CLOUD STRATEGY FOR HIGHER EDUCATION

Building a Common Solution

November 2014

Authors

Asbed Bedrossian
Beth Ann Bergsmark
Bob Carozzoni
Mike Chapple
Alan Crosswell
Lisa Davis
Guy Falsetti
Patton Fast
Ryan Frazier
Brad Greer
Jim Jokl
Charlie Leonhardt
Mark McCahill
Sharif Nijim
Scott Siler
Oren Sreebny
Bruce Vincent
Bob Winding
Steve Zoppi

Contents

Our Shared Vision

Higher education IT is in the midst of an exciting transformation. The economies of scale, resiliency, flexibility and agility provided by cloud computing are rendering the construction and maintenance of on-premises data centers obsolete. We believe that over the next decade, the availability and advantage of new technology models will result in a substantial decrease in the use of on-premises data centers. In this document, we outline a “Cloud First” strategy for higher education IT that moves from a traditional data center model to one centered on the public cloud and cloud-based services.

What are Cloud Services?

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction. It is more than simply using someone else’s data center, building on-premises virtualization capability, or shifting to managed services for applications. The cloud offers a radically different approach. It provides us with the opportunity, perhaps even the mandate, to transform our organizations from *builders of unique solutions* to *providers of IT services*.

Cloud services are commodities. Ranging from infrastructure building blocks to full application suites, cloud services benefit from economies of scale through common, integrated solutions that meet the needs of many customers. By offering their products on a global scale, cloud service providers are able to drive efficiency and develop innovative solutions in an unprecedented manner. Acting on our own, or as a consortium, higher education institutions simply do not have the financial or human resources to compete in this increasingly commoditized market.

Why the Public Cloud?

There are several different models of cloud services. In a public cloud approach, service providers offer software, platforms and/or infrastructure for use by the general public on hardware maintained by the service provider in that provider’s data centers. Virtualized on-premises environments, or “private clouds” take the opposite approach, pooling resources within an organization but not leveraging hardware or data centers shared with other customers. Hybrid and community cloud approaches combine aspects of public and private models, using some limited resource pooling.

Many institutions of higher education have already embraced the private cloud model, using virtualization technology to pool computing resources within campus data centers. In many cases, institutions run substantial portions of their infrastructure on private cloud virtualization platforms. We have achieved significant improvements in agility, cost and performance by migrating from physical servers to virtualized systems.

Embracing the public cloud offers higher education an even greater opportunity to improve. This transformation is the next logical pivot point in the history of higher education information technology. The capacity, resilience, agility, pricing models and staff development opportunities found in the public cloud offer unprecedented opportunity for improving the service provided to our campuses.

Capacity

Relative to the needs of any particular consumer, the public cloud provides a practically unlimited set of resources. Our institutions now have the ability to immediately provision compute capacity, storage and other services in extremely large quantities, consume those resources for as long as necessary, and immediately de-provision them when they are no longer needed. The public cloud model eliminates the need for over-provisioning resources to meet future demand. For example, developers can provision a multi-server environment to test new applications without having to go through traditionally long hardware procurement processes.

Resiliency

The infrastructure that powers cloud services is highly durable, often being spread across multiple geographic locations, alleviating concerns related to the reliability and disaster resiliency of on-premises data centers. IT architects designing public cloud solutions can easily deploy services across multiple data centers, dramatically boosting service reliability. For example, an institution’s website may be simultaneously hosted in data centers across North America and Europe, providing fault tolerance should one region go offline. The use of cloud-based resources allows off-campus users to continue to access institutional resources in the event of an on-campus outage.

Agility and Speed

Public cloud computing offers significant improvements in agility and speed of deployment over traditional, on-premises physical implementations. Servers and related infrastructure can be spun up in a matter of minutes rather than days or weeks. Applications running in the public cloud often use a rapid release model, deploying new features incrementally. Upgrades are much smaller in size than the traditional major disruptive upgrade of conventional software that happens every few years.

The use of public cloud resources also facilitates quickly processing computing workloads. For example, a researcher performing a complex data analysis can provision significant computing resources for a short period of time to quickly complete the workload. This is not economically viable in an on-premises model.

Pricing Model

The public cloud offers flexible financial options. The “pay as you go” model allows an institution to pay only for those services that are actually consumed and purchase them on demand. Under this model, the cost associated with computing resources may be terminated immediately when the resources are no longer required. In cases where computing requirements are predictable, public cloud vendors offer the opportunity to pre-purchase capacity at a substantial discount. In addition, public cloud vendors have a history of price decreases over the past decade as they have grown in size.

Cloud vendors also provide steeply discounted access to "spot instances" that allow users to harness excess capacity with the knowledge that their job may be terminated without notice. This option may be appealing for certain types of workloads.

Staff Development

The most precious resource an organization has is its workforce. While moving from an on-premises data center to the public cloud may have implications for staff, it can afford them the opportunity to develop new skills. The ability to easily test new scenarios in a public cloud model with low risk and low cost facilitates the development of a culture of creativity and innovation.

Cloud First Strategy

We propose that institutions adopt the "Cloud First" strategy shown in Figure 1. Under this philosophy, all new services are deployed in the cloud when possible. Concurrently, organizations may begin migrating existing services to cloud models on as ambitious a schedule as resources permit.

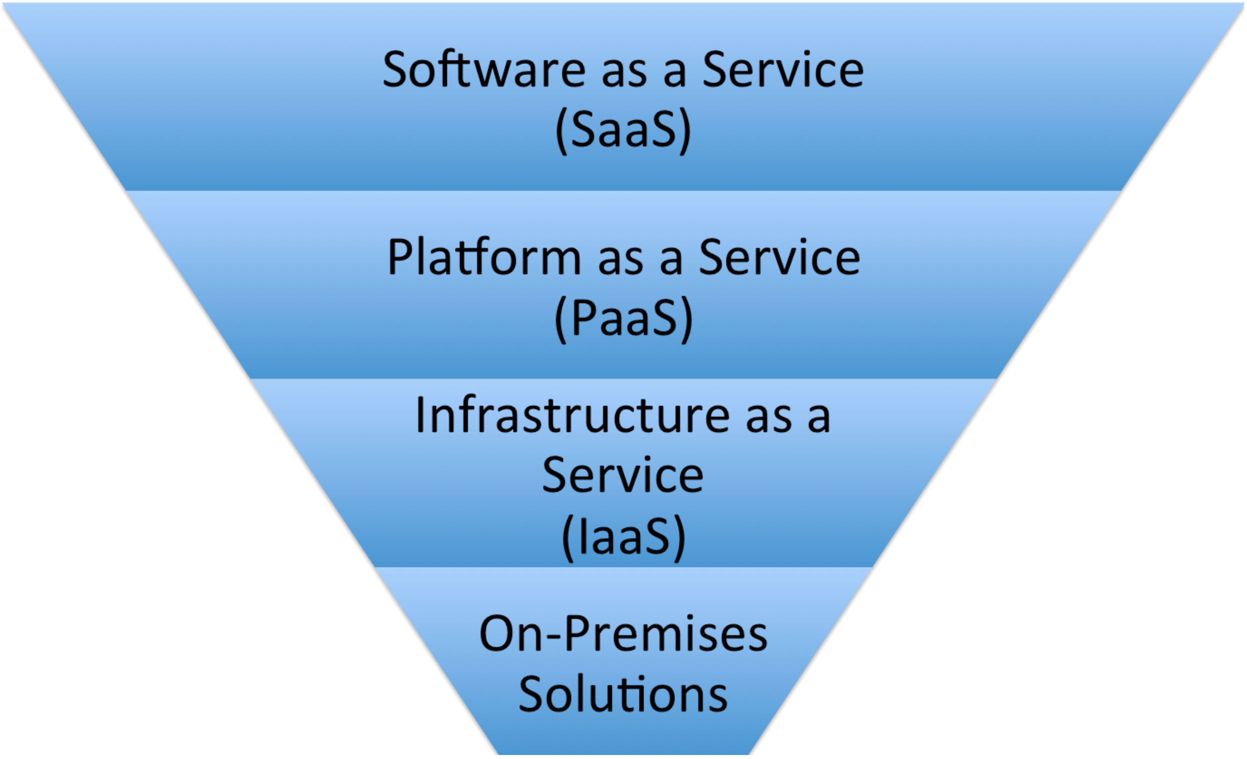

Software as a Service (SaaS)
Platform as a Service (PaaS)
Infrastructure as a Service (IaaS)
On-Premises Solutions

Figure 1: Cloud First approach to IT service deployment

The Cloud First strategy focuses the value of limited IT resources on delivering the most business value to the University. By using public cloud services to deliver commodity IT Infrastructure as a Service (IaaS) or Platforms as a Service (PaaS), we free up our staff to focus on developing and supporting non-commodity, differentiated services built upon those platforms. In the case of many application services, public cloud Software as a Service (SaaS) delivers greater capabilities than traditional application suites, and, in fact, are replacing many software vendors' traditional applications.

We further believe that institutions should consider cloud deployment options in the prioritized manner shown in Figure 1. SaaS deployments should be the model of choice, followed by PaaS options. If an institution must turn to solutions involving infrastructure managed by the institution, the first priority should be on adopting open source or commercial products and only custom-developing software when there is no viable alternative. In those cases, institutions should first seek to deploy solutions in an IaaS environment and only turn to on-premises options when that is not feasible.

Further elaborating on this approach, we offer the following guiding principles for consideration:

- Cloud services should be the first option for any new services or when evaluating alternatives or revisions to current services. Identify reasons for not moving to the cloud rather than why you should move. Factors such as the age of applications or the data center should be taken into consideration.

- When evaluating applications or platforms, favor those that can be run on cloud infrastructure, even if the initial implementation will be run on premises. This will help future-proof that application.

- When evaluating cloud services, aim to select services that run as high up the stack as possible (see Figure 1). This means selecting SaaS over PaaS and PaaS over IaaS. This enables the most effective use of IT staff resources and allows for taking full advantage of vendor architectures and support.

- Proper procedures must be taken for ensuring the security of University information and complying with all applicable regulations. By developing and applying rigorous data classification and security standards, appropriate technical and legal safeguards can be established. University counsel and procurement services should be involved in all cloud service agreements for University services, even in cases where no payment is involved.

- Consideration should be given to integration with existing on-premises and other cloud services, including identity management, networking, storage, etc. Not all cloud implementations require integration, but decisions to not integrate should be made deliberately. Preference should be given to systems that have common functional integration capabilities, such as web service APIs.

Cloud Services Value Proposition

Universities have been making progressively greater use of cloud services over the past five years. Email and collaboration tools are broadly adopted. We believe that the current client presents us with the opportunity to begin much broader adoption of a wider variety of cloud offerings. The challenge facing higher education CIOs is communicating the advantages, challenges, and risks associated with a cloud-first strategy to university leadership. In so doing, CIOs have to work within their institution to develop a business case in order to support and fund this exciting transformation. The move to the cloud requires a transition from a primarily on-premises, capital expenditure model to a mixed cloud/on-premises expenditure model that is predominantly operational.

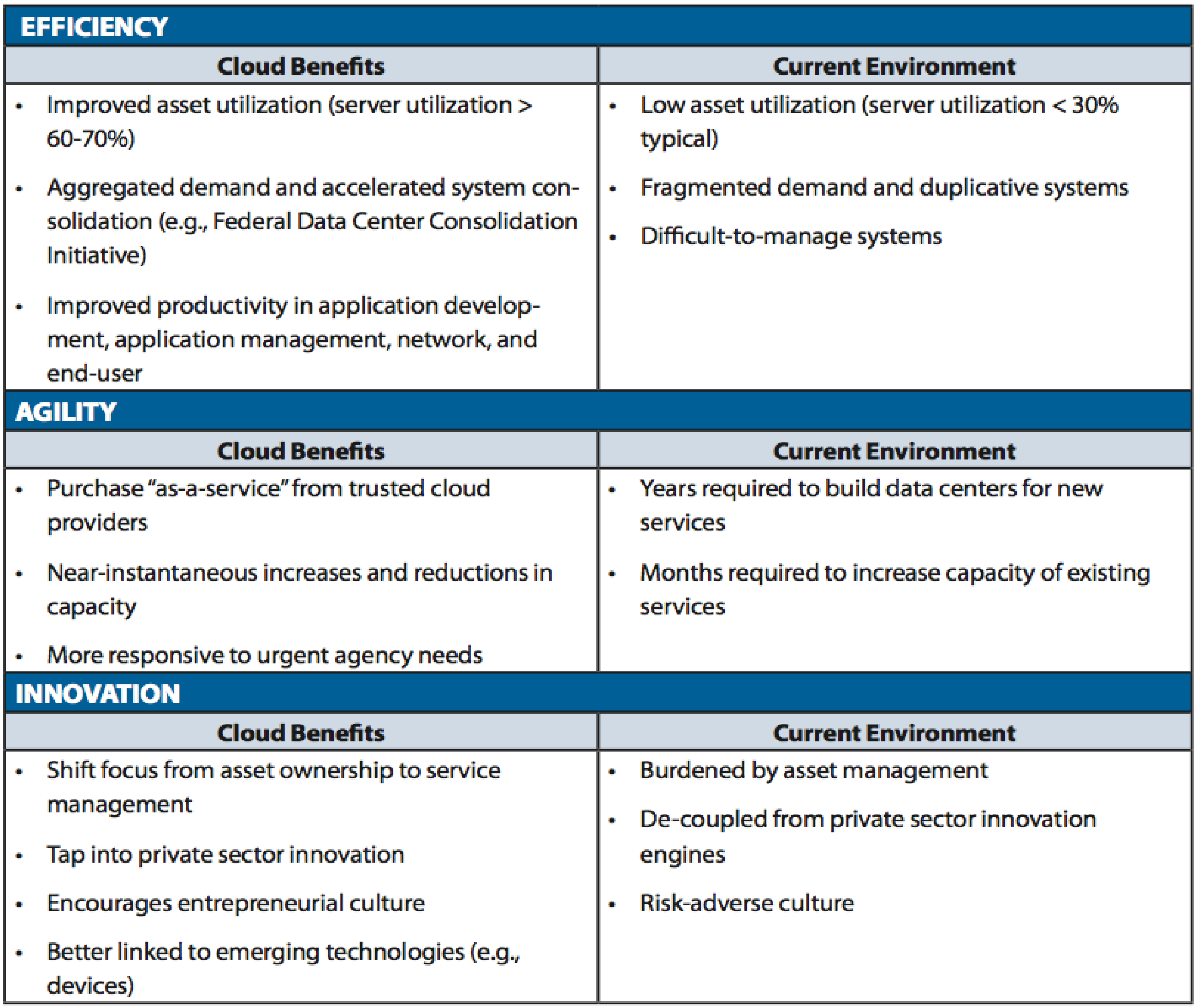
As we make this transition, higher education can benefit from the course charted and lessons learned by both the federal government and the private sector over the past five years.

