

某局点H3C FlexServer R390服务器阵列失败数据丢失的经验案例

RAID 固件升级 SW_Bundle AHS 周锋 2017-06-27 发表

某局点一台H3C FlexServer R390服务器，安装有7块硬盘，其中6块硬盘做RAID 10，1块硬盘配置成热备盘。阵列失败，数据丢失，无法正常进入系统。

开机自检时能看到如下的告警信息：

```
1792-Slot 0 Drive Array - Valid Data Found in Write-Back Cache.
Data will automatically be written to drive array.
1779-Slot 0 Drive Array - Replacement drive(s) detected OR previously failed
drive(s) now appear to be operational:
Port 1l: Box 2: Bay 2
Port 2l: Box 2: Bay 5
Logical drive(s) disabled due to possible data loss.
Select "F1" to continue with logical drive(s) disabled
Select "F2" to accept data loss and to re-enable logical drive(s)
(RESUME = "F1" OR "F2" KEY) [default = "F1" in 45 seconds] **TIMED OUT**
```

```
1716-Slot 0 Drive Array - Unrecoverable Media Errors Detected on Drives
during previous Rebuild or Background Surface Analysis (ARM) scan.
Errors will be fixed automatically when the sector(s) are overwritten.
Backup and Restore recommended.
```

分析日志发现问题如下：

1. IML记录有大量的介质错误，如下：

```
Critical,1192,29197,0x0013,Drive Array,,,05/30/2017 09:10:00,4: Internal Storage Enclosure Device Failure (Bay 5, Box 2, Port 2l, Slot 0)
Critical,1192,29231,0x0013,Drive Array,,,05/30/2017 09:10:00,5: Internal Storage Enclosure Device Failure (Bay 2, Box 2, Port 1l, Slot 0)
Repaired,1192,29234,0x0013,Drive Array,,,05/30/2017 09:10:00,4: Internal Storage Enclosure Device Failure (Bay 5, Box 2, Port 2l, Slot 0)
Repaired,1192,29274,0x0013,Drive Array,,,05/30/2017 09:10:00,5: Internal Storage Enclosure Device Failure (Bay 2, Box 2, Port 1l, Slot 0)
Caution,1193,933,0x000A,POST Message,,,05/30/2017 11:03:00,6: POST Error: 1792-Slot X Drive Array - Valid Data Found in Cache Module. Data will automatically be written to drive array.
Caution,1193,934,0x000A,POST Message,,,05/30/2017 11:03:00,7: POST Error: 1779-Slot X Drive Array - Replacement drive(s) detected OR previously failed drive(s) now appear to be operational.
Caution,1193,935,0x000A,POST Message,,,05/30/2017 11:03:00,8: POST Error: 1716-Slot X Drive Array - Unrecoverable Media Errors Detected on Drives during previous Rebuild or Background Surface Analysis (ARM) scan. Errors will be fixed automatically when the sector(s) are overwritten.
```

2. 分析ADU日志能发现当前的阵列配置信息情况是使用P420i阵列卡将bay1-bay6硬盘配置RAID 10，组建Array A，logical drive 1；bay1和bay4；bay2和bay5；bay3和bay6组成RAID 1组互为镜像，然后3个RAID 1组再组成一个RAID 0阵列。bay7硬盘是做热备的，上面报错的bay2和bay5硬盘刚好在同一个RAID 1组内，具体如下：

```
Big Drive Assignment Map 0x3f 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00
```

Position Device Status

```
-----
0 Physical Drive (500 GB SAS) 1l:2:1 Informational
1 Physical Drive (500 GB SAS) 1l:2:2 Informational
2 Physical Drive (500 GB SAS) 1l:2:3 Informational
3 Physical Drive (500 GB SAS) 1l:2:4 Informational
4 Physical Drive (500 GB SAS) 2l:2:5 Informational
5 Physical Drive (500 GB SAS) 2l:2:6 Informational
```

Fault Tolerance Mode 10 (0x0002)

