Configuring Resiliency and Management Interface Bonding

To help protect against loss of assets due to network failure, COS Release 3.5.1 supports data resiliency:

- At the node level, enabling recovery of data lost due to drive failure within a node.
- At the cluster level, enabling recovery of lost data sue to node failure within a cluster.

In addition, COS supports bonding of two management interface ports to provide redundant connectivity in the event of a network interface card (NIC) failure.

This section describes these features and provides instructions for configuring them through the COS setup file or (for data resiliency only) through the COS Service Manager GUI.

Configuring Resiliency

COS provides resiliency at the node or cluster level using one of two methods: mirroring or erasure coding. Resiliency at the node level is achieved using either local mirroring (LM) or local erasure coding (LEC), which is the default. Similarly, resiliency at the cluster level is achieved using either remote mirroring (RM), which is the default, or distributed erasure coding (DEC).

You can configure resiliency either directly, by updating the COS file /arroyo/test/aftersetupfile, or indirectly, by creating and applying asset redundancy policies through the COS Service Manager GUI. COS also allows for the configuration of mixed resiliency policies through the GUI.

Asset redundancy policies are assigned at the service endpoint level. Because an endpoint can only be associated with a single cluster and policy, asset redundancy policies are always applied at the cluster level, rather than at the node level. When you apply a policy to an endpoint, the COS AIC client writes the appropriate line(s) to the setup file for each node in the cluster.

Note We recommend using either remote mirroring or DEC, but not both. COS 3.5.1 does not support migration from one scheme to another while preserving stored content.

An endpoint can be configured in the GUI, but it must remain in Disabled state until it is associated with both a cluster and an asset redundancy policy. There is no default policy, policy type, or rule associated with and endpoint in the GUI. After an asset redundancy policy is applied to an endpoint, you can change to a different policy simply by applying it to the endpoint instead.

About Mirroring

Mirroring configures the disks in a node, or the nodes in a cluster, to hold a specified number of exact copies of the object data. The number of copies can be set from 1 to 4, with 1 representing the original object data only, and 4 representing the object data plus 3 copies. Thus, for example, a value of 2 specifies one copy plus the original object data.

When viewed or updated in the GUI, the value representing the number of copies does not include the original object data. Thus, you would set this value to 2 in the GUI to specify two copies in addition to the original. This causes the value 3 to be written to the setup file.

About Erasure Coding

Erasure coding is a method of data protection in which data is redundantly encoded, divided into blocks, and distributed or *striped* across an array of different locations or storage media. With local erasure coding, the data is striped across disks. With distributed erasure coding, the data is striped across nodes.

The goal of erasure coding is to allow data corrupted due to drive or node failure to be reconstructed using data stored in a part of the array that was not corrupted. Erasure coding is functionally similar to RAID, but offers relatively faster data reconstruction at the cost of relatively higher CPU utilization. Data recovery is performed as a low-priority background task to mitigate any performance impact.

Configuring Resiliency Using the GUI

To configure resiliency through the COS Service Manager GUI:

Step 1 Log in to the GUI as described in Using the COS Service Manager GUI, page A-2.

Step 2 On the COS Home page, click to select the service to be configured in the Service Summary section.

The Cloud Object Stores page opens displaying the service definition and its service endpoints.

Step 3 Click the check box to the left of a service endpoint and click the **Edit** icon to enable it for editing.

Step 4 Choose the desired resiliency policy from the Asset Redundancy Policy drop-down list for the endpoint.

Note

- If the desired asset redundancy policy does not appear, it may not yet exist. To check, choose **Service Domain Objects > Asset Redundancy Policies** and review the list of existing policies. If necessary, click the **Add Row** button and create a new policy.
- The GUI does not currently support configuration of the M:N (data:parity) block values needed for distributed erasure coding. To configure non-default values, select new values as described in Finding M:N Values, page B-5 and edit (or if not present, create) the COS file **/arroyo/test/aftersetupfile** to hold the new configuration.

