

Openstack裸金属节点跨网段访问镜像节点不通问题分析

openstack

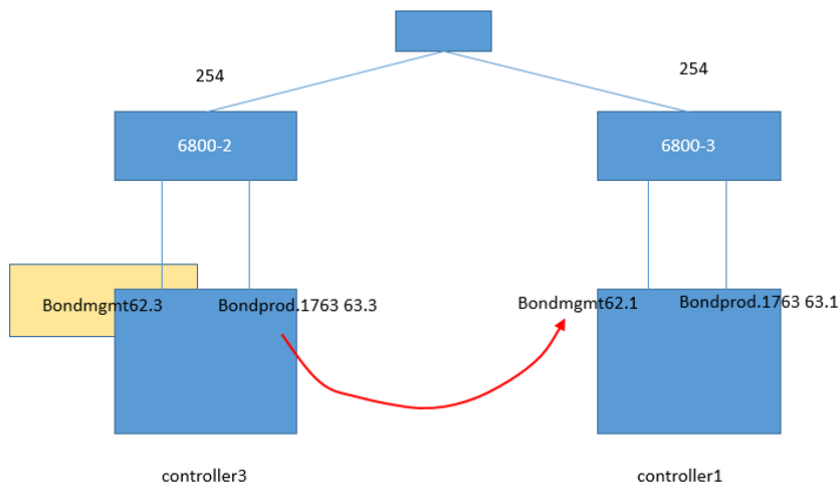
分布式网关

Fabric1.5

陈恺 2018-09-28 发表

组网及说明

组网说明



用controller3模仿裸金属节点在Inspect阶段访问openstack控制节点上的管理API，Inspect阶段的Inspect网络为10.15.63.0/24网段，管理API为10.15.62.0/24网段。

问题描述

问题现象

Controller 3 ping -I 10.15.63.3 10.15.62.1 Controller 1不通。（Linux ping -I为带源地址）

裸金属节点不能通过Ironic API节点去获取镜像，裸金属业务异常。

过程分析

问题分析

通过在Controller 1节点上转包分析，ICMP报文已经通过管理API到达Controller 1，但是62.1网卡没有回应。

No.	Time	Source	Destination	Protocol	Length	Info
270	0.303739	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=277/5377, ttl=64 (no response found)
1358	0.303738	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=278/5633, ttl=64 (no response found)
2265	2.114919	10.15.62.1	10.15.62.27	ICMP	98	Echo (ping) request id=0x6e45, seq=1/256, ttl=64 (reply in 2265)
2266	2.114991	10.15.62.27	10.15.62.1	ICMP	98	Echo (ping) reply id=0x6e45, seq=1/256, ttl=64 (request in 2265)
2581	2.303755	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=279/5889, ttl=64 (no response found)
3174	3.114781	10.15.62.1	10.15.62.27	ICMP	98	Echo (ping) request id=0x6e45, seq=2/512, ttl=64 (reply in 3174)
3176	3.114836	10.15.62.27	10.15.62.1	ICMP	98	Echo (ping) reply id=0x6e45, seq=2/512, ttl=64 (request in 3174)
3505	3.303744	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=280/6145, ttl=64 (no response found)
3947	4.115741	10.15.62.1	10.15.62.27	ICMP	98	Echo (ping) request id=0x6e45, seq=3/768, ttl=64 (reply in 3947)
3948	4.115803	10.15.62.27	10.15.62.1	ICMP	98	Echo (ping) reply id=0x6e45, seq=3/768, ttl=64 (request in 3947)
4023	4.303743	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=281/6401, ttl=64 (no response found)
4790	5.115789	10.15.62.1	10.15.62.27	ICMP	98	Echo (ping) request id=0x6e45, seq=4/1024, ttl=64 (reply in 4791)
4791	5.115843	10.15.62.27	10.15.62.1	ICMP	98	Echo (ping) reply id=0x6e45, seq=4/1024, ttl=64 (request in 4790)
5140	5.303741	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=282/6657, ttl=64 (no response found)
5879	6.303745	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=283/6913, ttl=64 (no response found)
6921	7.303746	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=284/7169, ttl=64 (no response found)
7751	8.303761	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=285/7425, ttl=64 (no response found)
9383	9.303741	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=286/7681, ttl=64 (no response found)
1081	10.303743	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=287/7937, ttl=64 (no response found)
1141	11.303743	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=288/8193, ttl=64 (no response found)
1221	12.303743	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=289/8449, ttl=64 (no response found)
1321	13.303753	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=290/8705, ttl=64 (no response found)
1371	14.303748	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=291/8961, ttl=64 (no response found)
1481	15.303758	10.15.63.3	10.15.62.1	ICMP	98	Echo (ping) request id=0x2e83, seq=292/9217, ttl=64 (no response found)

