Initial Report and Recommendations on
Research Cloud Computing for the CU Boulder Campus
November, 2016

1. Charge

In September 2016, Vice Chancellor for Research Terri Fiez convened a group of subject matter experts from around the CU Boulder campus to look at how cloud computing might be included in the campus research computing structure. The charge to the group included:

 (1) Create a plan for a cloud computing service model and cloud deployment model in support of research, especially research incorporating big data analytics. The plan will include requirements gathering to support the launching of any technical directions;
(2) Identify a service and deployment model in support of instruction and training;
(3) Generate a plan for an organization and staffing that provides open and equitable access to interested users;
(4) Explore a sustainable financial model for the operation and support of cloud resources; and
(5) Design guidelines and policies that are amenable to the adoption of cloud computing for researchers at CU Boulder in the future.

This initial report outlines a vision for research and education cloud computing, definition of success for the effort, results of a benchmarking exercise, identification of possible organizational structures, and recommendations for next steps, including a workflow mapping exercise that has already begun.

2. Working Group Members

Ken Anderson, Computer Science	Larry Levine, OIT
Jim Dykes, IBS	Kurt Maute, AES
Orrie Gartner, OIT	Joe McManus, ITP
Dirk Grunwald, Computer Science	Michael Paul, Information Science
Sangtae Ha, Computer Science	Ben Shapiro, ATLAS/Computer Science
David Hamrick, OIT	Doug Smith, College of Engineering
Thomas Houser, OIT/RC	
Brian Johnson, NSIDC	Supported by: Ligea Ferraro, OIT
David Kohnke, Leeds	Facilitator: Emily CoBabe-Ammann, RIO

3. Vision

The University will offer a multi-provider public cloud deployment service in a hybrid configuration providing cloud services for research and teaching & learning needs with a central point of contact acting as the broker of these cloud services.

With definitions as follows:

> Multi-provider: Different cloud vendor providers are available for use (examples in this space include AWS, Azure, Google Cloud). We know, both from our peers and listening to campus, no single cloud provider will meet all needs and multiple options must be available.
>
> Public cloud: Not on premise
>
> Hybrid configuration: manage both on premise cloud resources and public cloud resources as on premise compute is in abundance and still appropriate for many use cases (Cloud Opportunistic Strategy).[i]
>
> Cloud services[ii]: Compute service model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction (NIST, 2011)
>
> Central point of contact acting as a broker: A central point on campus can take the request and based on a number of factors, cost just being one of them, provide the various cloud compute options with a potential recommendation on which cloud service is best for the requested need.

Education, teaching and learning were added to the vision. Throughout discussions, it became evident that the two could not be effectively separated. It should be noted that addition take two distinct tracks: 1) overall education computing for teaching on campus, regardless of content area and 2) education and training within computer sciences and other disciplines where students design, implement and test both new hardware and new software solutions.

*4. Definition of Success for Research/Education Cloud Computing*

The working group was asked to define what success looked like, both in overall terms and in specific requirements. The definitions are not especially surprising but do outline some specific set of values the group held with regards to the role

that research computing (cloud or otherwise) has in computing on camps. Success of the chosen approach should include:

- An emphasize *effectiveness* and *efficiency*, where upwards of 60-80% of campus needs are met and the computing philosophy is that the campus manages the exceptions, not the rules.
- Allow and anticipate the convergence of High Performance Computing (HPC) and cloud computing, so solutions, including standards and analytics, should be **'platform agnostic'**.
- Seamless, standardized interfaces between internal and external data, regardless of campus platform.
- A community-defined service support infrastructure with clear boundaries, expectations and metrics.
- A community-defined portfolio of standard education and training offerings.

The group identified three areas of critical consideration for whichever cloud computing model is put forward to meet campus research and education needs.