Public Sector Adoption

In September 2009, the federal government announced its Cloud Computing Initiative. The basis for this initiative is the realization that through effective use of cloud computing, the government can capitalize on opportunities to reduce waste, increase data center efficiency and utilization, and lower operating costs.

In February of 2011, the *Federal Cloud Computing Strategy*[1] was published from the office of Vivek Kundra, U.S. CIO, announcing a Cloud First Strategy. Kundra identifies 25% of the $80 billion federal IT spent as candidates for moving to the cloud. That document does an excellent job of concisely summarizing the value proposition the cloud presents, shown here in Figure 2:

[1] Kundra, Vivek. "Federal Cloud Computing Strategy." (n.d.): n. pag. 11 Feb. 2011. Web. 29 Aug. 2014. <https://cio.gov/wp-content/uploads/downloads/2012/09/Federal-Cloud-Computing-Strategy.pdf>.



EFFICIENCY	
Cloud Benefits	**Current Environment**
• Improved asset utilization (server utilization > 60-70%) • Aggregated demand and accelerated system consolidation (e.g., Federal Data Center Consolidation Initiative) • Improved productivity in application development, application management, network, and end-user	• Low asset utilization (server utilization < 30% typical) • Fragmented demand and duplicative systems • Difficult-to-manage systems
AGILITY	
Cloud Benefits	**Current Environment**
• Purchase "as-a-service" from trusted cloud providers • Near-instantaneous increases and reductions in capacity • More responsive to urgent agency needs	• Years required to build data centers for new services • Months required to increase capacity of existing services
INNOVATION	
Cloud Benefits	**Current Environment**
• Shift focus from asset ownership to service management • Tap into private sector innovation • Encourages entrepreneurial culture • Better linked to emerging technologies (e.g., devices)	• Burdened by asset management • De-coupled from private sector innovation engines • Risk-adverse culture

Figure 2

Private Sector Adoption

Not surprisingly, the private sector is adopting the cloud at a rate which far outpaces both higher education and the federal government. Private sector experience bears witness to aggressive planning and significant realized cost savings. According to Stephen Orban[2], CIO of Dow Jones & Company, the firm stands to realize a $100 million dollar savings by converting 75% of on-premises infrastructure to the cloud over the course of three years. This strategy will allow the company to consolidate its forty existing physical data centers into six, with the remainder of their workload handled on Amazon Web Services.

[2] "AWS Case Study: Dow Jones." *Amazon Web Services, Inc.* N.p., Nov. 2013. Web. 29 Aug. 2014. <http://aws.amazon.com/solutions/case-studies/dow-jones/>.

According to Ashish Kelkar[3], Director of Infrastructure Strategy and Analytics at Facebook, an on-premises infrastructure becomes operationally viable only when public cloud expenditures exceed $300,000 per month. Kelkar's philosophy focuses on the ability to devote organizational energy to core mission activities while taking advantage of the commoditization of the data center. By way of comparison, it would cost approximately $60,000/month, or one-fifth of Kelkar's cost threshold, to operate a 1,000 server, 500 TB environment in the public cloud.

In another example of public sector adoption, Lionsgate acknowledges that there is a financial benefit to public IaaS in running its SharePoint and SAP workloads. According to Theresa Miller[4], Executive Vice President, Information Technology, Lionsgate will save approximately fifty percent by using the public cloud compared to a traditional hosting facility.

Recommendations for Higher Education

Universities are already reaping the benefits from the use of cloud services. The pace at which each university broadly adopts the cloud is going to depend upon institutional readiness, risk tolerance, existing contractual commitments, and budget constraints. The benefits of cloud adoption are already being realized on our campuses, by the United States government, and in the private sector. Many of the hard lessons have been learned, paving the way for a smoother adoption experience by higher education.

Over the long term, we believe that using the cloud is most likely to result in significant savings when compared with on-premises solutions. However, cost savings are not one of the primary motivations for this move and we would endorse a Cloud First strategy for higher education even if it were only cost-neutral, when compared to on-premises computing. For practical examples from within our community, please refer to the financial analyses in the Use Cases section of this document.

We believe that current pricing models for IaaS providers contain room for additional price cuts and could create situations where the direct costs of operating a service in the cloud exceed the direct costs of running it on-premises. Because of this, we expect cloud prices will continue to fall. Amazon Web Services currently dominates the IaaS market, and is facing growing competition from Google and Microsoft. When Google dropped the cost of storage in March 2014, Amazon responded the following day with a 50% drop in its storage pricing.

Our business cases must account for the existing significant investments our institutions have made in on-premises data centers. The cost savings associated with shrinking the physical footprint will only be realized over the long term. Due to the local considerations at each university, these benefits are difficult to calculate and are institutionally dependent. Moving to the cloud will require both an initial and ongoing investment in order to train staff and fund initial exploration. University CIOs will likely need to allocate funds for this purpose within ongoing budget constraints.

[3] Kelkar, Ashish. "When, If Ever, Should You Move Out of the Public Cloud?" Defrag 2013. Omni Interlocken Resort, Broomfield, CO. 30 July 2014. Address.

[4] "AWS Case Study: LIONSGATE." *Amazon Web Services, Inc.* N.p., n.d. Web. 29 Aug. 2014. <http://aws.amazon.com/solutions/case-studies/lionsgate/>.

In order to fully leverage cloud technology, senior IT leadership will need to elevate the financial discussion to the university level and remove cross-subsidies that distort the financial incentives of campus technology decision makers. Effective IT governance processes are another possible avenue for minimizing the impact of campus financial distortions.

While authoring the business case, each university must weigh the costs and benefits of cloud computing within their unique institutional context. Although we believe that cloud prices continue to have room to fall, the intangible benefits of the cloud (e.g., geographic diversity of services, fault tolerance, enhanced security and compliance, and automation) can make it more attractive.

The pace of cloud adoption will be governed by the tradeoff between the value of those intangible benefits and the speed of price reductions as global competition in cloud services expands. IT leaders should carefully consider cloud trends and likely future outcomes before making new investments in core infrastructure such as data center, large scale storage or compute infrastructure, and other technology with a longer-term return on investment.

Culture Change

Shifting to a Cloud First strategy is a major organizational transformation that, to be successful, requires a significant culture change. Many deeply technical, highly valued employees can feel threatened by this change. We must appreciate that, while physical complexity decreases with the adoption of cloud services, logical complexity increases. A tremendous amount of intellectually challenging work remains to successfully engage and motivate our staff.

It is the role of IT leaders to energize the organization, articulate the strategy in a publicly visible way, and ensure that team members see a personal path forward. Although jobs may change, there is enough automation and integration work to keep even the most technical staff highly challenged and engaged. Given the expected rate of change in IaaS, we will expect greater agility from our organizations as we respond to and take advantage of shifts in the cloud IaaS landscape.

Cloud Mindset

The cloud allows us to adopt a more agile and flexible approach to IT. The ease with which infrastructure configurations can be created and tested facilitates the ability for organizations to fail quickly, evaluate outcomes, refocus, and adjust. This is vital because we are in the early days of IaaS, and we expect the continued evolution of cloud vendors and their offerings. In the future, we may wish to move workloads between cloud vendors and avoid vendor lock-in. A "cloud mindset" implies that our systems are designed for transparent migration to take advantage of what will be increasingly competitive pricing between cloud vendors.

Designing for cloud server and storage portability is significantly different from optimizing systems to run using an on-premises infrastructure. In addition, some of the tools and infrastructure we have depended on for on-premises solutions are ill-suited for IaaS. For example, an on-premises solution would likely assume that user storage space should be provided via a centrally managed file system mounted by the server. This is unlikely to be a viable approach for a cloud-centric system, particularly if we expect to migrate between cloud providers.

Consumer Mindset

Perhaps the biggest organizational mindset shift applies to technologists. Traditionally, IT staff think of themselves as inventors and creators. In a cloud environment, we must become innovators and integrators. We become consumers at lower levels of the infrastructure stack. For example, email has become a commodity, to the point where most institutions have moved to hosted email over the past five years.

As we adopt the cloud, we can increase our focus on tasks that contribute more directly to our institutional missions. We can outsource non-core competencies. For instance, universities do not write their own statistical software packages or operating systems because other organizations can do so more efficiently and effectively. The same is true when it comes to building infrastructure-laden data centers.

The People Side of Change

The adoption of a cloud-first strategy can unsettle our teams, and it is important that IT leaders define trajectories for staff where there might otherwise be fear, uncertainty and doubt. Typically, the people whose career path will be most affected by the transition will have the strongest objection to the adoption of IaaS.

For example, good system administrators/operators develop an arsenal of scripts over time to automate repetitive tasks. As the ability to increase the ratio of systems managed per person increases, there will be a heightened focus on the ability to automate. Leadership needs to recognize this and invest in their system administrators, updating their skills to include modern provisioning, configuration, orchestration, and reporting tools.

Career paths and the people side of change is something each institution will have to address. While some individuals will view a move to the cloud as a potentially devastating end to their IT career, they have an opportunity to make a transformational change that will help prevent them from becoming obsolete. IT professionals have seen similar transitions before in such areas as keypunch operation, mainframe support and Novell administration. As the industry grows and matures, individuals can re-train or find ways in which to apply previously developed skills and abilities in new situations and environments.

Moving to the cloud may reduce the total number of roles in our IT groups, but increase the need for highly skilled “full stack” cloud engineers. Individuals will be needed who can pinpoint, diagnose and resolve performance bottlenecks in applications. Organizational leaders must foster a culture of curiosity and innovation, encouraging and rewarding staff for developing innovative uses of the cloud.

Possible career paths for those most affected by cloud transitions might include a System Administrator or Developer becoming a DevOps Engineer, or Product or Layer Specific Engineer developing into a Cloud Engineer who can work up and down the stack managing a greater level of abstraction

End User Impact

As the majority of IT services migrate off campus, users who are accustomed to having systems tailored to meet the most arcane business processes may go through a culture shock. Suddenly, product features are at the behest of the service provider and the broader community. We can request enhancements from vendors, but they may not act on those requests. If they are willing to make institution-specific customizations, they will likely do so at significant cost to the requesting institution.

In many cases, users must be connected to the Internet for cloud services to work. Some users prefer to work offline and others may find an online requirement challenging from a workflow perspective. The portability and interoperability of data and information in the cloud has different concerns: How much data will be stored? How fast does the data need to move, where and how often? How does “Big” data compare with “Not So Big” data?

Changes to cloud services happen so often as to be almost constant. New features and user interface changes frequently change as services evolve. This is primarily true in SaaS, not as much in PaaS or IaaS.

Users will also notice that cloud-based applications may have different performance characteristics than those hosted internally. Support and issue response will change as well if the support model of a cloud provider does not match the level of service we provided on our campuses.

The Service Automation Mindset

In order to make effective use of IaaS, an organization has to adopt an automate-first mindset. Instead of approaching servers as individual works of art with artisan application configurations, we must think in terms of service-level automation. From operating system through application deployment, organizations need to realize the ability to automatically instantiate services entirely from source code control repositories.

This shift requires embracing and investing in configuration management at the service level, not the component level. Thus, system administrators, system engineers, and developers need to be charged with service automation as a team. This is a departure for organizations where discipline-specific objectives are historically diverse. The key to success is to fully embrace infrastructure as code (software defined servers, storage, and networks) and to align and commit to invest in automation.

Complementing the shift toward automation of the system build process are developments that allow creation of "containerized workloads" which can be automatically deployed both on-premises and across diverse cloud vendors' infrastructure. For example, Docker containers provide a lightweight virtualization technology that bundles an application and OS together to insulate the workload inside the container from the system running it. By running inside a container you are not tied to the OS version of the vendor's cloud. The Docker containers are created via a script that configures both the operating system and application inside the container. Once the container has been built, it can be deployed across heterogeneous systems without modification or patching, a complete contrast with typical artisan system administrator practice.

Additional Considerations

While it off-loads traditional roles centered on system administration and design/build customizations to the vendor, SaaS presents new challenges for IT and business owners looking to leverage the move to streamline business processes, shut down vertically siloed data and processes, achieve holistic data-driven decision capability and take advantage of the new agility.

Here are some ideas to keep in mind when planning for a transition to the cloud:

- Central IT is often positioned as the **only neutral broker** to align new business processes across multiple domains, understanding the complexity of the configuration range, and applying security principles uniformly across the data integrations. In SaaS ERP offerings, the business process is in the system and is bringing forward the need for new IT/Business

Analyst roles with domain specific knowledge to bridge the gap.

- As the solution for a “gap” is no longer building a customization or modification, the **solution architect** who can look across multiple products, understand integration requirements, and construct bridges forward becomes critical.

- Old skills still apply, but in new places. Configuration management can be a complex role in many SaaS systems where **rapid release cycles and data dependencies** are critical. Configuration management, change control release tracking and other standard operating procedures that existed in operations areas are necessary in the application domains. Dedicated configuration management/control and testers are necessary to maintain data integrity.

- Legacy on-premises systems were not originally designed for numerous integrations at real-time speed. While some SaaS applications do present the pitfall of a cloud “black box”, others are embracing open APIs and push the data out into more and more feature rich extensions. This promotes the need for **data architects, integration and orchestration experts, and security analysts** focused on understanding the data model. Traditional integration specialists will broaden their view across multiple dimensions and need to work within a unified framework to leverage the capabilities. This also includes navigating dependencies on performance optimization, understanding licensing models, and developing security roles to move data correctly.

