

Red Hat Enterprise Linux 6

Überblick über die Cluster-Suite

Überblick über das Hochverfügbarkeits-Add-On für Red Hat Enterprise
Linux 6

Ausgabe 0



Rechtlicher Hinweis

Copyright © 2010 Red Hat, Inc. and others.

The text of and illustrations in this document are licensed by Red Hat under a Creative Commons Attribution–Share Alike 3.0 Unported license ("CC-BY-SA"). An explanation of CC-BY-SA is available at <http://creativecommons.org/licenses/by-sa/3.0/>. In accordance with CC-BY-SA, if you distribute this document or an adaptation of it, you must provide the URL for the original version.

Red Hat, as the licensor of this document, waives the right to enforce, and agrees not to assert, Section 4d of CC-BY-SA to the fullest extent permitted by applicable law.

Red Hat, Red Hat Enterprise Linux, the Shadowman logo, JBoss, MetaMatrix, Fedora, the Infinity Logo, and RHCE are trademarks of Red Hat, Inc., registered in the United States and other countries.

Linux® is the registered trademark of Linus Torvalds in the United States and other countries.

Java® is a registered trademark of Oracle and/or its affiliates.

XFS® is a trademark of Silicon Graphics International Corp. or its subsidiaries in the United States and/or other countries.

MySQL® is a registered trademark of MySQL AB in the United States, the European Union and other countries.

All other trademarks are the property of their respective owners.

1801 Varsity Drive
Raleigh, NC 27606-2072 USA
Phone: +1 919 754 3700
Phone: 888 733 4281
Fax: +1 919 754 3701

Zusammenfassung

Überblick über die Red Hat Cluster Suite liefert einen Überblick über das Hochverfügbarkeits-Add-On für Red Hat Enterprise Linux 6.

Inhaltsverzeichnis

Einführung

1. Dokumentkonventionen

- 1.1. Typografische Konventionen
- 1.2. Konventionen für Seitenansprachen (engl.: pull-quotes)
- 1.3. Anmerkungen und Warnungen

1. Überblick über das Hochverfügbarkeits-Add-On

- 1.1. Cluster-Grundlagen
- 1.2. Einführung in das Hochverfügbarkeits-Add-On
- 1.3. Cluster-Infrastruktur
 - 1.3.1. Cluster-Management
 - 1.3.2. Lock-Management
 - 1.3.3. Fencing
 - 1.3.4. Konfigurations-Management

- 1.4. Hochverfügbarkeitsdienst-Management
- 1.5. Cluster Administrations-Tools

A. Revisionsverlauf

Einführung

Dieses Dokument bietet einen umfassenden Überblick über das Hochverfügbarkeits-Add-On für Red Hat Enterprise Linux 6.

Auch wenn es sich bei den Informationen in diesem Dokument um einen Überblick handelt, sollten Sie über ausreichende Kenntnisse von Red Hat Enterprise Linux verfügen und die Konzepte von Server-Rechnern verstehen, um ein gutes Verständnis der Informationen zu erlangen.

Werfen Sie einen Blick auf die folgenden Quellen für weitere Informationen zur Verwendung von Red Hat Enterprise Linux:

- ▶ *Red Hat Enterprise Linux Installationshandbuch* — Liefert Informationen bezüglich der Installation von Red Hat Enterprise Linux 6.
- ▶ *Red Hat Enterprise Linux Bereitstellungshandbuch* — Liefert Informationen bezüglich des Einsatzes, der Konfiguration und der Administration von Red Hat Enterprise Linux 6.

Werfen Sie einen Blick auf die folgenden Quellen für weitere Informationen zu diesen und verwandten Produkten für Red Hat Enterprise Linux 6:

- ▶ *Konfiguration und Verwaltung des Hochverfügbarkeits-Add-Ons* Liefert Informationen zur Konfiguration und Verwaltung des Hochverfügbarkeits-Add-Ons (auch als Red Hat Cluster bekannt) für Red Hat Enterprise Linux 6.
- ▶ *Logical Volume Manager Administration* — Liefert eine Beschreibung des Logical Volume Manager (LVM), inklusive Informationen zum Einsatz von LVM in einer Cluster-Umgebung.
- ▶ *Global File System 2: Konfiguration und Administration* — Liefert Informationen zur Installation, Konfiguration und Pflege von Red Hat GFS (Red Hat Global File System 2), die Bestandteil des Resilient-Storage-Add-Ons ist.
- ▶ *DM-Multipath* — Liefert Informationen bezüglich des Device-Mapper Multipath Features von Red Hat Enterprise Linux 6.
- ▶ *Linux-Virtual-Server Administration* — Liefert Informationen zur Konfiguration von Hochleistungssystemen und -diensten mit dem Red Hat Lastverteilungs-Add-On (bisher als Linux Virtual Server [LVS] bekannt).
- ▶ *Versionshinweise* — Liefert Informationen zu aktuellen Releases von Red Hat Produkten.

Dieses und weitere Red Hat-Dokumente stehen als HTML-, PDF- und RPM-Versionen auf der Red Hat Enterprise Linux-Dokumentations-CD und online unter <http://docs.redhat.com/> zur Verfügung.

1. Dokumentkonventionen

Dieses Handbuch verwendet mehrere Konventionen, um bestimmte Wörter und Phrasen hervorzuheben und Aufmerksamkeit auf spezifische Informationen zu lenken.

In PDF- und Papiaerausgaben verwendet dieses Handbuch Schriftbilder des [Liberation-Fonts](#)-Sets. Das Liberation-Fonts-Set wird auch für HTML-Ausgaben verwendet, falls es auf Ihrem System installiert ist. Falls nicht, werden alternative, aber äquivalente Schriftbilder angezeigt. Beachten Sie: Red Hat Enterprise Linux 5 und die nachfolgende Versionen beinhalten das Liberation-Fonts-Set standardmäßig.

1.1. Typografische Konventionen

Es werden vier typografische Konventionen verwendet, um die Aufmerksamkeit auf spezifische Wörter und Phrasen zu lenken. Diese Konventionen und die Umstände, unter denen sie auftreten, sind folgende:

Nichtproportional Fett

Dies wird verwendet, um Systemeingaben hervorzuheben, einschließlich Shell-Befehle, Dateinamen und Pfade. Es wird ebenfalls zum Hervorheben von Tasten und Tastenkombinationen verwendet. Zum Beispiel:

Um den Inhalt der Datei **my_next_bestselling_novel** in Ihrem aktuellen Arbeitsverzeichnis zu sehen, geben Sie den Befehl **cat my_next_bestselling_novel** in den Shell-Prompt ein und drücken Sie **Enter**, um den Befehl auszuführen.

Das oben aufgeführte Beispiel beinhaltet einen Dateinamen, einen Shell-Befehl und eine Taste. Alle werden nichtproportional fett dargestellt und alle können, dank des Kontextes, leicht unterschieden werden.

Die Tastenkombination kann von einer Taste durch ein Pluszeichen, das alle Teile der Tastenkombination miteinander verbindet, unterschieden werden. Zum Beispiel:

Drücken Sie **Enter**, um den Befehl auszuführen.

Press **Ctrl+Alt+F2** to switch to a virtual terminal.

