



Sun Cluster Konzepthandbuch für Solaris OS

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95054
U.S.A.

Teilenr.: 819-2058
August 2005, Version A

Copyright 2005 Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 U.S.A. Alle Rechte vorbehalten.

Das in diesem Dokument beschriebene Produkt ist urheberrechtlich geschützt und wird unter Lizenzen vertrieben, die seine Verwendung, Vervielfältigung, Verteilung und Dekompilierung einschränken. Das Produkt oder Dokument darf weder vollständig noch in Teilen in irgendeiner Form oder irgendwelchen Mitteln ohne vorherige schriftliche Einwilligung von Sun und gegebenenfalls seiner Lizenzgeber reproduziert werden. Die Software anderer Hersteller, einschließlich der Schriftentechnologie, ist urheberrechtlich geschützt und von Lieferanten von Sun lizenziert.

Teile des Produkts können aus Berkeley BSD-Systemen stammen, die von der University of California lizenziert sind. UNIX ist eine eingetragene Marke in den Vereinigten Staaten und anderen Ländern und wird ausschließlich durch die X/Open Company Ltd. lizenziert.

Sun, Sun Microsystems, das Sun-Logo, docs.sun.com, AnswerBook, AnswerBook2, Sun Cluster, SunPlex, Sun Enterprise, Sun Enterprise 10000, Sun Enterprise SyMON, Sun Management Center, Solaris, Solaris Datenträger-Manager, Sun StorEdge, Sun Fire, SPARCstation, OpenBoot und Solaris sind in den USA und anderen Ländern Warenzeichen von Sun Microsystems Inc. Sämtliche SPARC-Marken werden unter Lizenz verwendet und sind Marken oder eingetragene Marken von SPARC International Inc. in den Vereinigten Staaten und anderen Ländern. Produkte mit der SPARC-Marke basieren auf einer von Sun Microsystems Inc. entwickelten Architektur. ORACLE, Netscape

Die grafischen Benutzeroberflächen von OPEN LOOK und Sun™ wurden von Sun Microsystems Inc. für seine Benutzer und Lizenznehmer entwickelt. Sun erkennt dabei die von Xerox Corporation geleistete Forschungs- und Entwicklungsarbeit auf dem Gebiet der visuellen oder grafischen Benutzeroberflächen für die Computerindustrie an. Sun ist Inhaber einer einfachen Lizenz von Xerox für die Xerox Graphical User Interface (grafische Benutzeroberfläche von Xerox). Mit dieser Lizenz werden auch die Sun-Lizenznehmer abgedeckt, die grafische OPEN LOOK-Benutzeroberflächen implementieren und sich ansonsten an die schriftlichen Sun-Lizenzvereinbarungen halten.

U.S. Government Rights – Commercial software. Regierungsbenutzer unterliegen der standardmäßigen Lizenzvereinbarung von Sun Microsystems, Inc. sowie den anwendbaren Bestimmungen der FAR und ihrer Zusätze.

DIE DOKUMENTATION WIRD "AS IS" BEREITGESTELLT, UND JEGLICHE AUSDRÜCKLICHE ODER IMPLIZITE BEDINGUNGEN, DARSTELLUNGEN UND HAFTUNG, EINSCHLIESSLICH JEGLICHER STILLSCHWEIGENDER HAFTUNG FÜR MARKTFÄHIGKEIT, EIGNUNG FÜR EINEN BESTIMMTEN ZWECK ODER NICHTÜBERTRETUNG WERDEN IM GESETZLICH ZULÄSSIGEN RAHMEN AUSDRÜCKLICH AUSGESCHLOSSEN.

Copyright 2005 Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 U.S.A. Tous droits réservés.

Ce produit ou document est protégé par un copyright et distribué avec des licences qui en restreignent l'utilisation, la copie, la distribution, et la décompilation. Aucune partie de ce produit ou document ne peut être reproduite sous aucune forme, par quelque moyen que ce soit, sans l'autorisation préalable et écrite de Sun et de ses bailleurs de licence, s'il y en a. Le logiciel détenu par des tiers, et qui comprend la technologie relative aux polices de caractères, est protégé par un copyright et licencié par des fournisseurs de Sun.

Des parties de ce produit pourront être dérivées du système Berkeley BSD licenciés par l'Université de Californie. UNIX est une marque déposée aux États-Unis et dans d'autres pays et licenciée exclusivement par X/Open Company, Ltd.

Sun, Sun Microsystems, le logo Sun, docs.sun.com, AnswerBook, AnswerBook2, et Solaris sont des marques de fabrique ou des marques déposées, de Sun Microsystems, Inc. aux États-Unis et dans d'autres pays. Toutes les marques SPARC sont utilisées sous licence et sont des marques de fabrique ou des marques déposées de SPARC International, Inc. aux États-Unis et dans d'autres pays. Les produits portant les marques SPARC sont basés sur une architecture développée par Sun Microsystems, Inc. ORACLE, Netscape

L'interface d'utilisation graphique OPEN LOOK et Sun™ a été développée par Sun Microsystems, Inc. pour ses utilisateurs et licenciés. Sun reconnaît les efforts de pionniers de Xerox pour la recherche et le développement du concept des interfaces d'utilisation visuelle ou graphique pour l'industrie de l'informatique. Sun détient une licence non exclusive de Xerox sur l'interface d'utilisation graphique Xerox, cette licence couvrant également les licenciés de Sun qui mettent en place l'interface d'utilisation graphique OPEN LOOK et qui en outre se conforment aux licences écrites de Sun.

CETTE PUBLICATION EST FOURNIE "EN L'ETAT" ET AUCUNE GARANTIE, EXPRESSE OU IMPLICITE, N'EST ACCORDEE, Y COMPRIS DES GARANTIES CONCERNANT LA VALEUR MARCHANDE, L'APTITUDE DE LA PUBLICATION A REpondre A UNE UTILISATION PARTICULIERE, OU LE FAIT QU'ELLE NE SOIT PAS CONTREFAISANTE DE PRODUIT DE TIERS. CE DENI DE GARANTIE NE S'APPLIQUERAIT PAS, DANS LA MESURE OU IL SERAIT TENU JURIDIQUEMENT NUL ET NON AVENU.



050816@12762



Inhalt

Vorwort 7

1	Einführung und Überblick	13
	Einführung in das Sun Cluster-System	14
	Drei Sichtweisen des Sun Cluster-Systems	15
	Hardwareinstallation und Service	15
	Systemverwaltung	16
	Anwendungsentwickler	17
	Sun Cluster-Systemaufgaben	18
2	Schlüsselkonzepte für Hardwaredienstleister	21
	Hardware- und Softwarekomponenten des Sun Cluster-Systems	21
	Cluster-Knoten	22
	Softwarekomponenten für Cluster-Hardwaremitglieder	23
	Multihost-Geräte	24
	Multi-Initiator-SCSI	25
	Lokale Platten	26
	Wechseldatenträger	26
	Cluster-Interconnect	27
	Öffentliche Netzwerkschnittstellen	27
	Client-Systeme	28
	Konsolenzugriffsgeräte	28
	Verwaltungskonsole	29
	SPARC: Sun Cluster-Topologien für SPARC	30
	SPARC: Geclusterte-Paare-Topologie für SPARC	30
	SPARC: Pair+N-Topologie für SPARC	31

SPARC: N+1 (Star)-Topologie für SPARC	32
SPARC: N*N (Scalable)-Topologie für SPARC	33
x86: Sun Cluster-Topologien für x86	34
x86: Geclusterte-Paar-Topologie für x86	34
3 Schlüsselkonzepte für Systemverwalter und Anwendungsentwickler	37
Verwaltungsschnittstellen	38
Cluster-Zeit	38
Hoch verfügbares Framework	39
Cluster-Mitgliedschaftsüberwachung (Cluster Membership Monitor, CMM)	40
Failfast-Mechanismus	41
Cluster-Konfigurations-Repository (Cluster Configuration Repository, CCR)	41
Globale Geräte	42
Geräte-IDs und DID-Pseudotreiber	42
Plattengerätegruppen	43
Plattengerätegruppen-Failover	44
Multiport-Plattengerätegruppen	45
Globaler Namensraum	47
Beispiel für lokale und globale Namensräume	48
Cluster-Dateisysteme	48
Verwenden von Cluster-Dateisystemen	49
HASToragePlus Ressourcentyp	50
Syncdir-Einhängeoption	50
Plattenpfadüberwachung	51
DPM - Überblick	51
Überwachen von Plattenpfaden	53
Quorum und Quorum-Geräte	55
Informationen zur Quorum-Stimmenanzahl	56
Informationen zum Fehlerschutz	57
Failfast-Mechanismus für den Fehlerschutz	58
Informationen zu Quorum-Konfigurationen	59
Erfüllen der Anforderungen für Quorum-Geräte	59
Einhalten von Empfehlungen für Quorum-Geräte	60
Empfohlene Quorum-Konfigurationen	61
Untypische Quorum-Konfigurationen	63
Unzulässige Quorum-Konfigurationen	63
Datendienste	65
Datendienstmethoden	68

Failover-Datendienste	68
Skalierbare Datendienste	68
Lastausgleichsverfahren	70
Failback-Einstellungen	72
Fehler-Monitore der Datendienste	72
Entwickeln neuer Datendienste	73
Eigenschaften von Scalable-Diensten	73
Datendienst-API und API der Datendienst-Entwicklungsbibliothek-API	74
Verwenden des Cluster-Interconnect für den Datendienstverkehr	75
Ressourcen, Ressourcengruppen und Ressourcentypen	76
Ressourcengruppen-Manager (RGM)	77
Zustände und Einstellungen für Ressourcen und Ressourcengruppen	77
Ressourcen- und Ressourcengruppeneigenschaften	79
Datendienst-Projektconfiguration	79
Bestimmen der Anforderungen der Projektconfiguration	82
Einstellen virtueller Speicherbegrenzungen für Prozesse	83
Failover-Szenarien	84
Öffentliche Netzwerkadapter und Internet Protocol (IP) Network Multipathing	89
SPARC: Unterstützung der dynamischen Rekonfiguration	91
SPARC: Allgemeine Beschreibung der dynamischen Rekonfiguration	92
SPARC: Erwägungen zur DR von CPU-Geräten im Cluster	93
SPARC: Erwägungen zur DR von Speichern im Cluster	93
SPARC: Erwägungen zur DR von Platten- und Bandlaufwerken im Cluster	93
SPARC: Erwägungen zur DR von Quorum-Geräten im Cluster	94
SPARC: Erwägungen zur DR von Cluster-Interconnect-Schnittstellen im Cluster	94
SPARC: Erwägungen zur DR von öffentlichen Netzwerkschnittstellen im Cluster	94
4 Fragen und Antworten	97
FAQs zur hohen Verfügbarkeit	97
FAQs zu Dateisystemen	98
FAQs zur Datenträgerverwaltung	99
FAQs zu Datendiensten	100
FAQs zum öffentlichen Netzwerk	101
FAQs zu Cluster-Mitgliedern	102
FAQs zum Cluster-Speicher	103
FAQs zum Cluster-Interconnect	103

FAQs zu Client-Systemen	104
FAQs zur Verwaltungskonsole	104
FAQs zu Terminal-Konzentrator und System Service Processor	105
Index	109

Vorwort

Das *Sun™ Cluster Konzepthandbuch für Solaris OS* enthält konzeptionelle und Referenzinformationen conceptual für das SunPlex™-System auf SPARC®- und x86-basierten Systemen.

Hinweis – In diesem Dokument bezieht sich der Begriff "x86" auf die Intel 32-Bit-Familie von Mikroprozessorchips sowie auf kompatible, von AMD hergestellte Mikroprozessorchips.

Zum SunPlex-System gehören alle Hardware- und Softwarekomponenten, aus denen die Sun Cluster-Lösung besteht.

Dieses Dokument ist für erfahrene Systemverwalter bestimmt, die mit der Sun Cluster-Software vertraut sind. Verwenden Sie dieses Dokument weder als Planungs- noch als Presales-Unterlage. Vor der Lektüre dieses Dokuments sollten die Systemanforderungen feststehen, und Sie sollten im Besitz der geeigneten Geräte und Software sein.

Damit Sie die in diesem Handbuch beschriebenen Konzepte verstehen, sollten Sie mit dem Solaris™Betriebssystem und mit Volume-Manager-Software vertraut sein, die mit dem Sun Cluster-System verwendet wird.

Hinweis – Sun Cluster-Software läuft auf zwei Plattformen, SPARC und x86. Die Informationen in diesem Dokument beziehen sich auf beide Plattformen, wenn nicht in einem eigenen Kapitel, Abschnitt, Anmerkung, Unterpunkt, Abbildung, Tabelle oder Beispiel anderweitige Angaben erfolgen.

Typografische Konventionen

Die folgende Tabelle beschreibt die in diesem Buch verwendeten typografischen Änderungen.

TABELLE P-1 Typografische Konventionen

Schriftart oder Symbol	Bedeutung	Beispiel
AaBbCc123	Die Namen von Befehlen, Dateien, Verzeichnissen; Bildschirmausgabe.	Bearbeiten Sie Ihre <code>.login</code> -Datei. Verwenden Sie <code>ls -a</code> , um eine Liste aller Dateien zu erhalten. Rechnername% Sie haben eine neue Nachricht.
AaBbCc123	Die Eingaben des Benutzers, im Gegensatz zu den Bildschirmausgaben des Computers	Rechner_name% su Passwort:
<i>AaBbCc123</i>	Befehlszeilen-Variable: durch einen realen Namen oder Wert ersetzen	Der Befehl zum Entfernen einer Datei lautet <code>rm filename</code> .
<i>AaBbCc123</i>	Buchtitel, neue Begriffe und Ausdrücke, die hervorgehoben werden sollen.	Lesen Sie hierzu Kapitel 6 im <i>Benutzerhandbuch</i> . Führen Sie eine <i>Patch-Analyse</i> durch. Speichern Sie die Datei <i>nicht</i> . [Einige Hervorhebungen werden online in Fettbuchstaben dargestellt.]

Beispiele für Shell-Eingabeaufforderungen in Befehlen

Die folgende Tabelle zeigt die Standard-Systemeingabeaufforderung und die Superuser-Eingabeaufforderung für die C-Shell, die Bourne-Shell und die Korn-Shell.

TABELLE P-2 Shell-Eingabeaufforderungen

Shell	Eingabeaufforderung
C Shell-Eingabeaufforderung	system%
C Shell-Superbenutzer-Eingabeaufforderung	system#
Bourne Shell- und Korn Shell-Eingabeaufforderung	\$
Bourne Shell- und Korn Shell-Superbenutzer-Eingabeaufforderung	#

Verwandte Dokumentation

Informationen zu verwandten Sun Cluster-Themen finden Sie in der Dokumentation, die in der folgenden Tabelle genannt ist. Die vollständige Sun Cluster-Dokumentation ist unter <http://docs.sun.com> verfügbar.

Thema	Dokumentation
Überblick	<i>Sun Cluster Überblick für das Betriebssystem Solaris</i>
Konzepte	<i>Sun Cluster Konzepthandbuch für Solaris OS</i>
Hardware-Installation und -Verwaltung	<i>Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS</i> Einzelne Hardwareverwaltungshandbücher
Softwareinstallation	<i>Sun Cluster Handbuch Softwareinstallation für Solaris OS</i>
Datendienstinstallation und -verwaltung	<i>Sun Cluster Data Services Planning and Administration Guide for Solaris OS</i> Einzelne Datendiensthandbücher
Datendienstentwicklung	<i>Sun Cluster Entwicklerhandbuch Datendienste für Solaris OS</i>
Systemverwaltung	<i>Sun Cluster Handbuch Systemverwaltung für Solaris OS</i>
Fehlermeldungen	<i>Sun Cluster Error Messages Guide for Solaris OS</i>
Befehle und Funktionen	<i>Sun Cluster Reference Manual for Solaris OS</i>

Eine vollständige Liste der Sun Cluster-Dokumentation finden Sie in den Versionshinweisen zu Ihrer Sun Cluster-Software unter <http://docs.sun.com>.

Zugriff auf die Online-Dokumentation von Sun

Über die Website docs.sun.comSM erhalten Sie Zugriff auf die technische Online-Dokumentation von Sun. Sie können das Archiv unter docs.sun.com durchsuchen oder nach einem bestimmten Buchtitel oder Thema suchen. Die URL lautet `http://docs.sun.com`.

Bestellen von Sun-Dokumentation

Ausgewählte Produktdokumentationen bietet Sun Microsystems auch in gedruckter Form an. Eine Liste der Dokumente und Hinweise zum Bezug finden Sie unter "Buy printed documentation" auf der Website `http://docs.sun.com`.

Hilfe anfordern

Wenden Sie sich im Falle von Problemen bei der Installation und Verwendung des Sun Cluster-Systems an Ihren Dienstleister, und geben Sie folgende Informationen an:

- Ihren Namen und E-Mail-Adresse (ggf.)
- Firmennamen, Adresse, Telefonnummer
- Modell und Seriennummern des Systems
- Versionsnummer des Betriebssystems (z.B. Solaris 9)
- Versionsnummer der Sun Cluster-Software (z. B. 3.1 8/05)

Sammeln Sie für Ihren Dienstleister mithilfe folgender Befehle Informationen zu allen Knoten.

Befehl	Funktion
<code>prtconf -v</code>	Zeigt die Größe des Systemspeichers an und gibt Informationen zu Peripheriegeräten zurück.
<code>psrinfo -v</code>	Zeigt Informationen zu Prozessoren an.

Befehl	Funktion
showrev -p	Gibt die installierten Korrekturversionen zurück.
SPARC: prtdiag -v	Zeigt Informationen zu Systemdiagnosen an.
scinstall -pv	Zeigt die Sun Cluster-Software- und Paketversion an.
scstat	Liefert einen Schnappschuss des Cluster-Status.
scconf -p	Listet die Cluster-Konfigurationsinformationen auf.
scrgadm -p	Zeigt Informationen zu installierten Ressourcen, Ressourcengruppen und Ressourcentypen an.

Halten Sie zudem den Inhalt der Datei `/var/adm/messages` bereit.

Produktschulung

Sun Microsystems bietet zahlreiche Schulungen unter Anleitung oder Selbstlernkurse zu vielen Sun-Technologien an. Informationen zu den von Sun angebotenen Schulungskursen und zur Anmeldung erhalten Sie auf der Schulungswebsite von Sun Microsystems unter <http://training.sun.com/>.

Einführung und Überblick

Das Sun Cluster-System ist eine integrierte Lösung aus Hardware und Sun Cluster-Software zur Erstellung von hoch verfügbaren und Scalable-Diensten.

Sun Cluster Konzepthandbuch für Solaris OS liefert die benötigten konzeptionellen Informationen für die wichtigste Zielgruppe der Sun Cluster-Dokumentation. Zu dieser Zielgruppe gehören

- Dienstleister, die Cluster-Hardware installieren und warten,
- Systemverwalter, die die Sun Cluster-Software installieren, konfigurieren und verwalten.
- Anwendungsentwickler, die Failover- und Scalable-Dienste für Anwendungen entwickeln, die derzeit nicht im Sun Cluster-Produkt enthalten sind.

Verwenden Sie dieses Handbuch zusammen mit der restlichen Sun Cluster-Dokumentationsreihe, um einen vollständigen Überblick über das Sun Cluster-System zu erhalten.

Dieses Kapitel

- Bietet eine Einführung in das Sun Cluster-System auf höchster Ebene.
- Beschreibt verschiedene Standpunkte der Sun Cluster-Zielgruppen.
- Gibt Schlüsselkonzepte an, deren Verständnis für das Arbeiten mit dem Sun Cluster-System erforderlich ist.
- Ordnet Schlüsselkonzepte der Sun Cluster-Dokumentation zu, die Informationen zu bestimmten Verfahrensweisen enthält.
- Zuordnung Cluster-spezifischer Aufgaben zur Dokumentation, in der die Verfahren zur Ausführung dieser Aufgaben enthalten sind.

Einführung in das Sun Cluster-System

Mit dem Sun Cluster-System wird das Solaris-Betriebssystem auf ein Cluster-Betriebssystem erweitert. Ein Cluster oder Plex ist eine Reihe von locker gekoppelten Computerknoten, die sich für den Client als ein einziges Netzwerk bzw. eine einzige Anwendung mit Datenbanken, Webdiensten und Dateidiensten darstellt.

Jeder Cluster-Knoten ist ein Standalone-Server, der seine eigenen Prozesse ausführt. Diese Prozesse kommunizieren miteinander und treten (für einen Netzwerk-Client) als ein einziges System auf. Im Zusammenspiel stellen sie Benutzern Anwendungen, Systemressourcen und Daten zur Verfügung.

Ein Cluster bietet im Vergleich zu traditionellen Einzelserversystemen mehrere Vorteile. Zu diesen Vorteilen gehört die Unterstützung von Failover- und Scalable-Diensten, die Fähigkeit zum modularen Wachsen und ein niedriger Anschaffungspreis im Vergleich zu traditionellen fehlertoleranten Hardwaresystemen.

Ziele des Sun Cluster-Systems:

- Reduzieren oder Beseitigen der Systemausfallzeiten aufgrund von Softwarefehlern oder Hardwareausfällen.
- Sicherstellen der Verfügbarkeit von Daten und Anwendungen für Endbenutzer, unabhängig von der Art des Fehlers, der normalerweise zum Ausfall eines Einzelserversystems führt.
- Erhöhen der Durchsatzleistung von Anwendungen, indem Dienste durch die Aufnahme weiterer Knoten in den Cluster auf zusätzliche Prozessoren skalierbar gemacht werden.
- Verbessern der Systemverfügbarkeit, weil Sie das System pflegen können, ohne den ganzen Cluster herunterfahren zu müssen.

Weitere Informationen zur Fehlertoleranz und zur hohen Verfügbarkeit erhalten Sie unter „Herstellen der Hochverfügbarkeit von Anwendungen mithilfe von Sun Cluster“ in *Sun Cluster Überblick für das Betriebssystem Solaris*.

Antworten auf häufig gestellte Fragen zur hohen Verfügbarkeit erhalten Sie unter [„FAQs zur hohen Verfügbarkeit“](#) auf Seite 97.

Drei Sichtweisen des Sun Cluster-Systems

In diesem Abschnitt werden drei verschiedene Sichtweisen des Sun Cluster-Systems, die entsprechenden Schlüsselkonzepte und die relevante Dokumentation für die jeweilige Sichtweise beschrieben. Diese Sichtweisen sind charakteristisch für folgende Fachleute:

- Hardware-Installations- und Servicepersonal
- Systemverwalter
- Anwendungsentwickler

Hardwareinstallation und Service

Für Hardwarespezialisten stellt sich das Sun Cluster-System als eine Sammlung aus Standard-Hardware dar, zu der Server, Netzwerke und Speicher gehören. Diese Komponenten werden mit Kabeln verbunden, so dass jede Komponente eine Sicherung hat und kein Single Point of Failure (Ausfallpunkt) mehr gegeben ist.

Schlüsselkonzepte – Hardware

Hardware-Servicespezialisten müssen die folgenden Cluster-Konzepte verstehen.

- Cluster-Hardwarekonfiguration und Verkabelung
- Installation und Wartung (Hinzufügen, Entfernen, Ersetzen):
 - Netzwerkschnittstellen-Komponenten (Adapter, Verbindungspunkte, Kabel)
 - Schnittstellenkarten für Platten
 - Platten-Arrays
 - Plattenlaufwerke
 - Verwaltungskonsole und Konsolenzugriffgerät
- Konfiguration der Verwaltungskonsole und des Konsolenzugriffgeräts

Weitere konzeptionelle Informationen zur Hardware

Die folgenden Abschnitte enthalten relevante Informationen zu den vorstehenden Schlüsselkonzepten:

- „Cluster-Knoten “ auf Seite 22
- „Multihost-Geräte “ auf Seite 24
- „Lokale Platten “ auf Seite 26

- „Cluster-Interconnect “ auf Seite 27
- „Öffentliche Netzwerkschnittstellen “ auf Seite 27
- „Client-Systeme “ auf Seite 28
- „Verwaltungskonsole “ auf Seite 29
- „Konsolenzugriffsgeräte “ auf Seite 28
- „SPARC: Geclusterte-Paare-Topologie für SPARC “ auf Seite 30
- „SPARC: N+1 (Star)-Topologie für SPARC “ auf Seite 32

Sun Cluster-Dokumentation für Hardware-Spezialisten

Das folgende Sun Cluster-Dokument enthält Verfahren und Informationen zu Konzepten für den Hardwareservice:

Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS

Systemverwaltung

Für Systemverwalter ist das Sun Cluster-System ein Satz aus verkabelten Servern (Knoten) und gemeinsam genutzten Speichergeräten. Der Systemverwalter betrachtet Software, die spezielle Aufgaben erfüllt:

- In die Solaris-Software integrierte Cluster-Spezialsoftware zum Überwachen der Konnektivität zwischen den Cluster-Knoten,
- Spezialsoftware zum Überwachen des Zustands der auf den Cluster-Knoten laufenden Benutzer-Anwendungsprogramme,
- Datenträgerverwaltungs-Software zum Konfigurieren und Verwalten von Platten,
- Cluster-Spezialsoftware, die allen Knoten Zugriff auf alle Speichergeräte ermöglicht, auch auf solche, die nicht direkt mit Platten verbunden sind,
- Cluster-Spezialsoftware, dank der Dateien auf allen Knoten angezeigt werden, als seien sie lokal mit diesem Knoten verbunden.

