

LOTC Disk Replication Consistency

OPERATING INSTRUCTIONS

Copyright

© Ericsson AB 2016, 2017. All rights reserved. No part of this document may be reproduced in any form without the written permission of the copyright owner.

Disclaimer

The contents of this document are subject to revision without notice due to continued progress in methodology, design and manufacturing. Ericsson shall have no liability for any error or damage of any kind resulting from the use of this document.

Trademark List

All trademarks mentioned herein are the property of their respective owners. These are shown in the document LDE Trademark Information.



Contents

1	Alarm Description	1
2	Procedure	3
2.1	Handle Alarm LOTC Disk Replication Consistency	3





1 Alarm Description

The alarm is raised when the control node pair has operated with inconsistent data for more than 20 minutes. The control node pair is in a non-redundant mode when its data is inconsistent.

Table 1 LOTC Disk Replication Consistency Alarm Causes

Alarm Cause	Description	Fault Reason	Fault Location	Impact
Inconsistent data in control node pair for more than 20 minutes	The control nodes have operated with inconsistent data for more than 20 minutes despite having connection with each other	Network failure leading to communication problems between the control nodes	Network	Both controllers take the primary role and no data is transferred between the nodes
		Disk mirroring failure leading to disk inconsistency between the control nodes	Disk failure	No immediate impact. If the node with the current data goes down, the cluster does not have a controller node containing redundant data which it can fail over to.
		Hardware failure on the secondary control node	Secondary control node	If one of the controller nodes is down, the cluster does not have a controller node with consistent data which it can fail over to

Note: This alarm can appear as a result of a maintenance activity.





2 Procedure

2.1 Handle Alarm LOTC Disk Replication Consistency

Prerequisites

- This instruction references the following documents:
 - *Data Collection Guideline*
 - *LOTC Disk Replication Communication*
- No tools are required.
- The following condition must apply:
 - The alarm is raised.

Steps

1. Is the alarm raised during initial installation or replacement of a control node?

Yes: Proceed with Step 3.

No: Continue with the next step.

2. Is the Disk Replication Communication alarm raised?

Yes: The consistency fault is most likely caused by a communication problem. Further actions are outside the scope of this instruction. Follow the procedure in *LOTC Disk Replication Communication* to clear the LOTC Disk Replication Communication alarm.

No: Continue with the next step.

3. Log on to the host to access a Linux® shell:

```
ssh <user>@<hostname> -p 7022
```

The hostname is part of alarm attribute `Source`.

4. Check which drbd version you are running:

```
cat /proc/drbd
```

```
version: 8.4.2 (api:1/proto:86-101)
GIT-hash: 7ad5f850d711223713d6dcadc3dd48860321070c build by root@lixia, 2012-09-19 16:40:30
0: cs:SyncSource ro:Primary/Secondary ds:UpToDate/UpToDate C r---n-
```

Is drbd version: 8.* ?



Yes: continue with next step.

No: proceed with Step 7.

5. Check that disk synchronization is ongoing:

```
cat /proc/drbd
```

The following is an example output:

```
version: 8.4.2 (api:1/proto:86-101)
GIT-hash: 7ad5f850d711223713d6dcadc3dd48860321070c build by root@lixia, 2012-09-19 16:40:30
0: cs:SyncSource ro:Primary/Secondary ds:UpToDate/Inconsistent C r--n-
ns:371204 nr:0 dw:252 dr:373273 al:11 bm:31 lo:0 pe:6 ua:1 ap:0 ep:1 wo:f oos:68164
[=====>...] sync'ed: 85.2% (68164/438684)K
finish: 0:00:05 speed: 11,692 (11,576) K/sec
```

Note: Synchronization is ongoing if the value after `synced:` is increasing.

On each control node, file `/proc/drbd` provides detailed information about the disk synchronization progress.

A complete disk synchronization is performed, which can take up to five hours to complete. The time to complete a disk synchronization depends on the following:

- How much data that has not been synchronized
- Disk size, disk speed, and network speed

During disk synchronization, the control nodes are not redundant.

6. Is the value after `synced:`, increasing and eventually reaching 100%?

Yes: Proceed with Step 10.

No: Proceed with Step 9.

