Huawei Storage Troubleshooting Skills

2020/9/9

Security Level:

HUAWEI

Foreword

- To understand basic troubleshooting skills is very important activity in routine maintenance.
- This course introduces how to troubleshooting storage common issues.

HUAWEI

Contents

Troubleshooting Fundamental Rules

- Analyze from “Top” to “Bottom”
  - ✓ Completely collect all the issue symptoms. Which of the hosts? When? How? What change did before?
  - ✓ Check storage alarm and indicator
  - ✓ For emergency issue, involve Huawei support or local office right after
- Analyze High Severity Alarm First
  - ✓ Analyze the Critical alarm first, then Major and Minor alarm.
- Analyze Common and Latest Alarm First
  - ✓ Analyze the latest alarm with higher priority. If there’re multiple alarms, check the scope of the issue, for example, all the alarms related to HyperMetro feature.

HUAWEI

Analyze Alarm

Refer to product documentation to check the alarm and error code.

For example:

12344630 2019-07-22 17:34:19 DST **0xF00CF005F** Fault Major Unrecovered None Controller (Controller Enclosure CTE0, controller A, item 03057201, SN 210305720110XXXXXXXX) is faulty. Error code: **0x4000cf4d**.Collect all related information and contact technical support engineers to replace the controller.

https://support.huawei.com/enterprise/en/doc/EDOC1000138404?idPath=7919749%7C251366268%7C250389224%7C251366266%7C21538251

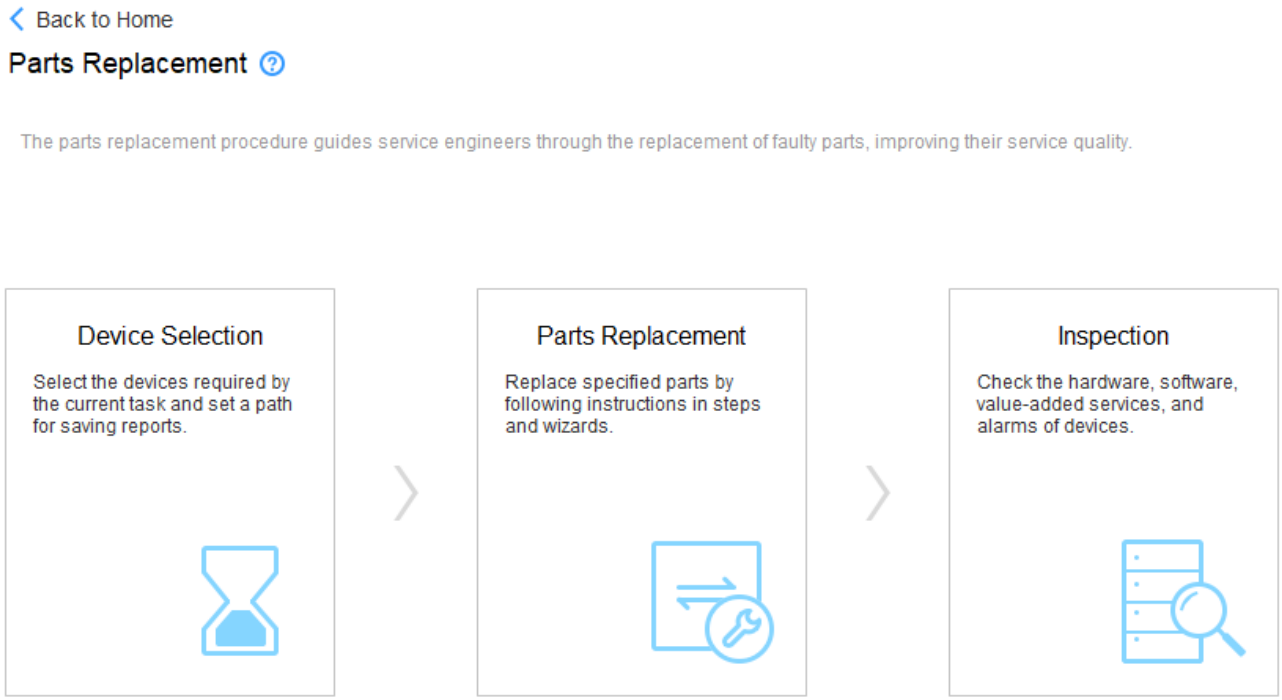
0x4000CF4D

Error Code	Description	Possible Causes	Handling Suggestion	Applicable Model
0x4000CF4D	A system disk on a control module is about to fail.	A system disk on a control module is about to fail.	Replace the controller.	Applicable to all models.