Schlüsselkonzepte –Systemverwaltung

Systemverwalter müssen die folgenden Konzepte und Prozesse verstehen:

- Interaktion zwischen Hardware- und Softwarekomponenten
- Allgemeiner Ablauf der Cluster-Installation und -Konfiguration, einschließlich:
 - Installieren des Solaris-Betriebssystems
 - Installieren und Konfigurieren der Sun Cluster-Software
 - Installieren und Konfigurieren eines Datenträger-Managers
 - Installieren und Konfigurieren der Anwendungssoftware für den Cluster-Betrieb

- Installieren und Konfigurieren der Sun Cluster-Datendienstsoftware
- Cluster-Verwaltungsverfahren für das Hinzufügen, Entfernen, Ersetzen und Pflegen der Cluster-Hardware- und Softwarekomponenten
- Konfigurationsänderungen zur Leistungssteigerung

Weitere konzeptionelle Informationen für Systemverwalter

Die folgenden Abschnitte enthalten relevante Informationen zu den vorstehenden Schlüsselkonzepten:

- „Verwaltungsschnittstellen “ auf Seite 38
- „Cluster-Zeit “ auf Seite 38
- „Hoch verfügbares Framework “ auf Seite 39
- „Globale Geräte “ auf Seite 42
- „Plattengerätegruppen “ auf Seite 43
- „Globaler Namensraum “ auf Seite 47
- „Cluster-Dateisysteme “ auf Seite 48
- „Plattenpfadüberwachung “ auf Seite 51
- „Informationen zum Fehlerschutz “ auf Seite 57
- „Datendienste “ auf Seite 65

Sun Cluster-Dokumentation für Systemverwalter

Die folgenden Sun Cluster-Dokumente enthalten Verfahren und Informationen zu Konzepten für die Systemverwaltung:

- *Sun Cluster Handbuch Softwareinstallation für Solaris OS*
- *Sun Cluster Handbuch Systemverwaltung für Solaris OS*
- *Sun Cluster Error Messages Guide for Solaris OS*
- *Sun Cluster 3.1 8/05 Versionshinweise für Solaris OS*
- *Sun Cluster 3.0-3.1 Release Notes Supplement*

Anwendungsentwickler

Das Sun Cluster-System stellt *Datendienste* für Anwendungen wie Oracle, NFS, DNS, Sun™ Java System Web Server, Apache Web Server (on SPARC based systems), and Sun Java System Directory Server bereit. Datendienste werden erstellt, indem Standard-Anwendungen für die Steuerung durch die Sun Cluster-Software konfiguriert werden. Die Sun Cluster-Software stellt Konfigurationsdateien und Verwaltungsmethoden zum Starten, Stoppen und Überwachen der Anwendungen zur Verfügung. Wenn Sie einen neuen Failover- oder Scalable-Dienst erstellen müssen, können Sie mit der Anwendungsprogrammierungsschnittstelle (API) von Sun Cluster und der DSET-API (API der Datendienst-Grundlagentechnologie) die Konfigurationsdateien und Verwaltungsmethoden erstellen, die zur Ausführung der Anwendung als Datendienst auf dem Cluster erforderlich sind.

Schlüsselkonzepte – Anwendungsentwicklung

Anwendungsentwickler müssen Folgendes verstehen:

- Die Eigenschaften der Anwendung, um festzustellen, ob diese zur Ausführung als Failover- oder Scalable-Dienst eingerichtet werden kann.
- Die Sun Cluster-API, DSET-API und den "generischen" Datendienste. Die Entwickler müssen entscheiden, welches Tool sich am besten für das Schreiben von Programmen oder Skripts eignet, mit denen ihre Anwendung für die Cluster-Umgebung konfiguriert werden soll.

Weitere konzeptionelle Informationen für Anwendungsentwickler

Die folgenden Abschnitte enthalten relevante Informationen zu den vorstehenden Schlüsselkonzepten:

- „Datendienste “ auf Seite 65
- „Ressourcen, Ressourcengruppen und Ressourcentypen “ auf Seite 76
- Kapitel 4

Sun Cluster-Dokumentation für Anwendungsentwickler

In den folgenden Sun Cluster-Dokumenten sind Verfahren und zugehörige Informationen für Anwendungsentwicklungskonzepte enthalten:

- *Sun Cluster Entwicklerhandbuch Datendienste für Solaris OS*
- *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*

Sun Cluster-Systemaufgaben

Für alle Sun Cluster-Systemaufgaben ist ein gewisser konzeptioneller Hintergrund erforderlich. Die nachstehende Tabelle gibt einen Überblick auf höchster Ebene über die Aufgaben und die Dokumentation, in der die Aufgabenschritte beschrieben sind. Die den Konzepten gewidmeten Abschnitte in diesem Buch beschreiben, wie die Konzepte den Aufgaben zugeordnet sind.

TABELLE 1–1 Task Map: Zordnung von Benutzeraufgaben zu Dokumentation

Schritt	Anweisungen
Installieren von Cluster-Hardware	<i>Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS</i>
Installieren der Solaris-Software auf dem Cluster	<i>Sun Cluster Handbuch Softwareinstallation für Solaris OS</i>
System TaskSPARC: Installieren der Sun™ Management Center-Software	<i>Sun Cluster Handbuch Softwareinstallation für Solaris OS</i>
Installieren und Konfigurieren der Sun Cluster-Software	<i>Sun Cluster Handbuch Softwareinstallation für Solaris OS</i>
Installieren und Konfigurieren der Datenträgerverwaltungs-Software	<i>Sun Cluster Handbuch Softwareinstallation für Solaris OS</i> Ihre Dokumentation zur Datenträgerverwaltung
Installieren und Konfigurieren der Sun Cluster-Datendienste	<i>Sun Cluster Data Services Planning and Administration Guide for Solaris OS</i>
Warten der Cluster-Hardware	<i>Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS</i>
Verwalten der Sun Cluster-Software	<i>Sun Cluster Handbuch Systemverwaltung für Solaris OS</i>
Verwalten der Datenträgerverwaltungs-Software	<i>Sun Cluster Handbuch Systemverwaltung für Solaris OS</i> und die Dokumentation zur Datenträgerverwaltung
Verwalten von Anwendungssoftware	Ihre Anwendungsdokumentation
Problemidentifizierung und empfohlene Benutzeraktionen	<i>Sun Cluster Error Messages Guide for Solaris OS</i>
Erstellen eines neuen Datendienstes	<i>Sun Cluster Entwicklerhandbuch Datendienste für Solaris OS</i>

Schlüsselkonzepte für Hardwaredienstleister

Dieses Kapitel beschreibt die Schlüsselkonzepte bezüglich der Hardwarekomponenten einer Sun Cluster-Systemkonfiguration. Folgende Themen werden behandelt:

- „Cluster-Knoten “ auf Seite 22
- „Multihost-Geräte “ auf Seite 24
- „Lokale Platten “ auf Seite 26
- „Wechseldatenträger “ auf Seite 26
- „Cluster-Interconnect “ auf Seite 27
- „Öffentliche Netzwerkschnittstellen “ auf Seite 27
- „Client-Systeme “ auf Seite 28
- „Konsolenzugriffsgeräte “ auf Seite 28
- „Verwaltungskonsole “ auf Seite 29
- „SPARC: Sun Cluster-Topologien für SPARC “ auf Seite 30
- „x86: Sun Cluster-Topologien für x86 “ auf Seite 34

Hardware- und Softwarekomponenten des Sun Cluster-Systems

Diese Informationen richten sich in erster Linie an Hardwaredienstleister. Diese Konzepte helfen Dienstleistern dabei, die Beziehungen zwischen den Hardwarekomponenten zu verstehen, bevor sie Cluster-Hardware installieren, konfigurieren oder warten. Auch für Cluster-Systemverwalter können diese Informationen als Hintergrund für das Installieren, Konfigurieren und Verwalten von Cluster-Software nützlich sein.

Ein Cluster besteht aus mehreren Hardwarekomponenten einschließlich:

- Cluster-Knoten mit lokalen Platten (nicht gemeinsam genutzt)
- Multihost-Speicher (von Knoten gemeinsam benutzte Platten)

- Wechseldatenträger (Bänder und CD-ROMs)
- Cluster-Interconnect
- Öffentliche Netzwerkschnittstellen
- Client-Systeme
- Verwaltungskonsole
- Konsolenzugriffsgeräte

Das Sun Cluster-System ermöglicht Ihnen das Kombinieren dieser Komponenten in einer Vielzahl von Konfigurationen. Diese Konfigurationen werden in den folgenden Abschnitten beschrieben.

- „SPARC: Sun Cluster-Topologien für SPARC “ auf Seite 30
- „x86: Sun Cluster-Topologien für x86 “ auf Seite 34

Eine Darstellung einer Beispiel-Cluster-Konfiguration mit zwei Knoten finden Sie unter „Sun Cluster-Hardwareumgebung“ in *Sun Cluster Überblick für das Betriebssystem Solaris*.

Cluster-Knoten

Ein Cluster-Knoten ist ein Computer, auf dem sowohl das Solaris-Betriebssystem als auch die Sun Cluster-Software ausgeführt wird. Ein Cluster-Knoten ist somit entweder ein aktuelles Mitglied des Clusters (ein *Cluster-Mitglied*) oder ein potenzielles Mitglied.

- Die SPARC: Sun Cluster-Software unterstützt bis zu sechzehn Knoten in einem Cluster. Informationen zu den unterstützten Knotenkonfigurationen finden Sie unter „SPARC: Sun Cluster-Topologien für SPARC “ auf Seite 30.
- Die x86: Sun Cluster-Software unterstützt zwei Knoten in einem Cluster. Informationen zu den unterstützten Knotenkonfigurationen finden Sie unter „x86: Sun Cluster-Topologien für x86 “ auf Seite 34.

Cluster-Knoten sind im Allgemeinen an mindestens ein Multihost-Gerät angeschlossen. Nicht an Multihost-Geräte angeschlossene Knoten verwenden das Cluster-Dateisystem zum Zugriff auf die Multihost-Geräte. In einer Scalable-Dienste-Konfiguration können die Knoten beispielsweise Anfragen abwickeln, ohne direkt mit den Multihost-Geräten verbunden zu sein.

Daneben teilen sich Knoten in Konfigurationen paralleler Datenbanken den gleichzeitigen Zugriff auf alle Platten.

- Informationen zum gleichzeitigen Plattenzugriff finden Sie unter „Multihost-Geräte “ auf Seite 24.
- Weitere Informationen zu Konfigurationen paralleler Datenbanken finden Sie unter „SPARC: Geclusterte-Paare-Topologie für SPARC “ auf Seite 30 und „x86: Geclusterte-Paar-Topologie für x86 “ auf Seite 34.

Alle Knoten im Cluster sind unter einem gemeinsamen Namen zusammengefasst - dem Cluster-Namen - der für den Cluster-Zugriff und die Cluster-Verwaltung verwendet wird.

Öffentliche Netzwerkadapter verbinden die Knoten mit den öffentlichen Netzwerken und ermöglichen den Zugriff der Clients auf den Cluster.

Die Mitglieder eines Clusters kommunizieren mit den anderen Knoten im Cluster über ein oder mehrere physisch unabhängige Netzwerke. Diese Gruppe physisch unabhängiger Netzwerke wird als *Cluster-Interconnect* bezeichnet.

Jeder Knoten im Cluster nimmt wahr, wenn ein anderer Knoten dem Cluster beiträgt oder diesen verlässt. Daneben nimmt jeder Knoten sowohl die lokalen Ressourcen als auch die Ressourcen auf den anderen Cluster-Knoten wahr.

Knoten innerhalb desselben Clusters sollten eine vergleichbare Verarbeitungs-, Speicher- und E/A-Kapazität aufweisen, so dass ein Failover ohne nennenswerten Leistungsabfall möglich ist. Aufgrund der Möglichkeit eines Failover muss jeder Knoten genügend überschüssige Kapazität aufweisen, um die Aufgaben aller Knoten zu übernehmen, für die er als Sicherheits- oder Sekundärknoten festgelegt ist.

Jeder Knoten bootet sein eigenes, individuelles Root-Dateisystem (/).

Softwarekomponenten für Cluster-Hardwaremitglieder

Auf einem Knoten, der als Cluster-Mitglied eingesetzt werden soll, muss folgende Software installiert sein:

- Solaris-Betriebssystem
- Sun Cluster-Software
- Datendienstanwendung
- Datenträgerverwaltung (Solaris Volume Manager™ oder VERITAS Volume Manager)

Ausgenommen sind Konfigurationen, die Hardware-RAID (Hardware Redundant Array of Independent Disks) verwenden. Diese Konfiguration erfordert möglicherweise keine Datenträgerverwaltungs-Software wie Solaris Volume Manager oder VERITAS Volume Manager.

- Im *Sun Cluster Handbuch Softwareinstallation für Solaris OS* finden Sie Informationen zur Installation des Solaris-Betriebssystems, von Sun Cluster und der Datenträgerverwaltungs-Software.
- Im *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* finden Sie Informationen zur Installation und Konfiguration der Datendienste.
- In [Kapitel 3](#) finden Sie konzeptionelle Informationen zu den zuvor genannten Softwarekomponenten.

Die folgende Abbildung bietet einen Überblick über die Softwarekomponenten, die gemeinsam die Sun Cluster-Softwareumgebung bilden.

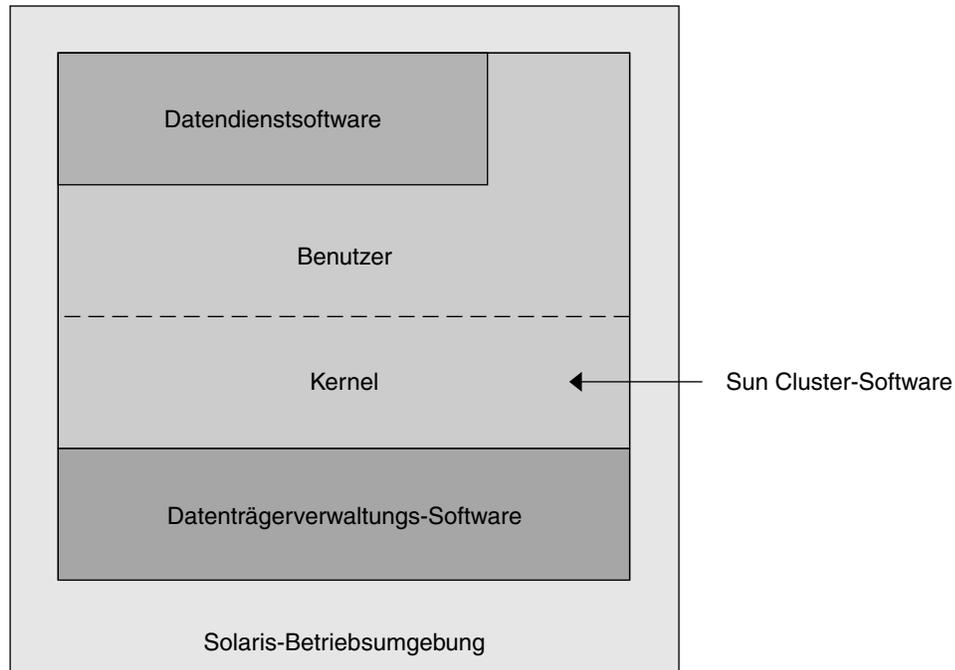


ABBILDUNG 2-1 Beziehung zwischen den Sun Cluster-Softwarekomponenten

In [Kapitel 4](#) finden Sie Fragen und Antworten zu den Cluster-Mitgliedern.

Multihost-Geräte

Platten, die mit mehr als einem Knoten gleichzeitig verbunden werden können, sind Multihost-Geräte. Durch die Multihost-Speicherung sind die Platten in der Sun Cluster-Umgebung hoch verfügbar. Die Sun Cluster-Software erfordert die Multihost-Speicherung zur Schaffung eines Quorums in Clustern mit zwei Knoten. Cluster mit mehr als zwei Knoten erfordern keine Quorum-Geräte. Weitere Informationen zum Quorum finden Sie unter „[Quorum und Quorum-Geräte](#)“ auf Seite 55.

Multihost-Geräte haben folgende Eigenschaften.

- Toleranz einzelner Knotenausfälle.
- Speicherung von Anwendungsdaten, Anwendungsbinärdateien und Konfigurationsdateien.
- Schutz vor Knotenversagen. Wenn Clients Daten über einen Knoten anfordern und dieser Knoten ausfällt, werden die Anfragen umgeschaltet und über einen anderen Knoten abgewickelt, der mit denselben Platten direkt verbunden ist.

- Globaler Zugriff über einen Primärknoten, der die Platten "unterstützt" oder direkter gleichzeitiger Zugriff über lokale Pfade. Die einzige Anwendung, die zurzeit den direkten gleichzeitigen Zugriff nutzt, ist Oracle Real Application Clusters Guard.

Ein Datenträger-Manager sorgt bei gespiegelten oder RAID-5-Konfigurationen für die Datenredundanz der Multihost-Geräten. Zurzeit unterstützt Sun Cluster die Programme Solaris Volume Manager und VERITAS Volume Manager, die nur in SPARC-basierten Clustern als Datenträger-Manager verwendet werden können, sowie den RDAC RAID-5-Hardwarecontroller auf verschiedenen Hardware-RAID-Plattformen.

Die Kombination aus Multihost-Geräten und Plattenspiegelung sowie Striping schützt vor Knoten- und Einzelplattenversagen.

In [Kapitel 4](#) finden Sie Fragen und Antworten zur Multihost-Speicherung.

Multi-Initiator-SCSI

Dieser Abschnitt betrifft nur SCSI-Speichergeräte und keine Glasfaserkanal-Speicher, die für Multihost-Geräte eingesetzt werden.

Auf einem Standalone-Server steuert der Serverknoten die SCSI-Busaktivitäten durch die SCSI-Host-Adapterschaltung, die diesen Server mit einem bestimmten SCSI-Bus verbindet. Diese SCSI-Hostadapterschaltung wird als *SCSI-Initiator* bezeichnet. Die Schaltung initiiert alle Busaktivitäten für diesen SCSI-Bus. Die SCSI-Standardadresse der SCSI-Host-Adapter in Sun-Systemen ist 7.

Cluster-Konfigurationen teilen ihren Speicher unter Verwendung von Multihost-Geräten auf mehrere Serverknoten auf. Wenn der Cluster-Speicher aus einpolig geerdeten oder differenziellen SCSI-Geräten besteht, wird die Konfiguration als Multi-Initiator-SCSI bezeichnet. Wie aus dieser Terminologie hervorgeht, ist mehr als ein SCSI-Initiator auf dem SCSI-Bus vorhanden ist.

Die SCSI-Spezifikation erfordert, dass jedes Gerät auf einem SCSI-Bus eine eindeutige SCSI-Adresse aufweist. (Der Host-Adapter ist ebenfalls ein Gerät im SCSI-Bus.) Die Standard-Hardwarekonfiguration führt in einer Multi-Initiator-Umgebung zu einem Konflikt, weil alle SCSI-Host-Adapter standardmäßig auf 7 gesetzt werden.

Lassen Sie zur Beseitigung dieses Konflikts auf jedem SCSI-Bus einen der SCSI-Host-Adapter mit der SCSI-Adresse 7, und stellen Sie für die anderen Host-Adapter ungenutzte SCSI-Adressen ein. Eine zweckmäßige Planung schreibt vor, dass sowohl die derzeit ungenutzten als auch die potenziell ungenutzten Adressen zu den "ungenutzten" SCSI-Adressen gehören. Ein Beispiel für zukünftige, ungenutzte Adressen ist das Aufstocken des Speichers durch das Installieren neuer Laufwerke an freien Laufwerk-Steckplätzen.

In den meisten Konfigurationen ist 6 die verfügbare SCSI-Adresse für einen zweiten Host-Adapter.

Sie können die ausgewählten SCSI-Adressen für diese Host-Adapter ändern, indem Sie eines der folgenden Tools zum Einstellen der `scsi-initiator-id`-Eigenschaft verwenden:

- `eeeprom(1M)`
- Das OpenBoot PROM in einem SPARC-basierten System
- Das SCSI-Dienstprogramm, das Sie optional nach dem Booten der BIOS in einem x86-basierten System aufrufen

Sie können diese Eigenschaft global für einen Knoten oder für jeden Host-Adapter einzeln einstellen. Anweisungen zum Festlegen einer eindeutigen `scsi-initiator-id` für jeden SCSI-Hostadapter finden Sie im *Sun Cluster 3.0-3.1 With SCSI JBOD Storage Device Manual for Solaris OS*.

Lokale Platten

Lokale Platten sind Platten, die nur mit einem einzigen Knoten verbunden sind. Lokale Platten sind daher nicht vor Knotenversagen geschützt (nicht hoch verfügbar). Alle Platten, einschließlich lokaler Platten, sind im globalen Namespace enthalten und als *globale Geräte* konfiguriert. Aus diesem Grund sind die Platten selbst von allen Cluster-Knoten aus sichtbar.

Sie können die Dateisysteme auf lokalen Platten für andere Knoten verfügbar machen, indem Sie sie unterhalb eines globalen Einhängpunkts platzieren. Versagt der Knoten, an dem derzeit eines dieser globalen Dateisysteme eingehängt ist, kann kein Knoten auf dieses Dateisystem zugreifen. Mithilfe eines Datenträger-Managers können Sie diese Platten spiegeln, so dass ein Fehler nicht dazu führt, dass kein Zugriff mehr auf diese Dateisysteme besteht; Datenträger-Manager schützen aber nicht vor Knotenversagen.

Im Abschnitt „*Globale Geräte*“ auf Seite 42 finden Sie weitere Informationen zu globalen Geräten.

Wechseldatenträger

Wechselmedien wie Band- und CD-ROM-Laufwerke werden in einem Cluster unterstützt. Im Allgemeinen werden diese Geräte auf dieselbe Weise installiert, konfiguriert und betrieben wie in einer Umgebung ohne Cluster. Diese Geräte werden in Sun Cluster als globale Geräte konfiguriert, sodass von jedem Knoten im Cluster auf jedes Gerät zugegriffen werden kann. Informationen zur Installation und Konfiguration von Wechseldatenträgern finden Sie im *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*.

Im Abschnitt „*Globale Geräte*“ auf Seite 42 finden Sie weitere Informationen zu globalen Geräten.

Cluster-Interconnect

Der *Cluster-Interconnect* ist die reale Konfiguration der Geräte, die für die Übertragung von Cluster-privaten Kommunikationen und Datendienstkommunikationen zwischen Cluster-Knoten eingesetzt wird. Der Interconnect wird umfassend für Cluster-interne Kommunikationen verwendet und kann dadurch die Leistung begrenzen.

Nur Cluster-Knoten können mit dem Cluster-Interconnect verbunden werden. Das Sun Cluster-Sicherheitsmodell setzt voraus, dass nur Cluster-Knoten physischen Zugriff auf den Cluster-Interconnect haben.

Alle Knoten müssen mit dem Cluster-Interconnect über mindestens zwei redundante, real unabhängige Netzwerke oder Pfade verbunden sein, um einen Single Point of Failure (Ausfallpunkt) zu vermeiden. Sie können mehrere real unabhängige Netzwerke (bis zu sechs) zwischen zwei beliebigen Knoten einrichten.

Der Cluster-Interconnect besteht aus drei Hardwarekomponenten: Adapter, Verbindungspunkten und Kabeln. Die nachstehende Liste beschreibt jede dieser Hardwarekomponenten.

- **Adapter** – Die Netzwerk-Schnittstellenkarten, die sich in jedem Cluster-Knoten befinden. Ihr Namen setzt sich aus dem Gerätenamen, unmittelbar gefolgt von einer realen Einheitsnummer, zum Beispiel qfe2, zusammen. Manche Adapter haben nur eine reale Netzwerkverbindung, andere hingegen, wie die qfe-Karte, haben mehrere reale Verbindungen. Manche enthalten sowohl Netzwerk- als auch Speicherschnittstellen.

