存储配置 孔樱【技术大咖】 2020-06-25 发表

组网及说明

我以 Legacy T-Class 3PAR 在线迁移到 V-Class 组网案例

配置步骤

#3PAR Peer Motion Procedure

我们用T*Class 和 V*Class 之间做数据迁移。其中T_Class InformOS: 3.1.3 MU1 V-Class 3.2.2 MU4 两套存储端口设置为以下:

9:1 RCIP Port

ZONE 配置

verbatim:root> alicreate "V400PM1", "20:34:00:02:ac:00:1b:ae" verbatim:root> alicreate "V400PM2", "21:34:00:02:ac:00:1b:ae" verbatim:root> alicreate "T400PM1", "22:22:00:02:ac:00:17:2e" verbatim:root> alicreate "T400PM2", "22:32:00:02:ac:00:17:2e" verbatim:root> zonecreate

"T400PM1_V400PM1", "T400PM1;V400PM1" verbatim:root> zonecreate "T400PM2_V400PM2", "T4 00PM2;V400PM2" verbatim:root> cfgadd "CPZC_Storage",

"T400PM1_V400PM1;T400PM2_V400PM2" verbatim:root> cfgenable CPZC_Storage You are about to enable a new zoning configuration. This action will replace the old zoning configuration with the cur rent configuration selected. If the update includes changes to one or more traffic isolation zones, the update may result in localized disruption to traffic on ports associated with the traffic isolation zone ch anges Do you want to enable "CPZC_Storage" configuration (yes, y, no, n): [no] y zone config "CPZC _Storage" is in effect Updating flash ... verbatim:root>

在 V-Class 取两个 free port (Ready) 设置为 peer port 我们这里那 0: 3: 4 和 1:3:4 两个端口。 V400 cli% controlport offline 0:3:4 Are you sure you want to run controlport offline on port 0:3:4? sele ct q=quit y=yes n=no: y V400 cli% controlport offline 1:3:4 Are you sure you want to run controlport offline on port 1:3:4? select q=quit y=yes n=no: y V400 cli% controlport config peer -ct point 0:3:4 Are you sure you want to run controlport config peer -ct point on port 0:3:4? select q=quit y=yes n=no: y V 400 cli% controlport config peer -ct point 1:3:4 Are you sure you want to run controlport config peer -ct point on port 1:3:4? select q=quit y=yes n=no: y V400 cli% showport N:S:P Mode State ----Node_W WN---- Port_WWN/HW_Addr- Type Protocol Label Partner FailoverState 0:3:4 initiator ready 2FF702 02AC001BAE 20340202AC001BAE free FC - - - none 1:3:4 initiator pending_reset 2FF70202AC001B AE 21340202AC001BAE free FC - - -

特别注意一点这时设置为 Peer Port之后 端口 WWPN 发生变化了。

我们接下来把 T-Class 的 2:2:2 和 3:2:2 设置为 host port 这两个端口分别和 V-Class 的0:3:4 和 1:3:4 划了ZONE的。

T400 cli% controlport offline 2:2:2 Are you sure you want to run controlport offline on port 2:2:2? select t q=quit y=yes n=no: y T400 cli% controlport offline 3:2:2 Are you sure you want to run controlport offline on port 3:2:2? select q=quit y=yes n=no: y T400 cli%

这里要强调注意一点: 如果将端口设置为 Peer port 之后这个端口的 WWPN 会发生改变

我们将 T-Class 的端口类型改为 Host Port 后 WWPN就变回 00 而不是 02, 这时要重新调整ZONE。 T400 cli% showport N:S:P Mode State ----Node_WWN---- -Port_WWN/HW_Addr- Type Protocol Labe I Partner FailoverState 2:2:2 target ready 2FF70002AC00172E 22220002AC00172E free FC - 3:2:2 n one 3:2:2 target ready 2FF70002AC00172E 23220002AC00172E free FC - 2:2:2 none 重新调整 ZONE

verbatim:root> alicreate "V400PM1", "20:34:02:02:ac:00:1b:ae" verbatim:root> alicreate "V400PM2", "21:34:02:02:ac:00:1b:ae" verbatim:root> zonecreate "T400PM1_V400PM1", "T400PM1;V400PM1" v erbatim:root> zonecreate "T400PM2_V400PM2", "T400PM2;V400PM2" verbatim:root> cfgenable CP ZC_Storage

