




Release Notes for SUSE Linux Enterprise High Availability Extension 11 Services Pack 4





Release Notes for SUSE Linux Enterprise High Availability Extension 11 Services Pack 4

These release notes apply to all SUSE Linux Enterprise High Availability Extension 11 Services Pack 4 based products (e.g. for x86, x86_64, Itanium, Power and System z). Some sections may not apply to a particular architecture or product. Where this is not obvious, the respective architectures are listed explicitly in these notes. Instructions for installing SUSE Linux Enterprise High Availability Extension can be found in the README file on the CD.

Manuals can be found in the docu directory of the installation media. Any documentation (if installed) can be found in the /usr/share/doc/ directory of the installed system.

This SUSE product includes materials licensed to SUSE under the GNU General Public License (GPL). The GPL requires SUSE to provide the source code that corresponds to the GPL-licensed material. The source code is available for download at <http://www.suse.com/download-linux/source-code.html> .

Also, for up to three years after distribution of the SUSE product, upon request, SUSE will mail a copy of the source code. Requests should be sent by e-mail to mailto:sle_source_request@suse.com  or as otherwise instructed at <http://www.suse.com/download-linux/source-code.html> . SUSE may charge a reasonable fee to recover distribution costs.

Publication date: 2015-06-19 , Version: Version 11.4.6 (2015-05-12)

SUSE Linux Products GmbH

Maxfeldstr. 5

90409 Nürnberg

GERMANY

<http://www.suse.com/documentation> 

Contents

- 1 Purpose and News 1**
- 2 Features and Versions 3**
- 3 Changed Functionality in SUSE Linux Enterprise High Availability Extension 11 SP4 6**
- 4 Deprecated Functionality in SUSE Linux Enterprise High Availability Extension 11 SP4 7**
- 5 Infrastructure, Package and Architecture Specific Information 8**
- 6 Other Updates 12**
- 7 Update-Related Notes 13**
- 8 Supported Deployment Scenarios SUSE Linux Enterprise High Availability Extension 11 SP4 14**
- 9 Known Issues in SUSE Linux Enterprise High Availability Extension 11 SP4 16**
- 10 Further Notes on Functionality 17**

- 11 Support Statement for SUSE Linux Enterprise High Availability Extension 11 SP4 19**
- 12 Technical Information 20**
- 13 Miscellaneous 21**
- 14 More Information and Feedback 22**

1 Purpose and News

1.1 Purpose

SUSE Linux Enterprise High Availability Extension is an affordable, integrated suite of robust open source clustering technologies that enable enterprises to implement highly available Linux clusters and eliminate single points of failure.

Used with SUSE Linux Enterprise Server, it helps firms maintain business continuity, protect data integrity, and reduce unplanned downtime for their mission-critical Linux workloads.

SUSE Linux Enterprise High Availability Extension provides all of the essential monitoring, messaging, and cluster resource management functionality of proprietary third-party solutions, but at a more affordable price, making it accessible to a wider range of enterprises.

It is optimized to work with SUSE Linux Enterprise Server, and its tight integration ensures customers have the most robust, secure, and up to date high availability solution. Based on an innovative, highly flexible policy engine, it supports a wide range of clustering scenarios.

With static or stateless content, the High Availability cluster can be used without a cluster file system. This includes web-services with static content as well as printing systems or communication systems like proxies that do not need to recover data.

Finally, its open source license minimizes the risk of vendor lock-in, and its adherence to open standards encourages interoperability with industry standard tools and technologies.

1.2 What's New in SUSE Linux Enterprise High Availability Extension 11 SP4

In Service Pack 4, a number of improvements have been added, some of which are called out explicitly here. For the full list of changes and bugfixes, refer to the change logs of the RPM packages. These changes are in addition to those that have already been added with Service Pack 1, 2, and 3. Here are some highlights:

SUSE Linux Enterprise High Availability Extension 11 Service Pack 4 includes a new feature called `pacemaker_remote`. It allows nodes not running the cluster stack (`pacemaker+corosync`) to integrate into the cluster and have the cluster manage their resources just as if they were a real cluster node. This feature makes it ideal to deploy large scale SAP deployment by supporting worker nodes as scale-out clustering option. With this feature, you can deploy a HA cluster with up to 40+ nodes.

SP4 improves the usability in HAWK by introducing new templates and wizards. Oracle and DB2 related templates and wizards make SP4 easier to use to boost your database availability.

2 Features and Versions

This section includes an overview of some of the major features and new functionality provided by SUSE Linux Enterprise High Availability Extension 11 SP4.