Ein Netzwerkadapter mit mehreren Schnittstellen kann zu einem Single Point of Failure werden, wenn der ganze Adapter ausfällt. Für maximale Verfügbarkeit sollten Sie Ihren Cluster so planen, dass der einzige Pfad zwischen zwei Knoten nicht von einem einzigen Netzwerkadapter abhängt.

- **Verbindungspunkte** – Die Schalter, die außerhalb der Cluster-Knoten liegen. Verbindungspunkte übernehmen Pass-Through- und Umschaltfunktionen, damit Sie mehr als zwei Knoten zusammenschließen können. In einem Zwei-Knoten-Cluster benötigen Sie keine Verbindungspunkte, weil die Knoten mit redundanten realen Kabeln, die an die redundanten Adapter auf jedem Knoten angeschlossen sind, direkt miteinander verbunden werden können. Für Konfigurationen mit mehr als zwei Knoten sind normalerweise Verbindungspunkte erforderlich.
- **Kabel** – Die physischen Verbindungen, die Sie zwischen zwei Netzwerkadaptern oder zwischen einem Adapter und einem Verbindungspunkt installieren.

In [Kapitel 4](#) finden Sie Fragen und Antworten zum Cluster-Interconnect.

Öffentliche Netzwerkschnittstellen

Die Verbindung der Clients mit dem Cluster erfolgt über die öffentlichen Netzwerkschnittstellen. Jede Netzwerkadapterkarte kann an ein oder mehrere öffentliche Netzwerke angeschlossen sein, je nachdem, ob die Karte mehrere

Hardwarechnittstellen hat. Sie können Knoten mit mehreren öffentlichen Netzwerkschnittstellenkarten einrichten, die so konfiguriert sind, dass mehrere Karten aktiv sind und dadurch gegenseitig als Failover-Sicherung dienen. Wenn einer der Adapter ausfällt, wird die Internet Protocol (IP) Network Multipathing-Software aufgerufen, um die defekte Schnittstelle durch einen anderen Adapter in der Gruppe zu ersetzen.

Es gibt keine hardwarespezifischen Erwägungen im Zusammenhang mit der Cluster-Bildung für öffentliche Netzwerke.

In [Kapitel 4](#) finden Sie Fragen und Antworten zu öffentlichen Netzwerken.

Client-Systeme

Client-Systeme beinhalten Workstations oder andere Server, die über das öffentliche Netzwerk auf den Cluster zugreifen. Client-seitige Programme verwenden Daten oder andere Dienste, die von serverseitigen Anwendungen bereitgestellt werden, die im Cluster ausgeführt werden.

Client-Systeme sind nicht hoch verfügbar. Daten und Anwendungen auf dem Cluster sind hoch verfügbar.

In [Kapitel 4](#) finden Sie Fragen und Antworten zu Client-Systemen.

Konsolenzugriffsgeräte

Sie benötigen Konsolenzugriff auf alle Cluster-Knoten. Verwenden Sie eines der folgenden Geräte, um auf die Konsole zuzugreifen:

- Den Terminal-Konzentrator, den Sie zusammen mit der Cluster-Hardware erworben haben
- Den System Service Processor (SSP) auf Sun Enterprise E10000-Servern (für SPARC-basierte Cluster)
- Den Systemcontroller auf Sun Fire™-Servern (ebenfalls für SPARC-basierte Cluster)
- Ein anderes Gerät, das auf jedem Knoten auf `ttya` zugreifen kann.

Bei Sun ist nur ein unterstützter Terminal-Konzentrator verfügbar, dessen Verwendung optional ist. Der Terminal-Konzentrator ermöglicht den Zugriff auf `/dev/console` auf jedem Knoten über ein TCP/IP-Netzwerk. Damit kann auf Konsolenebene von einer Remote-Workstation im Netzwerk aus auf jeden Knoten zugegriffen werden.

Der System Service Processor (SSP) ermöglicht den Konsolenzugriff für Sun Enterprise E1000-Server. Der SSP ist ein Computer in einem Ethernet-Netzwerk, der zur Unterstützung des Sun Enterprise E1000-Servers konfiguriert wurde. Der SSP dient als Verwaltungskonsolle für den Sun Enterprise E1000-Server. Mit der Sun Enterprise E1000-Netzwerkkonsolenfunktion kann jede Workstation im Netzwerk eine Host-Konsolensitzung öffnen.

Weitere Möglichkeiten für den Konsolenzugriff bieten Terminal-Konzentratoren, der `tip(1)`-Zugriff über den seriellen Anschluss eines anderen Knotens sowie Terminals ohne Verarbeitungsfunktionen. Sie können Tastaturen und Monitore von Sun™ oder andere Geräte für den seriellen Port verwenden, wenn Ihr Hardwaredienstleister diese unterstützt.

Verwaltungskonsolle

Als *Verwaltungskonsolle* für den aktiven Cluster können Sie eine dedizierte UltraSPARC®-Workstation oder einen Sun Fire V65x-Server verwenden. In der Regel installieren Sie auf der Verwaltungskonsolle Verwaltungssoftware wie das Cluster Control Panel (CCP) und das Sun Cluster-Modul für das Sun Management Center-Produkt (ausschließlich für die Verwendung mit SPARC-basierten Clustern). Mit `cconsole` im CCP können Sie mehr als eine Knotenkonsolle gleichzeitig verbinden. Weitere Informationen zur Verwendung des CCP finden Sie in Kapitel 1, „Einführung in die Verwaltung von Sun Cluster“ in *Sun Cluster Handbuch Systemverwaltung für Solaris OS*.

Die Verwaltungskonsolle ist kein Cluster-Knoten. Sie setzen die Verwaltungskonsolle für den Remote-Zugriff auf Cluster-Knoten entweder über öffentliche Netzwerke oder optional über einen netzwerkbasieren Terminal-Konzentrator ein. Wenn Ihr Cluster die Sun Enterprise E1000-Plattform umfasst, müssen Sie sich über die Verwaltungskonsolle der SSP anmelden und über den Befehl `netcon(1M)` eine Verbindung herstellen.

In der Regel konfigurieren Sie die Knoten ohne Monitore. Anschließend greifen Sie an der Verwaltungskonsolle über eine `telnet`-Sitzung auf die Konsolle des Knotens zu. Die Verwaltungskonsolle ist mit einem Terminal-Konzentrator und dieser wiederum mit dem seriellen Anschluss des Knotens verbunden. Bei einem Sun Enterprise E1000-Server stellen Sie die Verbindung über den System Service Processor her. Weitere Informationen hierzu finden Sie unter „Konsolenzugriffsgeräte“ auf Seite 28.

Sun Cluster erfordert keine dedizierte Verwaltungskonsolle. Die Verwendung einer Verwaltungskonsolle hätte jedoch folgende Vorteile:

- Sie ermöglicht eine zentralisierte Cluster-Verwaltung durch das Gruppieren von Konsolen- und Verwaltungstools auf demselben Rechner.
- Sie sorgt für eine potenziell schnellere Problemlösung durch Ihren Hardware-Dienstleister.

In Kapitel 4 finden Sie Fragen und Antworten zur Verwaltungskonsolle.

SPARC: Sun Cluster-Topologien für SPARC

Eine Topologie ist das Verbindungsschema, nach dem die Cluster-Knoten mit den in der Sun Cluster-Umgebung verwendeten Speicherplattformen verbunden sind. Sun Cluster unterstützt jede Topologie, die folgende Richtlinien einhält.

- Eine Sun Cluster-Umgebung unterstützt in SPARC-basierten Systemen maximal sechzehn Knoten in einem Cluster, unabhängig von den Speicherkonfigurationen, die Sie implementieren.
- Ein gemeinsam genutztes Speichergerät kann mit so vielen Knoten verbunden werden, wie vom Speichergerät unterstützt werden.
- Gemeinsam genutzte Speichergeräte müssen nicht mit allen Knoten im Cluster verbunden sein. Diese Speichergeräte müssen jedoch mit mindestens zwei Knoten verbunden sein.

Die Sun Cluster-Software schreibt für die Konfiguration eines Clusters keine spezifischen Topologien vor. Die Beschreibung der folgenden Topologien soll das Vokabular zur Erläuterung eines Cluster-Verbindungsschemas erschließen. Diese Topologien zeigen typische Verbindungsschemata.

- Geclustertes Paar
- Paar+N
- N+1 (Stern)
- N*N (Scalable)

Die folgenden Abschnitte enthalten Beispielgrafiken für jede Topologie.

SPARC: Geclusterte-Paare-Topologie für SPARC

Eine Geclusterte-Paare-Topologie besteht aus zwei oder mehreren Knotenpaaren, die unter einem einzigen Cluster-Verwaltungs-Framework arbeiten. In dieser Konfiguration wird ein Failover nur zwischen einem Paar ausgeführt. Dabei sind jedoch alle Knoten über den Cluster-Interconnect verbunden und arbeiten unter der Sun Cluster-Softwaresteuerung. Mit dieser Topologie können Sie eine parallele Datenbankanwendung auf einem Paar und eine Failover- oder Scalable-Anwendung auf einem anderen Paar ausführen.

Mithilfe des Cluster-Dateisystems können Sie auch eine Konfiguration mit zwei Paaren erstellen. Dabei führen mehr als zwei Knoten einen Scalable-Dienst oder eine parallele Datenbank aus, selbst wenn nicht alle Knoten mit den Platten verbunden sind, auf denen die Anwendungsdaten gespeichert werden.

Die nachstehende Abbildung zeigt eine Geclusterte-Paare-Konfiguration.

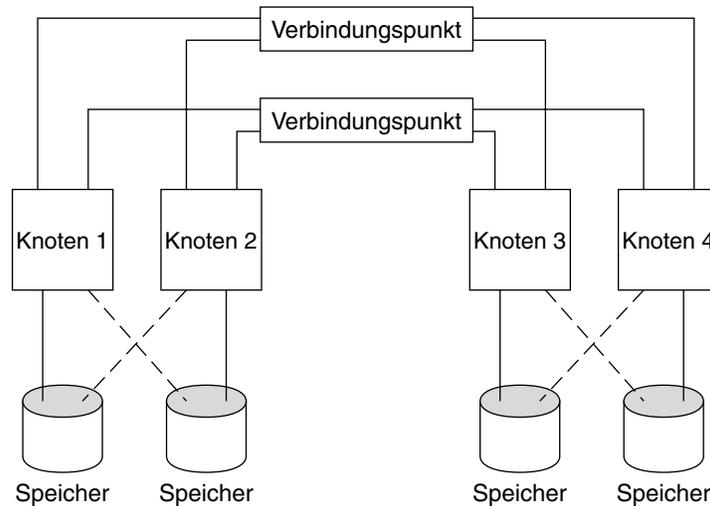


ABBILDUNG 2-2 SPARC: Geclusterte-Paar-Topologie

SPARC: Pair+N-Topologie für SPARC

Die Paar+N-Topologie umfasst ein direkt mit einem gemeinsam genutzten Speicher verbundenes Knotenpaar und einen weiteren Knotensatz, der den Cluster-Interconnect für den Zugriff auf den gemeinsam genutzten Speicher einsetzt – sie sind nicht direkt miteinander verbunden.

Die nachstehende Abbildung stellt eine Paar+N-Topologie dar, in der zwei der vier Knoten (Knoten 3 und Knoten 4) den Cluster-Interconnect für den Speicherzugriff verwenden. Diese Konfiguration kann um zusätzliche Knoten erweitert werden, die keinen direkten Zugriff auf den gemeinsam genutzten Speicher haben.

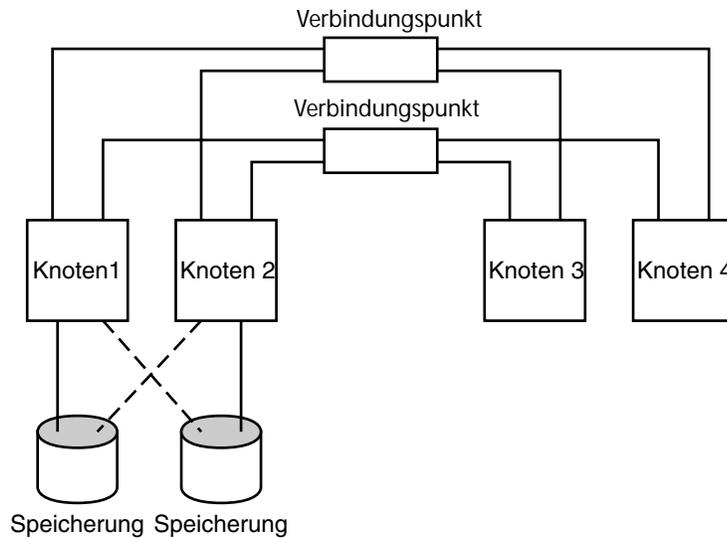


ABBILDUNG 2-3 Paar+N-Topologie

SPARC: N+1 (Star)-Topologie für SPARC

Eine N+1-Topologie umfasst mehrere Primärknoten und einen Sekundärknoten. Sie müssen die Primärknoten und den Sekundärknoten nicht identisch konfigurieren. Die Primärknoten stellen aktiv Anwendungsdienste zur Verfügung. Der Sekundärknoten muss sich nicht im Leerlauf befinden, während er auf den Ausfall eines Primärknotens wartet.

Der Sekundärknoten ist der einzige Knoten in der Konfiguration, der real mit dem ganzen Multihostspeicher verbunden ist.

Wenn ein Fehler in einem Primärknoten auftritt, führt Sun Cluster ein Failover der Ressourcen auf den Sekundärknoten durch, auf dem die Ressourcen bleiben, bis sie (entweder automatisch oder manuell) wieder auf den Primärknoten umgeschaltet werden.

Der Sekundärknoten muss immer über eine ausreichende CPU-Überkapazität verfügen, um die Last übernehmen zu können, falls einer der Primärknoten ausfällt.

Die nachstehende Abbildung stellt eine N+1-Konfiguration dar.

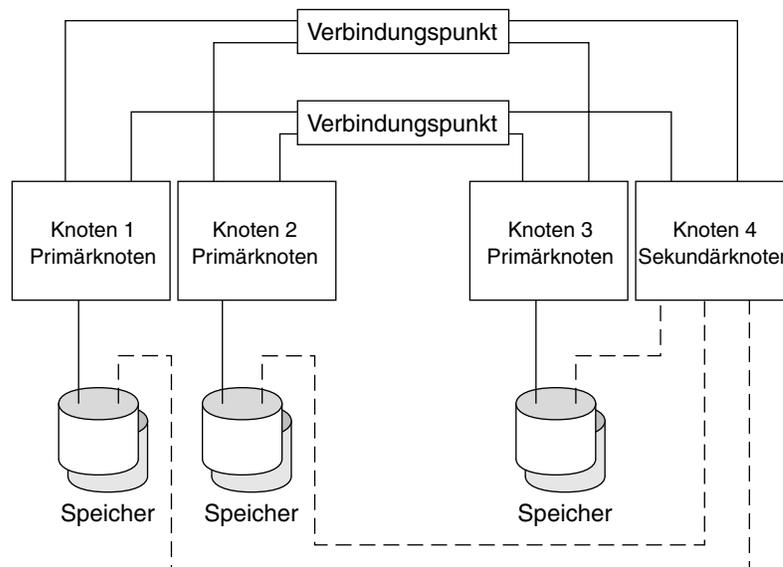


ABBILDUNG 2-4 SPARC: N+1-Topologie

SPARC: N*N (Scalable)-Topologie für SPARC

Mit einer N*N-Topologie können alle gemeinsam genutzten Speichergeräte in einem Cluster mit allen Knoten im Cluster verbunden werden. Mit dieser Topologie können hoch verfügbare Anwendungen ohne Abfall der Dienstqualität von einem Knoten auf einen anderen umgeleitet werden. Bei einem Failover kann der neue Knoten über einen lokalen Pfad auf das Speichergerät zugreifen, anstatt den privaten Interconnect zu verwenden.

Die nachstehende Abbildung stellt eine N*N-Konfiguration dar.

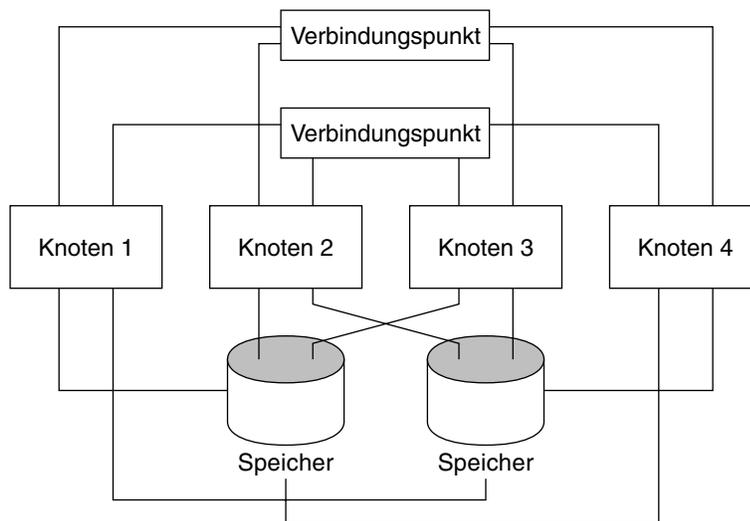


ABBILDUNG 2-5 SPARC: N*N-Topologie

x86: Sun Cluster-Topologien für x86

Eine Topologie ist das Verbindungsschema, nach dem die Cluster-Knoten mit den im Cluster verwendeten Speicherplattformen verbunden sind. Die Sun Cluster-Software unterstützt jede Topologie, die folgende Richtlinien einhält.

- Sun Cluster unterstützt in x86-basierten Systemen zwei Knoten in einem Cluster.
- Gemeinsam genutzte Speichergeräte müssen mit beiden Knoten verbunden sein.

Sun Cluster schreibt für die Konfiguration eines Clusters keine spezifischen Topologien vor. Die Beschreibung der folgenden Geclusterte-Paare-Topologie soll das Vokabular zur Erläuterung eines Cluster-Verbindungsschemas erschließen. Es handelt sich um die einzig mögliche Topologie für Cluster aus x86-basierten Knoten. Diese Topologie stellt ein typisches Verbindungsschema dar.

Der folgende Abschnitt enthält eine Beispielgrafik der Topologie.

x86: Geclusterte-Paar-Topologie für x86

Eine Geclusterte-Paare-Topologie besteht aus zwei oder mehreren Knotenpaaren, die unter einem einzigen Cluster-Verwaltungs-Framework arbeiten. In dieser Konfiguration wird ein Failover nur zwischen einem Paar ausgeführt. Dabei sind

jedoch alle Knoten über den Cluster-Interconnect verbunden und arbeiten unter der Sun Cluster-Softwaresteuerung. Mit dieser Topologie können Sie eine parallele Datenbankanwendung auf einem Paar und eine Failover- oder Scalable-Anwendung auf einem anderen Paar ausführen.

Die nachstehende Abbildung zeigt eine Geclusterte-Paare-Konfiguration.

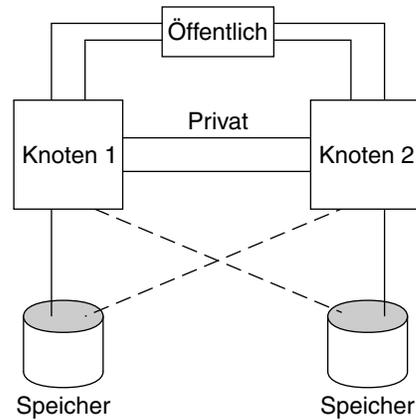


ABBILDUNG 2-6 x86: Geclusterte-Paar-Topologie

Schlüsselkonzepte für Systemverwalter und Anwendungsentwickler

In diesem Kapitel werden die Schlüsselkonzepte im Zusammenhang mit den Softwarekomponenten eines Sun Cluster-Systems beschrieben. Zu den behandelten Themen gehören:

- „Verwaltungsschnittstellen “ auf Seite 38
- „Cluster-Zeit “ auf Seite 38
- „Hoch verfügbares Framework “ auf Seite 39
- „Globale Geräte “ auf Seite 42
- „Plattengerätegruppen “ auf Seite 43
- „Globaler Namensraum “ auf Seite 47
- „Cluster-Dateisysteme “ auf Seite 48
- „Plattenpfadüberwachung “ auf Seite 51
- „Quorum und Quorum-Geräte “ auf Seite 55
- „Datendienste “ auf Seite 65
- „Verwenden des Cluster-Interconnect für den Datendienstverkehr “ auf Seite 75
- „Ressourcen, Ressourcengruppen und Ressourcentypen “ auf Seite 76
- „Datendienst-Projektconfiguration “ auf Seite 79
- „Öffentliche Netzwerkadapter und Internet Protocol (IP) Network Multipathing “ auf Seite 89
- „SPARC: Unterstützung der dynamischen Rekonfiguration “ auf Seite 91

Diese Informationen richten sich in erster Linie an Systemverwalter und Anwendungsentwickler, die mit der Sun Cluster-API und dem SDK arbeiten. Cluster-Systemverwalter können diese Informationen als Vorbereitung für das Installieren, Konfigurieren und Verwalten von Cluster-Software nutzen. Anwendungsentwickler können die Informationen nutzen, um die Cluster-Umgebung zu verstehen, in der sie arbeiten.

Verwaltungsschnittstellen

Sie können auswählen, wie Sie das Sun Cluster-System von unterschiedlichen Benutzeroberflächen aus installieren, konfigurieren und verwalten möchten. Sie können Systemverwaltungsaufgaben entweder über die grafische Benutzeroberfläche (Graphical User Interface, GUI) von SunPlex-Manager oder über die dokumentierte Befehlszeilenschnittstelle ausführen. Im obersten Bereich der Befehlszeilenschnittstelle befinden sich mehrere Dienstprogramme wie `scinstall` und `scsetup`, um bestimmte Installations- und Konfigurationsaufgaben zu vereinfachen. Zum Sun Cluster-System gehört auch ein Modul, das als Teil von Sun Management Center läuft und eine GUI für bestimmte Cluster-Aufgaben zur Verfügung stellt. Dieses Modul ist nur für SPARC-basierte Cluster verfügbar. Unter „Verwaltungstools“ in *Sun Cluster Handbuch Systemverwaltung für Solaris OS* finden Sie vollständige Beschreibungen der Verwaltungsschnittstellen.

Cluster-Zeit

Die Zeit muss zwischen allen Knoten eines Clusters synchronisiert sein. Ob die Zeitsynchronisierung der Cluster-Knoten mit einer externen Zeitquelle erfolgt, hat für den Cluster-Betrieb keine Bedeutung. Das Sun Cluster-System verwendet das NTP-Protokoll (Network Time Protocol) zum Synchronisieren der Uhren zwischen den Knoten.

Im Allgemeinen verursacht eine Änderung der Systemuhr um eine Sekundenfraktion keine Probleme. Wenn Sie jedoch `date (1)`, `rdate (1M)` oder `xntpdate (1M)` (interaktiv oder innerhalb von cron-Skripten) auf einem aktiven Cluster ausführen, können Sie eine weit über einer Sekundenfraktion liegende Zeitänderung erzwingen, um die Systemuhr mit der Zeitquelle zu synchronisieren. Diese erzwungene Änderung kann zu Problemen bei den Zeitmarken für Dateiänderungen oder beim NTP-Dienst führen.

Wenn Sie das Solaris-Betriebssystem auf jedem Cluster-Knoten installieren, können Sie die Standardeinstellung für Zeit und Datum für den Knoten ändern. Im Allgemeinen können Sie die werkseitigen Standardeinstellungen akzeptieren.

Wenn Sie die Sun Cluster-Software mit `scinstall (1M)` installieren, wird in einem Prozessschritt auch das NTP für den Cluster konfiguriert. Sun Cluster-Software stellt eine Vorlagendatei zur Verfügung, `ntp.cluster` (siehe `/etc/inet/ntp.cluster` auf einem installierten Clusterknoten), die eine Peer-Beziehung zwischen allen Cluster-Knoten erstellt. Ein Knoten wird als "bevorzugter" Knoten festgelegt. Die Knoten werden über ihre privaten Hostnamen identifiziert und die

Zeitsynchronisierung erfolgt über den Cluster-Interconnect. Anweisungen zur Konfiguration des Clusters über NTP erhalten Sie in Kapitel 2, „Installieren und Konfigurieren der Sun Cluster-Software“ in *Sun Cluster Handbuch Softwareinstallation für Solaris OS*.