HUAWEI

Component Replacement Method

- Tool-Based CRU and FRU Replacement
  - Follow the wizard to check system status before replacement, and inspect after replacement. Especially, risky disk, controller, enclosure replacement



HUAWEI

Contents

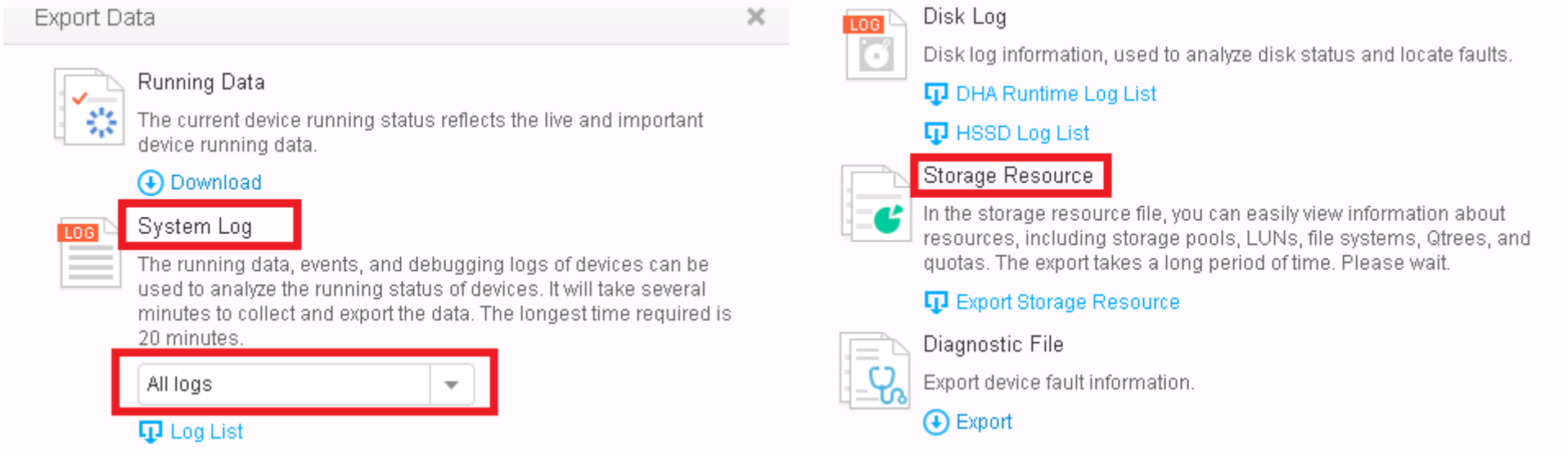
Collect Log by Device Manager

- Collect “System Log” -> “All logs” for most of the failure diagnose
  - ✓ Running Data: Device alarm and history event
  - ✓ System Log(Recent logs): Recent storage log without diagnose information.
  - ✓ **System Log(All logs): Complete storage log with diagnose information, including Running Data, Disk Log, etc.**
  - ✓ Disk Log: Disk health assessment information and faulty disk SMART
  - ✓ Storage Resource: Storage pool, LUN, file system, Quota tree, and quota information sort to excel
  - ✓ Diagnostic File: Internal module diagnose information, including hardware and software

