Einführung in SUSE Linux Enterprise High Availability

WAS?

In diesem Artikel werden die Funktionen, die Architektur, die Kernkonzepte und die Vorteile von SUSE Linux Enterprise High Availability erläutert.

WARUM?

Erfahren Sie mehr über SUSE Linux Enterprise High Availability, um zu entscheiden, ob Sie es verwenden möchten, oder bevor Sie zum ersten Mal einen Cluster einrichten.

AUFWAND

Etwa 20 Minuten Lesedauer.

Veröffentlicht: 11.12.2025

Inhalt

1 Was ist SUSE Linux Enterprise High Availability? 3

2 Kernkonzepte 7

3 Überblick über die Architektur 12

4 Installationsoptionen 16

5 Weitere Informationen 16

6 Rechtliche Hinweise 17

A GNU Free Documentation License 17

Hochverfügbarkeit – Glossar 25

1 Was ist SUSE Linux Enterprise High Availability?

Bei SUSE® Linux Enterprise High Availability handelt es sich um eine integrierte Suite mit Open Source-Clustering-Technologien. Ein *Hochverfügbarkeits*-Cluster ist eine Gruppe von Servern (*Knoten*), die zusammenarbeiten, um die höchstmögliche Verfügbarkeit von Daten, Anwendungen und Diensten sicherzustellen. Fällt ein Knoten aus, werden die *Ressourcen*, die auf diesem Knoten ausgeführt wurden, mit geringer oder ohne Ausfallzeit auf einen anderen Knoten verschoben. Sie können auch manuell Ressourcen zwischen den Knoten verschieben, um die Last auszugleichen oder um Wartungsaufgaben mit minimaler Ausfallzeit auszuführen. Bei Clusterressourcen kann es sich um Websites, E-Mail-Server, Datenbanken, Dateisysteme, virtuelle Rechner und andere serverbasierte Anwendungen oder Dienste handeln, die dem Benutzer stets zur Verfügung stehen müssen.

1.1 Produktverfügbarkeit

SUSE Linux Enterprise High Availability ist mit den folgenden Produkten verfügbar:

SUSE Linux Enterprise Server

High Availability ist als Erweiterung verfügbar. Dafür ist ein zusätzlicher Registrierungscode erforderlich.

SUSE Linux Enterprise Server für SAP-Anwendungen

High Availability ist als Modul enthalten. Es ist kein zusätzlicher Registrierungscode erforderlich.

1.2 Funktionen und Vorteile

SUSE Linux Enterprise High Availability eliminiert Single Points of Failure und verbessert so die Verfügbarkeit und Verwaltbarkeit wichtiger Ressourcen. Dadurch gewährleisten Sie Geschäftskontinuität, schützen die Datenintegrität und reduzieren ungeplante Ausfallzeiten bei kritischen Workloads.

Mehrere Clustering-Optionen

SUSE Linux Enterprise High Availability-Cluster können auf unterschiedliche Weise konfiguriert werden:

- Lokale Cluster: Ein einzelner Cluster an einem Standort (z. B. alle Knoten befinden sich in einem Rechenzentrum). Die Netzwerklatenz ist minimal. Auf den Speicher wird in der Regel von allen Knoten synchron zugegriffen.
- Metro-Cluster („ausgedehnte“ Cluster): Ein einzelner Cluster, der sich über mehrere Gebäude oder Rechenzentren erstrecken kann, wobei alle Standorte über Fibre Channel verbunden sind. Die Netzwerklatenz ist in der Regel gering. Der Speicher wird häufig durch Spiegelung oder synchrone Reproduktion reproduziert.
- Hybride Cluster: Virtuelle Server können zusammen mit physischen Servern geclustert werden. Dies verbessert die Dienstverfügbarkeit und die Ressourcennutzung.

Wichtig: Keine Unterstützung für gemischte Architekturen

Cluster mit gemischten Architekturen werden nicht unterstützt. Alle Knoten im selben Cluster müssen dieselbe Prozessorplattform haben: AMD64/Intel 64, IBM Z oder POWER.

Flexibles und skalierbares Ressourcenmanagement

SUSE Linux Enterprise High Availability unterstützt Cluster mit bis zu 32 Knoten. Ressourcen können automatisch auf einen anderen Knoten verschoben werden, wenn der aktuelle Knoten ausfällt, oder sie können manuell verschoben werden, um Hardwarefehler zu beheben oder die Workload auszugleichen. Ressourcen können auch so konfiguriert werden, dass sie zu einem bestimmten Zeitpunkt wieder auf die reparierten Knoten verschoben werden. Der Cluster kann Ressourcen auf der Grundlage konfigurierbarer Regeln anhalten und starten.

Breite Palette an Ressourcenagenten

Der Cluster verwaltet die Ressourcen über *Ressourcenagenten* (RAs). SUSE Linux Enterprise High Availability unterstützt viele verschiedene Ressourcenagenten, die für die Verwaltung bestimmter Anwendungs- oder Diensttypen konzipiert sind, einschließlich Apache, IPv4, IPv6, NFS und viele mehr. Im Lieferumfang sind auch Ressourcenagenten für Anwendungen von Drittanbietern wie IBM WebSphere Application Server enthalten.

Speicher- und Datenreproduktion

SUSE Linux Enterprise High Availability unterstützt Fibre Channel- oder iSCSI-SANs (Storage Area Networks), sodass Sie Serverspeicher bei Bedarf dynamisch oder neu zuweisen können. Außerdem ist er mit GFS2 und dem Cluster-Logical Volume Manager (Cluster-LVM) ausgestattet. Verwenden Sie zur Datenreproduktion DRBD*, um die Daten einer Ressource vom aktiven Knoten auf einen Standby-Knoten zu spiegeln.

Unterstützung für virtualisierte Umgebungen

SUSE Linux Enterprise High Availability unterstützt das gemischte Clustering von physischen und virtuellen Linux-Servern. Der Cluster kann Ressourcen erkennen und verwalten, die auf virtuellen Servern und physischen Servern ausgeführt werden, und er kann auch virtuelle KVM-Rechner als Ressourcen verwalten.

Fehlerbehebungsfunktionen

Im Lieferumfang von SUSE Linux Enterprise High Availability ist Relax-and-Recover (ReaR) enthalten, ein Fehlerbehebungs-Framework, mit Sie Systeme sichern und wiederherstellen können.

Benutzerfreundliche Verwaltungswerkzeuge

SUSE Linux Enterprise High Availability umfasst Werkzeuge für die Konfiguration und Verwaltung:

- Die *CRM-Shell* (`crmsh`) ist eine Befehlszeilenschnittstelle für die Installation und Einrichtung von High Availability-Clustern, die Konfiguration von Ressourcen und die Ausführung von Überwachungs- und Verwaltungsaufgaben.
- *Hawk* ist eine webbasierte grafische Oberfläche zur Überwachung und Verwaltung von High Availability-Clustern. Darauf kann mit einem grafischen Webbrowser von jedem Linux- oder Nicht-Linux-Rechner aus, der eine Verbindung mit den Clusterknoten herstellen kann, zugegriffen werden.

1.3 Wie funktioniert High Availability?

Die folgende Abbildung zeigt einen Cluster mit drei Knoten. Auf jedem der Knoten ist ein Webserver installiert, der zwei Websites hostet. Alle Daten, Grafiken und Inhalte für diese Websites sind auf einem freigegebenen Festplatten-Subsystem gespeichert, das mit allen Knoten verbunden ist.