Alternativ können Sie einen oder mehrere NTP-Server außerhalb des Clusters einrichten und die `ntp.conf`-Datei ändern, um diese Konfiguration wiederzugeben.

Im Normalbetrieb sollten Sie die Cluster-Zeit niemals anpassen müssen. Wenn die Zeit bei der Installation des Solaris-Betriebssystems jedoch nicht richtig eingestellt wurde und Sie sie ändern möchten, finden Sie das erforderliche Verfahren in Kapitel 7, „Verwalten des Clusters“ in *Sun Cluster Handbuch Systemverwaltung für Solaris OS*.

Hoch verfügbares Framework

Das Sun Cluster-System macht alle Komponenten auf dem "Pfad" zwischen Benutzern und Daten hoch verfügbar, einschließlich der Netzwerkschnittstellen, der Anwendungen selbst, des Dateisystems und der Multihostgeräte. Im Allgemeinen ist eine Cluster-Komponente hoch verfügbar, wenn sie trotz eines einzelnen Fehlers (Software oder Hardware) im System in Betrieb bleibt.

Die nachstehende Tabelle zeigt die Arten von Sun Cluster -Komponentenfehlern (bei Hardware und Software) und die entsprechende Form der Wiederherstellung, die in das hoch verfügbare Framework integriert ist

TABELLE 3-1 Ebenen der Sun Cluster-Fehlererkennung und Wiederherstellung

Fehlerhafte Cluster-Komponente	Softwarewiederherstellung	Hardwarewiederherstellung
Datendienst	HA-API, HA-Framework	N/V
Öffentlicher Netzwerkadapter	Internet Protocol (IP) Network Multipathing	Mehrfache öffentliche Netzwerkadapterkarten
Cluster-Dateisystem	Primär- und Sekundärknotenreplikate	Multihostgeräte
Gespiegeltes Multihostgerät	Datenträgerverwaltung (Solaris Volume Manager und VERITAS Volume Manager, das nur in SPARC-basierten Clustern verfügbar ist)	RAID-5-Hardware (zum Beispiel Sun StorEdge™ A3x00)
Globales Gerät	Primär- und Sekundärknotenreplikate	Mehrere Pfade zum Gerät, Cluster-Transportverbindungs- punkte

TABELLE 3-1 Ebenen der Sun Cluster-Fehlererkennung und Wiederherstellung
(Fortsetzung)

Fehlerhafte Cluster-Komponente	Softwarewiederherstellung	Hardwarewiederherstellung
Privates Netzwerk	HA-Transportsoftware	Mehrere private, hardwareunabhängige Netzwerke
Knoten	CMM, Failfast-Treiber	Mehrere Knoten

Das Hochverfügbarkeits-Framework der Sun Cluster -Software erkennt einen Knotenfehler schnell und erstellt einen neuen, gleichwertigen Server für die Framework-Ressourcen auf einem anderen Knoten im Cluster. Zu keiner Zeit sind alle Framework-Ressourcen nicht verfügbar. Framework-Ressourcen, die von einem abgestürzten Knoten nicht betroffen sind, sind während der Wiederherstellung voll verfügbar. Darüber hinaus sind die Framework-Ressourcen des ausgefallenen Knotens wieder verfügbar, sobald sie wiederhergestellt sind. Eine wiederhergestellte Framework-Ressource muss nicht warten, bis die Wiederherstellung aller anderen Framework-Ressourcen abgeschlossen ist.

Die meisten hoch verfügbaren Framework-Ressourcen werden für die Anwendungen (Datendienste), die sie verwenden, transparent wiederhergestellt. Die Semantik des Framework-Ressourcenzugriffs bleibt bei einem Knotenfehler vollständig erhalten. Die Anwendungen können einfach nicht erkennen, dass der Framework-Ressourcenserver auf einen anderen Knoten verschoben wurde. Der Ausfall eines einzelnen Knotens ist für die Programme auf verbleibenden Knoten vollständig transparent, indem die zu dem Knoten gehörenden Dateien, Geräte und Plattendatenträger verwendet werden. Diese Transparenz besteht, sofern ein alternativer Hardwarepfad zu den Platten von einem anderen Knoten aus vorhanden ist. Ein Beispiel ist die Verwendung von Multihostgeräten mit Ports für mehrere Knoten.

Cluster-Mitgliedschaftsüberwachung (Cluster Membership Monitor, CMM)

Alle Knoten müssen eine konsistente Einigung hinsichtlich der Cluster-Mitgliedschaft erzielen, um die Daten vor Beschädigung zu schützen. Der CMM koordiniert gegebenenfalls eine Cluster-Rekonfiguration der Cluster-Dienste (Anwendungen) infolge eines Fehlers.

Der CMM erhält Informationen über die Konnektivität mit anderen Knoten aus der Cluster-Transportschicht. Der CMM verwendet während einer Rekonfiguration den Cluster-Interconnect zum Austauschen von Statusinformationen.

Nach dem Erkennen einer Änderung bei der Cluster-Mitgliedschaft führt der CMM eine synchronisierte Konfiguration des Clusters aus. Bei einer synchronisierten Konfiguration können Cluster-Ressourcen auf Grundlage der neuen Mitgliedschaft im Cluster neu verteilt werden.

Im Unterschied zu früheren Sun Cluster Softwareversionen wird der CMM vollständig im Kernel ausgeführt.

Weitere Informationen darüber, wie sich der Cluster vor der Partitionierung in mehrere getrennte Cluster schützt, finden Sie unter „[Informationen zum Fehlerschutz](#)“ auf Seite 57.

Failfast-Mechanismus

Wenn der CMM ein kritisches Knoten-Problem erkennt, fordert er das Cluster-Framework auf, den Knoten zwangsweise herunterzufahren (Panik) und aus der Cluster-Mitgliedschaft zu entfernen. Der Mechanismus für diesen Vorgang wird als *Failfast* bezeichnet. Ein Failfast bewirkt das Herunterfahren eines Knotens auf zwei Arten.

- Wenn ein Knoten den Cluster verlässt und dann versucht, einen neuen Cluster ohne Quorum zu starten, wird er vom Zugriff auf die gemeinsam genutzten Platten "geschützt". Weitere Details zur Verwendung von Failfast finden Sie unter „[Informationen zum Fehlerschutz](#)“ auf Seite 57.
- Wenn einer oder mehr Cluster-spezifische Dämonen ausfallen (`clexecd`, `rpc.pmf`, `rgmd` oder `rpc.ed`) wird der Fehler vom CMM erkannt und der Knoten gerät in Panik.

Wenn der Ausfall eines Cluster-Dämons einen Knoten in Panik versetzt, wird in der Konsole für diesen Knoten eine Meldung angezeigt, die in etwa folgendermaßen aussieht.

```
panic[cpu0]/thread=40e60: Failfast: Aborting because "pmfd" died 35 seconds ago.  
409b8 cl_runtime: __0FZsc_syslog_msg_log_no_argsPviTCPcCTB+48 (70f900, 30, 70df54, 407acc, 0)  
%l0-7: 1006c80 000000a 000000a 10093bc 406d3c80 7110340 0000000 4001 fbfo
```

Nach der Panik kann der Knoten neu booten und versuchen, dem Cluster wieder beizutreten. Alternativ kann der Knoten in Clustern aus SPARC-basierten Systemen an der OpenBoot™ PROM (OBP)-Eingabeaufforderung bleiben. Die nächste Aktion des Knotens wird durch die Einstellung des `auto-boot?`-Parameters festgelegt. Sie können `auto-boot?` mit `eeprom(1M)` an der OpenBoot PROM `ok`-Eingabeaufforderung einstellen.

Cluster-Konfigurations-Repository (Cluster Configuration Repository, CCR)

Das CCR verwendet für Aktualisierungen einen Zwei-Phasen-Commit-Algorithmus: Eine Aktualisierung muss auf allen Cluster-Mitgliedern erfolgreich abgeschlossen werden, sonst wird die Aktualisierung zurückgenommen. Das CCR verwendet den Cluster-Interconnect für die Umsetzung der verteilten Aktualisierungen.



Caution – Das CCR besteht zwar aus Textdateien, aber Sie dürfen die CCR-Dateien unter keinen Umständen manuell bearbeiten. Jede Datei enthält einen Prüfsummeneintrag, um die Konsistenz zwischen den Knoten zu sichern. Eine manuell Aktualisierung der CCR-Dateien kann dazu führen, dass ein Knoten oder der ganze Cluster nicht mehr funktioniert.

Das CCR stellt mithilfe des CMM sicher, dass ein Cluster nur mit einem festgelegten Quorum läuft. Das CCR ist dafür zuständig, die Datenkonsistenz im ganzen Cluster zu überprüfen, eine ggf. erforderliche Wiederherstellung durchzuführen und Aktualisierungen für die Daten bereitzustellen.

Globale Geräte

Das Sun Cluster-System verwendet *Globale Geräte*, um einen cluster-weiten, hoch verfügbaren Zugriff auf alle Geräte innerhalb eines Clusters von jedem Knoten aus zur Verfügung zu stellen. Wenn ein Knoten ausfällt, während er Zugriff auf ein globales Gerät gewährt, erkennt die Sun Cluster-Software normalerweise einen anderen Pfad zu dem Gerät und leitet den Zugriff auf diesen Pfad um. Zu den globalen Geräten bei Sun Cluster gehören Platten, CD-ROMs und Bänder. Die einzigen globalen Multiport-Geräte, die von der Sun Cluster-Software unterstützt werden, sind jedoch Platten. Das bedeutet, dass CD-ROM- und Bandgeräte derzeit keine hoch verfügbaren Geräte sind. Die lokalen Platten auf jedem Server sind ebenfalls keine Multiport-Geräte und deswegen nicht hoch verfügbar.

Der Cluster weist jeder Platte, jedem CD-ROM-Laufwerk und jedem Bandgerät im Cluster automatisch eine einmalige ID zu. Diese Zuweisung ermöglicht einen konsistenten Zugriff auf jedes Gerät von jedem Cluster-Knoten aus. Der Namensraum globaler Geräte ist im Verzeichnis `/dev/global` enthalten. Weitere Informationen finden Sie unter „Globaler Namensraum“ auf Seite 47.

Globale Multiport-Geräte stellen mehrere Pfade zu einem Gerät zur Verfügung. Da Multihostplatten zu einer Plattengerätegruppe gehören, die von mehreren Knoten gehostet wird, werden sie hoch verfügbar gemacht.

Geräte-IDs und DID-Pseudotreiber

Die Sun Cluster-Software verwaltet globale Geräte über einen so genannten Geräte-ID-Pseudotreiber (DID-Pseudotreiber). Dieser Treiber wird verwendet, um automatisch jedem Gerät im Cluster einschließlich Multihostplatten, Bandlaufwerken und CD-ROM-Laufwerken eine einmalige ID zuzuweisen.

Der DID-Pseudotreiber ist fester Bestandteil der Zugriffsfunktion auf globale Geräte des Clusters. Der DID-Treiber testet alle Knoten des Clusters, erstellt eine Liste von einmaligen Plattengeräten und weist jeder Platte eine einmalige Geräteklassen- und Gerätenummer zu, die auf allen Knoten des Clusters konsistent ist. Der Zugriff auf globale Geräte erfolgt mithilfe der einmaligen, vom DID-Treiber zugewiesenen Geräte-ID anstelle der herkömmlichen Solaris-Geräte-IDs wie `c0t0d0` für eine Platte.

Dieser Ansatz stellt sicher, dass jede auf Platten zugreifende Anwendung (wie Datenträger-Manager oder Anwendungen, die im raw-Modus betriebene Geräte verwenden) im ganzen Cluster einen konsistenten Pfad verwenden. Diese Konsistenz ist besonders bei Multihostplatten wichtig, weil die lokale Geräteklassen- und Geräteummern für jedes Gerät von Knoten zu Knoten unterschiedlich sein kann und sich damit auch die Solaris-Konventionen zur Gerätebenennung ändern. Node 1 kann eine Multihostplatte zum Beispiel als `c1t2d0` identifizieren, während Node 2 dieselbe Platte vollkommen anders, nämlich als `c3t2d0`, identifiziert. Der DID-Treiber weist einen globalen Namen zu, wie `d10`, den der Knoten stattdessen verwendet, und gibt jedem Knoten eine konsistente Zuordnung zur Multihostplatte.

Sie aktualisieren und verwalten die Geräte-IDs mithilfe von `scdidadm(1M)` und `scgdevs(1M)`. Weitere Informationen entnehmen Sie bitte den folgenden Man Pages:

- `scdidadm(1M)`
- `scgdevs(1M)`

Plattengerätegruppen

Im Sun Cluster-System müssen alle Multihostgeräte von der Sun Cluster-Software gesteuert werden. Sie erstellen zunächst die Plattengruppen des Datenträger-Managers – entweder Solaris Volume Manager-Plattensätze oder VERITAS Volume Manager-Plattengruppen (nur in SPARC-basierten Clustern verfügbar) – auf den Multihost-Platten. Dann registrieren Sie die Datenträger-Manager-Plattengruppen als *Plattengerätegruppen*. Eine Plattengerätegruppe ist ein globaler Gerätetyp. Zudem erstellt die Sun Cluster-Software automatisch eine im raw-Modus betriebene Plattengerätegruppe für jedes Platten- und Bandgerät im Cluster. Diese Cluster-Gerätegruppen bleiben jedoch im Offline-Zustand, bis Sie darauf als globale Geräte zugreifen.

Die Registrierung liefert dem Sun Cluster-System Informationen zu den Pfaden zwischen Knoten und Datenträger-Manager-Plattengruppen. An diesem Punkt werden Datenträger-Manager-Plattengruppen innerhalb des Clusters global zugänglich. Wenn mehrere Knoten in eine Plattengerätegruppe schreiben können (die Gruppe unterstützen können), werden die in dieser Plattengerätegruppe gespeicherten Daten hoch verfügbar. Die hoch verfügbare Plattengerätegruppe kann zum Hosten von Cluster-Dateisystemen verwendet werden.

Hinweis – Plattengerätegruppen sind von Ressourcengruppen unabhängig. Ein Knoten kann eine Ressourcengruppe (die eine Gruppe von Datendienstprozessen darstellt) unterstützen, während ein anderer die Plattengruppen unterstützen kann, auf welche die Datendienste zugreifen. Das optimale Vorgehen besteht aber darin, die Plattengerätegruppe zur Speicherung bestimmter Anwendungsdaten und die Ressourcengruppe mit den Anwendungsressourcen (Anwendungsdaemon) auf demselben Knoten zu speichern. Weitere Informationen über die Zuordnung zwischen Ressourcengruppen und Plattengerätegruppen finden Sie unter „Relationship Between Resource Groups and Disk Device Groups“ in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

Wenn ein Knoten eine Plattengerätegruppe verwendet, wird die Datenträger-Manager-Plattengruppe "global", weil sie Multipathing für die zugehörigen Platten unterstützt. Jeder mit den Multihostplatten real verbundene Cluster-Knoten stellt einen Pfad zur Plattengerätegruppe zur Verfügung.

Plattengerätegruppen-Failover

Auf alle Plattengerätegruppen in einem Gehäuse kann über einen alternativen Pfad zugegriffen werden, wenn der derzeitige Knoten-Master der Gerätegruppe ausfällt, weil ein Plattengehäuse an mehrere Knoten angeschlossen ist. Der Ausfall des Knotens, der als Gerätegruppen-Master fungiert, beeinträchtigt den Zugriff auf die Gerätegruppe nur während der Zeitspanne, die zur Ausführung der Wiederherstellungs- und Konsistenzprüfung erforderlich ist. Während dieser Zeit werden alle Anforderungen (für die Anwendung transparent) gesperrt, bis das System die Gerätegruppe verfügbar macht.

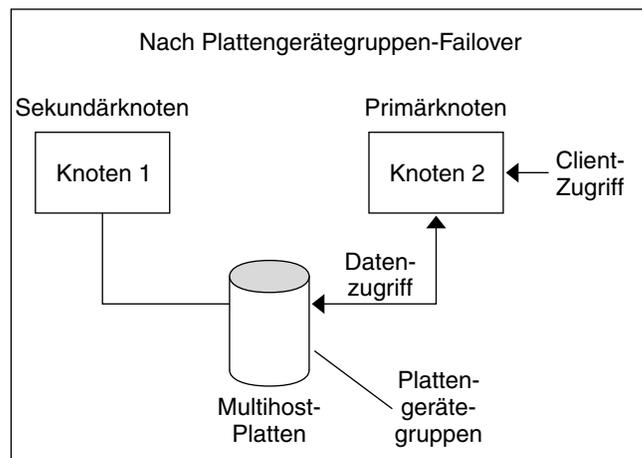
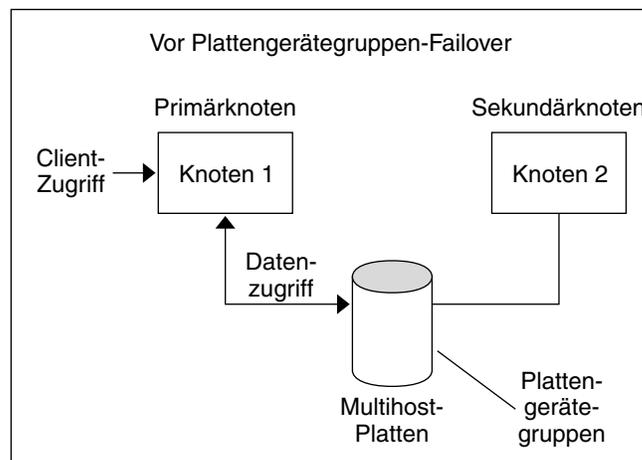


ABBILDUNG 3-1 Plattengerätegruppe vor und nach Failover

Multiport-Plattengerätegruppen

In diesem Abschnitt werden Plattengruppeneigenschaften beschrieben, mit denen Sie Leistung und Verfügbarkeit in einer Multiport-Plattenkonfiguration ausgleichen können. Die Sun Cluster-Software stellt Ihnen zwei Eigenschaften zur Verfügung, mit denen Sie eine Multiport-Plattenkonfiguration konfigurieren können: `preferred` und `numsecondaries`. Mit der `preferred`-Eigenschaft können Sie die Reihenfolge steuern, in der die Knoten versuchen, die Steuerung im Fall eines Failover zu übernehmen. Mit der `numsecondaries`-Eigenschaft legen Sie eine gewünschte Anzahl von Sekundärknoten für eine Gerätegruppe fest.

Ein hoch verfügbarer Dienst gilt als ausgefallen, wenn der Primärknoten ausfällt und kein Sekundärknoten die Rolle des Primärknotens übernehmen kann. Wenn ein Dienst-Failover erfolgt und die `preferenced`-Eigenschaft auf `true` eingestellt ist, folgen die Knoten bei der Auswahl eines Sekundärknotens der Reihenfolge in der Knotenliste. Die mit der `preferenced`-Eigenschaft eingerichtete Knotenliste definiert die Reihenfolge, in der die Knoten versuchen, die Steuerung als Primärknoten oder den Übergang von Spare-Knoten zu Sekundärknoten zu übernehmen. Sie können den Vorrang eines Gerätedienstes mit dem Dienstprogramm `scsetup(1M)` dynamisch ändern. Der Vorrang der abhängigen Dienste, zum Beispiel eines globalen Dateisystems, stimmt exakt mit dem Vorrang des Gerätedienstes überein.

Im Normalbetrieb führt der Primärknoten Checkpoint-Vorgänge für Sekundärknoten durch. In einer Multiport-Plattenkonfiguration führt das Durchführen des Checkpoint-Vorgangs an jedem Sekundärknoten zu einer Leistungsverminderung des Clusters und zu Speicherüberlastung. Die Spare-Knoten-Unterstützung wurde implementiert, um die Checkpoint-bedingte Leistungsminderung und Speicherüberlastung zu minimieren. Standardmäßig weist Ihre Plattengerätegruppe einen Primär- und einen Sekundärknoten auf. Die restlichen verfügbaren Anbieterknoten werden Spare-Knoten. Bei einem Failover wird der Sekundärknoten zum Primärknoten, und der Knoten mit der höchsten Priorität in der Knotenliste wird zum Sekundärknoten.

Die gewünschte Anzahl von Sekundärknoten kann auf jede beliebige Ganzzahl zwischen 1 und der Anzahl der betriebsbereiten Anbieterknoten in der Gerätegruppe, die keine Primärknoten sind, gesetzt werden.

Hinweis – Wenn Sie Solaris Volume Manager verwenden, müssen Sie zuerst die Plattengerätegruppe erstellen, bevor Sie die Eigenschaft `numsecondaries` auf einen anderen Wert als den Standardwert einstellen können.

Die standardmäßig eingestellte gewünschte Anzahl von Sekundärknoten für Gerätedienste ist Eins. Die tatsächliche Anzahl der vom Replikat-Framework verwalteten Sekundäranbieter entspricht der gewünschten Anzahl, es sei denn, die Anzahl der betriebsbereiten Anbieter, die keine Primärknoten sind, ist kleiner als die gewünschte Anzahl. Sie müssen die `numsecondaries`-Eigenschaft ändern und die Knotenliste erneut überprüfen, wenn Sie Knoten zur Konfiguration hinzufügen oder daraus entfernen. Die Verwaltung der Knotenliste und der gewünschten Anzahl von Sekundärknoten verhindert Konflikte zwischen der konfigurierten und der tatsächlichen, vom Framework erlaubten Anzahl von Sekundärknoten.

- (Solaris Volume Manager) Sie verwalten das Hinzufügen oder Entfernen von Knoten aus Ihrer Konfiguration mit dem `metaset(1M)`-Befehl für Solaris Volume Manager-Gerätegruppen in Verbindung mit den Einstellungen für die Eigenschaften `preferenced` und `numsecondaries`.
- (Veritas Volume Manager) Sie verwalten das Hinzufügen oder Entfernen von Knoten aus Ihrer Konfiguration mit dem `scconf(1M)`-Befehl für VxVM-Plattengerätegruppen in Verbindung mit den Einstellungen für die

Eigenschaften `preferenced` und `numsecondaries`.

- Verfahrenstechnische Informationen über das Ändern der Plattengerätegruppen-Eigenschaften finden Sie unter „Überblick über das Verwalten von Cluster-Dateisystemen“ in *Sun Cluster Handbuch Systemverwaltung für Solaris OS*.

Globaler Namensraum

Der Sun Cluster-Softwaremechanismus zur Aktivierung globaler Geräte ist der *globale Namensraum*. Der globale Namensraum beinhaltet die `/dev/global/`-Hierarchie sowie die Datenträger-Manager-Namensräume. Der globale Namensraum gibt sowohl Multihostplatten als auch lokale Platten wieder (und jedes andere Cluster-Gerät wie CD-Rom-Laufwerke und Bänder) und stellt mehrere Failover-Pfade für die Multihostplatten zur Verfügung. Jeder real an die Multihostplatten angeschlossene Knoten stellt für alle Knoten im Cluster einen Pfad zum Speicher zur Verfügung.

In der Regel befinden sich die Namensräume des Datenträger-Managers für Solaris Volume Manager in den Verzeichnissen `/dev/md/Plattensatz/dsk` (und `rdsk`). Die Namensräume des Datenträger-Managers für Veritas VxVM befinden sich in den Verzeichnissen `/dev/vx/dsk/disk-group` und `/dev/vx/rdsk/disk-group`. Diese Namensräume bestehen aus Verzeichnissen für jeden Solaris Volume Manager-Plattensatz bzw. jede VxVM-Plattengruppe auf dem ganzen Cluster. Jedes dieser Verzeichnisse enthält einen Geräteknoten für jedes Metagerät oder jeden Datenträger in diesem Plattensatz bzw. in dieser Plattengruppe.

Im Sun Cluster -System wird jeder Geräteknoten im lokalen Datenträger-Manager-Namensraum durch eine symbolische Verknüpfung mit einem Geräteknoten im Dateisystem `/global/.devices/node@nodeID` ersetzt, wobei `nodeID` als Ganzzahl für den Knoten im Cluster steht. Die Sun Cluster-Software zeigt die Datenträger-Manager-Geräte auch weiterhin als symbolische Verknüpfungen an ihren Standard-Speicherorten an. Sowohl der globale Namensraum als auch der Datenträger-Manager-Namensraum sind auf jedem Cluster-Knoten verfügbar.

Der globale Namensraum bietet unter anderem folgende Vorteile:

- Jeder Knoten bleibt ziemlich unabhängig, und es gibt nur geringfügige Änderungen im Geräteverwaltungsmodell.
- Geräte können selektiv global gemacht werden.
- Verknüpfungsgeneratoren von Drittherstellern arbeiten weiterhin.
- Nach Vergabe eines lokalen Gerätenamens erfolgt eine einfache Zuordnung, um den entsprechenden globalen Namen zu erhalten.