- **Performance optimization, performance tuning and predictive modeling** become critical to achieve the return on investment of moving to the cloud. Likewise, network engineering and routing become even more essential to reach the goal of friction-free data.

- As the quantity of relationships with cloud businesses increase, **vendor management also increases in importance** including areas of standardized contracts to address sensitive data issues; concerns about data ownership and transportability; active tracking of vendor roadmaps, changes and influences; and the ability to address cost modeling and billing matters.

Use Cases

CASE 1: University of Notre Dame: Main and Emergency Websites (www.nd.edu)

Motivation

Notre Dame evaluated the scalability of its main and emergency websites and found them unable to support an anticipated load of 5000 concurrent visitors. The service was also “desk checked” for resilience in the event of an on-campus IT failure affecting the web service. Time to restore was in the multi-hour range involving several departments and activation of the service at an alternate location in Indianapolis. The objective was to provide updated communications on the website within fifteen minutes of a failure. The existing on-premises solution could not meet these objectives.

Expected Benefit

We expected that moving the main web site to a public cloud service could provide the scalability we needed and the ability to run the service from several data centers. Running the website at multiple locations would provide the desired availability and decouple web communication from local events affecting Notre Dame’s main campus. Utilizing a public cloud would also make it possible to combine nominal and emergency operation, eliminating the need to maintain an alternate facility and eliminating any associated cutover time.

Implementation

Notre Dame moved its main, emergency, and selected sub-sites to Amazon Web Services (AWS) in January of 2013. The implementation included running the web farm as auto-scaled instances in three availability zones (AZ). Scaling to several times our anticipated load is now automatic. The service was implemented in Elastic Compute Cloud (EC2) Classic which is now a legacy environment at AWS.

The service had to be refactored in some minor ways. To facilitate auto-scaling, the publishing mechanism had to be redesigned to allow newly instantiated servers to update their content before accepting requests. A “tools” server is required to interact with the AWS API to change auto-scaling policy.

Since the service is run in three AZs, we were able to decommission the Indianapolis site. Cutover was simplified and now University Marketing and Communications (the department that manages the main website) simply needs to have Internet connectivity to update content. There is no time needed to cut over. Moving the service to a public cloud allowed us to achieve all of our objectives at approximately half the cost of the on-premises solution.

Lessons Learned

- There are some technical challenges migrating to public cloud environments.
    - Some of the tools and design approaches are different.
    - It’s helpful to work with experts in the infrastructure to build the required skill level.

    - On premise solutions have often grown organically and need to be decoupled from campus services and/or refactored to take full advantage of cloud technology.
- Accessing websites by their web URL (www.nd.edu) and also the top level Internet domain name (nd.edu) can't be directly accomplished unless DNS requests for the domain are all sent to AWS Route53.
    - The Internet standards require the top domain of a zone to resolve to an IP Address record rather than a common name (CNAME) that can be used for redirection which can also affect internal top level domains like biology.nd.edu.
- AWS Trusted Advisor recommends smaller instance sizes than was preferred. However, Trusted Advisor doesn't take network performance into consideration.

Financial Analysis

Over the past year of operation, the average monthly cost of operating the website has been approximately $515, including approximately $474 of EC2 charges and $41 of data transfer costs. This includes the purchase of one-year reserved instances on EC2. The annual cost of operating the website in AWS is approximately $6,180.

Prior to migrating to AWS, Notre Dame operated web servers in two geographically diverse data centers (South Bend, IN and Indianapolis, IN). The annualized cost of the physical web servers needed to operate these websites was $17,170. Migrating to AWS resulted in dramatic increases in scalability and resiliency while achieving cost savings exceeding 60%. This conservative cost savings estimate includes only the cost of the physical servers and does not include power, cooling, labor or network charges.

CASE 2: University of Notre Dame: Campus Independent Authentication, Authorization, and Accounting (AAA)

Motivation

Most institutions have implemented a number of critical services that are hosted by third parties on the Internet. Gmail, Box, and Sakai are common examples. These services utilize campus single sign-on or central authentication services. While this is very beneficial from a convenience and security standpoint, it links the availability of the service to both the hosting provider and the campus Internet and authentication system. This reduces availability and is confusing to the user.

If there is a problem with the central authentication system, an off-campus user may not be able to access an institutional Box account, for example, while having full access to their personal account. They may not associate this with a campus network maintenance event. To solve this problem, Notre Dame implemented a redundant AAA system in AWS cloud. This system answers authentication requests if the main campus authentication service is unavailable.

Expected Benefit

Eliminate the dependency of hosted solutions on campus authentication services.

Implementation

Notre Dame implemented its cloud based AAA service in AWS. The services consist of CAS, Shibboleth, directory servers and a domain controller. The CAS, Shibboleth and directory servers are in an independent Virtual Private Cloud (VPC) which is peered to a common services VPC that houses the domain controllers. The common services VPC has a customer gateway Virtual Private Network (VPN) used from domain synchronization. The domain controller is the credential store.

A key aspect of the implementation is the use of route53 to enable automatic failover and restoration when campus authentication services become unreachable. This was accomplished by creating a new DNS zone (identity.nd.edu) for which route53 is authoritative. Route53 can use health checks to monitor service availability and alter its response based on the result. In the off-campus DNS view, login.nd.edu refers to a host in identity.nd.edu which Route53 can re-write.

On-campus requests always resolve to on-campus authentication services. However, off-campus requests resolve to the on-campus services through Route53 when they are available and to off-campus services when campus services are unreachable by the Route53 health check.

The result is automatic failover in less than a minute upon campus AAA failure and automatic return to using campus services around five minutes after the service becomes reachable.

Lessons Learned

- AWS has a wide range of services that make some fairly complex scenarios easy to implement.
- Aside from the basic VMs for the servers, AWS had almost no additional cost.

Financial Analysis

This project was a service enhancement/augmentation. For this reason, there is no direct cost savings. However, to attempt to do something similar would have required separate infrastructure at an independent location with the ability to direct load based on service availability across disparate networks. The flexibility of the AWS environment and the ability of Route53 to alter DNS based on service health allowed us to send traffic to the currently available authentication service without the need to purchase network hardware, IP space, or house equipment at alternate facilities.

The AWS implementation consists of two tiers of services. A foundational infrastructure consisting of domain controllers that serve as the credential store and provide domain services as shared infrastructure. The domain consists of two M3.medium EC2 instances.

The second tier directly supports the authentication and authorization functions and consists of two CAS/Shibboleth servers and two LDAP directory servers. Each of these servers is an M3.medium EC2 instance. Each pair runs in separate availability zones and the front facing CAS/Shibboleth servers are load balanced by an AWS Elastic Load Balancer.

The cost of this environment is approximately $200.00/month including the amortization of one-time costs over the year, assuming 1 year reserved instances. One-third of this cost is common infrastructure serving multiple projects bringing the actual operational costs closer to $150/month.

CASE 3: Columbia University: Google Apps for Education

Motivation

Columbia's motivation to implement Google Apps for Education (GAE) was twofold: supply a robust, vendor-provided cloud service to replace central IT's existing email service and provide an alternative to decentralized mail services. With approximately thirty percent of University email account owners forwarding their email to systems outside of Columbia, it was clear that our customers were not satisfied with the current systems.

Columbia's default email, CubMail (Horde IMP web mail/Cyrus IMAP), was introduced in 2000 and supported nearly 80,000 users. It offered basic email capabilities and was on aging infrastructure. Central IT also runs a subscription based Exchange environment, offering email and calendar. With approximately 2,500 users, it was back level (Exchange 2007) and required a major upgrade. Throughout the campus, over 15 Exchange, Lotus Notes, Cyrus, and Zimbra environments were managed by local school and administrative departments. In addition, GAE domains had been implemented by the Alumni Association, University Libraries and affiliated institutions (Barnard College, Teachers College, and The [K-8] School at Columbia).

Expected Benefit

Our primary goal was to modernize CubMail while taking advantage of vendor-provided cloud services. This would allow us to increase the value and breadth of our services while also lowering recurring operating costs.

We saw an opportunity to "follow our users" by offering a version of one of the most popular cloud email services under the protection of a GAE contract. This would also allow for improved collaboration beyond basic email (calendar, docs, etc.) as well as improving our ability to ensure continued communications in the event of a campus or data center disaster.

From an infrastructure perspective, we anticipated avoiding further investments in the current email systems (e.g. CubMail, Exchange) and consolidating the 15+ email/calendaring systems across Columbia.

Implementation

Columbia participated with nine CSG peers to develop and run an RFP which eventually led to selection of GAE and Microsoft Office 365 (O365) as finalists. Columbia proceeded with several other issuing group institutions and agreed to a final contract for GAE. (Other schools decided on O365 only or used both GAE and O365).

Because Google did not sign a Business Associate Agreement (BAA) with Columbia, the Columbia University Medical Center and certain other departments were out of scope. We are now in the process of negotiating a BAA, but have so far been unhappy with the results, leading the Medical Center to decide to go with Office 365 instead.

Vendors were engaged to confirm the initial feasibility and architecture assessments as well as assist with the GAE deployment. We also used a vendor to assist with data migrations, initial website content, training and on-site support.

The initial implementation included Gmail, calendar and contacts. GAE accounts were provisioned (as an enhancement to a homegrown identity management system) and Cyrus data was migrated to the GAE accounts for faculty, staff, and students.

CAS was implemented for authentication and a device password schema was implemented for clients and mobile devices. Third party data loss prevention tools were implemented for Gmail and Drive.

Columbia has not yet finished migrating users off Cyrus, especially those within the HIPAA Covered Entity. Columbia has upgraded the central Exchange service to Exchange 2010 and has several departments that have implemented or are planning implementation of Office 365. The areas of HIPAA coverage and Google vs. Microsoft remains in a state of flux.

Prior to implementing Google Drive, a policy statement governing the use of Drive in an academic setting was developed and agreed to by IT, the Provost and the Office of Disability Services.

Lessons Learned

- Decide early whether custom training and documentation are required or if Google's can be used. Also, determine branding for training and documentation.
- Ensure knowledge transfer from vendor to core team so that data migrations can be performed between large scale engagements.
- During the planning phase of the project, determine the appropriate level of customer engagement required for data migrations while maintaining the scope and pace needed by central IT.
- Determine the support model for desktop clients, remote devices and web browser.
- Taking advantage of the speed with which Google releases improvements may require either realigning staff or obtaining new staff. A person or team should own the relationship with Google.
- Engage the local (school or departmental) IT personnel to help with the transition. Many of them are the primary interface to the end user and have established relationships with high-touch customers.
- Establish a dictionary for the project so that all constituents use consistent terminology (e.g. alias, send as, and alternate name).

Financial Analysis

Had GAE not been implemented, predicted costs to manage processing, purchase infrastructure and storage for internally run email systems (Cyrus, CubMail, and Exchange) were predicted to start at $500K per year. In three years with storage usage conservatively doubling, the TCO was predicted to rise to $1M per year. A more aggressive prediction with storage doubling annually led to a forecast of $3.9M per year.

Expenses for the Google Apps for Education implementation were spent on the initial feasibility study and analysis, the CAS implementation, and multiple data migrations that included support, communications and training. One vendor was used for the CAS implementation and a second vendor was used for the GAE feasibility study and migrations.

Approximately $40K was spent on CAS and over a three-year period $1M was spent on the GAE analysis and migrations. Note that the number of migration phases rather than the number of accounts migrated per phase is the largest driver in the total expense.

CASE 4: Georgetown University: Platform as a Service (PaaS) – Acquia/Drupal and Force.Com

Motivation

Georgetown University began reducing and centralizing 500+ websites in 2009 on the FatWire/Oracle Sites CMS platform. In 2012, only 50% of sites had migrated and over 150 remained on the original homegrown CMS platform. The roadmap for FW/Oracle projected significant capital investment to upsize and upgrade the environment as well as customization requirements potentially adding substantial cost. In addition, stakeholders were unhappy with the complexity of the tools and several legacy/shadow applications that did map into the CMS platform still existed in the legacy environment.

GU decided to move mid-migration to Drupal to leverage the open source platform, gain strides in more agile and mobile responsive tools and provide a more readily adoptable platform. GU analyzed the choice between building an on-premises Drupal implementation and Acquia's PaaS Drupal offering in the cloud. Acquia's platform was selected to gain speed to launch time, leverage a pre-built, scalable solution and rebalance the limited internal staff resources.

Force.com was chosen to address several legacy shadow/home applications that did not map into CMS/university public web templates, were deemed business critical and were not yet able to transition into the existing enterprise application cloud services. Force.com offered a more secure environment, modernized user interface, pre-built tool capabilities, standardized integrations, and a faster development platform than reconstructing or upgrading systems using PHP and other languages.

Expected Benefit

We expected that moving the web environment to Acquia/Drupal and adopting Force.com for application development would speed our progress, and address scalability and security requirements.

Using the combination of Acquia and Drupal, the speed to launch allowed GU time to develop Drupal expertise while improving performance monitoring and employing a scalable architecture. With Acquia and Drupal, GU had the ability to scale the resources needed up or down, and only pay for what we used. We also discovered that the flexible architecture of the Acquia/Drupal combination supported our large, multi-site environment.

With Force.com, we could speed up the time to develop applications since the focus was on clicks, not code. We also benefitted from the standard integration model and more secure access controls. Force.com normalizes the "shadow" apps with a common user interface, dashboards, data model, and objects as well as developing a bridge to later migration into ERP systems when that functionality becomes available (e.g. WorkDay).

Implementation

GU implemented a large multi-site environment to serve 90% of the sites from one code base and a shared template structure with separate databases. In addition, GU provided 10% of the schools that would differ from the University template with a separate doc root and dedicated resources in the GU Acquia/Drupal cloud environment.

The project has been successful with high adoption and satisfaction from customers, no downtime, and fast response for planned upsizing. Pending actions include migration of the top-level university site in the next six months following visual redesign. The intended architecture is a separate environment in Acquia/Drupal partitioned from the large school-based multi-site to support the higher traffic as well as burst/surge capability for emergency or high profile sites.

