HAUTE DISPONIBILITE LOGICIELLE ET BLUEMIND

You can set up a High Availability system that integrates with BlueMind.

This page provides recommendations and information about the BlueMind system required to be able to integrate the mail solution into a high availability infrastructure.

The third-party software solutions mentioned here are provided for illustration purposes only. This list is not comprehensive.

GETTING THE SYSTEM READY

Note: the two servers involved must follow the hardware sizing recommendations defined in the following section: Hardware Sizing

Storage space

The contents you want to share between the two servers can be shared either on a separate shared storage space such as a SAN (*Storage Area Network*), or through data replication between two separate storage spaces.

Replication-based high availability can cause major issues with access to shared disk resources which may occur during loss of service. The most typical issue with resource access and with potentially disastrous consequences occurs in a *split-brain* situation.

The *cyrus-imap* component does not support NFS-based storage. As a result, regardless of the type of replicated storage you choose, you need a *block-device*-based storage using technologies such as Fibre Channel or iSCSI for the data this component handles (/var/spool/cyrus and /var/lib/cyrus).

Data to be made available between both servers

The data located in the following directories must be made visible by both servers and its access must be managed by the HA handling system:

- /var/spool/bm-docs
- /var/spool/bm-elasticsearch
- /var/spool/bm-hsm
- /var/spool/cyrus
- /var/spool/postfix
- /var/spool/sieve
- /var/spool/bluemind-pkgs

The cyrus database located in the following directory must also be added to this data:

- /var/lib/cyrus
- /var/lib/postgresql

This data must therefore be located in a storage space -- SAN storage, GFS cluster, etc – that allows the passive server to access the data during switchovers.

REMINDER: /var/spool/cyrus MUST NOT be stored on an NFS mount.

Network

To work properly, BlueMind must be accessible through a single URL/IP. We therefore recommend that you use a system that is capable of handling floating (or virtual) IP addresses.

BlueMind's front-end access URL MUST always be the same.

Monitoring scripts

Please see our Monitoring page.

SETTING UP HIGH AVAILABILITY

If you are not using STONITH (see below), you must not enable automatic changeover otherwise you may end up with a split-brain and corrupted data (see box in the dedicated paragraph) which will not be covered by BlueMind support.

Data and services that need to be managed by HA

High availability-based synchronization of BlueMind configuration files

BlueMind's configuration files that must be synchronized in real time by the HA handling system are located under /etc

The following files must also be synchronized:

- /usr/share/bm-elasticsearch/config/elasticsearch.yml
- /etc/aliases
- /etc/aliases.db
- /etc/sysctl.conf
- /etc/ssl/certs/bm_cert.pem
- /var/lib/bm-ca/ca-cert.pem

Here are a few examples of how to synchronize configuration files in real time:

- incron, based on inotify, allows you to launch jobs depending on a file's status for example. The official documentation is available on the vendor's website.
- files can be copied by *rsync over ssh* for example, as shown on this website.
- other tools include lsyncd and csync2

Managing the BlueMind update

The key steps for updating a High Availability-based deployment of BlueMind are described below:

- Before you start the BlueMind update, disable the high availability handling services.
- Update the packages on both servers.
- Next, **on the main server** with the public IP address **only**, carry out the post-installation configuration as described in: Post-installation Configuration.

STONITH

STONITH, which stands for *Shoot The Other Node In The Head*, is a fencing or node isolation technique in cluster management. Its purpose is to shut down a server's failed cluster remotely – either through software or by directly cutting off its power supply.

This is done at the hardware infrastructure level.

This safety system strongly lowers the risk of corrupted data in the event of complex service failures, e.g. a split-brain, which leads both servers to consider themselves the sole master and attempt to access the shared storage resource at the same time. With data replication-based high availability, the risk of data corruption is high.

This technique can for instance be implemented using IPMI tools (*Intelligent Platform Management Interface*). IPMI is a server management interface specification whose implementations include freeIPMI, OpenIPMI, ipmitool, ...

As far as hardware is concerned, implementation can be done on dedicated hardware or using iDRAC cards for DELL equipment.