Beispiel für lokale und globale Namensräume

Die nachstehende Tabelle zeigt die Zuordnungen zwischen lokalen und globalen Namensräumen für eine Multihostplatte, `c0t0d0s0`.

TABELLE 3–2 Zuordnungen von lokalen und globalen Namensräumen

Komponente bzw. Pfad	Lokaler Knoten-Namensraum	Globaler Namensraum
Logischer Solaris-Name	<code>/dev/dsk/c0t0d0s0</code>	<code>/global/.devices/node@nodeID/dev/dsk/c0t0d0s0</code>
DID-Name	<code>/dev/did/dsk/d0s0</code>	<code>/global/.devices/node@nodeID/dev/did/dsk/d0s0</code>
Solaris Volume Manager	<code>/dev/md/diskset/dsk/d0</code>	<code>/global/.devices/node@nodeID/dev/md/diskset/dsk/d0</code>
SPARC: VERITAS Volume Manager	<code>/dev/vx/dsk/disk-group/v0</code>	<code>/global/.devices/node@nodeID/dev/vx/dsk/disk-group/v0</code>

Der globale Namensraum wird automatisch bei der Installation erzeugt und bei jedem Neubooten der Rekonfiguration aktualisiert. Sie können den globalen Namensraum auch durch Ausführen des `scgdevs (1M)`-Befehls generieren.

Cluster-Dateisysteme

Das Cluster-Dateisystem hat folgende Merkmale:

- Die Speicherorte für den Dateizugriff sind transparent. Ein Prozess kann eine Datei an einem beliebigen Speicherort im System öffnen. Die Prozesse können die Datei auf allen Knoten mithilfe desselben Pfadnamens suchen.

Hinweis – Wenn das Cluster-Dateisystem Dateien liest, wird die Zugriffszeit für diese Dateien nicht aktualisiert.

- Mit Kohärenzprotokollen wird die UNIX-Dateizugriffsemantik auch dann bewahrt, wenn mehrere Knoten gleichzeitig auf die Datei zugreifen.
- Umfassendes Caching wird zusammen mit Bulk-E/A-Bewegungen ohne Zwischenspeicherung (Zero-Copy) eingesetzt, um Dateidaten effizient zu verschieben.
- Das Cluster-Dateisystem stellt mit den `fcntl (2)`-Schnittstellen eine hoch verfügbare kooperative Dateisperrfunktion zur Verfügung. Anwendungen, die auf mehreren Cluster-Knoten ausgeführt werden, können den Datenzugriff

synchronisieren, indem sie die kooperative Dateisperrung auf ein Cluster-Dateisystem anwenden. Dateisperrungen werden sofort von allen den Cluster verlassenden Knoten und von allen gesperrten und fehlgeschlagenen Anwendungen aufgehoben.

- Kontinuierlicher Datenzugriff ist auch bei Ausfällen gesichert. Anwendungen sind von den Ausfällen nicht betroffen, solange noch ein Pfad zu den Platten funktionsfähig ist. Diese Garantie gilt für den Plattenzugriff im raw-Modus und für alle Dateisystemvorgänge.
- Cluster-Dateisysteme sind vom zugrunde liegenden Dateisystem und von der Datenträgerverwaltung unabhängig. Cluster-Dateisysteme machen jedes unterstützte, auf den Platten vorhandene Dateisystem global.

Sie können ein Dateisystem mit `mount -g` global oder mit `mount` lokal auf einem globalen Gerät einhängen.

Programme können auf die Dateien in einem Cluster-Dateisystem von jedem Knoten im Cluster aus und mit demselben Dateinamen zugreifen (z. B. `/global/foo`).

Ein Cluster-Dateisystem wird auf allen Cluster-Mitgliedern eingehängt. Sie können ein Cluster-Dateisystem nicht auf einer Untermenge von Cluster-Mitgliedern einhängen.

Ein Cluster-Dateisystem ist kein anderer Typ von Dateisystem. Die Clients überprüfen das zugrunde liegende Dateisystem (z. B. UFS).

Verwenden von Cluster-Dateisystemen

Im Sun Cluster-System sind alle Multihostplatten in Plattengerätegruppen eingebunden. Hierbei kann es sich um Solaris Volume Manager-Plattensätze, VxVM-Plattengruppen oder Platten handeln, die nicht von einem softwarebasierten Datenträger-Manager gesteuert werden.

Damit ein Cluster-Dateisystem hoch verfügbar ist, muss der zugrunde liegende Plattenspeicher mit mehreren Knoten verbunden sein. Deswegen ist ein lokales Dateisystem (ein auf der lokalen Platte eines Knotens gespeichertes Dateisystem), das zu einem Cluster-Dateisystem gemacht wird, nicht hoch verfügbar.

Sie können Cluster-Dateisysteme genauso wie andere Dateisysteme einhängen:

- **Manuell** – Verwenden Sie den Befehl `mount` mit den Einhängoptionen `-g` oder `-o global`, um das Cluster-Dateisystem von der Befehlszeile aus einzuhängen. Beispiel:

```
SPARC: # mount -g /dev/global/dsk/d0s0 /global/oracle/data
```

- **Automatisch** – Erstellen Sie einen Eintrag in der Datei `/etc/vfstab` mit einer `global`-Einhängeoption, um das Cluster-Dateisystem beim Booten einzuhängen. Dann können Sie im `/global`-Verzeichnis auf allen Knoten einen Einhängpunkt erstellen. Das `/global`-Verzeichnis ist ein empfohlener Speicherplatz, keine

Anforderung. Hier eine Beispielzeile für ein Cluster-Dateisystem aus einer `/etc/vfstab`-Datei:

```
SPARC: /dev/md/oracle/dsk/d1 /dev/md/oracle/rdisk/d1 /global/oracle/data ufs 2 yes global,logging
```

Hinweis – Solange die Sun Cluster-Software kein Benennungsverfahren für Cluster-Dateisysteme vorschreibt, können Sie die Verwaltung durch das Erstellen eines Einhängpunktes für alle Cluster-Dateisysteme unter demselben Verzeichnis vereinfachen, wie `/global/disk-device-group`. Weitere Informationen erhalten Sie in der *Sun Cluster 3.1 9/04 Software Collection for Solaris OS (SPARC Platform Edition)* und im *Sun Cluster Handbuch Systemverwaltung für Solaris OS*.

HASStoragePlus Ressourcentyp

Der HASStoragePlus-Ressourcentyp ist darauf ausgelegt, nicht globale Dateisystemkonfigurationen wie UFS und VxFS hoch verfügbar zu machen. Mit HASStoragePlus integrieren Sie das lokale Dateisystem in die Sun Cluster-Umgebung und machen das Dateisystem hoch verfügbar. HASStoragePlus stellt zusätzliche Dateisystemfunktionen wie Prüfen, Einhängen und erzwungenes Aushängen zur Verfügung, mit denen Sun Cluster bei lokalen Dateisystemen ein Failover durchführen kann. Zum Ausführen eines Failover-Vorgangs muss sich das lokale Dateisystem auf globalen Plattengruppen mit aktivierten Affinitäts-Switchover befinden.

Unter „Enabling Highly Available Local File Systems“ in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* finden Sie Informationen zum Konfigurieren des HASStoragePlus-Ressourcentyps.

HASStoragePlus kann auch zum Synchronisieren beim Start von Ressourcen und Plattengerätegruppen, von denen die Ressourcen abhängen, eingesetzt werden. Weitere Informationen finden Sie unter „[Ressourcen, Ressourcengruppen und Ressourcentypen](#)“ auf Seite 76.

Syncdir-Einhängeoption

Die syncdir-Einhängeoption kann für auf UFS basierende Cluster-Dateisysteme verwendet werden. Sie erzielen jedoch eine wesentlich bessere Leistung, wenn Sie syncdir nicht angeben. Bei Angabe von syncdir sind die Einträge garantiert POSIX-konform. Wenn syncdir nicht angegeben wird, verhält sich das System wie ein NFS-Dateisystem. So kann es beispielsweise ohne syncdir vorkommen, dass Sie fehlende Speicherkapazität zum Beispiel erst beim Schließen der Datei feststellen. Mit syncdir (und POSIX-Verhalten) wäre das Fehlen von Speicherplatz während des Schreibvorgangs erkannt worden. Es treten allerdings nur selten Probleme auf, wenn Sie syncdir nicht angeben.

Wenn Sie einen SPARC-basierten Cluster verwenden, hat VxFS keine Einhängoption, die der Einhängoption syncdir für UFS entspricht. Das VxFS-Verhalten entspricht dem UFS-Verhalten, wenn die syncdir-Einhängeoption nicht angegeben wurde.

Häufig gestellte Fragen zu globalen Geräten und Cluster-Dateisystemen finden Sie unter „FAQs zu Dateisystemen“ auf Seite 98.

Plattenfadüberwachung

Die derzeitige Version der Sun Cluster-Software unterstützt die Plattenfadüberwachung (DPM). Dieser Abschnitt liefert konzeptionelle Informationen zu DPM, zum DPM-Dämon und zu den Verwaltungstools für die Plattenfadüberwachung. Verfahrenstechnische Informationen zur Überwachung, zum Aufheben der Überwachung und zum Prüfen des Plattenfadstatus finden Sie im *Sun Cluster Handbuch Systemverwaltung für Solaris OS*.

Hinweis – DPM wird auf Knoten mit Versionen vor Sun Cluster 3.1 10/03 nicht unterstützt. Verwenden Sie während einer Aufrüstung keine DPM-Befehle. Nach der Aufrüstung müssen alle Knoten online sein, um die DPM-Befehle verwenden zu können.

DPM - Überblick

DPM verbessert die gesamte Zuverlässigkeit der Failover- und Switchover-Vorgänge durch die Überwachung der Verfügbarkeit des Plattenpfades für den Sekundärknoten. Mit dem `scdpm`-Befehl überprüfen Sie die Verfügbarkeit des von einer Ressource verwendeten Plattenpfades, bevor die Ressource umgeschaltet wird. Mit dem vom `scdpm`-Befehl zur Verfügung gestellten Optionen können Sie die Plattenpfade zu einem einzelnen oder zu allen Knoten im Cluster überwachen. Weitere Informationen zu Befehlszeilenoptionen finden Sie in der Online-Dokumentation `scdpm(1M)`.

Die DPM-Komponenten werden aus dem `SUNWscu`-Paket installiert. Das `SUNWscu`-Paket wird im Rahmen der Sun Cluster-Standardinstallation installiert. Details zur Installationsoberfläche finden Sie in der Online-Dokumentation `scinstall(1M)`. In der nachstehenden Tabelle wird der Standardspeicherort für die Installation der DPM-Komponenten beschrieben.

Standort	Komponente
Daemon	<code>/usr/cluster/lib/sc/scdpm</code>
Befehlszeilenschnittstelle	<code>/usr/cluster/bin/scdpm</code>

Standort	Komponente
Gemeinsam genutzte Bibliotheken	/user/cluster/lib/libscdpm.so
Dämon-Statusdatei (zur Laufzeit erstellt)	/var/run/cluster/scdpm.status

Ein Multithread-DPM-Dämon wird auf jedem Knoten ausgeführt. Der DPM-Dämon (`scdpm`) wird durch ein `rc.d`-Skript gestartet, wenn ein Knoten bootet. Wenn ein Problem auftritt, wird der Dämon von `pmfd` verwaltet und startet automatisch neu. In der folgenden Liste wird die Arbeitsweise von `scdpm` beim ersten Start beschrieben.

Hinweis – Beim Start wird der Status für jeden Plattenpfad auf `UNKNOWN` initialisiert.

1. Der DPM-Dämon sammelt Plattenpfad- und Knotennameninformationen aus der vorherigen Statusdatei oder aus der CCR-Datenbank. Weitere Informationen zum CCR finden Sie unter „[Cluster-Konfigurations-Repository \(Cluster Configuration Repository, CCR\)](#)“ auf Seite 41. Nach dem Start eines DPM-Dämons können Sie den Dämon zwingen, die Liste der überwachten Platten aus einer angegebenen Datei zu lesen.
2. Der DPM-Dämon initialisiert die Kommunikationsschnittstelle, um Anforderungen von Komponenten zu beantworten, die außerhalb des Dämons liegen, wie die Befehlszeilenschnittstelle.
3. Der DPM-Dämon pingt jeden Plattenpfad der überwachten Liste alle 10 Minuten mit dem Befehl `scsi_inquiry` an. Die Einträge werden gesperrt, um die Kommunikationsschnittstelle daran zu hindern, auf einen Eintrag zuzugreifen, der gerade geändert wird.
4. Der DPM-Dämon benachrichtigt das Sun Cluster-Ereignis-Framework und protokolliert den neuen Pfadstatus über den UNIX `syslogd(1M)`-Mechanismus.

Hinweis – Alle mit dem Dämon zusammenhängenden Fehler werden von `pmfd (1M)` aufgezeichnet. Alle API-Funktionen geben 0 für erfolgreich und -1 für einen Fehler zurück.

Der DPM-Dämon überwacht die Verfügbarkeit des logischen Pfades, der über Multipath-Treiber wie Sun StorEdge Traffic Manager, HDLM und PowerPath sichtbar ist. Die jeweiligen realen Pfade, die von diesen Treibern verwaltet werden, werden nicht überwacht, weil der Multipath-Treiber einzelne Ausfälle vor dem DPM-Dämon verbirgt.

Überwachen von Plattenpfaden

In diesem Abschnitt werden zwei Methoden zur Überwachung von Plattenpfaden in Ihrem Cluster beschrieben. Die erste Methode wird mit dem `scdpm`-Befehl zur Verfügung gestellt. Mit diesem Befehl können Sie die Plattenpfade im Cluster überwachen, die Überwachung aufheben oder deren Status anzeigen. Dieser Befehl dient auch dem Drucken der Liste von ausgefallenen Platten und der Überwachung der Plattenpfade aus einer Datei.

Die zweite Methode zur Überwachung von Plattenpfaden im Cluster wird von der grafischen Benutzeroberfläche (Graphical User Interface, GUI) von SunPlex-Manager zur Verfügung gestellt. SunPlex-Manager gibt Ihnen einen topologischen Überblick über die überwachten Plattenpfade im Cluster. Die Ansicht wird alle 10 Minuten aktualisiert, um Informationen zur Anzahl der fehlgeschlagenen Pings zu liefern. Zur Verwaltung der Plattenpfade verwenden Sie die Informationen der GUI von SunPlex-Manager zusammen mit dem `scdpm(1M)`-Befehl. Informationen über SunPlex-Manager finden Sie in Kapitel 10, „Verwaltung von Sun Cluster mithilfe der grafischen Benutzeroberflächen“ in *Sun Cluster Handbuch Systemverwaltung für Solaris OS*.

Verwenden des `scdpm`-Befehls zur Plattenpfadüberwachung

Der `scdpm(1M)`-Befehl stellt DMP-Verwaltungsbefehle zur Verfügung, mit denen Sie folgende Aufgaben ausführen können:

- Überwachen eines neuen Plattenpfades,
- Aufheben der Überwachung eines Plattenpfades,
- Neues Lesen der Konfigurationsdaten aus der CCR-Datenbank,
- Lesen der Platten, die überwacht werden sollen oder deren Überwachung aufgehoben werden soll, von einer angegebenen Datei aus,
- Berichten des Status eines Plattenpfades oder aller Plattenpfade im Cluster,
- Drucken aller Plattenpfade, auf die von einem Knoten aus zugegriffen werden kann.

Geben Sie den `scdpm(1M)`-Befehl mit dem Plattenpfad-Argument von einem beliebigen aktiven Knoten aus, um DPM-Verwaltungsaufgaben auf dem Cluster auszuführen. Das Plattenpfad-Argument besteht immer aus einem Knotennamen und einem Plattennamen. Der Knotenname ist nicht erforderlich und wird standardmäßig auf `all` gesetzt, wenn kein Knotenname angegeben wurde. In der nachstehenden Tabelle werden die Benennungskonventionen für den Plattenpfad beschrieben.

Hinweis – Die Verwendung des globalen Plattenpfadnamens wird dringend empfohlen, weil der globale Plattenpfadname auf dem ganzen Cluster konsistent ist. Der UNIX-Plattenpfadname ist nicht auf dem ganzen Cluster konsistent. Der UNIX-Plattenpfad für eine Platte kann von einem Cluster-Knoten zum anderen unterschiedlich sein. Der Plattenpfad kann auf einem Knoten `c1t0d0` und auf einem anderen Knoten `c2t0d0` sein. Verwenden Sie bei UNIX-Plattenpfadnamen den Befehl `scdidadm -L`, um die UNIX-Plattenpfadnamen den globalen Plattenpfadnamen zuzuordnen, bevor Sie DPM-Befehle ausgeben. Weitere Informationen finden Sie in der Online-Dokumentation `scdidadm(1M)`.

TABELLE 3-3 Beispiele für Plattenpfadnamen

Namenstyp	Beispiel Plattenpfadname	Beschreibung
Globaler Plattenpfad	<code>schost-1:/dev/did/dsk/d1</code>	Plattenpfad <code>d1</code> auf dem Knoten <code>schost-1</code>
<code>all:d1</code>	Plattenpfad <code>d1</code> auf allen Cluster-Knoten	
UNIX-Plattenpfad	<code>schost-1:/dev/rdisk/c0t0d0s0</code>	Plattenpfad <code>c0t0d0s0</code> auf dem Knoten <code>schost-1</code>
<code>schost-1:all</code>	Alle Plattenpfade auf dem Knoten <code>schost-1</code>	
Alle Plattenpfade	<code>all:all</code>	Alle Plattenpfade auf allen Cluster-Knoten

Verwenden von SunPlex-Manager zur Plattenpfadüberwachung

Mit SunPlex-Manager können Sie folgende grundlegende DPM-Verwaltungsaufgaben durchführen:

- Überwachen eines Plattenpfades,
- Aufheben der Überwachung eines Plattenpfades,
- Anzeigen des Status aller Plattenpfade im Cluster

Verfahrenstechnische Informationen zur Plattenpfadverwaltung mit SunPlex-Manager finden Sie in der Online-Hilfe zu SunPlex-Manager.

Quorum und Quorum-Geräte

Dieses Kapitel umfasst die folgenden Themen:

- „Informationen zur Quorum-Stimmenanzahl “ auf Seite 56
- „Informationen zum Fehlerschutz “ auf Seite 57
- „Informationen zu Quorum-Konfigurationen “ auf Seite 59
- „Erfüllen der Anforderungen für Quorum-Geräte “ auf Seite 59
- „Einhalten von Empfehlungen für Quorum-Geräte “ auf Seite 60
- „Empfohlene Quorum-Konfigurationen “ auf Seite 61
- „Untypische Quorum-Konfigurationen “ auf Seite 63
- „Unzulässige Quorum-Konfigurationen “ auf Seite 63

Hinweis – Eine Liste der Geräte, die von der Sun Cluster-Software als Quorum-Geräte unterstützt werden, erhalten Sie bei Ihrem Sun-Dienstanbieter.

Da Cluster-Knoten Daten und Ressourcen gemeinsam nutzen, darf ein Cluster nie in getrennte Partitionen aufgeteilt werden, die zur gleichen Zeit aktiv sind. Mehrere aktive Partitionen können zur Beschädigung von Daten führen. Durch den Cluster-Mitglied-Monitor (CMM) und den Quorum-Algorithmus wird gewährleistet, dass immer höchstens eine Instanz desselben Clusters in Betrieb ist, auch wenn der Cluster-Interconnect partitioniert ist.

Eine Einführung in Quorum und CMM finden Sie unter „Cluster-Mitgliedschaft“ in *Sun Cluster Überblick für das Betriebssystem Solaris*.

Bei Cluster-Partitionen ergeben sich zwei Arten von Problemen:

- Split Brain
- Amnesie

Split Brain tritt auf, wenn der Cluster-Interconnect zwischen Knoten verloren geht und der Cluster partitioniert wird. Jede Partition "glaubt", dass sie die einzige Partition ist, da der Knoten in der einen Partition nicht mit dem Knoten in der anderen Partition kommunizieren kann.

Zur Amnesie kommt es, wenn der Cluster nach dem Herunterfahren mit Cluster-Konfigurationsdaten neu startet, die älter als die Daten zum Zeitpunkt des Herunterfahrens sind. Dieses Problem kann auftreten, wenn Sie den Cluster an einem Knoten starten, der nicht Teil der letzten funktionierenden Cluster-Partition war.

Die Sun Cluster-Software vermeidet das Ausreten von Split Brain und Amnesie durch folgende Vorgänge:

- Zuweisung einer Stimme für jeden einzelnen Knoten
- Festlegung einer Stimmenmehrheit für einen betriebsbereiten Cluster

Eine Partition mit der Mehrheit der Stimmen erhält ein *Quorum* und erhält die Erlaubnis zum Arbeiten. Dieser auf einer Stimmenmehrheit basierende Mechanismus vermeidet Split Brain und Amnesie für den Fall, dass mehr als zwei Knoten in einem Cluster konfiguriert wurden. Das Zählen der Knotenstimmen allein reicht jedoch nicht aus, wenn mehr als zwei Knoten in einem Cluster konfiguriert wurden. In einem Zwei-Knoten-Cluster ist zwei eine Mehrheit. Wenn ein solcher Zwei-Knoten-Cluster partitioniert wird, benötigen beide Partitionen jeweils eine externe Stimme, um ein Quorum zu erhalten. Diese externe Stimme wird von einem *Quorum-Gerät* beigesteuert.

Informationen zur Quorum-Stimmenanzahl

Verwenden Sie den Befehl `scstat -q`, um folgende Informationen festzulegen:

- Gesamte konfigurierte Stimmen
- Aktuell vorhandene Stimmen
- Für Quorum erforderliche Stimmen

Weitere Informationen zu diesem Befehl finden Sie unter `scstat(1M)`.

Beide Knoten und Quorum-Geräte steuern Stimmen für den Cluster bei, um ein Quorum zu bilden.

Ein Knoten steuert je nach Status eine unterschiedliche Stimmenanzahl bei:

- Ein Knoten hat eine Stimmenanzahl von *ins*, wenn er gestartet wird und ein Cluster-Mitglied wird.
- Ein Knoten verfügt über *null* Stimmen, wenn er installiert wird.
- Wenn ein Systemadministrator einen Knoten in den Wartungszustand versetzt, hat der Knoten eine Stimmenanzahl von *null*.

Die von Quorum-Geräten beigesteuerten Stimmen basieren auf der Anzahl der Stimmen, die mit dem Gerät verbunden sind. Wenn Sie ein Quorum-Gerät konfigurieren, weist die Sun Cluster-Software dem Quorum-Gerät eine Stimmenanzahl von $N-1$ zu. Hierbei ist N die Anzahl der mit dem Quorum-Gerät verbundenen Stimmen. Ein Quorum-Gerät, das zum Beispiel mit zwei Knoten mit einer Stimmenanzahl von nicht Null verbunden ist, hat einen Quorum-Zählwert von Eins (Zwei minus Eins).

Ein Quorum-Gerät steuert Stimmen bei, wenn *eine* der beiden folgenden Bedingungen erfüllt ist:

- Bei mindestens einem der Knoten, an die das Quorum-Gerät derzeit angeschlossen ist, handelt es sich um ein Cluster-Mitglied.
- Mindestens einer der Knoten, an die das Quorum-Gerät derzeit angeschlossen ist, wird gerade gestartet und der entsprechende Knoten war ein Mitglied der Cluster-Partition, die zuletzt Eigentümer des Quorum-Geräts war.

Quorum-Geräte werden während der Cluster-Installation oder zu einem späteren Zeitpunkt mithilfe der in Kapitel 5, „Verwalten des Quorums“ in *Sun Cluster Handbuch Systemverwaltung für Solaris OS* beschriebenen Verfahren konfiguriert.

Informationen zum Fehlerschutz

Ein wichtiges Thema bei Clustern ist ein Fehler, der zur Partitionierung des Clusters führt (als *Split Brain* bezeichnet). In diesem Fall können nicht mehr alle Knoten miteinander kommunizieren, sodass einzelne Knoten oder Knoten-Teilsätze ggf. versuchen, Einzel- oder Untermengen-Cluster zu bilden. Jede Untermenge oder Partition kann davon "überzeugt" sein, alleinigen Zugriff auf die Multihostgeräte und die Eigentümerschaft zu haben. Wenn mehrere Knoten versuchen, auf die Platten zu schreiben, können Datenfehler auftreten.