Experiments & Sandbox (Agile/Flexible)	Hardware "touch"
Processes/ Data Integration	
Education	
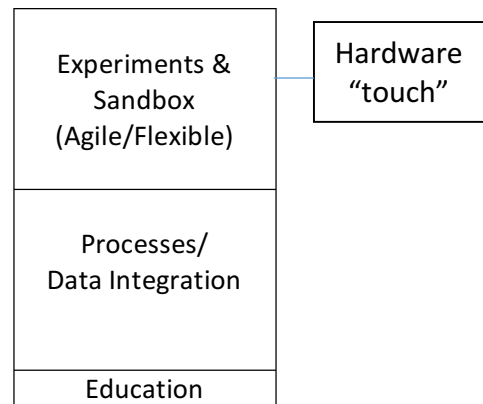
- *Flexible and fast* sandboxing/test beds, including the ability to connect to and test newly developed and prototyped hardware solutions. (CS R&D Space)
- Capacity for day-to-day research computing and data integration needs of the campus as a whole. This capacity is characterized by *stability, speed, well-defined workflows and a library of analytic tools.* (Data synthesis/data analytics)
- Increasingly, CU faculty are looking to the cloud to manage their classroom computing needs. Offering coordinated cloud solutions for those needs is viewed as an asset.

There is the recognition that what the cloud can do today is not what it will be able to do tomorrow. The group anticipates that more and more research computing will be done in the cloud as we move forward. In addition, many of our external partners have invested heavily in cloud solutions. The campus model, though a hybrid approach, should be able to anticipate that utilization of the cloud is likely to increase over time.

5. *Benchmarking*

Since incorporation of cloud computing into research computing structures is relatively new, the working group looked at several leading universities to explore how they defined, organized and implemented their cloud strategy.

The working group was particularly focused on commonalities or particular approaches that might allow the campus to meet the 'success' list. Findings include:

- Many universities exploring using the cloud started with islands of research or research clusters that implemented and experimented with cloud solutions on their own. These islands, in some case, transformed over some to a more-or-less coherent campus-level approach. Some invested in on premise cloud computing and some have moved completely to external cloud computing.
- Some of this speaks to the difficulty of overcoming institutional barriers and the challenge of getting faculty to buy-in.
- Most universities implemented a cloud model that included the same Process+Experimental+Education approach advocated for here.
- Funding models for the cloud efforts vary. Financial incentives are a challenge.
    - Many universities follow the same financial approach CU does, with regards to hardware, with lower overhead rates on equipment, but not on cloud services.

6. *Initial Next Step: Mapping Current User Needs and Workflow Types*

The working group quickly identified an initial critical step of mapping current user needs, based on their current workflow types. From there, it can be assessed what percentage of workflow volume and types can currently be done in the cloud. This is a critical step for developing a funding model and support structure for the effort.

Dirk Grunwald (CS) and Jim Dykes (IBS) have agreed to lead the effort, which is anticipated to have an 8-week timeline. A subsequent meeting of a subgroup focused on this mapping concluded that a 3-prong approach would be taken:

- A short survey of research and education needs will be sent to all faculty, instructors and researchers on campus. This will give some breadth to

the group's understanding of research/education computer usage and workflows on campus.

- Detailed interviews with targeted 'power users' and IT leads on campus, to develop a depth of understanding of the complexities of research and education computing needs.
- Focus groups will be convened in early January, drawn from respondents to the faculty survey. These focus groups will be designed to more fully understand the institutional needs from across campus.

Status: The survey has been drafted and is being reviewed by campus survey experts. It is anticipated to go to faculty, instructors and researchers before Thanksgiving. Regardless, survey will remain open approximately three weeks, and analysis and discussion of result with the committee will most likely occur early January.

7. Recommendations for Follow-On Work

The working group, in general, supported the notion of building cloud computing out from the current research computing framework, duties split between Thomas Hauser (consulting) and Orrie Gartner (sys admin).

