

知识库文章：000463942

VPLEX：从 5.5 u2 版起，ESXi 主机上随机出现临时性存储设备连接断开和/或性能下降 (000463942)

主要产品：VPLEX 系列

产品：VPLEX 系列、VPLEX Geo、VPLEX Local、VPLEX Metro、VPLEX VS1、VPLEX VS2、VPLEX Virtual Edition、VPLEX GeoSynchrony 5.1、VPLEX GeoSynchrony 5.1 Patch 1、VPLEX GeoSynchrony 5.1 Patch 2、VPLEX GeoSynchrony 5.1 Patch 3、VPLEX GeoSynchrony 5.1 Patch 4

版本：4 文章类型：故障修复 目标受众：级别 30 = 客户 上次发布时间：2016 年 11 月 30 日星期三 17:09:36 (格林威治标准时间)

摘要： 由于 VMFS 心跳信号插槽上发生 ATS 失配，ESXi 主机尝试重获设备的控制权。为此，主机将在包含该 VMFS 的 LUN 上发起 SCSI 设备重置。此 LUN 上的所有活动的 I/O 将会中止，SCSI 设备将会重置。VMKernel 日志中将显示连接临时中断。

问题： ESXi 主机在一段短时间内断开到 VMFS 数据存储区的连接。在此期间，数据存储区的任何虚拟机可能出现故障或出现 I/O 错误。

需要在 VMFS3 或 VMFS5 数据存储区上使用比较和写入 (SCSI 操作代码 89) 获得心跳信号锁定的 ESXi 主机请求，会因为“(ATS) Miscompare during verify operation” (验证操作期间 ATS 不匹配) 而失败。

由于 VMFS 心跳信号插槽上的这种 ATS (原子测试和设置) 失配，ESXi 主机会尝试重新获得设备的控制权。为此，主机将在包含该 VMFS 的 LUN 上发起 SCSI 设备重置。

此 LUN 上的所有活动的 I/O 将会中止，SCSI 设备将会重置。

在 NMP 和 PowerPath 中都可能发生 ATS 失配。

您在主机 VMKernel 日志中将看到类似于以下的事件：

```
2015-09-30T22:13:55.516Z cpu1:33645)ScsiDeviceIO: 2338: Cmd(0x413686250680) 0x89, CmdSN 0x12b from world 32949 to dev "naa.600014400000000XXXXXXXXXXXXXXXXXXXX" failed H:0x0 D:0x2 P:0x0 Valid sense data:0xe 0x1d 0x0.
```

对于 NMP 主机您可能还会看到：

```
2015-09-30T22:13:55.516Z cpu1:33645)NMP: nmp_ThrottleLogForDevice:2321: Cmd 0x89 (0x413686450680, 32949) to dev "naa.600014400000000XXXXXXXXXXXXXXXXXXXX" on path "vmhba2:C0:T5:L13" Failed: H:0x0 D:0x2 P:0x0 Valid sense data: 0xe 0x1d 0x0. Act:NONE
```

这些事件意味着 VPLEX 向主机 (验证操作期间失配) 返回感知数据 **0E/1D/00**，以响应由 VMware 硬件辅助锁定机制 (ATS) 使用的 SCSI 操作代码 89 (比较和写入)。

原子测试和设置 (ATS)

这是一种锁定机制，它在进行元数据更新时取代 VMFS 卷上的 SCSI 保留。基本上，ATS 锁定可视为一种磁盘扇区修改机制，若修改成功，将允许 ESXi 主机执行对 VMFS 的元数据更新。这包括在资源调配期间向 VMDK 分配空间，因为元数据中的某些特征需要更新，以体现新的文件大小。有趣的是，在 VAAI 初始版本中，ATS 原语在每个存储阵列上必须以不同的方式实现，所以您的 ATS 操作码因供应商而异。ATS 现在是标准 T10，并使用操作码 0x89 (比较和写入)。