Caution When manually setting M:N (or any other) values that must persist, be sure to use aftersetupfile and not setupfile. The settings in setupfile can be overwritten by changes made via the GUI.

Configuring Local Mirroring Manually

To configure local mirroring on a node manually:

Step 1 Open (or if not present, create) the COS file **/arroyo/test/aftersetupfile** for editing.

Step 2 Include the line **vault local copy count** in the file and set the value to **2**, **3**, or **4** as appropriate.

Note Setting the value to **1** simply maintains the original data and creates no additional copies.

Step 3 Disable local erasure coding by setting **allow vault raid** to **0** (or simply omit or remove this line).

Example

```
# CServer core configuration. Changes to this file require a server reboot.
    serverid 1
    groupid  3333
    arrayid  6666
    . . . .
    allow vault raid 0
    vault local copy count 2
    vault mirror copies 2
    allow server raid 0
    allow tcp traffic 1
    . . . .
    er_enable 0
    rtp_enable 0
```

Configuring Local Erasure Coding Manually

To enable local erasure coding manually:

Step 1 Open (or if not present, create) the COS file **/arroyo/test/aftersetupfile** for editing.

Step 2 Set **allow vault raid** to **1** to enable LEC.

Step 3 Disable local mirroring by setting **vault local copy count** to **0** (or simply omit or remove this line).

Example

```
# CServer core configuration. Changes to this file require a server reboot.
    serverid 1
    groupid  3333
    arrayid  6666
    . . . .
    allow vault raid 1
    vault local copy count 0
    vault mirror copies 2
    allow server raid 0
    allow tcp traffic 1
    . . . .
    er_enable 0
    rtp_enable 0
```

Migrating from LM to LEC Manually

To migrate a service endpoint from local mirroring to local erasure coding:

Step 1 Temporarily leave local mirroring enabled for the service endpoint.

Step 2 Enable local erasure coding for the service endpoint and allow it to establish the needed parity for each data object.

Step 3 When parity is established, disable local mirroring.

Configuring Remote Mirroring Manually

To enable and configure remote mirroring manually:

Step 1 Open (or if not present, create) the COS file **/arroyo/test/aftersetupfile** for editing.

Step 2 Set **vault mirror copies** to the value **2**, **3**, or **4** as appropriate to enable remote mirroring. The value you enter specifies the object data plus the number of exact copies desired.

Note Setting the value to **1** simply maintains the original data and creates no additional copies.

Step 3 Disable distributed erasure coding by setting **allow server raid** to **0** (or simply omit or remove this line).

Example

```
# CServer core configuration. Changes to this file require a server reboot.
    serverid 1
    groupid  3333
    arrayid  6666
    . . . .
    allow vault raid 0
    vault local copy count 2
    vault mirror copies 2
    allow server raid 0
    allow tcp traffic 1
    . . . .
    er_enable 0
    rtp_enable 0
```

Configuring Distributed Erasure Coding Manually

To enable and configure distributed erasure coding manually:

Step 1 Open (or if not present, create) the COS file **/arroyo/test/aftersetupfile** for editing.

Step 2 Set **allow server raid** to **1** and add the following lines immediately below:

- **target server raid data blocks <value>**

  This controls the number of data blocks used. The default <value> is 8, and the valid range is 1-18.

- **target server raid parity blocks <value>**

  This controls the number of parity blocks used. The default <value> is 1, and the valid range is 1-18.

Note See Finding M:N Values, page B-5 to determine appropriate data block and parity block values.

Step 3 Disable remote mirroring by setting **vault mirror copies** to **0** (or simply omit or remove this line).