2.1 Resource Management

2.1.1 Data Replication—Distributed Remote Block Device (DRBD)

Data replication is part of a disaster prevention strategy in most large enterprises. Using network connections data is replicated between different nodes to ensure consistent data storages in case of a site failure.

Data replication is provided in SUSE Linux Enterprise High Availability Extension 11 SP4 with DRBD. This software based data replication allows customers to use different types of storage systems and communication layers without vendor lock-in. At the same time, data replication is deeply integrated into the operating system and thus provide ease-of-use. Features related to data replication and included with this product release are:

- YaST setup tools to assist initial setup
- Fully synchronous, memory synchronous or asynchronous modes of operation
- Differential storage resynchronization after failure
- Bandwidth of background resynchronization tunable
- Shared secret to authenticate the peer upon connect
- Configurable handler scripts for various DRBD events
- Online data verification

With these features data replication can be easier configured and used. And with improved storage resynchronization recovery times will be decreased significantly.

The distributed replicated block device (DRBD) version included supports active/active mirroring, enabling the use of services such as cLVM2 or OCFS2 on top.

2.1.2 IP Load Balancing—Linux Virtual Server (LVS)

Linux Virtual Server (LVS) is an advanced IP load balancing solution for Linux. IP load balancing provides a high-performance, scalable network infrastructure. Such infrastructure is typically used by enterprise customers for web servers or other network related service workloads.

With LVS network requests can be spread over multiple nodes to scale the available resources and balance the resulting workload. By monitoring the compute nodes, LVS can handle node failures and redirect requests to other nodes maintaining the availability of the service.

2.1.3 Distributed Lock Manager (DLM)

The DLM in SUSE Linux Enterprise High Availability Extension 11 SP4 supports both TCP and SCTP for network communications, allowing for improved cluster redundancy in scenarios where network interface bonding is not feasible.

2.2 Other Changes and Version Updates

2.2.1 Support coredumps with STONITH Enabled and Timeouts for Kdump

The `kdumpcheck` STONITH plugin did not work as expected. This plug-in checks if a kernel dump is in progress on a node. If so, it returns `true`, and acts as if the node has been fenced. This avoids fencing a node that is already down but doing a dump, which takes some time.

Use the `stonith:fence_kdump` resource agent (provided by the package `fence-agents`) to monitor all nodes with the `kdump` function enabled. In `/etc/sysconfig/kdump`, configure `KDUMP_POSTSCRIPT` to send a notification to all nodes when the `kdump` process is finished. The node that does a `kdump` will restart automatically after `kdump` has finished.

Do not forget to open a port in the firewall for the `fence_kdump` resource. The default port is `7410`.

2.2.2 IPVS support for iptables -m

Setting up Linux Virtual Server (LVS) in combination with SMTP did not work due to missing iptables -m ipvs support. IPVS (IP Virtual Server) is used for Linux Virtual Server.

The iptables package has been updated to a higher version and now includes the match support of iptables. To make it work, you need the probe kernel module xt_ipvs . It is provided by the cluster-network-* package.

3 Changed Functionality in SUSE Linux Enterprise High Availability Extension 11 SP4

4 Deprecated Functionality in SUSE Linux Enterprise High Availability Extension 11 SP4

5 Infrastructure, Package and Architecture Specific Information

5.1 Architecture Independent Information

5.1.1 Security

5.1.1.1 Improved pssh -P Output

Using pssh with the -P option now prints the host name in front of each line received.

5.1.2 Network

5.1.2.1 crmsh: Manage Multiple Resources as One Using Resource Tags

It may be desired to start and stop multiple resources all at once, without having explicit dependencies between those resources.

This feature adds resource tags to Pacemaker. Tags are collections of resources that do not imply any colocation or ordering constraints, but can be referenced in constraints or when starting and stopping resources.

5.1.2.2 Enable Colocating Resources Without Further Dependency Between Them

Sometimes, it is desired that two resources should run on the same node. However, there should be no further dependency implied between the two resources, so that if one fails, the other one can keep running.

This can be accomplished using a third, dummy resource which both resources depend on in turn. To make it easier to create a configuration like this, the command "assist weak-bond" has been added to crmsh.

5.1.2.3 Avoid Starting openais at Boot

In the past you have had to start the cluster at boot unconditionally, if you want to make sure that the cluster stops when the server stops.

Now it is possible to the cluster later.

The openais service is still in sysconfig, but additionally, we now have a parameter START_ON_BOOT=Yes/No in /etc/sysconfig/openais.

- If START_ON_BOOT=Yes (default), the openais service will start at boot.
- If START_ON_BOOT=No, the openais service will not start at boot. Then you can start it manually whenever you want to start it.