对于在支持 VAAI 的阵列上格式化的 VMFS5 数据存储区，使用 ATS 进行心跳信号锁定。VAAI 启用的 VMFS5 上不应再有任何 SCSI 保留。即使存在争用，也将继续使用 ATS。在非 VAAI 阵列上，继续使用 SCSI 保留在 VMFS5 中建立关键扇区。

在 VPLEX 固件日志上可看到以下事件：

主机 HBA 注销 (**tach/38**, **stdf/18**) 和重新登录 (**tach/37**, **stdf/17**)：

```
128.221.252.67/cpu0/log:5988:W/"0060165f237510728-2":23768:<6>2015/09/18 06:53:24.45: (A0-FC01): login with
0x1234567890123456 (nPortId 0x012345) type TGT is closing.
128.221.252.68/cpu0/log:5988:W/"0060165e9a38102140-2":46706:<6>2015/09/18 06:53:24.45: (B0-FC01): login
with 0x1234567890123456 (nPortId 0x012345) type TGT is closing.
128.221.252.68/cpu0/log:5988:W/"0060165e9a38102140-2":46707:<4>2015/09/18 06:53:24.45: connection lost. IT:
[EXAMPLESERVER_HBA1 (0x1234567890123456) B0-FC01 (0x50001442b035
128.221.252.67/cpu0/log:5988:W/"0060165f237510728-2":23769:<4>2015/09/18 06:53:24.45: connection lost. IT:
[EXAMPLESERVER_HBA1 (0x1234567890123456) A0-FC01 (0x50001442a035

128.221.252.68/cpu0/log:5988:W/"0060165e9a38102140-2":46708:<6>2015/09/18 06:53:24.45: (B0-FC01): login
with 0x1234567890123456 (nPortId 0x012345) type TGT is ready.
128.221.252.67/cpu0/log:5988:W/"0060165f237510728-2":23770:<6>2015/09/18 06:53:24.45: (A0-FC01): login with
0x1234567890123456 (nPortId 0x012345) type TGT is ready.
128.221.252.67/cpu0/log:5988:W/"0060165f237510728-2":23771:<4>2015/09/18 06:53:24.45: connection
established. IT: [EXAMPLESERVER_HBA1 (0x1234567890123456) A0-FC01 (0x50001
128.221.252.68/cpu0/log:5988:W/"0060165e9a38102140-2":46709:<4>2015/09/18 06:53:24.45: connection
established. IT: [EXAMPLESERVER_HBA1 (0x1234567890123456) B0-FC01 (0x50001
```

由于主机 HBA 重置 (登出/登入) 而接收到注册状态更改通知 (RSCN) (tach/42)，通知前带有字符串 “TGT_LGN_FR”：

```
128.221.252.68/cpu0/log:5988:W/"0060165e9a38102140-2":46710:<6>2015/09/18 06:53:26.50: (B0-FC00): finished
discovery in 58.650 msec, reason to start: TGT_LGN_FROM_UNKNOWN_NP D, result: succeeded
128.221.252.67/cpu0/log:5988:W/"0060165f237510728-2":23772:<6>2015/09/18 06:53:26.52: (A0-FC00): finished
discovery in 61.626 msec, reason to start:
TGT_LGN_FROM_UNKNOWN_NPID:RSCN_RECEIVED
, result: succeeded
```

应当也有正常的 **RSCN_RECEIVED** 消息 (tach/42)：

```
128.221.252.68/cpu0/log:5988:W/"0060165e9a38102140-2":46711:<6>2015/09/18 06:53:26.64: (B0-FC01): finished
discovery in 63.309 msec, reason to start: RSCN_RECEIVED, result:
128.221.252.67/cpu0/log:5988:W/"0060165f237510728-2":23773:<6>2015/09/18 06:53:26.66:
(A0-FC01): finished discovery in 62.665 msec, reason to start: RSCN_RECEIVED, result:
```