=-----V400 cli% showport N:S:P Mode State ----Nod e_WWN---- -Port_WWN/HW_Addr- Type Protocol Label Partner FailoverState 0:3:4 initiator ready 2F F70202AC001BAE 20340202AC001BAE peer FC - - - 1:3:4 initiator ready 2FF70202AC001BAE 213 40202AC001BAE peer FC - - -

我们在 T-Class 上看看端口类型,已经变为 Host Port

====== T400 cli% showport N:S:P Mode State ----Node_WWN---- -Port_WWN/HW_Addr- Type Protocol Label Partner FailoverState 2:2:2 target ready 2FF70002AC00172E 22220002AC00172E host FC - 3:2:2 none 3:2:2 target ready 2FF70002AC0017 2E 23220002AC00172E host FC - 2:2:2 none

这时我们是有VV映射给 T-Class 通过主机端口 2:5:1,3:5:1

T400 cli% showvlun -host BL460G6 Active VLUNs Lun VVName HostName -Host_WWN/iSCSI_Nam e- Port Type Status ID 0 pmvv1 BL460G6 1000000C9902674 2:5:1 host active 1 0 pmvv1 BL460G6 1000000C9902674 3:5:1 host active 1 ------- 2 total VLUN Templates Lun VVName HostName -Host_WWN/iSCSI_Name- Port Type 0 pmvv1 BL460G6 -------------- host ------- 1 total T400 cli%

我们回到 V-Class 看 T-Class 都有哪些 LUN、

V400 cli% showtarget -lun all ---Size(MB)--- Lun ----Node_WWN----- Lun_WWN----- Device_Type -VSize- -Rsvd- --MFR--- -State- -Vendor_Specific- 254 23220002AC00172E 50002ACFFFFF172E ses _device 0 0 3PARdata -- unknown 254 22220002AC00172E 50002ACFFFFF172E ses_device 0 0 3P ARdata -- unknown V400 cli%

下面我在配置Peer Motion之前,先确认配置的 Peer Port并没有VV 映射过来。(0:3:4,1:3:4 没有VV映射)

V400 cli% showvlun -host BL460G6 Active VLUNs Lun VVName HostName -

接下来我们需要将 Peer Port 创建主机 (在错误一侧存储配置,端口显示"---"状态)

V400 cli% createhost pmhost 20340202AC001BAE 21340202AC001BAE Setting default host person a 2 (Generic-ALUA) V400 cli% V400 cli% showhost Id Name Persona -WWN/iSCSI_Name- Port 13 p mhost Generic-ALUA 20340202AC001BAE --- 21340202AC001BAE --- V400 cli%

上面做法是错误的,原来不应该在目标端存储(V-Class)创建主机,而应该是在源端存储 T-Class 创建主机。

T400 cli% createhost pmhost 21340202AC001BAE 20340202AC001BAE Setting default host person a 2 (Generic-ALUA) T400 cli% showhost ld Name Persona -WWN/iSCSI_Name- Port 15 pmhost Gen eric-ALUA 21340202AC001BAE 3:2:2 20340202AC001BAE 2:2:2 T400 cli%

然后我们需要在 T-Class 源端存储将需要同步的 VV 映射给新创建的主机 pmhost

T400 cli% createvlun pmvv1 auto pmhost Warning: Virtual volume pmvv1 is already exported. Unless this volume is being exported to a clustered environment, creation of multiple exports could lead to da ta corruption. To avoid accidental creation of multiple exports, consider setting the virtual volume policy to restrict export to a single host. Do you still wish to continue with the export? select q=quit y= yes n=no: y T400 cli% showvlun Active VLUNs Lun VVName HostName -Host_WWN/iSCSI_Name-Port Type Status ID 0 pmvv1 BL460G6 1000000C9902674 2:5:1 host active 1 0 pmvv1 BL460G6 10 00000C9902674 3:5:1 host active 1 0 pmvv1 pmhost 20340202AC001BAE 2:2:2 host active 1 0 pmv v1 pmhost 21340202AC001BAE 3:2:2 host active 1 T400 cli% showvlun Active VLUNs Lun VVName HostName -Host_WWN/iSCSI_Name-Port Type Status ID 0 ky_winvv1.r BL460G7 5001438021DD2 AC8 2:5:1 host active 258 0 ky_winvv1.r BL460G7 5001438021DD2AC8 3:5:1 host active 258 0 pmvv 1 pmhost 20340202AC001BAE 2:2:2 host standby 1 0 pmvv1 pmhost 21340202AC001BAE 3:2:2 host s

