiveness, informing cost policy, improving property accountability metrics, assessing operations and sustainment of weapon systems, estimating technology development costs, and understanding the cost implications of applying robotics.

Ian P. Cook is a management systems analyst for the RAND Corporation. He has worked for numerous government clients in the areas of national security, research and development of technology programs, and information security. His particular focus has been on cost-benefit analysis and program evaluation.

Isaac R. Porche III is a senior engineer at the RAND Corporation. His areas of expertise include cybersecurity; network and communication technology; intelligence, surveillance, and reconnaissance; information assurance; big data; cloud computing; and computer network defense.

Daniel Gonzales is a senior physical scientist at the RAND Corporation. His areas of expertise include command, control, and communications and intelligence (C3I) systems; information technology (IT); and new IT system architectures, such as service-oriented architectures.

The RAND Corporation is a nonprofit institution that helps improve policy and decision-making through research and analysis. RAND focuses on the issues that matter most, such as health, education, national security, international affairs, law and business, the environment, and more. As a nonpartisan organization, RAND operates independent of political and commercial pressures. We serve the public interest by helping lawmakers reach informed decisions on the nation's pressing challenges. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors. **RAND**® is a registered trademark.

www.rand.org

© Copyright 2014 RAND Corporation

PE-113-A (2014)

CHILDREN AND FAMILIES

EDUCATION AND THE ARTS

ENERGY AND ENVIRONMENT

HEALTH AND HEALTH CARE

INFRASTRUCTURE AND TRANSPORTATION

INTERNATIONAL AFFAIRS

LAW AND BUSINESS

NATIONAL SECURITY

POPULATION AND AGING

PUBLIC SAFETY

SCIENCE AND TECHNOLOGY

TERRORISM AND HOMELAND SECURITY

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

This electronic document was made available from www.rand.org as a public service of the RAND Corporation.

Support RAND

Browse Reports & Bookstore

Make a charitable contribution

For More Information

Visit RAND at www.rand.org

Explore the RAND Corporation

View document details

Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND electronic documents to a non-RAND website is prohibited. RAND electronic documents are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see RAND Permissions.