Proposed Commons Credits Model Pilot Service Provider Conformance Requirements – 12/22/2015 Version

Definitions:

1. Digital Object: An electronic artifact, including, but not limited to data, software, metadata, and/or workflows that can be stored or manipulated in an electronic information system.
2. Digital Object Steward: The individual or organization that created and/or controls a digital object and that has formal responsibility for its security, integrity and/or availability.
3. Investigator: A user who interacts with the Commons.
4. Coordinating Center: The organization (including subcontractors, where relevant) that distributes computing resources (Credits) to stewards and prospective users of digital objects for use with Providers.
5. Provider: An organization that makes a conformant cloud infrastructure available to users of the Commons and accepts NIH Commons Credits
6. Reseller: an entity which provides capabilities as a result of reselling or providing access to another Provider's capabilities.
7. FISMA: Federal Information Security Management Act (44 USC § 3541 *et seq*) enacted as Title III of the E-Government Act of 2002, defines federal agency responsibilities for Information Assurance.
8. NIST: National Institute of Standards and Technology
9. IaaS: Infrastructure as a Service, based on NIST definitions[1]
10. PaaS: Platform as a Service, based on NIST definitions
11. SaaS: Software as a Service, based on NIST definitions
12. REST: Representational State Transfer; an implementation independent protocol for exchanging information over networks.
13. SLA: Service level agreement
14. CPU: Central Processing Units
15. VM: Virtual Machines
16. FTP: File Transfer Protocol
17. SFTP: Secure (SSH) Files Transfer Protocol.

General Requirements:

1. All providers must support direct access to IaaS as per the requirements defined elsewhere in these requirements. More advanced offerings (PaaS and SaaS) may also be included in a Provider's offering for reducing the effort needed for developing or running computational (or visualization) tools so long as these capabilities are accessible to recipients of credits and the general public in ways that do not require the use of these PaaS or SaaS offerings. The only exceptions to this requirement are interfaces specified in these conformance standards.
2. Resellers: A reseller of services can act as a conformant provider so long as the provider upon which they operate their service layer is able to meet the conformance requirements, including and the reseller is able to meet the service-layer portion of those requirements (e.g., interface).

[1] The NIST Definition of Cloud Computing, Peter Mell and Timothy Grance. NIST Special Publication 800-145

Resellers must also explicitly offer IaaS, but can do so by proxy and association with the underlying service entity.

3. Business Relationships with the Coordinating Center: The Provider must set up an appropriate business relationship with the NIH designated coordinating center that will be distributing NIH resources/credits to stewards of digital objects. The exact nature of the relationship is to be determined by the Provider and the coordinating center.
4. Credit Distribution Model: The provider must accept the financial mechanism by which the Government intends to deliver payment and to provide monthly on pre-defined and mutually agreeable reporting of Commons user metrics for those utilizing their services.
5. Validation of Conformance: Providers will need to document compliance with conformance standards prior to receiving approval to receive Commons credits from the coordinating center. Provider self-certifications will be bound by the False Claims Act (31 USC §§ 3729-3733) and the general provision about making false statements to the government encapsulated in 18 USC § 1001 and is punishable by fine and/or imprisonment. During the current pilot period (October 1, 2015 to September 30, 2018), providers will not need to re-certify. Providers will need to recertify upon the transition to regular operations and thereafter recertify on a bi-annual basis.
6. Changes to NIH requirements: The NIH will post proposed changes to the requirements for conformant clouds by an appropriate mechanism (to be determined) for comment by interested parties. Approved changes will become effective six (6) months from the date of approval. All approved Providers will retain their ability to participate in the Commons during the three (3) years of proposed Pilot activities. Conformant providers, once approved, will be grandfathered against changes to the conformance requirements for the remainder of the pilot. However in order to retain their accreditation, they will need to meet the current requirements that are effective at the time of any post-pilot activities.
7. General access considerations: In order to be part of the Commons, Providers must make their services available to the broad research community. Thus, a cloud that is inaccessible outside of that organization will not be considered conformant, since it does not make the digital objects contained within that cloud available to the broad research community.
8. Business relationships and liability: Digital Object Stewards and other investigators that interact with the Commons will do so under a business relationship with the Provider(s); the government will not be a party to these agreements. Similarly, the government and Providers will not participate in a direct relationship for the purposes of the distribution of resources; rather resources will be distributed and managed by a third party (the coordinating center) with whom the government will have a contractual relationship. The government therefore accepts no liability for the actions of investigators in the Commons. Providers are encouraged to seek appropriate counsel and take appropriate steps to understand their potential liability associated with the actions of investigators in the Commons. The government will not define the terms and conditions which Providers choose to offer their license agreements, or equivalent, to investigators. Providers are free to define terms and conditions, so long as they are (a) consistently applied, (b) are part of the regular account provisioning process, (c) do not violate federal or other applicable law and regulation, (d) do not assert intellectual property or other title on data or other digital objects, and (e) do not otherwise violate the conformance requirements.
9. Investigator access: Any academic investigator must be able to retrieve and download data they are authorized and approved to access from the Commons without using credits or other forms