HUAWEI

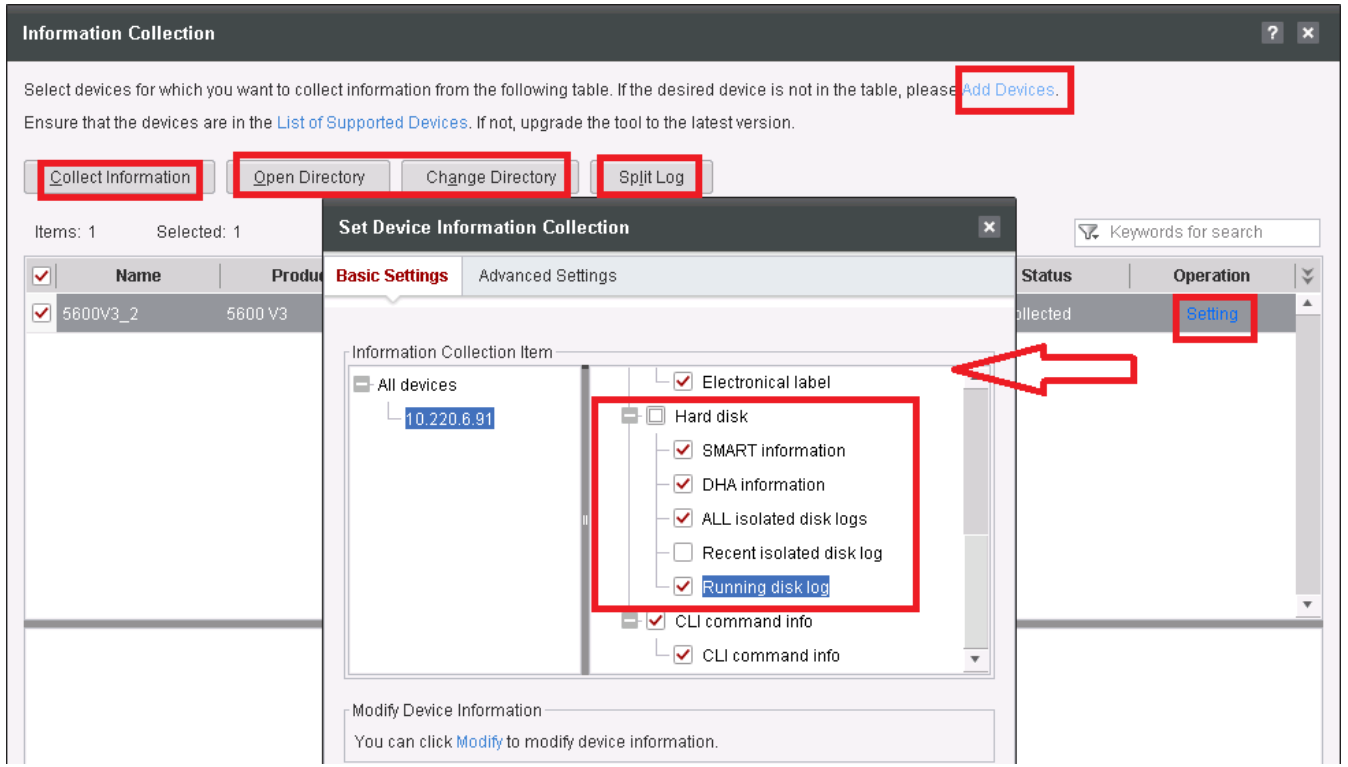
Collect Log by Device Manager(Continue)