Lessons Learned

Acquia/Drupal

- Higher Ed is always "unique"; we encountered complexity in addressing security certificates across several hundred sites and our internal DNS for sites using sub-zone domains with multiple services rather than one site. While this is the "norm" for higher education, it is not a common situation for commercial and government clients.
- Scalability is only cost effective if you use it. At GU, we have found that the core 300+ multi-site environment represents 60% of the total costs while the 3 partitioned separate code/base doc roots for sites that differ from the University core represent 40% of the cost. The separate sites are brokered and managed by central IT but funded by the specific schools.
- Acquia's PaaS environment does not have automatic burst capability for high availability sites beyond pre-provisioning resources. The model still requires scaling to the high point for emergency and high traffic or unpredictable sites.

Force.com

- We needed to be aware of licensing complexities and requirements as well as vendor pricing model complexity. Enterprise-wide licensing would reduce barriers to development and leveraging the platform.

General

- We needed to identify plans for bridge capability as opposed to shadow systems for functions not yet available in the product base. Customizations and modifications are no longer an option, which we view as a positive. We need an alternate plan or business process re-alignment for needs not met by the new system.

- There is a greater need for configuration documentation or standard operating procedures for changes to prepare for fast release cycles as well as more business analysts and testers are needed on the central IT team than before to address agile changes.
- There is a greater need to engage with the vendor to influence product road maps.

CASE 5: Columbia University: HPC Cluster Deployment Options

Motivation

Examine the costs associated with setting up and running a high performance computing (HPC) cluster in different deployment scenarios. Three deployment options are examined: Columbia University Data Center, NYSERNet Syracuse Data Center, and Amazon Cloud.

Expected Benefit

The HPC Cluster Deployment analysis will provide a clear answer as to the direction Columbia should pursue to offer the best combination of services, functionality and cost

Implementation

Certain assumptions were made to constrain the analysis to most common cases. The work required to perform system administration and support for end users are both assumed to be constant. Purchased equipment was assumed to have a life of four years. For various reasons, taxes, indirect recovery costs, the cost of head nodes, and scheduling software costs are not included in the analysis.

Three types of clusters were the focus of this study: Simple, Infiniband, and Mixed. The simple cluster is comprised of 72 identical nodes. The Infiniband cluster consists of 72 identical nodes with Infiniband FDR interconnect added to each node. Apart from the addition of Infiniband, this cluster is otherwise identical to the Simple Cluster. The mixed cluster is intended to represent a slightly upgraded version of the existing Yeti Cluster as it was initially purchased. The only difference is that the CPU has been upgraded to the one used in the Yeti expansion round, the Intel E5-2650 v2.

Lessons Learned

- Columbia and NYSERNet are close in price for all options.
- Amazon has a competitive price for the Simple Cluster even before considering dynamic resizing and spot pricing. However, if memory is an important requirement then it gets a lot more expensive.
- Amazon’s “Infiniband Cluster” is just a higher-performing network option than the default and is not actually Infiniband. The performance of this cluster would probably be far inferior to a true Infiniband deployment and an argument could be made that it shouldn’t even be listed here.
- Amazon storage is relatively cheap but the performance (IOPS) would need to be tested to see if it can match the physical systems.

- Relatively minor changes in requirements or billing models could easily result in large price changes. In addition, some of the assumptions underlying the model may be shown to be completely incorrect and turn the results upside-down.

Financial Analysis

The combined four year costs for each cluster option combined with each storage option.

Cluster	Storage	Columbia	NYSERNet	Amazon
Simple Cluster	50 TB Storage	$1,480,478	$1,422,158	$929,620
Simple Cluster	200 TB Storage	$1,874,072	$1,776,008	$1,168,420
Infiniband Cluster	50 TB Storage	$1,629,174	$1,601,094	$929,620
Infiniband Cluster	200 TB Storage	$2,022,768	$1,954,944	$1,168,420
Mixed Cluster	50 TB Storage	$1,986,972	$2,109,228	
Mixed Cluster	200 TB Storage	$2,380,566	$2,463,078	

Migration Plan

Introduction

Institutional computing environments have evolved over the course of many years, and support a very heterogeneous set of applications, services, and architectures which prove difficult to disentangle in preparation for the move to a new, cloud-based architecture. As a result, most migration plans and approaches will proceed over an extended period of time (measured in years), and will rarely be a focused, one-time project that is implemented and finished in a compact time frame. Furthermore, because a cloud architecture is significantly different than a traditional, on-premises environment and is, in many instances, contracted instead of built, there is a need to develop and modify policies, create new technical foundations, retrain or invest in staff with new skills, reorient operational practices, and modify existing organizational structures as part of a migration plan.

To properly guide these activities, which will fundamentally change almost all areas of an IT organization, execution should occur in the context of a defined strategy with appropriate planning and governance. However, even this will need to evolve from traditional approaches to account for both the extended length of time of the change internally and the increasing pace of vendor and service innovation externally. These forces will demand a more adaptive and iterative approach that can morph as needed to address the needs at any particular migration phase.

This guide to migration approaches details a phased approach for implementation as well as governance, security and technical foundations. It is deliberately written at a high level and will not include extensive implementation or technical details. Also, the duration and pacing which an institution will follow for their migration is purposely not included, as those will be driven by the institution's distinct strategy and the particulars of their current operating environment.

Areas of Consideration during Migration

Strategy & Planning

To implement a cloud vision through its various phases of migration, institutions will need to ensure they properly define the strategy they will use to achieve this vision, and then undertake the appropriate level of planning to follow this strategy.

There are multiple strategies that can be followed, so it is not as important to consider the strategy itself, but more about clearly defining and articulating what the strategy is. This will largely define the resources committed to achieving the goals of each phase, which in turn influences the pacing, scope, and potentially the sequencing of each migration phase. The planning element will in turn support the strategy, and seek to better understand what will be required to achieve the goals of each migration phase.

Developing the right approach to both strategy and planning is likely to be a challenge for organizations. Historically, higher education institutions have taken very deliberate approaches with an emphasis on collaboration and consensus. While collaborative approaches will continue to be critical given the wide impact and uncertainty involved in many choices, unless deliberation and consensus can be achieved in short time cycles there is a high risk of “analysis paralysis.” Instead, strategy and planning should be viewed as dynamic and iterative, with the constant being the underlying vision which is being pursued, and clear governance structures which will allow decisions to be made quickly and reviewed often.

One of the key advantages of using the cloud is the agility it provides to institutions. The cost of failure is dramatically lowered and failing fast becomes desirable. IT staff can create and decommission entire environments in minutes, resulting in a very low cost of experimentation and rendering long planning processes less necessary. The cloud provides the ability to iterate fast, try many different approaches to technical issues and move those that work into production.

Governance

Effective governance structures will be critical to supporting and enabling a transition to cloud services in a way that minimizes the negative impacts of disruption to the organization. As a result, there will be a need for differing forms of governance, both during and across particular migration phases. Some elements of governance will be best served by traditional executive-centered models, while others will require the creation of more organic and network-organized structures. Regardless, these structures will be key to ensuring continued forward momentum by being able to rapidly understand the current state, the forward vision, and potential methods, and then choosing the next set of methods with a commitment to frequent reviews as the environment changes.

Security, Compliance, and Policy

Approaches to security, compliance, and policy will be highly dependent upon existing structures and maturity within the organization. In most respects, the fundamental principles in these areas are not significantly changed, though the need for clear and effective approaches will be greatly expanded. As a result, organizations with mature practices in this area will principally need to be engaged in adapting these to cloud-based services, whereas those with ad-hoc approaches will require more effort to first codify their policies, and do so in a way which can encompass cloud services as well.

Technical Foundations

Implementing cloud solutions requires defining and building various technical components which will be critical to effectively building new applications, migrating existing applications, and supporting and operating services in the cloud. The presence of these foundations can significantly ease the transition and migration, while their absence or delayed implementation has the potential of creating new technical debt which will need to be mitigated or addressed in the future. While some of these foundations will be purely technical implementations, others will be in the form of architectural guidelines and design patterns which make adoption and ongoing operations more scalable.

Operations

As services migrate to the cloud, appropriate operational practices and support systems need to be developed. Depending upon strategy, these may be extensions of existing operations, or they may be developed as "cloud native" solutions which are operated in parallel to existing offerings. Building these capabilities at appropriate times will be critical to achieve projected returns and expected benefits from cloud-based services. Absence of these capabilities may result in cost structures and staffing needs which grow linearly with scale as each new environment or application re-creates operational constructs. Work streams in the other areas will help identify and define operational components and how they will be implemented, with the focus of this area on how they are sustained and utilized by the organization in an ongoing fashion.

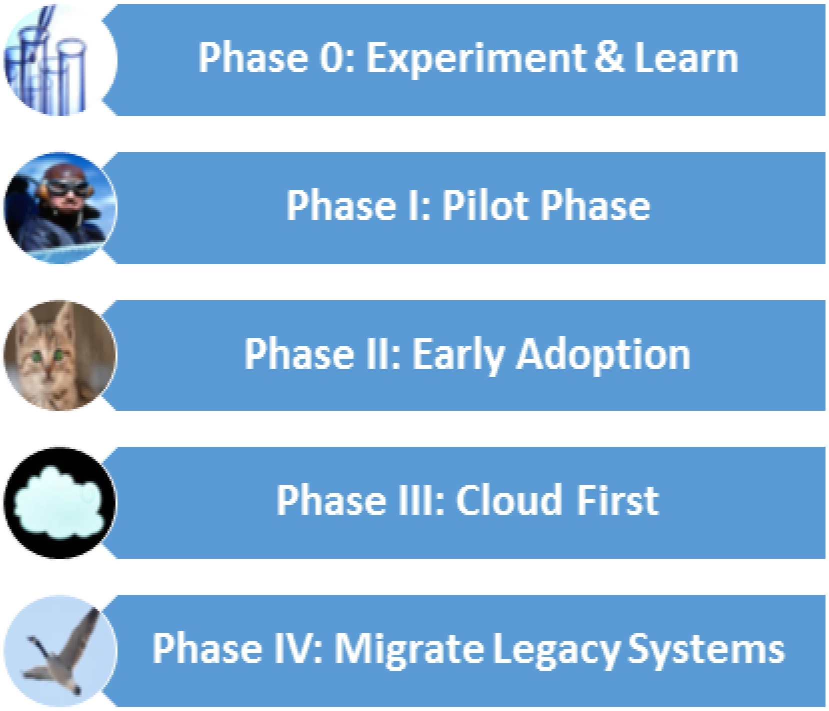
Organizational Impacts

As with operations, work in other areas will have a strong impact on how the organization will need to adapt to manage capability development, migrations, and steady-state operations. In addition to the cultural elements discussed in the Culture Change section, there will be very practical needs in training, organizational structure evolution, and the eventual retirement of legacy organizations.

Migration Pathway

The framework we present consists of several distinct phases representing different stages of cloud adoption. While these stages are presented as a linear progression, actual adoption programs may not create or observe clear, recognized boundaries between phases, so it is likely that they will often overlap or proceed as a gradual transition rather than with discrete steps. As a result, the individual activities that need to happen may begin earlier or later than specified here, depending upon particular needs, strategy, and execution path. Furthermore, in environments that are more distributed or lack strong centralized control, pacing may be uneven or may not align uniformly to this progression, resulting in multiple migration streams, or activities occurring from different phases simultaneously. This is another reason why having governance and strategy approaches which are both adaptive and inclusive is an important consideration.

Here are the phases that comprise the migration pathway:



Phase 0: Experiment & Learn

Phase I: Pilot Phase

Phase II: Early Adoption

Phase III: Cloud First

Phase IV: Migrate Legacy Systems

Phase 0: Experiment & Learn
In this phase, the organization will begin learning about cloud service offerings, typically in an ad hoc fashion and driven at the local level, with a focus on development and test environments or "proof of concept" implementations. This will happen most successfully and organically in organizations with a strong culture of innovation, while others will require more intentional encouragement and support for staff.

In this phase, there may not be a need for formal governance, although communities of practice and coordinating groups can help ensure learning is shared and common needs are identified. Formal and informal training to introduce cloud concepts should be offered and, where possible, tied to early projects. To further encourage and support active use, enterprise agreements with vendors and new or updated security policies should be developed. At a technical level, networking requirements and approaches should start to be reviewed, as these will become critical needs of early adopters and will be among the first challenges groups will encounter.

Phase I: Pilot Phase (3-5 projects)
In the Pilot Phase, organizations will begin to more systematically explore how to use the cloud. They will develop organizational excitement, commitment, and strategies that will allow them, with appropriate buy-in, to move into the more focused and intentional cloud migrations of later phases. Projects should be "low hanging fruit" with a focus on implementations that align to core cloud use cases, be obvious wins with enthusiastic

executive sponsorship, have limited political and financial risk, and have simple technical requirements, particularly with respect to integration with internal systems. Governance structures will start to emerge, though maintaining informal communities of practice is key to supporting organizational growth. A series of focused initiatives to build the necessary policies as well as organizational, operational and technical foundations will begin in this phase and accelerate into the next phase.

Existing Web workloads are ideal candidates for this early phase of adoption such as web applications that are not linked to sensitive institutional data. Cloud IaaS is often optimized for this kind of workload so it is easier to implement and gain benefits that are difficult to implement otherwise, like auto-scalability and operating multiple data centers. Web workloads, like public websites, can utilize simpler architecture and security controls. This enables faster deployment while minimizing data security risks and building skills to implement the more complex architectures that will be required to support a broader range of applications. Backups and disaster recovery may also be appealing candidates for initial cloud experimentation.

Phase II: Early Adoption (10-15 projects with broad coverage)

As organizations move into the Early Adoption phase, organizational commitment has grown and a clear strategy with the initial outline of a migration plan has been developed. This phase will see the initial development of organizational structures, as well as service design and development to support the growing needs of cloud environments. Projects in this phase should be selected to learn more and help serve the development of capabilities by evolving staff, service models, policies, and technical components. These projects will become increasingly complicated, and may be selected because of the capabilities they help develop, the cost commitments or capital investments they avoid, or the institutional buy-in they help create. By the end of this phase, most of the policies as well as the organizational, operational, and technical foundations should be in place. The necessary executive and organizational buy-in to support the broad adoption of the next phase should also be developed at this time.