SCSI 操作码 89 (比较和写入) 主机中止 (**stdf/10**，状态码以 “89” 开头)：

```
firmware.log_20150123073924.1:128.221.252.67/cpu0/log:5988:W/"0060165f237510728-2":235 11:35:43.15:
stdf/10 Scsi Tmf [Abort Task] on fcp ITLQ: [EXAMPLESERVER_HBA1 (0x1234567 A0-FC01 (0x50001442a0353d01)
0x7d000000000000 0x2ed] vol VIRTUAL_VOLUME_NAME_vol taskElapsedTime(usec) 7996921 dormantQCnt 5
enabledQCnt 1 status 890000000000100:0
firmware.log_20150123073924.1:128.221.252.67/cpu0/log:5988:W/"0060165f237510728-2":235 11:35:44.15:
stdf/10 Scsi Tmf [Abort Task] on fcp ITLQ: [EXAMPLESERVER_HBA1 (0x1234567 A0-FC01 (0x50001442a0353d01)
0x7d000000000000 0x55d] vol VIRTUAL_VOLUME_NAME_vol taskElapsedTime(usec) 929473 dormantQCnt 6
enabledQCnt 1 status 890000000000100:0
```

事件解释：

tach/38：FC 登录正在关闭。

stdf/18：因从结构注销或分离而使 FCP 启动器端口与目标端口的连接中断时，即生成此日志消息。

tach/37：FC 登录准备好为 IO 提供服务。

tach/42：最近完成的 FC 发现过程的摘要。

stdf/10：主机在中止一个 IO。如果主机仍不满意，则主机将上报到“逻辑单元重置”和“目标重置” TMF。

原因：实际上有两个问题，它们均由 ATS 心跳信号功能触发：

1. 许多阵列供应商 (包括 EMC 在内) 在 ESXi 5.5u2 中引入的 ATS 心跳信号功能上都有问题。

注意：根据 VMware KB 2113956 可知，此问题影响 ESX 版本 VMware ESXi 5.5.x 和 VMware ESXi 6.0.x，但未提供所有具体版本，所以本知识库文章假设所有安装有 5.5u2 及更高版本和所有 vSphere 6.0 版本的 ESXi 主机都会受到影响。

VMware vSphere 版本 5.5.0 Update 2 (内部版本 2068190) 及更高版本和 vSphere 6.0 及更高版本使用原子测试和设置 (ATS) 锁定 VMFS 心跳。在 5.5.0 u2 之前的版本中，使用 SCSI-2 非持久保留实现此目的。

主机通过定期执行到给定卷上其心跳信号的 I/O，来表明其活性。因此，如果在此时间段内看不到主机的心跳插槽上有任何活动，我们就可以得出结论，该主机已失去与该卷的连接。

ATS 心跳 I/O 的超时值非常低，这会导致主机断开连接和应用程序中断，进而导致失去与磁盘的连接和/或主机上的连接性能降级。

于是，主机就会在心跳信号插槽上注册此失败，并在发出重置命令时中止 LUN 上的所有活动 IO。此 LUN 上所有挂起的 I/O 将失败，并显示主机感知数据 8 (H:0x8 SCSI 重置)。