我们这时可以在客户端主机上创建 LVM

[root@rhel7 ~]# multipath -ll mpathe (350002ac00887172e) dm-6 3PARdata,VV size=30G features=" 1 queue_if_no_path" hwhandler="0" wp=rw `-+- policy="round-robin 0" prio=50 status=active |- 1:0:0: 0 sdh 8:112 active ready running `- 1:0:5:0 sdi 8:128 active ready running [root@rhel7 ~]# 为了便于识别,我在 /etc/multipath.conf 将这个 VV 起一个别名pmvv

[root@rhel7 etc]# multipath -II pmvv (350002ac00887172e) dm-6 3PARdata,VV size=30G features=" 1 queue_if_no_path" hwhandler="0" wp=rw `-+- policy="round-robin 0" prio=50 status=active |- 1:0:0: 0 sdh 8:112 active ready running `- 1:0:5:0 sdi 8:128 active ready running [root@rhel7 etc]# [root@rh el7 etc]# pvcreate /dev/mapper/pmvv Physical volume "/dev/mapper/pmvv" successfully created. [root@rhel7 etc]# vgcreate pmvg /dev/mapper/pmvv Volume group "pmvg" successfully created [root @rhel7 etc]# lvcreate -I 100%FREE pmvg pmlv Physical Volume "pmlv" not found in Volume Group " pmvg". [root@rhel7 etc]# lvcreate -I 100%VG -n pmlv pmvg Logical volume "pmlv" created. [root@rhel7 etc]# mkfs.ext4 /dev/pmvg/pmlv mke2fs 1.42.9 (28-Dec-2013) Discarding device blocks: f ailed - Input/output error Filesystem label= OS type: Linux Block size=4096 (log=2) Fragment size=4 096 (log=2) Stride=4 blocks, Stripe width=4096 blocks 1966080 inodes, 7860224 blocks 393011 bloc ks (5.00%) reserved for the super user First data block=0 Maximum filesystem blocks=2155872256 2 40 block groups 32768 blocks per group, 32768 fragments per group 8192 inodes per group

Superblock backups stored on blocks: 32768, 98304, 163840, 229376, 294912, 819200, 884736, 160 5632, 2654208, 4096000 Allocating group tables: done Writing inode tables: done Creating journal (3 2768 blocks): done Writing superblocks and filesystem accounting information: done [root@rhel7 etc] # [root@rhel7 etc]# mount -o discard /dev/pmvg/pmlv /pmvv [root@rhel7 etc]# df -kh Filesystem Size Used Avail Use% Mounted on /dev/mapper/pmvg-pmlv 30G 45M 28G 1% /pmvv [root@rhel7 etc]#

这时我们在目标端存储通过 showtarget -lun all 可以看到源端存储映射过来的 VV,可以看到卷类型,卷大小。

V400 cli% showtarget -lun all ---Size(MB)--- Lun ----Node_WWN----- Lun_WWN----- Device_Type -VSize- -Rsvd- --MFR--- -State- -Vendor_Specific- 0 23220002AC00172E 50002AC00887172E disk 3 0720 512 3PARdata -- tpvv 254 23220002AC00172E 50002ACFFFFF172E ses_device 0 0 3PARdata -- unknown 0 22220002AC00172E 50002AC00887172E disk 30720 512 3PARdata -- tpvv 254 22220002AC00172E 50002ACFFFFF172E ses_device 0 0 3PARdata -- unknown V400 cli% 但这时我们看到 State 状态还是为空,这样在目标端存储 admitvv 就会提示找不到源存储 VV WWN V400 cli% admitvv pmvv1:50002AC00887172E:50202AC00887172E Remote VV WWN 50002AC00887172E does not exist 0 VVs admitted. V400 cli%

这个时候我们必须要强制扫描一次,通过 showtarget -lun all 看到对端VV的状态为 discovered 才正确 V400 cli% showtarget -force -rescan V400 cli% showtarget -lun all ---Size(MB)--- Lun ----