- Collect “Storage Resource” to get NAS related configuration



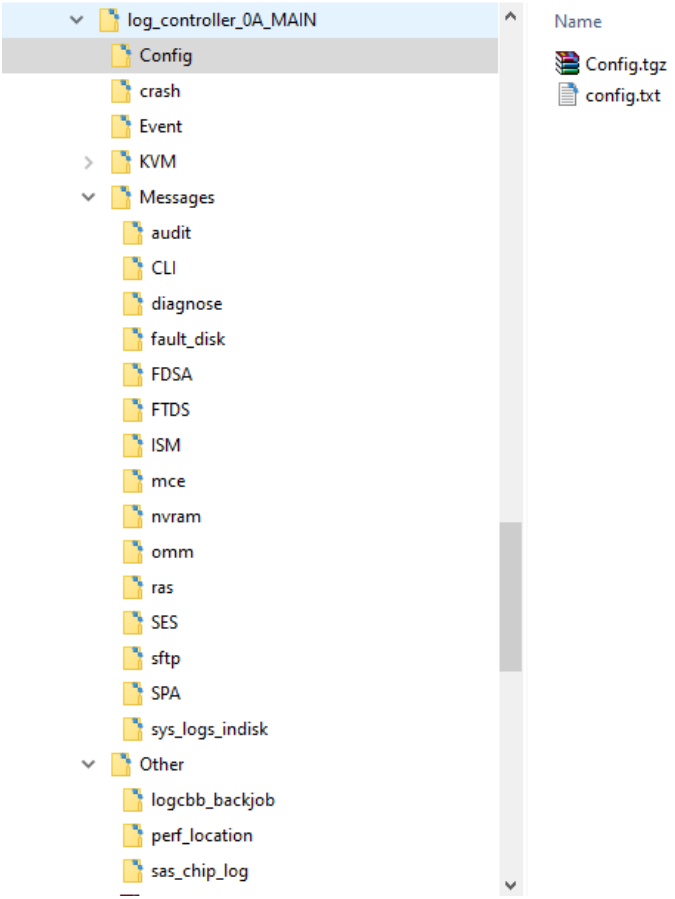
Collect Log by SmartKit

- ✓ For deep diagnose for issues like high disk failure rate, need to collect Running disk log and all disk SMART by SmartKit.



Storage Log Structure

- ✓ Config -> config.txt: collect on master controller, including all hardware and software configuration(status)
- ✓ Event -> Event.txt: collect on master controller, including current alarm and history events
- ✓ Messages -> sys_logs_indisk: history OS log, like Linux /var/log/messages, including NVRAM dump log for controller abnormal reset



HUAWEI

Basic Log Analysis

✓Check hardware status and software configuration by config.txt

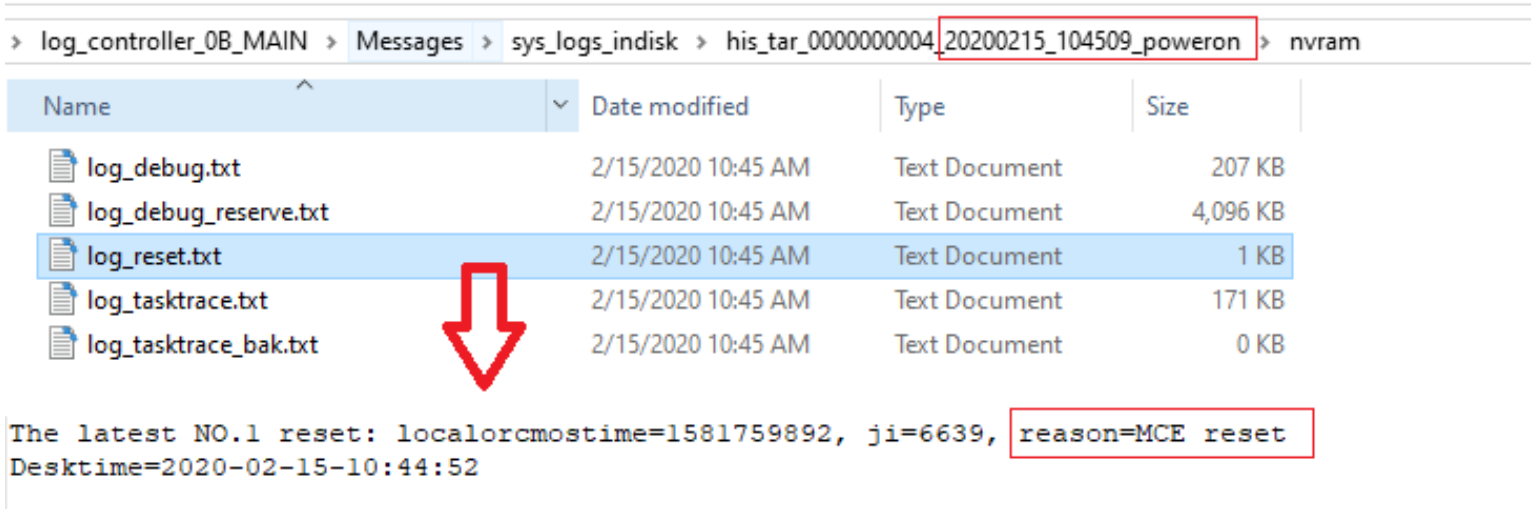
```
SFP Info:
    Vendor: Hisense
    Model: LTF8502-BC+
    SN: N3383022418
    Health Status: Normal
    Running Status: Link Up
    Type: Multi Mode
    Working Rate(Mbps): 10000
    Temperature(C): 31
    RxPowerReal(uW): 576.1
    RXPowerMax(uW): 1258.9
    RXPowerMin(uW): 64.6
    TxPowerReal(uW): 587.7
    TXPowerMax(uW): 794.3
    TXPowerMin(uW): 186.2
    Item: --
    ExternalModel: --
    Rev: --
```

```
Initial Capacity: 1048576(KB)
Compression Enabled: Yes
Compression Method: Fast
Dedup Enabled: Yes
Byte_by_byte Comparison Enabled: No
Intelligent Dedup Enabled: Yes
Dedup Metadata Sample Ratio: 2
Dedup Running Status: Yes
Dedup Saved Capacity: 53130968(KB)
Dedup Saved Ratio(%): 11
Compression Saved Capacity: 129415074(KB)
Compression Saved Ratio(%): 28
Total Saved Capacity: 182546042(KB)
Total Saved Ratio(%): 39
Original Total Capacity: 450223161(KB)
Thresholds Percent(%): 90
Thresholds Switch: Off
Vstore ID: --
```

