

知 3PAR persistent port 功能可能会触发部分主机口性能瓶颈

存储系统 张思群 2023-05-18 发表

组网及说明

3PAR Primera Alletra9000

问题描述

存储出现性能问题，经过检查高延迟发生在特定的前端端口上，因为该端口上收发的数据量为其他端口的两倍，已经达到端口处理能力阈值，造成了端口的拥堵。

```
ss20800 cli% statport -ni -host -iter 1
17:14:36 05/17/2023 r/w I/O per second          KBytes per sec          Svt ms          IOSz KB
Port      D/C      Cur   Avg  Max      Cur   Avg  Max      Cur   Avg      Cur   Avg Qlen
-----
5:5:3     Data    t    4103 4103 4103  707103 707103 707103  0.78  0.78  172.3 172.3  1
5:5:4     Data    t   4092 4092 4092  749326 749326 749326  1.35  1.35  183.1 183.1  2
6:5:1     Data    t     78  78  78    6068  6068  6068  0.38  0.38  78.3  78.3  0
6:5:2     Data    t     76  76  76    5934  5934  5934  0.36  0.36  78.0  78.0  0
6:5:4     Data    t   4811 4811 4811  959051 959051 959051  1.65  1.65  199.3 199.3  8
7:5:1     Data    t     77  77  77    5147  5147  5147  0.39  0.39  67.2  67.2  0
7:5:2     Data    t     75  75  75    5988  5988  5988  0.40  0.40  79.8  79.8  0
7:5:3     Data    t   9612 9612 9612 1609200 1609200 1609200  43.88 43.88 167.4 167.4 450
7:5:4     Data    t   4816 4816 4816  775861 775861 775861  1.49  1.49  161.1 161.1  7
-----
28        Data    t  54915 54915          9976003 9976003          8.67  8.67  181.7 181.7  483
```

主机端多路径配置正常，能够在存储路径上实现负载均衡，所以端口数据量超阈值有其他原因

过程分析

存储当前的状态为端口6:5:3故障暂时断开连线，由于存储开启了persistent port 功能，所以在6:5:3故障期间wwn已经failed over到同一个节点对的对应端口上（7:5:3）

N:S:P	Mode	State	----Node_WWN----	-Port_WWN/HW_Addr-	Type	Protocol	Label	Partner	FailoverState
6:5:1	target	ready	2FF70002AC01FB2E	26510002AC01FB2E	host	FC	-	7:5:1	none
6:5:2	target	ready	2FF70002AC01FB2E	26520002AC01FB2E	host	FC	-	7:5:2	none
6:5:3	target	loss_sync	2FF70002AC01FB2E	26530002AC01FB2E	free	FC	-	7:5:3	failed_over
6:5:4	target	ready	2FF70002AC01FB2E	26540002AC01FB2E	host	FC	-	7:5:4	none
6:9:1	peer	offline	-	480FCFA2E421	free	IP	IP6	-	-
7:0:1	initiator	ready	50002ACFF701FB2E	50002AC70101FB2E	disk	SAS	DP-1	-	-
7:0:2	initiator	ready	50002ACFF701FB2E	50002AC70201FB2E	disk	SAS	DP-2	-	-
7:0:3	initiator	ready	50002ACFF701FB2E	50002AC70301FB2E	disk	SAS	DP-3	-	-
7:0:4	initiator	loss_sync	50002ACFF701FB2E	50002AC70401FB2E	free	SAS	DP-4	-	-
7:1:1	initiator	ready	50002ACFF701FB2E	50002AC71101FB2E	disk	SAS	DP-1	-	-
7:1:2	initiator	ready	50002ACFF701FB2E	50002AC71201FB2E	disk	SAS	DP-2	-	-
7:1:3	initiator	ready	50002ACFF701FB2E	50002AC71301FB2E	disk	SAS	DP-3	-	-
7:1:4	initiator	loss_sync	50002ACFF701FB2E	50002AC71401FB2E	free	SAS	DP-4	-	-
7:2:1	initiator	ready	50002ACFF701FB2E	50002AC72101FB2E	disk	SAS	DP-1	-	-
7:2:2	initiator	ready	50002ACFF701FB2E	50002AC72201FB2E	disk	SAS	DP-2	-	-
7:2:3	initiator	ready	50002ACFF701FB2E	50002AC72301FB2E	disk	SAS	DP-3	-	-
7:2:4	initiator	loss_sync	50002ACFF701FB2E	50002AC72401FB2E	free	SAS	DP-4	-	-
7:5:1	target	ready	2FF70002AC01FB2E	27510002AC01FB2E	host	FC	-	6:5:1	none
7:5:2	target	ready	2FF70002AC01FB2E	27520002AC01FB2E	host	FC	-	6:5:2	none
7:5:3	target	ready	2FF70002AC01FB2E	27530002AC01FB2E	host	FC	-	6:5:3	active
7:5:4	target	ready	2FF70002AC01FB2E	27540002AC01FB2E	host	FC	-	6:5:4	none
7:9:1	peer	offline	-	480FCFA2E41D	free	IP	IP7	-	-

该功能可以保障主机端感应不到存储端口故障，操作系统的多路径软件不会因为存储物理端口故障而发生切换。

但这个功能激活后，路径数量不变，多路径软件会保持端口间负载均衡，将6:5:3和7:5:3的数据都通过物理端口7:5:3路径一并发送，这就导致了7:5:3端口的数据量翻倍，最终达到端口处理能力的阈值。

解决方法

关闭persistent port功能，并将6:5:3failover back，触发主机多路径切换，使原本6:5:3的数据均分到其他有效端口上

1. 关闭persistent port功能
2. 重启故障端口controlport rst x:x:x