The next phase will require an expansion of support infrastructure or an extension of existing support infrastructure to the cloud. This includes monitoring, logging, security scans, system administrator remote access, and system management tools. Now is the time to consider different architecture design patterns that permit the hosting of more complex applications and cover both centrally and departmentally managed applications.

Cloud IaaS changes the role of central IT. It can enable a greater expansion of departmentally developed and managed systems. Central IT can establish a new role for itself by serving departments as the center of expertise in managing cloud-based systems. To facilitate this, develop a shared accountability support and systems management model between central and distributed IT. Obviously, security and proper systems management is key to robust system delivery. By adopting a shared responsibility model, central IT can provide the benefits of deep security and operational management expertise to departmental IT without limiting the agility that often leads departments to pursue their own development efforts.

Cloud infrastructure can be implemented in several ways to facilitate a shared responsibility model. A shared services environment can be constructed to host common services such as authentication, authorization, domain and system management services, logging, DNS, etc. A common database facility can also be provided to house databases with broad access requirements. These services can be maintained by central IT and made available to both central and distributed IT through security and access controls. Central and departmental application "zones" can then leverage these facilities in accordance with existing policies. These shared services can either be implemented in the cloud directly or as an extension of the cloud to the on-premises data center through virtual private networks. Depending on the mechanisms used to isolate the various components of this architecture, separate accounting for each component may be provided by the vendor. This greatly simplifies cost sharing as departmental expenses are readily separated from central IT costs.

During this phase, institutions should also begin to develop expertise in cloud platform cost optimization. In addition to developing core technical expertise, institutions must have the business acumen to efficiently deploy workloads into the cloud.

Phase III: Cloud First for New Initiatives and Major Upgrades

The third phase will mark a growing acceleration towards cloud adoption, with an emphasis on full-scale implementations, particularly for all new systems or scheduled lifecycle upgrades to existing systems. During this period, organizational change management and ongoing capability development and optimization will be the key focus.

With most foundational elements addressed in earlier phases, technical needs will focus on the advanced tooling and operational needs as implementation and support scale increases. Training will become more specialized, and a greater emphasis will focus upon retraining core staff from legacy operations areas.

Towards the end of this phase, planning will focus upon final migration strategies for legacy applications, addressing both technical and organizational decommissioning. At this point, the preferred order of implementing systems or services will be:

- 1st Choice: SaaS
- 2nd Choice: PaaS
- 3rd Choice: IaaS
- 4th Choice: Things we own

Phase IV: Migrate Legacy Systems

In this final phase, all remaining legacy systems, and organizational, technical, and infrastructure components required to support those systems will be addressed. A variety of strategies and approaches can be used, including bulk "lift and shift" to cloud infrastructure, outsourcing, co-location services, or service decommissioning/attrition. Technical investments may need to be made at this phase to expedite migration and allow for the rapid disposition of legacy services. In some instances, financial or unusual technical limitations may necessitate continuing to serve needs with local infrastructure, though over time this will become a very small proportion of IT services, and will be addressed in ways that do not look like current data center and infrastructure support service models.

Focused discipline in disposing of remaining services will be absolutely critical to achieving the anticipated returns from migration, as there is a risk of needing to retain services with high fixed costs to support a very small number of applications and services. This will apply to physical infrastructure components such as data centers, infrastructure storage, compute platforms, operations management and other software licenses that were used to support and deliver services, as well as staffing roles used to support these environments. Effective planning in earlier stages in developing an application migration and organizational transition strategy will help ease the challenges associated with this phase and prevent stranded investments of both capital and people.

General Application and Service Migration Tasks

Security and Compliance Assessment

You should involve your security specialists and auditors early in the process to ensure efforts are compatible with your institution’s security and compliance requirements. Information security can be a daunting issue if not properly understood and analyzed. Hence, it is important that you understand your risks, threats (and likelihood of those threats), and then based on sensitivity of your data, classify data assets into different categories. This will help identify which datasets (or databases) to move directly to the cloud and which ones may require additional security controls if they are moved.

It is also important to understand these important basics regarding cloud providers:

- You should consider the sensitivity of your data, and decide if and how you will encrypt your data while it is in transit and while it is at rest.
- You can set highly granular permissions to manage access of a user within your organization to specific service operations, data, and resources in the cloud for greater security control.
- Security should be designed in at every layer of the environment. Institutions should require the use of two-factor authentication where possible and use a least privilege design model.

Technical and Functional Assessment

A technical assessment is required to understand which applications are more suited to the cloud architecturally and strategically. At some point, enterprises determine which applications to move into the cloud first, which applications to move later, and which applications should remain in-house.

Enterprise architects should ask the following questions:

- Which business applications should move to the cloud first?
- Does the cloud provide all of the infrastructure building blocks we require?
- Can we reuse our existing resource management and configuration tools?

- Can we use cloud-native architecture, such as spinning up and down entire environments on demand?
- How can we get rid of support contracts for hardware, software, and network?
- Is the application/solution/license supported by our chosen cloud provider? If not, how can we re-platform/re-architect the application to work in the chosen environment?

Migrating Licensed Products

It is important to iron out licensing concerns during the assessment phase. Cloud IaaS vendors work with many third-party Independent Software Vendors to smooth the migration path as much as possible. Licensing models vary, ranging from Bring Your Own License (BYOL) to purchasing servers with licensing "baked in" and to expanding existing enterprise license agreements to cloud providers.

Direct "Forklift" Migration

The simplest mechanism to migrate systems is to use a network architecture in the cloud which closely resembles the on-premises data center structure. This might be considered the first step in a large scale migration and will not leverage all the advantages of public cloud infrastructure. Using this structure, a network-by-network migration can be achieved, and even a server-by-server migration is possible. The main assumptions in this process are the preservation or forwarding of existing IP space and the tunneling of all traffic through existing on-premises network security controls. This approach allows you to take advantage of potential cost savings on compute and storage services and avoid capital equipment replacement cycles or expansion without having to re-engineer potentially complex network and security controls.

It is still important to assess cloud-specific security requirements. Although the bulk of security controls remain unchanged, you may need new isolation mechanisms for the subnets that are migrated to the cloud. These isolation mechanisms, including infrastructure account and access management, must be understood to ensure that the security configuration of the data center remains unchanged when migrated. It is also important to maintain operational systems and backups in separate accounts and to use multi-factor access controls on all privileged accounts.

The following table shows the spectrum of migration steps and the affected on-premises infrastructure.

Server/ Subnet Migration	**Network Security Migration**	**Database/ Shared Storage Migration**	**Network Refactoring**	**Service Refactoring**
Compute/ Local Storage	Firewalls/ Load balancers	Central Storage/ Shared Storage	Migration of Public Services to Amazon Web Services IP space, DNS, Virtual Private Cloud/ Subnet organization	Incorporate managed services as appropriate.

Planning to move data to cloud infrastructure requires an understanding of the storage options provided. Where possible, it’s preferable to use PaaS rather than IaaS since PaaS is higher up the abstraction stack.

Migrating shared storage may present some issues depending on your current implementation. If you’re using Network-Attached Storage (NAS) and presenting servers with Common Internet File Service (CIFS) or Network File System (NFS), you need to understand how that storage is being used. Some questions that need to be answered: What servers are using CIFS mounts and how? What servers are using NFS mounts and how? Can the storage be migrated to an object store like Simple Storage Service (S3)?

Categorize databases to determine what can move as a service. Moving a collection of applications and their associated databases may be easier than moving component parts of the service. Look at how databases are organized; they may need to be deconsolidated. Identify show stoppers that need to change. For example, you might replace a technology that does not support public cloud deployments, such as Oracle Real Application Clusters (RAC), with a cloud native alternative, such as a managed database service. You will also need to understand how backups, snapshots, and data restores to development and test systems map from current practice to the cloud.

If applications are moved prior to database migration, then you will need to determine latency and/or fragility of database connections. One method of accomplishing this might be to move a load-balanced server to the cloud, test latency and connection issues, then include the server in the production load pool to validate with production transactions.

There will be services that don’t make sense to move to the cloud or that need to have at least some on-premises component (e.g., DHCP, DNS, and Authentication). Typically these are the basic network services that enable the operation of end-user systems, labs, podia, etc. There also may be specialized systems or storage that cannot be moved. Identify these systems and services, then determine how to vacate them from any subnets that will be migrated.

Inventory and right size services where you can before moving them. In many cases, as systems evolve over time, they may become very inefficient in their use of compute, memory, and storage resources. Often, this is a result of sunk equipment costs and technical debt. Migrating systems without evaluating or adjusting resource needs will limit the financial benefit of migration.

Forklift migration assumes that all of the support systems move to the cloud as well, including monitoring and alerting, logging, and systems management. There may be an opportunity to optimize the support systems by mapping internal tools to those provided by the cloud vendor. Many of the fundamental management tools will be better provided by the cloud vendor as an integrated part of the environment.

Service/Application based Migration

The shared services infrastructure developed in Phase II serves as the foundation for Service/Application migration. This infrastructure architecture can be augmented in two ways to support services and applications. Just as there is the notion of a shared service zone, there is an analogous common applications zone. This concept can be used where applications share common resources like:

- Application or presentation layer server farms
- VPN connectivity to the on-premises data center
- Common security requirements, like campus access restrictions or broad public access
- Shared services zone
- Shared database zone

Applications that don’t fit this model can be cast into “independent application” zones. While this would be less common, it may be useful for applications that have a high degree of isolation, perhaps only needing access to the cloud-hosted common services zone. This simplifies the security controls on these applications and makes it easier to control or isolate lower level dependencies like disconnecting the application from campus DNS or using a different mechanism for IP addressing. Departmental applications are potentially a special case of independent application. Departmental applications may need more diverse connectivity, but specific security policies can be applied in accordance with the shared responsibility agreement.

Using these structures, applications can be migrated into the appropriate zones.

Refactoring Applications/Services

The previously described migration methodologies facilitate deployment of applications to cloud infrastructure but don’t fully take advantage of the facilities that are available. Cloud IaaS offers the ability to have unparalleled scalability and availability. However, applications often have to be designed to take advantage of these features. To take advantage of scalability, applications need to completely externalize state and data so that the auto-scaling of compute instances can properly initialize prior to accepting requests. It is also essential to design cloud native applications with horizontal scalability and loose coupling in mind. Solutions designed for vertical scaling hit limits far sooner than those designed with horizontal scaling. Using queuing services to create loose coupling makes applications more scalable, highly available, and fault tolerant.

Cloud IaaS providers can facilitate high availability by enabling applications to run with loads distributed across several geographically separated data centers. To take advantage of this, each tier of the application must be able to be run in independent network address spaces. Refactoring applications in this way also simplifies maintenance and systems management. When applications can both scale and be resilient to component failures, near non-stop operation is possible. This avoids scheduled maintenance outages and facilitates automated maintenance.

Risk Assessment

The use of cloud services is a completely new operating model for higher education IT. As such, it involves different types of risks than we traditionally consider and requires different approaches to existing known risks. During the development of this strategy, task force members consulted with representatives from CSG member schools and solicited input regarding the risks of operating in a cloud environment. Participants identified over thirty specific risks that were then grouped by members of the task force into four major categories: security, operational, financial and legal/compliance. Task force members then refined and consolidated the risk statements and documented strategies that institutions might use to mitigate those risks.

Security Risks

This section covers the confidentiality, integrity and availability risks that exist to institutions and individuals particular to using cloud providers. That is, most existing cyber threats also apply to using the cloud, but will not be listed here.

Infrastructure Security

The use of cloud service providers transfers some or all responsibility for managing infrastructure security from an institution's IT staff to the cloud provider.

Risk Description: Campuses become dependent upon IaaS providers for security of underlying infrastructure when using these vendors in an ongoing and operational basis. While this is of significant benefit from the support and implementation perspectives, it dramatically reduces the insight of the institution when monitoring infrastructure security. In addition, it provides the vendor with low level access to the underlying infrastructure.

Risk Mitigation: Institutions must carefully select cloud providers with a proven track record of securing their infrastructures. This should include initial and periodic reviews of security controls, including, but not limited to, review of a security assessment prepared by an independent third party. It is worth noting that many cloud providers bring significant security expertise to the table and may offer a higher level of security control than on-premises options. For example, cloud providers serve a wide variety of customers and must maintain compliance with a large number of security requirements. Those combined requirements typically meet or exceed the requirements of any individual customer.

In addition, institutions should integrate cloud infrastructure providers with the institution's existing security management processes to the extent possible. For example, cloud provider audit logs should be enabled at an appropriate level of detail and those log entries should be consumed by the institution's existing security log management processes. This provides important forensic response visibility and enhances troubleshooting capability.

The bottom line is that, in a public cloud deployment, institutions must partner with providers in a shared responsibility security model. The provider is responsible for the portions of the infrastructure that they manage (data centers, hypervisors, networking, etc.) and the institution is responsible for the portions they manage (OS patching, firewall ports, etc.). Many aspects of IaaS security, such as OS patching, firewall rules, and intrusion detection, are the same as in on-premises situation.

Data Storage and Transmission

Storing and transferring sensitive information in the cloud creates additional potential for security incidents.

Risk Description: Each time sensitive information touches a new service, that exposure increases the attack surface for that information. Storing sensitive information in both cloud and on-premises infrastructure provides two potential environments for an attacker to exploit in an effort to gain access to the information.

Risk Mitigation: Encryption should be used to secure all transmissions of sensitive information over public networks. Additionally, where feasible, encryption should be used to protect stored sensitive information, whether that information is stored in on-premises or cloud platforms. Organizations adopting encryption as a security control must remain cognizant of the fact that acting on data typically requires decrypting. Therefore, particular scrutiny should be given to the technical and business processes that interact with sensitive information. Additionally, institutions must understand how the decryption keys used to access this data are managed, with particular attention to whether the cloud vendor is able to obtain them.

Incident Response

Vendor incident response programs may not include adequate notification and disclosure of potential security incidents.

Risk Description: While many vendors have incident response processes, they are not necessarily effective relative to the needs of the consuming institution. For example, in the case of a security breach on shared computing infrastructure, it is often not possible to understand the scope of a breach. In this type of case, the conservative assumption is that a data breach could have occurred for all tenants. The practical consequence of that situation is the provider being flooded by audit requests all at once. Additionally, the audit logs will likely have multiple customers' transactions listed, which is an additional complexity to disentangle and potential data exposure of its own.