HUAWEI

Basic Log Analysis(Continue)

✓Check abnormal reset



HUAWEI

Contents

Query and Set Network Configuration

- ✓ Query network configuration by “show port *XXX* …”, associate by TAB
- ✓ Change management and maintenance port configuration by “change system ma*XXX* …”
- ✓ Check port type(bond, logic), MTU, etc.

```
developer:/>show port general logic_type=Maintenance_Port
ETH port:
------------  Maintenance Port:------------

  ID                     Health Status  Running Status  Type             IPv4 Address    IPv6 Address  MAC               Role  Working
Rate(Mbps)  Enabled  Max Speed(Mbps)
  ---------------------  -------------  --------------  ---------------  --------------  ------------  ----------------  ----  --------
----------  -------  ---------------
  CTE0.SMM0.MAINTENANCE  Normal         Link Down       Maintenance Port  172.31.128.101  --            04:b0:e7:a8:3c:11  --    --
            Yes      1000
  CTE0.SMM1.MAINTENANCE  Normal         Link Down       Maintenance Port  172.31.128.102  --            04:b0:e7:a8:3c:13  --    --
            Yes      1000
developer:/>change system maintenance_ip eth_port_id=CTE0.SMM1.MAINTENANCE ip=10.220.6.12 mask=255.255.255.0
WARNING: You are about to change the IP address of the maintenance port. This operation will interrupt the connection between the mainten
ance engineer and the system.
Suggestion: Before you perform this operation, ensure that you have entered an available IP address.
Have you read warning message carefully?(y/n)y

Are you sure you really want to perform the operation?(y/n)y
Error: The specified IP address and the management network port's IP address reside on the same network segment.
Suggestion: Enter another IP address.
developer:/>
```

HUAWEI

Network Diagnose Command

“ip a” and “ifconfig”: check network configuration

“ip rule” and “route”: check route configuration

telnet: test specific port service status

ping: test network connectivity

“sshtoremoteExt *X*” or “sshtoremote: login other controller by heartbeat port

```
Storage: minisystem> ping -f -s 9000 -I eth-b1 10.220.6.71
PING 10.220.6.71 (10.220.6.71) from 10.220.6.91 eth-b1: 9000(9028) bytes of data.
..^
--- 10.220.6.71 ping statistics ---
26723 packets transmitted, 26722 received, 0% packet loss, time 15008ms
rtt min/avg/max/mdev = 0.483/0.508/2.105/0.027 ms, pipe 2, ipg/ewma 0.561/0.517 ms
Storage: minisystem> ip rule
0:      from all lookup local
1:      from 192.168.27.225 lookup 26384
100:    from 192.168.27.225 blackhole
101:    from 10.220.6.91 lookup 998
101:    from 172.31.128.101 lookup 996
32766:  from all lookup main
32767:  from all lookup default
Storage: minisystem>
```