7. Check that disk synchronization is ongoing

```
drbdsetup events2 --statistics --now
```

The following is an example output:

```
exists resource name:drbd0 role:Primary suspended:no write-ordering:flush
exists connection name:drbd0 peer-node-id:1 conn-name:node2-vc11 connection:Connected role:Secondary congested:no
exists device name:drbd0 volume:0 minor:0 disk:UpToDate size:10485760 read:9102377 written:3720 al-writes:5 =>
bm-writes:0 upper-pending:0 lower-pending:10 al-suspended:no blocked:no
exists peer-device name:drbd0 peer-node-id:1 conn-name:node2-vc11 volume:0 replication:SyncSource =>
peer-disk:Inconsistent resync-suspended:no received:0 sent:8813432 out-of-sync:1672332 pending:0 unacked:10
exists -

exists resource name:drbd0 role:Primary suspended:no write-ordering:flush
exists connection name:drbd0 peer-node-id:1 conn-name:node2-vc11 connection:Connected role:Secondary congested:no
exists device name:drbd0 volume:0 minor:0 disk:UpToDate size:10485760 read:9207989 written:3720 al-writes:5 =>
bm-writes:0 upper-pending:0 lower-pending:4 al-suspended:no blocked:no
exists peer-device name:drbd0 peer-node-id:1 conn-name:node2-vc11 volume:0 replication:SyncSource =>
peer-disk:Inconsistent resync-suspended:no received:0 sent:8919044 out-of-sync:1568768 pending:2 unacked:4
exists -
```




Note: Synchronization is ongoing if the value after `out-of-sync:` is decreasing.

On each control node, command `drbdsetup events2 --statistics --now` provides detailed information about the disk synchronization progress.

A complete disk synchronization is performed, which can take up to five hours to complete. The time to complete a disk synchronization depends on the following:

- How much data that has not been synchronized
- Disk size, disk speed, and network speed

During disk synchronization, the control nodes are not redundant.

8. Is the value after `out-of-sync:` decreasing and eventually reaching 0?

Yes: Continue with the next step.

No: Proceed with Step 10

9. Is the value after `synced:`, increasing and eventually reaching 100%?

Yes: Continue with the next step.

No: Proceed with Step 11.

10. Is the alarm cleared?

Yes: Proceed with Step 16.

No: Proceed with Step 14.

Step 14

11. Identify the Distributed Replicated Block Device (DRBD) interfaces as follows:

- a. Get the name of the interface (`eth<x>`):

```
cat /etc/cluster/nodes/this/networks/internal/primary/interface/name
```

The following is an example output:

```
eth0
```

- b. Get the IP address (`<ip>`):

```
cat /etc/cluster/nodes/this/networks/internal/primary/address
```



The following is an example output:

```
169.254.43.11
```

- c. Get the network mask (<netmask>):

```
cat /etc/cluster/nodes/this/networks/internal/primary/network/netmask
```

The following is an example output:

```
255.255.255.0
```

12. Check the log `/var/log/messages` for recent system log messages indicating DRBD interface-related issues, for example (to show the last 1000 lines in the log):

```
tail -1000 /var/log/messages
```

The following is an example output in a faulty situation with network interface issues:

```
Aug 26 12:17:52 SC-1 kernel: [ 277.720545] hrtimer: interrupt took 572013 ns
Aug 26 12:32:50 SC-1 kernel: [ 1175.612842] tipc: Resetting bearer <eth:eth0>
Aug 26 12:32:50 SC-1 dhcpd: receive_packet failed on eth0: Network is down
Aug 26 12:32:50 SC-1 syslog-ng[1810]: I/O error occurred while writing; fd='6', error='Network =>
is unreachable (101)'
Aug 26 12:32:50 SC-1 syslog-ng[1810]: Connection broken; time_reopen='10'
Aug 26 12:32:59 SC-1 ntpd[2240]: sendto(192.0.2.10) (fd=23): Network is unreachable
Aug 26 12:33:00 SC-1 syslog-ng[1810]: Connection failed; error='Network is unreachable (101)'
Aug 26 12:33:00 SC-1 syslog-ng[1810]: Initiating connection failed, reconnecting; time_reopen='10'
Aug 26 12:33:10 SC-1 syslog-ng[1810]: Connection failed; error='Network is unreachable (101)'
Aug 26 12:33:10 SC-1 syslog-ng[1810]: Initiating connection failed, reconnecting; time_reopen='10'
Aug 26 12:33:20 SC-1 syslog-ng[1810]: Connection failed; error='Network is unreachable (101)'
Aug 26 12:33:20 SC-1 syslog-ng[1810]: Initiating connection failed, reconnecting; time_reopen='10'
```

Note: In this output, `eth0` is the interface used by the DRBD.

13. Are there any issues with the network interface used for the DRBD?

Yes: The consistency fault is most likely caused by a communication problem. Further actions are outside the scope of this instruction. Follow the procedure in *LOT Disk Replication Communication* to clear the LOT Disk Replication Communication alarm, which is to be raised within 20 minutes if not raised yet.

No: Continue with the next step.

14. Perform data collection, refer to *Data Collection Guideline*.
15. Consult the next level of maintenance support. Further actions are outside the scope of this instruction.
16. Job is completed.