Installation and usage instructions for the Call Recorder Apresa in a High Availability setup with two nodes

Introduction

The purpose of Apresa High Availability is to keep the recording system functioning and available for use, even when a single hardware failure occurs. The setup consists of two nodes (two computers) with a network link in between. The primary node records calls. The secondary node is in stand-by mode, until it takes over when the primary fails.

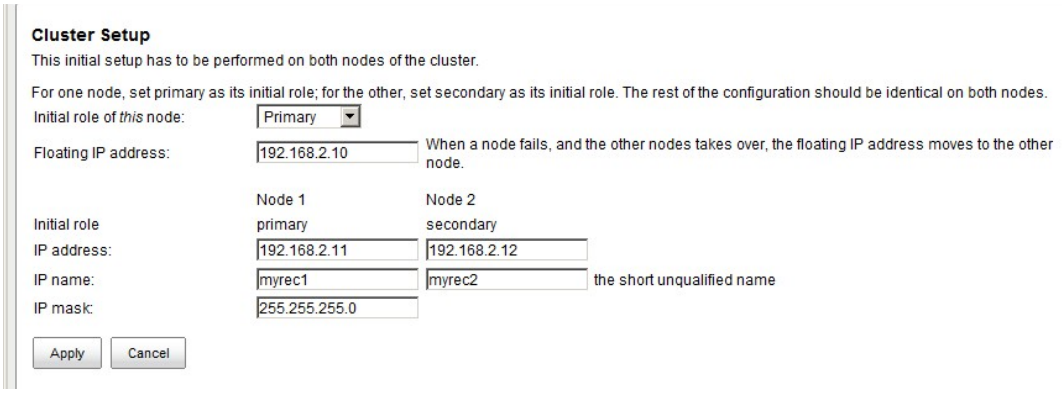
Installation

Two computers (virtualized or not) are needed. Both system should have an identical hard disk of the same type and size.

VoIP data that need to be recorded needs to be sent (mirrored/spanned) to both nodes.

Install the software using the general Apresa CD installation image, and select the option **HA-Cluster** in the menu. Please note that installing Apresa will destroy existing data on the hard disk. Existing installations cannot be upgraded to HA. Instead, a fresh installation is required, because a special hard disk layout is needed.

After CD installation, connect to the system as usual at 192.168.55.55 (the default) using a web browser. After login, a setup screen is displayed to fill in the initial parameters. Always start with the primary node. An example configuration is shown below.



If you use a cluster IP mask other than 255.255.255.0, and you are using the installer with Apresa version 8.5.0 and Debian 8, the following command needs to be run in the system shell before starting the cluster setup:

```
sudo /usr/share/apresa/aset IPmask THEMASK
```

And fill in THEMASK with the IP mask of the cluster.

After clicking Apply, the system will do the initial cluster setup, which can take a few minutes. After completion, it should be possible to logon to the primary node. The cluster setup of the secondary should only be started after:

1) the setup of the primary has completed successfully,
2) normal logon is possible on the primary,
3) and the primary and secondary node have a network connection between them.

Initially the DRBD recordings drive needs to synchronize (from the primary to the secondary). Depending on the size of the hard disk, this will take several hours. While this synchronization is busy, the system can be used, but it is not yet ready to perform a fail-over.

Network connectivity

As calls are recorded and stored on the hard disk, this data is also sent over the network, from the primary to the secondary. The amount of data will depend on the number of simultaneous calls that are recorded. It is important that the network connection is fast enough to handle this. If many calls are recorded, it is necessary to use at least a gigabit connection.

It is important that the network connection between the two nodes is maintained at all times. If network connectivity is lost, it is likely that both nodes will conclude that the other node is down, both nodes will become primary, resulting in a so-called “split-brain”, which might require manual action to resolve. To avoid this problem, the cluster should first be put in maintenance mode.

Start of a node

Since Apresa version 9.2.0, when a node starts, it will not automatically join the cluster. This allows for an administrator to inspect the node first, to investigate for failure, evaluate its status, and to initiate recovery actions (See: *Recovery from a single-node failure*).
To join a node to the cluster, issue the following command on the command line:

```
sudo service pacemaker start
```

It can take a minute for the node to come online. If a node is in an incorrect state, and the node is joined to the cluster without repairing it, a two-node failure could occur.