Node_WWN----- Lun_WWN----- Device_Type -VSize- -Rsvd- --MFR--- -State--- -Vendor_Specific-0 23220002AC00172E 50002AC00887172E disk 30720 512 3PARdata discovered tpvv 254 2322000 2AC00172E 50002ACFFFF172E ses_device 0 0 3PARdata -- unknown 0 22220002AC00172E 50002AC00887172E disk 30720 512 3PARdata discovered tpvv 254 22220002AC00172E

50002ACFFFF172E ses_device 0 0 3PARdata -- unknown V400 cli%

我们这个时候才再次提交这个卷,提交成功

V400 cli% admitvv pmvv1:50002AC00887172E 1 VVs admitted. V400 cli%

我们在目标端存储查看该 VV 状态为 peer类型

%

接下来我们在目标端存储将该 VV 映射给服务器

V400 cli% createvlun pmvv1 auto BL460G6 V400 cli% showvlun -host BL460G6 Active VLUNs Lun V VName HostName -Host_WWN/iSCSI_Name- Port Type Status ID 0 pmvv1 BL460G6 10000000C99 02674 0:2:1 host active 2 0 pmvv1 BL460G6 10000000C9902674 0:2:2 host active 2 0 pmvv1 BL460 G6 1000000C9902674 1:2:1 host active 2 0 pmvv1 BL460G6 1000000C9902674 1:2:2 host active 2 -------

最后我们要执行 importvv 动作就是将 T-Class 上的 pmvv1 迁移到 V-Class 的 pmvv1, 在迁移同时我要 对这个VV对应的 LVM 持续写入 20G 随机数据,看在写入数据同时做数据迁移有什么变化。

dd if=/dev/urandom of=/pmvv/pmvv.urandom bs=1M count=20000

[root@rhel7 ~]# dd if=/dev/urandom of=/pmvv/pmvv.urandom bs=1M count=20000 & [1] 29578 [root @rhel7 ~]#

再执行importvv 时要求再次提交,因为这时的State状态为discovered,应该正确为 admitted. V400 cli% importvv -tpvv FC_r5 pmvv1 Importing VVs pmvv1 select q=quit y=yes n=no: y Rescan ne eded on host BL460G6(1000000C9902674) on port 0:2:2 for volume 50002AC00887172E V400 cli % V400 cli% showtarget -force -rescan V400 cli% showtarget -lun all ---Size(MB)--- Lun ----Node_W WN---- Device_Type -VSize -Rsvd- --MFR--- State-- -Vendor_Specific- 0 2322000 2AC00172E 50002AC00887172E disk 30720 14848 3PARdata admitted tpvv 254 23220002AC00172E 50002ACFFFFF172E ses_device 0 0 3PARdata -- unknown 0 22220002AC00172E 50002AC00887172E disk 30720 14848 3PARdata admitted tpvv 254 2222002 AC00172E 50002ACFFFFF172E ses_device 0 0 3PARdata -- unknown 0 22220002AC00172E 50002ACFFFFF172E ses_device 0 0 3PARdata -- unknown V400 cli% 我们看主机的路径多了一倍,这时是必须要在主机这端做一个扫描,否则总报Rescan..... [root@rhel7 ~]# multipath -II pmvv (350002ac00887172e) dm-6 3PARdata,VV size=30G features="1 queue_if_no_path" hwhandler="0" wp=rw `-+- policy="round-robin 0" prio=50 status=active |- 1:0:0:0 sdh 8:112 active ready running |- 1:0:5:0 sdi 8:128 active ready running |- 1:0:1:0 sdj 8:144 active read dy running |- 1:0:2:0 sdk 8:160 active ready running |- 1:0:3:0 sdl 8:176 active ready running `- 1:0:4:0

sdm 8:192 active ready running [root@rhel7 ~]#

我们再次提交,成功。

V400 cli% importvv -tpvv FC_r5 pmvv1 Importing VVs pmvv1 select q=quit y=yes n=no: y Task