5.2 Systems Management

5.2.1 Hawk Wizards for Common Configurations

Multiple wizards have been added to hawk to ease configuration, including cLVM + MD-RAID and DRBD.

5.2.2 Hawk Wizard for DB2 HADR

Hawk now includes a wizard for configuring a cluster resource for the DB2 HADR database.

5.2.3 Hawk Wizard for DB2

Hawk now includes a wizard for configuring a cluster resource for the DB2 database.

5.2.4 Hawk Wizard for Configuring the Oracle Database

Hawk now includes a wizard for configuring cluster resources for the Oracle database.

5.2.5 Hawk Wizards for Configuring Common Scenarios

Hawk now includes additional wizards for configuring common cluster configurations.

5.2.6 crmsh: Enable Anonymous Shadow CIBs

When scripting cluster configuration changes, the scripts are more robust if changes are applied with a single commit. Creating a shadow CIB to collect the changes makes this easier, but has previously required naming the shadow CIB.

crmsh now allows the creation of shadow CIBs without explicitly specifying a name. The name will be determined automatically and will not clash with any other previously created shadow CIBs.

5.3 AMD64/Intel64 64-Bit (x86_64) and Intel/AMD 32-Bit (x86) Specific Information

5.3.1 System and Vendor Specific Information

5.3.1.1 Additional Relax-and-Recover Version 1.16 (rear116)

When Relax-and-Recover version 1.10.0 does not work the newer version 1.16 could help.

In addition to Relax-and-Recover version 1.10.0 that is still provided in the RPM package "rear" we provide Relax-and-Recover version 1.16 as additional totally separated RPM package "rear116".

A new separated package name rear116 is used so that users where version 1.10.0 does not support their particular needs can manually upgrade to version 1.16 but on the other hand users who have a working disaster recovery procedure with version 1.10.0 do not need to upgrade. Therefore the package name contains the version and the packages conflict with each other to avoid that an installed version may get accidentally replaced with another version.

When you have a working disaster recovery procedure, do not change it!

For each rear version upgrade you must carefully re-validate that your particular disaster recovery procedure still works for you.

See in particular the section "Version upgrades" at https://en.opensuse.org/SDB:Disaster_Recovery  .

6 Other Updates

7 Update-Related Notes

This section includes update-related information for this release.

8 Supported Deployment Scenarios SUSE Linux Enterprise High Availability Extension 11 SP4

The SUSE Linux Enterprise High Availability Extension stack supports a wide range of different cluster topologies.

Local and Metro Area (stretched) clusters are supported as part of a SUSE Linux Enterprise High Availability Extension subscription. Geographical clustering requires an additional Geo Clustering for SUSE Linux Enterprise High Availability Extension subscription.

8.1 Local Cluster

In a local cluster environment, all nodes are connected to the same storage network and on the same network segment; redundant network interconnects are provided. Latency is below 1 millisecond, and network bandwidth is at least 1 Gigabit/s.

Cluster storage is fully symmetric on all nodes, either provided via the storage layer itself, mirrored via MD Raid1, cLVM2, or replicated via DRBD.

In a local cluster all nodes run in a single corosync domain, forming a single cluster.

8.2 Metro Area Cluster

In a Metro Area cluster, the network segment can be stretched to a maximum latency of 15 milliseconds between any two nodes (approximately 20 miles or 30 kilometers in physical distance), but fully symmetric and meshed network inter-connectivity is required.

Cluster storage is assumed to be fully symmetric as in local deployments.

As a stretched version of the local cluster, all nodes in a Metro Area cluster run in a single corosync domain, forming a single cluster.

8.3 Geographical Clustering

A Geo scenario is primarily defined by the network topology; network latency higher than 15 milliseconds, reduced network bandwidth, and not fully interconnected subnets. In these scenarios, each site by itself must satisfy the requirements of and be configured as a local or metropolitan cluster as defined above. A maximum of three sites are then connected via Geo Clustering for SUSE Linux Enterprise High Availability Extension; for this, direct TCP connections between the sites must be possible, and typical latency should not exceed 1 second.

Storage is typically asymmetrically replicated by the storage layer, such as DRBD, MD Raid1, or vendor-specific solutions.

DLM, OCFS2, and cLVM2 are not available across site boundaries.

9 Known Issues in SUSE Linux Enterprise High Availability Extension 11 SP4

9.1 Linux Virtual Server Tunnelling Support

The LVS TCP/UDP load balancer currently only works with Direct Routing and NAT setups. IP-over-IP tunnelling forwarding to the real servers does not currently work.