Smart Array P420i in Embedded Slot : SAS Array A : Logical Drive 1 : Mirror/Parity Group Information

Paired Drive 0x0003 0x0004 0x0005 0x0000 0x0001 0x0002 0x0006 0x0007 0x0008 0x0009 0x000a
0x000b 0x000c 0x000d 0x000e 0x000f 0x0010
0x0011 0x0012 0x0013 0x0014 0x0015 0x0016 0x0017 0x0018 0x0019 0x001a 0x001b 0x001c 0x00
1d 0x001e 0x001f 0x0020 0x0021
0x0022 0x0023 0x0024 0x0025 0x0026 0x0027 0x0028 0x0029 0x002a 0x002b 0x002c 0x002d 0x00
2e 0x002f 0x0030 0x0031 0x0032
0x0033 0x0034 0x0035 0x0036 0x0037 0x0038 0x0039 0x003a 0x003b 0x003c 0x003d 0x003e 0x00
3f 0x0040 0x0041 0x0042 0x0043
0x0044 0x0045 0x0046 0x0047 0x0048 0x0049 0x004a 0x004b 0x004c 0x004d 0x004e 0x004f 0x005
0 0x0051 0x0052 0x0053 0x0054
0x0055 0x0056 0x0057 0x0058 0x0059 0x005a 0x005b 0x005c 0x005d 0x005e 0x005f 0x0060 0x006
1 0x0062 0x0063 0x0064 0x0065
0x0066 0x0067 0x0068 0x0069 0x006a 0x006b 0x006c 0x006d 0x006e 0x006f 0x0070 0x0071 0x007
2 0x0073 0x0074 0x0075 0x0076
0x0077 0x0078 0x0079 0x007a 0x007b 0x007c 0x007d 0x007e 0x007f 0x0080 0x0081 0x0082 0x008
3 0x0084 0x0085 0x0086 0x0087
0x0088 0x0089 0x008a 0x008b 0x008c 0x008d 0x008e 0x008f 0x0090 0x0091 0x0092 0x0093 0x009
4 0x0095 0x0096 0x0097 0x0098
0x0099 0x009a 0x009b 0x009c 0x009d 0x009e 0x009f 0x00a0 0x00a1 0x00a2 0x00a3 0x00a4 0x00a
5 0x00a6 0x00a7 0x00a8 0x00a9
0x00aa 0x00ab 0x00ac 0x00ad 0x00ae 0x00af 0x00b0 0x00b1 0x00b2 0x00b3 0x00b4 0x00b5 0x00b
6 0x00b7 0x00b8 0x00b9 0x00ba
0x00bb 0x00bc 0x00bd 0x00be 0x00bf 0x00c0 0x00c1 0x00c2 0x00c3 0x00c4 0x00c5 0x00c6 0x00c7
0x00c8 0x00c9 0x00ca 0x00cb
0x00cc 0x00cd 0x00ce 0x00cf 0x00d0 0x00d1 0x00d2 0x00d3 0x00d4 0x00d5 0x00d6 0x00d7 0x00d
8 0x00d9 0x00da 0x00db 0x00dc
0x00dd 0x00de 0x00df 0x00e0 0x00e1 0x00e2 0x00e3 0x00e4 0x00e5 0x00e6 0x00e7 0x00e8
0x00e9 0x00ea 0x00eb 0x00ec 0x00ed
0x00ee 0x00ef 0x00f0 0x00f1 0x00f2 0x00f3 0x00f4 0x00f5 0x00f6 0x00f7 0x00f8 0x00f9 0x00fa 0x0
0fb 0x00fc 0x00fd 0x00fe
0x00ff

