12/19/2019

IBM® TS7700 Series
Grid Resiliency Improvements User's Guide
Version 1.1

Takeshi Nohta
nohta@jp.ibm.com
IBM Systems Development, IBM Japan

Contents

Summary of Changes 3
1 Introduction to Grid Resiliency Improvements 4
2 Definitions and Keywords 5
3 Supported Code Level and Configuration 5
4 Local Cluster Fence 6
4.1 Local Cluster Fence Mechanism 6
4.2 Local Cluster Fence Monitored Symptoms 6
5 Remote Cluster Fence 7
5.1 How to determine if the remote cluster is SBND 8
5.1.1 Diag Data for Remote Cluster Fence function 8
5.1.2 Thresholds and TIME options to trigger Remote Cluster Fence action 10
5.1.3 TIME option "DELAY" 13
5.1.4 LI REQ, DIAGDATA, RESET command 14
5.2 How to fence remote SBND cluster 15
5.2.1 Healthy Cluster Agreement Rule 15
5.2.2 Remote Cluster Fence Actions 17
5.2.3 Apply Remote Cluster Fence Actions 18
6 Manual Cluster Fence 20
7 Immediate Takeover against a fenced cluster 21
8 Cluster Unfence Operation 21
8.1 Cluster Unfence Operation from MI 21
8.2 Mixed Code or Standalone Configuration consideration 23
9 Customer Notification 23
9.1 Operator Messages 23
9.2 Management Interface 25
9.2.1 Grid Summary/Fence Mode Page 25
9.2.2 Events 27
9.2.3 Tasks 29
9.3 LI REQ 30
9.4 Call Home/Service Information Message (SIM) 31
10 Library Request Commands 31
10.1 LI REQ, <comp lib>, FENCE, {ENABLE|DISABLE} 32
10.2 LI REQ, <comp lib>, FENCE, THRESHLD, {SCRVOAVG|PRIVOAVG|VCAVG|TOKAVG|TMO|ERR|EVALWIN}, <value> 32
10.3 LI REQ, <comp lib>, FENCE, TIME, {DELAY|CONSCNT}, <value> 33
10.4 LI REQ, <dist lib>, FENCE, ACTION, PRI, {NONE|ALERT|OFFLINE|REBOOT|REBOFF} 34
10.5 LI REQ, <dist lib>, FENCE, ACTION, SEC, {ENABLE|DISABLE} 34
10.6 LI REQ, <dist lib>, FENCE, ACTION, AIXDUMP, {ENABLE|DISABLE} 34
10.7 LI REQ, <comp lib>, FENCE, SHOW 35
10.8 LI REQ Error Text 40
11 Disclaimers 42

Summary of Changes

- Version 1.0 – First Release (Supported code level is R4.1.2 (8.41.200.x)).
- Version 1.1 – Added a new local fence reason code supported at R5.0 (8.50.x.x).

1 Introduction to Grid Resiliency Improvements

The IBM TS7700 Series is the latest in the line of tape virtualization products that has revolutionized the way mainframe customers run system z tape operations. A TS7700 Grid is made up of two or more TS7700 clusters interconnected through Ethernet connections and it is designed and implemented as a business continuance solution with implied enterprise resiliency. When a cluster in the Grid has a problem, the multi-cluster Grid should accommodate the outage and have the ability to continue the operation even if the state of the Grid is degraded.

In the last year, a number of customers have experienced grid-wide problems that are due to a single cluster in the Grid experiencing a problem. The issues we have seen are problems in one cluster that causes it to be sick or unhealthy, but not completely dead (Sick But Not Dead (SBND)). Then the peer clusters are greatly affected, and customer jobs end up being affected (long mount time, failed sync mode writes… much more than degraded).

Each issue which causes the symptom in grid-wide has a root cause and the issue can be fixed one by one. But resiliency should accommodate symptoms themselves and better handling unexpected issues to prevent negative impacts to the Grid when it occurs so that the symptoms don't hurt the production processing.

Grid Resiliency Improvements are the functions to identify the symptoms and make the Grid more resilient when a single cluster experiences a problem by removing the sick or unhealthy cluster from the Grid, either explicitly or implicitly through different methods. By removing it, the rest of peer clusters can then treat it as "dead" and avoid further handshakes with it until it can be recovered.

Figure 1 shows a typical failure scenario which Grid Resiliency improvements should resolve.

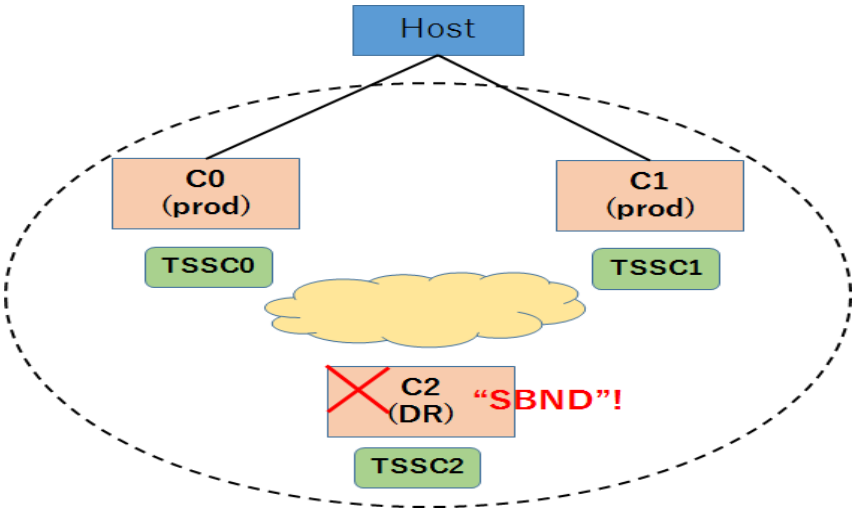

Figure 1. Grid Resiliency scenario and purpose

It is 3-way Grid (C0/C1/C2) and C0/C1 are production clusters while C2 is a DR cluster:

1. C2 becomes SBND (Sick But Not Dead), but C0/C1 still see C2 as "functional" then they still try to communicate with C2.
2. C2 is SBND so all communication request from C0/C1 to C2 are delayed or timed out.

3. In the end, the production jobs to C0/C1 starts being affected until C2 is isolated from Grid (i.e. shutdown/reboot…).

Grid Resiliency improvements are the functions or abilities which are introduced in R4.1.2 (8.41.200.xx or above) to remove a single SBND cluster (C2 in Figure 1) from a Grid automatically or manually.

The following sections will describe the functions and all its supported features. Throughout this white paper, it's assumed the reader has a good understanding of the TS7700 Grid functionality.

2 Definitions and Keywords

The following definitions and keywords are used throughout this document.

- **GR:** Grid Resiliency
- **LI REQ**: z/OS Host Command Library Request command
- **Diagnostic Data (Diag Data)**: The elapsed time, timeout, counts and error counts between two or more clusters in in TS7700 Grid configuration. LI REQ DIAGDATA was introduced in R3.2 PGA2 or above to provide Diag Data.
- **Fence:** Temporarily remove a cluster from the Grid
- **Fence Action:** How to remove the cluster from Grid temporarily (either of "force offline", "reboot", "reboot and stay at offline" or "isolate from Grid" ("alert" is also a part of the fence actions").
- **Local Fence:** A cluster does fence by itself.
- **Remote Fence:** A cluster does fence another (remote/peer) cluster in the Grid.
- **Manual Fence:** User does fence the local/remote cluster manually from Management Interface (MI)
- **Unfence:** Get the fenced cluster back into the Grid
- **SBND:** Sick But Not Dead. SBND cluster means an unhealthy cluster which still looks live from the peer clusters in the Grid.

3 Supported Code Level and Configuration

GR functions consist of 2 major sub-functions, local and remote cluster fence. Manual fence function is also supported.

- Local Fence is supported once the local cluster is at R4.1.2 or above. It's enabled on a standalone configuration as well as mixed code configuration. It's always enabled and cannot be disabled.

- Remote Fence is supported when all clusters in the Grid are at R4.1.2 or above. It's not supported under mixed code configuration. Only Grid (not standalone) configuration is supported. It can be enabled or disabled through LI REQ. The default is disabled.

- Manual Fence is supported once the local cluster is at R4.1.2 or above. New MI panel to fence the target cluster is provided. Manual remote fence is supported only when all clusters in the Grid are at R4.1.2 or above. Manual local fence is supported on a standalone or under mixed code configuration as well.

4 Local Cluster Fence

This chapter explains how GR local cluster fence function works.

4.1 Local Cluster Fence Mechanism

In the prior code levels (R4.1.1 (8.41.100.x) or below), TS7700 has some local fence mechanisms. One of them is, for example, the component heartbeat check described. Figure 2 (left side) shows the basic heartbeat check mechanism:

- Each component reports the heartbeat periodically.
- If the heartbeat from any component which is supervised by another component (High Availability (HA) component) has been missing for a certain period, the local cluster is rebooted.

This mechanism has been working expectedly but there are two areas to improve in R4.1.2:

- ✓ There have been situations where a component can be SBND, but it is still able to continue the heartbeat, then SBND condition is not detected by this mechanism.
- ✓ The cluster reboots without any notification to the host/MI etc., then it may be difficult to figure out why the local cluster is/was rebooted until IBM support is involved.

Therefore, R4.1.2 GR local cluster fence mechanism is improved as described in Figure 2 (right side):

- Each component monitors its own state. Also, some components monitor other components' state. When a component detects SBND condition, it's reported to HA component with the appropriate fence action and the local cluster is fenced.
- The notification is sent to the host and IBM support through the operational messages, call home etc. prior to fencing the local cluster.

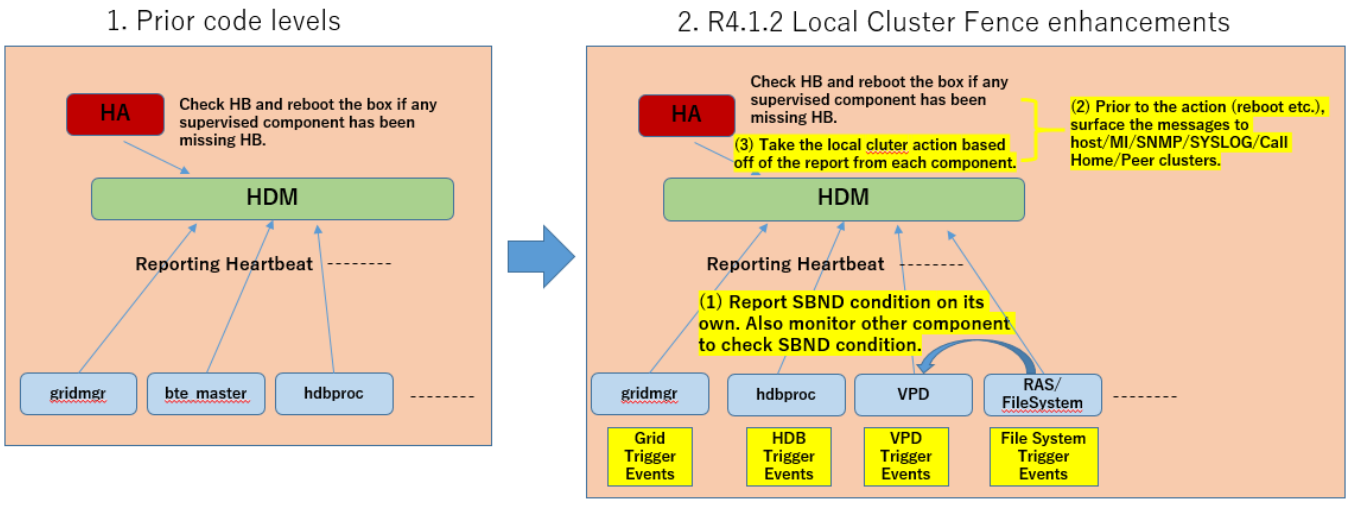

Figure 2. Local Cluster Fence mechanism

4.2 Local Cluster Fence Monitored Symptoms

In R4.1.2, the following symptoms are monitored and initiate the local fence action (reboot the local cluster by itself):

- Access check of the files on the cache subsystem
  if the periodical file access check under 2 file systems TS7700 manages has failed for more than 20 minutes.
- The number of database connections
  if any internal component holds more than 700 DB connections.
- The number of database transaction timeouts
  if 5 or more DB transactions are consistently timing out and more than 10 timeouts has popped up.
- Loss of hardware redundancy in the cache filesystems
  if the cache filesystems hardware redundancy is lost and the system cannot ensure data integrity
- The process to control TS7700 VPD (Vital Product Data)
  if the component which manages TS7700 Vital Product Data has been missing (i.e. failed to respawn when it needs to be restarted).
- The number of the handlers for the communication between the components
  if the utilization of the communication handles between the internal components exceeds 90 %.
- Component heartbeat status
  if any component has been missing the heartbeat for more than 15 minutes.
- Waiting I/O to the cache file system
  if the cache filesystem I/O wait symptom has occurred for more than 600 seconds.
- Stuck file system access
  if the access to the cache file systems TS7700 manages has been stuck

5 Remote Cluster Fence

This chapter explains how GR remote cluster fence function works. When a cluster becomes SBND, it's optimal for the local cluster to fence itself by applying the local cluster fence action. But there may be any unknown/unexpected reasons and the current local cluster fence mechanism may not be able to detect the SBND condition. The remote cluster fence is a new mechanism introduced in R4.1.2 to detect the SBND symptoms which are caused by the remote cluster and trigger the remote cluster fence action from a peer cluster.

Figure 2 shows the basic concept of the remote cluster fence mechanism:

1. C2 becomes SBND, but it cannot trigger the local cluster fence action.
2. A peer cluster (C0) detects SBND symptoms against C2.
   - Several checks are executed to verify the SBND condition on C2.
3. If all checks are positive, C0 triggers the remote cluster fence action against C2.