Der Fehlerschutz schränkt den Knotenzugriff auf die Multihostgeräte ein, indem der Zugriff auf die Platten real verhindert wird. Wenn ein Knoten den Cluster verlässt (aufgrund eines Ausfalls oder Partitionierung), wird mit dem Fehlerschutz sichergestellt, dass der Knoten keinen Zugriff mehr auf die Platte hat. Nur aktuelle Mitgliederknoten haben Zugriff auf die Platten. Das sichert die Datenintegrität.

Plattengerätedienste stellen Failover-Funktionen für Dienste zur Verfügung, die Multihostgeräte verwenden. Wenn ein Cluster-Mitglied, das zurzeit als Primärknoten (Eigentümer) der Plattengerätegruppe dient, ausfällt oder nicht mehr erreichbar ist, wird ein neuer Primärknoten ausgewählt. Der neue Primärknoten ermöglicht den weiteren Zugriff auf die Plattengerätegruppe nach einer geringfügigen Unterbrechung. Während dieses Vorgangs muss der alte Primärknoten den Zugriff auf die Geräte freigeben, bevor der neue Primärknoten gestartet werden kann. Wenn ein Mitglied jedoch aus dem Cluster ausscheidet und nicht mehr erreichbar ist, kann der Cluster diesen Knoten nicht zur Freigabe der Geräte auffordern, für die er als Primärknoten fungierte. Sie brauchen also ein Mittel, mit dessen Hilfe die funktionsfähigen Mitglieder die Steuerung und den Zugriff auf die globalen Geräte der ausgefallenen Mitglieder übernehmen können.

Das Sun Cluster-System verwendet SCSI-Plattenreservierungen zur Implementierung des Fehlerschutzes. Mit den SCSI-Reservierungen werden die Multihostgeräte vor den ausgefallenen Knoten "geschützt" und der Zugriff auf diese Platten wird verhindert.

SCSI-2-Plattenreservierungen unterstützen eine Reservierungsform, die entweder allen mit der Platte verbundenen Knoten den Zugriff gestattet (wenn keine Reservierung vorliegt) oder den Zugriff auf einen einzelnen Knoten beschränkt (der Knoten, der die Reservierung aufweist).

Wenn ein Cluster-Mitglied erkennt, dass ein anderer Knoten nicht mehr über den Cluster-Interconnect kommuniziert, leitet er ein Fehlerschutzverfahren ein, um den anderen Knoten am Zugriff auf die gemeinsam genutzten Platten zu hindern. Wenn dieser Fehlerschutz eintritt, gerät der geschützte Knoten in Panik und eine Meldung zum "Reservierungskonflikt" wird auf seiner Konsole angezeigt.

Die Entdeckung, dass ein Knoten nicht länger Mitglied des Clusters ist, bewirkt eine SCSI-Reservierung auf allen Platten, die von diesem und anderen Knoten gemeinsam genutzt werden. Der geschützte Knoten ist sich möglicherweise nicht "bewusst", dass er geschützt wird, und wenn er versucht, auf eine der gemeinsam genutzten Platten zuzugreifen, entdeckt er die Reservierung und gerät in Panik.

Failfast-Mechanismus für den Fehlerschutz

Der Mechanismus, durch den das Cluster-Framework sicherstellt, dass ein ausgefallener Knoten nicht neu booten und auf gemeinsam genutzten Speicher schreiben kann, wird als *Failfast* bezeichnet.

Knoten, die Cluster-Mitglieder sind, aktivieren kontinuierlich ein spezifisches ioctl, `MHIOCENFAILFAST`, für die Platten, auf die sie zugreifen. Hierzu gehören auch die Quorum-Platten. Dieses ioctl ist eine Direktive für den Plattentreiber. Das ioctl gibt einem Knoten die Möglichkeit, selbst in Panik zu geraten, wenn er durch eine Reservierung eines anderen Knotens nicht auf die Platte zugreifen kann.

Das ioctl `MHIOCENFAILFAST` veranlasst den Treiber, die Fehlerrückgabe jedes Lese- und Schreibvorgangs, den ein Knoten an die Platte ausgibt, auf den Fehlercode `Reservation_Conflict` zu überprüfen. Das ioctl gibt im Hintergrund regelmäßig einen Testvorgang an die Platte aus, um sie auf die Meldung `Reservation_Conflict` zu überprüfen. Sowohl der Vordergrund- als auch der Hintergrund-Flussaufzeichnungspfad geraten in Panik, wenn `Reservation_Conflict` zurückgegeben wird.

Bei SCSI-2-Platten sind die Reservierungen nicht dauerhaft – sie werden beim erneuten Booten von Knoten gelöscht. Bei SCSI-3-Platten mit PGR (Persistent Group Reservation) werden die Reservierungsinformationen auf der Platte gespeichert und bleiben auch nach dem Booten von Knoten erhalten. Der Failfast-Mechanismus funktioniert stets auf dieselbe Weise, ob Sie nun SCSI-2- oder SCSI-3-Platten verwenden.

Wenn ein Knoten die Konnektivität mit anderen Knoten im Cluster verliert und nicht zu einer Partition gehört, die ein Quorum erzielen kann, wird er erzwungenermaßen von einem anderen Knoten aus dem Cluster entfernt. Ein anderer Knoten, der Teil der Partition ist, die ein Quorum erzielen kann, belegt die gemeinsam genutzten Platten mit Reservierungen. Wenn der Knoten, der über kein Quorum verfügt, versucht, auf die gemeinsam genutzten Platten zuzugreifen, wird ein Reservierungskonflikt gemeldet und der Knoten gerät, bewirkt durch den Failfast-Mechanismus, in Panik.

Nach der Panik kann der Knoten neu booten und versuchen, dem Cluster wieder beizutreten oder in Clustern aus SPARC-basierten Systemen an der OpenBoot™ PROM (OBP)-Eingabeaufforderung bleiben. Welche Aktion eingeleitet wird, bestimmt die Einstellung des `auto-boot?`-Parameters. Sie können den Parameter `auto-boot?` mit `eeprom(1M)` an der OpenBoot PROM-Eingabeaufforderung `ok` in einem SPARC-basierten Cluster einstellen. Alternativ dazu können Sie diesen Parameter auch mit dem SCSI-Dienstprogramm einstellen, das Sie nach dem Booten des BIOS in einem x86-basierten Cluster ausführen können.

Informationen zu Quorum-Konfigurationen

Die folgende Liste enthält Informationen zu Quorum-Konfigurationen:

- Quorum-Geräte können Benutzerdaten enthalten.
- In einer N+1-Konfiguration, in der N Quorum-Geräte mit einem der 1 bis N Knoten und dem N+1-Knoten verbunden sind, überlebt der Cluster den Ausfall aller 1 bis N Knoten oder eines der N/2-Knoten. Bei dieser Verfügbarkeit wird vorausgesetzt, dass das Quorum-Gerät ordnungsgemäß funktioniert.
- In einer N-Knotenkonfiguration, in der ein einzelnes Quorum-Gerät mit allen Knoten verbunden ist, kann der Cluster den Ausfall aller N-1-Knoten überleben. Bei dieser Verfügbarkeit wird vorausgesetzt, dass das Quorum-Gerät ordnungsgemäß funktioniert.
- In einer N-Knoten-Konfiguration, bei der ein einzelnes Quorum-Gerät mit allen Knoten verbunden wird, kann der Cluster den Ausfall des Quorum-Geräts überleben, sofern alle Cluster-Knoten verfügbar sind.

Beispiele für Quorum-Konfigurationen, die Sie besser vermeiden sollten, finden Sie unter „Unzulässige Quorum-Konfigurationen“ auf Seite 63. Beispiele für empfohlene Quorum-Konfigurationen finden Sie unter „Empfohlene Quorum-Konfigurationen“ auf Seite 61.

Erfüllen der Anforderungen für Quorum-Geräte

Sie müssen folgende Anforderungen erfüllen. Wenn Sie diese Anforderungen missachten, kann die Verfügbarkeit Ihres Clusters beeinträchtigt werden.

- Stellen Sie sicher, dass die Sun Cluster-Software Ihr Gerät als Quorum-Gerät unterstützt.

Hinweis – Eine Liste der Geräte, die von der Sun Cluster-Software als Quorum-Geräte unterstützt werden, erhalten Sie bei Ihrem Sun-Dienstanbieter.

Die Sun Cluster-Software unterstützt zwei Arten von Quorum-Geräten:

- Gemeinsam genutzte Multihost-Platten, die SCSI-3-PGR-Reservierungen unterstützen
- Gemeinsam genutzte Dual-Host-Platten, die SCSI-2-Reservierungen unterstützen
- In einer Konfiguration mit zwei Knoten müssen Sie mindestens ein Quorum-Gerät konfigurieren, um sicherzustellen, dass ein einzelner Knoten weiterbesteht, wenn der andere Knoten ausfällt. Informationen hierzu finden Sie in [Abbildung 3-2](#).

Beispiele für Quorum-Konfigurationen, die Sie besser vermeiden sollten, finden Sie unter „Unzulässige Quorum-Konfigurationen“ auf Seite 63. Beispiele für empfohlene Quorum-Konfigurationen finden Sie unter „Empfohlene Quorum-Konfigurationen“ auf Seite 61.

Einhalten von Empfehlungen für Quorum-Geräte

Ermitteln Sie mithilfe der folgenden Informationen, welche Quorum-Konfiguration am besten für Ihre Topologie geeignet ist:

- Verfügen Sie über ein Gerät, das an alle Knoten des Clusters angeschlossen werden kann?
 - Falls ja, konfigurieren Sie dieses Gerät als Ihr Quorum-Gerät. Es ist *nicht* nötig, dass Sie ein weiteres Quorum-Gerät konfigurieren, da Ihre Konfiguration bereits optimal ist.



Achtung – Wenn Sie diese Anforderung ignorieren und ein weiteres Quorum-Gerät konfigurieren, wird die Verfügbarkeit des Clusters durch das zusätzliche Quorum-Gerät verringert.

- Falls nicht, konfigurieren Sie Ihr Gerät bzw. Ihre Geräte mit zwei Anschlüssen.
- Stellen Sie sicher, dass die Gesamtanzahl der Stimmen, die von Quorum-Geräten bereitgestellt werden, grundsätzlich geringer ist als die Gesamtanzahl der von Knoten bereitgestellten Stimmen. Ist dies nicht der Fall, können die Knoten keinen Cluster bilden, wenn alle Platten verfügbar sind – auch wenn sämtliche Knoten funktionsfähig sind.

Hinweis – In bestimmten Umgebungen sollten Sie die allgemeine Cluster-Verfügbarkeit reduzieren, um sie Ihren Bedürfnissen anzupassen. Ignorieren Sie in diesen Fällen die genannte Empfehlung. Wenn diese Empfehlung jedoch nicht wahrgenommen wird, wird die Verfügbarkeit insgesamt beeinträchtigt. Die unter „Untypische Quorum-Konfigurationen“ auf Seite 63 beschriebene Konfiguration beispielsweise ist weniger verfügbar: Die Quorum-Stimmen übersteigen hier die Knotenstimmen. Der Cluster ist so konfiguriert, dass der gesamte Cluster ausfällt, wenn der Zugriff auf den von Nodes A und Node B gemeinsam genutzten Speicher verloren geht.

Unter „Untypische Quorum-Konfigurationen“ auf Seite 63 finden Sie eine Ausnahme für diese Empfehlung.

- Geben Sie ein Quorum-Gerät für sämtliche Knotenpaare an, die auf ein gemeinsam genutztes Speichergerät zugreifen. Diese Quorum-Konfiguration beschleunigt den Fehlerschutzvorgang. Informationen hierzu finden Sie unter „Quorum in

Konfigurationen mit mehr als zwei Knoten " auf Seite 61.

- Im Allgemeinen steigt die Cluster-Verfügbarkeit, falls durch das Hinzufügen eines Quorum-Geräts die Gesamtanzahl der Cluster-Stimmen gerade wird.
- Quorum-Geräte verlangsamen die Neukonfiguration nach dem Beitritt eines Knotens bzw. dem Ausfall eines Knotens geringfügig. Fügen Sie aus diesem Grund nicht mehr Quorum-Geräte als erforderlich hinzu.

Beispiele für Quorum-Konfigurationen, die Sie besser vermeiden sollten, finden Sie unter „Unzulässige Quorum-Konfigurationen " auf Seite 63. Beispiele für empfohlene Quorum-Konfigurationen finden Sie unter „Empfohlene Quorum-Konfigurationen " auf Seite 61.

Empfohlene Quorum-Konfigurationen

Dieser Abschnitt enthält Beispiele für empfohlene Quorum-Konfigurationen. Beispiele für Quorum-Konfigurationen, die besser vermieden werden sollten, finden Sie unter „Unzulässige Quorum-Konfigurationen " auf Seite 63.

Quorum in Zwei-Knoten-Konfigurationen

Zwei Quorum-Stimmen sind erforderlich, damit ein Zwei-Knoten-Cluster gebildet werden kann. Diese beiden Stimmen können von den beiden Cluster-Knoten oder von einem Knoten und einem Quorum-Gerät stammen.

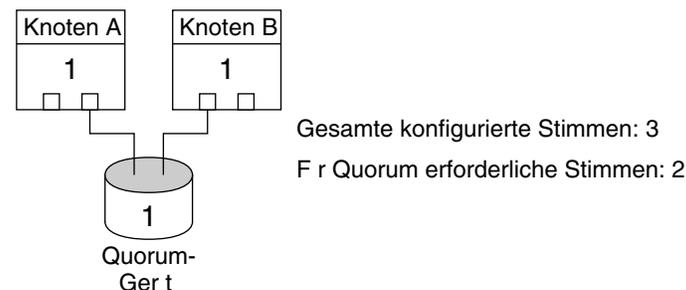
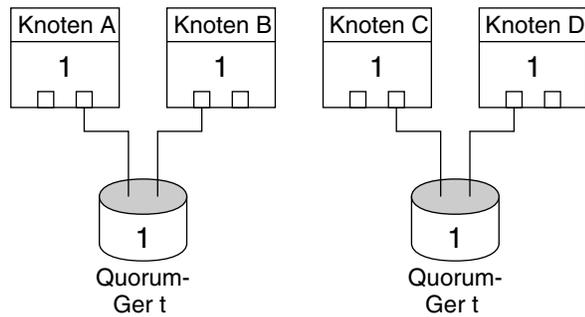


ABBILDUNG 3-2 Zwei-Knoten-Konfiguration

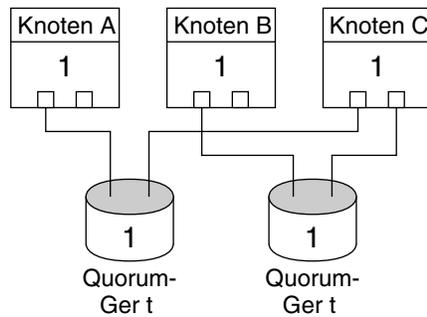
Quorum in Konfigurationen mit mehr als zwei Knoten

Cluster mit mehr als zwei Knoten können auch ohne ein Quorum-Gerät konfiguriert werden. In diesem Fall können Sie den Cluster nicht ohne eine Knotenmehrheit im Cluster starten.



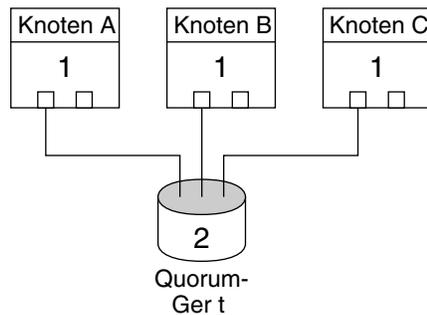
Gesamte konfigurierte Stimmen: 6
 Für Quorum erforderliche Stimmen: 4

In dieser Konfiguration muss jedes Paar für beide Paarknoten verfügbar sein, um zu überleben.



Gesamte konfigurierte Stimmen: 5
 Für Quorum erforderliche Stimmen: 3

In dieser Konfiguration sind die Anwendungen in der Regel für die Ausführung auf Knoten A und Knoten B konfiguriert und verwenden Knoten C als Hot-Spare.



Gesamte konfigurierte Stimmen: 5
 Für Quorum erforderliche Stimmen: 3

In dieser Konfiguration können ein oder mehrere Knoten und das Quorum-Gerät einen Cluster bilden.

Untypische Quorum-Konfigurationen

In [Abbildung 3-3](#) wird vorausgesetzt, dass auf **Node A** und **Node B** sehr wichtige Anwendungen (beispielsweise Oracle-Datenbanken) ausgeführt werden. Wenn **Node A** und **Node B** nicht verfügbar sind und nicht auf gemeinsam genutzte Daten zugreifen können, sollten Sie den gesamten Cluster herunterfahren. Anderenfalls ist diese Konfiguration nicht optimal, da sie keine Hochverfügbarkeit bietet.

Informationen zu Empfehlungen, für die diese Ausnahme gilt, finden Sie unter „[Einhalten von Empfehlungen für Quorum-Geräte](#)“ auf [Seite 60](#).

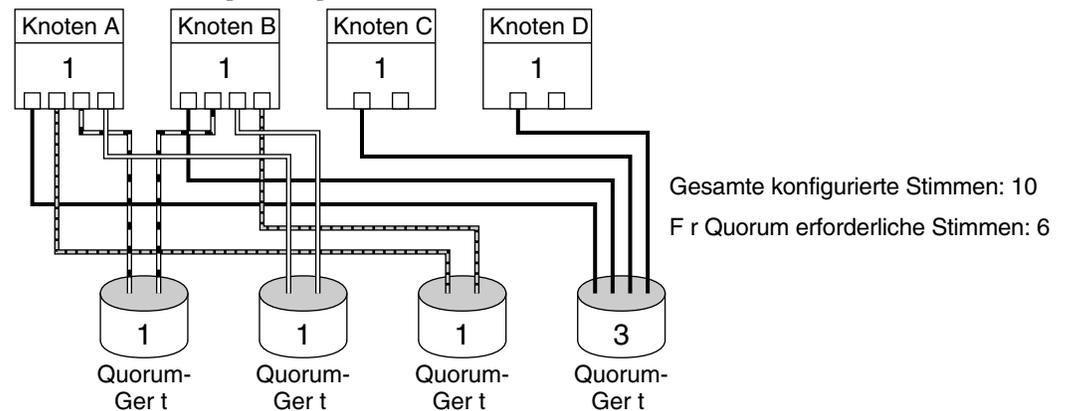
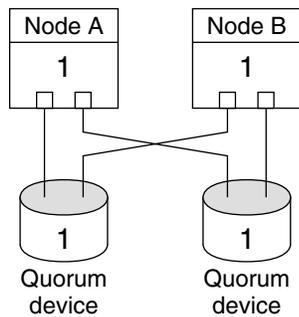


ABBILDUNG 3-3 Untypische Konfiguration

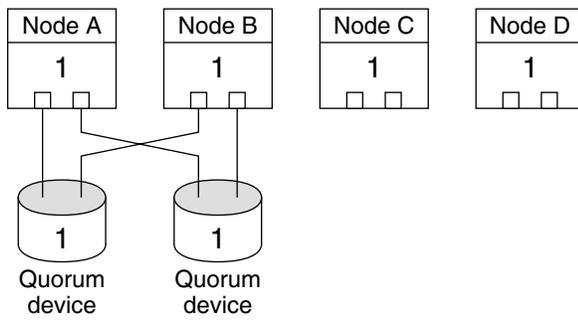
Unzulässige Quorum-Konfigurationen

Dieser Abschnitt enthält Beispiele für Quorum-Konfigurationen, die Sie vermeiden sollten. Beispiele für empfohlene Quorum-Konfigurationen finden Sie unter „[Empfohlene Quorum-Konfigurationen](#)“ auf [Seite 61](#).



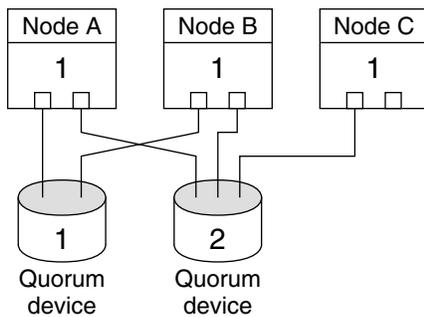
Total configured votes: 4
 Votes required for quorum: 3

This configuration violates the best practice that quorum device votes should be strictly less than votes of nodes.



Total configured votes: 6
 Votes required for quorum: 4

This configuration violates the best practice that you should not add quorum devices to make total votes even. This configuration does not add availability.



Total configured votes: 6
 Votes required for quorum: 4

This configuration violates the best practice that quorum device votes should be strictly less than votes of nodes.

Datendienste

Der Begriff *Datendienst* beschreibt eine Anwendung wie Sun Java System Web Server oder Oracle, die eher zur Ausführung auf einem Cluster als auf einem Einzelserver konfiguriert wurde. Ein Datendienst besteht aus einer Anwendung, spezialisierten Sun Cluster-Konfigurationsdateien und Sun Cluster -Verwaltungsmethoden, mit denen die nachfolgenden Anwendungsaktionen gesteuert werden.

- Beginn
- Anhalten
- Überwachen und Ergreifen von Korrekturmaßnahmen

Informationen zu Datendienst-Typen finden Sie unter „Datendienste“ in *Sun Cluster Überblick für das Betriebssystem Solaris*.

In *Abbildung 3-4* wird eine Anwendung, die auf einem einzigen Anwendungsserver ausgeführt wird (Einzelservermodell) mit derselben Anwendung auf einem Cluster (Cluster-Servermodell) verglichen. Der einzige Unterschied zwischen den beiden Konfigurationen besteht darin, dass die Cluster-Anwendung ggf. schneller ausgeführt wird und eine bessere Hochverfügbarkeit zeigt.

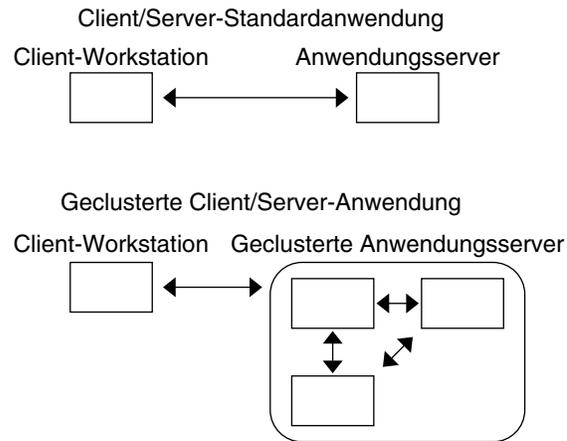


ABBILDUNG 3-4 Standardkonfiguration versus Client/Server-Konfiguration als Cluster

Beim Einzelservermodell konfigurieren Sie die Anwendung für den Zugriff auf den Server über eine bestimmte öffentliche Netzwerkschnittstelle (einen Hostnamen). Der Hostname ist diesem realen Server zugeordnet.

Beim geclusterten Servermodell ist die öffentliche Netzwerkschnittstelle ein *logischer Hostname* oder eine *gemeinsam genutzte Adresse*. Der Begriff *Netzwerkressourcen* bezeichnet sowohl die logischen Hostnamen als auch die gemeinsam genutzten Adressen.

Für bestimmte Datendienste müssen Sie entweder logische Hostnamen oder gemeinsam genutzte Adressen als Netzwerkschnittstellen angeben. Logische Hostnamen und gemeinsam genutzte Adressen sind nicht austauschbar. Für andere Datendienste können Sie wahlweise logische Hostnamen oder gemeinsam genutzte Adressen angeben. Details zum jeweils erforderlichen Schnittstellentyp finden Sie in den Angaben zur Installation und Konfiguration für den jeweiligen Datendienst.

Eine Netzwerkressource ist keinem spezifischen realen Server zugeordnet. Netzwerkressourcen können zwischen den realen Servern migrieren.

Eine Netzwerkressource ist zunächst einem einzigen Knoten zugeordnet, dem *Primärknoten*. Wenn der Primärknoten ausfällt, werden Netzwerkressource und Anwendungsressource mit einem Failover-Vorgang auf einen anderen Cluster-Knoten (einen Sekundärknoten) umgeleitet. Wenn die Netzwerkressource ein Failover durchführt, wird die Anwendungsressource nach einer kurzen Verzögerung auf dem Sekundärknoten weiter ausgeführt.

In *Abbildung 3-5* wird das Einzelservermodell mit dem geclusterten Servermodell verglichen. Beachten Sie, dass eine Netzwerkressource (in diesem Beispiel ein logischer Hostname) in einem geclusterten Servermodell zwischen mindestens zwei Cluster-Knoten wechseln kann. Die Anwendung ist zur Verwendung dieses logischen Hostnamens anstelle eines Hostnamens konfiguriert, der einem bestimmten Server zugeordnet ist.