Position Device Association Status

0 Physical Drive (500 GB SAS) 11:2:1 Physical Drive (500 GB SAS) 11:2:4 Informational
1 Physical Drive (500 GB SAS) 11:2:2 Physical Drive (500 GB SAS) 21:2:5 Informational
2 Physical Drive (500 GB SAS) 11:2:3 Physical Drive (500 GB SAS) 21:2:6 Informational
3 Physical Drive (500 GB SAS) 11:2:4 Physical Drive (500 GB SAS) 11:2:1 Informational
4 Physical Drive (500 GB SAS) 21:2:5 Physical Drive (500 GB SAS) 11:2:2 Informational
5 Physical Drive (500 GB SAS) 21:2:6 Physical Drive (500 GB SAS) 11:2:3 Informational
6 Physical Drive (500 GB SAS) 21:2:7 Physical Drive (500 GB SAS) 21:2:7 Informational

3. 阵列失败的情况是bay5硬盘发现被拔掉，导致logical drive降级，不长时间bay2硬盘又有被拔掉的记录，由于bay2和bay5在同一个RAID 1组内，同时和其他硬盘组成RAID 10，所以导致阵列失败，逻辑驱动器失败，bay7这个热备盘也在随后被发现有拔除记录，具体如下：

Critical,1192,29211,Smart Array,Physical drive removed, ,0x00,05/30/2017 09:10:03,[05/30 10:45:21]
Hot-plug drive removed, Port=21 Box=2 Bay=5 SN=9XF2L38300009411DFVH
Critical,1192,29212,Smart Array,Physical drive failure, ,0x00,05/30/2017 09:10:03,[05/30 10:45:21] P
hysical drive failure, Port=21 Box=2 Bay=5 reason=0x14
Caution,1192,29213,Smart Array,Logical drive status changed, ,0x00,05/30/2017 09:10:03,[05/30 10:
45:21] State change, logical drive 0, new state=DEGRADED
Caution,1192,29214,Smart Array,Logical drive status changed, ,0x00,05/30/2017 09:10:03,[05/30 10:
45:26] State change, logical drive 0, new state=NEEDS_REBUILD
Caution,1192,29215,Smart Array,Logical drive status changed, ,0x00,05/30/2017 09:10:03,[05/30 10:
45:26] State change, logical drive 0, new state=REBUILDING
Caution,1192,29216,Smart Array,Physical drive inserted, ,0x00,05/30/2017 09:10:03,[05/30
10:45:43] Hot-plug drive inserted, Port=21 Box=2 Bay=5 SN=9XF2L38300009411DFVH
Caution,1192,29217,Smart Array,Logical drive status changed, ,0x00,05/30/2017 09:10:03,[05/30 10:
45:43] State change, logical drive 0, new state=NEEDS_REBUILD
Critical,1192,29218,Smart Array,Physical drive removed, ,0x00,05/30/2017 09:10:03,[05/30 10:45:43]
Hot-plug drive removed, Port=11 Box=2 Bay=2 SN=9XF2L2JE000094141M37

Critical,1192,29219,Smart Array,Physical drive failure, ,0x00,05/30/2017 09:10:03,[05/30 10:45:43] Physical drive failure, Port=11 Box=2 Bay=2 reason=0x14
Caution,1192,29220,Smart Array,Logical drive exchanged media, ,0x00,05/30/2017 09:10:03,[05/30 10:45:43] Media exchanged detected, logical drive 0
Caution,1192,29221,Smart Array,Logical drive status changed, ,0x00,05/30/2017 09:10:03,[05/30 10:45:43] State change, logical drive 0, new state=FAILED
Caution,1192,29222,Smart Array,Rebuild complete despite uncorrectable media errors, ,0x00,05/30/2017 09:10:03,[05/30 10:45:45] Rebuild URE, LDrv=0 LBA=0x0005E3800-0x0005E4FFF
Caution,1192,29239,Smart Array,Physical drive inserted, ,0x00,05/30/2017 09:10:08,[05/30 10:45:57] Hot-plug drive inserted, Port=11 Box=2 Bay=2 SN=9XF2L2JE000094141M37
Critical,1192,29314,Smart Array,Physical drive removed, ,0x00,05/30/2017 09:11:18,[05/30 10:46:36] Hot-plug drive removed, Port=21 Box=2 Bay=7 SN=9XF2L2BM00009413GJFD
Critical,1192,29315,Smart Array,Physical drive failure, ,0x00,05/30/2017 09:11:18,[05/30 10:46:36] Physical drive failure, Port=21 Box=2 Bay=7 reason=0x14
Caution,1192,29316,Smart Array,Physical drive inserted, ,0x00,05/30/2017 09:11:18,[05/30 10:46:57] Hot-plug drive inserted, Port=21 Box=2 Bay=7 SN=9XF2L2BM00009413GJFD