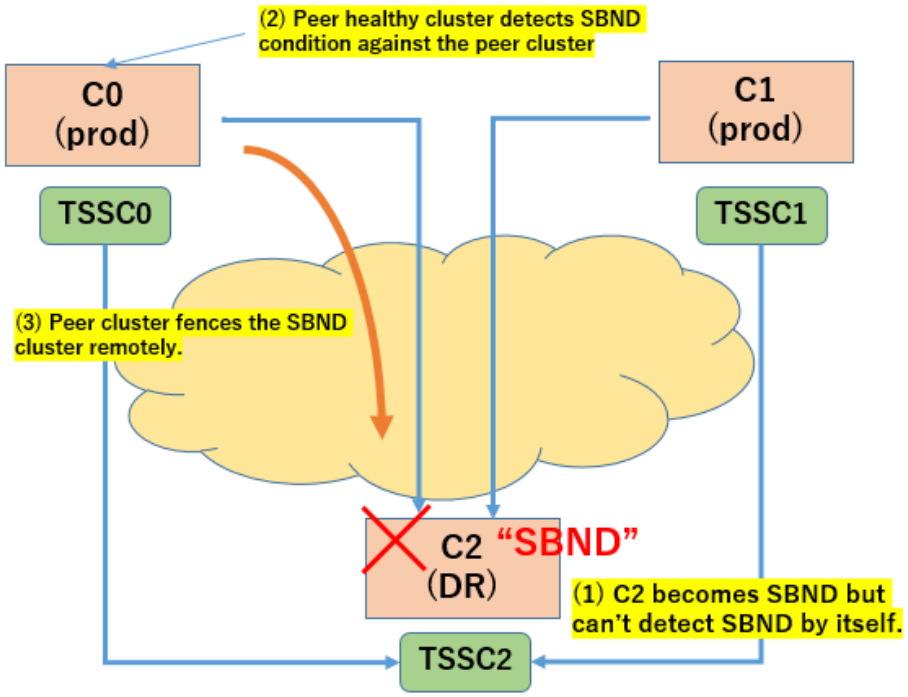

Figure 3. Remote Cluster Fence mechanism

From the next sections, the following 2 basic concepts of the Remote Cluster Fence function are explained:

1. **How to determine if the remote cluster is SBND**
2. **How to fence the remote SBND cluster**

5.1 How to determine if the remote cluster is SBND

This section explains how to determine if the remote cluster is SBND and trigger the remote cluster fence action.

5.1.1 Diag Data for Remote Cluster Fence function

Remote Cluster Fence function uses Diag Data which was released in R3.3 PGA2 (8.33.2.9) to determine if the remote cluster is SBND or not. Please refer to DIAGDATA Guidance white paper (http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102701) to understand what Diag Data is, but Figure 4 explains Diag Data and how it can be correlated with SBND cluster briefly.

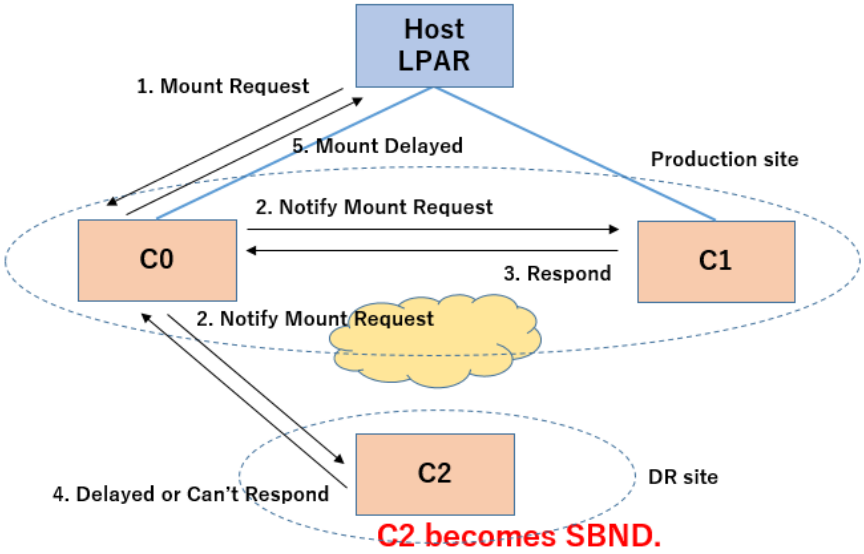

Figure 4. How the customer I/O is affected by SBND cluster

In Figure 4, it's 3-way Grid and a host zLPAR is connected to C0/C1 as production clusters and C2 is a DR cluster:

1. zLPAR issues a mount request to C0.
2. C0 receives the mount request and notifies C1/C2 of the mount request.
3. C0 times the response of the mount request from C1/C2. This is the source data of Diag Data. Once C0 receives the response from C1/C2, the mount request is processed.
   Note: There are multiple handshake
4. Given C2 becomes SBND, the response starts being delayed or timed out

As explained above, the statistics of Diag Data can be used to evaluate the remote cluster's state indirectly. If the statistics has been showing consistently delayed or timed out, the remote cluster may be becoming SBND.

Figure 5 shows LI REQ DIAGDATA output example. Remote Cluster Fence function uses the statistics of Diag Data in the red rectangles. Diag Data of the volume open in the mount processing, volume close in the demount processing and the internal token handshakes are used.

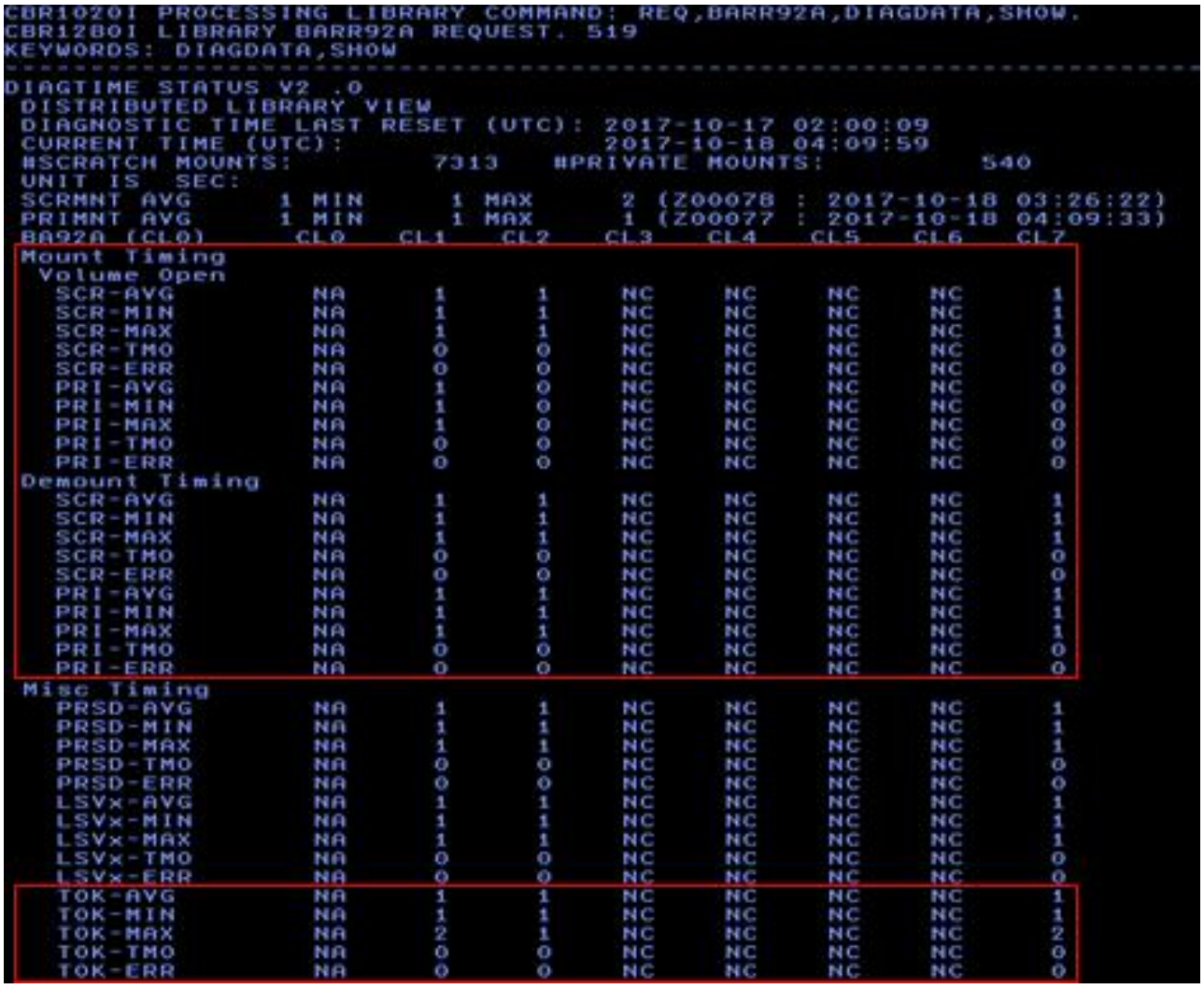

Figure 5. LI REQ DIAGDATA output example
(Diag Data used for Remote Cluster Fence functions are in the red rectangles)

5.1.2 Thresholds and TIME options to trigger Remote Cluster Fence action

Remote Cluster Fence function checks the statistics of Diag Data and compares them to the thresholds which can be set through LI REQ to figure out if the remote cluster needs to be considered as SBND. Also, its TIME options are also taken into account. Figure 6 shows the new supported LI REQ (FENCE, SHOW) to provide the current thresholds and TIME options of a Grid:

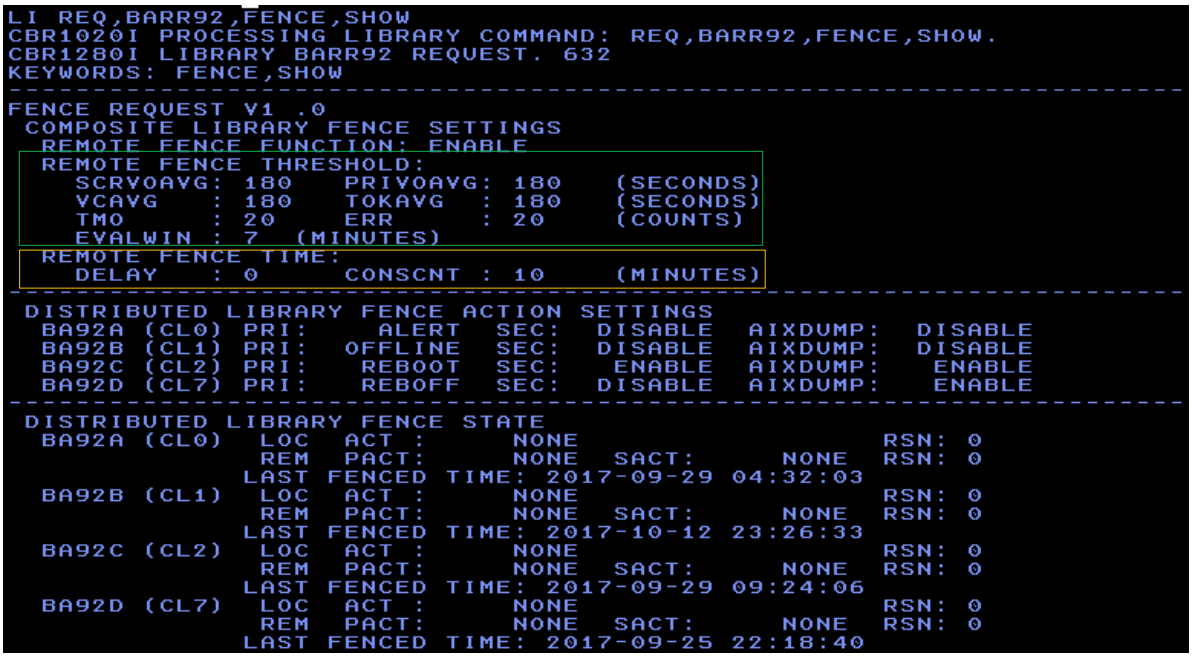

Figure 6. LI REQ, FENCE, SHOW output example (thresholds/TIME options)

The values in the green rectangle are the thresholds and the ones in the yellow rectangle are TIME options of the Remote Cluster Fence function.

The following thresholds can be set through LI REQ:

- SCRVOAVG: The average elapsed time of the volume open/mount request peer handshakes for scratch mounts (Default = 180 seconds. 0 – 1200 can be set (0 means no check))
- PRIVOAVG: The average elapsed time of the volume open/mount request peer handshakes for private mounts (Default = 180 seconds. 0 – 1200 can be set (0 means no check))
- VCAVG: The average elapsed time of the volume close/demount (Rewind Unload) request peer handshakes (all mount types) (Default = 180 seconds. 0 – 1200 can be set (0 means no check))
- TOKAVG: The average elapsed time of token request handshakes (with the timeout value less than 3 minutes) (Default = 120 seconds. 0 – 180 can be set (0 means no check))
- TMO: The timeout count of all major peer handshakes (all mounts/demounts/tokens/others) (Default = 20 count. 0 – 1000 can be set (0 means no check))
- ERR: The error count of all major peer handshakes (all mounts/demounts/tokens/others) (Default = 20 count. 0 – 1000 can be set (0 means no check))
- EVALWIN: The diag data evaluation window (Default = 7 minutes (1 – 30 can be set))

Note: The minimum value (1) can be set to all the above thresholds. However, defining very aggressive low values could lead to a "False Remote Cluster Fence" due to heavy workload conditions.

SCRVOAVG and PRIVOAVG are the thresholds for the volume open/mount request. The default timeout value of the peer handshake is 20 minutes.
VCAVG is the threshold of the volume close/demount (Rewind Unload) request. The default timeout value of the peer handshake is also 20 minutes. This threshold is applied to both private and scratch volume demount request.
TOKAVG is the threshold of various token request handshakes which are used in Grid. They have different timeout values but Diag Data only times the token handshakes with 3 minutes timeout value.

TKM and ERR are the timeout counts while SCROAVG/PRIOAVG/VCAG/TOKAVG are the average elapsed time. The timeout counts include the handshakes which are executed during the mounts/demounts/tokens/others.