HUAWEI

Network Capture

- tcpdump.sh
  - Only capture packet header, no content
  - Limited packet number to write file(avoid out of memory)
  - Print out the network capture for long term capture

```
Storage: minisystem> tcpdump.sh src 10.220.6.115 -i eth-b1
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth-b1, link-type EN10MB (Ethernet), capture size 64 bytes
19:23:01.454480 IP 10.220.6.115.51896 > 10.220.6.91.22: . ack 3921768283 win 255
19:23:01.509356 IP 10.220.6.115.51896 > 10.220.6.91.22: . ack 145 win 255
19:23:01.556212 IP 10.220.6.115.51896 > 10.220.6.91.22: . ack 273 win 254
19:23:01.591425 IP 10.220.6.115.52135 > 10.220.6.91.8088: P 3940353117:3940355746(2629) ack 4040524245 wi
19:23:01.591663 IP 10.220.6.115.51896 > 10.220.6.91.22: . ack 577 win 253
Storage: minisystem> tcpdump.sh src 10.220.6.115 -i eth-b1 -w test.pcap -c 16000
Command will run in the background, and the result will be saved to /OSM/coffer_log/log/tcpdump/test.pcap
tcpdump: listening on eth-b1, link-type EN10MB (Ethernet), capture size 64 bytes
Storage: minisystem>
```

HUAWEI

New Command for Dorado V6

In Dorado V6, all the network diagnose command encapsulated by “net.sh”

```
Storage: minisystem> net.sh ip rule
0:      from all oif vrf-inner lookup 100
0:      from all iif vrf-inner lookup 100
10:     from all lookup local
101:    from 172.31.128.102 lookup 996
101:    from 51.10.42.132 lookup 998
200:    from all oif vrf-srv lookup 101
200:    from all iif vrf-srv lookup 101
1000:   from all lookup [l3mdev-table]
32766:  from all lookup main
32767:  from all lookup default
```

```
Storage: minisystem> net.sh ifconfig eth4
eth4: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 1500
        inet 51.10.42.132  netmask 255.255.224.0  broadcast 51.10.63.255
        inet6 2992::24  prefixlen 64  scopeid 0x0<global>
        inet6 fe80::aaf5:acff:fe25:926c  prefixlen 64  scopeid 0x20<link>
        ether a8:f5:ac:25:92:6c  txqueuelen 1000  (Ethernet)
        RX packets 38162671  bytes 3013521172 (2.8 GiB)
        RX errors 0  dropped 137475  overruns 0  frame 0
        TX packets 37640  bytes 2672972 (2.5 MiB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
```

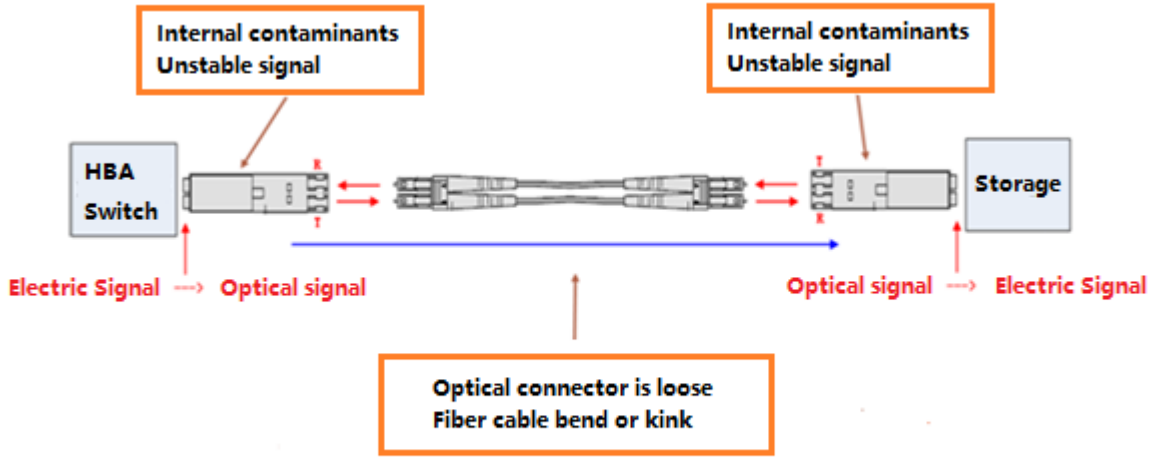
```
Storage: minisystem> net.sh ip route show table 998
default via 51.10.32.1 dev eth4 onlink
51.10.32.0/19 dev eth4 proto kernel scope link src 51.10.42.132
172.31.0.0/16 dev eth4.4081 proto kernel scope link src 172.31.128.102
```