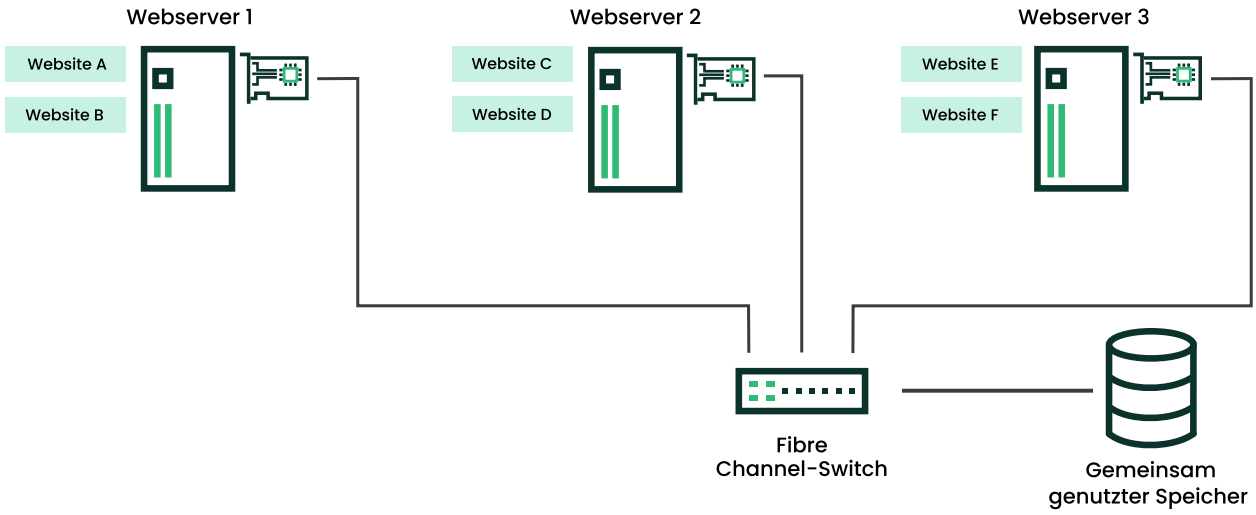

ABBILDUNG 1: CLUSTER MIT ZWEI KNOTEN

Im normalen Clusterbetrieb befindet sich jeder Knoten in ständiger Kommunikation mit den anderen Knoten im Cluster und überprüft zur Fehlererkennung in regelmäßigen Abständen die Ressourcen.

Die folgende Abbildung zeigt, wie der Cluster Ressourcen verschiebt, wenn Webserver 1 aufgrund von Hardware- oder Softwareproblemen ausfällt.

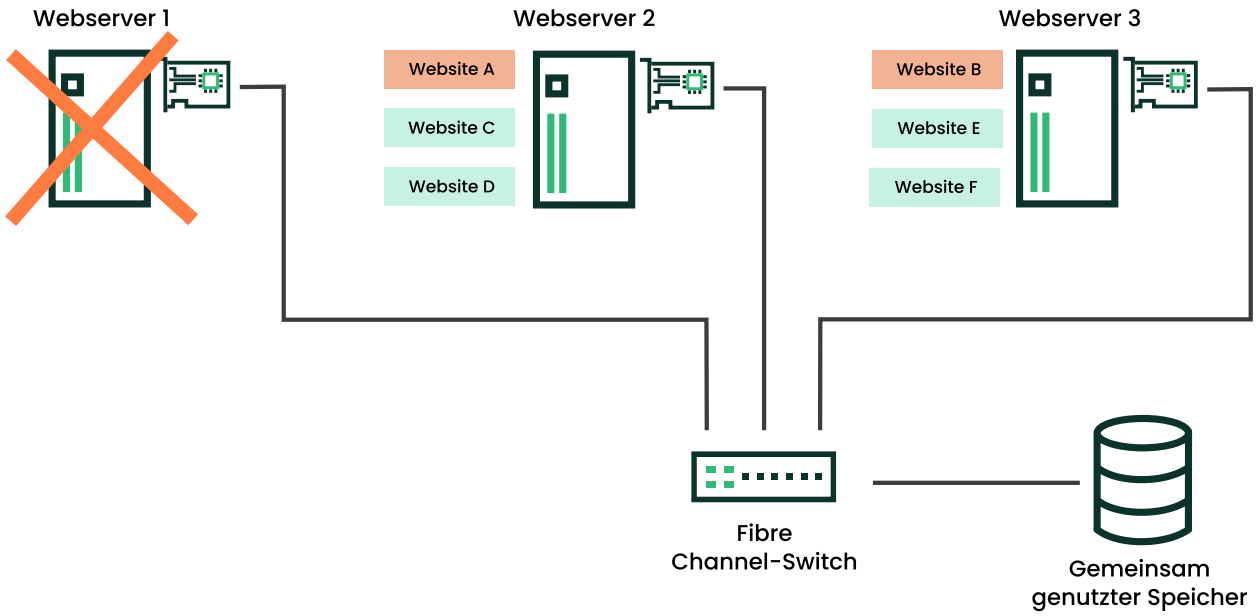

ABBILDUNG 2: CLUSTER MIT DREI KNOTEN NACH AUSFALL EINES KNOTENS

Website A wird auf Webserver 2 und Website B auf Webserver 3 verschoben. Die Websites sind weiterhin verfügbar und werden gleichmäßig auf die verbleibenden Clusterknoten verteilt.

Als Webserver 1 ausfiel, hat die High Availability-Software die folgenden Aufgaben ausgeführt:

1. Es wurde ein Fehler festgestellt und überprüft, ob Webserver 1 wirklich ausgefallen ist.
2. Die gemeinsam genutzten Datenverzeichnisse, die auf Webserver 1 eingebunden waren, wurden auf Webserver 2 und 3 neu eingebunden.
3. Die auf Webserver 1 ausgeführten Anwendungen wurden auf Webserver 2 und Webserver 3 neu gestartet.
4. Die Zertifikate und IP-Adressen wurden an die Webserver 2 und 3 übertragen.

In diesem Beispiel erfolgte der Failover schnell und die Benutzer konnten innerhalb von Sekunden wieder auf die Websites zugreifen – und zwar in der Regel, ohne sich erneut anmelden zu müssen.

Wenn Webserver 1 in den normalen Betriebszustand zurückkehrt, können Website A und Website B automatisch wieder auf Webserver 1 verschoben werden „(Failback“). Dadurch entsteht eine gewisse Ausfallzeit. Alternativ können Sie die Ressourcen so konfigurieren, dass nur zu einem bestimmten Zeitpunkt ein Failback erfolgt, der eine minimale Unterbrechung des Diensts verursacht.

2 Kernkonzepte

In diesem Abschnitt werden die Kernkonzepte von SUSE Linux Enterprise High Availability erläutert.

2.1 Cluster und Knoten

Ein *Hochverfügbarkeits*-Cluster ist eine Gruppe von Servern, die zusammenarbeiten, um die Verfügbarkeit von Anwendungen oder Diensten sicherzustellen. Server, die als Mitglieder des Clusters konfiguriert sind, werden als *Knoten* bezeichnet. Fällt ein Knoten aus, können die auf ihm ausgeführten Ressourcen auf einen anderen Knoten im Cluster verschoben werden. SUSE Linux Enterprise High Availability unterstützt Cluster mit bis zu 32 Knoten. Cluster fallen jedoch in der Regel in eine von zwei Kategorien: Cluster mit zwei Knoten oder Cluster mit einer ungeraden Anzahl von Knoten (in der Regel drei oder fünf).

2.2 Kommunikationskanäle

Die interne Clusterkommunikation wird von *Corosync* übernommen. Die Corosync Cluster Engine ist ein Gruppenkommunikationssystem, das Messaging-, Mitgliedschafts- und Quorum-Informationen zum Cluster bereitstellt. In SUSE Linux Enterprise High Availability 16 verwendet Corosync „kronosnet“ (`knet`) als Standard-Transportprotokoll.

Es wird dringend empfohlen, mindestens zwei Kommunikationskanäle für den Cluster zu konfigurieren. Die bevorzugte Methode ist die Verwendung von Netzwerkgeräte-Bonding. Wenn Sie kein Netzwerk-Bonding verwenden können, können Sie in Corosync einen redundanten Kommunikationskanal einrichten (auch bekannt als zweiter „Ring“).