TS7700 stores the statistics of Diag Data periodically. Then the remote cluster is considered as SBND when the statistics of Diag Data are evaluated and have crossed the above configured thresholds. But the remote cluster fence action is actually triggered only when they have exceeded the thresholds over a moving window of time for a configured consecutive number of minutes. The threshold "EVALWIN" and one of TIME options "CONSCNT" determine the window and consecutive durations:

- EVALWIN (evaluation window): Diag Data evaluation window (Default = 7 minutes (1 – 30 can be set))
- CONSCNT (consecutive fence condition count): The consecutive fence check condition count (minute) prior to kicking the actual remote cluster fence action (Default = 10 minutes (0 – 60 can be set))

All Diag Data information is analyzed each minute and only the last "EVALWIN" period of time is evaluated to determine if any of the specific configured and enabled thresholds have been crossed during that window. Then, only if the last "CONSCNT" counts (minutes) of "EVALWIN" checks all show one or more specific thresholds were triggered consecutively, the remote cluster fence will be initiated.

Figure 7 shows how EVALWIN and CONSCNT are used to evaluate the statistics of Diag Data.

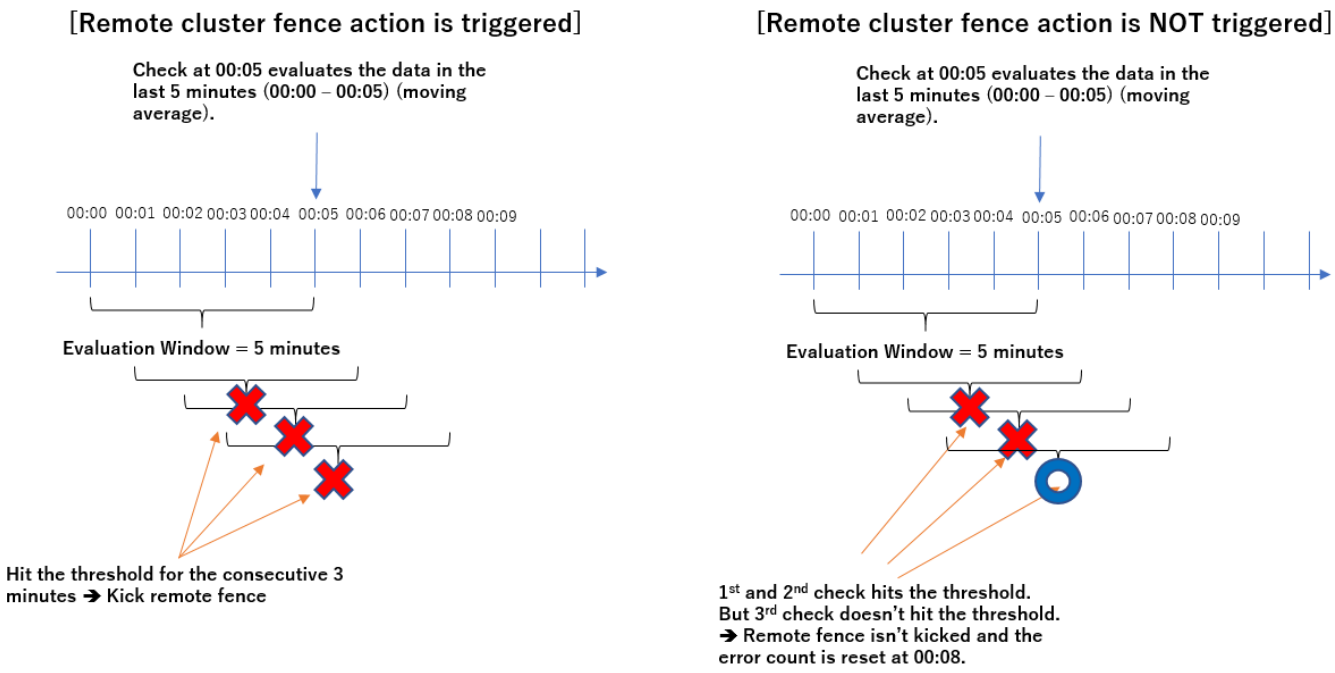

Figure 7. How EVALWIN and CONSCNT are used to evaluate the statistics of Diag Data

Diag Data is gathered during all runtime peer cluster handshakes. Remote Cluster Fence function evaluates the snapshots of Diag Data once every minute. As an example, Figure 7 shows the examples that the snapshots are taken and valuated at 00:00, 00:01, 00:02 and so on, and EVALWIN = 5 minutes and CONSCNT = 3 minutes are used. Looking at the left side in the figure. At 00:06, the snapshots for the last 5 minutes (the statistics data from 00:01 to 00:05) are evaluated and if any of the threshold

events are true, a flag is set to indicate the remote cluster is SBND. But the remote cluster fence action is not yet triggered. The same check has been running and if any of the threshold events are also true at 00:07 and 00:08, the check has crossed the threshold for the consecutive 3 minutes (counts), then the remote cluster fence action is triggered against the remote SBND cluster. On the other hand, the right side in the figure shows the same error symptoms at 00:06 and 00:07, but at 00:08, none of the threshold events are true. In that case, the remote cluster fence action is not triggered and the error flag is reset at 00:08.

It's not required to have crossed the same threshold event for the consecutive "CONSCNT" interval in order to trigger the remote cluster fence action. For example:

- ✓ At 00:06, SCRVOAVG threshold has been crossed.
- ✓ At 00:07, VCAVG threshold has been crossed.
- ✓ At 00:08, TMO threshold has been crossed.

The different threshold events have been true for the consecutive interval, then the remote cluster fence action is still triggered.

TS7700 checks the total counts of the volume open/mount request peer handshakes for scratch and private mounts as well as the volume close/demount (Rewind Unload) request peer handshakes, which are evaluated by SCRVOAVG/PRIVOAVG/VCAVG thresholds. If the total count in "EVALWIN" duration doesn't reach the minimum count (10 handshakes), TS7700 considers the data for the type is not enough to evaluate and just skips the remote cluster fence check against the data.

5.1.3 TIME option "DELAY"

There are 2 settings which can be set in TIME options. "CONSCNT" is explained in the previous section and another "DELAY" option is explained in this section.

- ➢ DELAY: The delay in minute to execute the actual remote cluster fence since the target remote cluster is determined as SBND (Default = 0 minutes (0 – 60 can be set))

Figure 8 shows when the actual remote cluster fence action is kicked when TIME option "DELAY" is set to 0 and 15 minutes.

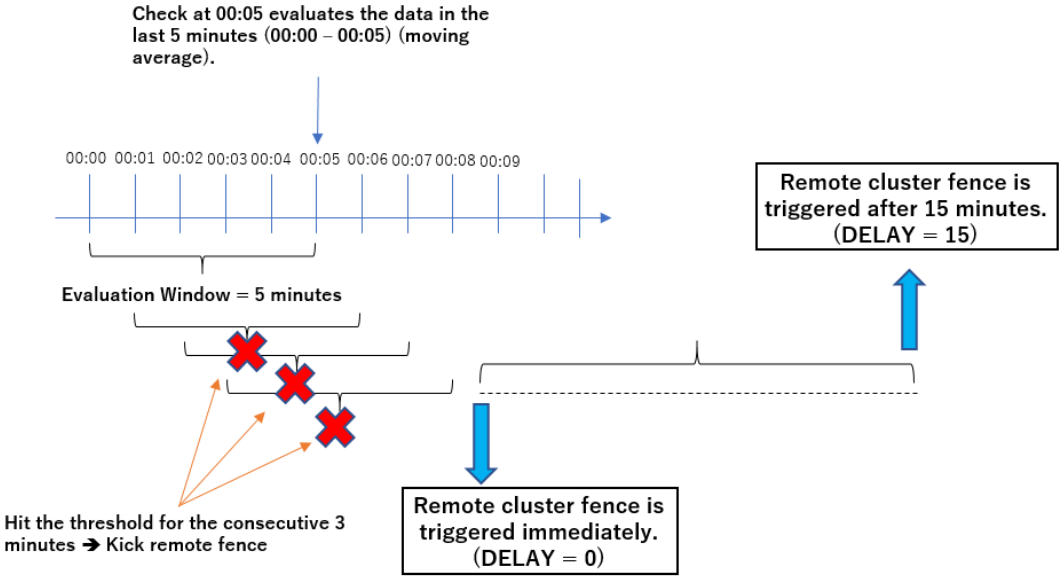

Figure 8. How TIME option "DELAY" works

In Figure 8, any of threshold events become true for the consecutive CONSCNT (3) minutes at 00:08, then the remote cluster fence action needs to be triggered. If "DELAY" is set to 0, the action is triggered immediately. But if "DELAY" is set to 15 minutes, the action is triggered after 15 minutes (at 00:23).

This delay may be set when the manual operations/interventions, such as varying devices offline attached to the remote cluster which will be fenced in "DELAY" time), are required.

When this "DELAY" is set to other than 0, an operator message is surfaced before entering "DELAY" interval so that the user is notified of the fence action which will be applied in "DELAY" minutes.

When the manual remote cluster fence action is applied or the remote cluster fence action is "ALERT", no delay is applied although "DELAY" is set to other than 0.

As explained in the sections 5.1.2 and 5.1.3, here is the summary of the threshold and TIME options usage of "EVALWIN", "CONSCNT" and "DELAY" usage:

- "EVALWIN" is used to check how much of past data should be evaluated against the thresholds (moving average/total timeout/error count) for one-minute check.
- "CONSCNT" is used to determine how many consecutive "EVALWIN" checks must detect threshold condition was crossed before triggering the configured fence action.
- "DELAY" is used to configure how must of delay should occur prior to applying the remote cluster fence function after the remote cluster is determined to be fenced.

5.1.4 LI REQ, DIAGDATA, RESET command

Diag Data is reset per each cluster by using LI REQ, DIAGDATA, RESET command. When it's reset, the statistics of Diag Data which is used by the remote cluster fence function is also reset.

The remote cluster fence function uses the statistics for the entire evaluation window configured by "EVALWIN" threshold. Once the statistics is reset, the remote cluster fence function will need to wait

for "EVALWIN" duration to store enough data to evaluate the remote cluster's state. Then the evaluation will not start until "EVALWIN" duration has passed since the reset.

5.2 How to fence remote SBND cluster

This section explains how to fence the remote SBND cluster once the remote cluster is determined as SBND.

5.2.1 Healthy Cluster Agreement Rule

Once a remote cluster is determined as SBND based off of the thresholds/EVALWIN/CONSCNT settings by a healthy cluster in Grid, the heathy cluster always asks the rest of all clusters except the remote SBND cluster to check if they also think the remote cluster is surely SBND. Only when the rest of all clusters agree that the target remote cluster is SBND, the remote cluster fence action is kicked. For example, if it's 8-way Grid and C0 determines C7 is SBND, C0 asks C1 to C6 if C7 is viewed as SBND from them as well. Only when C1 to C6 all think C7 is SBND, the remote cluster fence action is kicked. If any of the clusters doesn't agree, is in service, is already fenced, or unavailable (offline etc.), the remote cluster fence action is not kicked.

Figure 9 shows both scenarios that the remote cluster fence action is applied and denied based off of the agreement and disagreement from the rest of healthy clusters. C0 detects SBND symptoms against C2 in 3-way Grid:

(1) C0 detects SBND symptoms against C2.
(2) C0 asks C1 to check if C1 agrees to fence C2.
[Cluster Fence is applied (in yellow)]
(3) C1 checks C2 state locally and remotely if required and agrees to fence C2.
(4) C0 kicks the remote cluster fence action against C2.

[Cluster Fence is denied (in gray)]
(3)' C1 checks C2 state locally and remotely if required and disagrees to fence C2.
(4)' C0 doesn't kick the remote cluster fence action and resets the internal error count against C2.

When C1 is asked from C0 to check if C2 is SBND, C1 does the following checks:

- If the statistics of Diag Data on C1 against C2 already shows SBND symptoms, C1 agrees to fence C2.
- If the statistics doesn't show SBND symptom against C2 (it could happen when no or little I/O is applied from zLPAR to C1), C1 attempts a converged health check to C2. If the health check detects SBND symptoms, C2 is viewed as SBND from C1 viewpoint.
  Note: The converged health check performs some token handshake, dummy mount/demounts to the remote target cluster and the local token check on the remote target cluster.

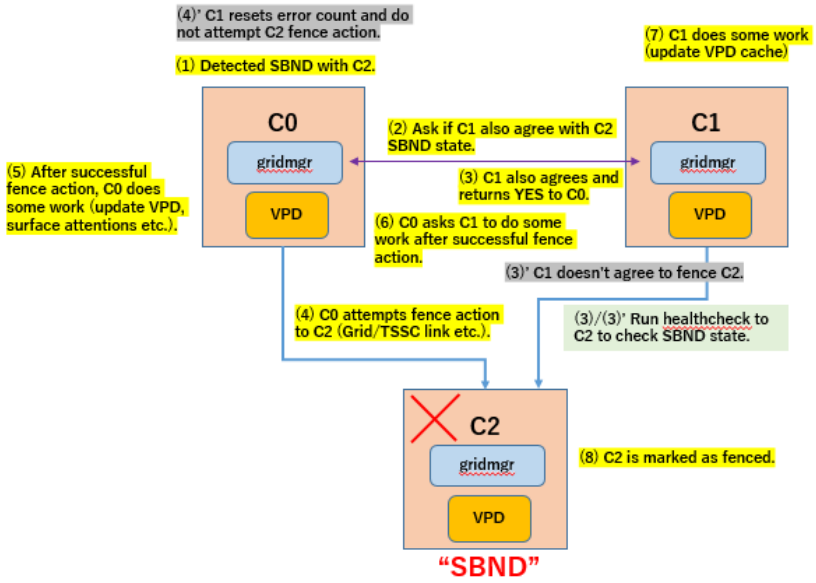

Figure 9. Healthy cluster agreement rule (3-way or more Grid)

However, on 2-way Grid, there is no remote healthy cluster to ask if the target cluster is SBND or not. Figure 10 shows how the agreement rule is applied under 2-way Grid:

(1) C0 detects SBND symptoms against C1.
(2) C0 asks C1 to check how C1 views C0.
[Cluster Fence is applied (in yellow)]
(3) C1 checks C1 state locally and remotely if required and agrees to fence C2.
(4) C0 kicks the remote cluster fence action against C2.