Example

```
# CServer core configuration. Changes to this file require a server reboot.
    serverid 1
    groupid  3333
    arrayid  6666
    . . . .
    allow vault raid 0
    vault local copy count 2
    vault mirror copies 0
    allow server raid 1
    target server raid data blocks 8
    target server raid parity blocks 1
    allow tcp traffic 1
    . . . .
    er_enable 0
    rtp_enable 0
```

Finding M:N Values

To configure DEC, you must specify the number of data blocks (M) and parity blocks (N) used for data encoding. Table B-1 shows the corresponding data-to-parity block (M:N) values for a given number of nodes in a cluster and for a given degree of resiliency desired for the cluster.

Note COS does not currently support configuration of new M:N (data:parity) block values through the COS Service Manager GUI. If you need to configure new M:N values, you must do so in the **aftersetup** file.

In this table:

- **Nodes** is the number of nodes in the cluster.
- **RF** is the desired *resiliency factor*, or number of nodes that can fail without data loss.
- **Min** is the minimum number of nodes required to achieve a given resiliency factor.

The ratios appearing in the cells of the table are M:N values, where M is the number of data blocks and N is the number of parity blocks needed to achieve the desired resiliency factor for a given node count.

To use the table to find the M:N values for a cluster:

Step 1 In the Nodes column, locate the **row** corresponding to the number of nodes in the cluster.

Note For COS 3.5.1, you must select the M:N configuration based upon the initial nodes in the cluster. COS does not support adding nodes to a cluster after DEC is configured for the cluster. (Automatic tuning of M:N values is under consideration as a future enhancement.)

Step 2 Locate the **column** in the table whose header represents the desired RF value for the cluster.

Step 3 Find the corresponding **M:N value** at the intersection of the row and column just located.

Step 4 Configure DEC using **M** as the number of data blocks and **N** as the number of parity blocks.

Table B-1 Possible M:N Values for DEC

Nodes	Min = 1 RF = 0	Min = 3 RF = 1	Min = 5 RF = 2	Min = 7 RF = 3	Min = 9 RF = 4
1	1:0	—	—	—	—
2	1:0	—	—	—	—
3	1:0	1:1	—	—	—
4	1:0	2:1	—	—	—
5	1:0	3:1	2:2	—	—
6	1:0	4:1	3:2	—	—
7	1:0	5:1	4:2	3:3	—
8	1:0	6:1	5:2	4:3	—
9	1:0	7:1	6:2	5:3	4:4
10	1:0	8:1	7:2	6:3	5:4
11	1:0	8:1	8:2	7:3	6:4
12	1:0	8:1	8:2	8:3	7:4
13	1:0	8:1	8:2	9:3	8:4
14	1:0	8:1	8:2	10:3	9:4
15	1:0	8:1	8:2	11:3	10:4
16	1:0	8:1	8:2	12:3	11:4
17	1:0	8:1	8:2	12:3	12:4
18	1:0	8:1	8:2	12:3	12:4
19	1:0	8:1	8:2	12:3	12:4
20	1:0	8:1	8:2	12:3	12:4

Configuring Management Interface Bonding

COS supports the ability to bond two NICs so that they appear to the host node as a single logical management interface. With management interface bonding, two ports on a node are defined as a primary-backup pair.

- For the C3160, the designated ports are eth0 and eth3.
- For the CDE 465, the designated ports are eth0 and eth1.

Note COS 3.5.1 supports bonding of management interfaces for resiliency, but not for improved performance.

If both of the NICs in the node are bonded, the management link is not lost if either logical interface or its physical link or lost. The management link is also maintained if either physical link is disconnected and then reconnected, and then the other physical link is disconnected. Additionally, the management interface bonding feature itself remains enabled if the node is rebooted.

Existing COS nodes that use one management interface can be upgraded to COS Release 3.5.1 without being reinitialized. The cosinit routine provides the option to add management bonding functionality to COS nodes being upgraded to this release.

To Configure Bonding Manually

To configure management interface bonding manually:

Step 1 Open (or if not present, create) the COS file **/arroyo/test/aftersetupfile** for editing.

Step 2 Add the line **management bond <value>**, where <value> is **0** to disable or **1** to enable the feature.

When enabled, one NIC serves as the primary management interface and the other as the backup interface.