UC Berkeley's Future-looking Cloud Strategy

June 12, 2017
William Allison
UC Berkeley

Introduction

Cloud computing, like the personal computing revolution and the meteoric growth of the internet, brings new capabilities for Berkeley's faculty, staff and students. Cloud technology now in use at Berkeley allows researchers to instantaneously provision unlimited compute power for large-scale data processing, gives professors the ability to bring this power into a lecture hall[1], and lets University staff automate tedious administrative processes. Once confined to a mainframe across campus that spit out paychecks once a month, this computing power is now on every desk, in every palm and pocket, in every unit or department, and every classroom, helping us advance the University's mission of teaching, research, and service.

2017 cloud technology is still rapidly evolving, but most vendors who provide infrastructure and services critical to UC Berkeley are rapidly shifting their product development to focus to cloud services. Most applications running in the datacenter will be moved to the cloud as the data center becomes more expensive and less practical over the next ten years. As we all know, there is no shortage of raves about the potential of cloud computing,[2] punctuated with occasional spurts of caution[3]. At Berkeley, we know technology's benefits, and understand that its ubiquity and ever-expansive potential changes the risks and increases the costs of IT relative to other spending. Our goal in developing a strategy for this shift is to continue harnessing opportunities of cloud technology while managing the risks, in order to corral this shift to best serve the University mission.

The applications remaining in the data center now include everything from large enterprise systems to hundreds of diverse applications owned by researchers and departments across campus. "Moving to the cloud" over the next ten years will differ for stakeholders depending on criteria like the readiness of cloud solutions, existing vendor contacts, financial resources, and the technical challenges. Some application owners will delay moving to the cloud when analysis shows that holding back is a better strategic choice. For others, "moving to the cloud" can mean anything from buying an application from a vendor, to a complex, labor-intensive re-architecting of an application now in the data-center, to the simple re-hosting of an application on cloud infrastructure – a so-called "lift-and-shift".

---

[1] Holgraf et al, 2012 "Portable learning environments for hands-on computational instruction" https://arxiv.org/pdf/1703.04900.pdf

[2] Thomas Friedman is so impressed with the cloud he renames it - as he describes: "or as I call it, the supernova, is creating a release of energy that is amplifying all different forms of power -- the power of machines, of individuals, of flows of ideas, and of humanity as a whole, to unprecedented levels".

[3] On the other hand, one respondent in a University survey about cloud use wrote: "Once I saw a cloud that looked like the Sta Puft marshmallow man, but these days they mostly look like dragons sitting on stashes of gold."

As with every new technology, and especially for cloud computing, roles on campus in relation to technology will change. For one, all or part of the computer stack is removed from local IT control. So, in some instances of Software-as-a-Service central IT will not be involved at all -- not with running the technology, building and security patching the application, managing the datacenter where it is hosted, or allocating storage for applications. This will give functional owners, for example a departmental dean or a central campus administrator, more autonomy to purchase the technology that meets the needs of the University. They may purchase a SaaS service, using self-service guidance from IT and procurement, to assure that it meets University, state and federal security, privacy, accessibility and contractual standards.

Because of the breadth of options and the complicated underlying cloud architectures by which different systems interact, a cohesive University IT strategy is important. This is preferable to the alternative, a more vendor-friendly but institutionally expensive organic process where independent unilateral decisions lead to fragmentation and undermine the goals driving cloud technology adoption. In the absence of strategic coordination, local decisions can inadvertently and negatively impact University capabilities and budgets.

For this reason, UC Berkeley leadership should direct stakeholders to coordinate cloud technology activities across campus. This means changes in how the University funds and measures technology investments, to the processes and policies used to procure technology, in the types of skills needed to deploy effective cloud technology solutions, and in the type of IT support most effective for cloud technology and critical cloud-enabling technology infrastructure. This document is a high-level overview of the actions the University should make over the coming few years to continue modernizing IT and embracing cloud technologies[4].

Measurement, Metrics and Finance

As cloud computing becomes the primary mode of delivering IT services, all of our systems will become still more interconnected and interdependent, even as functional owners become more autonomous. The use of metrics to measure legacy and cloud IT services is a key to guiding the implementation of a programmatic cloud strategy, and ultimately for assessing the effectiveness of campus technology decisions.