Risk Mitigation: The primary control for this risk is including adequate contractual language in cloud vendor agreements that specifies the circumstances, procedures and content of security incident notifications. Institutions should also maintain an ongoing relationship with the vendor that facilitates a productive discourse on security issues. Sample contract language that may be used for security incident response:

> Upon Vendor (which includes Contractors/Agents) becoming aware of:
>
> (i) any unlawful or unauthorized access to any Customer Data stored on equipment used by or on behalf of Vendor or in facilities used by or on behalf of Vendor;
>
> (ii) any unlawful or unauthorized access to any such equipment used by or on behalf of Vendor or in facilities used by or on behalf of Vendor that has resulted in, or Vendor reasonably expects to result in, loss, disclosure or alteration of Customer Data or any such equipment or facilities; or
>
> (iii) any incident for which Vendor is unable to promptly determine whether any unlawful or unauthorized access to any Customer Data stored on any such equipment used by or on behalf of Vendor or in facilities used by or on behalf of Vendor has occurred (each a "Security Incident"),
>
> Vendor will:
>
> (1) promptly notify Customer of the Security Incident in a timely manner to meet the state breach notification laws applicable to Customer;
>
> (2) promptly investigate the Security Incident and provide Customer with detailed information about the Security Incident; and
>
> (3) to the extent that a Security Incident is not caused by the negligence or willful misconduct or illegal act of Customer or any End User or by the breach of this Agreement by Customer or a breach by an End User of the Terms of Service, take reasonable steps to mitigate the effects and to minimize any damage resulting from the Security Incident.
>
> Following the occurrence of a Security Incident, Vendor will take prompt and appropriate corrective action aimed at preventing the reoccurrence of a similar Security Incident in the future.

Ease of Deployment

The ease of provisioning cloud services provides many individuals within the organization with the capability of replicating servers and storage, potentially increasing the attack surface.

Risk Description: The simplicity of deployment for cloud services facilitates the rapid provisioning of virtual servers and storage. One aspect of having physical computing assets (which must be procured and provisioned) is that there are more opportunities to vet a deployment decision. It's also a question of scale; a deployment of a thousand servers in the cloud is not visible, while a thousand servers in racks in an on-campus data center will likely attract attention. The same is true of storage. Given the large bandwidth capabilities of major research institutions and the virtually limitless capacity of most cloud providers, an institution's most valuable data can be replicated in a matter of hours.

Risk Mitigation: Best practices must be established and followed to avoid accidental security risks due to the ease of deployment. Institutions should adopt sound configurations and change management practices that effectively control the provisioning and consumption of cloud resources. For example, institutions may wish to leverage the APIs offered by cloud providers to inventory running instances and log API calls. In combination, those tools provide valuable change management and auditing capabilities.

Control Verification

It may be difficult to perform audits of cloud providers to validate that controls are in place as designed.

Risk Description: Most enterprise-class cloud providers undergo annual audits of varying depth and sophistication, however, the results of these audits are normally only available in terms of certifications. The actual findings of the audits are generally proprietary information that cloud providers are unwilling to share with consuming institutions.

Risk Mitigation: Institutions should include contractual provisions requiring cloud providers to perform service-appropriate audits and/or risk assessments on a regular basis. Customers should have access to those reports, with the understanding that some material may require redaction for security purposes. Institutions should also adopt business practices that monitor the receipt of audit reports and ensure they are reviewed by appropriate subject matter experts. Institutions may wish to include contract language specifying that they retain the right to perform their own audits, although actually performing those audits may be cost prohibitive.

Institutions may choose to adopt language in cloud service contracts similar to the provisions below:

> Vendor will audit the security of its systems, applications and data centers from which Vendor provides service to Customer ("Service Locations"). Such audit:
>
> (1) will be performed at least annually and after the occurrence, if any, of a Security Incident;
> (2) will be performed according to all applicable industry security standards;
> (3) will be performed by third party security professionals at Vendor's election and expense; and
> (4) will result in the generation of an audit report. The audit report will address the control procedures used by Vendor, including specifically an assessment of whether
>
> > (A) the control procedures were suitably designed to provide reasonable assurance that the stated internal control objectives would be achieved if the procedures operated as designed, and
> > (B) the control procedures operated effectively at all times during the reporting period.
>
> Following the generation of each audit report, Vendor will, on a confidential need-to-know basis, provide Customer with access to a redacted version thereof so that Customer can reasonably verify Vendor's compliance with its security obligations under this Agreement. Vendor may redact only information from the audit report that may compromise the security of Vendor's information technology environment or the confidentiality of any third-party confidential information.

Operational Risks

Shifting to a public cloud IaaS approach changes many of the operational details of information technology organizations. Institutions must understand these operational changes and their associated risks.

Disaster Recovery and Business Continuity

In an IaaS model, vendors become critical components of business continuity and disaster recovery plans.

Risk Description: Service level agreements (SLAs) provided by cloud vendors may not exceed or even meet the expectations of campus customers. In addition, when a single cloud provider is used, the institution becomes vulnerable to loss of data and/or services in the event that the provider suffers a catastrophic failure.

Risk Mitigation: Institutions must understand the SLAs offered by cloud providers and align them with legitimate business requirements of the institution. In many cases, the existing SLAs offered by cloud providers will meet or exceed campus expectations. For example, Microsoft and Amazon, the two major vendors in IaaS, both offer a 99.95% SLA for virtual machines and a 99.9% SLA for storage availability. In rare cases where these guarantees are not sufficient, institutions may establish additional controls to increase availability, such as the use of servers within multiple availability zones and/or regions run by the same service provider or the use of virtual servers from other providers.

The risk associated with a catastrophic failure of an IaaS provider is of extremely high impact, but very low likelihood provided the institution is using a major IaaS provider. IaaS environments are designed to be highly resilient and such a catastrophic failure, if it were to occur, would impact thousands of customers around the world and a substantial portion of the Internet. Institutions should, nevertheless, include this scenario in their disaster recovery planning. To protect against this risk, institutions may wish to host backups of critical systems and data either on premises or with a secondary IaaS provider.

Configuration/Change Management

Vendors may not follow appropriate configuration and change management processes to prevent service issues.

Risk Description: By nature, cloud services often change on a “forced rollout” basis. Changes in service may break processes or workflow without providing institutions the ability to reverse or delay changes. Vendors may make changes without providing advance notice or adequate time to test or validate the change. This is particularly true for SaaS products where vendors may feel free to roll out changes without notice.

Risk Mitigation: Institutions should consider change management carefully when choosing cloud service vendors. In some cases (again, particularly SaaS products), the deployment of features with little advance notice may be acceptable. For example, many institutions now use cloud-based email services where this practice is the norm and users are generally accepting of the pace of change, recognizing that it facilitates the addition of new features.

In an IaaS environment, uncontrolled changes affecting services provided to customers would be quite risky and potentially disruptive. For this reason, institutions should choose a cloud vendor with a mature service management strategy that allows for gradual change at the customer's pace.

As an example, Amazon Web Services introduced their second generation of virtual machines (EC2 VPC) in March 2013. At the time of this writing in August 2014, Amazon has not taken any steps to force existing customers to migrate to the new platform. New customers are automatically placed on EC2 VPC but existing customers are still offered the old platform under the name EC2 Classic. This approach allows customers to adopt new features that would potentially be disruptive, on their own timetables.

Institutions should consult with their candidate cloud providers to determine whether they implement most changes in a way that does not cause customer disruption. Planned changes should be performed only with prior notice to customers. Furthermore, cloud providers should follow strong internal change management practices, including logging, auditing, review and approval.

Vendor Viability/Lock-in

The dependence of institutions on IaaS vendors results in extreme impact if the vendor fails and difficulty migrating workloads among vendors.

Risk Description: The dependence of institutions on cloud providers introduces two related risks. First, the viability of the vendor as an ongoing business concern is of paramount importance to the institution. If the vendor suddenly ceases operations, the impact on the institution could be severe. Second, the use of vendor-specific functionality may make it more difficult to port workloads from one vendor to another.

Risk Mitigation: The risk of a critical vendor suddenly ceasing operations is similar to that discussed in the Disaster Recovery and Business Continuity section above. The risk is of extremely high impact, but very low likelihood, provided the institution is using a major IaaS provider. Such a catastrophic failure, if it were to occur, would impact thousands of customers around the world and a substantial portion of the Internet. Institutions should, nevertheless, include this scenario in their disaster recovery planning. To protect against this risk, institutions may wish to host backups of critical systems and data either on premises or with a secondary IaaS provider.

When designing IaaS deployments, institutions should carefully note any services utilized that are not offered by the majority of IaaS providers. Alternatives for these services should be noted in the institution's disaster recovery plan. For example, if an institution makes use of a proprietary DNS service offered by one IaaS provider, the disaster recovery plan may include provisions for building a DNS service with similar capability using the virtual server capability of another provider.

Institutions should be careful not to design to a lowest common denominator offered by all cloud providers, as this could hamper innovation and speed of development. Vendors may offer API-based methods to retrieve data for migration to other platforms.

Institutions should also examine the notification periods for contract termination (both with and without cause) contained within their cloud service agreements and ensure they meet institutional requirements.

Identity and Access Management

Identity and access management (IAM) integrations with vendors may add complexity and exposure of credentials depending on integration options.

Risk Description: If the identity and access management integration options offered by an IaaS provider require steps that would potentially grant the provider access to user and/or administrator credentials, those credentials may be compromised for use on unrelated services.

Risk Mitigation: Institutions should select known secure means for integrating IAM services with cloud providers. Administrator accounts should always be protected with two-factor authentication and designs should favor approaches where the cloud provider does not gain direct access to user credentials, such as SAML.

Third Party Relationships

Many SaaS vendors rely upon third party IaaS providers to deliver services to customers.

Risk Description: The nature of the current cloud marketplace is such that many SaaS providers make use of one of the major IaaS providers as part of their service delivery architecture. In those cases, the institution's use of the cloud service is dependent upon the IaaS provider. Also, institutions seeking to implement third party solutions in the cloud may find that application providers do not support cloud deployments.

Risk Mitigation: When contracting with a SaaS provider, institutions should gain an understanding of the technical architecture underlying the provider's offerings. In cases where the provider is using an IaaS supplier, the institution should include this in their assessment of the exposure they have to that IaaS provider.

If an institution adopts a "Cloud First" approach, this strategy should be embedded in procurement processes. The ability and willingness of third party providers to operate within that strategy should be a critical element in the vendor selection process. A vendor's inability or unwillingness to implement in the cloud should be treated the same way as any other inability to comply with technology standards.

Financial Risks

The use of cloud services presents financial challenges to IT organizations. While it is likely that the implementation of an IaaS approach will reduce expenditures for most institutions, the costs incurred are variable operational costs rather than the fixed capital expenditures that institutions expect for technology maintenance and upgrades.

Costs and Billing

The costs of IaaS computing are variable, "pay as you go" costs that fluctuate with varying consumption.

Risk Description: The traditional financial models of IT are designed around the concept of provisioning resources through large capital expenditures that take place periodically and provide capacity needed for multiple years. IaaS providers use a consumption-based approach that allows for just-in-time provisioning and de-provisioning, but replaces those capital expenditures with operational costs that may vary from month-to-month.

Risk Mitigation: Institutions should carefully monitor cloud costs. One or more individuals from the IT group's business office should be included on cloud deployment teams to develop an understanding of the billing models. These individuals can also adapt business and budgeting practices to accommodate the variable costs of cloud computing. In addition, institutions may choose to adopt a third-party financial management package to assist with the analysis and optimization of costs.

Financial management of the environment is a component of both cloud and hybrid approaches and is a skill that must be developed. In an on-premises environment, procurement controls serve to monitor costs but the cloud requires a new set of skills and procedures. Specifically, staff must be trained on the differences between cloud cost models and the ways that instance sizing and contract types (on demand, spot and reserved instances) affect final cost.

Financial Complexity

In a cloud model, where services are available on demand, engineers may lack an understanding of the financial implications of their actions.

Risk Description: IaaS service models are complex and it may be difficult for individuals performing an action to completely understand the financial implications of those actions before committing institutional funds. This is especially true when the individuals performing those actions are technologists with little or no business training.

Risk Mitigation: As with the previous risk, institutions should include one or more individuals from the IT group's business office on cloud deployment teams to develop an understanding of the billing models as well as adapt business and budgeting practices to accommodate the variable costs of cloud computing. In addition, institutions may choose to adopt a third-party financial management package to assist with the analysis and optimization of costs. Finally, technologists and technical managers should receive education on the cost models of cloud computing and the impacts their actions may have.

Legal and Compliance Risks

The use of cloud services introduces and/or highlights a number of legal and compliance risks that must be addressed. Institutions must ensure that their use of cloud services both complies with any legal, regulatory, or contractual obligations and provides sufficient legal protection for the institution.

Regulatory Compliance

The use of cloud services introduces new complexities that may prevent institutional compliance with legal and regulatory obligations.

Risk Description: The complex nature of cloud services, combined with the fact that data may be stored, processed and transmitted using resources that are geographically distant from campus may introduce regulatory and legal compliance issues.

Risk Mitigation: Campuses should evaluate each of their regulatory obligations to determine whether it is possible to leverage a particular cloud service and maintain compliance with that obligation. Particular attention should be paid to any regulatory provisions that specify the physical location(s) where data may be stored, processed or transmitted. In an IaaS approach, institutions will likely retain the majority of responsibility for implementing security controls that reside above the infrastructure layer.

Institutions of higher education seeking to move regulated activities to cloud providers should develop compliance plans specific to their regulatory environment and specific use case(s). Mitigation strategies for regulations commonly affecting higher education include:

- Credit card storage, processing and transmission must be done in compliance with the Payment Card Industry Data Security Standard (**PCI DSS**). Any IaaS vendor handling credit card data should be a validated service provider and appear on the Visa Global Registry of Service Providers. In addition, institutions should carefully delineate the division of compliance responsibilities with their service provider and must ensure that the institution's use of those services is otherwise compliant with PCI DSS.
- Institutions working with Protected Health Information as either covered entities under the Health Insurance Portability and Accountability Act (**HIPAA**) or under a Business Associate Agreement (BAA) must ensure that their activities comply with HIPAA regulations. In most cases, this will include entering into a formal BAA with the service provider.
- If service providers will handle student educational records, the Family Education Records and Privacy Act (**FERPA**) may apply. Contractual language with such service providers should include language designating the provider as an authorized school official under FERPA and spelling out the provider's security and privacy responsibilities.