While the mapping exercise (above) is the first critical step, there are several types of information that needs to be gathered:

- A survey of public cloud opportunities, including:
  - An understanding of cloud investments from NSF, NOAA and other federal agencies
  - An assessment of capabilities, support and costing of the major cloud services (including Amazon Web Services, Google, IBM)
- Understanding cloud cost vs. usage, including:
  - Contracting pathways with public cloud providers
  - Costs of purchase, maintenance and support for internal cloud efforts

Once these data are in hand, a set of larger strategic efforts should be undertaken on campus, by convening appropriately broad groups across campus. These efforts include:

1. Identify the campus risk posture for cloud computing, particularly in the realm of public vs. private cloud. (OIT, Campus Leadership)

2. Define a core set of services to be offered through OIT, with other campus entities. (OIT, Institutes, RIO, campus representation)
3. Define a core set of education and training to support those services, offered through OIT, in partnership with other campus entities. (OIT, Institutes, RIO, campus representation).
4. Develop an organization and staffing plan to support research cloud computing, including downscaled efforts in HPC as a result of cloud deployment (OIT).[iii]
5. Develop an efficacious costing model for cloud research computing that includes cost savings over time as some HPC resources move to cloud computing, any institutional investments in the effort, and a billing model for internal and external users. (OIT, Institutes, OCG, RIO)

Once these efforts are completed, by late Spring 2017, a campus-level working group can develop a deployment plan, with timetable and costs.[iv]

---

[i] This graphic is one way to represent cloud opportunistic at CU Boulder.
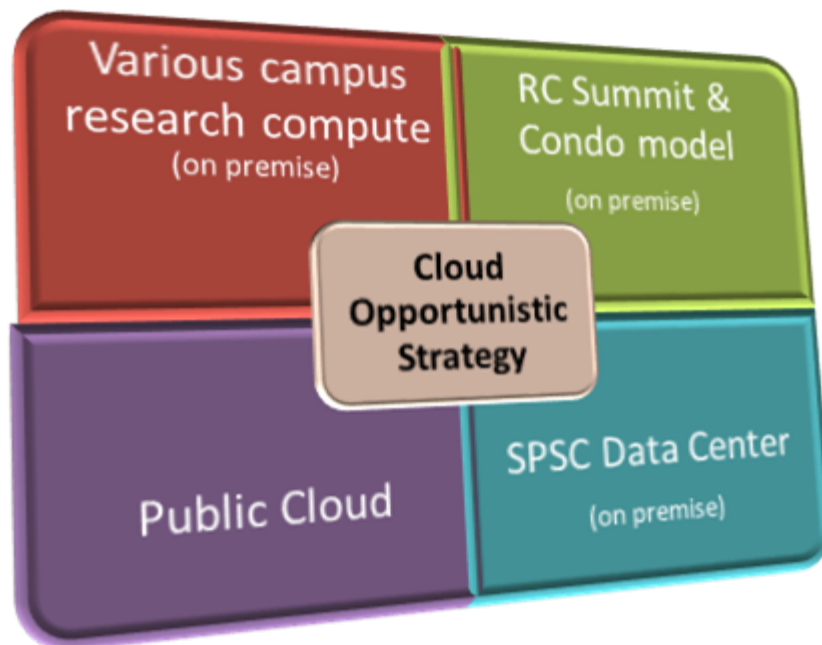
[ii] There are blurry boundaries between research, T&L, and ERP/administrative cloud services, which may be the same service in many instances, perhaps administered differently. Any cloud service designed for a particular campus domain (research, T&L, administration) if well designed will be extensible to other domains.

[iii] Public cloud and on premise cloud are not zero sum and one does not by default achieve a service or goal better than the other. Both are part of a larger strategy. Cost savings is one factor to consider in moving on premise services to a public cloud. Cost savings is always a goal,

but overall a total ROI is considered in offering a service, i.e., a method of achieving a service may not be the least expensive, but may be justified based on necessary outcomes. [iv] We will likely want to create one or more project charters, perhaps to address immediate verses longer range service needs, in order to plan scope, schedule, and cost.