Engage Campus Leaders to Minimize Redundancy and Assess Options

University leaders must assist with incentives to minimize redundancy and aid assessment of myriad options. Technology decisions around cloud computing should be driven by data about the costs, benefits and risks of tradeoffs. Equally important are values-oriented considerations on the role of technology at the University, such as where and how IT contracts may trade privacy for lower vendor prices.

[4] A related document contains the University's modernization progress of the past decade, along with the more detailed considerations that drove these recommendations.

Develop Baseline Metrics to Enable Cloud Decision-Making

Metrics should be used to gauge progress and return on investment in the transition to cloud computing:

1. IST should develop a set of baseline metrics for the IT services currently offered via the Warren Hall data center. Measures should include customers, costs, and patterns of changing use.
2. Cloud pricing and technology options are very much in flux, and metrics must be used consistently for the University to assess different options. Analysis can show whether moving to the cloud will lower costs or improve capabilities, and can be used to prioritize which services to transition to cloud options. Once a service is moved to the cloud, metrics can be used to inform its effectiveness compared to previous and future alternatives.

Establish Effective Cloud Funding and Cost Recovery Beyond Recharge

1. Supported by leadership, central IT should rework recharge rates to decouple most labor costs from hardware and software expenses so that on-premises data center recharge services can be accurately compared with cloud alternatives (whose prices do not include labor).
2. The University must provide a baseline level of consulting, funded without need for recharge, to ensure support and consistent adoption of common architectures (consistent, reusable patterns or templates that can be used by anyone on campus to build or configure their IT application and infrastructure). Having IT experts with institutional knowledge funded on a common-good basis will also help distributed campus IT use common architectures.

Process, Policy and Procurement

We need to protect our interests as an institution. "Moving to the cloud" can sometimes suggest a singular problem-solving technology solution - "the cloud". Even among the three main modes of cloud computing however, Software as a Service (SaaS), Platform-as-a-Service (PaaS) and Infrastructure-as-a-Service (IaaS) -- there are many vendors and different business models with varying degrees of alignment with UC Berkeley's interests. The procurement process needs to facilitate distributed decision-making as simply as possible, upholding our institutional values and responsibilities for things like security and accessibility while enabling a coherent IT strategy. This means different considerations depending on the mode of cloud computing and the type of application. [See chart on page 8.] We recommend some specific policies and practices:

Data Use and Governance

The University should establish a Data Use and Governance policy. Some companies price cloud services inexpensively (or may even make them free) when they plan to monetize University information transmitted to the service whose value the institution may not yet have assessed. This is mostly an issue with SaaS, since PaaS and IaaS vendors typically generate revenue more directly by selling customers computing cycles. The policy should establish boundaries on the transfer or commercialization of institutional information to private companies. It should also establish privacy safeguards for individuals when the institution expects them to use a cloud service. Finally, the policy should set criteria for how to make tradeoffs with vendors on functionality and cost in light of the value we set on our institutional and personal data.

A Few Words About SaaS vs PaaS and IaaS

SaaS is generally the preferred option for cloud applications, because it replaces software coding, hosting, security, maintenance, etc., with a web application entirely managed by an external company. University systems, however, involve many large enterprise systems and specialized infrastructure applications that are not a fit for SaaS, not yet offered as SaaS (or as products not ready for prime time). For the cases where SaaS would be a fit, we usually recommend strategically delaying replacement of data center hosted applications or the launch of new applications until commercial SaaS matures. Unlike businesses, the "bleeding edge" is not where the University needs or wants to be for cloud adoption.

In some circumstances, functional owners who have access to resources and technical expertise choose to build or deploy a needed application that cannot wait via PaaS or IaaS. [See chart on p. 8] Both PaaS (such as Salesforce) and IaaS, require more IT expertise than SaaS. IaaS involves either completely building an application (along with associated cloud infrastructure), or deploying a vendor application on cloud infrastructure instead of in the data center. Since IaaS affords scale, flexibility (pay for the services you use) it is a go-to technology in support of research and academic applications and some strategic or critical applications that do not lend themselves to SaaS. Functional sponsors of IaaS applications must be clear-eyed in assessing the risks of IaaS solutions, and understand what it takes to ensure an application's safe, reliable ongoing operation.

