Federated Cloud Information Publishing

David Wallom

Introduction

Currently within large integrated distributed computing systems there are a small number of different information sources publishing information and a fairly small number of consumers in terms of different services. With the introduction of cloud or virtualisation as a fundamental layer on top of which other community specific services are then launched there are in some cases a smaller number of providers though a new class of consumers are then defined.

**How fed cloud uses the information system?**

How a cloud system will utilise information that is published will depend on the methodology behind federation that is employed. There are currently two accepted methods for federation that we will term in this document as either Loose or Tight.

**Loose Federation**

Here the utilisation of different cloud systems and crossing of cloud boundaries are controlled totally by the user of the system. They define where services will be run through a functional equivalent of a metascheduler for clouds.

**Tight Federation**

Within this the user interacts with only a single cloud system which then expands its services through ‘bursting’ into other cloud based services that are able to satisfy the users requirements. The cloud operator defines the different clouds with which they are interoperating and federating with.

**Information Types**

There are three different types of information that are required to be available;

**Static**, information that requires an intervention from the service provider to change, e.g. endpoint names, total available capacity, management interface, authorisation mechanisms. As its name suggests this data changes rarely if ever. Within the current production e-infrastructure the production registry is the GOCDB system, allowing for static information to be machine consumable.

**Dynamic**: This is information for which the end service consumers actions are able to affect the values of the information published. Examples include, instantaneously available resources including both what type of resource can be run on the system but also who or what those resource are available to. The current instantaneous status of the system within the production e-infrastructure is through the BDII system.

**Information System for Cloud as an EGI Resource Type**

It is important to differentiate where within the system we would want to locate ‘the/a cloud’ within EGI. There are currently three different types of service within the EGI model. Firstly the Core services. These are ubiquitous across communities. Secondly we then have Shared services. These are end user facing, though mainly setup for one or another community are able to be shared by the resource provider (mainly research organisations such as universities who must support multiple different types of community) and finally dedicated services. These are setup by and operated for a single community with no sharing between communities of these services. Within the remit of this document the author would consider cloud to be a shared service, operated by an entity for the benefit of providing user facing services.

Within this activity the content of the information system is defined by the questions that you are trying to ask of the system. Here are a sample of the questions that the authors believe are important, the first group are requests on static information, the second dynamic;

1. What is the name of the resource and what type of interface can I use to manage instances on the resource?
2. What are the AuthN and AuthZ rules that operate on your cloud?
3. What instances are already installed on the resource and am I allowed to upload my own instances?
4. If I am able to upload instances what format of instances does the resource accept?
5. Is there a data interface available and if so what is it?
6. What is the overall size of the resource?
7. Are instance templates defined that limit the choice of instance scales I am able to run?
8. What type of virtual network can I establish on the resource?
9. Does the resource support cloud scalability through managed bursting to another external provider?

The following are questions on the dynamic information;

1. I have a virtual instance that requires X,Y,Z resources, does your cloud have A>X, B>Y,C>Z resource available?
2. My instance is short lived is its utilisation of resources going to be captured in the information system such that overprovisioning will/will not occur?
3. What is the charging scheme and how much will using your cloud cost?

**Answers to Questions…**

1)

Resource Name: <NAME>

Resource Management Interface Type <AWS/OCCI/Other>

2)

AuthN and AuthZ rules: <????>

3)

Instance Upload; <Y/N>

4)

Instance Format: <OVF,Other>

5)

Data Interface: <y/N>

Data Interface Type <S3, Other>

6)

Resource Capacity; <# virtual core ><total storage volume>

7)

Templates: <Y/N>

Type of Templates: <Template defn>

8)

Internal virtual network interface

Internal:

Scenario 3

. Update rate of the information define its nature.

. Three information systems: Trade registry (i.e. GOCDB) static

information provided by the provider saying what should be there. A sort

of place to publish SLAs. Second is BDII: overwiew of the status of the

site semi-static information. 3rd is the monitoring system - real time

system.

. BDII - Cloud provider should implement an information service to

publish cloud information to BDII.

. Information to publish through GLUSCHEMA2. Define our own schema? How

difficult is to integrate a new schema to BDII?

. Information system should answer to the questions? Which is the best

resource provider to satisfy my requirements?

. Issue of scaling: people will use the provider that works and they

will not move to antoher provider.

What information:

. Availability of resources?

. Do you accept images?

. Do you accept my identity?

. Templates?

. Scalibility? -> Capability: Federation (Yes/No), Type (VM, Storage, Network)

. Managmement interfaces

. What type of storage

. Image catalog

. Service catalog

. which information is retrieved, is it dynamic or static?

. at which frequency

. is a caching mechanism used

. is published information accessible to a limited set of consumers?

. how frequently does dynamic information change?

PLEASE DESCRIBE USE CASES

2) what happens if the Infosys is not reachable

. how the Infosys endpoints are discovered

. is the information system needed or can be replaced by gocdb

3) how is topological information retrieved/published in fed clouds?

4) Any other requirements for the information system: missing information, enhancements needed

PLEASE DESCRIBE USE CASES i.

5) how many queries on average per day do you foresee (through web pages and through PI) etc..