HUAWEI

Contents

Sources of Bit Error

- Fibre cable quality
- SFP quality
- HBA/Connect board quality
- Compatility: for example, unsupported SFP, switch, etc.



Higher speed, higher signal quality

HUAWEI

FC Bit Error Diagnose

- Bit Error alarm trigger condition: bit error count increase in 3 continuous check cycles
- Basic diagnose process:
  - Check SFP TX and RX power, whether under 50% of normal SFP
    1. If TX power low, replace local SFP
    2. If RX power low, check remote SFP and fiber cable
  - Check the regularity of bit error
    - ✓ One port or multiple ports have bit error, maybe switch pass through issue frame without verification
    - ✓ Whether report alarm at specific period, maybe related to workload
  - Deep analyze storage log and SAN switch log

HUAWEI

Contents

Common Startup Failures

- Enter Rescue Mode
  - 3 abnormal reset in 30 minutes, like OOM reset, MCE reset, Oops reset, Watchdog reset, Panic reset, etc.
- Self-Check Failed
  - Hardware failure, like PSU, BBU, coffer disk failure, etc.
  - Incorrect SAS cabling
  - Plug out controllers at the same time, cause “dirty” cache loss
  - Inconsistent cluster configuration, for example, scale-out cluster only power on one controller enclosure

HUAWEI

Common Diagnose Process

- Check boot mode: showbootmode
  - resumebootmode
  - Collect log for root cause analysis
- Check fail reason of self-check
  - showsystrace/sys.sh showtrace
  - showsystrace *X*/sys.sh showtrace *X*
  - sys.sh showflowtrace

HUAWEI

Common Diagnose Process(Continue)

Storage: minisystem> **sys.sh showtrace**

admin:/diagnose>sys showtrace

Date Time	FlowId	Setup	RunCnt	FailCnt	Status
2020-09-01 16:42:26	**2**	**CLS_POWER_ON**	**1**	**1**	**Failure**
2020-09-01 16:42:34	5	CLS_LINK_CHECK	3	0	Success
2020-09-01 16:41:03	8	NODE_POWER_ON	1	0	Success

Storage: minisystem> **sys.sh showtrace 2**

admin:/diagnose>sys showtrace 2

CLS_POWER_ON :

TotalRunCnt	TotalFailCnt	CurStatus
1	1	Failure

Description:

Power on node bitmap(3).

id	date time	second	current trace
0000	**2020-09-01 16:42:26**	**206.601879**	**PowerOn: node clear failed info**

---> FAIL ACTION: RecoveryProfile: NtfClsUtil(CCDB DLM C-CLS)

Node: (1); Result: (4).

---> NODE POWER ON FAILED ERROR CODE INFO:

Node: (0); Controller Enslocure CTE0, Controller A; ErrCode: (0x4000C938).