Der erste Absatz hebt die jeweilige Taste hervor, die gedrückt werden soll. Der zweite Absatz hebt zwei Tastenkombinationen hervor (jeweils ein Satz von drei Tasten, wobei jeder Satz gleichzeitig gedrückt wird).

Falls Quellcode diskutiert wird, werden Klassennamen, Methoden, Funktionen, Variablennamen und Rückgabewerte, die innerhalb eines Abschnitts erwähnt werden, wie oben gezeigt **nichtproportional fett** dargestellt. Zum Beispiel:

Zu dateiverwandten Klassen zählen **filesystem** für Dateisysteme, **file** für Dateien und **dir** für Verzeichnisse. Jede Klasse hat ihren eigenen Satz an Berechtigungen.

Proportional Fett

Dies kennzeichnet Wörter oder Phrasen, die auf einem System vorkommen, einschließlich Applikationsnamen, Text in Dialogboxen, beschriftete Schaltflächen, Bezeichnungen für Auswahlkästchen und Radio-Buttons, Überschriften von Menüs und Untermenüs. Zum Beispiel:

Wählen Sie **System** → **Einstellungen** → **Maus** in der Hauptmenüleiste aus, um die **Mauseinstellungen** zu öffnen. Klicken Sie im Reiter **Tasten** auf das Auswahlkästchen **Mit links bediente Maus** und anschließend auf **Schließen**, um die primäre Maustaste von der linken auf die rechte Seite zu ändern (d.h., um die Maus auf Linkshänder anzupassen).

To insert a special character into a **gedit** file, choose **Applications** → **Accessories** → **Character Map** from the main menu bar. Next, choose **Search** → **Find...** from the **Character Map** menu bar, type the name of the character in the **Search** field and click **Next**. The character you sought will be highlighted in the **Character Table**. Double-click this highlighted character to place it in the **Text to copy** field and then click the **Copy** button. Now switch back to your document and choose **Edit** → **Paste** from the **gedit** menu bar.

Der oben aufgeführte Text enthält Applikationsnamen, systemweite Menünamen und -elemente, applikationsspezifische Menünamen sowie Schaltflächen und Text innerhalb einer GUI-Oberfläche. Alle werden proportional fett dargestellt und sind anhand des Kontextes unterscheidbar.

Nichtproportional Fett Kursiv oder **Proportional Fett Kursiv**

Egal ob nichtproportional fett oder proportional fett, ein zusätzlicher Kursivdruck weist auf einen ersetzbaren oder variablen Text hin. Kursivdruck kennzeichnet Text, der nicht wörtlich eingegeben wird,

oder angezeigter Text, der sich je nach gegebenen Umständen ändert. Zum Beispiel:

Um sich mit einer Remote-Maschine via SSH zu verbinden, geben Sie an einem Shell-Prompt **ssh *username@domain.name*** ein. Falls die Remote-Maschine **example.com** ist und Ihr Benutzername auf dieser Maschine John lautet, geben Sie also **ssh john@example.com** ein.

Der Befehl **mount -o remount *file-system*** hängt das angegebene Dateisystem wieder ein. Um beispielsweise das **/home**-Dateisystem wieder einzuhängen, verwenden Sie den Befehl **mount -o remount /home**.

Um die Version des derzeit installierten Pakets zu sehen, verwenden Sie den Befehl **rpm -q *package***. Die Ausgabe sieht wie folgt aus: ***package-version-release***.

Note the words in bold italics above — username, domain.name, file-system, package, version and release. Each word is a placeholder, either for text you enter when issuing a command or for text displayed by the system.

Neben der Standardbenutzung für die Darstellung des Titels eines Werks, zeigt der Kursivdruck auch die erstmalige Nutzung eines neuen und wichtigen Begriffs an. Zum Beispiel:

Publican ist ein *DocBook* Publishing-System.

1.2. Konventionen für Seitenansprachen (engl.: pull-quotes)

Ausgaben des Terminals und Auszüge aus dem Quellcode werden visuell vom umliegenden Text hervorgehoben.

Eine Ausgabe, die an das Terminal gesendet wird, wird in den Schrifttyp **nichtproportional Roman** gesetzt und demnach wie folgt präsentiert:

```
books      Desktop  documentation  drafts  mss    photos  stuff  svn
books_tests Desktop1  downloads      images  notes  scripts  svgs
```

Auszüge aus dem Quellcode werden ebenfalls in den Schrifttyp **nichtproportional Roman** gesetzt, doch wird zusätzlich noch die Syntax hervorgehoben:

```
package org.jboss.book.jca.ex1;

import javax.naming.InitialContext;

public class ExClient
{
    public static void main(String args[])
        throws Exception
    {
        InitialContext iniCtx = new InitialContext();
        Object          ref    = iniCtx.lookup("EchoBean");
        EchoHome        home   = (EchoHome) ref;
        Echo            echo    = home.create();

        System.out.println("Created Echo");

        System.out.println("Echo.echo('Hello') = " + echo.echo("Hello"));
    }
}
```

1.3. Anmerkungen und Warnungen

Zu guter Letzt verwenden wir drei visuelle Stile, um die Aufmerksamkeit auf Informationen zu lenken, die andernfalls vielleicht übersehen werden könnten.



Anmerkung

Eine Anmerkung ist ein Tipp, ein abgekürztes Verfahren oder ein alternativer Ansatz für die vorliegende Aufgabe. Das Ignorieren von Anmerkungen sollte keine negativen Auswirkungen haben, aber Sie verpassen so vielleicht einen Trick, der Ihnen das Leben vereinfachen könnte.



Wichtig

Die Wichtig-Schaukästen lenken die Aufmerksamkeit auf Dinge, die sonst leicht übersehen werden können: Konfigurationsänderungen, die nur für die aktuelle Sitzung gelten oder Dienste, für die ein Neustart nötig ist, bevor eine Aktualisierung wirksam wird. Das Ignorieren von Wichtig-Schaukästen würde keinen Datenverlust verursachen, kann aber unter Umständen zu Ärgernissen und Frustration führen.



Warnung

Eine Warnung sollte nicht ignoriert werden. Das Ignorieren von Warnungen führt mit hoher Wahrscheinlichkeit zu Datenverlust.

