OpenNebula - Request #4876
Allow clusters to share hosts

10/21/2016 04:49 PM - John Noss

Status:	Pending	**Start date:**	10/21/2016
Priority:	Normal	**Due date:**	
Assignee:		**% Done:**	0%
Category:		**Estimated time:**	0.00 hour
Target version:			
Pull request:			

Description

Currently clusters can share vnets and datastores, but not hosts. It would be nice if hosts could belong to more than one cluster, to solve the case of a heterogeneous infrastructure where not all resources are shared, but some are.

When launching a vm, the chosen vnet and image datastore should be able to uniquely specify which system datastore and which host to use based on the clusters that contain those vnets/datastores/hosts - in heterogeneous environments sometimes a host belongs to more than one overlapping set of those vnet/datastore groupings, hence needing to be in multiple clusters.

Note, this is similar to request http://dev.opennebula.org/issues/4464 for migration between clusters (which works great, thank you!), but, having the ability to have overlapping clusters would solve the issue more generally, and also allow to select system datastores.

(Note also, another path to do this is by pairing ceph image and system datastores, but this method solves it for other datastore types as well).

Our setup is like a venn diagram, summarized like this:

Cluster 1: Hosts A, B share vnets 1, 2, 3, system datastores sys1 and sys2, and image datastores img1, img2
Cluster 2: Hosts C, D share vnets 2, 3, 4, system datastores sys2 and sys3, and image datastores img2, img3

The desired behavior is:

case 1. if I launch a vm onto vnet 1 (from any image datastore), it goes on hosts A or B (this currently works as desired) and gets system datastore sys1 (this is currently only possible by setting DS scheduling rules on every vm template -- without manually setting these DS Scheduling requirements, the overall oned scheduling rules apply and this might land on either sys1 or sys2; this is undesirable as sys2 has lower performance due to spanning)

case 2. if I launch a vm onto vnet 2 or 3, from an image in datastore img1, it should go to hosts A, B (this works as desired) and get system datastore sys1 (again this is only possible by setting DS scheduling requirements)

case 3. if I launch a vm onto vnet 2 or 3, from an image in datastore img2, it should launch on hosts A, B, C, or D (which works currently, and, it can be migrated between the resulting clusters successfully as per 4464) and it should also launch into sys2 (which is currently only possible by DS scheduling. Furthermore, if it launches into sys1 or sys3, it won't be able to migrate across the clusters as desired.)

It would be nice to be able to have a 3rd cluster that spanned hosts A, B, C and D, like this:

Desired cluster setup:

Cluster 1: hosts A, B, vnets 1,2,3, system datastore sys1 and image datastore img1
Cluster 2: hosts C, D, vnets 2,3,4, system datastore sys3 and image datastore img3
Cluster 3: hosts A,B,C,D vnets 2,3 and system datastore sys2 and image datastore img2

This more precisely defines where a vm should land based on the selected vnets/image datastores, and allows the automatic selection of the proper system datastore without requiring adding DS scheduling rules to every template. If a vm uses vnet 1, it launches into cluster 1 and gets system datastore sys1 (case 1); use vnet 2 or 3 and image from img1, launch into cluster 1 and sys1 (case 2); use vnet 2 or 3 and image from img2, launch into cluster 3, get system datastore sys2 (case 3).

The extra benefit is that when an admin makes changes to the clusters (adding vnets, or datastores) the scheduling will take that into account without requiring additional metadata to be added to the vnets or datastores, or any DS scheduling rules to be updated.

History

#1 - 10/21/2016 05:04 PM - John Noss

Note see also http://dev.opennebula.org/issues/4877 for a different way to address this, for ceph datastores, by pairing ceph image and system datastores

#2 - 11/01/2016 09:45 AM - Ruben S. Montero

- Tracker changed from Feature to Request

Hi

OpenNebula is designed as you described, the datastore and vnets determine the set of suitable hosts. The more natural solution was to fix the host and let the vnet and ds cluster membership constraint the set of possible clusters. Freeing also host cluster membership would complicate scheduling, apart from preventing some cluster-wide attributes to be inherited by the host.

AUTOMATIC_DS_REQUIREMENTS are set by oned, and DS_REQUIREMENTS is very useful to implement different storage policies (similar to VMware DRS)

To implement the use case described in the issue, you need to cluster hosts with access to the same vnets and datastores.

Cheers