[Cluster Fence is denied (in gray)]
(3)' C1 checks C2 state locally and remotely if required and disagrees to fence C2.
(4)' C0 doesn't kick the remote cluster fence action and resets the internal error count against C2.

When C1 is asked from C0 to check if C1 views C0 as SBND, the same methods described at 3-way or more Grid case is applied.

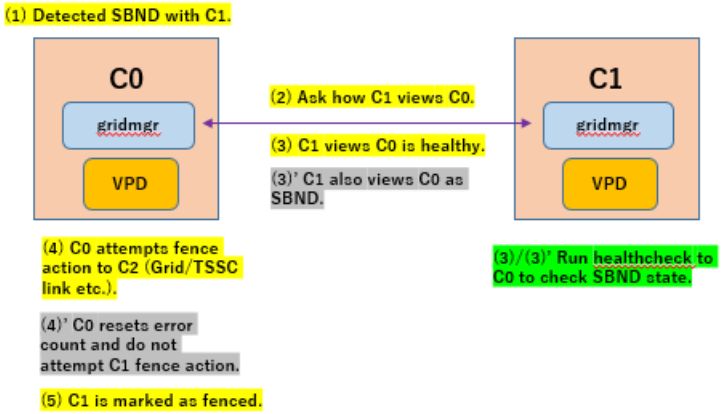

Figure 10. Healthy cluster agreement rule (2-way Grid)

5.2.2 Remote Cluster Fence Actions

There are two types of the remote cluster fence actions, primary and secondary fence actions. In order to fence a remote SBND cluster, TS7700 always applies the primary fence action first, then secondary fence action if enabled and required.

[Primary Fence Action]
Any of the five primary fence actions below can be set per each cluster through LI REQ:

- NONE: No remote cluster fence action is applied.
- ALERT: Only alerts (operator messages, event creation etc.) are reported. The cluster is allowed to be in Grid even after it's determined as SBND. This action may be selected if it's preferable to control when and how the SBND cluster should to be fenced manually.
- REBOOT: Reboot the remote cluster. The remote cluster may be able to come up to online automatically if the reboot successfully completes. This action may be selected if it's preferable to reboot the SBND cluster and allow it back to be operational without further investigation by IBM service personnel.
- OFFLINE: Put the remote cluster into offline state.
- REBOFF: Reboot the remote cluster and leave it offline state.
  OFFLINE and REBOFF actions may be selected if IBM service personnel should be involved for the further investigation prior to recovering the fenced cluster back to the operational state.

[Secondary Fence Action]
The secondary fence action can be enabled or disabled per each cluster through LI REQ and it's applied under the following conditions if enabled:

- All attempts to apply the primary fence action to the remote SBND cluster fails.

or

- The primary fence action (OFFLINE/REBOOT/REBOFF) has been successfully applied, but the remote SBND cluster hasn't been offline for more than 20 minutes.

Note: When the secondary fence action is applied successfully, both primary and secondary fence actions are set against the remote cluster to indicate both fence actions were applied.

When the secondary fence action is applied successfully, the remote SBND cluster is isolated from Grid by shutting down the Grid link ports (1415/1416/350) to the remote SBND cluster on all the rest of healthy clusters in the Grid (Note: the Grid link ports for the support usage (ssh/ping) still usable).

[AIX System Dump Option]
In addition to the primary and secondary remote cluster fence actions, another option to determine if AIX dump is taken when REBOOT/REBOFF action is applied can be set per each cluster through LI REQ as well. AIXDUMP option can be enabled or disable. If it's enabled, AIX dump is taken for the data capture when the remote fence action REBOOT/REBOFF as well as local cluster fence action (reboot) are applied successfully.

Figure 11 shows the new supported LI REQ (FENCE, SHOW) to provide the current remote cluster fence action options per each cluster:

```
LI REQ,BARR92,FENCE,SHOW
CBR1020I PROCESSING LIBRARY COMMAND: REQ,BARR92,FENCE,SHOW.
CBR1280I LIBRARY BARR92 REQUEST. 632
KEYWORDS: FENCE,SHOW
------------------------------------------------------------------------
FENCE REQUEST V1 .0
 COMPOSITE LIBRARY FENCE SETTINGS
  REMOTE FENCE FUNCTION: ENABLE
  REMOTE FENCE THRESHOLD:
    SCRVOAVG: 180    PRIVOAVG: 180    (SECONDS)
    VCAVG   : 180    TOKAVG  : 180    (SECONDS)
    TMO     : 20     ERR     : 20     (COUNTS)
    EVALWIN : 7  (MINUTES)
  REMOTE FENCE TIME:
    DELAY   : 0      CONSCNT : 10     (MINUTES)
------------------------------------------------------------------------
 DISTRIBUTED LIBRARY FENCE ACTION SETTINGS
  BA92A (CL0) PRI:    ALERT  SEC:  DISABLE  AIXDUMP:  DISABLE
  BA92B (CL1) PRI:  OFFLINE  SEC:  DISABLE  AIXDUMP:  DISABLE
  BA92C (CL2) PRI:   REBOOT  SEC:   ENABLE  AIXDUMP:   ENABLE
  BA92D (CL7) PRI:   REBOFF  SEC:  DISABLE  AIXDUMP:   ENABLE
------------------------------------------------------------------------
 DISTRIBUTED LIBRARY FENCE STATE
  BA92A (CL0)  LOC  ACT :     NONE                      RSN: 0
               REM  PACT:     NONE  SACT:     NONE      RSN: 0
              LAST FENCED TIME: 2017-09-29 04:32:03
  BA92B (CL1)  LOC  ACT :     NONE                      RSN: 0
               REM  PACT:     NONE  SACT:     NONE      RSN: 0
              LAST FENCED TIME: 2017-10-12 23:26:33
  BA92C (CL2)  LOC  ACT :     NONE                      RSN: 0
               REM  PACT:     NONE  SACT:     NONE      RSN: 0
              LAST FENCED TIME: 2017-09-29 09:24:06
  BA92D (CL7)  LOC  ACT :     NONE                      RSN: 0
               REM  PACT:     NONE  SACT:     NONE      RSN: 0
              LAST FENCED TIME: 2017-09-25 22:18:40
```

Figure 11. LI REQ, FENCE, SHOW output example (fence actions)

The values in the red rectangle are the remote cluster fence actions of the Remote Cluster Fence function.

The following thresholds can be set through LI REQ:

- PRI: The primary remote cluster fence action of the cluster. "ALERT", "OFFLINE", "REBOOT", "REBOFF" should be displayed.
- SEC: The secondary remote cluster fence action of the cluster. "ENABLE" or "DISABLE should be displayed.
- AIXDUMP: AIX dump option. "ENABLE" or "DISABLE should be displayed.

5.2.3 Apply Remote Cluster Fence Actions

This section explains how the configured remote cluster fence action is applied. The remote cluster fence action is attempted using a few methods. Figure 11 shows the variations in the methods to apply the remote cluster fence action. In Figure 12, C0 detects SBND symptoms against C2 and C1 also agrees to fence C2, then C0 starts attempting the remote cluster fence action against C2. The primary remote cluster fence action "OFFLINE" is configured and the secondary remote cluster fence action is enabled against C2:

(1) The primary remote fence action "OFFLINE" is applied to C2 through Grid link port 1415/1416.
(2) If (1) fails, the same primary fence action is applied to C2 through Grid link port 350.
(3) If (2) fails, the same primary fence action is applied to C2 through TSSC link if configured.
(4) If (1) – (3) all fail, the secondary fence action is applied if enabled.
(5) If C2 is still not fenced/isolated from Grid, the manual fence action from MI can be applied.
Note1: If (1) – (3) primary fence action is applied successfully but C2 hasn't been offline for more than 20 minutes, the secondary fence action is applied if enabled.
Note2: Applying the primary remote fence action via Grid/TSSC link has a timeout. The timeout values are set to 3.5 minutes ((1) and (2)) and 15 minutes (3).

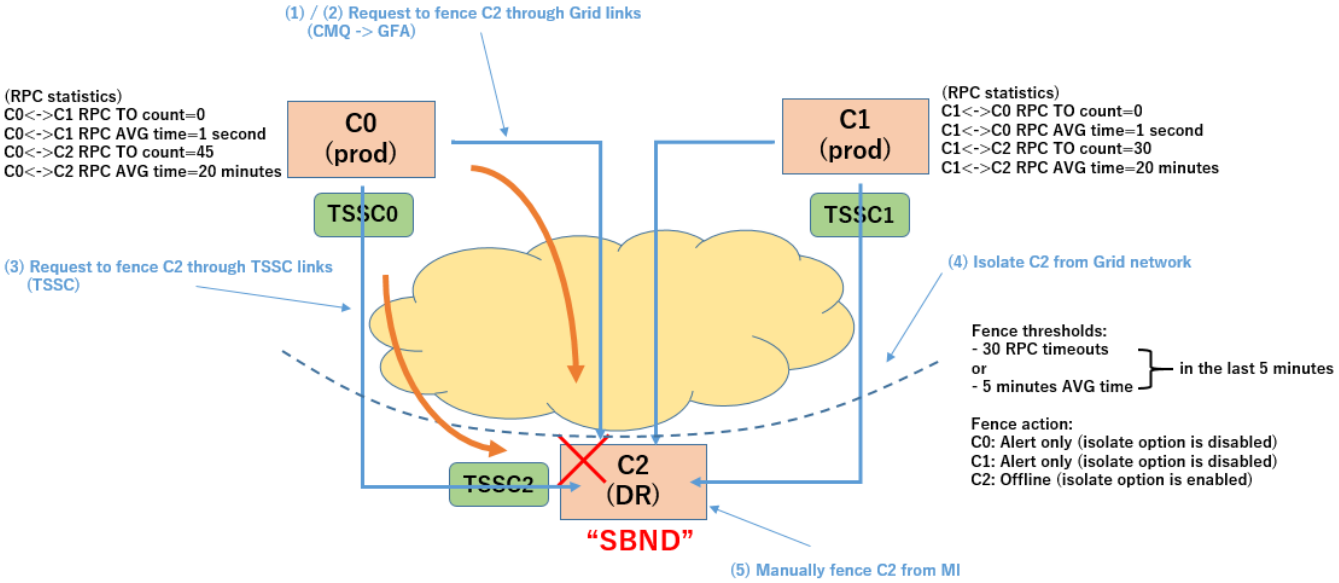

Figure 12. How the remote cluster fence action is attempted

As explained in chapter 5, here is the summary of the steps/sequences when the remote cluster fence action is applied to the remote SBND cluster:

Any healthy cluster in Grid detects SBND symptom against a remote cluster based off of Diag Data snapshots analysis (thresholds/"EVALWIN"/"CONSCNT" settings are used).
The healthy cluster starts driving the primary remote cluster fence action and asks other healthy clusters if they agree to fence the remote SBND cluster (the mechanism is slightly different between 2-way and 3-way or more Grid).
Once all other healthy clusters agree to fence the remote SBND cluster, the healthy cluster waits for "DELAY" minutes if it's configured (if any of them disagree, the internal error count is reset and the remote cluster fence action is not invoked).
After "DELAY" minutes is passed (or "DELAY" is not configured), the healthy cluster kicks the primary fence action through Grid link port 1415/1416, 350 then TSSC link.
If the primary fence action is successfully applied, the healthy cluster which issued the action waits for up to 20 minutes to see the target SBND cluster becomes offline state. If the target SBND cluster doesn't become offline after 20 minutes, the secondary fence action is initiated if enabled.
If all attempts to apply the primary fence action fail, the healthy cluster doesn't wait for 20 minutes but applies the secondary fence action if enabled

- "EVALWIN" is used to check how much of past data should be evaluated against the thresholds (moving average/total timeout/error count) for one-minute check.
- "CONSCNT" is used to determine how many consecutive "EVALWIN" checks must detect threshold condition was crossed before triggering the configured fence action.
- "DELAY" is used to configure how must of delay should occur prior to applying the remote cluster fence function after the remote cluster is determined to be fenced.

6 Manual Cluster Fence

TS7700 detects SBND symptoms on the local or against peer cluster, then the local/remote cluster fence action is applied based off of the defined/configured fence action. On the other hand, a cluster can be also locally and remotely fenced from Management Interface (MI) manually.

Manual fence is a "force" mode, then no healthy cluster agreement is required to kick the remote cluster fence action (i.e. the remote cluster fence action is forcibly applied even though the other healthy clusters don't think the target cluster is SBND). Figure 13 shows MI panels of the manual cluster fence operation:

- Actions pull down menu in Grid Summary page provides "Fence cluster" menu.
- The pop-up window provides the options:
  - Fence local cluster or remote cluster
  - If "Fence remote cluster" is selected, Cluster pull-down menu provides the remote cluster names.
  - Action pull-down menu provides the fence action:
    - ✧ Alert/Force Offline/Reboot/Reboot and stay Offline/Isolate from Grid

    Note: Isolate from Grid is the same with the secondary remote cluster fence action and it can be selected only when "Fence remote cluster" is selected (it's not available with "Fence local cluster" option).
- Manual Fence cluster operation creates a task to track the progress.