Kapitel 1. Überblick über das Hochverfügbarkeits-Add-On

Das Hochverfügbarkeits-Add-On ist ein Cluster-System, das Verlässlichkeit, Skalierbarkeit und Verfügbarkeit für kritische Produktionsdienste. Die folgenden Abschnitte liefern eine Beschreibung der Komponenten und Funktionen des Hochverfügbarkeits-Add-On auf hohem Niveau:

- ▶ [Abschnitt 1.1, „Cluster-Grundlagen“](#)
- ▶ [Abschnitt 1.2, „Einführung in das Hochverfügbarkeits-Add-On“](#)
- ▶ [Abschnitt 1.3, „Cluster-Infrastruktur“](#)
- ▶ [Abschnitt 1.4, „Hochverfügbarkeitsdienst-Management“](#)
- ▶ [Abschnitt 1.5, „Cluster Administrations-Tools“](#)

1.1. Cluster-Grundlagen

Ein Cluster besteht aus zwei oder mehreren Computern (auch *Knoten* oder *Mitglieder* genannt), die bei der Durchführung einer Aufgabe zusammenarbeiten. Es gibt vier Haupttypen von Clustern:

- ▶ Speicher-Cluster
- ▶ Hochverfügbarkeits-Cluster
- ▶ Lastverteilungs-Cluster
- ▶ Hochleistungs-Cluster

Speicher-Cluster stellen ein Image mit konsistentem Dateisystem zwischen Servern in einem Cluster zur Verfügung. Dies erlaubt den Servern, gleichzeitig auf ein einzelnes gemeinsam genutztes Dateisystem lesend und schreibend zuzugreifen. Ein Speicher-Cluster vereinfacht die Speicheradministration, indem es die Installation und das Patchen von Applikationen auf ein Dateisystem reduziert. Auch eliminiert ein Speicher-Cluster den Bedarf für redundante Kopien von Applikationsdaten durch ein clusterweites Dateisystem und vereinfacht Backup und Notfallwiederherstellung. Das Hochverfügbarkeits-Add-On stellt Speicher-Clustering in Verbindung mit Red Hat GFS2 (als Teil des Resilient-Storage-Add-On) zur Verfügung.

Hochverfügbarkeits-Cluster bieten kontinuierliche Verfügbarkeit von Diensten, indem sie einzelne Ausfallpunkte eliminieren und Dienste von einem Cluster-Knoten auf einen anderen transferieren, falls ein Knoten nicht funktionsfähig ist. Typischerweise schreiben und lesen Dienste Daten in einem Hochverfügbarkeits-Cluster (via eingehängte Dateisysteme, die lesbar/beschreibbar sind). Aus diesem Grund muss ein Hochverfügbarkeits-Cluster die Integrität von Daten gewährleisten, wenn ein Cluster-Knoten die Kontrolle eines Dienstes von einem anderen Cluster-Knoten übernimmt. Knotenausfälle in einem Hochverfügbarkeits-Cluster sind von Clients außerhalb des Clusters nicht sichtbar. (Hochverfügbarkeits-Cluster werden manchmal auch als Ausfallsicherungs-Cluster bezeichnet.) Das Hochverfügbarkeits-Add-On stellt Hochverfügbarkeits-Clustering im Rahmen der Hochverfügbarkeitsdienst-Management-Komponente **rgmanager** zur Verfügung.

Lastverteilungs-Cluster versenden Anfragen von Netzwerk-Diensten an mehrere Cluster-Knoten, um die Auslastung durch die Anfrage gleichmäßig zwischen den Cluster-Knoten zu verteilen. Lastverteilung bietet eine kostengünstige Skalierbarkeit, da die Anzahl der Knoten mit den Auslastungsanforderungen abgeglichen werden kann. Falls ein Knoten in einem Lastverteilungs-Cluster nicht mehr funktionsfähig ist, erkennt die Lastverteilungs-Software diesen Ausfall und leitet Anfragen an andere Cluster-Knoten weiter. Knoten-Ausfälle in einem Lastverteilungs-Cluster sind nicht ersichtlich von Clients außerhalb des Clusters. Lastverteilung steht mit dem Lastverteilungs-Add-On zur Verfügung.

Hochleistungs-Cluster verwenden Cluster-Knoten zur Durchführung von gleichzeitigen Berechnungen. Ein Hochleistungs-Cluster erlaubt Applikationen, parallel zu arbeiten und steigert daher die Leistungsfähigkeit der Applikationen. (Hochleistungs-Cluster werden auch als Computer-Cluster oder Grid-Computing bezeichnet.)



Anmerkung

Die Cluster-Typen, die in diesem vorausgehenden Text zusammengefasst sind, geben Grundkonfigurationen wieder. Ihre Anforderungen erfordern ggf. eine Kombination der beschriebenen Cluster.

1.2. Einführung in das Hochverfügbarkeits-Add-On

Das Hochverfügbarkeits-Add-On ist ein integriertes Set von Software-Komponenten, das in einer Vielzahl von Konfigurationen eingesetzt werden kann, um Ihren Anforderungen in Sachen Leistung, Hochverfügbarkeit, Lastverteilung, Skalierbarkeit, gemeinsamen Dateizugriff und Wirtschaftlichkeit gerecht zu werden.

Das Hochverfügbarkeits-Add-On besteht aus den folgenden Hauptkomponenten:

- ▶ Cluster-Infrastruktur — Bietet grundlegende Funktionen für Knoten für die Zusammenarbeit als ein Cluster: Verwalten der Konfigurationsdatei, der Zugehörigkeit, der Sperrung und der Datenabgrenzung (Fencing).
- ▶ Hochverfügbarkeitsdienst-Management — Bietet die Übertragung von Diensten von einem Knoten auf einen anderen, falls ein Knoten nicht mehr funktionsfähig ist.
- ▶ Cluster Administrations-Tools — Konfigurations- und Management-Tools zum Einrichten, Konfigurieren und Verwalten des Hochverfügbarkeits-Add-Ons. Die Tools sind zur Verwendung mit den Cluster-Infrastruktur-Komponenten, den Hochverfügbarkeits-Komponenten und den Dienst-Management-Komponenten und Speicher gedacht.

Sie können das Hochverfügbarkeits-Add-On durch die folgenden Komponenten ergänzen:

- ▶ Red Hat GFS2 (Global File System 2) — Als Teil des Resilient-Storage-Add-Ons bietet es ein Cluster-Dateisystem zur Verwendung mit der Red Hat Cluster Suite. GFS2 ermöglicht mehreren Knoten, Speicher auf einem Blocklevel gemeinsam zu nutzen, als ob der Speicher lokal mit jedem Cluster-Knoten verbunden wäre.
- ▶ Cluster Logical Volume Manager (CLVM) — Als Teil des Resilient-Storage-Add-Ons bietet es die Verwaltung von Datenträgern eines Cluster-Speichers.
- ▶ Lastverteilungs-Add-On — Routing-Software, die IP-Lastverteilung bietet. Das Lastverteilungs-Add-On läuft als ein Paar redundanter virtueller Server, die Client-Anfragen gleichmäßig auf reale Server hinter den virtuellen Servern verteilen.

1.3. Cluster-Infrastruktur

Die Hochverfügbarkeits-Add-On Cluster-Infrastruktur liefert die Grundfunktionen für eine Gruppe von Computern (auch *Knoten* oder *Mitglieder* genannt), damit diese in einem Cluster zusammenarbeiten können. Sobald ein Cluster unter Verwendung der Cluster-Infrastruktur gebildet wird, können Sie weitere Komponenten verwenden, um Ihren Clustering-Anforderungen gerecht zu werden (z.B. das Einrichten eines Cluster für den gemeinsamen Dateizugriff auf einem GFS2-Dateisystem oder das Einrichten von Dienst-Übertragungen bei einem Ausfall). Die Cluster-Infrastruktur führt die folgenden Funktionen durch:

- ▶ Cluster-Management
- ▶ Lock-Management
- ▶ Fencing
- ▶ Cluster-Konfigurations-Management

1.3.1. Cluster-Management

Cluster-Management verwaltet das Cluster-Quorum und die Cluster-Zugehörigkeit. CMAN (eine Abkürzung für Cluster-Manager) führt das Cluster-Management im Rahmen des Hochverfügbarkeits-Add-Ons für Red Hat Enterprise Linux 6 durch. CMAN ist ein verteilter Cluster-Manager und läuft auf jedem Cluster-Knoten. Cluster-Management wird auf alle Knoten im Cluster verteilt.