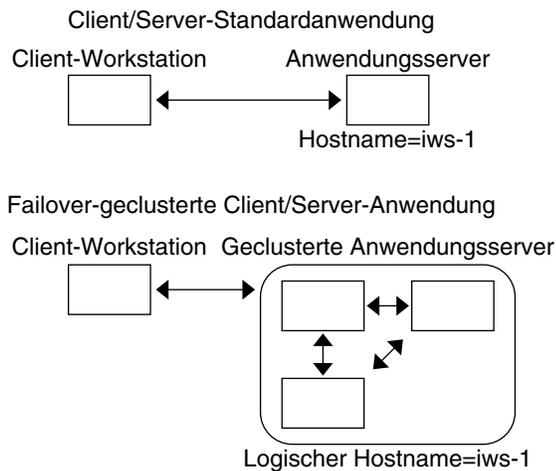


ABBILDUNG 3-5 Festgelegter Hostname versus logischer Hostname

gemeinsam genutzte Adresse ist zunächst ebenfalls einem einzigen Knoten zugeordnet. Dieser Knoten wird als globaler Schnittstellenknoten (GIF-Knoten) bezeichnet. Eine gemeinsam genutzte Adresse (die so genannte *globale Schnittstelle*) wird als einzige Netzwerkschnittstelle zum Cluster eingesetzt.

Der Unterschied zwischen dem Modell logischer Hostname und dem Modell Scalable-Dienst besteht darin, dass beim letztgenannten auf jedem Knoten die gemeinsam genutzte Adresse auf der Schleifenschnittstelle ebenfalls aktiv konfiguriert ist. Dank dieser Konfiguration können mehrere Instanzen eines Datendienstes auf mehreren Knoten gleichzeitig aktiv sein. Der Begriff "Scalable-Dienst" bedeutet, dass Sie die CPU-Leistung für die Anwendung durch Hinzufügen zusätzlicher Cluster-Knoten erhöhen können und die Leistung verbessert wird.

Wenn der globale Schnittstellenknoten ausfällt, kann die gemeinsam genutzte Adresse auf einem anderen Knoten gestartet werden, auf dem ebenfalls eine Instanz der Anwendung ausgeführt wird (wobei dieser andere Knoten zum globalen Schnittstellenknoten wird). Oder die gemeinsam genutzte Adresse wird mit einem Failover auf einen anderen Cluster-Knoten verschoben, auf dem die Anwendung zuvor nicht ausgeführt wurde.

In [Abbildung 3-6](#) wird die Einzelserverkonfiguration mit der Cluster-Konfiguration mit Scalable-Diensten verglichen. Beachten Sie, dass die gemeinsam genutzte Adresse in der Scalable-Dienst-Konfiguration auf allen Knoten vorhanden ist. Ähnlich wie bei der Verwendung eines logischen Hostnames für Failover-Datendienste wird die Anwendung zur Verwendung dieser gemeinsam genutzten Adresse anstelle eines einem bestimmten Server zugeordneten Hostnames konfiguriert.

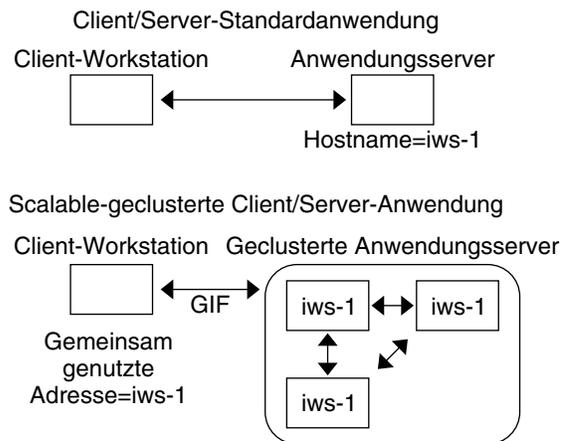


ABBILDUNG 3-6 Festgelegter Hostname versus gemeinsam genutzte Adresse

Datendienstmethoden

Die Sun Cluster-Software stellt eine Reihe von Dienstverwaltungsmethoden zur Verfügung. Diese Methoden werden vom Ressourcengruppen-Manager (RGM) gesteuert und eingesetzt, um die Anwendung auf den Cluster-Knoten zu starten, anzuhalten und zu überwachen. Mit diesen Methoden, der Cluster-Framework-Software und den Multihostgeräten können die Anwendungen als Failover- oder Scalable-Datendienste verwendet werden.

Der RGM verwaltet auch Ressourcen im Cluster einschließlich der Anwendungsinstanzen und Netzwerkressourcen (logische Hostnamen und gemeinsam genutzte Adressen).

Zusätzlich zu den Methoden der Sun Cluster-Software stellt das Sun Cluster-System auch eine API und verschiedene Datendienst-Entwicklungstools zur Verfügung. Mit diesen Tools können Anwendungsprogrammierer die Datendienstmethoden entwickeln, die sie zum Ausführen anderer Anwendungen als hoch verfügbare Datendienste mit der Sun Cluster-Software benötigen.

Failover-Datendienste

Wenn der Knoten ausfällt, auf dem der Datendienst ausgeführt wird (der Primärknoten), migriert der Dienst ohne Benutzereingriff auf einen anderen funktionsfähigen Knoten. Failover-Dienste verwenden eine *Failover-Ressourcengruppe*, die Ressourcen für Anwendungsinstanzen und Netzwerkressourcen enthält (*logische Hostnamen*). Logische Hostnamen sind IP-Adressen, die auf einem Knoten als aktiv konfiguriert werden können. Später werden sie auf dem Originalknoten automatisch dekonfiguriert und auf einem anderen Knoten konfiguriert.

Für Failover-Datendienste werden die Anwendungsinstanzen nur auf einem einzigen Knoten ausgeführt. Wenn der Fehler-Monitor einen Fehler erkennt, versucht er entweder die Instanz auf demselben Knoten erneut zu starten oder die Instanz auf einem anderen Knoten zu starten (Failover). Das Ergebnis hängt davon ab, wie der Datendienst konfiguriert wurde.

Skalierbare Datendienste

Der Scalable-Datendienst kann aktive Instanzen auf mehreren Knoten ausführen. Scalable-Dienste verwenden die beiden folgenden Ressourcengruppen:

- Eine *Scalable-Ressourcengruppe* enthält die Anwendungsressourcen.
- Eine *Failover-Ressourcengruppe* enthält die Netzwerkressourcen (*gemeinsam genutzte Adressen*), von denen der Scalable-Dienst abhängt.

Die Scalable-Ressourcengruppe kann auf mehreren Knoten online sein, so dass mehrere Instanzen dieses Dienstes gleichzeitig ausgeführt werden können. Die Failover-Ressourcengruppe, welche die gemeinsam genutzten Adressen hostet, ist jeweils nur auf einem Knoten online. Alle Knoten, die einen Scalable-Dienst hosten, verwenden dieselbe gemeinsam genutzte Adresse, um den Dienst zu hosten.

Dienstanforderungen erreichen den Cluster über eine einzige Netzwerkschnittstelle (die globale Schnittstelle). Diese Anforderungen werden auf Grundlage mehrerer vordefinierter Algorithmen, die in den *Lastausgleichsverfahren* festgelegt sind, an die Knoten verteilt. Der Cluster kann die Dienstlast mithilfe des Lastausgleichsverfahrens zwischen mehreren Knoten ausgleichen. Mehrere globale Schnittstellen können auf anderen Knoten vorhanden sein, die weitere, gemeinsam genutzte Adressen hosten.

Bei Scalable-Diensten werden die Anwendungsinstanzen auf mehreren Knoten gleichzeitig ausgeführt. Wenn der Knoten mit der globalen Schnittstelle ausfällt, wird die globale Schnittstelle auf einen anderen Knoten verschoben. Wenn eine Anwendungsinstanz bei der Ausführung fehlschlägt, versucht die Instanz, auf demselben Knoten neu zu starten.

Wenn eine Anwendungsinstanz nicht auf demselben Knoten neu gestartet werden kann und ein anderer, nicht verwendeter Knoten für die Ausführung des Dienstes konfiguriert ist, wird der Dienst im Rahmen eines Failover-Vorgangs auf den nicht verwendeten Knoten verschoben. Andernfalls wird der Dienst weiter auf den restlichen Knoten ausgeführt und führt möglicherweise zu einer Reduzierung des Dienstdurchsatzes.

Hinweis – Der TCP-Zustand für jede Anwendungsinstanz bleibt auf dem Knoten mit der Instanz, nicht auf dem Knoten mit der globalen Schnittstelle. Deswegen hat ein Ausfall des Knotens mit der globalen Schnittstelle keine Auswirkung auf die Verbindung.

Abbildung 3-7 zeigt ein Beispiel für eine Failover- und eine Scalable-Ressourcengruppe und die Abhängigkeiten, die zwischen den beiden für Scalable-Dienste bestehen. Dieses Beispiel enthält drei Ressourcengruppen. Die Failover-Ressourcengruppe enthält Anwendungsressourcen für hoch verfügbare DNS und Netzwerkressourcen, die sowohl vom hoch verfügbaren DNS als auch vom hoch verfügbaren Apache Web Server genutzt werden (nur in SPARC-basierten Clustern verwendet). Die Scalable-Ressourcengruppen enthalten nur Anwendungsinstanzen des Apache Web Servers. Beachten Sie, dass Ressourcengruppenabhängigkeiten zwischen den Scalable- und den Failover-Ressourcengruppen (durchgezogene Linien) bestehen. Außerdem hängen alle Apache-Anwendungsressourcen von der Netzwerkressource `schost-2` ab, die eine gemeinsam genutzte Adresse ist (gestrichelte Linien).

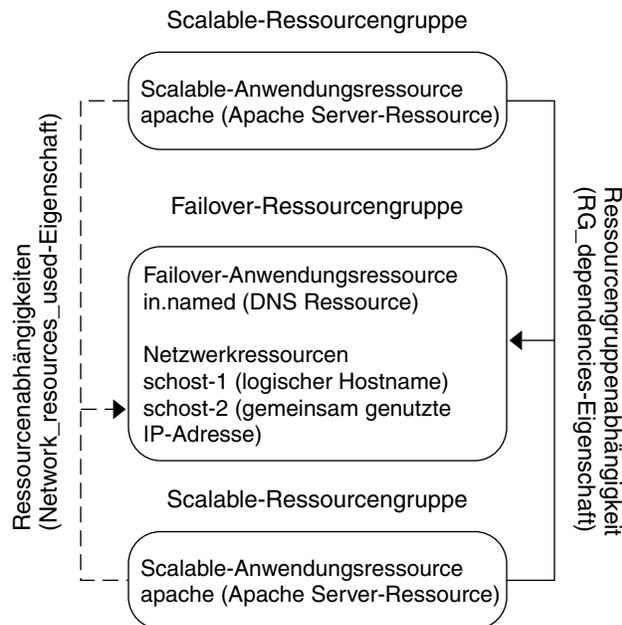


ABBILDUNG 3-7 SPARC: Beispiel für Failover- und Scalable-Ressourcengruppen

Lastausgleichsverfahren

Der Lastausgleich verbessert die Leistung der Scalable-Dienste sowohl bei der Antwortzeit als auch beim Durchsatz. Es gibt zwei Klassen von Scalable-Datendiensten.

- Rein
- Sticky

Bei einem *reinen* Dienst, kann jede Instanz Client-Anforderungen beantworten. Bei einem *Sticky*-Dienst kann ein Client Anforderungen an dieselbe Instanz senden. Diese Anforderungen werden nicht an andere Instanzen umgeleitet.

Ein reiner Dienst verwendet ein gewichtetes Lastausgleichsverfahren. Bei diesem Lastausgleichsverfahren werden Client-Anforderungen standardmäßig gleichmäßig auf die Serverinstanzen im Cluster verteilt. Angenommen, jeder Knoten in einem Drei-Knoten-Cluster hat eine Gewichtung von 1. Jeder Knoten bedient 1/3 der Anforderungen aller beliebigen Clients im Auftrag des Dienstes. Gewichtungen können jederzeit vom Verwalter über die `scrgadm (1M)`-Befehlsschnittstelle oder über die SunPlex-Manager-GUI geändert werden.

Ein Sticky-Dienst hat zwei Typen, *normal-sticky* und *Platzhalter-Sticky*. Mit Sticky-Diensten können Sitzungen auf Anwendungsebene gleichzeitig über mehrere TCP-Verbindungen ausgeführt werden, um den Zustandsspeicher (Anwendungssitzungszustand) gemeinsam zu nutzen.

Mit normalen Sticky-Diensten kann ein Client den Zustand zwischen mehreren gleichzeitigen TCP-Verbindungen teilen. Der Client wird als "sticky" gegenüber der Serverinstanz bezeichnet, die einen einzigen Port abhört. Der Client hat die Sicherheit, dass alle seine Anforderungen an dieselbe Serverinstanz gehen. Voraussetzung hierfür ist, dass die Instanz aktiv und zugänglich bleibt und das Lastausgleichsverfahren nicht geändert wird, solange der Dienst online ist.

Beispiel: Ein Webbrowser auf dem Client stellt mithilfe von drei verschiedenen TCP-Verbindungen eine Verbindung mit einer gemeinsam genutzten IP-Adresse an Port 80 her. Die Verbindungen tauschen jedoch zwischengespeicherte Sitzungsinformationen über den Dienst aus.

Eine Verallgemeinerung eines Sticky-Verfahrens erstreckt sich auf mehrere Scalable-Dienste, die im Hintergrund Sitzungsinformationen bei derselben Instanz austauschen. Wenn diese Dienste im Hintergrund Sitzungsinformation bei derselben Instanz austauschen, wird der Client als "sticky" bezeichnet, weil mehrere Serverinstanzen auf demselben Knoten unterschiedliche Ports abhören. .

Beispiel: Ein E-Commerce-Kunde füllt seinen Einkaufskorb mit Artikeln mithilfe von HTTP an Port 80, wechselt dann jedoch zum Senden sicherer Daten mit SSL auf Port 443, um die Artikel im Einkaufskorb mit Kreditkarte zu bezahlen.

Platzhalter-Sticky-Dienste verwenden dynamisch zugewiesene Port-Nummern, erwarten jedoch trotzdem, dass die Client-Anforderungen an denselben Knoten geleitet werden. Der Client ist "Sticky-Platzhalter" bei Ports, die dieselbe IP-Adresse aufweisen.

Ein gutes Beispiel dieses Verfahrens ist der passive FTP-Modus. Beispiel: Ein Client stellt an Port 21 eine Verbindung mit einem FTP-Server her. Anschließend wird er vom Server angewiesen, eine neue Verbindung mit einem Listener-Port-Server im dynamischen Port-Bereich herzustellen. Alle Anforderungen an diese IP-Adresse werden an denselben Knoten weitergeleitet, den der Server dem Client über die Steuerinformationen angegeben hat. .

Für jedes dieser Sticky-Verfahren ist das gewichtete Lastausgleichsverfahren standardmäßig gültig. Daher wird die ursprüngliche Client-Anforderung an die vom Lastausgleich vorgegebene Instanz geleitet. Nachdem der Client eine Affinität zum Knoten mit der ausgeführten Instanz aufgebaut hat, werden zukünftige Anforderungen unter bestimmten Bedingungen an diese Instanz geleitet. Der Knoten muss zugänglich sein und das Lastausgleichsverfahren darf sich nicht geändert haben.

Weitere Details zu spezifischen Lastausgleichsverfahren werden nachstehend erläutert.

- Gewichtet. Die Last wird entsprechend den angegebenen Gewichtungswerten auf mehrere Knoten verteilt. Dieses Verfahren wird mit dem `LB_WEIGHTED`-Wert für die `Load_balancing_weights`-Eigenschaft eingestellt. Wenn die Gewichtung für einen Knoten nicht ausdrücklich angegeben ist, beträgt die Standardgewichtung für diesen Knoten Eins.
Das gewichtete Verfahren leitet einen gewissen Prozentsatz des Datenverkehrs von Clients auf einen bestimmten Knoten. Bei X =Gewichtung und A =Gesamtgewichtung aller aktiven Knoten ist davon auszugehen, dass etwa X/A aller neuen Verbindungen zu einem aktiven Knoten geleitet werden. Allerdings muss dafür die Gesamtanzahl der Verbindungen groß genug sein. Dieses Verfahren befasst sich nicht mit einzelnen Anforderungen.
Beachten Sie, dass es sich nicht um ein Round-Robin-Verfahren handelt. Bei einem Round-Robin-Verfahren würde jede Anforderung eines Clients an einen unterschiedlichen Knoten geleitet. Beispielsweise würde die erste Anforderung an Knoten 1, die zweite Anforderung an Knoten 2 geleitet und so weiter.
- Sticky. Bei diesem Verfahren ist der Satz an Ports bei der Konfiguration der Anwendungsressourcen bekannt. Dieses Verfahren wird mit dem `LB_STICKY`-Wert für die `Load_balancing_policy`-Ressourceneigenschaft eingestellt.
- Sticky-Platzhalter. Dieses Verfahren ist eine Obermenge des normalen "Sticky"-Verfahrens. Bei einem durch die IP-Adresse identifizierten Scalable-Dienst werden die Ports vom Server zugewiesen (und sind vorher nicht bekannt). Die Ports können sich ändern. Dieses Verfahren wird mit dem `LB_STICKY_WILD`-Wert für die `Load_balancing_policy`-Ressourceneigenschaft eingestellt.

Failback-Einstellungen

Ressourcengruppen wechseln beim Failover von einem Knoten auf einen anderen. Wenn ein solcher Failover auftritt, wird der ursprüngliche Sekundärknoten zum neuen Primärknoten. Die Failback-Einstellungen legen die Aktionen fest, die ausgeführt werden, sobald der ursprüngliche Primärknoten wieder online gebracht wird. Die Optionen sind, entweder den ursprünglichen Primärknoten wieder zum Primärknoten zu machen (Failback) oder den aktuellen Primärknoten als solchen zu belassen. Sie geben die gewünschte Option mit der Einstellung der `Failback`-Ressourcengruppeneigenschaft an.

Bei bestimmten Instanzen kann die Failback-Einstellung zu einer reduzierten Verfügbarkeit der Ressourcengruppe führen, wenn der Originalknoten mit der Ressource ausfällt und mehrmals neu bootet.

Fehler-Monitore der Datendienste

Jeder Sun Cluster-Datendienst stellt einen Fehler-Monitor zur Verfügung, der regelmäßig den Datendienst auf einwandfreies Funktionieren prüft. Ein Fehler-Monitor prüft, ob der bzw. die Anwendungsdämon(en) ausgeführt und die

Clients bedient werden. Aufgrund der von den Testsignalen zurückgegebenen Informationen können vordefinierte Aktionen wie ein Neustart eines Dämons oder das Einleiten eines Failovers ausgelöst werden.

Entwickeln neuer Datendienste

Sun stellt Vorlagen für Konfigurationsdateien und Verwaltungsmethoden zur Verfügung, mit denen Sie unterschiedliche Anwendungen zu Failover- oder Scalable-Diensten innerhalb eines Clusters machen können. Falls Sun die Anwendung, die Sie als Failover- oder Scalable-Dienst ausführen möchten, nicht anbietet, haben Sie eine Alternative. Verwenden Sie eine Sun Cluster-API oder die DSET-API, um die Anwendung als Failover- oder Scalable-Dienst zu konfigurieren. Es können jedoch nicht alle Anwendungen als Scalable-Dienst programmiert werden.

Eigenschaften von Scalable-Diensten

Sie können anhand einer Reihe von Kriterien feststellen, ob eine Anwendung als Scalable-Dienst ausgeführt werden kann. Um festzustellen, ob Ihre Anwendung als Scalable-Dienst ausgeführt werden kann, lesen Sie den Abschnitt „Analysieren der Eignung einer Anwendung“ in *Sun Cluster Entwicklerhandbuch Datendienste für Solaris OS*. Diese Kriterien werden nachfolgend zusammengefasst.

- Erstens: Ein solcher Dienst besteht aus einer oder mehreren Server *instanzen*. Jede Instanz wird auf einem anderen Knoten des Clusters ausgeführt. Zwei oder mehr Instanzen desselben Dienstes können nicht auf demselben Knoten ausgeführt werden.
- Zweitens: Wenn der Dienst einen externen logischen Datenspeicher zur Verfügung stellt, müssen Sie mit Bedacht vorgehen. Der gleichzeitige Zugriff auf diesen Speicher durch mehrere Serverinstanzen muss synchronisiert werden, um den Verlust von Aktualisierungen oder das Lesen von Daten zu verhindern, wenn diese gerade geändert werden. Beachten Sie, dass der Begriff "extern" verwendet wird, um diesen Speicher vom Arbeitsspeicherstatus zu unterscheiden. Der Begriff "logisch" weist darauf hin, dass der Speicher als einzelne Entität angezeigt wird, obwohl er möglicherweise repliziert wurde. Außerdem hat dieser logische Datenspeicher die Eigenschaft, dass eine Aktualisierung des Speichers durch eine beliebige Instanz sofort von anderen Instanzen "erkannt" wird.

Das Sun Cluster-System stellt einen solchen externen Speicher über das Cluster-Dateisystem und die globalen Partitionen im raw-Modus zur Verfügung. Angenommen, ein Dienst schreibt neue Daten in eine externe Protokolldatei oder ändert die dort vorhandenen Daten. Wenn mehrere Instanzen dieses Dienstes ausgeführt werden, hat jede Instanz Zugriff auf dieses externe Protokoll und jede

kann gleichzeitig auf dieses Protokoll zugreifen. Jede Instanz muss den Zugriff auf dieses Protokoll synchronisieren, andernfalls kommt es zu Interferenzen zwischen den Instanzen. Der Dienst kann die normale Solaris-Dateisperrung über `fcntl(2)` und `lockf(3C)` verwenden, um die gewünschte Synchronisierung zu erreichen.

Ein weiteres Beispiel für diese Art des Speichers ist eine Backend-Datenbank, wie beispielsweise Hochverfügbarkeits-Oracle Real Application Clusters Guard für SPARC-basierte Cluster oder Oracle. Bei diesem Typ von Backend-Datenbankserver ist die Synchronisierung über Datenbankabfragen oder Aktualisierungstransaktionen integriert. So müssen Serverinstanzen nicht für jede einzelne Instanz eine eigene Synchronisierung durchführen.

Der Sun IMAP-Server ist ein Beispiel für einen Dienst, bei dem es sich nicht um einen Scalable-Dienst handelt. Der Dienst aktualisiert einen Speicher, aber dieser Speicher ist privat. Wenn mehrere IMAP-Instanzen in diesen Speicher schreiben, werden die jeweiligen Einträge überschrieben, weil die Aktualisierungen nicht synchronisiert sind. Der IMAP-Server muss für die Synchronisierung eines gleichzeitigen Zugriffs umgeschrieben werden.

- Beachten Sie schließlich, dass Instanzen über private Daten verfügen können, die von den Daten anderer Instanzen getrennt sind. In einem solchen Fall muss der Dienst den gleichzeitigen Zugriff nicht selbst synchronisieren, weil die Daten privat sind und nur diese Instanz sie bearbeiten kann. In diesem Fall müssen Sie darauf achten, diese privaten Daten nicht unter dem Cluster-Dateisystem zu speichern, weil die Möglichkeit eines globalen Zugriffs besteht.

Datendienst-API und API der Datendienst-Entwicklungsbibliothek-API

Das Sun Cluster-System bietet folgende Informationen, um Anwendungen hochverfügbar zu machen:

- Datendienste, die als Bestandteil des Sun Cluster-Systems bereitgestellt werden.
- Eine Datendienst-API,
- Eine Entwicklungsbibliotheks-API für Datendienste
- Einen "generischen" Datendienst.

Im *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* wird die Installation und Konfiguration der Datendienste beschrieben, die im Lieferumfang des Sun Cluster-Systems enthalten sind. Im *Sun Cluster 3.1 9/04 Software Collection for Solaris OS (SPARC Platform Edition)* wird beschrieben, wie andere Anwendungen für die Hochverfügbarkeit im Sun Cluster-Framework konfiguriert werden.

Mit den Sun Cluster-APIs können Anwendungsentwickler Fehler-Monitore sowie Skripts zum Starten und Anhalten von Datendienstinstanzen entwickeln. Mit diesen Tools kann eine Anwendung als Failover- oder Scalable-Datendienst eingerichtet werden. Das Sun Cluster-System stellt einen "generischen" Datendienst zur

Verfügung. Verwenden Sie diesen generischen Datendienst, um schnell die erforderlichen Start- und Stoppmethoden zu generieren und den Datendienst als Failover- oder Scalable-Dienst zu implementieren.

