*→ zhipeng has joined*

*← gampel1 has quit (Quit: Leaving.)*

*→ gampel has joined*

joehuang

hello

saggi

hi

gampel is in another meeting.

He'll be here shortly

are we waiting for anyone else?

because he said that we can start without him.

zhipeng

I think we could start *:)*

joehuang

ok

hello, Steve

joehuang

let's discuss how multi-cascade service will work together, it'll be the key in the design

saggi

I thought we are going to postpone this. I thought we are going to talk about the bottom cascade service.

joehuang

what's the major feature of the bottom cascade service

#startmeeting

saggi

connectivity and and event sending

joehuang

#startmeeting tricircle

saggi

joehuang: We don't have the meeting bot on this channel. I don't think this will work.

joehuang

no meeting robot

zhipeng

joehuang yes

joehuang

ok

let's start

event sending should be careful especially if there are lots of event sending at the same time

for example, many bottom openstack failed and restart at the same time

saggi

joehuang: The general idea is that the bottom service will do the polling since it's closer to that services and only propegate changes to the top service.

joehuang

just image data center power off and power on suddenly

saggi

We can batch updates. It's all a question of latency

joehuang

have you read the mail I sent in the mail-list

I can't find it in the lab

saggi

joehuang: which one?

joehuang

the bottom service one. not lots of mail sent these days

saggi

on openstack-dev?

joehuang

yes

saggi

<http://lists.openstack.org/pipermail/openstack-dev/2015-August/071162.html>

It's not just for polling it's an additional feature for optimization.

The general purpose is to have a site-local agent for cascade related operations.

like setting up connectivity

joehuang

how to set up connectivity

to create external network for Neutron?

connecting the external network as one overlay L2 network?

saggi

yes

joehuang

polling status and sync to top openstack is just to move the task to bottom service, it's has been mentioned in the last mail

ok

currently, we use softether VPN software to create overlay L2 network for cross neutron conectivity in CJK project

saggi

joehuang: I don't know how portable it is across neutron backends.

joehuang

not only L2 but also L3.

if we want portable solution, we need to define a general purpose API interface in the bottom cascade service

saggi

Yes, gampel has more information. But it's a general extension based on network backend

so we support DB based backends and not just discovery based backends.

joehuang

can we make these API as extension in Neutron?

saggi

It's a combination. We will need to have some extensions in Neutron as it needs to know about those \*fake\* ports that point to the gateway.

joehuang

Adding one more service in bottom openstack will lead to distribution package change. only the distribution with the package can work with tricircle

*\* irenab is lurking*

saggi

joehuang: Yes, but that is true for the top as well

If they don't package tricircle they can't have the top service.

joehuang

for "fake" port, we'll register one BP for these, and it's discussed before Vancouver summit, and some Neutron core supports that

saggi

Good

joehuang

the port is not tied to host, but some VTEP endpoint, so the virtual port could be created, and communicate with the port inside neutron

saggi

As long as the packet finds it's way to the l2gw *:)*

the bottom cascade service configures the l2gw to know what site it should be forwarded to and how.

since only the cascade service as a cross site view

without that Neutron will have to know about the cross site topology

joehuang

similar idea. this topic will be discussed the hackthon in Xian this week, some core/engineer from neutron community will attend the hackthon

saggi

I don't think it's neutron's job to know how to move a packet cross site. Just to get it to the gw

joehuang

not move packet, but how to make port tied to VTEP ip, not the host

saggi

who configures the VTEP?

joehuang

for single openstack, it's the responisility of upper layer software, could be cascading OpenStack, or other cloud broker, or horizon

saggi

And that is what the bottom cascade is for. To get that information from the top and configure it locally.

joehuang

not get info from the top, but the top indicate the bottom. the bottom should work passivly except staus publish

saggi

how?

joehuang

even for status publish, the top should register a callback to the bottom. and the bottom will call the callback

cascade service task flow.

cascade service task flow should indicate the bottom

saggi

You need to let the bottom layer know that port XXX should be sent thought tunnel YYY with the tag ZZZ.

Neutron doesn't configure tunnels. So how will those be created?

joehuang

there are two types of cross neutron neutworking

one is provider networking for neutron, for example, the overlay L2 networking betweet two neutron's external network

the other is tenant level overlay L2 networking

I guess you mixed these two type of networking

saggi

I'm trying to show that neutron doesn't do overlay L2 networking betweet two neutron's. And that is why we need to bottom openstack.

*→ gampel1 has joined*

joehuang

do you mean tenant level L2 networking?

*← gampel has quit (Ping timeout: 240 seconds)*

*← saggi has quit (Excess Flood)*

*→ saggi has joined*

joehuang

hi, Saggi, I think we need to use mail-list to describe your thought in detail.

hi, Saggi, I think we need to use mail-list to describe your thought in detail.

saggi

joehuang: 1 sec, having connectivity issues

gampel1

hi I am sorry i was in another meeting until now

irenab

guys, is there list of topics for discussion that you follow?