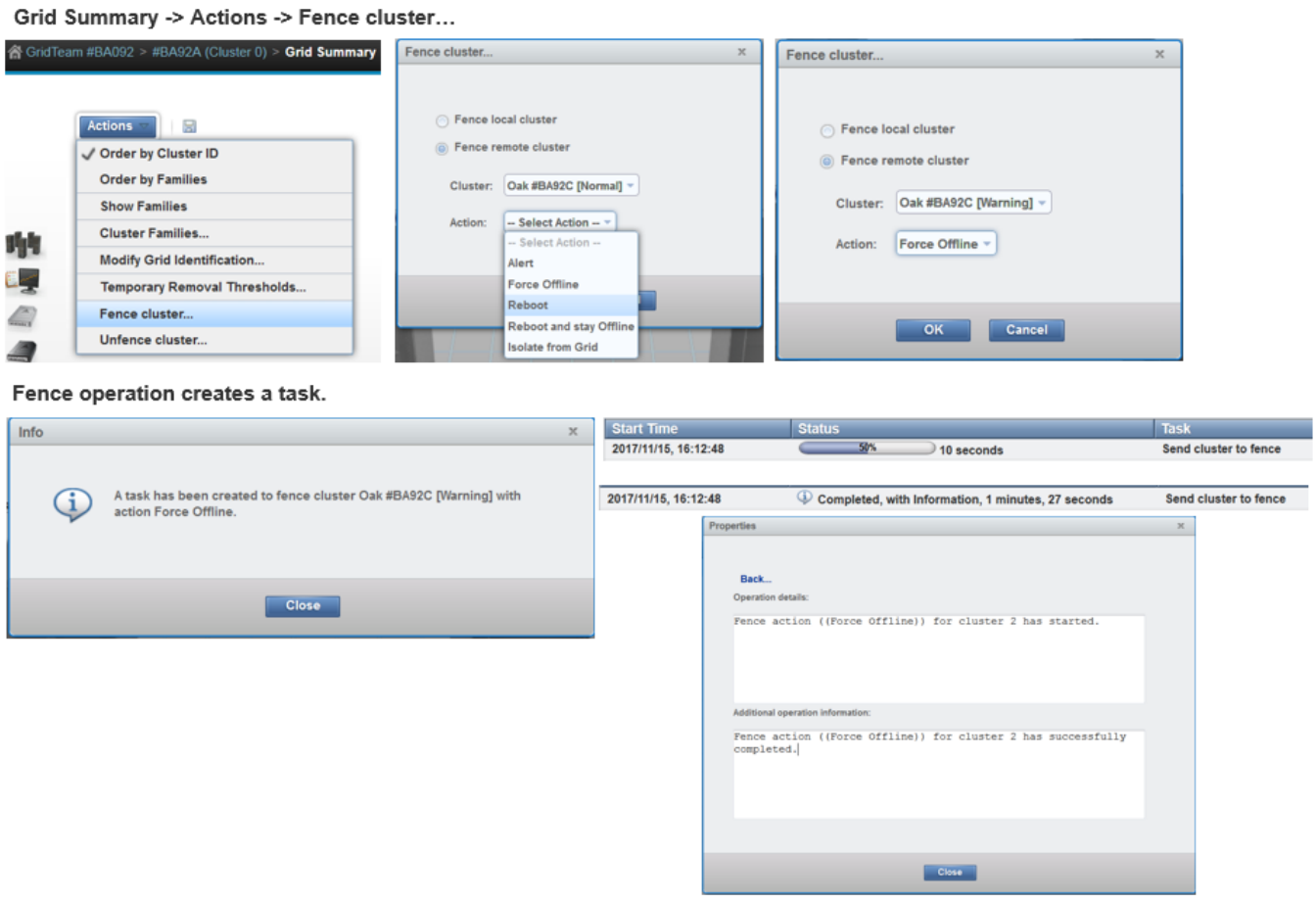

Figure 13. MI manual cluster fence menu

It is still possible to fence multiple clusters at a time manually. But it must be avoided and only one cluster should be fenced at a time and it's unfenced prior to attempting the manual fence action against a different cluster in Grid.

When a target cluster was fenced by "ALERT" fence action locally or remotely, it's possible to override the action to Force Offline/Reboot/Reboot and stay Offline/Isolate by manual remote cluster fence action.

7 Immediate Takeover against a fenced cluster

When AOTM (Autonomic Ownership Takeover Mode) is configured and it detects the remote cluster outage via TSSC link, the ownership takeover mode is triggered based off of the autonomic ownership takeover mode configuration (grace and retry period). The default grace and retry period is 25 and 5 minutes and their minimum values are 10 and 5 minutes.

When it is recognized that a remote cluster is fenced and offline, AOTM configured ownership takeover mode is initiated quickly by skipping the grace period as well as shortening the retry period:

- If the target cluster is fenced (locally or remotely), the grace period is skipped and any configured AOTM setting is initiated immediately against the cluster. The retry period is shortened to 1 minute.
- If the local or primary remote cluster fence action is applied, the AOTM handshake through the TSSC link is skipped (but it's still checked if all clusters view the target cluster as "offline").
- If a secondary remote cluster fence action is applied, then the AOTM handshake through the TSSC link is still initiated given the remote cluster could still be running.

Table 1 is the summary of grace/retry period as well as TSSC handshake.

	Skip grace period/shorten retry period	Skip TSSC handshake through TSSC link
Local cluster fence action is applied	Yes	Yes
Only primary remote cluster fence action is applied	Yes	Yes
Remote secondary cluster fence action is applied	Yes	No

Table 1. Immediate Takeover

8 Cluster Unfence Operation

This chapter explains the cluster unfence operation.

8.1 Cluster Unfence Operation from MI

The fenced cluster (locally or remotely) will either automatically or manually (through MI) exit the fenced state.

- If the fence action is "REBOOT" (the local cluster fence action is always "REBOOT") and the cluster becomes online without any error (i.e. reboot may have fixed the SBND symptom), the cluster is automatically unfenced when it becomes online. No manual cluster unfence operation is required.
- If it's "REBOFF" or "OFFLINE", the cluster will be automatically unfenced once it's brought back online by using MI unfence cluster menu.
- If it's "ALERT", the cluster will still need to be unfenced by using MI unfence cluster menu.
- If the secondary remote cluster fence action is applied, the cluster must be manually unfenced by using MI unfence cluster menu independent of its reboot or online state. This is to prevent a problematic cluster from coming in/out of visibility. Only after the issue has been resolved, the cluster can be manually unfenced through MI.

Figure 14 shows MI panels of the manual cluster unfence operation:

- Actions pull down menu in Grid Summary page provides "Unfence cluster" option. This option is provided only when a fenced cluster exists in the Grid.
- The pop-up window provides the options to select the fenced cluster to unfence.
- Manual Unfence cluster operation creates a task to track the progress.

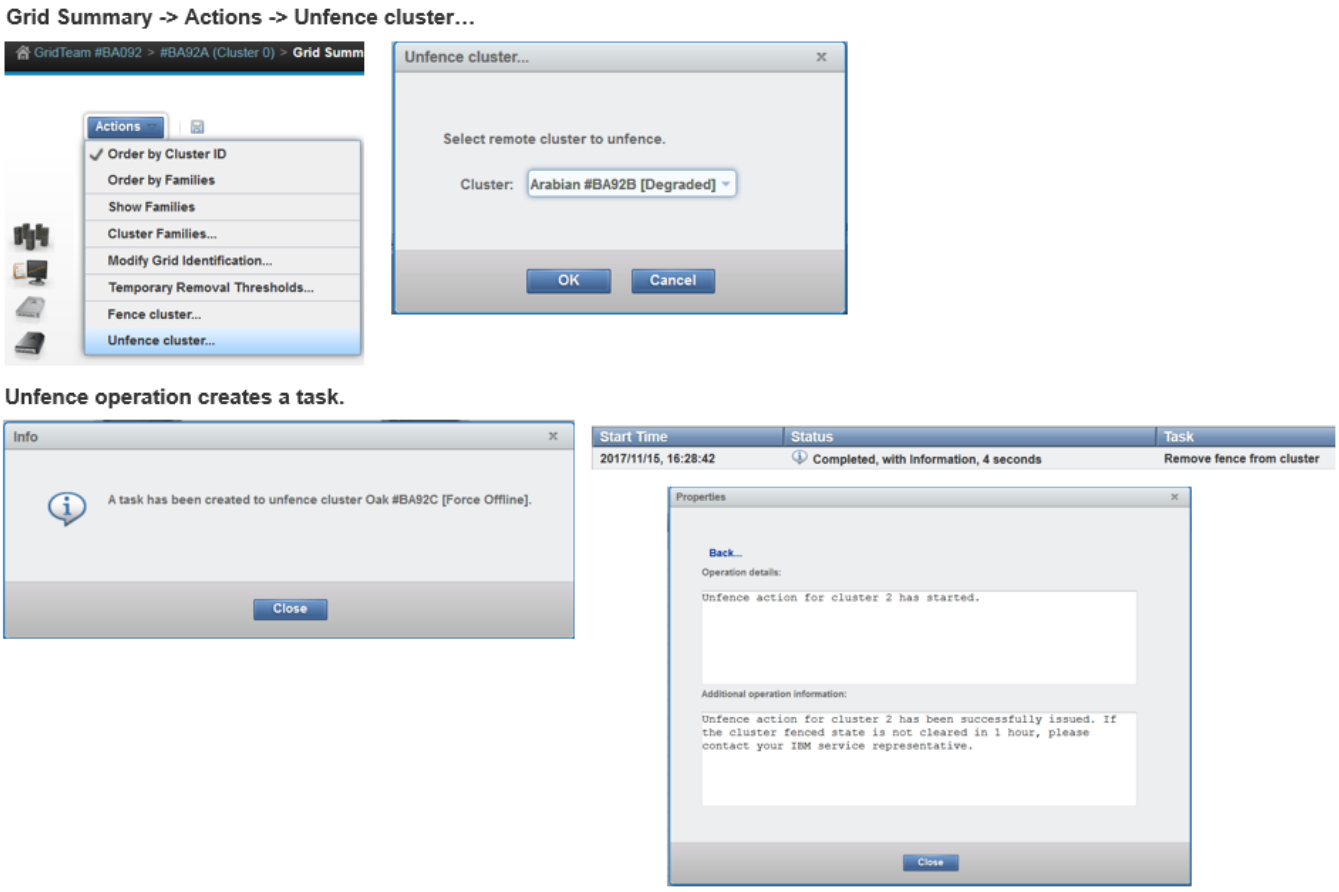

Figure 14. MI manual cluster unfence menu

The unfence cluster operation by using MI always needs to be executed from the online peer cluster. It cannot be executed from the fenced cluster (even with "ALERT" fence action).

The unfence cluster operation may return quickly although the actual unfence operation may still run in a background because it's returned once the unfence cluster operation is successfully initiated and the task is created.

After the unfence operation is successfully initiated but the target cluster hasn't become online in an hour, please check if other peer clusters are online. If they are online, please contact your IBM service representative to investigate the state of the unfence operation and target cluster.

8.2 Mixed Code or Standalone Configuration consideration

As it's mentioned in Chapter 3, the local cluster fence is enabled and applied under the mixed code as well as standalone configuration. When a fenced cluster is unfenced, the unfence operation must be always attempted from the online peer cluster. For example, if it's 2-way Grid (C0 and C1) and the local cluster fence action is applied on C1, the unfence cluster operation by using MI can be done from C0 only.

If it's a standalone (not Grid) configuration and the cluster was locally fenced in the following case, it is always required to call IBM support to unfence the cluster:

- Local cluster fence action "REBOOT" was applied but the cluster failed to come to online (i.e. reboot didn't fix SBND issue or it cannot become online due to other unexpected reasons).
- Local cluster fence action "ALERT"/"OFFLINE"/"REBOFF" was successfully applied (this can be done manually).

If it's a mixed code (Grid) configuration and there are more than 2 clusters with R4.1.2 or above exists, the locally fenced cluster can be unfenced from another online peer cluster. But if only 1 R4.1.2 cluster exists in Grid and it's fenced, it needs to be handled as same as mentioned on a standalone configuration above.

9 Customer Notification

This chapter explains what kind of the customer notification is surfaced when the cluster is fenced or unfenced.

9.1 Operator Messages

The following operator messages are surfaced when a cluster is fenced or unfenced:

- The primary remote cluster fence action has been applied:

*G0046 Library %s[*1] has crossed remote cluster fence threshold. %s[*2] has been reported Library %s[*3] has applied %s[*4] remote fence action*
**1: The target SBND cluster sequence number*
**2: The detected SBND symptom such as "TMO 20 counts continue in last 10 minutes" (TMO could be SCRVOAVG/PRIVOAVG/VCAVG/TOKAVG/TMO/ERR and 20 counts/10 minutes could vary based off of the configured thresholds). But if the manual remote cluster fence action is applied, "Manual local cluster fence has been issued" is set.*
**3: The healthy cluster sequence number which has applied this primary remote fence action*
**4: The applied remote cluster fence action. "ALERT", "OFFLINE", "REBOOT", "REBOFF" is set.*

- The primary remote cluster fence action has completed successfully:

*G0047 Library %s[*1] has successfully applied remote fence action %s[*2] to peer library %s[*3]*
**1: The healthy cluster sequence number which has applied this primary remote fence action*
**2: The applied remote cluster fence action. "ALERT", "OFFLINE", "REBOOT", "REBOFF" is set.*
**3: The target SBND cluster sequence number*

- The primary remote cluster fence action has been applied, but it failed:

*G0048 Library %s[*1] has failed to apply remote fence action %s[*2] to peer library %s[*3]*
**1: The healthy cluster sequence number which has applied this primary remote fence action*
**2: The applied remote cluster fence action. "ALERT", "OFFLINE", "REBOOT", "REBOFF" is set.*
**3: The target SBND cluster sequence number*

- The secondary remote cluster fence action has completed successfully:

*G0049 Library %s[*1] has successfully applied secondary fence action to peer library %s[*2]*
**1: The healthy cluster sequence number which has applied the secondary remote fence action*
**2: The target SBND cluster sequence number*

- The secondary remote cluster fence action has been applied, but it failed:

*G0050 Library %s[*1] has failed to apply secondary fence action to peer library %s[*2]*
**1: The healthy cluster sequence number which has applied the secondary remote fence action*
**2: The target SBND cluster sequence number*

- Cluster unfence operation has completed successfully:

*G0051 Library %s[*1] has been unfenced successfully*
**1: The cluster sequence number which has been unfenced successfully (i.e. it becomes online)*

- The local cluster fence action has been applied (on the local cluster):

*G0053 Library %s[*1] has applied %s[*2] local fence action. Reason: %s[*3]*
**1: The SBND cluster sequence number which has applied the local cluster fence action*
**2: The applied local cluster fence action. "ALERT", "OFFLINE", "REBOOT", "REBOFF" is set.*
**3: The reason of the local cluster fence action. The following sentences could be set:*

VPD daemon has been inactive
Cache access has been lost
Manual local cluster fence has been issued
Cluster will reboot soon because the cache access has been totally lost
Heartbeat timeout
CMQ issue has been detected
An application has surpassed the max allowed contexts (%d) by having %d contexts.
The database timer code has experienced excessive query timeouts.
I/O waiting too long to be written to cache filesystems.
Accessing file system has been stuck

- The local cluster fence action on the peer cluster has been detected (on the remote cluster):

*G0054 Library %s[*1] has detected remote library %s[*2] has applied %s[*3] local fence action. Reason: %s[*4]*
**1: The healthy cluster sequence number which has detected the local cluster fence action on the peer SBND cluster*
**2: The SBND cluster sequence number which has applied the local cluster fence action*

**3: The applied local cluster fence action. "ALERT", "OFFLINE", "REBOOT", "REBOFF" is set.*
**4: The reason of the local cluster fence action. The following sentences could be set:*
VPD daemon has been inactive
Cache access has been lost
Manual local cluster fence has been issued
Cluster will reboot soon because the cache access has been totally lost
Heartbeat timeout
CMQ issue has been detected
An application has surpassed the max allowed contexts (%d) by having %d contexts.
The database timer code has experienced excessive query timeouts.