Controller 1 上的prod.1763发送的ARP请求（63网卡发出来的），该ARP请求源地址是10.15.62.1，目的地址是10.15.63.3。这种源地址和目的地址不在同一个网段的ARP请求是特殊的。

No.	Time	Source	Destination	Protocol	Length	Info
1	0.000000	InsputE1_5e:d3:54	Broadcast	ARP	42	Who has 10.15.63.3? Tell 10.15.62.1
2	1.000989	InsputE1_5e:d3:54	Broadcast	ARP	42	Who has 10.15.63.3? Tell 10.15.62.1
3	2.002988	InsputE1_5e:d3:54	Broadcast	ARP	42	Who has 10.15.63.3? Tell 10.15.62.1
4	3.999998	InsputE1_5e:d3:54	Broadcast	ARP	42	Who has 10.15.63.3? Tell 10.15.62.1
5	5.000989	InsputE1_5e:d3:54	Broadcast	ARP	42	Who has 10.15.63.3? Tell 10.15.62.1
6	6.002989	InsputE1_5e:d3:54	Broadcast	ARP	42	Who has 10.15.63.3? Tell 10.15.62.1
7	8.000005	InsputE1_5e:d3:54	Broadcast	ARP	42	Who has 10.15.63.3? Tell 10.15.62.1
8	9.000987	InsputE1_5e:d3:54	Broadcast	ARP	42	Who has 10.15.63.3? Tell 10.15.62.1
9	10.002988	InsputE1_5e:d3:54	Broadcast	ARP	42	Who has 10.15.63.3? Tell 10.15.62.1
10	11.999998	InsputE1_5e:d3:54	Broadcast	ARP	42	Who has 10.15.63.3? Tell 10.15.62.1
11	13.000986	InsputE1_5e:d3:54	Broadcast	ARP	42	Who has 10.15.63.3? Tell 10.15.62.1
12	14.002988	InsputE1_5e:d3:54	Broadcast	ARP	42	Who has 10.15.63.3? Tell 10.15.62.1