2.3 Ressourcenmanagement

In einem High Availability-Cluster werden die Anwendungen und Dienste, die hochverfügbar sein müssen, als *Ressourcen* bezeichnet. Bei Clusterressourcen kann es sich um Websites, E-Mail-Server, Datenbanken, Dateisysteme, virtuelle Rechner und andere serverbasierte Anwendungen oder Dienste handeln, die Sie den Benutzern stets zur Verfügung stellen möchten. Sie können Ressourcen bei Bedarf starten, anhalten, überwachen und verschieben. Sie können auch angeben, ob bestimmte Ressourcen zusammen auf demselben Knoten ausgeführt werden sollen oder ob sie nacheinander gestartet und angehalten werden sollen. Wenn ein Clusterknoten ausfällt, gehen die auf ihm ausgeführten Ressourcen nicht verloren, sondern werden auf einen anderen Knoten *verschoben* (Failover).

In SUSE Linux Enterprise High Availability ist der *Cluster-Ressourcenmanager* (CRM) *Pacemaker*, der alle Clusterdienste verwaltet und koordiniert. Pacemaker verwendet *Ressourcenagenten* (RAs), um Ressourcen zu starten, anzuhalten und zu überwachen. Ein Ressourcenagent abstrahiert die von ihm verwaltete Ressource und präsentiert dem Cluster ihren Status. SUSE Linux Enterprise High Availability unterstützt viele verschiedene Ressourcenagenten, die für die Verwaltung bestimmter Anwendungs- oder Diensttypen konzipiert sind.

2.4 Knoten-Fencing

Bei einem *Split Brain-Szenario* werden Clusterknoten in zwei oder mehr Gruppen (oder *Partitionen*) unterteilt, die sich gegenseitig nicht erkennen können. Dies kann beispielsweise durch einen Hardware- oder Softwarefehler oder eine fehlerhafte Netzwerkverbindung erfolgen. Ein Split Brain-Szenario kann durch das *Fencing* (Zurücksetzen oder Ausschalten) eines oder mehre-

rer Knoten aufgelöst werden. Beim Fencing auf Knotenebene wird verhindert, dass ein ausgefallener Knoten auf gemeinsam genutzte Ressourcen zugreift und dass Clusterressourcen auf einem Knoten mit unsicherem Status ausgeführt werden.

SUSE Linux Enterprise High Availability verwendet STONITH als Fencing-Mechanismus für Knoten. Damit sie unterstützt werden, *müssen* alle SUSE Linux Enterprise High Availability-Cluster über mindestens ein STONITH-Gerät verfügen. Für kritische Workloads wird die Verwendung von zwei oder drei STONITH-Geräten empfohlen: Als STONITH-Gerät kann entweder ein physisches Gerät (ein Netzschalter) oder ein Softwaremechanismus (SBD in Kombination mit einem Watchdog) verwendet werden.

2.5 Quorum-Bestimmung

Wenn die Kommunikation zwischen einem oder mehreren Knoten und dem Rest des Clusters ausfällt (ein *Split Brain-Szenario*, entsteht eine *Clusterpartition*. Die Knoten können nur mit anderen Knoten in derselben Partition kommunizieren und erkennen die anderen Knoten nicht. Eine Clusterpartition hat ein *Quorum* (oder ist „quorumfähig“), wenn sie über die Mehrheit der Knoten (oder „Stimmen“) verfügt. Dies wird durch die *Berechnung des Quorums* bestimmt. Das Quorum muss berechnet werden, um das *Fencing* von nicht quorumfähigen Knoten zu ermöglichen.

Corosync berechnet das Quorum anhand der folgenden Formel:

```
N ≥ C/2 + 1

N = minimum number of operational nodes
C = number of cluster nodes
```

Ein Cluster mit fünf Knoten benötigt beispielsweise mindestens drei betriebsbereite Knoten, um das Quorum aufrechtzuerhalten.

Bei Clustern mit einer geraden Anzahl von Knoten, insbesondere bei Clustern mit zwei Knoten, kann es sein, dass die Anzahl der Knoten in jeder Partition gleich ist und daher keine Mehrheit besteht. Um diese Situation zu vermeiden, können Sie den Cluster so konfigurieren, dass er QDevice in Kombination mit QNetd verwendet. QNetd ist ein *Vermittler*, der außerhalb des Clusters ausgeführt wird. Er kommuniziert mit dem QDevice-Daemon, der auf den Clusterknoten ausgeführt wird, um eine konfigurierbare Anzahl von Stimmen für die Berechnung des Quorums bereitzustellen. Dadurch kann ein Cluster mehr Knotenausfälle verkraften, als die Standard-Quorum-Regeln zulassen.

2.6 Speicher- und Datenreproduktion

High Availability-Cluster können ein gemeinsam genutztes Festplatten-Subsystem enthalten, das über Fibre Channel oder iSCSI verbunden ist. Bei Ausfall eines Knotens werden die zuvor auf diesem ausgefallenen Knoten eingebundenen Festplattenverzeichnisse automatisch auf einem anderen Knoten im Cluster eingebunden. Auf diese Weise können die Benutzer kontinuierlich auf die Verzeichnisse des gemeinsamen genutzten Festplatten-Subsystems zugreifen. Zu den auf der gemeinsam genutzten Festplatte gespeicherten Inhalten können Daten, Anwendungen und Dienste gehören.

Die folgende Abbildung zeigt, wie eine typische Konfiguration eines Fibre Channel-Clusters aussehen könnte. Die grünen Linien stellen Verbindungen mit einem Ethernet-Netzschalter dar, der einen Knoten neu starten kann, wenn bei einer Ping-Anfrage ein Fehler auftritt.

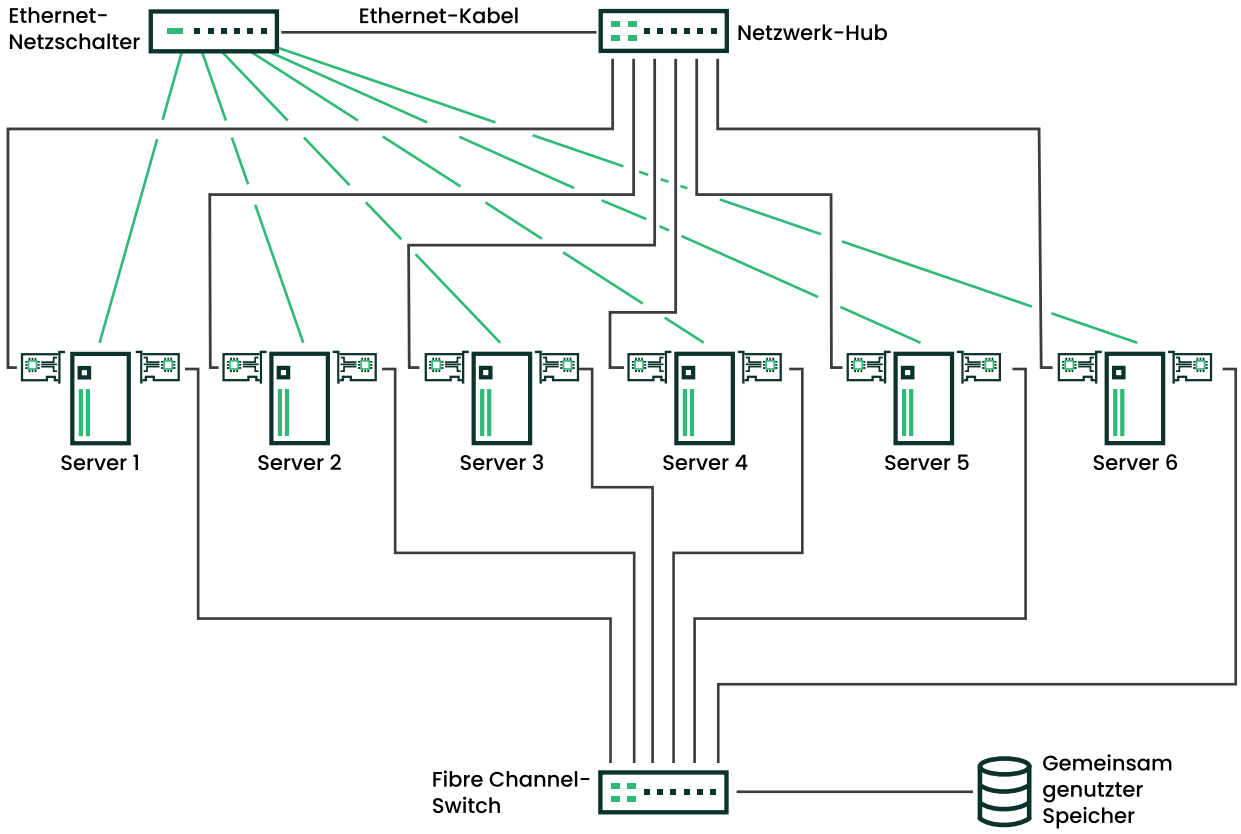

ABBILDUNG 3: TYPISCHE FIBRE CHANNEL-CLUSTERKONFIGURATION

Die folgende Abbildung zeigt, wie eine typische Konfiguration eines iSCSI-Clusters aussehen könnte.