Enable Safe, Effective Distributed Cloud Adoption

With hundreds of campus units purchasing thousands of IT services, it is not cost effective to have central IT perform architecture and security reviews for all of them. Units may hear of solutions from colleagues outside the University, and the selection of particular cloud solutions will not be fully within control of IT, who will only be able to officially support a subset of the actual vendor solutions used by campus. To facilitate unit autonomy and a self-managed review process for cloud services we recommend adapting the existing procedures[5], to facilitate approvals for University IaaS, PaaS and SaaS cloud services, establishing thresholds of impact, cost and risk that trigger more formal review. Self-service is an important part of our cohesive cloud strategy. The Cloud Resource Center (described below) will provide an initial cloud technology self-assessment that covers risk/security, accessibility for

[5] In this case, the Offsite Hosting Form can be adapted for this use.

people with disabilities, business continuity/reliability, product maturity, API/integration capability, data classification and management, and exit strategy.

Protect Institutional and Personal Data and Intellectual Property
For larger applications identified the initial self-assessment process, the IT Architecture and Infrastructure Committee (ITAIC) should oversee campus policies and required campus standards for IaaS and PaaS. In a world where one unit's technology choices can affect campus-wide costs and risks ITAIC guidance helps Berkeley hold the line.

Simplify Procurement and On-Ramps for MS, Google and AWS
To facilitate the campus' need for IaaS, University leadership should produce a simplified cloud procurement process that comports with law and policy requirements. Units will be expected to use one of the three top IaaS providers -- Amazon, Google, and Microsoft (see below), which will better enable central IT to support IaaS. Campus should establish operative agreements to streamline procurement with "on-ramps" in the form of common reference architectures, self-service deployment tools, and central IT support

Protect the University's Cost-Effective Logical Architecture
The University runs on hundreds of IT services that must interoperate and share data while being managed by many different units. Without a single integration strategy, the costs of adding or changing these services will spiral out of control, and the job of securing our institutional data and electronic assets will become impossible. The University should establish a policy that all enterprise systems must use the campus iHub (https://api-central.berkeley.edu/) to best leverage this common good resource, and to ensure maintenance of a simple logical data architecture for the University – which will aid the rapid adoption of cloud services.

People / Organization

As the University accelerates technology to the cloud, this shift will affect staff and faculty – and not just those in IT roles. As functional owners take more active stewardship of technology decisions and costs, they will also need to grapple with the implications of replacing staff-intensive practices with automation in order to leverage cloud computing's new capabilities, value, and economies of scale. They will need to be astute to the rapidly changing technology marketplace. For instance, right now cloud services seem cheaper partially because the labor required to operate them is not included in the list price - whereas the infrastructure services run by IT include all the operator labor costs. We recommend assisting this transition as follows:

Support Services for Campus Cloud Computing via a Cloud Resource Center
In order to promote smooth technology adoption and standardization of University cloud technology and to ensure all units have access to expertise as they make cloud technology choices, the CIO should launch a Cloud Resource Center (CRC). The CRC will enable autonomy being a clearinghouse for information, an associated online

information service via the main campus technology website, and a competency center staffed by technical and procurement experts from the Office of the CIO and IST. People working on cloud projects or looking for cloud-based IT services can access:

- CRC self-service pre-procurement options (such as the updated Offsite Hosting Form).
- On-line cloud technology resources and access to technical experts. We have received feedback and suggestions for the type of support units need to help people move to the cloud.
- A registry of UC Berkeley services running in the cloud (along with contact information for teams willing to share their experiences and lessons learned).
- A registry of approved UCOP and UC Berkeley cloud contracts.
- Discussion groups monitored by central IT as a place for free assistance.
- A list of University provided cloud technology training options (such as through the UC Extension and our campus Lynda.com agreement).
- Self-service tools to aid risk assessment and procurement, as well as standard tools and reference architectures to aid smart cloud investment decisions.