- Institutions conducting research or other activities regulated under **export control regulations** must take steps to ensure that service providers will not store, process or transmit information using facilities located in geographic regions that would violate the provisions of those regulations.
- Institutions considered financial institutions under the Gramm-Leach-Bliley Act (**GLBA**) should ensure that their security plans address cloud computing initiatives.

Institutions may be able to make use of encryption technology to limit the scope of compliance obligations.

Institutions considering cloud services for the use of regulated data should work closely with their legal counsel and/or compliance staff to ensure that the use of cloud services is consistent with regulatory obligations.

Protection of Intellectual Property

Vendor contract terms may not sufficiently guard the institution's intellectual property rights.

Risk Description: The generic terms offered by cloud computing vendors may not provide sufficient protection for the institution's intellectual property. Specifically, generic terms may grant providers an intellectual property interest in data stored, processed or transmitted by the provider.

Risk Mitigation: When negotiating contract terms with cloud providers, institutions should pay particular attention to language concerning rights to intellectual property that will be stored, processed or transmitted by the vendor. Institutions may wish to adopt contract language similar to the following clause:

> Except as expressly set forth herein, this Agreement does not grant either party any rights, implied or otherwise, to the other's content or any of the other's intellectual property. As between the parties, Customer owns all Intellectual Property Rights in Customer Data, and Vendor owns all Intellectual Property Rights in the services provided.

The specific language used to protect intellectual property may vary depending upon the unique needs of each institution.

Notification of Legal Process

Vendors may not notify customers of subpoenas or other legal issues that they receive demanding access to customer data.

Risk Description: As the custodians of institutional data, vendors may receive subpoenas, warrants, National Security Letters or other legal orders to surrender that data to the government or a third party. If vendors do not inform the institution when they receive such items, institutional information may be disclosed without the institution's knowledge or consent.

Risk Mitigation: Institutions have two primary opportunities to mitigate the risks associated with legal processes served directly on service providers requesting access to institutional data. First, the institution should include language in cloud service contracts requiring that the vendor notify the institution of legal process when permitted to do so by law. Institutions may wish to begin with this model clause:

> Vendor acknowledges that Customer is obligated to comply with the Family Educational Rights and Privacy Act (FERPA) and the Gramm–Leach–Bliley Financial Modernization Act (GLBA). Notwithstanding anything in this Agreement to the contrary, Vendor shall not use or disclose Customer Content or Customer Data, including education records as defined by FERPA and data regulated by GLBA, except as necessary:
>
> (i) to provide the Service Offerings to Customer and any End Users in accordance with the Documentation (and any such disclosures by Vendor in connection with this clause shall only be to Vendor Contractors/Agents who satisfy the definitions of "School Officials" with a "legitimate education interest" as those terms are defined in FERPA); or
>
> (ii) to comply with Applicable Law (including subpoenas) or a binding order of an Authority. Vendor will give Customer reasonable notice of any such request of a governmental or regulatory body (including any subpoena) to allow Customer to seek a protective order or other appropriate remedy (except to the extent Vendor's compliance with the foregoing would cause it to violate a binding order of an Authority or Applicable Law).

Please note that this model language does not include protected health information under HIPAA. Any institution working with HIPAA data should ensure that the separate Business Associates Agreement (BAA) addresses notification of legal process.

Institutions must recognize that vendors may receive requests for data that they are not permitted by law to reveal to the institution. Of particular concern are the "National Security Letters" issued by government agencies directly to service providers that require the disclosure of customer information while prohibiting providers to acknowledge the existence of the order. Institutions may mitigate this risk by using strong encryption to protect data stored in cloud services and, where appropriate, ensuring that the cloud provider does not have access to the decryption key. This approach would render the provider unable to comply and would force the requesting agency to approach the institution directly.

Software Licensing

Existing software license agreements may not permit running licensed software in an IaaS environment or may result in institutions being double-charged for the same software.

Risk Description: Many of the license agreements for software used by institutions either did not anticipate the use of cloud infrastructure or intentionally prohibits running the software in a cloud environment. Furthermore, cloud providers often bundle license fees (particularly for operating systems) into an hourly service charge. This may result in institutions paying twice for software that would normally be covered under a campus site license.

Risk Mitigation: Institutions should undertake a review of the license agreements for all software that will be run in an IaaS environment to determine whether it contains restrictions that limit the use of the software either by placing geographic restrictions on its use or by directly prohibiting the use of IaaS environments. In cases where restrictions are discovered, they may be the subject of negotiation with the vendor when renewing license agreements.

Institutions should also be cognizant of cases where they are double-charged for software licenses. In an IaaS environment, this may occur when the IaaS provider bundles license costs into the hourly rates for server utilization and the institution has an existing site license for that software. In cases where this occurs, institutions may choose to accept the charges in the interest of achieving broader savings or they may choose to mitigate the risk by re-architecting their use of the IaaS service in such a way that campus site licenses may be used.

Indemnification

Vendor may not sufficiently indemnify institution against legal risks relating to use of the vendor's services.

Risk Description: The language in a cloud provider's generic terms of service may not sufficiently indemnify the institution against risks arising from their use of the technology. Specifically, if the vendor's services infringe upon the intellectual property of others, the institution may be exposed to liability.

Risk Mitigation: Institutions should ensure that contracts with cloud service providers offer sufficient protection against intellectual property claims made by third parties against the cloud service itself. Institutions may choose to use language similar to the following in their cloud service agreements:

> Vendor will indemnify, defend, and hold harmless Customer from and against all liabilities, damages, and costs (including settlement costs and reasonable attorneys' fees) arising out of a third party claim that Vendor's technology used to provide the Services infringes or misappropriates any patent, copyright, trade secret or trademark of such third party.

As with other contractual issues, institutions should consult with legal counsel to help negotiate indemnification language that is appropriate to the service being provided.

Exit Considerations

Overview

Cloud computing enables IT professionals, business units and individuals to respond to, flex and grow with many of the complex needs of today's ever-changing work environments. The benefits of having ubiquitous, convenient, on-demand, shared pools of rapidly provisioned and released IT resources can be of great business value but requires considerable advance planning and strategy. There are new service providers emerging and imploding every day in this nascent industry and up-front preparation for migration, recovery or exit altogether, is part of any well-considered cloud strategy.

Even though it doesn't happen often, what happens when the need arises to change services or repatriate a service back into our internally managed environment? As is the case for engaging a cloud service provider, the means by which they provide their services, store data, and how they manage and maintain their environments cannot be ignored. This advance understanding is useful, if for no other purpose than to understand how, and if, those services need to be brought back in-house or moved to the latest and greatest provider.

Cloud security advisor, Rob Livingstone, says that moving into the Cloud is like flying a light aircraft--easy to take off, but a nightmare to land and get out of.[5]

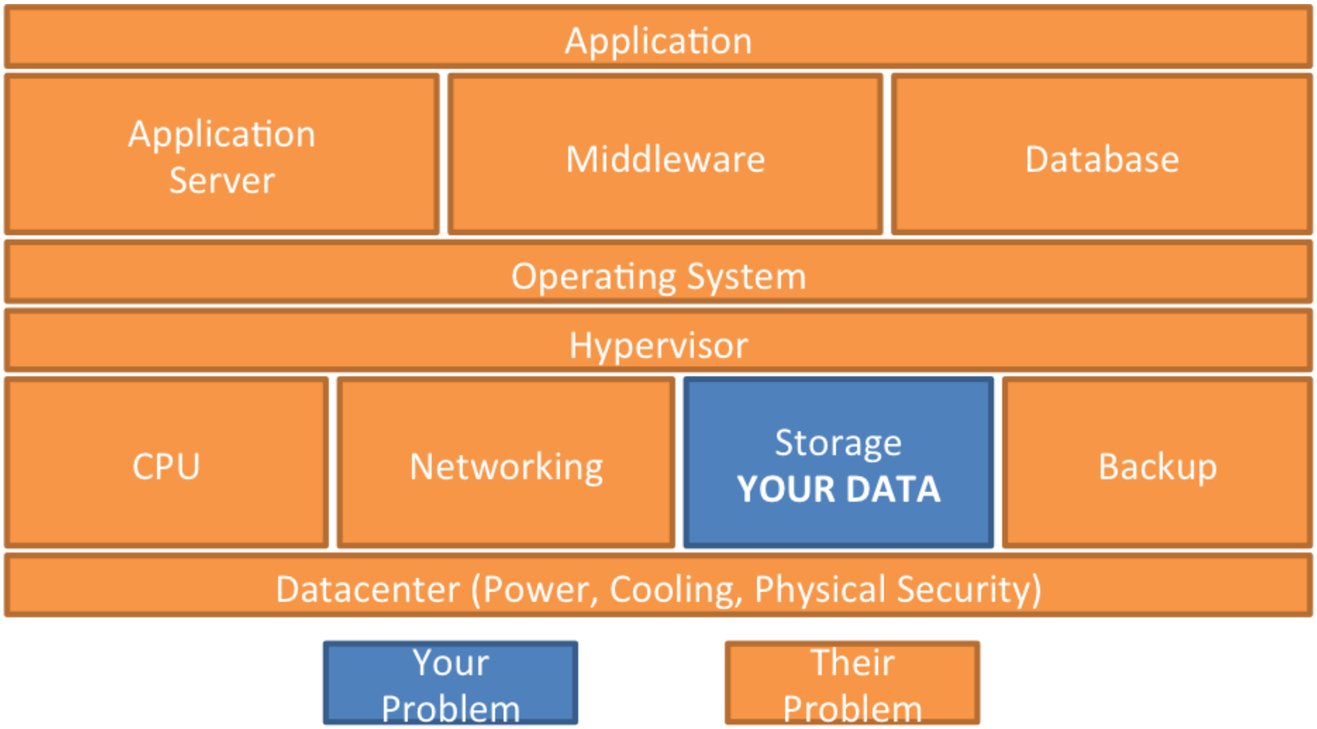
Gunnar Hellekson (of Red Hat's US Public Sector Group) notes: "The Federal Shared Services Implementation Guide, the agency blueprint to the cloud, makes it very clear that government entities engaging in cloud computing need a clear "exit strategy" for anything as a service. It might seem ridiculous to consider how one should migrate from a technology before it is even implemented, but when it comes to the cloud, being able to get your data out is just as important as getting it in. It's about choice and control."[6]

Types of Models - What to Consider

There are essentially three different models of cloud services consumed today. Depending on the type of service provided (IaaS, PaaS, or SaaS), the dimensions of the constructed exit strategy may be very different. (See Figure 1 in Our Shared Vision.)

[5] Barwick, Hamish. "Cloud Exit Strategy 101." CIO. 8 May 2012. Web. 14 Aug. 2014. <http://www.cio.com.au/article/423902/cloud_exit_strategy_101/>.

[6] Hellekson, Gunnar. "Do You Have a Cloud Exit Strategy? Here's One Clear Path. -- GCN." 27 Aug. 2013. Web. 14 Aug. 2014. <http://gcn.com/articles/2013/08/27/cloud-exit-strategy.aspx?m=1>.

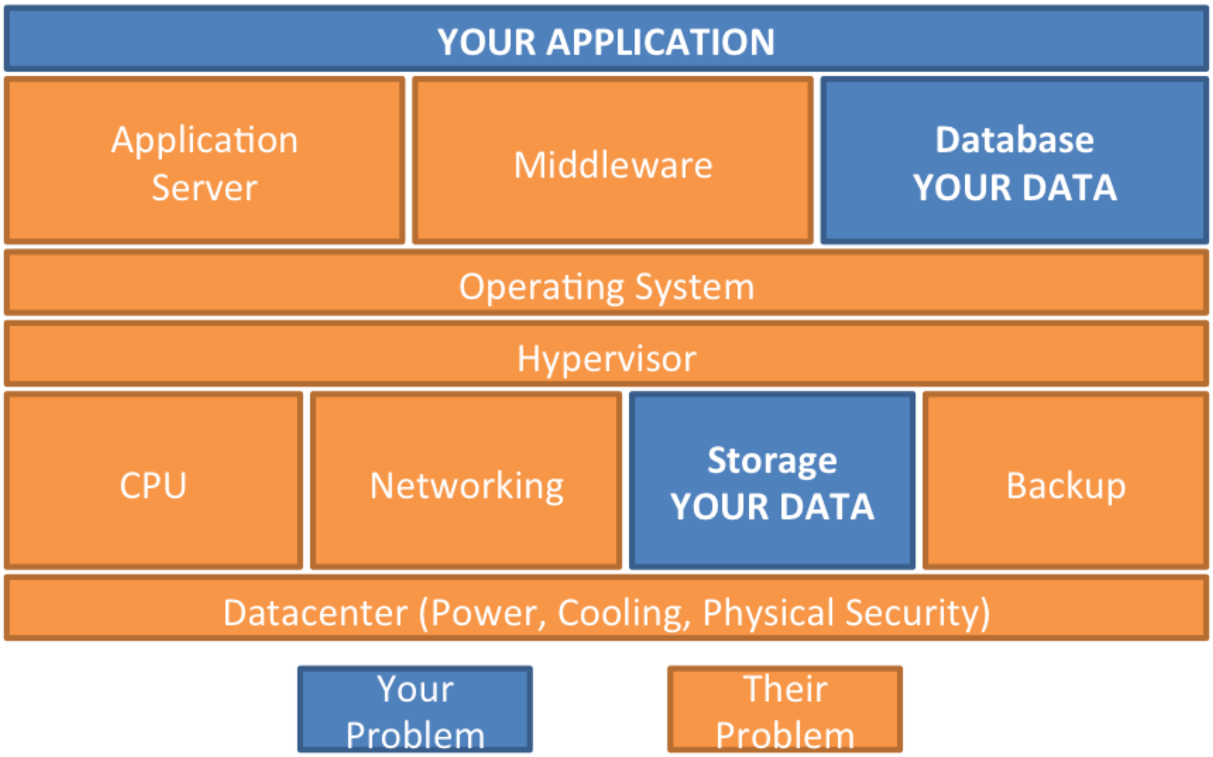


(Figure 2 - SaaS Ownership Model)

In the SaaS model above, the nature of planning is generally confined to the area in blue ("Your Data/Storage"). While this varies from provider to provider, this is the least complex environment to understand.