来自使用 NMP 的 ESXi 主机的消息示例：

```
2015-10-01T00:31:00.333Z cpu9:33645)NMP: nmp_ThrottleLogForDevice:2321: Cmd 0x89 (0x412e82aeed40, 32805) to dev
"naa.60001440000000XXXXXXXXXXXXXXXXXX" on path "vmhba2:C0:T5:L10" Failed: H:0x8 D:0x0 P:0x0 Possible sense data:
0x5 0x20 0x0. Act:EVAL
2015-10-01T00:31:00.333Z cpu9:33645)WARNING: NMP: nmp_DeviceRequestFastDeviceProbe:237: NMP device
"naa.60001440000000XXXXXXXXXXXXXXXXXX" state in doubt; requested fast path state update...
2015-10-01T00:31:00.333Z cpu9:33645)ScsiDeviceIO: 2338: Cmd(0x412e82aeed40) 0x89, CmdSN 0x72b97 from world 32805 to
dev "naa.60001440000000XXXXXXXXXXXXXXXXXX" failed H:0x8 D:0x0 P:0x0 Possible sense data: 0x5 0x20 0x0.
2015-10-01T00:31:01.333Z cpu9:33645)WARNING: NMP: nmp_DeviceRequestFastDeviceProbe:237: NMP device
"naa.60001440000000XXXXXXXXXXXXXXXXXX" state in doubt; requested fast path state update...
2015-10-01T00:31:01.333Z cpu9:33645)ScsiDeviceIO: 2338: Cmd(0x413686ad0b80) 0x89, CmdSN 0x72b9a from world 32805 to
dev "naa.60001440000000XXXXXXXXXXXXXXXXXX" failed H:0x8 D:0x0 P:0x0 Possible sense data: 0x5 0x24 0x0.
2015-10-01T00:31:01.406Z cpu9:33645)ScsiDeviceIO: 2307: Cmd(0x41368008ee80) 0x2a, CmdSN 0x8000005d from world
1655292 to dev "naa.60001440000000XXXXXXXXXXXXXXXXXX" failed H:0x8 D:0x0 P:0x0
2015-10-01T00:31:01.406Z cpu9:33645)ScsiDeviceIO: 2307: Cmd(0x413686778800) 0x2a, CmdSN 0x8000004d from world
1655292 to dev "naa.60001440000000XXXXXXXXXXXXXXXXXX" failed H:0x8 D:0x0 P:0x0
2015-10-01T00:31:01.406Z cpu9:33645)ScsiDeviceIO: 2307: Cmd(0x4136838cc140) 0x2a, CmdSN 0x80000049 from world
1655292 to dev "naa.60001440000000XXXXXXXXXXXXXXXXXX" failed H:0x8 D:0x0 P:0x0
2015-10-01T00:31:01.608Z cpu9:33645)ScsiDeviceIO: 2307: Cmd(0x4136848c5c00) 0x2a, CmdSN 0x80000065 from world
1655292 to dev "naa.60001440000000XXXXXXXXXXXXXXXXXX" failed H:0x8 D:0x0 P:0x0
2015-10-01T00:31:01.609Z cpu9:33645)ScsiDeviceIO: 2307: Cmd(0x4136836fde80) 0x2a, CmdSN 0x8000002c from world
1655292 to dev "naa.60001440000000XXXXXXXXXXXXXXXXXX" failed H:0x8 D:0x0 P:0x0
2015-10-01T00:31:01.811Z cpu9:33645)NMP: nmp_ThrottleLogForDevice:2321: Cmd 0x2a (0x4136804206c0, 1655292) to dev
"naa.60001440000000XXXXXXXXXXXXXXXXXX" on path "vmhba2:C0:T5:L10" Failed: H:0x8 D:0x0 P:0x0 Possible sense data:
0x0 0x0 0x0. Act:EVAL
2015-10-01T00:31:02.014Z cpu9:33645)ScsiDeviceIO: 2307: Cmd(0x4136848cb740) 0x28, CmdSN 0x72b98 from world 34950 to
dev "naa.60001440000000XXXXXXXXXXXXXXXXXX" failed H:0x8 D:0x0 P:0x0
2015-10-01T00:31:02.014Z cpu18:34950)HBX: 2832: Waiting for timed out [HB state abcdef02 offset 4161536 gen 297 stampUS
933180151199 uuid 551234ba-5123418f-0123-7123457d566e jml <FB 15038> drv 14.60] on vol 'VPLEX-VOLUME-NAME'
2015-10-01T00:31:02.015Z cpu9:33645)ScsiDeviceIO: 2307: Cmd(0x41368386e100) 0x2a, CmdSN 0x72b99 from world 32805 to
dev "naa.60001440000000XXXXXXXXXXXXXXXXXX" failed H:0x8 D:0x0 P:0x0
2015-10-01T00:31:05.675Z cpu9:33039)VMW_SATP_INV: satp_inv_UpdatePath:754: Failed to update path "vmhba3:C0:T5:L10"
state. Status Transient storage condition, suggest retry
```