joehuang

the topic is connectivity and and event sending of bottom cascade service

gampel1

if i understand correctly what saggi is saying is that we must have a "L2GW" for inter site connectivity

joehuang

correct. the L2GW should be part of Neutron, or the responsibility of bottom cascade service

gampel1

but the current l2GW is focused on connection to HW it seem that we will need to extend it to support inter site connectivity

joehuang

and just mention that for "fake" port, we'll register one BP for these, and it's discussed before Vancouver summit, and some Neutron core supports that

this topic will be discussed the hackthon in Xian this week, some core/engineer from neutron community will attend the hackthon

gampel1

fake port will be very good but even if you have the fake ports you will need to bind them to the "L2GW"

joehuang

yes

gampel1

Ok so we will need code that does the binding per openstack networking implementations

irenab

What do you envision the API of L2GW will lookm like?

joehuang

he "fake" port is not tied to host, but some VTEP endpoint, so the virtual port could be created, and communicate with the port inside neutron

the "fake" port is not tied to host, but some VTEP endpoint, so the virtual port could be created, and communicate with the port inside neutron

mechanism driver should be provided to handle the "fake" port which tied to VTEP, i.e. L2GW

gampel1

joe: I not sure i understand irenab:first connecting the tunnels and second we need to think how to bind the segmentations

joehuang

and neutron API interface should allow fake port creation

gampel1

fake port is great but still you will have to redirect the fake port traffic to the L2GW

irenab

joehuang: so there iwll be logical L2 gateway instance, right?

it will be connected to L2 network(s)

and VTEP end points (+ tag)

joehuang

fake port can be tied to VTEP IP but not host, so that the L2GW ip could be the VTEP ip of lots of remote port (which is fake port in local neutron)

irenab

what is fake port is going to serve?

joehuang

to irenab: yes, logical L2GW, reprensented by VTEP ip

gampel1

1. IP is L3, not L2

gampel1

2. how will the l2GW know to take ownership of the VTEP IP? (you need some kind of API and then it needs to bind it like a regular port into neutron, so I'm not sure what the advantage would be)

am I missing sonething?

joehuang

fake port is to server as a virtual port inside local neutron L2 network, and leverage the L2 population

gampel1

Please give an example of the entire flow

irenab

I think we should look on both operator and user flows, just to see what is deploy phase tasks and what is run time

gampel1

definitely

joehuang

ok. what's your idea on L2 connectivity across Neutron via L2GW

I can describe it in mail-list

gampel1

The idea is that L2GW binds to the port (real or fake, I don't care) and only then L2Population can automatically reach the L2GW. Without the binding operation I do not understand how this will work.

gampel1

We see a race condition between the port creation and the binding operation and due to this we think that the bottom service should initiate both the port creation and the binding operation (and in this way there is no race because one entity is in charge of the order of both operations)

zhiyuan

For example, OS1 has a vm in net1, when creating a vm in net1 in OS2, we will create a \*fake\* port in OS1, so OS1 is aware of the port in OS2 and can setup tunnel to OS2. That's how \*fake\* port is used in POC.

not sure \*fake\* port is necessary if we use L2GW

irenab

zhipeng: fake port is per L2 network?

gampel1

so you bind the fake port to the tunnel port ?

irenab

zhiyuan: sorry, the question *^^,* is for you :-)

zhiyuan

\*fake\* port is created in each cascaded OS in that network

zhipeng

irenab per OS instance

joehuang

fake port binding to tunnel port, if L2GW used, binding to L2GW

gampel1

so you mean there is binding on every compute node ?

joehuang

if no L2GW, it's point to point tunneling, so every compute node

if L2GW used, point to point tunneling among L2GWs

gampel1

exactly! this is why we need the GW (otherwise it does not scale)

zhiyuan

irenab: zhipeng also answers for me *:)*

gampel1

now, in order for L2GW to work, it has to bind to the ports io

in order to let Neutron know where the port is and then make sure that traffic would reach the right place (i.e. the gw)

gampel1

If we agree on this, the next thing to understand is that the API to do these operations does not exists in L2GW

and this is the first gap we need to cover

zhiyuan

so L2GW is project we would like to integrate?

joehuang

to enhance the L2GW or propose an extension in Neutron?

gampel1

Not sure about L2GW. Because the API is not complete we need to engage that community and see if they are willing to enhance the API, If yes, then that would be great, if not we would need to define a new inter-cloud gw project

joehuang

as you said, the L2GW project is for overlay network to physical network connectivity

irenab

aacoring to their wiki, its ‘In a general sense, L2 gateway is an entity or resource which bridges two L2 domains (or networks) to achieve one seamless L2 broadcast domain’.

gampel1

The reason we are at all considering L2GW as an option is because we see their problem as a subset of ours so we think it would make sense to have a single API. But if they only want to solve the specific problem of virtual to physical then we will have no choice

joehuang

if neutron accept the BP for virtual port( fake port), it would be better to enhance the Neutron

gampel1

I would like to try and focus on what irenab suggested before - focus on the APIs that we are missing

after we map the gaps then we can decide if L2GW is the right place or not

irenab

gampel1: I think priority should be to leverage l2gw

joehuang

great

I have to leave now

let's continue the work

gampel1

i agree lets try to define what we are missing in the API

lets continue in another day

joehuang

good

gampel1

joe when will be a good time we could do it earlier your time

joehuang

Friday, same time?

sorry, Friday is not working day for you

gampel1

yes it is not

joehuang

So next Monday?

gampel1

for us

Ok for me

irenab

ok for me

joehuang

ok

gampel1

what time ?

joehuang

bye

same time

UTC 8:00

gampel1

so 8UTC ?

Ok

zhiyuan

fine for me

saggi

I'm good with that

zhipeng

me too

gampel1

lets all try to think what we are missing in the API