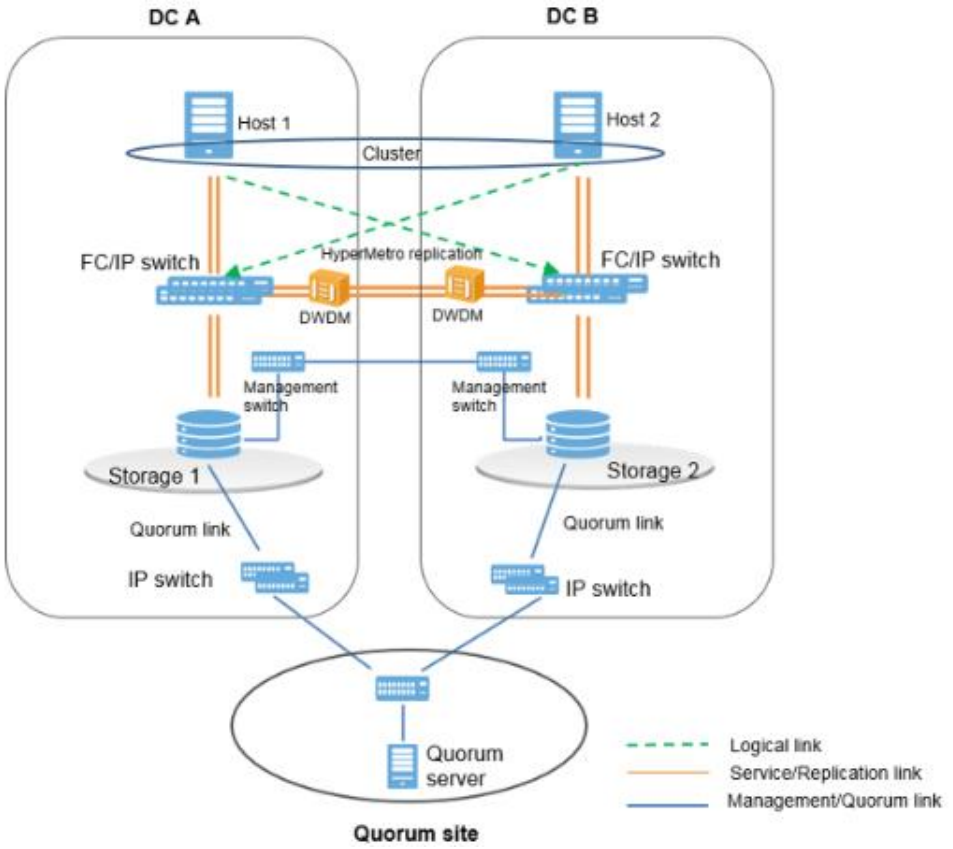
Node: (1); Controller Enslocure CTE0, Controller B; ErrCode: (0x4000C938).

HUAWEI

Contents

Understand HyperMetro Solution

- Isolated quorum network and replication network
  - Without quorum server or shared quorum/replication network will reduce service reliability
  - Dual quorum servers are working in active-standby mode, still need the standard networking
- HyperMetro load balance and local preferred mode
  - Decided by switchover mode on storage initiators for third-party multipathing
  - Decided by multipathing configuration for Huawei Ultrapath



HUAWEI

HyperMetro Pair Interruption

- Possible Reason
  - Replication network interruption
  - HyperMetro mirror write failed/timeout. For example, performance bottleneck, capacity used out.
- Collect log for HyperMetro pair interruption
  - All log for both of the storages to check the direct reason of interruption
  - Performance log for both of storages to check the performance statistics, possible bottleneck
  - Sometimes, also need host log, like VMware(vmsupport), Linux (/var/log/messages) to check error at host side

HUAWEI

HyperMetro Brain-Split

- Brain-split: the HyperMetro “cluster” can’t decide which node should continue service
  - Preferred site down in Static Priority Mode
  - Replication links and quorum links disconnected within 60 seconds in Quorum Mode. For example, after replication link, preferred site request arbitration immediately, non-preferred site request arbitration delayed. If the quorum links down, before arbitration result answered, both sites stop service because of brain-split.
  - Data still consistency in brain-split status
- Force synchronize the HyperMetro pair after confirm brain-split
  - Both sites are preferred site or non-preferred site is not brain-split
  - **Need to confirm with Huawei support before restore the pair**

HUAWEI

Thank you.

Bring digital to every person, home and organization for a fully connected, intelligent world.