Training

The current skillsets of campus IT staff are not optimized for cloud computing. We suggest offering training to help IT staff become effective internal advisers and consultants for cloud technologies. We can also offer tailored technical training on selected commercial cloud platforms, and general concepts around automation, cloud architecture and security.

DevOps

In order to deploy and support some the IaaS applications that we believe will be an inevitable part of our cloud strategy, we will need to adjust our culture to support DevOps modes of operations. Specialized technology and support architecture with an emphasis on automation and DevOps tools/culture for those things we need to build and run on IaaS or in the data center.

Adjust Cost Recovery Mechanisms Mapped to Changes in Critical IT Skills

We will need three categories of services to manage cloud technologies. We can simplify central IT cost recovery by experimenting with shifting some recharge to simpler, more straightforward approaches for some services.

1. Administrative IT: Centrally funded bundles of functional staff, application and infrastructure support for ERP systems such as BearBuy and PeopleSoft.
2. Central IT Trusted Advisers: Create a freemium model to fund technical experts in infrastructure, automation, and certain tools. They will create and maintain self-service frameworks and when necessary provide expertise that supports technology decisions and operations.
3. Traditional recharge for services still run by IT and to support campus-wide IT projects.

Technology and a More Programmatic Move to the Cloud

The University has opportunistically moved or deployed hundreds of large and small applications to the cloud over the previous ten years, projects like the SciQuest eProcurement system (BearBuy), campus e-mail, the Rec Sports appointment scheduling and the Cal Performances ticket system. We are underway with a new set of cloud migration efforts whose goal is to initiate a more systematic and programmatic move out of the data center. Current projects include:

- Moving the University website and the campus events calendar to an IaaS cloud provider, for lower costs, and greater resiliency and responsiveness.
- Developing a cloud-based service as an alternative to the IST low-end database service that is currently in deficit.
- Retiring the current campus mobile framework and adopting a SaaS responsive web or mobile platform to offer a better mobile experience for our students, faculty and alumni.
- Retiring the IST Web Farm and migrating customers to SaaS or IaaS web hosting options.

Reimagine the data center's functions for the cloud

The data center must be reimagined “in the cloud” over the next ten years. This is not a shift or re-creation of the campus data center’s capabilities on a single cloud vendor. Rather we will use the three main cloud providers (Amazon, Microsoft, and Google) to deliver the services and capabilities currently provided by IST in the data center. This could be described in an abstract sense as a multi-cloud, virtual data center. At this stage in the evolution of cloud computing it is impossible to predict which vendors will offer the best service in the future. Furthermore, among them there will be strong reasons for a particular unit (research, educational or administrative) to use one over another, depending on the requirements.

IST Infrastructure Technology Roadmap

IST should establish a technology infrastructure roadmap (what's needed to evolve IT over next 5 years) that specifically examines identity management, a data network risk assessment, API/integration, automation of all aspects of IT operation, and security infrastructure (such as VPN and security analytics for the big three IaaS providers).

Plan to Move Major Campus Enterprise Systems to SaaS Over the Long-term

We recommend that the Enterprise Applications and Data Committee (EADC), along with IT and functional leaders develop a long-range plan to sequence the move of all major University enterprise systems to SaaS. This means that Berkeley’s PeopleSoft HR system will be outsourced to UCPATH, SaaS options for BFS (PeopleSoft Financials) should be considered, and eventually PeopleSoft Campus Solutions (SIS) will also move to cloud technology. However, we do NOT recommend moving any campus enterprise systems such as PeopleSoft to IaaS, without clear benchmarks that make a strong case (for example, a 2-year return on investment).

Use Containers to Drive Common Cloud Architectures and Avoid Vendor Lock-in
Campus IT should adopt industry standards for campus IaaS projects with an emphasis on containerization, such as Kubernetes and Docker, to maximize portability. Central IT should create reference architectures shared via the Cloud Resource Center, for when PaaS or IaaS solutions are required. Reference architectures and implementations should work with all three major cloud IaaS vendors (Amazon, Microsoft and Google) and maximize the flexibility to use the most cost-effective service while avoiding lock-in and minimizing reduplicative work.

Chart: Considerations in the Selection of SaaS, PaaS, IaaS and Data center hosting