9.2 Samba CTDB Cluster Rolling Update Support

The CTDB resource should be stopped on all nodes prior to update. Rolling CTDB updates are not supported for this release, due to the risk of corruption on nodes running previous CTDB versions.

10 Further Notes on Functionality

10.1 Cluster-concurrent RAID1 Resynchronization

To ensure data integrity, a full RAID1 resync is triggered when a device is re-added to the mirror group. This can impact performance, and it is thus advised to use multipath IO to reduce exposure to mirror loss. Due to the need of the cluster to keep both mirrors uptodate and consistent on all nodes, a mirror failure on one node is treated as if the failure had been observed cluster-wide, evicting the mirror on all nodes. Again, multipath IO is recommended to reduce this risk.

In situations where the primary focus is on redundancy and not on scale-out, building a storage target node (using `md raid1` in a fail-over configuration or using `drbd`) and reexporting via iSCSI, NFS, or CIFS could be a viable option.

10.2 Quotas on OCFS2 Filesystem

To use quotas on ocfs2 filesystem, the filesystem has to be created with appropriate quota features: 'usrquota' filesystem feature is needed for accounting quotas for individual users, 'grpquota' filesystem feature is needed for accounting of quotas for groups. These features can be also enabled later on an unmounted filesystem using `tuneefs.ocfs2`.

For quota-tools to operate on the filesystem, you have to mount the filesystem with 'usrquota' (and/or 'grpquota') mount option.


When a filesystem has appropriate quota feature enabled, it maintains in its metadata how much space and files each user (group) uses. Since ocfs2 treats quota information as a filesystem internal metadata, there is no need to ever run `quotacheck(8)` program. Instead, all the needed functionality is built into `fsck.ocfs2` and the filesystem driver itself.

To enable enforcement of limits imposed on each user / group, run `quotaon(8)` program similarly as for any other filesystem.

Commands `quota(1)`, `setquota(8)`, `edquota(8)` work as usual with ocfs2 filesystem. Commands `repquota(8)` and `warnquota(8)` do not work with ocfs2 because of a limitation in the current kernel interface.

For performance reasons each cluster node performs quota accounting locally and synchronizes this information with a common central storage once per 10 seconds (this interval is tunable by `tunefs.ocfs2` using options `'usrquota-sync-interval'` and `'grpquota-sync-interval'`). Thus quota information need not be exact at all times and as a consequence user / group can slightly exceed their quota limit when operating on several cluster nodes in parallel.

11 Support Statement for SUSE Linux Enterprise High Availability Extension 11 SP4

Support requires an appropriate subscription from SUSE; for more information, see http://www.suse.com/products/server/services_support.html .

A Geo Clustering for SUSE Linux Enterprise High Availability Extension subscription is needed to receive support and maintenance to run geographical clustering scenarios, including manual and automated setups.

Support for the DRBD storage replication is independent of the cluster scenario and included as part of the SUSE Linux Enterprise High Availability Extension product and does not require the addition of a Geo Clustering for SUSE Linux Enterprise High Availability Extension subscription.

General Support Statement

The following definitions apply:

- L1: Installation and problem determination - technical support designed to provide compatibility information, installation and configuration assistance, usage support, on-going maintenance and basic troubleshooting. Level 1 Support is not intended to correct product defect errors.
- L2: Reproduction of problem isolation - technical support designed to duplicate customer problems, isolate problem areas and potential issues, and provide resolution for problems not resolved by Level 1 Support.
- L3: Code Debugging and problem resolution - technical support designed to resolve complex problems by engaging engineering in patch provision, resolution of product defects which have been identified by Level 2 Support.

SUSE will only support the usage of original (unchanged or not recompiled) packages.

12 Technical Information

13 Miscellaneous

14 More Information and Feedback

- Read the READMEs on the CDs.
- Get detailed changelog information about a particular package from the RPM:

```
rpm --changelog -qp <FILENAME>.rpm
```

<FILENAME>. is the name of the RPM.

- Check the ChangeLog file in the top level of CD1 for a chronological log of all changes made to the updated packages.
- Find more information in the docu directory of CD1 of the SUSE Linux Enterprise High Availability Extension CDs. This directory includes a PDF version of the High Availability Guide.
- http://www.suse.com/documentation/sle_ha/ contains additional or updated documentation for SUSE Linux Enterprise High Availability Extension 11.
- Visit <http://www.suse.com/products/> for the latest product news from SUSE and <http://www.suse.com/download-linux/source-code.html> for additional information on the source code of SUSE Linux Enterprise products.

Copyright © 2015 SUSE LLC.

Thanks for using SUSE Linux Enterprise High Availability Extension in your business.

The SUSE Linux Enterprise High Availability Extension Team.