在controller 1 所在的68交换机上debug ARP，可见已经收到ARP请求报文，但是网关的ARP代理不能进行处理。

```
Jan 28 10:15:53:011 2011 B1DC-RC-Leaf6800-03 ARP/7/ARP_RCV: Received an ARP message, operation: 1, sender MAC: 6c92-bf5e-d39, sender IP: 10.15.62.1, target MAC: 0000-0000-0000, target IP: 10.15.63.3
Jan 28 10:15:53:007 2011 B1DC-RC-Leaf6800-03 ARP/7/ARP_RCV: Received an ARP message, operation: 1, sender MAC: 6c92-bf5e-d39, sender IP: 10.15.62.1, target MAC: 0000-0000-0000, target IP: 10.15.63.3
Jan 28 10:15:54:009 2011 B1DC-RC-Leaf6800-03 ARP/7/ARP_RCV: Received an ARP message, operation: 1, sender MAC: 6c92-bf5e-d39, sender IP: 10.15.62.1, target MAC: 0000-0000-0000, target IP: 10.15.63.3
Jan 28 10:15:55:011 2011 B1DC-RC-Leaf6800-03 ARP/7/ARP_RCV: Received an ARP message, operation: 1, sender MAC: 6c92-bf5e-d39, sender IP: 10.15.62.1, target MAC: 0000-0000-0000, target IP: 10.15.63.3
Jan 28 10:15:57:007 2011 B1DC-RC-Leaf6800-03 ARP/7/ARP_RCV: Received an ARP message, operation: 1, sender MAC: 6c92-bf5e-d39, sender IP: 10.15.62.1, target MAC: 0000-0000-0000, target IP: 10.15.63.3
Jan 28 10:15:58:009 2011 B1DC-RC-Leaf6800-03 ARP/7/ARP_RCV: Received an ARP message, operation: 1, sender MAC: 6c92-bf5e-d39, sender IP: 10.15.62.1, target MAC: 0000-0000-0000, target IP: 10.15.63.3
Jan 28 10:15:59:011 2011 B1DC-RC-Leaf6800-03 ARP/7/ARP_RCV: Received an ARP message, operation: 1, sender MAC: 6c92-bf5e-d39, sender IP: 10.15.62.1, target MAC: 0000-0000-0000, target IP: 10.15.63.3
Jan 28 10:16:01:007 2011 B1DC-RC-Leaf6800-03 ARP/7/ARP_RCV: Received an ARP message, operation: 1, sender MAC: 6c92-bf5e-d39, sender IP: 10.15.62.1, target MAC: 0000-0000-0000, target IP: 10.15.63.3
Jan 28 10:16:02:009 2011 B1DC-RC-Leaf6800-03 ARP/7/ARP_RCV: Received an ARP message, operation: 1, sender MAC: 6c92-bf5e-d39, sender IP: 10.15.62.1, target MAC: 0000-0000-0000, target IP: 10.15.63.3
Jan 28 10:16:03:011 2011 B1DC-RC-Leaf6800-03 ARP/7/ARP_RCV: Received an ARP message, operation: 1, sender MAC: 6c92-bf5e-d39, sender IP: 10.15.62.1, target MAC: 0000-0000-0000, target IP: 10.15.63.3
Jan 28 10:16:05:007 2011 B1DC-RC-Leaf6800-03 ARP/7/ARP_RCV: Received an ARP message, operation: 1, sender MAC: 6c92-bf5e-d39, sender IP: 10.15.62.1, target MAC: 0000-0000-0000, target IP: 10.15.63.3
Jan 28 10:16:06:009 2011 B1DC-RC-Leaf6800-03 ARP/7/ARP_RCV: Received an ARP message, operation: 1, sender MAC: 6c92-bf5e-d39, sender IP: 10.15.62.1, target MAC: 0000-0000-0000, target IP: 10.15.63.3
Jan 28 10:16:07:011 2011 B1DC-RC-Leaf6800-03 ARP/7/ARP_RCV: Received an ARP message, operation: 1, sender MAC: 6c92-bf5e-d39, sender IP: 10.15.62.1, target MAC: 0000-0000-0000, target IP: 10.15.63.3
```

解决方法

在VCFC控制器对接Openstack 裸金属功能典型配置手册中，Inspect网络的ARP代答建议关闭。如果ARP代答选择了关闭，默认会配置成代理模式。而代理模式网关vsi interface不能处理源目地址不在同一个网段的地址。所以需要将Inspect网络仍然配置为代答方式，让目的终端直接处理ARP请求。

修改虚拟链路网络 ✕

VDS	VDS2
类型	VXLAN
* 名称	PXE
Segment ID	1763

隐藏高级设置

网关ARP探测速率(pps)	0
ARP泛洪	开启
ARP代答	开启
ARP上送控制器	关闭
RARP泛洪	开启
RARP上送控制器	关闭
DHCP上送控制器	关闭
未知单播报文抑制	关闭
未知组播报文抑制	关闭
广播报文抑制	关闭
☒ 共享网络	
分布式部署	☒ 开启 ☐ 关闭
DHCP中继	☐ 开启 ☒ 关闭

应用 取消