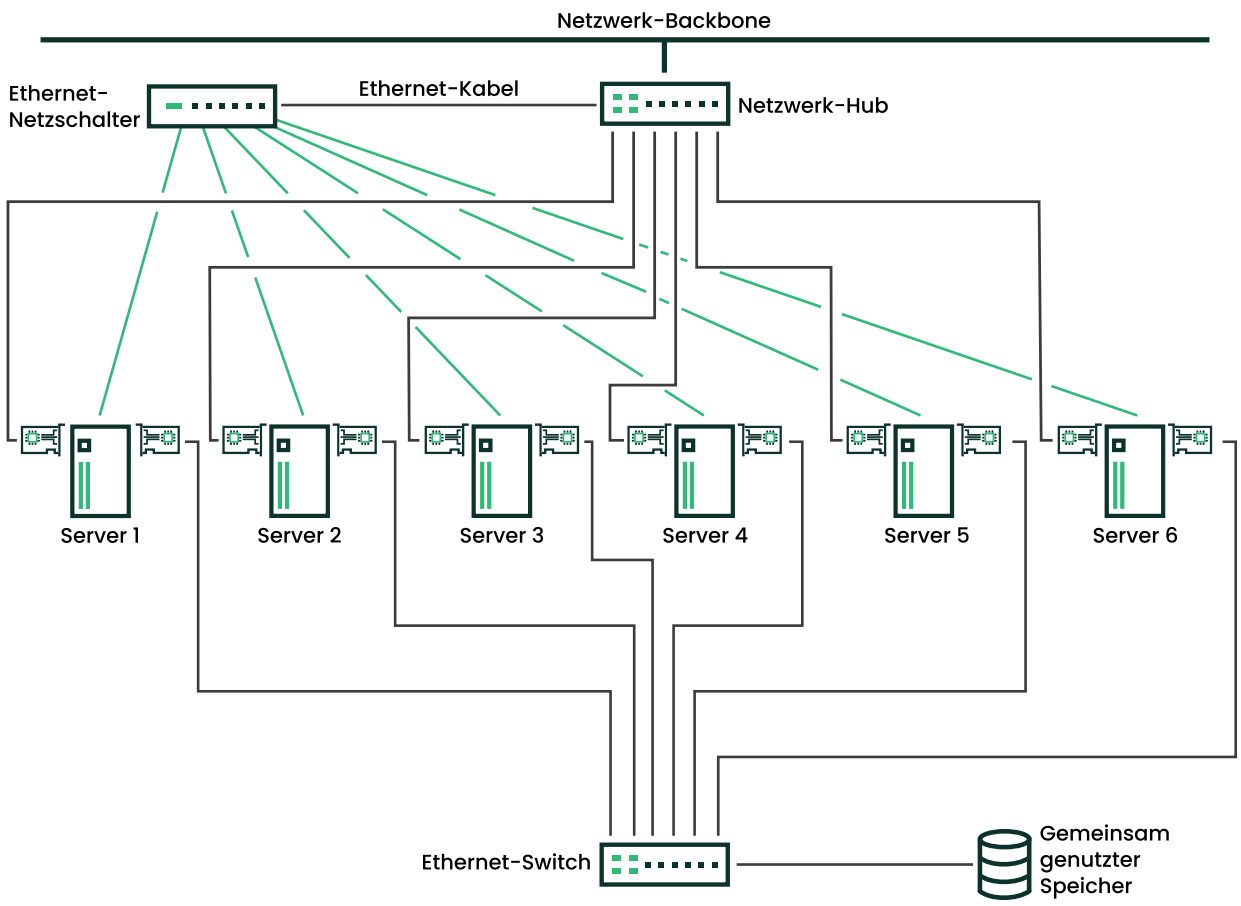

ABBILDUNG 4: TYPISCHE KONFIGURATION EINES ISCSI-CLUSTERS

Obwohl die meisten Cluster ein gemeinsam genutztes Festplatten-Subsystem enthalten, können Sie auch einen Cluster ohne ein gemeinsam genutztes Festplatten-Subsystem erstellen. Die folgende Abbildung zeigt, wie ein Cluster ohne ein gemeinsam genutztes Festplatten-Subsystem aussehen könnte.

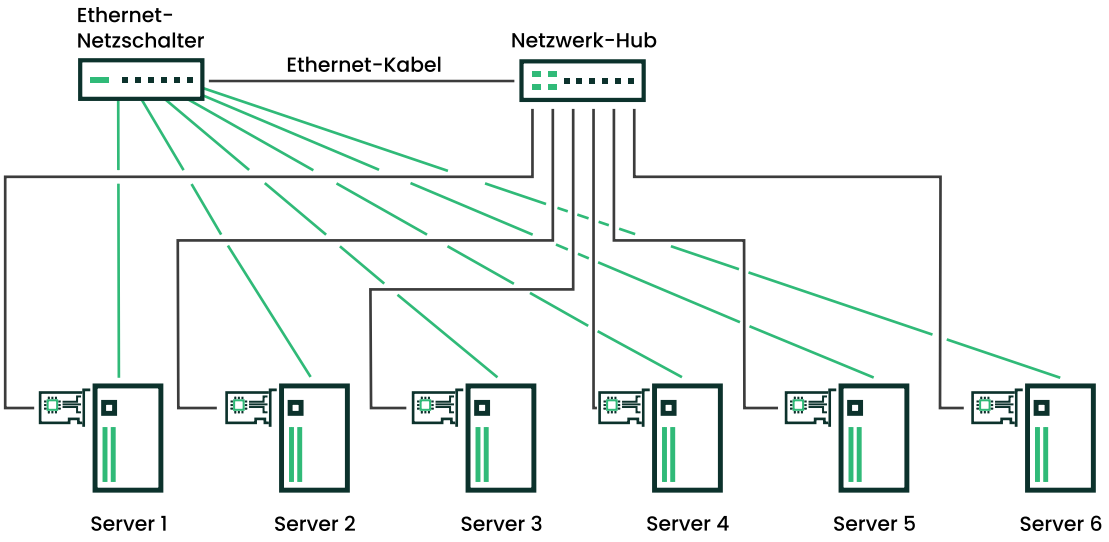

ABBILDUNG 5: TYPISCHE CLUSTERKONFIGURATION OHNE GEMEINSAM GENUTZTEN SPEICHER

3 Überblick über die Architektur

In diesem Abschnitt werden die Architektur eines SUSE Linux Enterprise High Availability-Clusters und die Interoperabilität der verschiedenen Komponenten erläutert.

3.1 Architekturschichten

SUSE Linux Enterprise High Availability verfügt über eine mehrschichtige Architektur. Die folgende Abbildung zeigt die verschiedenen Schichten und ihre jeweiligen Komponenten.