of payment, subject to the limits described in bullet 3 in the section on Networking and Connectivity, below. Any interested party must be able to retrieve data from the Commons, however, individuals not affiliated with an academic research institutions may be subjected to standard commercial terms for data download.

10. Providers must follow a defined protocol when attempting to raise rates for users of the Commons in this pilot. The maximum price that can be charged for a given service by a given Provider is the published price on the Commons Portal for that Provider for that service. Notification of price increases for any service listed on the Commons Portal must be made through the Coordination Center, who will update the prices on the Commons Portal upon request. This may be done up to once per month for any Provider. The lead time for processing by the Coordination Center will be no more than 1 week.

Interfaces:

1. Interfaces: all interface standards and specifications germane to Commons Credit Pilot operations should be published and available to the research community. Although the interface must be open source, there is no requirement for the software itself to be open source.
2. Data deposit interface: Providers must make a series of data deposit interfaces available to investigators. These interfaces include an interactive, web based interface, a REST based web service and such other services as the Provider feels will add value to the research community. The interfaces should provide an identifier (such as a URI) that can be used to access the digital object.
3. Data download: All Providers are required to support a set of simple data download interfaces. These must include an interactive, web based interface, an FTP interface and a REST based web service. Providers are free to provide additional interfaces at their discretion. Providers should note additional download requirements enumerated in the section Networking and Connectivity.
4. Management: Providers must provide a minimally-functional web based management console available that
    a. Enables investigators to actively manage access controls for their data, and to manage Credits provided to them by the Coordination Center, as approved by the NIH,
    b. Provides historical and current service metrics to investigators, indicative of their storage, compute and network usage.
    c. Generates configurable alerts based on threshold usage and cost
    d. Monitors the Provider's SLA
5. Computational Tools: Providers must make available a mechanism to launch pre-defined workflows or data analysis tools on data stored in the Commons. This requirement should be understood to make available a relatively simple mechanism to launch applications against any accessible data. Prospective cloud Providers should demonstrate how to launch an application in their environment against such data. It is expected that the research community will be the final arbiters of the utility of the provided mechanism by selecting Providers that support the most useful/simple mechanism when placing data into the Commons. This may also be described as having a means by which instances of software can be operated.
6. Search: Although Providers will not be expected to provide search capabilities, Providers must be willing to work with search entities defined by the NIH to enable appropriate indexing of content resident in the Commons.

Identifiers and Metadata:

1. Identifiers: Providers must provide a mechanism to apply a resolvable identifier at the granularity requested by the investigator. This identifier may be a URI, URL or other identifier, so long as it can be resolved into an address that can be used for appropriately authorized access to the digital object.

Compute and Storage:

1. Storage: Conformant Providers must provide
    a. A minimum of 5 Petabytes of persistent storage pool available at published commercial rates[2].
    b. Attached block storage for VMs up to I TB
    c. File-based storage capability
    d. Elastic capability to provision storage based on computational need triggers
    e. Notification service with selectable alerts for storage usage
2. Compute: A conformant Provider must have
    a. Minimum pool size of 100,000 cores (or equivalent measure) for allocation to VMs. These should be at published commercial rates[3]. Provider conformance package should provide a clear definition of their computing unit measure for reasonable comparison.
    b. Offer a selection of VMs for low, medium and high computational needs
    c. Capability to vertically or horizontally scale VMs based on computational need
    d. Notification service for compute with selectable alerts.