4. 分析每块硬盘的M&P记录,发现2块硬盘 (bay2, bay7) 有读写/恢复错误,同时有指向硬盘背板的bus faults记录, 1块硬盘 (bay5) 本身没有任何错误,只有bus faults记录,如下:

Smart Array P420i in Embedded Slot : Internal Drive Cage at Port 11 : Box 2 :Physical Drive (500 GB SAS) 11:2:2 : Monitor and Performance Statistics (Since Factory)

Serial Number 9XF2L2JE000094141M37
Firmware Revision HPD8
Product Revision HP MM0500FBFVQ
Reference Time 0x00156e40
Sectors Read 0x0000002195fb69f4
Read Errors Hard 0x00000000
Read Errors Retry Recovered 0x00000000
Read Errors ECC Corrected 0x0000000000000000
Sectors Written 0x0000000078debd2b
Write Errors Hard 0x00000000
Write Errors Retry Recovered 0x00000000
Seek Count 0xffffffff
Seek Errors 0xffffffff
Spin Cycles 0x00000000
Spin Up Time 0x0000
Performance Test 1 0x0000
Performance Test 2 0xffff
Performance Test 3 0xffff
Performance Test 4 0xffff
Reallocation Sectors 0xffffffff
Reallocated Sectors 0xffffffff
DRQ Time Outs 0xffff
Other Time Outs 0x0000
Drive Rebuild Count 0 (0x0000)
Spin Retries 65535 (0xffff)
Recovers Failed Read 0x0002
Recovers Failed Write 0x0000
Format Errors 0x0000
Self Test Failures 0xffff
Not Ready Failures 0x00000000
Remap Abort Failures 0xffffffff
IRQ Deglitch Count 4294967295 (0xffffffff)
Bus Faults 0x00000016
Hot Plug Count 1 (0x00000001)

Track Rewrite Errors 0xffff
Write Errors After Remap 0x0000
Background Firmware Revision 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00
Media Failures 0x0000
Hardware Errors 0x0000
Aborted Command Failures 0x0000
Spin Up Failures 0x0000
Bad Target Count 0 (0x0000)
Predictive Failure Errors 0x00000000

Smart Array P420i in Embedded Slot : Internal Drive Cage at Port 2I : Box 2 :Physical Drive (500 GB
SAS) 2I:2:5 : Monitor and Performance Statistics (Since Factory)

Serial Number 9XF2L38300009411DFVH
Firmware Revision HPD8
Product Revision HP MM0500FBFVQ
Reference Time 0x00156e40
Sectors Read 0x0000002193dd9f06
Read Errors Hard 0x00000000
Read Errors Retry Recovered 0x00000000
Read Errors ECC Corrected 0x0000000000000000
Sectors Written 0x0000000078deb745
Write Errors Hard 0x00000000
Write Errors Retry Recovered 0x00000000
Seek Count 0xffffffffffff
Seek Errors 0xffffffffffff
Spin Cycles 0x00000000
Spin Up Time 0x0000
Performance Test 1 0x0000
Performance Test 2 0xffff
Performance Test 3 0xffff
Performance Test 4 0xffff
Reallocation Sectors 0xffffffff
Reallocated Sectors 0xffffffff
DRQ Time Outs 0xffff
Other Time Outs 0x0000
Drive Rebuild Count 0 (0x0000)
Spin Retries 65535 (0xffff)
Recovers Failed Read 0x0000
Recovers Failed Write 0x0000
Format Errors 0x0000
Self Test Failures 0xffff
Not Ready Failures 0x00000000
Remap Abort Failures 0xffffffff
IRQ Deglitch Count 4294967295 (0xffffffff)
Bus Faults 0x00000016
Hot Plug Count 1 (0x00000001)
Track Rewrite Errors 0xffff
Write Errors After Remap 0x0000
Background Firmware Revision 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00
Media Failures 0x0000
Hardware Errors 0x0000
Aborted Command Failures 0x0000
Spin Up Failures 0x0000
Bad Target Count 0 (0x0000)
Predictive Failure Errors 0x00000000