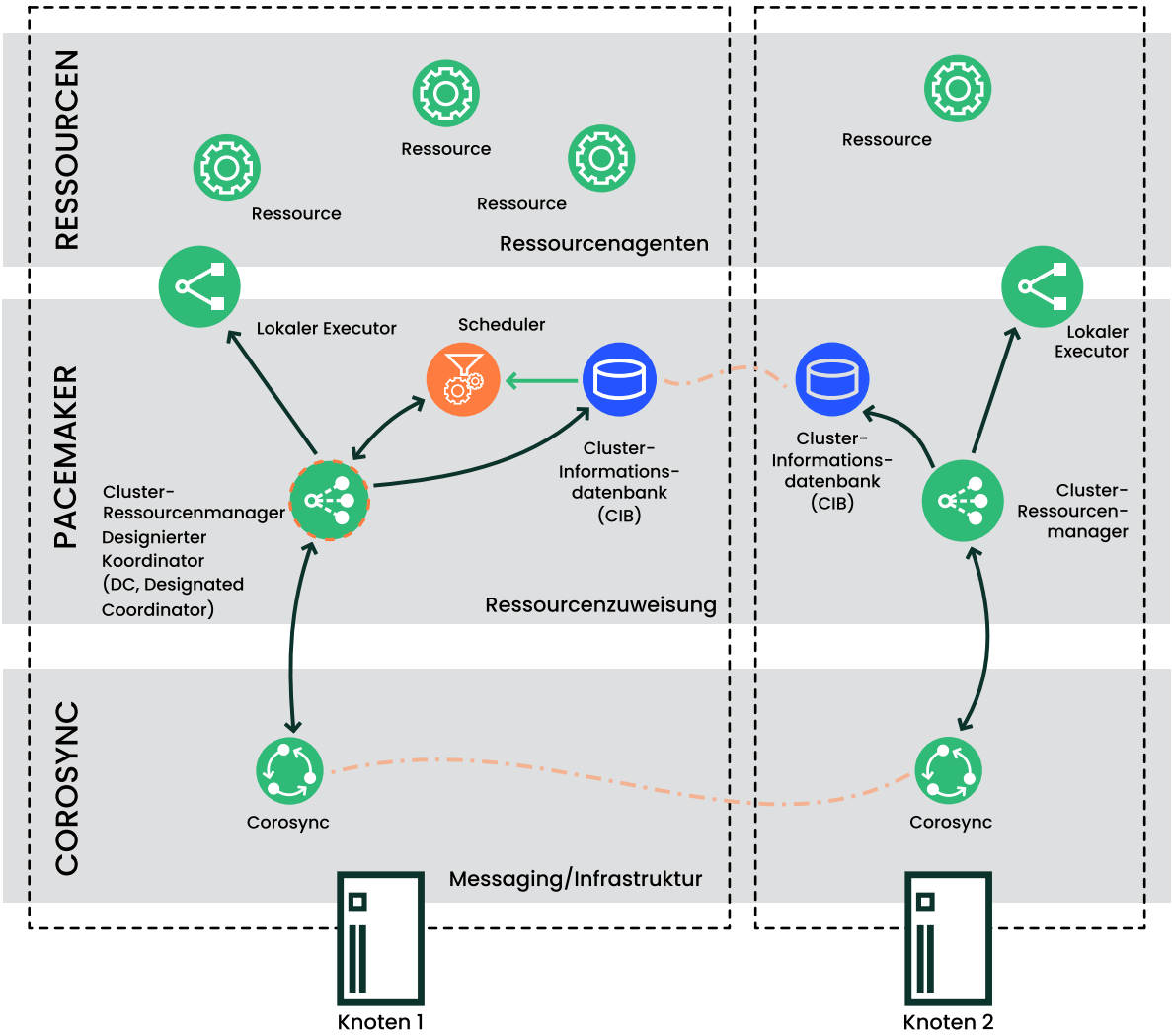

ABBILDUNG 6: ARCHITEKTUR EINES SUSE LINUX ENTERPRISE HIGH AVAILABILITY-CLUSTERS

Mitgliedschaft und Messaging (Corosync)

Die Corosync Cluster Engine ist ein Gruppenkommunikationssystem, das Messaging-, Mitgliedschafts- und Quorum-Informationen zum Cluster bereitstellt.

Cluster-Ressourcenmanager (Pacemaker)

Pacemaker ist der Cluster-Ressourcenmanager, der auf Ereignisse im Cluster reagiert. Ereignisse können Knoten sein, die dem Cluster beitreten oder ihn verlassen, der Ausfall von Ressourcen oder geplante Aktivitäten wie Wartungen. Der Daemon `pacemakerd` startet und überwacht alle anderen zugehörigen Daemons.

Die folgenden Komponenten sind auch Teil der Pacemaker-Schicht:

- **Cluster-Informationsdatenbank (CIB)**

Pacemaker verwaltet die Cluster-Informationsdatenbank (CIB) auf jedem Knoten. Dies ist eine XML-Darstellung der gesamten Clusterkonfiguration (einschließlich Clusteroptionen, -knoten, -ressourcen, -einschränkungen und der Beziehungen zueinander). Die CIB spiegelt auch den aktuellen Clusterstatus wider. Der CIB-Manager (`pacemaker-based`) sorgt dafür, dass die CIB im gesamten Cluster synchron bleibt, und verarbeitet das Lesen und Schreiben der Clusterkonfiguration und den Status.

- **Designierter Koordinator (DC, Designated Coordinator)**
  Der Daemon `pacemaker-controld` ist der Cluster-Controller, der alle Aktionen koordiniert. Dieser Daemon verfügt über eine Instanz auf jedem Clusterknoten, aber nur eine Instanz wird als designierter Koordinator gewählt. Der designierte Koordinator wird gewählt, wenn die Clusterdienste starten oder wenn der aktuelle designierte Koordinator ausfällt oder den Cluster verlässt. Der designierte Koordinator entscheidet, ob eine clusterweite Änderung vorgenommen werden muss, z. B. das Fencing eines Knotens oder das Verschieben von Ressourcen.

- **Scheduler**
  Der Scheduler wird auf jedem Knoten als Daemon `pacemaker-schedulerd` ausgeführt, ist aber nur auf dem designierten Koordinator aktiv. Wenn ein Clusterwechsel erforderlich ist, berechnet der Scheduler den erwarteten nächsten Status des Clusters und bestimmt, welche Aktionen geplant werden müssen, um diesen Status zu erreichen.

Lokaler Executor

Der lokale Executor befindet sich zwischen der Pacemaker-Schicht und der Ressourcenschicht auf jedem Knoten. Mit dem Daemon `pacemaker-execd` kann Pacemaker Ressourcen starten, anhalten und überwachen.

Ressourcen und Ressourcenagenten

In einem High Availability-Cluster werden die Dienste, die hochverfügbar sein müssen, als *Ressourcen* bezeichnet. Ressourcenagenten (RAs) sind Skripte, die Clusterressourcen starten, anhalten und überwachen.

3.2 Prozessfluss

Viele Aktionen, die im Cluster ausgeführt werden, führen zur einer clusterweiten Änderung, z. B. das Hinzufügen oder Entfernen einer Clusterressource oder das Ändern von Ressourcenbeschränkungen. Im folgenden Beispiel wird erläutert, was im Cluster geschieht, wenn Sie eine solche Aktion ausführen:

BEISPIEL 1: CLUSTERPROZESS BEIM HINZUFÜGEN EINER NEUEN RESSOURCE

1. Sie verwenden die CRM-Shell oder Hawk, um eine neue Clusterressource hinzuzufügen. Sie können dies von jedem Knoten im Cluster aus vornehmen. Durch das Hinzufügen der Ressource wird die CIB geändert.
2. Die Änderung der CIB wird auf allen Clusterknoten reproduziert.
3. Anhand der Informationen in der CIB berechnet `pacemaker-schedulerd` den idealen Status des Clusters und wie dieser erreicht werden sollte. Es gibt eine Liste mit Anweisungen an den designierten Koordinator weiter.
4. Der designierte Koordinator sendet Befehle über Corosync, die von den `pacemaker-controld`-Instanzen auf den anderen Knoten empfangen werden.
5. Jeder Knoten verwendet seinen lokalen Executor (`pacemaker-execd`), um Änderungen an den Ressourcen vorzunehmen. Der Daemon `pacemaker-execd` erkennt Cluster nicht und interagiert direkt mit den Ressourcenagenten.
6. Alle Knoten melden die Ergebnisse ihrer Vorgängen wieder an den designierten Koordinator.
7. Wenn ein Fencing erforderlich ist, ruft `pacemaker-fenced` den Fencing-Agenten auf, um das Fencing für den Knoten vorzunehmen.
8. Nachdem der designierte Koordinator zu dem Schluss gekommen ist, dass alle erforderlichen Vorgänge erfolgreich ausgeführt wurden, kehrt der Cluster in den Leerlauf zurück und wartet auf weitere Ereignisse.
9. Wenn ein Vorgang nicht wie geplant ausgeführt wurde, wird `pacemaker-schedulerd` erneut aufgerufen, wobei die neuen Informationen in der CIB gespeichert werden.