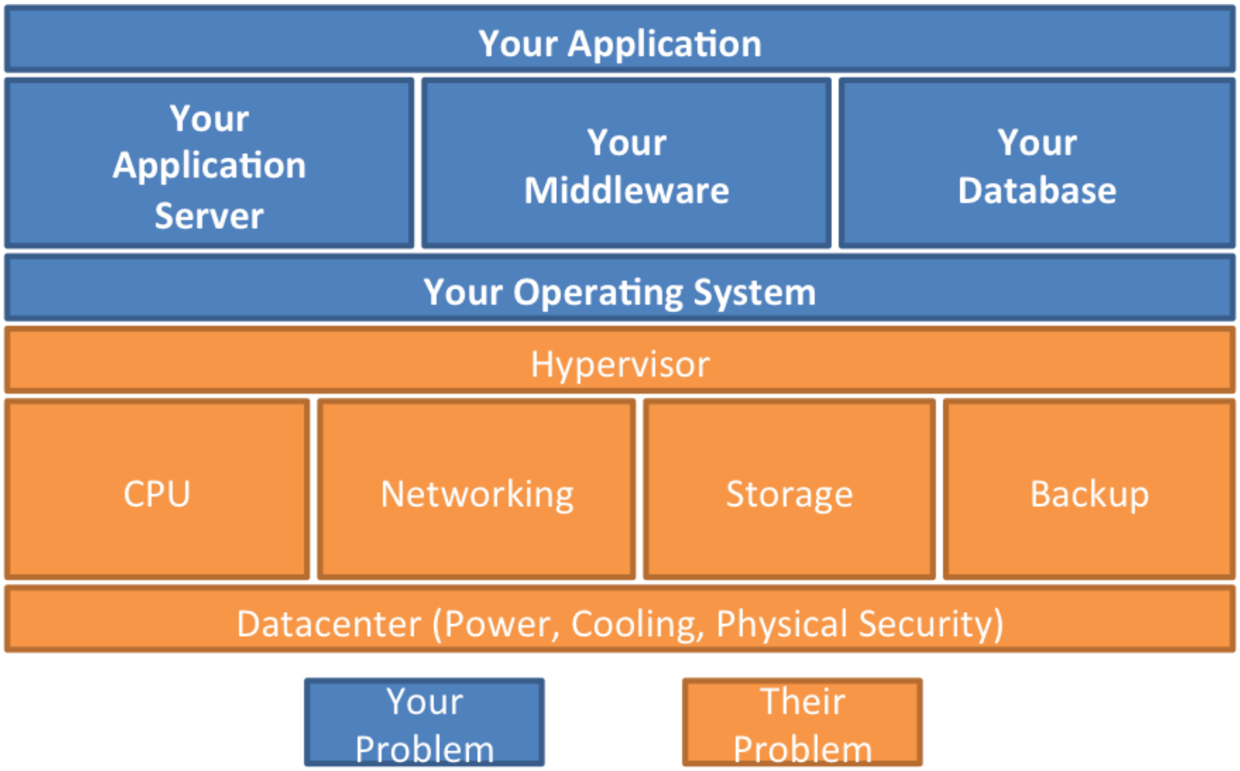
The ability to exit this type of environment can be very simple if the data is in a standard "tabular" or database format, but can be highly complex when the data is large-volume, unstructured documents (spreadsheets, documents, presentations, etc.) and may be protected by complex roles, permissions and/or metadata (as is the case with Box.com).

In this model, there is much greater control on the part of the service provider with respect to how those data are stored and manipulated. In many cases, these choices can impose a type of "lock in" because of their specialized use of access controls, metadata or perhaps a proprietary format in which they store data that may not easily migrate to, or be consumed by, other applications.



(Figure 3 - PaaS Ownership Model)

In the PaaS model, the scope of understanding required increases to include components of the application layer as well as the underlying data stores. Not only do customers concern themselves with storage (as in SaaS), but also with the application layers. In this PaaS model, customers are generally given the ability to configure, customize or implement business logic which best suits their business needs, but often these are done in languages which are proprietary to the platform (as is the case with Force.Com, the application development suite for Salesforce.com.)



(Figure 4 - IaaS Ownership Model)

The IaaS model presents the greatest level of complexity when considering an advance exit strategy. However, there is also much greater control on the part of the architect making technical choices. IaaS implementations can range from being completely self-contained, virtual machines to being fully-integrated "machine instances" which are tightly coupled to the IaaS provider's service layers (as can be the case with Amazon Web Services).

Having explored these three broad-ranging architectural options at a high-level, the next step is to examine situations that require an exit strategy and how to best position our institutions for a smooth transition.

Provider

In examining the provider itself for challenges, there are many possibilities that can provoke the need for migration or exit:

- Service Provider considerations
  - The provider's services may be less reliable, take a different direction than is required by the business, or may prove to be inadequate over time when a competitor's offering may be more suitable.
  - The business needs for local, statutory and/or regulatory compliance may exceed the capabilities of the Service Provider.
  - The service provider itself may become insolvent or unable to continue due to financial or technical catastrophe.

- Disintegrating relationship with provider

  Recent experiences have been noted in which changes to Terms of Service (ToS) or other evolutionary changes have caused the relationship with the service provider to crumble.

  - In one example, a service provider simply ceased operation because the cost of recovering from a breach of its Amazon Console was just too overwhelming.
  - In another example, a service provider increased its pricing to the point where it was no longer practical to have the service hosted in the cloud and required that the entire application be repatriated.

- Security or jurisdiction issues

  The legal jurisdiction of the provider should be understood especially if the service provider is operating in a different country from the institution.

- Changes (planned or unplanned)

  Because the solutions provided are so varied in scope, there are myriad possibilities to consider when evaluating possible changes in the service provider relationship.

  Among them are changes in:

  - Service Levels
  - Provider Ownership
  - Price
  - Terms and Conditions
  - Service Offering
  - Storage Location or Geographic Operation

Contracts

Contract provisions vary widely from provider to provider but understanding the terms they assert for exit provides the best place to start. Here are examples of the very different language expressed by two providers with respect to data ownership, a very common consideration:

Provider 1:

> *We will provide you with the same post-termination data retrieval assistance that we generally make available to all customers.*

Provider 2:

> *You will not have access to your data stored on the Services during a suspension or following termination. You have the option to create a snapshot or backup of your Cloud Servers or Databases, respectively, however, it is your responsibility to initiate the snapshot or backup and test your backup to determine the quality and success of your backups. You will be charged for your use of backup services as listed in your Order.*

There are other complex considerations when evaluating the contract basis for exit:

- Expiration of enterprise agreement or contract

Knowing exactly when and how the ramp-down occurs as well as the terms of asset egress must be identified *before the contract is signed*.

- Lack of support

  This is a consideration that is frequently overlooked when drafting the contract. Most providers have limited (if any) support for asset migration. If the reason for exit is due to lack of support, omission of a "support on termination" clause may make it difficult, if not impossible, to retrieve digital assets.

- Data, security, or privacy breach

  The event of any of these breaches presents unique problems too numerous to cover because everyone's digital assets are different. Service providers are, in most cases, asserting a lack of liability for any breach, or they offer a minimum "payback" for services. In many cases the Cloud Security Alliance Cloud Controls Matrix (Downloadable) serves as the best guide for disclosure by any cloud service provider.

- Provider's inability to stay competitive with industry features

  One of the most compelling reasons to move to a cloud service is to take advantage of a provider's particularly compelling version of a "solution to a (set of) problem(s)." Many providers offer a means of transferring assets out of one service into theirs but some do not. This form of "data lock-in" can make it difficult to move to another provider with better features.

- Namespace Protection

  The ability to protect or establish a namespace within the service may or may not exist. For example, if <research_institution>.edu were to subscribe to <service_provider>.com's solution and the institution name is embedded within their URL/URI scheme, there is frequently no protection preventing the name from being co-opted by another agency or subscriber.

- Repeated or prolonged outages

  Although services are becoming more robust over time, there are considerations to make. The value of the service being offered can be overwhelmed by outages or frequent service impairment. These are considerations that need to be drafted into the contracts before the service is deployed.

- Lack of remuneration for services lost

  As mentioned previously, the common language limits the liability to (perhaps) a month of billing as remuneration for a service outage, loss, breach or other catastrophe while some vendors offer nothing. Caveat Emptor.

- Change of internal leadership, strategy or corporate direction

  Many vendors, in the course of their corporate lifetime, change business strategies. In efforts to find the right market, many have remade themselves, been acquired, or failed completely. These are critical considerations. For example, a company that had considered its core business as "document management" decides to totally shift its approach to "web content management". These are completely different problem spaces and the company may no longer excel at the customer's original need, document management.

- Who quits whom?

  These seem simple at first, but the two most common events to be agreed upon in the contract phases are:

  - What if the vendor exits or changes ownership?
  - What if the customer exits or needs to change direction?

Data Management

Architectural Choices	
Asset Repatriation	
Types of Data	Volume (Dataset size)
	Velocity (Data volatility and need for speed)
	Variety (Varying forms of data)
	Veracity (How accurate are the data sources?)
Asset Restoration Process	Does the institution have the right means of restoring the data?
	Does the institution have the licenses for the container software?
	Does the institution have the skills necessary to understand how these data are organized and retrieved?
Production Restoration	
Compatibility of Formats	
Encrypted data containers and private key access	
Software containers (databases and licenses for correct versions)	
Metadata/document tagging	Re-application of tags (for compound objects like documents, worksheets, etc.)
	Re-application of embedded URLs, external references that cannot or will not move with the document
Access Controls (and associated hierarchies)	User Permissions
	Group Permissions
	Role Permissions
Hierarchical Nature of Document Organization	
Migration Considerations	
Egress fees and medium for delivery	Depends on the dataset size and complexity
Identity and access management (and associated namespaces) that are largely vendor specific.	

Application Management	
Business Logic	
Vendor's business logic lock-in more difficult that data lock-in	Business logic key differentiator that makes the services interesting
	No different from in-house products whose logic built from proprietary languages (e.g., on-premises ERPs such as SAP).
Institution's implementation of business logic (e.g., Force.com)	Has the institution been able to abstract a way of porting its institutional business logic in a way that is portable to other service platforms (e.g., from SalesForce CRM to SugarCRM)?

Other Considerations

If an institution decides to move its infrastructure out of the cloud and back in-house, then there are a number of considerations IT executives need to make.

- Does the institution possess the infrastructure required to run the applications?
- Can software licenses be re-established? Many cloud service providers use open-source software, but it may be modified to suit their commercial purpose (e.g. SugarCRM).
- Institutions may also need to build interfaces to existing systems and change operating processes.
- Advance Planning:
    - Build as a fully-abstracted implementation (software architecture) to provide any hope of gaining API independence.
    - Determine where backups are stored and how deep they go (Egress, transfer costs, and “Speed of Light”/“Speed of Net”/“Speed of FedEx” considerations).
    - What other data/services need to interact with this pool of data?
- Case Studies
    - AWS-based code hosting - “fatal exit” of http://www.codespaces.com/
    - Dedoose research data loss http://chronicle.com/blogs/wiredcampus/hazards-of-the-cloud-data-storage-services-crash-sets-back-researchers/52571

Sources

- Barwick, Hamish. "Cloud Exit Strategy 101." *CIO*. 8 May 2012. Web. 14 Aug. 2014. <http://www.cio.com.au/article/423902/cloud_exit_strategy_101/>.
- Hellekson, Gunnar. "Do You Have a Cloud Exit Strategy? Here's One Clear Path. -- GCN." 27 Aug. 2013. Web. 14 Aug. 2014. <http://gcn.com/articles/2013/08/27/cloud-exit-strategy.aspx?m=1>.

Acknowledgements

This strategy document reflects the input and effort of many individuals across the higher education community, particularly the members of the Common Solutions Group. We would like to extend our thanks to the members of the CSG Cloud Strategy Working Group:

Asbed Bedrossian
University of Southern California

Beth Ann Bergsmark
Georgetown University

Bob Carozzoni
Cornell University

Mike Chapple
University of Notre Dame

Alan Crosswell
Columbia University

Lisa Davis
Georgetown University

Patton Fast
University of Minnesota

Guy Falsetti
University of Iowa

Ryan Frazier
Harvard University

Brad Greer
University of Washington

Jim Jokl
University of Virginia

Charlie Leonhardt
Georgetown University

Mark McCahill
Duke University

Sharif Nijim
University of Notre Dame

Oren Sreebny
University of Chicago

Bruce Vincent
Stanford University

Bob Winding
University of Notre Dame

Steve Zoppi
Internet2

The Working Group is indebted to those who assisted us in our work. We would like to thank **Scott Siler** and **Michelle Sorensen**, of the University of Notre Dame. Scott served as our primary writer and editor and deserves full credit for the clarity of this document. Michelle served as our project manager, keeping a diverse group on task and organized.

We would also like to thank **Klara Jelinkova, Oren Sreebny** and **Tina Morris** of the University of Chicago for their generous offer to host our meeting and their gracious hospitality.

We also received tremendously helpful input from vendor partners. We offer our thanks to **Blake Chism, Steve Elliot, Ann Merrihew,** and **Leo Zhadanovsky** of Amazon Web Services as well as **Tyler Farmer** and **Mike Long** of Microsoft Azure.

Finally, we extend our thanks to all members of the Common Solutions Group who participated in the discussion during our May 2014 meeting at Notre Dame, our July 2014 workshop at the University of Chicago and our September 2014 meeting at Cornell.

This document represents the work of IT thought leaders in cloud expertise at the listed research universities. The views expressed are those of the authors and do not necessarily represent official views of their respective institutions.