Networking and Connectivity:

1. Commodity Internet: The Provider must provide connectivity to the commodity internet that can be configured to support networking with a minimum bandwidth of 40 Gb/sec[4].
2. Internet2: The Provider must provide connectivity to internet2 (http://www.internet2.edu/) that can be configured to support networking with a minimum bandwidth of 40 Gb/sec[5].
3. Download: The Provider must support free download of data for academic clients up to a rate of 10 TB/month per identified user through either commodity internet, internet2 or both. This will require a mechanism of tracking. Unlimited no-cost download policies will also satisfy this requirement.
4. Network pool: The Provider must provide a network pool composed of preconfigured devices such as virtual firewalls, physical network switches for ease in implementing redundant connectivity, load balancing or link aggregation.

Information Assurance:

1. FISMA Compliance: Although the NIH does not view the Commons as a Federal Information System under the definition of OMB Circular A-130, conformant Providers that will accept

---

[2] A published commercial rate in this case is a service that is advertised on the Commons Portal and available to users of Commons Credits at those defined prices.

[3] Ibid

[4] The government recognizes that networking in a public cloud environment is dependent on the configuration selected by the user of those services. Thus, the requirement is not to *provide* 40 GB/sec in all cases but that the environment can be *configured* by a user to support 40 GB/sec if needed. The government also recognizes that such capabilities would likely incur additional costs to the Provider, costs which can be passed on to investigators that require such service levels

[5] See previous footnote

sensitive data must meet FISMA requirements. The minimum standard is FISMA Low[6], although Providers are free to provide graduated environments that are certified to Low, Moderate and High (at the same or different price points) as desired, so long as the Provider has an available Low environment and the relative costs at the various levels of assurance are posted. Providers should consult relevant NIST publications for current standards. FedRAMP accreditation may be used as a proxy for this compliance line item.

2. Authorization and Accreditation (A&A): Providers must carry out the standard authorization and accreditation activities on their standard IaaS, PaaS and SaaS offerings, as well as such additional PaaS capabilities are provided to support their approval as part of the Commons.
3. Providers must provide a mechanism to assure the security of data in their IaaS, PaaS and SaaS environments. For example, a boundary firewall that limits access only to those with secure access (e.g. Walled Garden)

Authentication and Authorization:

1. General requirements: Providers must follow relevant NIST guidelines with regard to authentication and authorization, specifically NIST Special Publications 800-63 Rev 2, Electronic Authentication Guideline (http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-63-2.pdf),
2. Authentication: The Provider must allow and support authentication in a form that complies with good and standard industry practices.
3. Authentication: The Provider must allow and support authentication inCommon credentials (http://www.incommonfederation.org/).
4. Authorization: Providers must provide a straightforward mechanism to enable investigators to set permissions on data and other digital objects that have been stored in the Commons (see 'Interfaces' above). The minimum requirement is for a digital object to be made available either (a) only to the depositor, (b) to such individuals as defined by the investigator (who must know the relevant account names) or (c) publicly available to all users.

Desirable Features (preferred, but not required):

1. Continuity: Providers must have a published Continuity of Operations and Disaster Recovery Capability (COOP/DR) plan, and will submit this as part of the application. Where such a plan is part of a standard SLA, offerors may submit the SLA to meet this requirement.
2. Interfaces: All interface specifications must be published and licensed under an approved, non-viral, Open Source license. A list of approved licenses will be made available by the NIH Office of the Associate Director for Data Science, but the Apache license version 2.0 (http://www.apache.org/licenses/LICENSE-2.0) is the preferred license.
3. Interfaces: Providers must adhere to Open Cloud Computing Interfaces (OCCI) and/or Cloud Infrastructure Management Interface (CIMI) specifications for APIs.
4. Encryption service: an encryption must be provided for storing sensitive data, which may include capabilities for data in motion, data at rest, and data under computation.
5. Storage: perpetual or other long-term data maintenance contract terms must be available on commercial terms.

[6] FIPS PUB 200

6. Virtualization: options must be provided to support virtualization methods where the kernel of an operating system allows for multiple isolated user-space instances.
7. Messaging: options must be provided to use an available messaging system for facilitating information exchange between systems.
8. Processing: options for supporting batch and real time processing of tasks must be provided.