CMAN behält das Cluster-Quorum im Auge, indem er die Zahl der Cluster-Knoten überwacht. Falls mehr als die Hälfte der Knoten aktiv sind, besitzt das Cluster ein Quorum. Falls die Hälfte der Knoten (oder weniger) aktiv sind, besitzt das Cluster kein Quorum und sämtliche Cluster-Aktivitäten werden angehalten. Ein Cluster-Quorum verhindert das Auftreten eines "split-brain"-Zustands — ein Zustand, bei dem zwei Instanzen desselben Clusters laufen. Ein Split-Brain-Zustand würde jeder Cluster-Instanz erlauben, auf Cluster-Ressourcen ohne Kenntnis der anderen Cluster-Instanz zuzugreifen, was zu einer fehlerhaften Cluster-Integrität führt.

Ein Quorum wird durch die Kommunikation von Meldungen zwischen Cluster-Knoten via Ethernet bestimmt. Optional kann ein Quorum durch eine Kombination aus kommunizierenden Meldungen via Ethernet *und* durch eine Quorum-Platte bestimmt werden. Bei einem Quorum via Ethernet besteht dieses zu 50 Prozent aus den Knotenstimmen plus 1. Bei einem Quorum via Quorum-Platte besteht dieses aus benutzerspezifischen Bedingungen.



Anmerkung

Standardmäßig besitzt jeder Knoten eine Quorum-Stimme. Optional können Sie jeden Knoten so konfigurieren, dass er mehr als eine Stimme besitzt.

CMAN behält die Zugehörigkeit im Auge, indem er die Meldungen von anderen Cluster-Knoten überwacht. Wenn sich Cluster-Zugehörigkeiten ändern, informiert der Cluster-Manager die anderen Infrastruktur-Komponenten, die dann die entsprechenden Aktionen einleiten. Falls ein Cluster-Knoten eine Meldung nicht innerhalb einer vorgeschriebenen Zeitspanne übermittelt, entfernt der Cluster-Manager den Knoten aus dem Cluster und kommuniziert mit anderen Cluster-Infrastruktur-Komponenten, dass der Knoten kein Mitglied mehr ist. Andere Cluster-Infrastruktur-Komponenten wiederum bestimmen, welche Aktionen als Folge der Benachrichtigung, dass der Knoten nicht mehr länger ein Bestandteil des Clusters ist, einzuleiten sind. "Fencing" würde den Knoten, der kein Mitglied mehr ist, beispielsweise ausgrenzen.

1.3.2. Lock-Management

Lock-Management ist ein allgemeiner Cluster-Infrastruktur-Dienst, der einen Mechanismus für andere Cluster-Infrastruktur-Komponenten zur Synchronisierung ihres Zugriffs auf gemeinsam genutzte Ressourcen zur Verfügung stellt. In einem Red Hat-Cluster ist DLM (Distributed Lock Manager) der Lock-Manager. Wie sein Name schon vermuten lässt, ist der DLM ein verteilter Lock-Manager und läuft auf jedem Cluster-Knoten. Lock-Management wird auf alle Knoten im Cluster verteilt. GFS2 und CLVM verwenden Locks des Lock-Managers. GFS2 verwendet Locks des Lock-Managers, um den Zugriff auf Metadaten des Dateisystems (auf gemeinsam genutzten Speicher) zu synchronisieren. CLVM verwendet Locks des Lock-Managers, um Aktualisierungen auf den LVM-Datenträgern und Datenträgergruppen (ebenfalls auf gemeinsam genutzten Speicher) zu synchronisieren. Zusätzlich dazu verwendet **rgmanager** DLM zur Synchronisation von Zuständen von Diensten.

1.3.3. Fencing

Die Datenabgrenzung (Fencing) ist die Trennung eines Knotens vom gemeinsam genutzten Speicher des Clusters. Fencing unterbricht I/O auf gemeinsam genutzten Speicher und gewährleistet so die Datenintegrität. Die Cluster-Infrastruktur führt Fencing mit Hilfe des Fence-Daemon, **fenced**, aus.

Wenn CMAN bemerkt, dass ein Knoten ausgefallen ist, kommuniziert er mit den anderen Cluster-Infrastruktur-Komponenten, dass der Knoten ausgefallen ist. Sobald **fenced** von dem Ausfall unterrichtet ist, grenzt er den ausgefallenen Knoten ab. Die anderen Cluster-Infrastruktur-Komponenten bestimmen, welche Schritte einzuleiten sind — d.h. sie führen sämtliche Wiederherstellungsmaßnahmen durch, die durchgeführt werden müssen. So suspendieren beispielsweise DLM und GFS2 beim Erhalt der Information über einen Knoten-Ausfall sämtliche Aktivitäten, bis sie feststellen, dass der **fenced** den Fencing-Prozess für den ausgefallenen Knoten abgeschlossen hat. Nach der Bestätigung, dass der ausgefallene Knoten abgegrenzt ist, führen DLM und GFS2 Wiederherstellungsmaßnahmen durch. DLM gibt die Locks für den ausgefallenen Knoten frei und GFS2 stellt das Journal des ausgefallenen Knotens wieder her.

Das Fencing-Programm ermittelt anhand der Cluster-Konfigurationsdatei, welche Fencing-Methode verwendet werden soll. Zwei Schlüsselemente in der Cluster-Konfigurationsdatei definieren eine Fencing-Methode: Der Fencing-Agent und das Fencing-Gerät. Das Fencing-Programm startet einen Aufruf zu einem in der Cluster-Konfigurationsdatei angegebenen Fencing-Agent. Der Fencing-Agent grenzt den Knoten im Gegenzug via Fencing-Gerät ab. Nach Abschluss des Fencings, unterrichtet das Fencing-Programm den Cluster-Manager.

Das Hochverfügbarkeits-Add-On stellt eine Vielzahl von Fencing-Methoden zur Verfügung:

- » Power-Fencing — Eine Fencing-Methode, die einen Power-Kontroller verwendet, um einen nicht mehr funktionierenden Knoten auszuschalten.
- » Fibre-Channel-Switch-Fencing — Eine Fencing-Methode, die den Fibre-Channel-Port deaktiviert, der den Speicher mit einem nicht funktionsfähigen Knoten verbindet.
- » Sonstiges Fencing — Mehrere sonstige Fencing-Methoden, die I/O oder den Strom eines nicht funktionsfähigen Knotens deaktiviert, inklusive IBM-Bladecenters, PAP, DRAC/MC, ILO, IPMI, IBM RSA II und weitere.

[Abbildung 1.1, „Power-Fencing Beispiel“](#) zeigt ein Beispiel für Power-Fencing. In diesem Beispiel veranlasst das Fencing-Programm auf Knoten A den Power-Kontroller, Knoten D auszuschalten.