Start of the cluster

If both nodes are offline (shutdown), to bring the cluster online, the following procedure can be followed:
- Power on one of the nodes, and start pacemaker on this node
- Wait until the node becomes primary, and functional
- Start the other node, and follow the procedure for recovery from a single-node failure

Failover

A failover is when the primary node has failed, and the secondary node takes over. The secondary then becomes the new primary node. This happens automatically.

To trigger a failover with limited side effects: (to test it)
- On the primary, in the system shell, type: `sudo service pacemaker restart`
Hard power-off would also trigger a failover, but might have additional side effects.

During a fail-over, the web interface can be temporarily unavailable, although the call recording continues. When the failover is completed, the floating IP address has moved to the other node that has taken over, and the web interface should be available again.

During a fail-over, some calls might be recorded twice.

Recovery from a single-node failure

Recovering from a fail-over is not automatic. The failed node should be analyzed to find the cause of the failure.

To see the current status of the DRBD recordings data storage:
- Type: `cat /proc/drbd`
If it is good, it should tell it is Connected, and the roles should be Secondary/Primary. To recover from a problem with the DRBD recordings data storage, on the failing node:
- Type: `sudo drbdadm -- --discard-my-data connect r0`

On the good node:
- Type: `sudo drbdadm connect r0`

To ensure data integrity, the database of the failing node will usually get disabled automatically, and must be restored from the primary node. This can be done from the command line, as follows:
- Log on to the system shell (vidi account) of the failed node
- Type: `sudo /usr/share/apresa/recoverdb.sh IPADDR`
where IPADDR should be replaced with the IP address or IP name of the node that still works.

The two-node cluster protects again a failure of one node. It does not prevent against a double failure of both nodes. If one node has failed, it must first be brought back to good operation, before it can function as a good backup for the other node. To verify after a failover that the cluster is again in good health, check the system information page (Tools, System, System information), the items below H.A. Cluster.

Two-node failure

During a two-node failure, the H.A. cluster cannot provide its services. Normally, this should never happen. If there is a two-node failure*, and the only remaining failure is the database*, then to recover, go to the node that contains the latest correct data, and in the system shell, do the following: (as root)

```
crm configure property maintenance-mode=on
service postgresql stop
rm -f /database/db/main/recovery.conf
rm -f /var/lib/pgsql/tmp/PGSQL.lock
service postgresql start
crm resource cleanup postgresql
crm configure property maintenance-mode=off
```

There should be a number of seconds of pause between some of these commands, to allow it to be applied. When the node has been restored, follow the usual procedure to recover from a single-node failure, to recover the other node.

Maintenance

When maintenance is done on the nodes or its network, the cluster should be put in maintenance mode to avoid a split-brain or an unintended failover. This can be done on the command line:

```
sudo crm configure property maintenance-mode=on
```

The cluster status (`crm status`) should now display “unmanaged”. During maintenance mode, no automatic failover will be performed, but recording continues as normal. After the maintenance work is completed, the cluster maintenance mode should again be switched off, from the command line, as follows:

```
sudo crm configure property maintenance-mode=off
```

Call content encryption

After encryption has been enabled, to allow encryption to take place also on the secondary, playback decryption must be unlocked on the secondary at least once. To do so, click on the unlock link displayed in the web interface, when logged on as administrator. After that, restart the recording component on both the primary and secondary (menu Tools, System). This is a one-time procedure, and it is not necessary to repeat it for encryption to function. However, for playback to function, playback has to be unlocked after each reboot.

Normal Use

The floating IP address is the IP address that regular users should use to access the system.

Administration

Changes to the configuration of the system are done on the primary node. The web interface of the secondary node is available for log-on, but provides a mostly read-only view of the system, and some administrative functions. A limited number of settings that are machine-specific (such as licensing) can be configured per node, so also on the secondary. In the system settings, the settings that cannot be changed are grayed out. The IP addresses and names of the nodes should not be changed, because this will cause malfunction.

To display the cluster state for verification, from the Tools menu, choose Cluster (on any node). Also see the H.A. Cluster status on the system information page.

Software updates should be installed first on the primary node, and after it has finished, also on the secondary node.