注意：主机向设备报告连接问题。这不是物理连接问题，它是由于单次 LUN 重置导致失去到存储设备的路径而造成的。

2. 低于版本 1.1.58.0-1 的 QLogic qlnativefc HBA 驱动程序在 ATS 心跳方面有重大问题：

该驱动程序在解释 SCSI ATS 失败感知代码方面存在问题，返回此代码的情形是：阵列不能正确地将消息传回主机的多路径层 (NMP 或 PowerPath)，因此 ESXi 内核不知道 iss 命令，而且主机在 vmkernel.log 中未报告任何消息。

更改： 主机升级到 ESXi 版本 5.5.0 Update 2 (内部版本号 2068190) 或更高版本。

主机升级到 ESXi 版本 6.0 或更高版本。

解决方法：1. 许多阵列供应商 (包括 EMC 在内) 在 ESXi 5.5u2 中引入的 ATS 心跳信号功能上都有问题。

从 EMC 阵列的角度看，目前没有变通办法。客户可以接洽 VMware 以确认此问题，或者按照 KB 15034 中所述使用支持虚拟机的 EMCgrub 来确认此问题。当前，建议受影响的客户在 ESXi 上禁用 VAAI ATS 心跳信号功能。

参见 [VMware KB 2113956](#) 了解进一步的信息。

禁用此功能将使主机恢复到以前的 SCSI-2 保留模式。

2. QLogic qlnativefc HBA 驱动程序在 ATS 心跳信号方面有重大问题，该问题已经在驱动程序版本 1.1.58.0-1 中解决。

使用低于版本 1.1.58.0-1 的 qlnativefc 驱动程序的 HBA 应升级到 qlnativefc 1.1.58.0-1 版。

更新的 qlnativefc 1.1.58.0-1 驱动程序可从以下位置下载：

<https://my.vmware.com/web/vmware/details?downloadGroup=DT-ESXI55-QLOGIC-QLNATIVEFC-11580-1&productId=353>

注意： 请注意：仅当 VPLEX 因 VAAI ATS 心跳使用的 SCSI 操作代码 0x89 命令 (比较和写入) 而向主机返回失败 (感知数据 0E/1D/00) 时，适用该解决方案。如果 0x89 (比较和写入) 命令由于其他原因 (例如超时、主机中止等) 而失败，则此解决方法不适用，建议联系 EMC 支持部门。

此文章已升级为 EMC 社区网络 (ECN) 上的 HVC：<https://community.emc.com/docs/DOC-54138>

主要产品： VPLEX 系列

产品： VPLEX 系列、VPLEX Geo、VPLEX Local、VPLEX Metro、VPLEX VS1、VPLEX VS2、VPLEX Virtual Edition、VPLEX GeoSynchrony 5.1、VPLEX GeoSynchrony 5.1 Patch 2、VPLEX GeoSynchrony 5.1 Patch 3、VPLEX GeoSynchrony 5.1 Patch 4、VPLEX GeoSynchrony 5.2、VPLEX GeoSynchrony 5.2 Service Pack 1、VPLEX GeoSynchrony 5.2 Service Pack 1 Patch 1、VPLEX GeoSynchrony 5.2 Service Pack 1 Patch 2、VPLEX GeoSynchrony 5.2 Service Pack 1 Patch 3、VPLEX GeoSynchrony 5.3、VPLEX GeoSynchrony 5.3 Patch 1、VPLEX GeoSynchrony 5.3 Patch 2、VPLEX GeoSynchrony 5.3 Patch 4、VPLEX GeoSynchrony 5.4、VPLEX GeoSynchrony 5.4 Service Pack 1、VPLEX GeoSynchrony 5.4 Service Pack 1 Patch