Smart Array P420i in Embedded Slot : Internal Drive Cage at Port 2I : Box 2 :Physical Drive (500 GB
SAS) 2I:2:7 : Monitor and Performance Statistics (Since Factory)

Serial Number 9XF2L2BM00009413GJFD
Firmware Revision HPD8

Product Revision HP MM0500FBFVQ
Reference Time 0x00156e40
Sectors Read 0x00000000004056f
Read Errors Hard 0x00000001
Read Errors Retry Recovered 0x00000000
Read Errors ECC Corrected 0x0000000000000000
Sectors Written 0x000000000234999
Write Errors Hard 0x00000000
Write Errors Retry Recovered 0x00000000
Seek Count 0xffffffffffff
Seek Errors 0xffffffffffff
Spin Cycles 0x00000000
Spin Up Time 0x0000
Performance Test 1 0x0000
Performance Test 2 0xffff
Performance Test 3 0xffff
Performance Test 4 0xffff
Reallocation Sectors 0xffffffff
Reallocated Sectors 0xffffffff
DRQ Time Outs 0xffff
Other Time Outs 0x0000
Drive Rebuild Count 0 (0x0000)
Spin Retries 65535 (0xffff)
Recovers Failed Read 0x0000
Recovers Failed Write 0x0000
Format Errors 0x0000
Self Test Failures 0xffff
Not Ready Failures 0x00000000
Remap Abort Failures 0xffffffff
IRQ Deglitch Count 4294967295 (0xffffffff)
Bus Faults 0x00000016
Hot Plug Count 1 (0x00000001)
Track Rewrite Errors 0xffff
Write Errors After Remap 0x0000
Background Firmware Revision 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00
Media Failures 0x0000
Hardware Errors 0x0000
Aborted Command Failures 0x0000
Spin Up Failures 0x0000
Bad Target Count 0 (0x0000)
Predictive Failure Errors 0x00000000

5. 另外，发现阵列卡固件，BIOS和iLO 4固件均偏低，如下：

iLO (iLO Advanced License) iLO 4 v2.00p67 built on Jul 30 2014
System ROM 02/10/2014
Slot Controller Serial# Version Version Version Revision Revision

0 P420i 001438030013160 6.00 1.90 01.90.002.002 1 40

综上日志分析，若排除人为拔盘的操作，可以定位主要是硬盘背板的原因导致的阵列失败，同时可以确认2块硬盘（bay2， bay7）有问题，与bay2同一RAID 1组的bay5硬盘没有硬件错误， bay7是热备盘，所以如果更换硬盘背板解决连接稳定性后阵列数据是没有丢失的。

1. 更换硬盘背板，然后先拔掉bay2和bay7问题硬盘（拔掉这两个硬盘对阵列数据完整性没有影响）；
2. 重启机器，然后重新激活阵列后能进入系统，做好数据备份；
3. 同时更换掉bay2， bay7问题硬盘，然后使用最新的SW Bundle更新机器固件。

1. 从日志中找到阵列失败的时间点和具体硬盘如何组成的阵列对分析问题十分有帮助；
2. 针对阵列、存储、硬盘类问题需要收集全AHS和ADU日志；
3. 硬盘M&P的记录对分析硬盘是否有硬件问题以及硬盘背板是否正常非常有用。