[Abbildung 1.2, „Fibre-Channel-Switch-Fencing Beispiel“](#) zeigt ein Beispiel für Fibre-Channel-Switch-Fencing. In dem Beispiel veranlasst das Fencing-Programm auf Knoten A den Fibre-Channel-Switch, den Port für Knoten D zu deaktivieren, sodass der Knoten D getrennt wird.

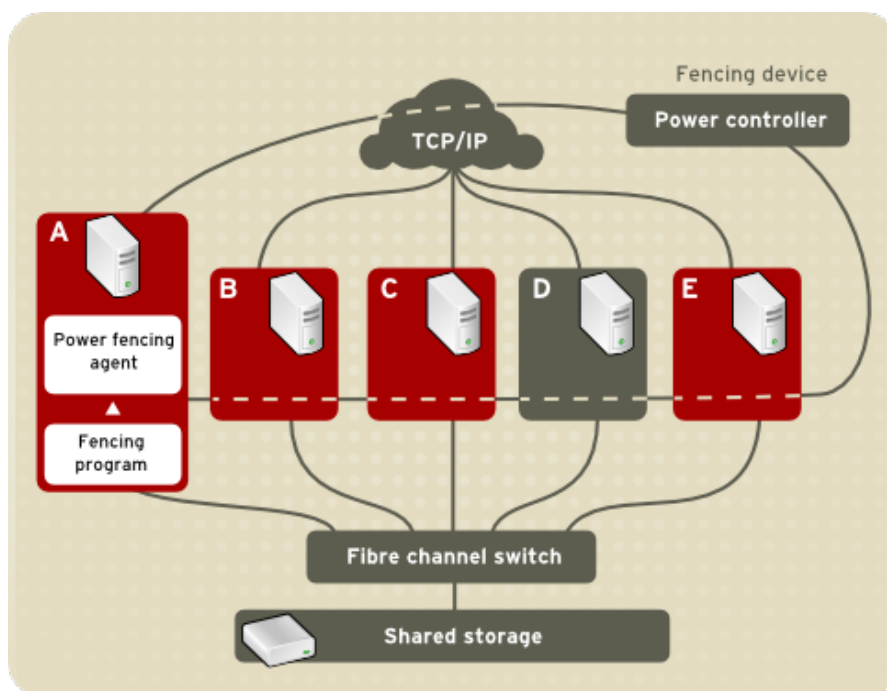


Abbildung 1.1. Power-Fencing Beispiel

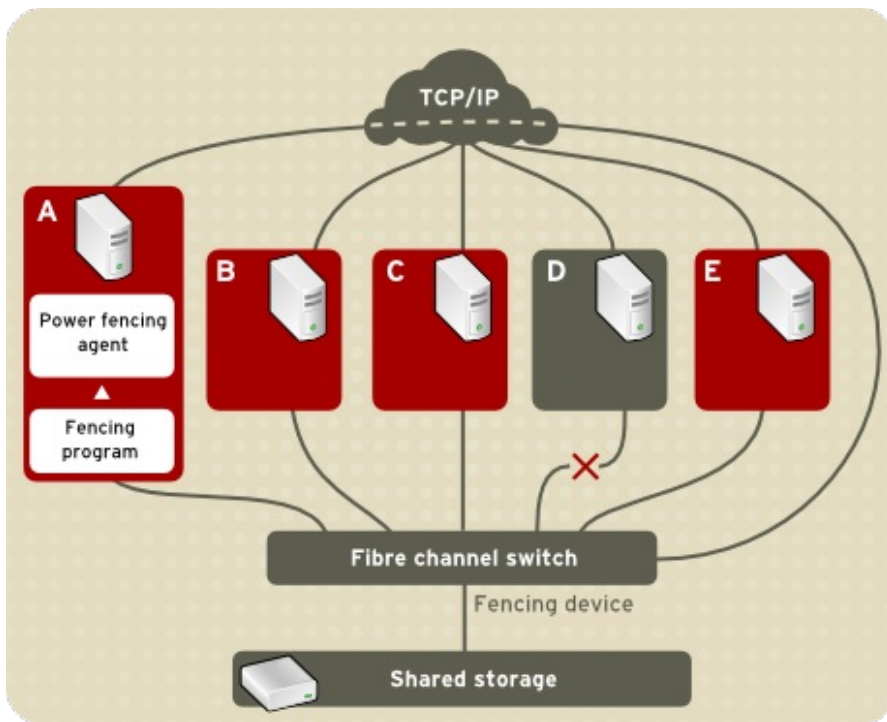


Abbildung 1.2. Fibre-Channel-Switch-Fencing Beispiel

Das Angeben einer Fencing-Methode beinhaltet das Bearbeiten einer Cluster-Konfigurationsdatei für die Zuweisung eines Namens für die Fencing-Methode, dem Fencing-Agent und dem Fencing-Gerät für jeden Knoten im Cluster.

Die Art und Weise, wie eine Fencing-Methode angegeben wird, hängt davon ab, ob ein Knoten entweder Dual-Stromversorgung oder mehrere Pfade zum Speicher besitzt. Falls ein Knoten eine duale Stromversorgung besitzt, dann muss die Fencing-Methode für den Knoten mindestens zwei Fencing-Geräte angeben — ein Fencing-Gerät für jede Stromversorgung (siehe [Abbildung 1.3, „Datenabgrenzung \(Fencing\) eines Knotens mit einer Dual-Stromversorgung“](#)). Genauso gilt, dass wenn ein Knoten mehrere Pfade zu Fibre-Channel-Speicher besitzt, die Fencing-Methode für den Knoten ein Fencing-Gerät für jeden Pfad zum Fibre-Channel-Speicher angeben muss. Wenn ein Knoten beispielsweise zwei Pfade zum Fibre-Channel-Speicher besitzt, sollte die Fencing-Methode zwei Fencing-Geräte angeben — einen für jeden Pfad zum Fibre-Channel-Speicher (siehe [Abbildung 1.4, „Datenabgrenzung \(Fencing\) eines Knotens mit Dual-Fibre-Channel-Verbindungen“](#)).

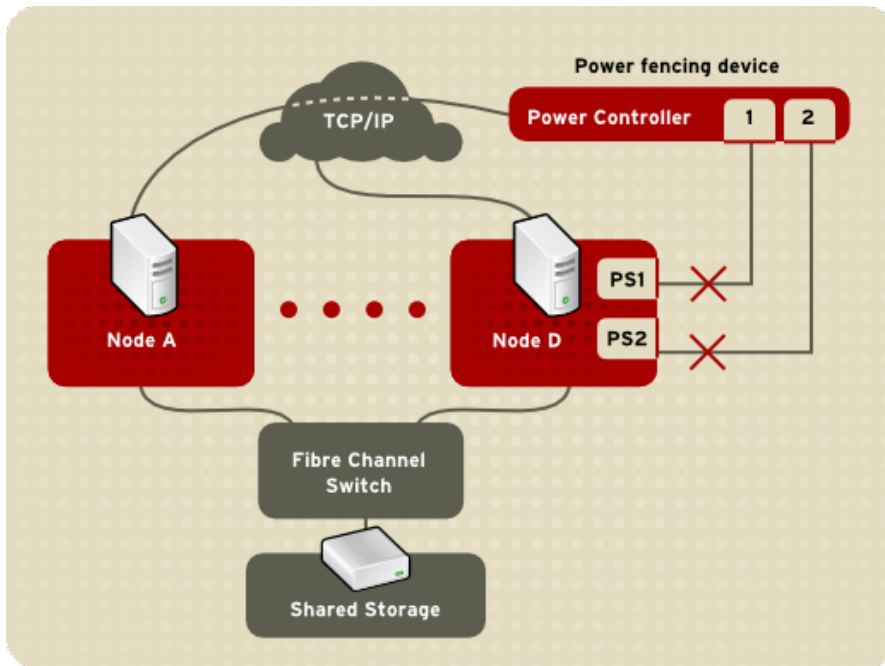


Abbildung 1.3. Datenabgrenzung (Fencing) eines Knotens mit einer Dual-Stromversorgung