4 Installationsoptionen

Dieser Abschnitt beschreibt die verschiedenen Optionen zum Installieren eines SUSE Linux Enterprise High Availability-Clusters und enthält Links zu den verfügbaren Installationshandbüchern.

In den folgenden Kurzanleitungen wird erklärt, wie Sie einen minimalen Cluster mit Standardeinstellungen einrichten:

Installing a basic two-node High Availability cluster (https://susedoc.github.io/doc-modular/main/html/HA-installing-three-node-cluster/index.html)

In dieser Kurzanleitung wird erklärt, wie Sie einen einfachen High Availability-Cluster mit zwei Knoten mit QDevice, festplattenlosem SBD und Software-Watchdog einrichten. Für diese Einrichtung ist QDevice erforderlich, damit festplattenloses SBD Split Brain-Szenarien für den Cluster mit zwei Knoten verarbeiten kann.

Installing a basic three-node High Availability cluster (https://susedoc.github.io/doc-modular/main/html/HA-installing-two-node-cluster/index.html)

In dieser Kurzanleitung wird erklärt, wie Sie einen einfachen High Availability-Cluster mit drei Knoten mit festplattenlosem SBD und Software-Watchdog einrichten. Für diese Einrichtung sind drei Knoten erforderlich, damit festplattenloses SBD Split Brain-Szenarien ohne Hilfe von QDevice verarbeiten kann.

5 Weitere Informationen

Weitere Informationen zu SUSE Linux Enterprise High Availability finden Sie in den folgenden Ressourcen:

https://clusterlabs.org/

Das Upstream-Projekt für viele der Komponenten von SUSE Linux Enterprise High Availability.

https://www.clusterlabs.org/pacemaker/doc/

Dokumentation zu Pacemaker. Informationen zu SUSE Linux Enterprise High Availability 16 finden Sie in der Dokumentation zu Pacemaker 3.

6 Rechtliche Hinweise

Copyright © 2006–2025 , SUSE LLC und Mitwirkende. Alle Rechte vorbehalten.

Es wird die Genehmigung erteilt, dieses Dokument unter den Bedingungen der GNU Free Documentation License, Version 1.2 oder (optional) Version 1.3 zu vervielfältigen, zu verbreiten und/oder zu verändern; die unveränderlichen Abschnitte hierbei sind der Urheberrechtshinweis und die Lizenzbedingungen. Eine Kopie dieser Lizenz (Version 1.2) finden Sie in Abschnitt „GNU Free Documentation License“.

Die SUSE Marken finden Sie in https://www.suse.com/company/legal/. Alle anderen Marken von Drittanbietern sind Besitz ihrer jeweiligen Eigentümer. Markensymbole (®, ™ usw.) kennzeichnen Marken von SUSE und ihren Tochtergesellschaften. Sternchen (*) kennzeichnen Marken von Drittanbietern.

Alle Informationen in diesem Buch wurden mit größter Sorgfalt zusammengestellt. Auch hierdurch kann jedoch keine hundertprozentige Richtigkeit gewährleistet werden. Weder SUSE LLC noch ihre Tochtergesellschaften noch die Autoren noch die Übersetzer können für mögliche Fehler und deren Folgen haftbar gemacht werden.

A GNU Free Documentation License

Copyright (C) 2000, 2001, 2002 Free Software Foundation, Inc. 51 Franklin St, Fifth Floor, Boston, MA 02110-1301 USA. Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or non-commercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input

to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or non-commercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties--for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See https://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

ADDENDUM: How to use this License for your documents

```
Copyright (c) YEAR YOUR NAME.
Permission is granted to copy, distribute and/or modify this document
under the terms of the GNU Free Documentation License, Version 1.2
or any later version published by the Free Software Foundation;
with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.
A copy of the license is included in the section entitled “GNU
Free Documentation License”.
```

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the “with…Texts.” line with this:

```
with the Invariant Sections being LIST THEIR TITLES, with the
Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.
```

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.

Hochverfügbarkeit – Glossar

aktiv/aktiv, aktiv/passiv

Wie die Ressourcen auf den Knoten ausgeführt werden. „Aktiv/Passiv“ bedeutet, dass die Ressourcen nur auf dem aktiven Knoten ausgeführt werden, aber auf den passiven Knoten verschoben werden können, wenn der aktive Knoten ausfällt. „Aktiv/Aktiv“ bedeutet, dass alle Knoten gleichzeitig aktiv sind und Ressourcen auf jedem Knoten des Clusters ausgeführt werden (und dorthin verschoben werden) können.

Auslastung

Teilt dem CRM mit, welche Kapazität eine bestimmte *Ressourcen* von einem Knoten benötigt.

Bestehender Cluster

Der Begriff *bestehender Cluster* bezieht sich auf jeden Cluster, der aus mindestens einem Knoten besteht. Ein bestehender Cluster verfügt über eine grundlegende *Corosync*-Konfiguration, die die Kommunikationskanäle definiert, jedoch noch nicht unbedingt über eine Ressourcenkonfiguration.

cluster

Ein *Hochverfügbarkeits*-Cluster ist eine Gruppe von Servern (physisch oder virtuell), die in erster Linie dazu dient, die höchstmögliche Verfügbarkeit von Daten und Anwendungsdiensten sicherzustellen. Er ist nicht zu verwechseln mit einem *Hochleistungs*-Cluster, der die Anwendungslast teilt, um schnellere Ergebnisse zu erzielen.

Cluster-Informationsdatenbank (CIB, Cluster Information Base):

Eine XML-Darstellung der gesamten Clusterkonfiguration und des Clusterstatus (Clusteroptionen, -knoten, -ressourcen, -einschränkungen und die Beziehungen zueinander). Der CIB-Manager (`pacemaker-based`) sorgt dafür, dass die CIB im gesamten Cluster synchron bleibt, und verarbeitet Änderungsanfragen.

Cluster-Logical Volume Manager (Cluster-LVM)

Der Begriff *Cluster-LVM* gibt an, dass der LVM in einer Clusterumgebung verwendet wird. Dies erfordert Konfigurationsanpassungen zum Schutz der LVM-Metadaten in einem gemeinsam genutzten Speicher.

Cluster-Ressourcenmanager (CRM)

Die Verwaltungseinheit, die für die Koordinierung aller nicht lokalen Interaktionen in einem High Availability-Cluster zuständig ist. SUSE Linux Enterprise High Availability verwendet *Pacemaker* als CRM. Es interagiert mit verschiedenen Komponenten: lokalen Executors auf ihrem eigenen Knoten und auf den anderen Knoten, nicht lokalen CRMs, administrativen Befehlen, der Fencing-Funktion und der Mitgliedschaftsschicht.

Clusterpartition

Eine Clusterpartition entsteht, wenn die Kommunikation zwischen einem oder mehreren Knoten und dem Rest des Clusters ausfällt. Die Knoten werden in Partitionen aufgeteilt, sind aber weiterhin aktiv. Sie können nur mit Knoten in derselben Partition kommunizieren und erkennen die anderen Knoten nicht. Dies wird als *Split Brain*-Szenario bezeichnet.

Clusterstapel

Die Gesamtheit der Softwaretechnologien und -komponenten, aus denen ein Cluster besteht.