2019-06-20 21:54:19 CST Storing task data for later restarts. 2019-06-20 21:54:19 CST Importing to CPG "FC_r5" as tpvv. 2019-06-20 21:54:19 CST Started region move of 256MB from (pmvv1.rmt.0:0 MB) to (pmvv1.ldv.0:0MB). 2019-06-20 21:54:19 CST Started region move of 256MB from (pmvv1.rm t.0:256MB) to (pmvv1.ldv.0:256MB). 2019-06-20 21:54:19 CST Started region move of 256MB from (pmvv1.rmt.0:512MB) to (pmvv1.ldv.0:512MB). 2019-06-20 21:54:19 CST Started region move of 256 MB from (pmvv1.rmt.0:512MB) to (pmvv1.ldv.0:768MB). 2019-06-20 21:54:19 CST Started region move of 256 MB from (pmvv1.rmt.0:512MB) to (pmvv1.ldv.0:768MB).

我们查看当前 V-Class 状态,为 Importing,需要等到结束。

V400 cli% showtarget -lun all ---Size(MB)--- Lun ----Node_WWN----- Lun_WWN----- Device_Type -VSize- -Rsvd- --MFR--- --State-- -Vendor_Specific- 0 23220002AC00172E 50002AC00887172E disk 30720 19456 3PARdata importing tpvv 254 23220002AC00172E 50002ACFFFFF172E ses_device 0 0 3PARdata -- unknown 0 22220002AC00172E 50002AC00887172E disk 30720 19456 3PARdata im porting tpvv 254 22220002AC00172E 50002ACFFFF172E ses_device 0 0 3PARdata -- unknown V4 00 cli% V400 cli% showvv ----Rsvd(MB)----- -(MB)-- Id Name Prov Type CopyOf BsId Rd --Detailed_S tate--- Adm Snp Usr VSize 1 .srdata full base --- 1 RW normal 0 0 81920 81920 0 admin full base --- 0 RW normal 0 0 2183 pmvv1 peer base --- 2183 RW importing,exclusive 0 0 30720 30720 直到所有region都迁移完成:

V400 cli% showtask -d 27085 ld Type Name Status Phase Step ------StartTime--

4080 V400 cli% V400 cli% showvlun Active VLUNs Lun VVName HostName -Host_WWN/iSCSI_Na me- Port Type Status ID 0 pmvv1 BL460G6 1000000C9902674 0:2:1 host active 2 0 pmvv1 BL460G 6 1000000C9902674 0:2:2 host active 2 0 pmvv1 BL460G6 1000000C9902674 1:2:1 host active 2 0 pmvv1 BL460G6 1000000C9902674 1:2:2 host active 2 V400 cli% showtarget -lun all ---Size(MB)--- Lun ----Node_WWN---- ---Lun_WWN---- Device_Type -VSize- -Rsvd- --MFR--- -State-- -Vendor_S pecific- 0 23220002AC00172E 50002AC00887172E disk 30720 31360 3PARdata imported tpvv 254 23220002AC00172E 50002AC00887172E disk 30720 31360 3PARdata -- unknown 0 22220002AC00172E 50002AC00887172E disk 30720 31360 3PARdata imported tpvv 254 22220002

AC00172E 50002ACFFFF172E ses_device 0 0 3PARdata -- unknown V400 cli%

上面的操作就完成了数据迁移。

检查主机端多路径状态

[root@rhel7 pmvv]# multipath -ll Jun 21 07:31:58 | sdc: couldn"t get asymmetric access state Jun 21 07:31:58 | sdd: couldn"t get asymmetric access state pmvv (350002ac0088d172e) dm-2 3PARdata,V V size=50G features="1 queue_if_no_path" hwhandler="0" wp=rw `-+- policy="round-robin 0" prio=50 status=active |- 1:0:0:0 sdc 8:32 active faulty running |- 1:0:1:0 sdd 8:48 active faulty running |- 1:0:2:0 sdj 8:144 active ready running |- 1:0:3:0 sdk 8:160 active ready running |- 1:0:4:0 sdl 8:176 active rea dy running `- 1:0:5:0 sdm 8:192 active ready running [root@rhel7 pmvv]# 以上操作都是在线操作,对业务没有任何影响。

配置关键点

我们将 T-Class 的端口类型改为 Host Port 后 WWPN就变回 00 而不是 02 ,这时要重新调整ZONE。 就是说如果将端口设置为 Peer Port 之后,该端口的WWPN就会改变。