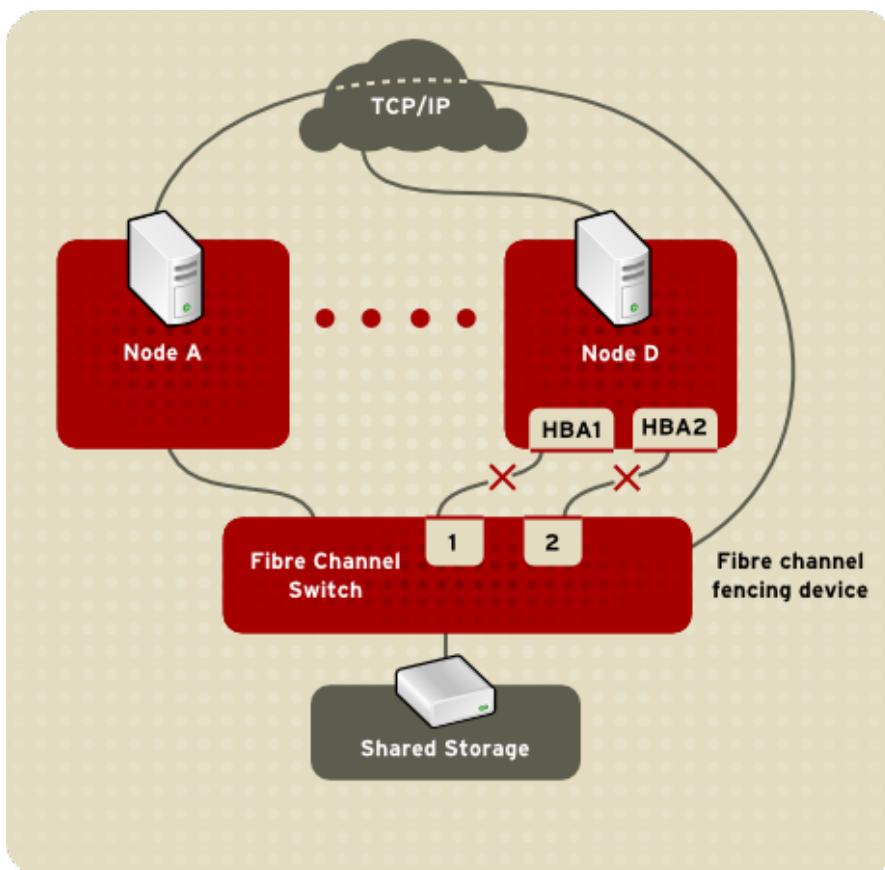


Abbildung 1.4. Datenabgrenzung (Fencing) eines Knotens mit Dual-Fibre-Channel-Verbindungen

Sie können einen Knoten mit einer oder mehreren Fencing-Methoden konfigurieren. Wenn Sie einen Knoten mit einer Fencing-Methode konfigurieren, ist dies die einzige Fencing-Methode, die für das Fencing dieses Knotens zur Verfügung steht. Wenn Sie einen Knoten mit mehreren Fencing-Methoden konfigurieren, werden die Fencing-Methoden von einer Fencing-Methode zu einer anderen *kaskadiert*, nach der in der Cluster-Konfigurationsdatei angegebenen Reihenfolge für Fencing-Methoden. Falls ein

Knoten ausfällt, wird er mit der ersten in der Cluster-Konfigurationsdatei für diesen Knoten angegebenen Fencing-Methode abgegrenzt. Falls die erste Fencing-Methode nicht erfolgreich ist, wird die zweite für diesen Knoten angegebene Fencing-Methode verwendet. Falls keine der Fencing-Methoden erfolgreich ist, startet der Fencing-Prozess erneut mit der ersten angegebenen Fencing-Methode und geht die Fencing-Methoden in der in der Cluster-Konfigurationsdatei angegebenen Reihenfolge solange durch, bis der Knoten abgegrenzt wurde.

1.3.4. Konfigurations-Management

In der Cluster-Konfigurationsdatei (`/etc/cluster/cluster.conf`) wird die Konfiguration des Hochverfügbarkeits-Add-Ons definiert. Sie ist eine XML-Datei, die die folgenden Cluster-Charakteristiken beschreibt:

- ▶ Cluster-Name — Definiert den Cluster-Namen, das Revisionslevel der Cluster-Konfigurationsdatei und grundlegende Fence-Timing-Eigenschaften, die verwendet werden, wenn ein Knoten einem Cluster beiträgt oder aus einem Cluster abgegrenzt wird.
- ▶ Cluster — Definiert jeden Knoten des Clusters, unter Angabe des Knoten-Namens, der Knoten-ID, der Zahl der Quorum-Stimmen und der Fencing-Methode für diesen Knoten.
- ▶ Fence-Gerät — Definiert Fence-Geräte im Cluster. Die Parameter variieren je nach Typ des Fence-Geräts. Für einen Power-Kontroller beispielsweise, der als Fence-Gerät benutzt wird, definiert die Cluster-Konfiguration den Namen des Power-Kontroller, dessen IP-Adresse, das Login und das Passwort.
- ▶ Verwaltete Ressourcen — Definiert die Ressourcen, die zum Erstellen von Cluster-Diensten benötigt werden. Verwaltete Ressourcen umfassen die Definition von Ausfallsicherungs-Domains, Ressourcen (zum Beispiel einer IP-Adresse) und Dienste. Zusammen definieren die verwalteten Ressourcen Cluster-Dienste und Ausfallsicherungsverhalten der Cluster-Dienste.

Die Cluster-Konfiguration wird automatisch basierend auf dem Cluster-Schema unter `/usr/share/cluster/cluster.rng` während des Systemstarts und beim erneuten Laden einer Konfiguration validiert. Sie können eine Cluster-Konfiguration auch jederzeit unter Verwendung des Befehls `ccs_config_validate` validieren.

Ein mit Kommentaren versehenes Schema steht zur Einsicht unter `/usr/share/doc/cman-X.Y.ZZ/cluster_conf.html` zur Verfügung (zum Beispiel `/usr/share/doc/cman-3.0.12/cluster_conf.html`).

Überprüfung der Validität der Konfiguration für die folgenden grundlegenden Fehler:

- ▶ XML-Validität — Überprüft, dass es sich bei der Konfigurationsdatei um eine gültige XML-Datei handelt.
- ▶ Konfigurationsoptionen — Überprüft, dass Optionen (XML-Elementen und -Attributen) gültig sind.
- ▶ Optionswerte — Überprüft, ob die Optionen gültige Daten enthalten (eingeschränkt).