Corosync

Corosync bietet zuverlässige Messaging-, Mitgliedschafts- und Quorum-Informationen zum Cluster. Dies wird von der Corosync Cluster Engine, einem Gruppenkommunikationssystem, übernommen.

`crmsh` (CRM-Shell)

Das Befehlszeilen-Dienstprogramm `crmsh` verwaltet den Cluster, die Knoten und die Ressourcen.

Csync2

Ein Synchronisierungswerkzeug zur Replikation von Konfigurationsdateien auf allen Knoten im Cluster.

Designierter Koordinator (DC, Designated Coordinator)

Der Daemon `pacemaker-controld` ist der Cluster-Controller, der alle Aktionen koordiniert. Dieser Daemon verfügt über eine Instanz auf jedem Clusterknoten, aber nur eine Instanz wird als designierter Koordinator gewählt. Der designierte Koordinator wird gewählt, wenn die Clusterdienste starten oder wenn der aktuelle designierte Koordinator ausfällt oder den Cluster verlässt. Der designierte Koordinator entscheidet, ob eine clusterweite Änderung vorgenommen werden muss, z. B. das Fencing eines Knotens oder das Verschieben von Ressourcen.

Disaster

Eine unerwartete Unterbrechung der kritischen Infrastruktur, die durch Natur, Menschen, Hardwarefehler oder Softwarefehler verursacht wird.

Distributed Lock Manager (DLM)

Distributed Lock Manager (DLM) koordiniert die Zugriffe auf gemeinsam genutzte Ressourcen in einem Cluster, z. B. durch die Verwaltung von Dateisperren in geclusterten Dateisystemen, um die Leistung und Verfügbarkeit zu verbessern.

DRBD

DRBD® ist ein Blockgerät zum Erstellen von High Availability-Clustern. Es repliziert Daten auf einem primären Gerät auf sekundäre Geräte in einer Art und Weise, mit der sichergestellt wird, dass alle Kopien der Daten identisch bleiben.

Failover

Tritt auf, wenn eine Ressource oder ein Knoten auf einem Rechner ausfällt und die betroffenen Ressourcen auf einen anderen Knoten verschoben werden.

Failover-Domäne

Eine benannte Teilmenge von Clusterknoten, die zur Ausführung einer Ressource berechtigt sind, wenn ein Knoten ausfällt.

Fehlerbehebung

Der Prozess, durch den eine Funktion nach einem Notfall wieder auf den normalen, stabilen Zustand zurückgesetzt wird.

Fehlerbehebungsplan

Eine Strategie zur Wiederherstellung nach einem Notfall mit möglichst geringen Auswirkungen auf die IT-Infrastruktur.

Fencing

Verhindert den Zugriff auf eine gemeinsam genutzte Ressource durch isolierte oder ausgefallene Clustermitglieder. Es gibt zwei Arten von Fencing: Fencing auf *Ressourcenebene* und Fencing auf *Knotenebene*. Beim Fencing auf Ressourcenebene wird der exklusive Zugriff auf eine Ressource sichergestellt. Beim Fencing auf Knotenebene wird verhindert, dass ein ausgefallener Knoten auf gemeinsam genutzte Ressourcen zugreift und dass Ressourcen auf einem Knoten mit unsicherem Status ausgeführt werden. Dies geschieht in der Regel durch Zurücksetzen oder Ausschalten des Knotens.

GFS2

Global File System 2 (GFS2) ist ein gemeinsam genutztes Festplattendateisystem für Linux-Computer-Cluster. GFS2 ermöglicht allen Knoten den direkten und gleichzeitigen Zugriff auf denselben gemeinsam genutzten Blockspeicher. GFS2 hat keinen getrennten Betriebsmodus und keine Client- oder Serverrollen. Alle Knoten in einem GFS2-Cluster fungieren als Peers. GFS2 unterstützt bis zu 32 Clusterknoten. Für die Verwendung von GFS2 in einem Cluster muss die Hardware den Zugriff auf den gemeinsam genutzten Speicher zulassen und es ist ein Lock Manager erforderlich, um den Zugriff auf den Speicher zu steuern.

Gleichzeitigkeitsverletzung

Eine Ressource, die eigentlich nur auf einem Knoten des Clusters ausgeführt werden sollte, wird auf mehreren Knoten ausgeführt.

group

Ressourcengruppen enthalten mehrere Ressourcen, die sich an einem gemeinsamen Ort befinden, nacheinander gestartet und in umgekehrter Reihenfolge angehalten werden müssen.

Hawk (HA-Webkonsole)

Eine benutzerfreundliche webbasierte Oberfläche zum Überwachen und Verwalten eines High Availability-Clusters von Linux- oder Nicht-Linux-Rechnern aus. Auf Hawk kann von jedem Rechner aus, der eine Verbindung mit den Clusterknoten herstellen kann, über einen grafischen Webbrowser zugegriffen werden.

Heuristik

QDevice unterstützt die Verwendung einer Reihe von Befehlen (*Heuristiken*), die lokal beim Start von Clusterdiensten, bei einer Änderung der Clustermitgliedschaft, bei einer erfolgreichen Verbindung mit dem *QNetd*-Server oder optional zu regelmäßigen Zeiten ausgeführt werden. Das Ergebnis wird in Berechnungen verwendet, um zu bestimmen, welche Partition über *Quorum* verfügen sollte.

Hochstufbarer Klon

Hochstufbare Klone sind eine besondere Art von *Klon*-Ressourcen, die hochgestuft werden können. Aktive Instanzen dieser Ressourcen werden in zwei Status unterteilt: hochgestuft und nicht hochgestuft (auch bekannt als „aktiv und passiv“ oder „primär und sekundär“).

Klon

Ein *Klon* ist eine identische Kopie eines vorhandenen Knotens, um die Bereitstellung mehrerer Knoten zu vereinfachen.

Im Kontext einer Cluster-*Ressourcen* ist ein Klon eine Ressource, die auf mehreren Knoten aktiv sein kann. Jede Ressource kann geklont werden, wenn ihr Ressourcenagent dies unterstützt.

knet (Kronosnet)

Eine Netzabstraktionsschicht, die Redundanz, Sicherheit, Fehlertoleranz und schnelles Failover bei Netzverbindungen unterstützt. In SUSE Linux Enterprise High Availability 16 ist *knet* das Standard-Transportprotokoll für die *Corosync*-Kommunikationskanäle.

Knoten

Jeder Server (physisch oder virtuell), der Mitglied eines Clusters ist.

Koexistenzeinschränkung

Eine Art *Ressourcenbeschränkung*, die angibt, welche Ressourcen auf einem Knoten zusammen ausgeführt werden können und welche nicht.

Lokaler Cluster

Ein einzelner Cluster an einem Standort (z. B. alle Knoten befinden sich in einem Rechenzentrum). Die Netzwerklatenz ist minimal. Auf den Speicher wird in der Regel von allen Knoten synchron zugegriffen.

Lokaler Executor

Der lokale Executor befindet sich zwischen *Pacemaker* und den Ressourcen auf jedem Knoten. Über den Daemon `pacemaker-execd` kann Pacemaker Ressourcen starten, anhalten und überwachen.

Meta-Attribute (Ressourcenoptionen)

Parameter, die dem *Cluster-Ressourcenmanager (CRM)* mitteilen, wie eine bestimmte *Ressourcen* zu behandeln ist. Sie können zum Beispiel die Priorität oder die Zielrolle einer Ressource definieren.

Metro-Cluster

Ein einzelner Cluster, der sich über mehrere Gebäude oder Rechenzentren erstrecken kann, wobei alle Standorte über Fibre Channel verbunden sind. Die Netzwerklatenz ist in der Regel gering. Der Speicher wird häufig durch Spiegelung oder synchrone Reproduktion reproduziert.