- The SBND symptom against remote cluster has been detected and the primary remote cluster fence action is triggered, but other peer cluster(s) doesn't agree the fence action (this operator message is reported only once):

*G0055 Library %s[*1] has crossed remote cluster fence threshold. %s[*2] has been reported. But library %s[*3] thinks library %s[*4] looks healthy. Remote cluster fence is not applied*
**1: The target SBND cluster sequence number*
**2: The detected SBND symptom. "TMO 20 counts continue in last 10 minutes" (TMO could be SCRVOAVG/PRIVOAVG/VCAVG/TOKAVG/TMO/ERR and 20 counts/10 minutes could vary based off of the configured thresholds).*
**3: A cluster sequence number which has disagreed to fence the target SBND cluster*
**4: The target SBND cluster sequence number*

- TIME option "DELAY" has started prior to kicking the actual remote cluster fence action:

*G0056 Library %s[*1] has crossed remote cluster fence threshold. %s[*2] has been reported. Library %s[*3] will apply %s[*4] remote fence action in %d[*5] minutes*
**1: The target SBND cluster sequence number*
**2: The detected SBND symptom. "TMO 20 counts continue in last 10 minutes" (TMO could be SCRVOAVG/PRIVOAVG/VCAVG/TOKAVG/TMO/ERR and 20 counts/10 minutes could vary based off of the configured thresholds).*
**3: The healthy cluster sequence number which has applied this primary remote fence action*
**4: The applied remote cluster fence action. "OFFLINE", "REBOOT", "REBOFF" is set.*
**5: The configured "DELAY" in minute*

Note: "*G0052 Unfence library %s has failed*" is defined and it's viewed from MI Notification Settings menu, but it's not used in R4.1.2 code level.

9.2 Management Interface

This section explains the customer notification on MI.

9.2.1 Grid Summary/Fence Mode Page

When a cluster is locally or remotely fenced, Grid Summary page shows the cluster fenced state.

From a healthy (online) cluster, a fenced cluster state can be seen in Grid Summary page. Figure 15 shows Grid Summary page example when a remote cluster is fenced.

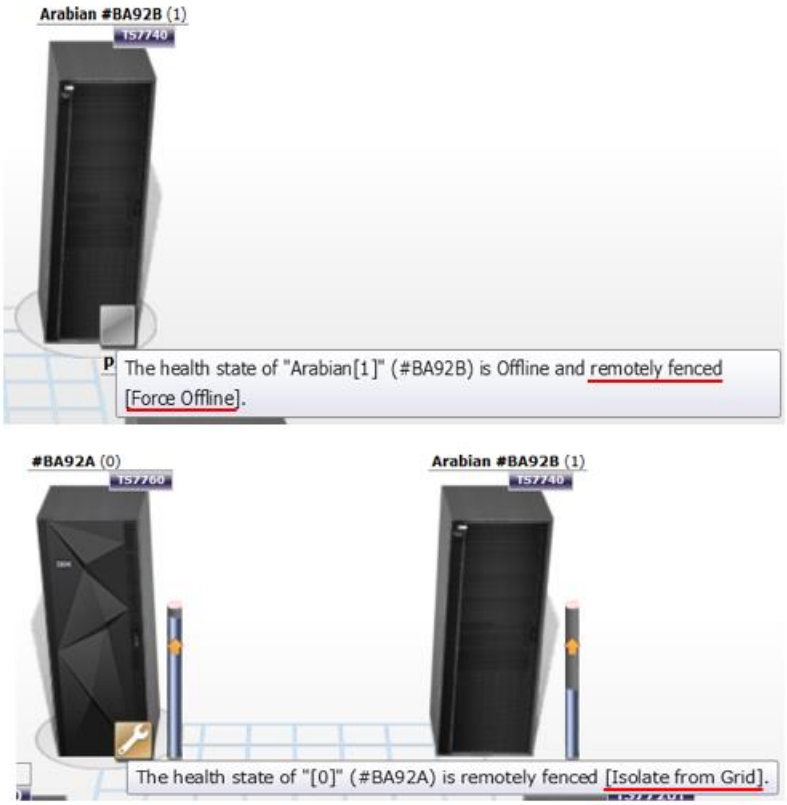

Figure 15. Grid Summary page example
(the remote cluster is fenced by the primary "OFFLINE" (top) and secondary fence action (bottom))

On the other hand, even though the cluster is still online, MI on a fenced cluster always shows "Fence Mode" page and normal MI page doesn't show up. This can happen when the fence action "ALERT" or secondary fence action is applied but the target SBND cluster is still online. When a fenced cluster is offline, MI shows "Cluster Nodes Online/Offline" page with the description that a fence action is applied to the local cluster. Figure 16 shows Cluster Nodes Online/Offline and Fence Mode page examples.

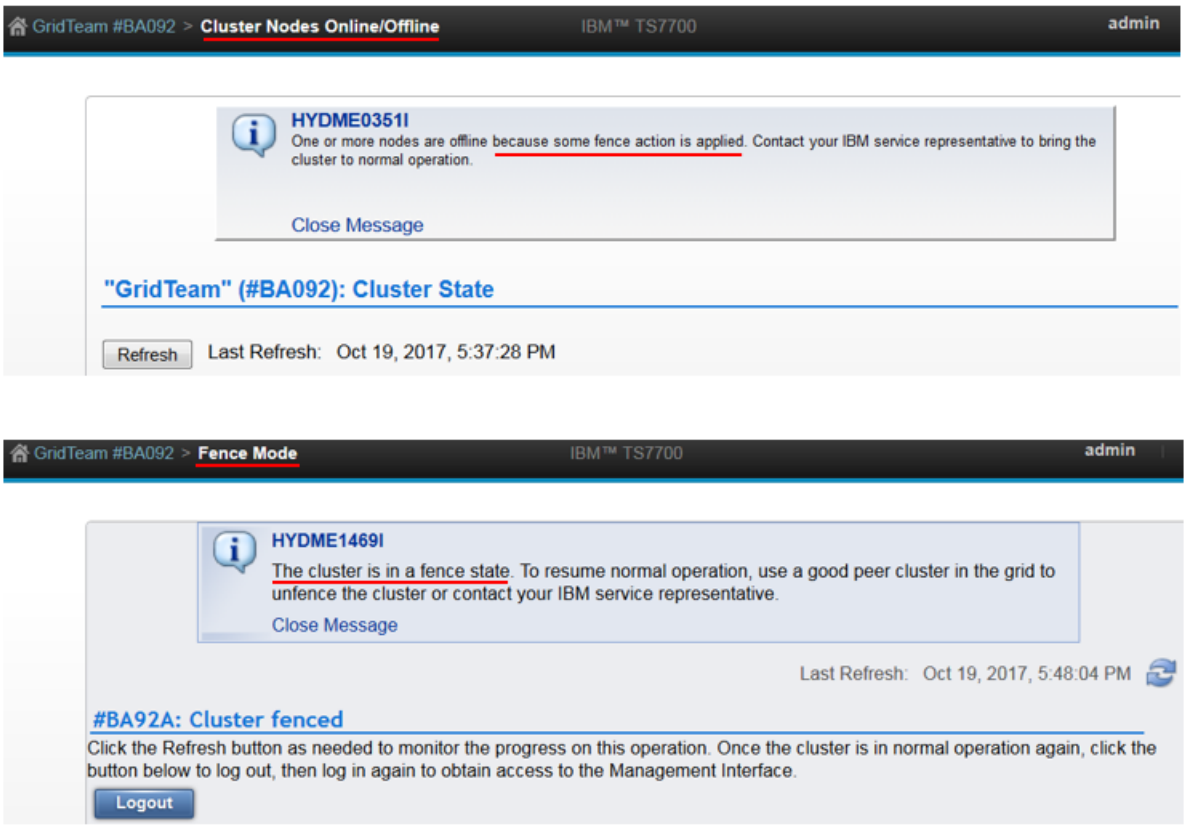

Figure 16. MI page example on a fenced cluster
(the cluster is fenced and offline (top) and online (bottom))

9.2.2 Events

The following events are created when a cluster is fenced successfully and it becomes inactive when the cluster is unfenced successfully:

- The local cluster fence action has been applied (on the local cluster):

*The cluster %d[*1] has applied %s[*2] local cluster fence action. Reason: %s[*3]*
**1: The SBND cluster ID which has applied the local cluster fence action*
**2: The applied local cluster fence action. "Alert", "Force Offline", "Reboot", "Reboot and stay offline" is set.*
**3: The reason of the local cluster fence action. The following sentences could be set:*

VPD daemon has been inactive
Cache access has been lost
Manual local cluster fence has been issued
Cluster will reboot soon because the cache access has been totally lost
Heartbeat timeout
CMQ issue has been detected
An application has surpassed the max allowed contexts (%d) by having %d contexts.
The database timer code has experienced excessive query timeouts.
I/O waiting too long to be written to cache filesystems
Accessing file system has been stuck

- The local cluster fence action on the peer cluster has been detected (on the remote cluster):

*The cluster %d[*1] has detected the remote cluster %d[*2] has applied %s[*3] local fence action. Reason: %s[*4]*
**1: The healthy cluster ID which has detected the local cluster fence action on the peer SBND cluster*
**2: The SBND cluster ID which has applied the local cluster fence action*
**3: The applied local cluster fence action. "Alert", "Force Offline", "Reboot", "Reboot and stay offline" is set.*
**4: The reason of the local cluster fence action*

- The remote cluster fence action has completed successfully:

*The cluster %d[*1] has successfully completed the remote fence action %s[*2] to the peer cluster %d[*3]. Reason: %s[*4]*
**1: The healthy cluster ID which has applied this remote fence action*
**2: The applied remote cluster fence action. "Alert", "Force Offline", "Reboot", "Reboot and stay offline", "Isolate from Grid" is set.*
**3: The target SBND cluster ID*
**4: The reason of the remote cluster fence action such as "TMO 20 counts continue in last 10 minutes" (TMO could be SCRVOAVG/PRIVOAVG/VCAVG/TOKAVG/TMO/ERR and 20 counts/10 minutes could vary based off of the configured thresholds). But if the manual remote cluster fence action is applied, "Manual local cluster fence has been issued" is set.*

- When a fenced cluster is unfenced successfully, the created active events which were created when the cluster was fenced are automatically marked inactive, and Event History shows the unfenced timestamp with the following sentence:

*The cluster %d[*1] has been unfenced successfully.*
**1: The cluster ID which has been unfenced successfully*

Figure 17 shows an example of an active event when a cluster is fenced:

Figure 17. Active event example (remote cluster fence action "OFFLINE" is applied)

9.2.3 Tasks

The following tasks are created when a cluster is fenced or unfenced manually:

When a manual local or remote cluster fence action is applied:

Task name: Send cluster to fence
*Operation details: Fence action (%s[*1]) for cluster %d[*2] has started.*
**1: The applied cluster fence action. "Alert", "Force Offline", "Reboot", "Reboot and stay offline", "Isolate from Grid" is set.*
**2: The target SBND cluster ID*

Then, If the fence action successfully completes, the task is updated:
*Additional operation information: Fence action (%s[*1]) for cluster %d[*2] has successfully completed.*
**1: The applied cluster fence action*
**2: The target SBND cluster ID*

If the fence action fails, the task is updated:
*Additional operation information: Fence action (%s[*1]) for cluster %d[*2] has failed due to %s[*3]. Contact your IBM service representative.*
**1: The applied cluster fence action*
**2: The target SBND cluster ID*
**3: In R4.1.2 "Unexpected reasons" is always set.*

When a manual cluster unfence action is applied:

Task name: Remove fence from cluster
*Operation details: Unfence action for cluster %d[*1] has started.*
**1: The target cluster ID to unfence*

Then, If the unfence action successfully issued (started), the task is updated:
*Additional operation information: Unfence action for cluster %d[*1] has been successfully issued. If the cluster fenced state is not cleared in 1 hour, please contact your IBM service representative.*
**1 The target cluster ID to unfence*

If the unfence action fails to start, the task is updated:
*Additional operation information: Unfence action for cluster %d[*1] has failed due to %s[*2]. Contact your IBM service representative.*
**1: The target cluster ID to unfence*
**2: In R4.1.2 "Unexpected reasons" is always set.*

Figure 18 shows an example of a task when a cluster is manually fenced:

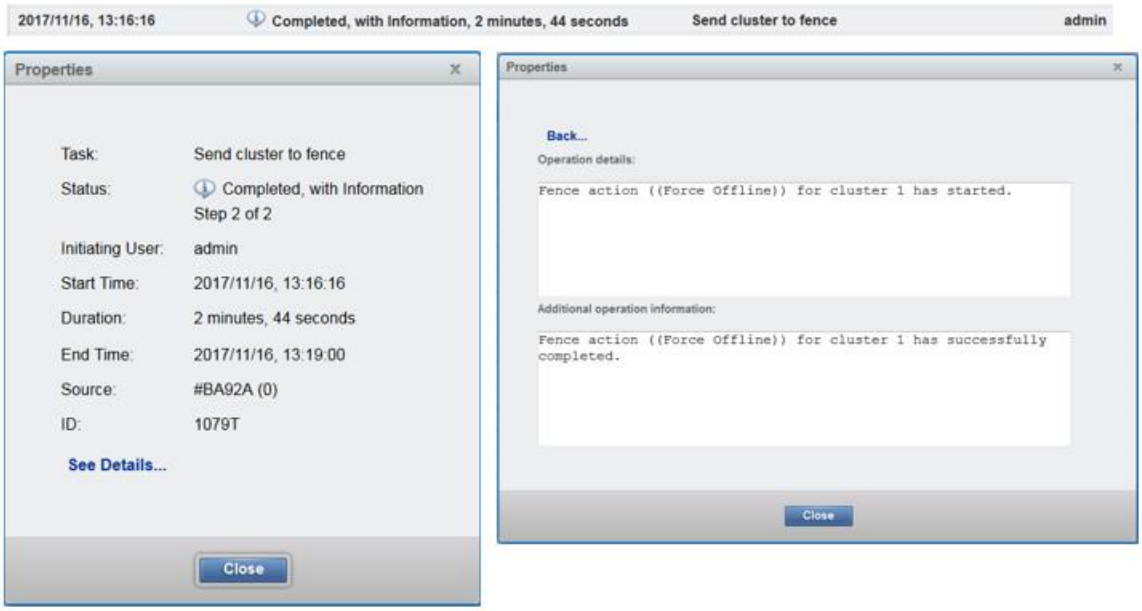

Figure 18. Task example (remote cluster fence action "OFFLINE" is manually applied)

9.3 LI REQ

New supported LI REQ (the keywords are FENCE, SHOW) provides the cluster fence state in Grid. Figure 18 shows LI REQ output example. The format and the fields are explained in Chapter 10.

```
CBR1020I PROCESSING LIBRARY COMMAND: REQ,BARR92,FENCE,SHOW.
CBR1280I LIBRARY BARR92 REQUEST. 143
KEYWORDS: FENCE,SHOW
------------------------------------------------------------------------
FENCE REQUEST V1 .0
 COMPOSITE LIBRARY FENCE SETTINGS
  REMOTE FENCE FUNCTION: ENABLE
  REMOTE FENCE THRESHOLD:
    SCRVOAVG: 10    PRIVOAVG: 10    (SECONDS)
    VCAVG   : 180   TOKAVG  : 180   (SECONDS)
    TMO     : 20    ERR     : 20    (COUNTS)
    EVALWIN : 7  (MINUTES)
  REMOTE FENCE TIME:
    DELAY   : 0     CONSCNT : 0     (MINUTES)
------------------------------------------------------------------------
 DISTRIBUTED LIBRARY FENCE ACTION SETTINGS
  BA92A (CL0) PRI:    ALERT  SEC:  DISABLE  AIXDUMP:  DISABLE
  BA92B (CL1) PRI:  OFFLINE  SEC:  DISABLE  AIXDUMP:  DISABLE
  BA92C (CL2) PRI:   REBOOT  SEC:   ENABLE  AIXDUMP:   ENABLE
  BA92D (CL7) PRI:   REBOFF  SEC:  DISABLE  AIXDUMP:   ENABLE
------------------------------------------------------------------------
 DISTRIBUTED LIBRARY FENCE STATE
  BA92A (CL0)  LOC  ACT :     NONE                    RSN: 0
               REM  PACT:     NONE  SACT:     NONE    RSN: 0
              LAST FENCED TIME: 2017-09-29 04:32:03
  BA92B (CL1)  LOC  ACT :     NONE                    RSN: 0
               REM  PACT:  OFFLINE  SACT:     NONE    RSN: 101
              LAST FENCED TIME: 2017-10-18 05:40:51
  BA92C (CL2)  LOC  ACT :     NONE                    RSN: 0
               REM  PACT:     NONE  SACT:     NONE    RSN: 0
              LAST FENCED TIME: 2017-10-18 08:44:00
  BA92D (CL7)  LOC  ACT :     NONE                    RSN: 0
               REM  PACT:     NONE  SACT:     NONE    RSN: 0
              LAST FENCED TIME: 2017-09-25 22:18:40
```

Figure 18. LI REQ, FENCE, SHOW output example

9.4 Call Home/Service Information Message (SIM)

A call home is surfaced to IBM support and the corresponding SIM alert is created when a cluster is locally and remotely fenced. The dump data is automatically taken when a cluster is fenced from the rest of healthy clusters in Grid, also when a fenced cluster is successfully unfenced, the cluster takes dump data automatically so that the necessary logs for the root cause analysis may not be overwritten.

10 Library Request Commands

This chapter provides all supported LI REQ commands and their expected outputs of GR functions.

Similar to other LI REQ commands, a host based library name must be provided when issuing the commands from the host. All LI REQ commands related to GR functions are supported across all TS7700 models so long as they are running the appropriate microcode level.

Table 2 provides all the supported keywords for LI REQ FENCE.

KW1	KW2	KW3	KW4	Description	Comp	Dist
FENCE	ENABLE/ DISABLE			Enable/disable remote cluster fence function.	Y	N
	THRESHLD	SCRVOAVG/ PRIVOAVG/ VCAVG/ TOKAVG/ TMO/ ERR/ EVALWIN	<value>	Set the thresholds to determine when the remote cluster fence action is triggered.	Y	N
	TIME	DELAY/ CONSCNT	<value>	Adjust the timing requirements used to determine when a remote cluster fence action is applied.	Y	N
	SHOW			Request information about the remote cluster fence function.	Y	N
	ACTION	PRI	NONE/ ALERT/ OFFLINE/ REBOOT/ REBOFF	Configure what primary action takes place against a remote cluster when a fence action is applied.	N	Y
		SEC	ENABLE/ DISABLE	Configure whether the secondary fence action (grid network isolation) is initiated against a remote cluster when the primary action is unsuccessful.	N	Y
		AIXDUMP	ENABLE/ DISABLE	Configure whether an AIX system dump is automatically initiated as part of the fencing action when REBOOT/REBOFF is initiated.	N	Y

Table 2. Supported LI REQ commands of GR functions

- Changing the options of GR remote cluster fence function can be done through LI REQ only. No MI panel to change the options is supported.
- Fencing/Unfencing a cluster can be done through MI only and no LI REQ to fence and unfence a cluster is supported.
- All LI REQ commands of GR functions are accepted only when all clusters in Grid are at R4.1.2 or above except 2 keywords below:
  - 2nd keyword = 'SHOW' (required to show the local cluster fence state)
  - 2nd keyword = 'ACTION' AND 3rd keyword = 'AIXDUMP' (required to set AIXDUMP option for the local cluster fence function)
- The 1st keyword is always "FENCE".

10.1 LI REQ, <comp lib>, FENCE, {ENABLE|DISABLE}

This LI REQ is used to enable and disable GR remote cluster fence function for the entire Grid. The command target is the composite library.

- Second keyword: "ENABLE" or "DISABLE"
- To use the remote cluster fence function, it must be enabled first. The default is disable. The local cluster fence function is always enabled.
- Other options can be changed even if the function is disabled.
- The settings of all options are preserved even after the function is disabled (they're persistent over offline/online (reboot)).
- Manual cluster fence operation can be done even if the function is disabled.

When the request successfully completes, the following text is returned:

```
FENCE REQUEST V1 .0
OPERATION SUCCESSFULLY COMPLETED
```

Line	Bytes	Name	Description
1	0:14	Header Info	'FENCE REQUEST V'
	15:16	Version	The version number for the response. The number if left justified and padded with blanks. Starts with 1.
	17	Dot	'.'
	18:19	Revision	The revision number for the response. The number is left justified and padded with blanks. Starts with 0.
	20:69	Blanks	
2	0:31	Success Info	'OPERATION SUCCESSFULLY COMPLETED'
	32:69	Blanks	

10.2 LI REQ, <comp lib>, FENCE, THRESHLD, {SCRVOAVG|PRIVOAVG|VCAVG|TOKAVG|TMO|ERR|EVALWIN}, <value>

This LI REQ is used to change the thresholds of the remote cluster fence function to determine when the remote cluster fence action needs to be triggered. The command target is the composite library.

- Second keyword: "THRESHLD"

- Third keyword: "SCRVOAVG", "PRIVOAVG", "VCAVG", "TOKAVG", "TMO", "ERR" and "EVALWIN" can be used to set the threshold.
- Fourth keyword: The value of the threshold specified in the third keyword.
  - SCRVOAVG: The average elapsed time of the volume open/mount request peer handshakes for scratch mounts
    Default = 180 seconds. 0 – 1200 can be set (0 means no check)
  - PRIVOAVG: The average elapsed time of the volume open/mount request peer handshakes for private mounts
    Default = 180 seconds. 0 – 1200 can be set (0 means no check)
  - VCAVG: The average elapsed time of the volume close request peer handshakes (all mount types)
    Default = 180 seconds. 0 – 1200 can be set (0 means no check)
  - TOKAVG: The average elapsed time of token request handshakes (with the timeout value less than 3 minutes)
    Default = 120 seconds. 0 – 180 can be set (0 means no check)
  - TMO: The timeout count of all major peer handshakes (all mounts/TKM others)
    Default = 20 count. 0 – 1000 can be set (0 means no check)
  - ERR: The error count of all major peer handshakes (all mounts/TKM others)
    Default = 20 count. 0 – 1000 can be set (0 means no check)
  - EVALWIN: The diag data evaluation window
    Default = 7 minutes (1 – 30 can be set)

When the request successfully completes, the following text is returned:

```
FENCE REQUEST V1 .0
OPERATION SUCCESSFULLY COMPLETED
```

10.3 LI REQ, <comp lib>, FENCE, TIME, {DELAY|CONSCNT}, <value>

This LI REQ is used to change the time options of the remote cluster fence function to determine how the remote cluster fence action is applied. The command target is the composite library.

- Second keyword: "TIME"
- Third keyword: "DELAY" and "CONSCNT" can be used to set the time option.
- Fourth keyword: The value of the time option specified in the third keyword.
  - DELAY: The delay in minute to execute the actual remote cluster fence action since the target remote cluster is determined as SBND.
    Default = 0 minute. 0 – 60 can be set (0 means no delay).
  - CONSCNT: : The consecutive fence check condition count prior to kicking the actual remote cluster fence action. The check is done once a minute.
    Default = 10 minutes. 0 – 60 can be set (when it's set to 0, it's treated as same as 1 minute).

When the request successfully completes, the following text is returned:

```
FENCE REQUEST V1 .0
OPERATION SUCCESSFULLY COMPLETED
```

10.4 LI REQ, <dist lib>, FENCE, ACTION, PRI, {NONE|ALERT|OFFLINE|REBOOT|REBOFF}

This LI REQ is used to change the primary remote cluster fence action per each cluster. The command target is the distributed library.

- Second keyword: "ACTION"
- Third keyword: "PRI"
- Fourth keyword: "NONE", "ALERT", "OFFLINE, "REBOOT" and "REBOFF" can be used to set the fence action
  - NONE: No remote cluster fence action is applied (default).
  - ALERT: Only alert (operator message/event/call home etc.) is applied. No attempt to fence the cluster is attempted. But it's still required to unfence the fenced cluster with ALERT action.
  - OFFLINE: The cluster becomes offline (force offline) once this fence action is successfully applied.
  - REBOOT: The cluster reboots once this fence action is successfully applied. If the cluster becomes online without any error, it's unfenced automatically.
  - REBOFF: The cluster reboots and stays at offline state once this fence action is successfully applied.
- When "NONE" or "ALERT" is set when the secondary remote cluster fence action is already enabled against the target distributed library, the secondary remote cluster fence action is disabled automatically.

When the request successfully completes, the following text is returned:

```
FENCE REQUEST V1 .0
OPERATION SUCCESSFULLY COMPLETED
```

10.5 LI REQ, <dist lib>, FENCE, ACTION, SEC, {ENABLE|DISABLE}

This LI REQ is used to enable and disable the secondary remote cluster fence action per each cluster. The command target is the distributed library.

- Second keyword: "ACTION"
- Third keyword: "SEC"
- Fourth keyword: "ENABLE" or "DISABLE"
- The default is disable.

When the request successfully completes, the following text is returned:

```
FENCE REQUEST V1 .0
OPERATION SUCCESSFULLY COMPLETED
```

10.6 LI REQ, <dist lib>, FENCE, ACTION, AIXDUMP, {ENABLE|DISABLE}

This LI REQ is used to enable and disable the option to AIX dump when the cluster is fenced by "REBOOT" or "REBOFF" fence action per each cluster. The command target is the distributed library.

- Second keyword: "ACTION"
- Third keyword: "AIXDUMP"
- Fourth keyword: "ENABLE" or "DISABLE"

- The default is disable.

When the request successfully completes, the following text is returned:

```
FENCE REQUEST V1 .0
OPERATION SUCCESSFULLY COMPLETED
```

10.7 LI REQ, <comp lib>, FENCE, SHOW

This LI REQ provides the settings of the remote cluster fence function and cluster fence state. The output contains the thresholds, time options, fence actions and current fence state for each cluster. The command target is the composite library.

- Second keyword: "SHOW"
- Third keyword: "AIXDUMP"
- Fourth keyword: "ENABLE" or "DISABLE"
- The default is disable.