1.4. Hochverfügbarkeitsdienst-Management

Hochverfügbarkeitsdienst-Management liefert die Fähigkeit, Hochverfügbarkeits-Cluster-Dienste mit Hilfe des Hochverfügbarkeits-Add-Ons zu erstellen und zu managen. Die Schlüsselkomponente für Hochverfügbarkeitsdienst-Management in einem Red Hat-Cluster, `rgmanager`, implementiert Ausfallsicherung für Standardapplikationen. In einem Red Hat-Cluster wird eine Applikation mit anderen Cluster-Ressourcen konfiguriert, um ein einen Hochverfügbarkeits-Cluster-Dienst zu bilden. Ein Hochverfügbarkeits-Cluster-Dienst kann zur Ausfallsicherung von einem Cluster-Knoten auf einen anderen wechseln, ohne signifikante Unterbrechung für Cluster-Clients. Cluster-Dienst-Ausfallsicherung kann dann in Kraft treten, wenn ein Knoten ausfällt, oder wenn ein Cluster-Systemadministrator den Dienst von einem Cluster-Knoten auf einen anderen verschiebt (beispielsweise für einen geplanten Ausfall eines Cluster-Knotens).

Um einen Hochverfügbarkeitsdienst zu erstellen, müssen Sie diesen in der Cluster-Konfigurationsdatei konfigurieren. Ein Cluster-Dienst besteht aus *Cluster-Ressourcen*. Cluster-Ressourcen bilden Blöcke, die Sie in der Cluster-Konfigurationsdatei erstellen und verwalten können — beispielsweise eine IP-Adresse, ein Skript zur Initialisierung einer Applikation oder eine via Red Hat GFS2 gemeinsam genutzte Partition.

Sie können einen Cluster-Dienst mit einer *Ausfallsicherungs-Domain* verknüpfen. Eine Ausfallsicherungs-Domain ist eine Teilmenge von Cluster-Knoten, die einen bestimmten Cluster-Dienst ausführen dürfen (siehe [Abbildung 1.5, „Ausfallsicherungs-Domains“](#)).



Anmerkung

Ausfallsicherungs-Domains werden für den Betrieb *nicht* benötigt.

Ein Cluster-Dienst kann zur Gewährleistung der Datenintegrität nur auf einem Cluster-Knoten gleichzeitig laufen. Sie können eine Ausfallsicherungspriorität in einer Ausfallsicherungs-Domain angeben. Die Angabe einer solchen Ausfallsicherungspriorität besteht aus der Zuweisung eines Prioritätslevels für jeden Knoten in einer Ausfallsicherungs-Domain. Das Prioritätslevel ermittelt die Reihenfolge der Ausfallsicherung — dabei wird bestimmt, auf welchen Knoten ein Cluster-Dienst im Falle eines Ausfalls wechseln soll. Falls Sie keine Ausfallsicherungspriorität angeben, kann ein Cluster-Dienst im Falle eines Ausfalls auf jeden beliebigen Knoten innerhalb seiner Ausfallsicherungs-Domain wechseln. Auch können Sie angeben, ob ein Cluster-Dienst so eingeschränkt werden soll, dass er nur auf Knoten seiner verknüpften Ausfallsicherungs-Domains läuft. (Ist ein Cluster-Dienst mit einer uneingeschränkten Ausfallsicherungs-Domain verknüpft, kann er auf jedem beliebigen Cluster-Knoten starten für den Fall, dass kein Mitglied der Ausfallsicherungs-Domain zur Verfügung steht.)

Unter [Abbildung 1.5, „Ausfallsicherungs-Domains“](#) ist die Ausfallsicherungs-Domain 1 so konfiguriert, dass die Ausfallsicherung innerhalb der Domain eingeschränkt ist. Aus diesem Grund kann Cluster-Dienst X im Falle eines Ausfalls nur zwischen Knoten A und Knoten B wechseln. Die Ausfallsicherungs-Domain 2 ist ebenfalls so konfiguriert, die Ausfallsicherung mit seiner Domain einzuschränken. Zusätzlich ist sie jedoch für Ausfallsicherungs-Priorität konfiguriert. Die Ausfallsicherungs-Domain 2 Priorität ist mit Knoten C als Priorität 1, Knoten B als Priorität 2 und Knoten D als Priorität 3 konfiguriert. Falls Knoten C ausfällt, wechselt Cluster-Dienst Y als nächstes zu Knoten B über. Falls er nicht auf Knoten B wechseln kann, versucht er, auf Knoten D zu wechseln. Die Ausfallsicherungs-Domain 3 ist ohne Priorität und ohne Einschränkungen konfiguriert. Falls der Knoten, auf dem Cluster-Dienst Z läuft ausfällt, versucht Cluster-Dienst Z auf einen der Knoten in der Ausfallsicherungs-Domain 3 zu wechseln. Wenn allerdings keiner dieser Knoten zur Verfügung steht, kann Cluster-Dienst Z auf jeden beliebigen Knoten im Cluster wechseln.