Netzwerkgeräte-Bonding (bevorzugt)

Beim Netzwerkgeräte-Bonding werden zwei oder mehr Netzwerkschnittstellen in einem einzigen Bond-Gerät kombiniert, um die Bandbreite zu erhöhen und/oder Redundanz bereitzustellen. Bei Verwendung von *Corosync* wird das Bond-Gerät nicht von der Clustersoftware verwaltet. Daher muss das Bond-Gerät auf jedem Clusterknoten konfiguriert werden, der möglicherweise darauf zugreifen muss.

Pacemaker

Pacemaker ist der *Cluster-Ressourcenmanager (CRM)* in SUSE Linux Enterprise High Availability, also das „Gehirn“, das auf Ereignisse im Cluster reagiert. Ereignisse können Knoten sein, die dem Cluster beitreten oder ihn verlassen, der Ausfall von Ressourcen oder geplante Aktivitäten wie Wartungen. Der Daemon `pacemakerd` startet und überwacht alle anderen zugehörigen Daemons.

Parameter (Instanzattribute)

Die Parameter bestimmen, welche Instanz eines Diensts die *Ressourcen* steuert.

Primitiv

Eine primitive Ressource ist der einfachste Typ einer Clusterressource.

QDevice

QDevice und *QNetd* beteiligen sich an *Quorum*-Entscheidungen. Der Daemon `corosync-qdevice` wird auf jedem Clusterknoten ausgeführt und kommuniziert mit QNetd, um eine konfigurierbare Anzahl an Stimmen bereitzustellen, sodass ein Cluster mehr Knotenausfälle bewältigen kann, als die Standard-Quorum-Regeln erlauben.

QNetd

QNetd ist ein *Vermittler*, der außerhalb des Clusters ausgeführt wird. Der Daemon `corosync-qnetd` stellt dem Daemon `corosync-qdevice` auf jedem Knoten eine Stimme bereit, damit dieser sich an Quorum-Entscheidungen beteiligen kann.

Quorum

Eine *Clusterpartition* ist als Quorum (*quorumfähig*) definiert, wenn sie über die Mehrheit der Knoten (oder „Stimmen“) verfügt. Das Quorum steht genau für eine Partition. Dies ist Teil des Algorithmus, um zu verhindern, dass mehrere nicht verbundene Partitionen oder Knoten („Split Brain“) fortgesetzt werden und Daten- sowie Dienstbeschädigungen verursachen. Das Quorum ist eine Voraussetzung für das Fencing, das wiederum sicherstellt, dass das Quorum eindeutig ist.

ReaR (Relax and Recover)

Ein Administrator-Werkzeugsatz zum Erstellen von *Fehlerbehebung*-Images.

Reihenfolgeneinschränkung

Eine Art *Ressourcenbeschränkung*, die die Reihenfolge der Aktionen definiert.

Ressourcen

Jede Art von Dienst oder Anwendung, die *Pacemaker* bekannt ist, z. B. eine IP-Adresse, ein Dateisystem oder eine Datenbank. Der Begriff *Ressource* wird auch für *DRBD* verwendet, wo er eine Gruppe aus Blockgeräten bezeichnet, die eine gemeinsame Verbindung für die Reproduktion nutzen.

Ressourcenagent (RA)

Ein Skript, das als Proxy für die Verwaltung einer *Ressourcen* fungiert (z. B. zum Starten, Anhalten oder Überwachen einer Ressource). SUSE Linux Enterprise High Availability unterstützt verschiedene Arten von Ressourcenagenten.

Ressourcenbeschränkung

Ressourcenbeschränkungen geben an, auf welchen Clusterknotenressourcen ausgeführt werden können, in welcher Reihenfolge Ressourcen geladen werden und von welchen anderen Ressourcen eine bestimmte Ressource abhängig ist.

Weitere Informationen finden Sie unter *Koexistenzeinschränkung*, *Standorteinschränkung* und *Reihenfolgeneinschränkung*.

Ressourcensatz

Als alternatives Format zum Definieren von Standort-, Koexistenz- oder Reihenfolgeneinschränkungen können Sie *Ressourcensätze* verwenden, in denen primitive Ressourcen in einem Satz gruppiert sind. Beim Erstellen einer Einschränkung können Sie mehrere Ressourcen angeben, für die die Einschränkung gelten soll.

Ressourcenvorlage

Um die Erstellung vieler Ressourcen mit ähnlichen Konfigurationen zu erleichtern, können Sie eine Ressourcenvorlage definieren. Nach dem Definieren kann in primitiven Ressourcen oder in bestimmten Arten von Einschränkungen auf sie verweisen werden. Wenn in einer primitiven Ressourcen auf eine Vorlage verwiesen wird, erbt die primitive Ressource alle in der Vorlage definierten Vorgänge, Instanzattribute (Parameter), Meta-Attribute und Nutzungsattribute.

SBD (STONITH Block Device)

SBD bietet einen *Fencing*-Mechanismus für Knoten durch den Austausch von Nachrichten über einen gemeinsam genutzten Blockspeicher. Alternativ kann es auch im festplattenlosen Modus verwendet werden. In jedem Fall ist auf jedem Knoten ein Hardware- oder Software-*Watchdog* erforderlich, um sicherzustellen, dass problembehaftete Knoten wirklich angehalten werden.

Scheduler

Der Scheduler ist als `pacemaker-schedulerd` implementiert. Wenn ein Clusterwechsel erforderlich ist, berechnet `pacemaker-schedulerd` den erwarteten nächsten Status des Clusters und bestimmt, welche Aktionen geplant werden müssen, um den nächsten Status zu erreichen.

Split Brain

Ein Szenario, bei dem die Clusterknoten in zwei oder mehr Gruppen aufgeteilt sind, die sich gegenseitig nicht erkennen (entweder durch einen Software- oder Hardwarefehler). *STONITH* verhindert ein Split Brain-Szenario, das den gesamten Cluster beeinträchtigt. Dies ist auch als Szenario mit *partitioniertem Cluster* bekannt.

Der Begriff *Split Brain* wird auch in *DRBD* verwendet, bedeutet aber, dass die Knoten unterschiedliche Daten enthalten.

SPOF (Single Point of Failure)

Eine beliebige Komponente eines Clusters, die bei einem Ausfall den Ausfall des gesamten Clusters auslöst.

Standort

Im Kontext eines gesamten Clusters kann sich der *Standort* auf den physischen Standort der Knoten beziehen (z. B. könnten sich alle Knoten im selben Rechenzentrum befinden). Im Kontext einer *Standorteinschränkung* bezieht sich der *Standort* auf die Knoten, auf denen eine Ressource ausgeführt werden kann oder nicht.

Standorteinschränkung

Eine Art *Ressourcenbeschränkung*, die die Knoten definiert, auf denen eine Ressource ausgeführt werden kann oder nicht.

STONITH

Ein Akronym für *Shoot the other Node in the Head* (Schieß dem anderen Knoten in den Kopf). Es bezieht sich auf den *Fencing*-Mechanismus, der einen problembehafteten Knoten abschaltet, um zu verhindern, dass er Probleme in einem Cluster verursacht. In einem *Pacemaker*-Cluster wird STONITH vom Fencing-Subsystem `pacemaker-fenced` verwaltet.

Umschaltung

Die geplante Verschiebung von Ressourcen auf andere Knoten in einem Cluster. Siehe auch *Failover*.

Vermittler

Ein *Vermittler* ist ein Rechner, der außerhalb des Clusters ausgeführt wird und eine zusätzliche Instanz für Clusterberechnungen bereitstellt. Zum Beispiel bietet *QNetd* eine Stimme an, damit sich *QDevice* an *Quorum*-Entscheidungen beteiligen kann.

Watchdog

Für *SBD (STONITH Block Device)* ist ein Watchdog erforderlich, um sicherzustellen, dass problembehaftete Knoten wirklich angehalten werden. SBD „füttert“ den Watchdog, indem es regelmäßig einen Dienstimpuls an ihn schreibt. Wenn SBD den Watchdog nicht mehr „füttert“, erzwingt die Hardware einen Neustart des Systems. Dies schützt vor Fehlern des SBD-Prozesses selbst, z. B. wenn er bei einem E/A-Fehler hängen bleibt.