August, 22, 2015

OpenStack Fabric Management Proposal

**Problem Statement:**

Today, OpenStack Block storage supports locally attached storage via LVM and external attached storage via iSCSI and Fiber Channel based storage fabrics. When storage is externally attached, Cinder provides the ability to automatically configure Fiber Channel Zones via the FC Zone Manager driver. Unfortunately, the interfaces to support these Fiber Channel zoning operations are tied directly into the Cinder attach workflows and do not support APIs that would allow OpenStack Storage administrators more insight into the details of the zone configurations (without leaving the OpenStack Control Plane).

Cinder also enables the ability to manually specify multiple Fiber Channel Virtual Fabrics that will be used for automated Fiber Channel Zone Configuration. However, this requires the Virtual Fabrics to be configured outside the OpenStack Control Plane by using fabric specific administration tools. At the moment, this would be either Cisco or Brocade specific tools.

When using externally attached iSCSI storage fabrics, there is no support built into OpenStack that allows dynamic configuration of Layer 2 VLAN segment and Layer 3 IP information. In order to deliver an OpenStack cloud solution that is multi-tenant capable, the OpenStack administrator must properly isolate traffic by scripting or manually setting up these configurations on each of the appropriate components in the iSCSI fabric. Each OpenStack distribution needs to replicate this configuration process.

[OPEN ISSUE] How does the existing support of Neutron Layer 2 support intersect with this new service? Do we completely defer to Neutron on Layer 2 and 3 management?

[TODO] Add FCoE section. Fibre Channel over Ethernet (FCoE) inherets the same provisioning characteristics from Fibre Channel, and as a result also inherits the same limitations listed above. However, FCoE creates an additional restriction on network environments as it requires the ability to provide lossless Layer 2 networking and specific Class of Service (COS) values. OpenStack distribution that intends to provide any-traffic-to-any-device capability using the same underlying physical topology needs to ensure guaranteed lossless Layer 2 support for FCoE provisioning.

Finally, for all externally attached storage fabrics: iSCSI, FCoE, and Fiber Channel there is limited support for easily enabling fabric specific plugins for advanced features to enable QoS Policies, Performance and Health metric collection, and other storage fabric specific capabilities. The current plugin interface is limited to Fiber Channel Zone Drivers and does not support Virtual Fabric operations or iSCSI or FCoE operations.

**Proposed Solution:**

**[OPEN ISSUE] Need to define the High Level Architecture for this approach. In particular, we need to clearly define:**

1. **Automated Fiber Channel Zoning – this is critical to ensure the Cinder community signs-off on the need to extract the FC Zone Manager**
2. **North bound APIs for all services**
3. **Internal OpenStack Service interfaces**
   1. **Library interface to Cinder**
   2. **Interface to Neutron – to pull Layer 2 and Layer 3 addresses.**
4. **Virtual Fabric Management Details – we need an abstraction that is defined in such a way that does not assume any specific Fabric implementation. Vendor specific Fabrics would simply provide plug-in drivers to support. This is one of the more complex parts of this proposal and needs more detail to ensure alignment going forward.**

[OPEN ISSUE] Do we believe dynamic creation of Virtual Fabrics is an important concept to support a more flexible tenant management capability?

[ACTION ITEMS] Need to get customer / operator feedback on the value of this service. Garth to work with Rao.

[ACTION ITEMS] Get the Gerrit review up with the high level architecture details and post in the Operator IRc channel.

Although the current storage fabric support in OpenStack is a good start, it is lacking support for several key capabilities that would further simplify the OpenStack administrator’s infrastructure management experience. This solution proposes that a NEW Storage Fabric Management service be created in OpenStack. It would consist of the following:

* APIs to perform CRUD operations on Fiber Channel Zones
* APIs to support configuration of Layer 2 segmentation and CoS selection for both FCoE and iSCSI, and Layer 3 IPs across the storage fabric for iSCSI. This will provide much more flexibility for managing iSCSI- and FCoE-based fabrics in a multi-tenant environment. This would also be consistent with the approach Manila is taking to support dynamic tenant isolation for File Share networks.

[OPEN ISSUE] How do we guarantee the traffic to be on the appropriate (loss-less) network? The bladed infrastructure (in particular C-class) has a way to wire up the server mezz cards to the appropriate external uplinks to the switching fabric. Is there a way to validate the dynamic configurations of Iscsi, FCoE, etc.?

* APIs to support CRUD operations on Virtual Fabrics.
* Plugin interfaces to enabled advanced Fabric Management capabilities across all supported Storage Fabrics (QoS, etc.)
* Horizon support for
  + Fiber Channel Zone CRUD
  + Fiber Channel Virtual Fabric CRUD
  + iSCSI Layer 2 and Layer 3 CRUD
  + FCoE Layer 2 CRUD

**[NOTE] Need to capture the demand (both current and future) for FCoE and its importance in the OpenStack domain.**

**Use Cases (working list):**

1. Enable OpenStack Storage Administrator to quickly configure Virtual Fabrics for Multi-Tenant Isolation, independent of what the underlying Storage fabric
2. Enable OpenStack Storage Administrator to enumerate all configured zones and identify the health and topology details (i.e. what VMs are dependent on each zone)
3. Enable OpenStack Storage Administrators to easily setup iSCSI Layer 2 and Layer 3 configurations in their Storage Fabric (i.e. initiators, targets, and storage switches)

Enable OpenStack Storage Administrators to easily set up FCoE Layer 2 lossless VLANs with appropriate COS values for FC Fabric login and provisioning.

What are we NOT trying to do:

1. The goal is NOT to replace or duplicate Neutron Layer 2 and Layer 3 address management. Leverage is the goal.