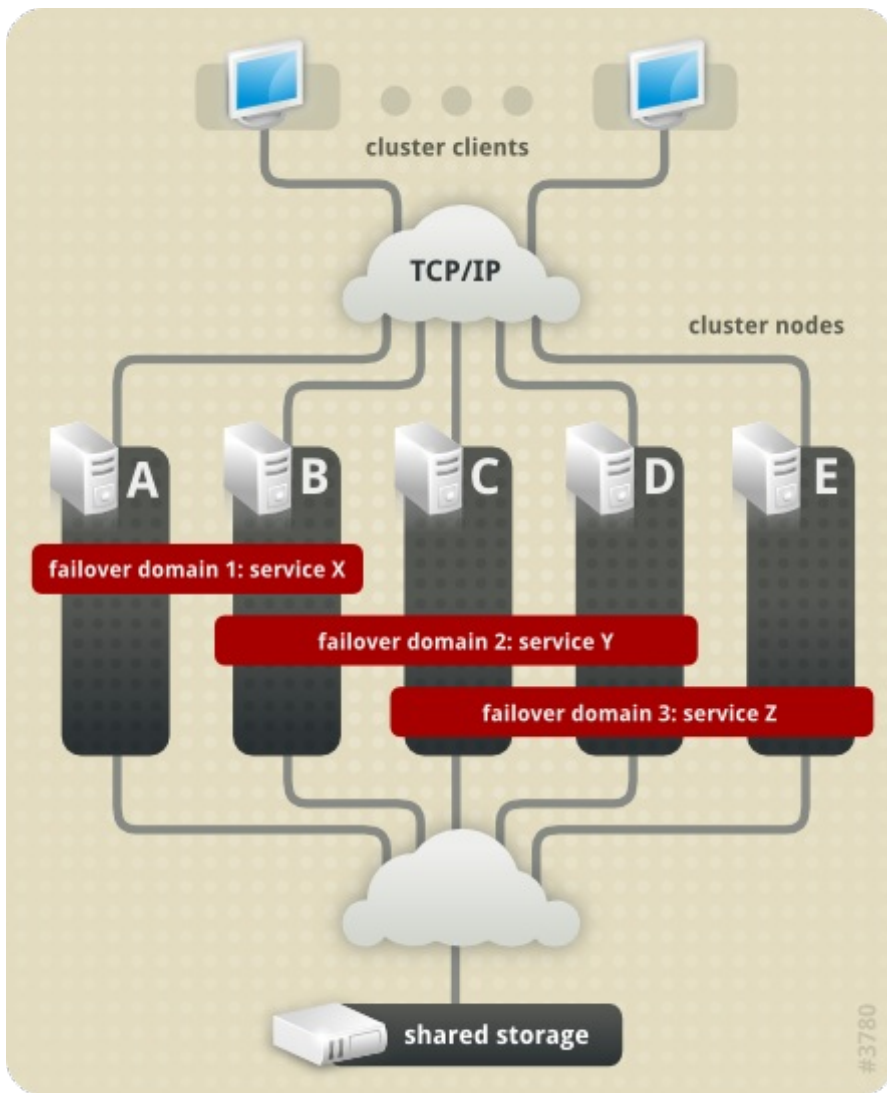


Abbildung 1.5. Ausfallsicherungs-Domains

[Abbildung 1.6, „Web-Server Cluster-Dienst Beispiel“](#) zeigt ein Beispiel eines Hochverfügbarkeits-Cluster-Dienstes, der ein Web-Server mit der Bezeichnung "content-webserver" ist. Es läuft auf Cluster-Knoten B und befindet sich in einer Ausfallsicherungs-Domain, die aus den Knoten A, B und D besteht. Zusätzlich ist die Ausfallsicherungs-Domain mit einer Ausfallsicherungspriorität konfiguriert, um im Falle eines Ausfalls auf den Knoten D zu wechseln, bevor auf Knoten A gewechselt wird und um die Ausfallsicherung nur auf Knoten aus dieser Ausfallsicherungs-Domain zu beschränken. Der Cluster-Dienst umfasst diese Cluster-Ressourcen:

- ▶ IP Adressen-Ressource — IP-Adresse 10.10.10.201.
- ▶ Eine Applikations-Ressource mit dem Namen "httpd-content" — Ein Init-Skript `/etc/init.d/httpd` für eine Web-Server-Applikation (legt `httpd` fest).
- ▶ Eine Dateisystem-Ressource — Red Hat GFS2, genannt "gfs2-content-webserver".

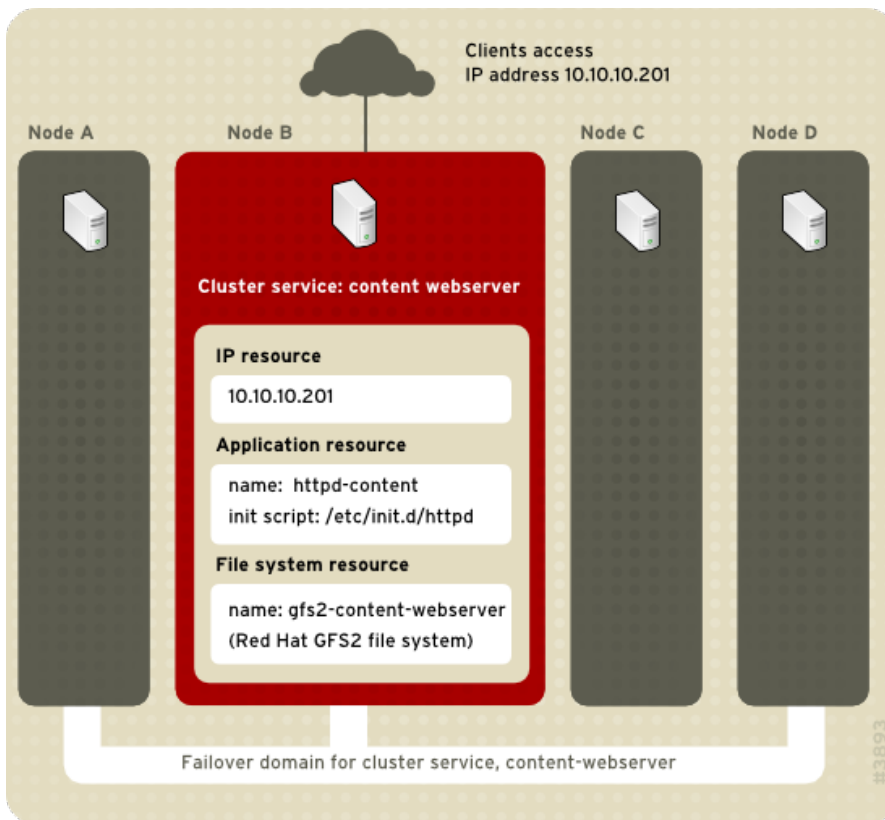


Abbildung 1.6. Web-Server Cluster-Dienst Beispiel

Clients greifen auf den Cluster-Dienst via IP-Adresse 10.10.10.201 zu, was eine Interaktion mit der Web-Server-Applikation, httpd-content, ermöglicht. Die Applikation "httpd-content" verwendet das "gfs2-content-webserver"-Dateisystem. Falls der Knoten B ausfallen sollte, würde der Cluster-Dienst "content-webserver" auf Knoten D wechseln. Falls Knoten D nicht verfügbar sein sollte, oder auch ausgefallen ist, würde der Dienst auf Knoten A wechseln. Es läge eine Ausfallsicherung ohne nennenswerte Unterbrechung für andere Cluster-Clients vor. Der Cluster-Dienst wäre von einem anderen Cluster-Knoten aus unter derselben IP-Adresse erreichbar, wie vor der Ausfallsicherung.

1.5. Cluster Administrations-Tools

Die Verwaltung von Red Hat Hochverfügbarkeits-Add-On-Software setzt sich aus der Verwendung von Konfigurations-Tools zur Definition der Beziehung zwischen den Cluster-Komponenten zusammen. Die folgenden Cluster-Konfigurations-Tools stehen im Rahmen des Red Hat Hochverfügbarkeits-Add-Ons zur Verfügung:

- **Conga** — Hierbei handelt es sich um eine umfangreiche Benutzeroberfläche zur Installation, Konfiguration und Verwaltung des Red Hat Hochverfügbarkeits-Add-Ons. Werfen Sie einen Blick auf *Konfiguration und Verwaltung des Hochverfügbarkeits-Add-Ons* für Informationen zur Konfiguration und Verwaltung des Hochverfügbarkeits-Add-Ons mit **Conga**.
- Kommandozeilen-Tools — Hierbei handelt es sich um eine Reihe Kommandozeilen-Tools zur Konfiguration und Verwaltung des Hochverfügbarkeits-Add-Ons. Werfen Sie einen Blick auf *Konfiguration und Verwaltung des Hochverfügbarkeits-Add-Ons* für Informationen zur Konfiguration und Verwaltung eines Clusters mit Kommandozeilen-Tools.



Anmerkung

system-config-cluster steht unter RHEL 6 nicht zur Verfügung.

Revisionsverlauf

Version 0-16 Rebuild for Publican 3.0	2012-07-18	Anthony Towns
Version 1.0-0 Erstmaliges Release	Wed Nov 10 2010	Paul Kennedy