Figure 19 shows an example of the output when the request successfully completes:

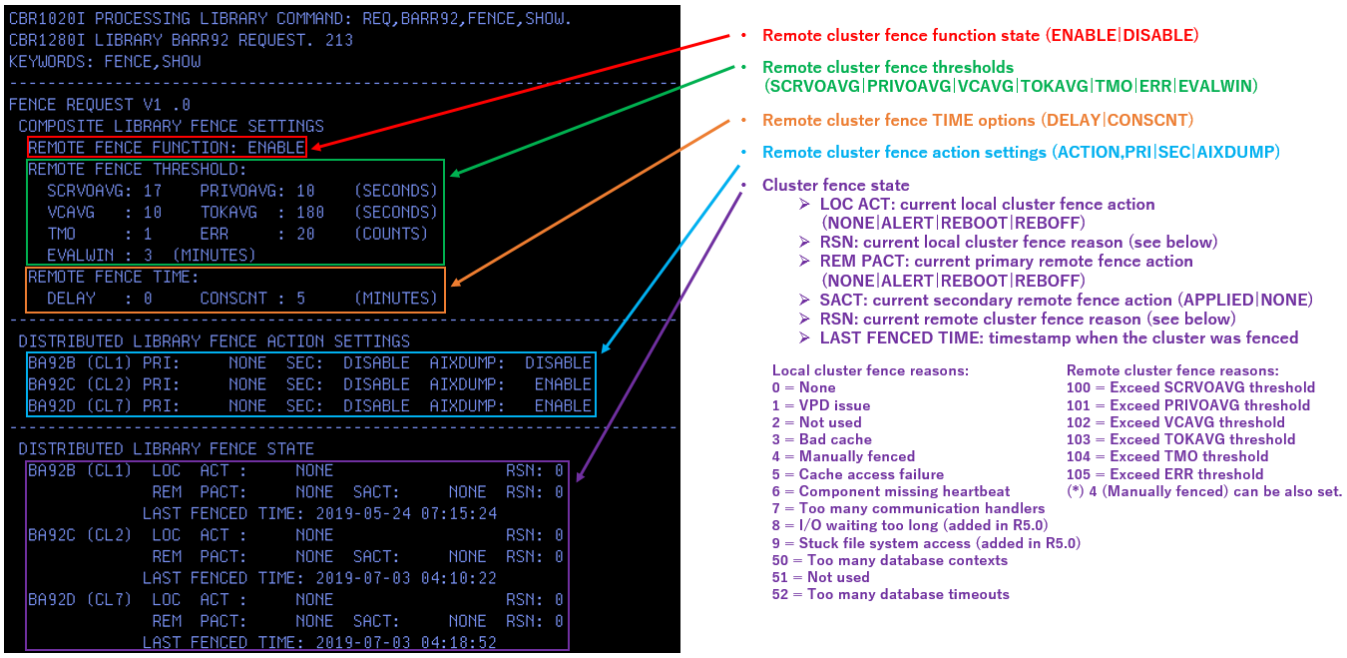

Figure 19. Example of FENCE, SHOW output and the field explanation

Line	Bytes	Name	Description
1	0:14	Header Info	'FENCE REQUEST V'
	15:16	Version	The version number for the response. The number if left justified and padded with blanks. Starts with 1.
	17	Dot	'.'
	18:19	Revision	The revision number for the response. The number is left justified and

Line	Bytes	Name	Description
			padded with blanks. Starts with 0.
	20:69	Blanks	
2	0:32	Header Info	' COMPOSITE LIBRARY FENCE SETTINGS'
	33:69	Blanks	
3	0:24	Header Info	' REMOTE FENCE FUNCTION: '
	25:31	The current remote cluster fence function status	This is the current remote cluster fence function status 'DISABLE': Remote cluster fence function is disabled. 'ENABLE ': Remote cluster fence function is enabled.
	32	Blank	
	33:69	Mixed code configuration status, or blanks	If this Grid is a mixed code configuration, the following statement is provided, otherwise all blanks: '(ONLY LOCAL FENCE FUNCTION SUPPORTED)'
4	0:24	Header Info	' REMOTE FENCE THRESHOLD:'
	25:69	Blanks	
5	0:13	Header Info	' SCRVOAVG: '
	14:17	SCRVOAVG threshold value	This is the current SCRVOAVG threshold value (the average elapsed time of the volume open/mount request peer handshakes for scratch mounts). The value is left justified and padded with blanks.
	18:19	Blanks	
	20:29	Header Info	'PRIVOAVG: '
	30:33	PRIVOAVG threshold value	This is the current PRIVOAVG threshold value (the average elapsed time of the volume open/mount request peer handshakes for private mounts). The value is left justified and padded with blanks.
	34:44	Header Info	' (SECONDS)'
	45:69	Blanks	
6	0:13	Header Info	' VCAVG : '
	14:17	VCAVG threshold value	This is the current VCAVG threshold value (the average elapsed time of the volume close/demount (Rewind Unload) request peer handshakes (all mount types)). The value is left justified and padded with blanks.
	18:19	Blanks	
	20:29	Header Info	'TOKAVG : '
	30:33	TOKAVG threshold value	This is the current TOKAVG threshold value (the average elapsed time of token request handshakes (with the timeout value less than 3 minutes)). The value is left justified and padded with blanks.
	34:44	Header Info	' (SECONDS)'
	45:69	Blanks	
7	0:13	Header Info	' TMO : '
	14:17	TMO threshold value	This is the current TMO threshold value (the timeout count of all major peer handshakes (all mounts/demounts/tokens/others)). The value is left justified and padded with blanks.
	18:19	Blanks	

Line	Bytes	Name	Description
	20:29	Header Info	'ERR : '
	30:33	ERR threshold value	This is the current ERR threshold value (the error count of all major peer handshakes (all mounts/demounts/tokens/others)). The value is left justified and padded with blanks.
	34:43	Header Info	' (COUNTS)'
	44:69	Blanks	
8	0:13	Header Info	' EVALWIN : '
	14:15	EVALWIN threshold value	This is the current EVALWIN threshold value (the diag data evaluation window). The value is left justified and padded with blanks.
	26:25	Header Info	' (MINUTES)'
	26:69	Blanks	
9	0:19	Header Info	' REMOTE FENCE TIME:'
	20:69	Blanks	
10	0:13	Header Info	' DELAY : '
	14:15	TIME DELAY option value	This is the current TIME DELAY option value (the delay in minute to execute the actual remote cluster fence since the target remote cluster is determined as SBND). The value is left justified and padded with blanks.
	16:19	Blanks	
	20:29	Header Info	'CONSCNT : '
	30:31	TIME CONSCNT option value	This is the current TIME CONSCNT option value (the consecutive fence check condition count (minute) prior to kicking the actual remote cluster fence action). The value is left justified and padded with blanks.
	32:44	Header Info	' (MINUTES)'
	45:69	Blanks	
11	0:69	Separator	All dash '-' characters
12	0:41	Header Info	' DISTRIBUTED LIBRARY FENCE ACTION SETTINGS'
	42:69	Blanks	
For each distributed library in the grid configuration a line is formatted as follows (each line includes the primary/secondary remote cluster fence actions as well as AIX dump option and it continues from the line 13 to 13 + (N – 1) (N-way Grid)):			
13	0:1	Blanks	
	2:6	Dist Lib Sequence Number	The distributed library sequence number.
	7:10	Header Info	' (CL'
	11	Cluster ID	The cluster ID of this distributed library (0 – 7).
	12:18	Header Info	') PRI: '
	19:26	PRI remote cluster fence action	This is the current primary remote cluster fence action defined to this cluster: ' NONE' No primary fence action is applied. ' ALERT' "ALERT" is applied.

Line	Bytes	Name	Description
			' OFFLINE' "OFFLINE" is applied. ' REBOOT' "REBOOT" is applied. ' REBOFF' "REBOFF" is applied. ' UNKNOWN' Unknown primary fence action is applied.
	27:33	Header Info	' SEC:'
	34:41	SEC remote cluster fence action	This is the current secondary remote cluster fence action defined to this cluster: ' DISABLE' Secondary remote cluster fence action is disabled. ' ENABLE' Secondary remote cluster fence action is enabled.
	42:52	Header Info	' AIXDUMP: '
	53:60	AIX dump option	This is the current AIX dump option defined to this cluster which determines to taken an AIX dump automatically when the primary fence action "REBOOT"/"REBOFF" or the local cluster fence action is applied: ' DISABLE' AIX dump is not taken automatically. ' ENABLE' AIX dump is taken automatically.
	61:69	Blanks	
M	0:69	Separator	All dash '-' characters (M = 13 +N) (N-way Grid)
M+1	0:31	Header Info	' DISTRIBUTED LIBRARY FENCE STATE'
	32:69	Blanks	
For each distributed library in the grid configuration a line is formatted as follows (each line includes the active (applied) local and primary/secondary remote cluster fence actions and reasons as well as the timestamp when the cluster has been fenced and it continues from the line (M+2) to (M+2) + (N – 1) (N-way Grid)):			
M+2	0:1	Blanks	
	2:6	Dist Lib Sequence Number	The distributed library sequence number.
	7:10	Header Info	' (CL'
	11	Cluster ID	The cluster ID of this distributed library (0 – 7).
	12:25	Header Info	') [LOC] ACT : '
	26:33	Active local cluster fence action	This is the current active (applied) local cluster fence action on this cluster: ' NONE' No primary fence action is applied. ' ALERT' "ALERT" is applied. ' OFFLINE' "OFFLINE" is applied. ' REBOOT' "REBOOT" is applied. ' REBOFF' "REBOFF" is applied. ' UNKNOWN' Unknown primary fence action is applied.
	34:56	Header Info	' RSN: '
	57:59	Active Local Cluster Fence	This is the current active (applied) local cluster fence reason on this

Line	Bytes	Name	Description
		Reason	cluster: 0 = No active local cluster fence action is applied. 1 = VPD issue 2 = Not used in R4.1.2 3 = Bad cache 4 = Manually fenced 5 = Cache access failure 6 = Component missing heartbeat 7 = Too many communication handlers 8 = I/O waiting too long (added in R5.0) 9 = Stuck file system access (added in R5.0) 50 = Too many database contexts 51 = Not used in R4.1.2 52 = Too many database timeouts
	60:69	Blanks	
M+3	0:25	Header Info	' [REM] PACT: '
	11	Cluster ID	The cluster ID of this distributed library (0 – 7).
	12:25	Header Info	') [LOC] ACT : '
	26:33	Active Primary Remote Cluster Fence Action	This is the current active (applied) primary remote cluster fence action against this cluster: ' NONE' No primary fence action is applied. ' ALERT' "ALERT" is applied. ' OFFLINE' "OFFLINE" is applied. ' REBOOT' "REBOOT" is applied. ' REBOFF' "REBOFF" is applied. ' UNKNOWN' Unknown primary fence action is applied.
	34:41	Header Info	' SACT: '
	42:49	Active Secondary Cluster Fence Action	This is the current active (applied) secondary remote cluster fence action against this cluster: ' NONE' No secondary fence action is applied. ' APPLIED' Secondary fence action is applied.
	50:56	Header Info	' RSN: '
	57:59	Active Remote Cluster Fence Reason	This is the current active (applied) remote cluster fence reason against this cluster: 100 = Exceed SCRVOAVG threshold 101 = Exceed PRIVOAVG threshold 102 = Exceed VCAVG threshold 103 = Exceed TOKAVG threshold 104 = Exceed TMO threshold

Line	Bytes	Name	Description
			105 = Exceed ERR threshold 4 = Manually fenced
	60:69	Blanks	
M+4	0:31	Header Info	` LAST FENCED TIME: '
	32:50	Last Fenced Time	This is the timestamp when this cluster was fenced locally or remotely. This timestamp is only updated when the cluster is successfully fenced. It's not updated (reset) when the fenced cluster is unfenced. The timestamp is formatted in UTC as follows: YYYY-MM-DD HH:MM:SS
	51:69	Blanks	

10.8 LI REQ Error Text

If LI REQ, FENCE commands which needs to be issued to a composite library was issued to a distributed library, the request fails with the error text:

```
"REQUEST INVALID FOR DISTRIBUTED LIBRARY"
```

If LI REQ, FENCE commands which needs to be issued to a distributed library was issued to a composite library, the request fails with the error text:

```
"REQUEST INVALID FOR COMPOSITE LIBRARY"
```

If any LI REQ other than LI REQ, FENCE, SHOW and ACTION, AIXDUMP is issued to a Grid which is a mixed code configuration, the request fails with the error text:

```
"MINIMUM CODE LEVEL FOR DOMAIN IS NOT MET"
```

If unsupported second/third/fourth keyword is specified, the request fails with the error text and the supported keywords are provided:

```
"INVALID 2ND/3RD/4TH KEYWORD"
" FENCE,[ENABLE|DISABLE]"
" FENCE,ACTION,PRI,[NONE|ALERT|OFFLINE|REBOOT|REBOFF]"
" FENCE,ACTION,SEC,[ENABLE|DISABLE]"
" FENCE,ACTION,AIXDUMP,[ENABLE|DISABLE]"
" FENCE,THRESHLD,[TMO|ERR|SCRVOAVG|PRIVOAVG|VCAVG|TOKAVG|EVALWIN],<Val>"
" FENCE,TIME,[DELAY|CONSCNT],<Val>"
" FENCE,SHOW"
```

If a value (XXXX) which is out of range is specified for LI REQ, FENCE, THRESHOLD/TIME commands, the request fails with the error text:

```
"THE VALUE OF 4TH KEYWORD XXXX IS OUT OF RANGE"
```

If any LI REQ, FENCE commands failed to get the configured fence actions, thresholds, time options or fenced cluster states, the request fails with the error text:

```
"FAILED TO GET FENCE DATA FROM VPD"
```

If any LI REQ, FENCE commands to set the thresholds, time options or fence actions fails, the request fails with the error text:

```
"FAILED TO SET FENCE DATA IN VPD"
```

If LI REQ, FENCE, ACT, SEC, ENABLE is attempted when the target cluster has "NONE" or "ALERT" primary remote cluster fence action, the request fails with the error text:
"CANNOT ENABLE SEC OPTION WITH PRI NONE/ALERT OPTION"

11 Disclaimers

Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice. This information could include technical inaccuracies or typographical errors. IBM may make improvements and/or changes in the product(s) and/or programs(s) at any time without notice.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business. Any reference to an IBM Program Product in this document is not intended to state or imply that only that program product may be used. Any functionally equivalent program, that does not infringe IBM's intellectually property rights, may be used instead. It is the user's responsibility to evaluate and verify the operation of any non-IBM product, program or service.

The information provided in this document is distributed "AS IS" without any warranty, either express or implied. IBM EXPRESSLY DISCLAIMS any warranties of merchantability, fitness for a particular purpose OR NON INFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted according to the terms and conditions of the agreements (*e.g.*, IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided. IBM is not responsible for the performance or interpretability of any non-IBM products discussed herein. The customer is responsible for the implementation of these techniques in its environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. Unless otherwise noted, IBM has not tested those products in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

The provision of the information contained herein is not intended to, and does not grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

Trademarks

The following are trademarks or registered trademarks of International Business Machines in the United States, other countries, or both.

IBM, TotalStorage, DFSMS/MVS, S/390, z/OS, and zSeries.

Other company, product, or service names may be the trademarks or service marks of others.