Verwenden des Cluster-Interconnect für den Datendienstverkehr

Ein Cluster benötigt mehrere Netzwerkverbindungen zwischen den Knoten, die den Cluster-Interconnect bilden. Sun Cluster verwendet mehrere Interconnects, um folgende Ziele zu erreichen:

- Gewährleisten von Hochverfügbarkeit
- Verbessern der Leistung

Für den internen Datenverkehr, wie beispielsweise Dateisystemdaten oder Scalable-Dienstdaten, werden Meldungen über alle verfügbaren Interconnects im Round-Robin-Verfahren gesendet. Der Cluster-Interconnect steht auch Anwendungen zur hoch verfügbaren Kommunikation zwischen den Knoten zur Verfügung. Eine verteilte Anwendung kann zum Beispiel Komponenten auf unterschiedlichen Knoten ausführen, die miteinander kommunizieren müssen. Indem der Cluster-Interconnect anstelle des öffentlichen Netzwerks verwendet wird, bleiben diese Verbindungen beim Versagen einer einzelnen Verknüpfung bestehen.

Um den Cluster-Interconnect für die Datenverbindung zwischen den Knoten zu verwenden, muss eine Anwendung die während der Sun Cluster-Installation konfigurierten privaten Hostnamen verwenden. Wenn zum Beispiel der private Hostname für `node 1` `clusternode1-priv` ist, sollten Sie diesen Namen verwenden, um mit `node 1` für die Kommunikation über den Cluster-Interconnect zu kommunizieren. TCP-Sockets, die mit diesem Namen geöffnet wurden, werden über den Cluster-Interconnect geleitet und können bei einem Netzwerkfehler transparent umgeleitet werden.

Da die privaten Hostnamen während der Sun Cluster-Installation konfiguriert werden können, kann der Cluster-Interconnect jeden beliebigen, während der Installation gewählten Namen verwenden. Den tatsächlichen Namen erhalten Sie über den `scha_cluster_get(3HA)`-Befehl mit dem Argument `scha_privatelink_hostname_node`.

Sowohl die Anwendungskommunikation als auch die Cluster-interne Kommunikation werden über alle Interconnects verteilt. Da die Anwendungen den Cluster-Interconnect mit dem internen Cluster-Datenverkehr teilen, hängt die für die Anwendung verfügbare Bandbreite von der Bandbreite ab, die für anderen Cluster-Datenverkehr verwendet wird. Bei einem Versagen werden der gesamte interne Datenverkehr und der Anwendungsdatenverkehr über alle verfügbaren Interconnects verteilt.

Jedem Knoten wird außerdem eine feste Pro-Knoten-Adresse zugewiesen. Diese Pro-Knoten-Adresse ist fest mit dem `clprivnet`-Treiber verknüpft. Die IP-Adresse ist dem privaten Hostnamen für den Knoten zugeordnet: `clusternode1-priv`. Informationen zum Sun Cluster-Treiber für private Netzwerke finden Sie in der Online-Dokumentation `clprivnet(7)`.

Wenn die Anwendung an allen Stellen konsistente IP-Adressen benötigt, konfigurieren Sie die Anwendung so, dass die Pro-Knoten-Adresse sowohl an den Client als auch an den Server angebunden ist. In diesem Fall scheinen alle Verbindungen von der Pro-Knoten-Adresse zu stammen und zu dieser zurückzukehren.

Ressourcen, Ressourcengruppen und Ressourcentypen

Datendienste verwenden mehrere Typen von *Ressourcen*: Anwendungen wie Sun Java System Web Server oder Apache Web Server verwenden Netzwerkadressen (logische Hostnamen und gemeinsam genutzte Adressen), von denen die Anwendungen abhängen. Anwendung und Netzwerkressourcen bilden eine grundlegende Einheit, die vom RGM verwaltet wird.

Datendienste sind Ressourcentypen. Sun Cluster HA für Oracle ist der Ressourcentyp `SUNW.oracle-server` und Sun Cluster HA für Apache ist der Ressourcentyp `SUNW.apache`.

Eine Ressource ist eine Instanz eines *Ressourcentyps*, der Cluster-weit definiert ist. Es sind mehrere Ressourcentypen definiert.

Netzwerkressourcen sind entweder `SUNW.LogicalHostname`- oder `SUNW.SharedAddress`-Ressourcentypen. Diese beiden Ressourcentypen sind von der Sun Cluster-Software vorregistriert.

Die Ressourcentypen `HASStorage` und `HASStoragePlus` werden zur Synchronisierung beim Starten von Ressourcen und Plattengerätegruppen verwendet, von denen die Ressourcen abhängen. Durch diese Ressourcentypen wird sichergestellt, dass die Pfade zu Einhängenpunkten im Cluster-Dateisystem, globalen Geräten und Gerätegruppennamen verfügbar sind, bevor ein Datendienst startet. Weitere Informationen finden Sie unter "Synchronisieren der Startvorgänge zwischen Ressourcengruppen und Plattengerätegruppe" im *Data Services Installation and Configuration Guide*. Der `HASStoragePlus`-Ressourcentyp ist ab Sun Cluster 3.0 5/02 verfügbar und bietet eine weitere Funktion, um lokale Dateisysteme hoch verfügbar zu machen. Weitere Informationen zu dieser Funktion finden Sie unter „`HASStoragePlus` Ressourcentyp“ auf Seite 50.

RGM-verwaltete Ressourcen werden in so genannten *Ressourcengruppen* zusammengefasst, so dass sie als Einheit verwaltet werden können. Eine Ressourcengruppe migriert als Einheit, wenn ein Failover oder ein Switchover für die Ressourcengruppe eingeleitet wird.

Hinweis – Wenn Sie eine Ressourcengruppe mit Anwendungsressourcen online bringen, wird die Anwendung gestartet. Die Datendienst-Startmethode wartet, bis die Anwendung aktiv ist und ausgeführt wird, bevor sie erfolgreich beendet wird. Die Feststellung, wann die Anwendung aktiv ist und ausgeführt wird, erfolgt auf dieselbe Weise, in der Datendienst-Fehler-Monitor feststellt, ob ein Datendienst Clients bedient. Im *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* erhalten Sie weitere Informationen über diesen Vorgang.

Ressourcengruppen-Manager (RGM)

Der RGM steuert Datendienste (Anwendungen) als Ressourcen, die mit *Ressourcentypen*-Implementierungen verwaltet werden. Diese Implementierungen werden entweder von Sun geliefert oder von einem Entwickler mithilfe einer generischen Datendienstvorlage, der API für die Datendienst-Entwicklungsbibliothek (DSDL-API) oder der Ressourcenverwaltungs-API (RMAPI), erstellt. Der Cluster-Verwalter erstellt und verwaltet Ressourcen in Containern, die als *Ressourcengruppen* bezeichnet werden. Der RGM stoppt und startet die Ressourcengruppen auf den ausgewählten Knoten je nach den Änderungen in der Cluster-Mitgliedschaft.

Der RGM verwaltet *Ressourcen* und *Ressourcengruppen*. Durch RGM-Aktionen wechseln Ressourcen und Ressourcengruppen zwischen dem Online- und Offline-Zustand. Eine umfassende Beschreibung dieser Zustände und der Einstellungen für Ressourcen und Ressourcengruppen finden Sie im Abschnitt [„Zustände und Einstellungen für Ressourcen und Ressourcengruppen“](#) auf Seite 77.

Informationen über das Starten von Solaris-Projekten unter der Steuerung des RGM finden Sie unter [„Datendienst-Projektkonfiguration“](#) auf Seite 791.

Zustände und Einstellungen für Ressourcen und Ressourcengruppen

Ein Verwalter verwendet statische Einstellungen für Ressourcen und Ressourcengruppen. Diese Einstellungen können nur über Verwaltungsaktionen geändert werden. Der RGM ändert die dynamischen "Zustände" von Ressourcengruppen. Diese Einstellungen und Zustände sind in der nachstehenden Liste beschrieben.

- **Verwaltet oder unverwaltet** – Diese Cluster-weiten Einstellungen gelten nur für Ressourcengruppen. Ressourcengruppen werden vom RGM verwaltet. Mit dem `scrgadm (1M)`-Befehl wird die Verwaltung einer Ressourcengruppe durch den RGM aktiviert bzw. deaktiviert. Diese Einstellungen werden bei einer Cluster-Neukonfiguration nicht geändert.

Wenn eine Ressourcengruppe zum ersten Mal erstellt wird, ist sie unverwaltet. Eine Ressourcengruppe muss verwaltet werden, bevor enthaltene Ressourcen aktiv werden können.

Bei einigen Datendiensten, zum Beispiel einem Scalable-Webserver, muss diese Arbeit vor dem Starten und nach dem Anhalten der Netzwerkressourcen erledigt werden. Dies wird durch die Datendienstmethoden zum Initialisieren (INIT) und Beenden (FINI) erledigt. Die INIT-Methoden werden nur ausgeführt, wenn die Ressourcengruppe, zu der die Ressource gehört, den Zustand "Verwaltet" hat.

Wenn der Status einer Ressourcengruppe von unverwaltet auf verwaltet gesetzt wird, werden alle registrierten INIT-Methoden für die in der Gruppe enthaltenen Ressourcen ausgeführt.

Wenn der Status einer Ressourcengruppe von verwaltet auf unverwaltet gesetzt wird, werden alle registrierten FINI-Methoden aufgerufen, um eine Bereinigung durchzuführen.

Am häufigsten werden INIT- und FINI-Methoden für Netzwerkressourcen für Scalable-Dienste eingesetzt. Sie können jedoch für alle beliebigen Initialisierungs- oder Bereinigungsaufgaben eingesetzt werden, die nicht von der Anwendung erledigt werden.

- **Aktiviert oder deaktiviert** – Diese Cluster-weiten Einstellungen gelten für Ressourcen. Der `scrgadm (1M)`-Befehl kann zur Aktivierung oder Deaktivierung einer Ressource verwendet werden. Diese Einstellungen werden bei einer Cluster-Neukonfiguration nicht geändert.

Bei einer normalen Einstellung ist eine Ressource aktiviert und wird aktiv im System ausgeführt.

Wenn die Ressource auf keinem Cluster-Knoten verfügbar sein soll, deaktivieren Sie die Ressource. Eine deaktivierte Ressource steht nicht zur allgemeinen Verwendung zur Verfügung.

- **Online oder offline** – Diese dynamischen Zustände gelten sowohl für Ressourcen als auch für Ressourcengruppen.

Die Zustände "online" und "offline" ändern sich bei den Cluster-Übergängen im Verlauf der Cluster-Rekonfigurationsschritte bei einem Switchover oder Failover. Außerdem können diese Zustände über Verwaltungsaktionen geändert werden. Mit dem Befehl `scswitch (1M)` können Sie den Online- bzw. Offline-Zustand einer Ressource oder Ressourcengruppe ändern.

Eine Failover-Ressource oder -Ressourcengruppe kann nur auf jeweils einem Knoten gleichzeitig online sein. Eine Scalable-Ressource oder -Ressourcengruppe kann gleichzeitig auf mehreren Knoten online und auf anderen offline sein. Während eines Switchovers oder Failovers werden Ressourcengruppen mit den darin enthaltenen Ressourcen auf einem Knoten offline genommen und auf einem anderen Knoten online gebracht.

Wenn eine Ressourcengruppe offline ist, sind auch alle ihre Ressourcen offline.
Wenn eine Ressourcengruppe online ist, sind auch alle ihre aktivierten Ressourcen online.

Ressourcengruppen können mehrere Ressourcen mit Abhängigkeiten zwischen Ressourcen enthalten. Aufgrund dieser Abhängigkeiten müssen die Ressourcen in einer bestimmten Reihenfolge online gebracht und offline genommen werden. Die Methoden, um die Ressourcen online zu bringen und offline zu nehmen, können für jede Ressource unterschiedlich lange dauern. Die Zustände online und offline können für Ressourcen einer einzigen Ressourcengruppe aufgrund von Ressourcenabhängigkeiten und Unterschieden bei Start- und Stopzeit bei einer Cluster-Rekonfiguration unterschiedlich sein.

Ressourcen- und Ressourcengruppeneigenschaften

Sie können Eigenschaftswerte für die Ressourcen und Ressourcengruppen Ihrer Sun Cluster-Datendienste konfigurieren. Standardeigenschaften sind für alle Datendienste gleich. Erweiterungseigenschaften sind für jeden Datendienst spezifisch. Einige Standard- und Erweiterungseigenschaften sind mit Standardeinstellungen konfiguriert, so dass Sie diese nicht ändern müssen. Andere müssen im Rahmen der Erstellung und Konfiguration von Ressourcen eingestellt werden. Die Dokumentation für jeden Datendienst gibt an, welche Ressourceneigenschaften eingestellt werden können und wie diese einzustellen sind.

Die Standardeigenschaften werden zur Konfiguration von Ressourcen- und Ressourcengruppeneigenschaften eingesetzt, die in der Regel von keinem bestimmten Datendienst abhängen. Die Standardeigenschaften werden in Anhang A, „Standard Properties“ in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* aufgeführt.

Die RGM-Erweiterungseigenschaften liefern Informationen wie zum Beispiel den Speicherort von binären Anwendungsdateien und Konfigurationsdateien. Sie ändern die Erweiterungseigenschaften beim Konfigurieren der Datendienste. Die Erweiterungseigenschaften werden im Handbuch zum Datendienst im beschrieben.

Datendienst-Projektkonfiguration

Datendienste können für einen Start mit einem Solaris-Projektnamen konfiguriert werden, wenn sie mit dem RGM online gebracht werden. Die Konfiguration ordnet einer vom RGM verwalteten Ressource oder Ressourcengruppe eine Solaris Projekt-ID zu. Durch die Zuordnung der Ressource oder Ressourcengruppe zu einer Projekt-ID können Sie anspruchsvolle Steuerfunktionen, die im Solaris-Betriebssystem verfügbar sind, für die Verwaltung von Arbeitslasten und Verbrauch in Ihrem Cluster einsetzen.

Hinweis – Diese Konfiguration können Sie nur dann durchführen, wenn Sie mit der aktuellen Version der Sun Cluster-Software unter Solaris 9 oder höher arbeiten.

Mithilfe der Solaris-Verwaltungsfunktion in einer Sun Cluster-Umgebung können Sie sicherstellen, dass die wichtigsten Anwendungen Priorität erhalten, wenn sie einen Knoten mit anderen Anwendungen teilen. Anwendungen können einen Knoten teilen, wenn Sie mit konsolidierten Diensten arbeiten oder weil Anwendungen bei einem Failover den Knoten gewechselt haben. Die Verwendung der hier beschriebenen Verwaltungsfunktion kann die Verfügbarkeit einer kritischen Anwendung verbessern, indem andere Anwendungen mit niedrigerer Priorität daran gehindert werden, zu viele Systemleistungen wie CPU-Zeit in Anspruch zu nehmen.

Hinweis – In der Solaris-Dokumentation zu dieser Funktion werden CPU-Zeit, Prozesse, Aufgaben und vergleichbare Komponenten als "Ressourcen" beschrieben. In der Sun Cluster-Dokumentation hingegen wird der Begriff "Ressourcen" zur Beschreibung von Entitäten verwendet, die vom RGM gesteuert werden. Im nachstehenden Abschnitt wird der Begriff "Ressource" für Sun Cluster-Entitäten verwendet, die vom RGM gesteuert werden. In diesem Abschnitt wird der Begriff "Leistungen" für CPU-Zeit, Prozesse und Aufgaben verwendet.

Dieser Abschnitt enthält eine konzeptuelle Beschreibung zur Konfiguration von Datendiensten, um die Prozesse in einem gegebenen Solaris 9 `project(4)` zu starten. In diesem Abschnitt werden auch mehrere Failover-Szenarien beschrieben. Des Weiteren enthält er Empfehlungen zur Planung, wie die im Solaris-Betriebssystem zur Verfügung gestellte Verwaltungsfunktion eingesetzt werden kann.

Detaillierte konzeptionelle und verfahrenstechnische Dokumentation zur Verwaltungsfunktion finden Sie im Kapitel 1, „Network Service (Overview)“ in *System Administration Guide: Network Services*.

Solaris-Verwaltungsfunktion in einem Cluster können Sie folgenden Prozess auf höchster Ebene einsetzen:

1. Konfigurieren von Anwendungen als Teil der Ressource.
2. Konfigurieren von Ressourcen als Teil einer Ressourcengruppe.
3. Aktivieren von Ressourcen in der Ressourcengruppe.
4. Aktivieren der Verwaltung der Ressourcengruppe.
5. Erstellen eines Solaris-Projekts für die Ressourcengruppe.
6. Konfigurieren von Standardeigenschaften, um den Ressourcengruppenamen dem unter Punkt 5 erstellten Projekt zuzuordnen.
7. Online-bringen der Ressourcengruppe.

Verwenden Sie zur Konfiguration der Standardeigenschaften `Resource_project_name` oder `RG_project_name` die `-y`-Option mit dem `scrgadm(1M)`-Befehl, um die Solaris-Projekt-ID der Ressource oder Ressourcengruppe zuzuordnen. Stellen Sie die Eigenschaftswerte für die Ressource oder die Ressourcengruppe ein. Die Definitionen der Eigenschaften finden Sie in Anhang A, „Standard Properties“ in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*. Eigenschaftsbeschreibungen finden Sie unter `r_properties(5)` und `rg_properties(5)`.

Der angegebene Projektname muss in der Projektdatenbank (`/etc/project`) vorhanden sein und der Root-Benutzer muss als Mitglied des genannten Projekts konfiguriert sein. Konzeptionelle Informationen zur Projektname-Datenbank finden Sie in Kapitel 2, „Projects and Tasks (Overview)“ in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*. Eine Beschreibung der Syntax der Projektdatei finden Sie unter `project(4)`.

Wenn der RGM Ressourcen oder Ressourcengruppen online bringt, startet er die verknüpften Prozesse unter dem Projektname.

Hinweis – Benutzer können die Ressource oder Ressourcengruppe jederzeit einem Projekt zuordnen. Der neue Projektname ist jedoch erst gültig, wenn die Ressource oder Ressourcengruppe mit dem RGM offline genommen und wieder online gebracht wurde.

Durch das Starten von Ressourcen und Ressourcengruppen unter dem Projektname können Sie die folgenden Funktionen zur Verwaltung der Systemleistungen in Ihrem Cluster konfigurieren.

- **Erweiterte Berechnung** – Ein flexibler Weg zur Aufzeichnung des Verbrauchs auf Aufgaben- oder Prozessbasis. Mit dieser Funktion können Sie die historische Nutzung nachverfolgen und den Leistungsbedarf für zukünftige Arbeitslasten beurteilen.
- **Steuerungen** – Ein Mechanismus zur Einschränkung von Systemleistungen. Damit kann verhindert werden, dass Prozesse, Aufgaben und Projekte große Anteile von bestimmten Systemleistungen verbrauchen.
- **Faire Nutzungsplanung (FSS)** – Damit besteht die Möglichkeit, die Zuweisung der verfügbaren CPU-Zeit auf die Arbeitslasten je nach ihrer Bedeutung zu steuern. Die Bedeutung der Arbeitslast wird durch die Anteile an CPU-Zeit ausgedrückt, die Sie jeder Arbeitslast zuweisen. Weitere Informationen erhalten Sie auf den folgenden Man Pages:
 - `dispadmin(1M)`
 - `prionctl(1)`
 - `ps(1)`
 - `FSS(7)`

- Pools – Damit besteht die Möglichkeit, Partitionen je nach Anwendungsanforderungen für interaktive Anwendungen zu nutzen. Pools können zur Partitionierung eines Servers eingesetzt werden, der eine Reihe unterschiedlicher Softwareanwendungen unterstützt. Die Verwendung von Pools führt zu einer besser vorhersehbaren Antwort für jede Anwendung.

Bestimmen der Anforderungen der Projektkonfiguration

Bevor Sie die Datendienste zur Verwendung der von Solaris zur Verfügung gestellten Steuermöglichkeiten in einer Sun Cluster-Umgebung konfigurieren, müssen Sie entscheiden, wie Sie die Ressourcen bei Switchover- oder Failover-Vorgängen steuern und verfolgen möchten. Identifizieren Sie vor der Konfiguration eines neuen Projekts die Abhängigkeiten innerhalb des Clusters. Ressourcen und Ressourcengruppen hängen zum Beispiel von Plattengerätegruppen ab.

Verwenden Sie die Ressourcengruppeneigenschaften `nodelist`, `failback`, `maximum primaries` und `desired primaries`, die mit `scrgadm(1M)` konfiguriert sind, um die Prioritäten in der Knotenliste für Ihre Ressourcengruppe zu identifizieren.

- Eine kurze Erläuterung der Abhängigkeiten in der Knotenliste zwischen Ressourcengruppen und Plattengerätegruppen finden Sie unter „Relationship Between Resource Groups and Disk Device Groups“ in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.
- Detaillierte Beschreibungen der Eigenschaften finden Sie unter `rg_properties(5)`.

Legen Sie die Prioritäten der Knotenliste für die Plattengerätegruppe mit den Eigenschaften `preferenced` und `failback` fest, die mit `scrgadm(1M)` und `scsetup(1M)` konfiguriert werden.

- Konzeptionelle Informationen zur Eigenschaft `preferenced` finden Sie unter „Multiport-Plattengerätegruppen“ auf Seite 45.
- Verfahrenstechnische Informationen finden Sie unter "How To Change Disk Device Properties" in „Verwalten von Plattengerätegruppen“ in *Sun Cluster Handbuch Systemverwaltung für Solaris OS*.
- Konzeptionelle Informationen zur Knotenkonfiguration und dem Verhalten von Failover- und Scalable-Datendiensten finden Sie unter „Hardware- und Softwarekomponenten des Sun Cluster-Systems“ auf Seite 21.

Wenn Sie alle Cluster-Knoten gleich konfigurieren, werden die Nutzungsgrenzen auf Primär- und Sekundärknoten identisch durchgesetzt. Die Konfigurationsparameter der Projekte müssen nicht für alle Anwendungen in den Konfigurationsdateien auf allen Knoten gleich sein. Auf alle der Anwendung zugeordneten Projekte muss zumindest die Projektdatenbank auf allen potenziellen Mastern der Anwendung

zugreifen können. Angenommen, der Master für Anwendung 1 ist *phys-schost-1* kann aber durch ein Switchover oder Failover auch *phys-schost-2* oder *phys-schost-3* sein. Auf allen drei Knoten (*phys-schost-1*, *phys-schost-2* und *phys-schost-3*) muss Zugriff auf das der Anwendung 1 zugeordnete Projekt bestehen.

Hinweis – Die Projektdatenbank-Informationen können sich in einer lokalen `/etc/project`-Datenbankdatei befinden oder in der NIS-Karte oder im LDAP-Verzeichnisdienst gespeichert sein.

Die Solaris-Umgebung ermöglicht eine flexible Konfiguration der Nutzungsparameter und Sun Cluster erlegt nur wenige Einschränkungen auf. Die Konfigurationsauswahl hängt von den Standortbedürfnissen ab. Beachten Sie die allgemeinen Richtlinien in den nachfolgenden Abschnitten, bevor Sie Ihre Systeme konfigurieren.

Einstellen virtueller Speicherbegrenzungen für Prozesse

Stellen Sie die `process.max-address-space`-Steuerung ein, um den virtuellen Speicher auf Prozessbasis zu begrenzen. Unter `rctladm(1M)` finden Sie detaillierte Informationen zur Einstellung des `process.max-address-space`-Wertes.

Konfigurieren Sie bei Verwendung von Verwaltungssteuerungen mit der Sun Cluster-Software die Speicherbegrenzung angemessen, um unnötige Failover-Vorgänge von Anwendungen und einen "Ping-Pong"-Effekt bei den Anwendungen zu vermeiden. Im Allgemeinen sollten Sie sich an folgende Richtlinien halten:

- Stellen Sie die Speicherbegrenzungen nicht zu niedrig ein.
Wenn eine Anwendung die Speicherbegrenzung erreicht, kann ein Failover erfolgen. Diese Richtlinie ist besonders für Datenbankanwendungen von Bedeutung, wo das Erreichen einer virtuellen Speicherbegrenzung unerwartete Folgen haben kann.
- Stellen Sie die Speicherbegrenzungen für Primär- und Sekundärknoten nicht identisch ein.
Identische Begrenzungen können einen Ping-Pong-Effekt auslösen, wenn eine Anwendung die Speichergrenze erreicht und ein Failover auf einen Sekundärknoten mit einer identischen Speicherbegrenzung erfolgt. Stellen Sie die Speicherbegrenzung auf dem Sekundärknoten etwas höher ein. Network MultipathingDer Unterschied bei den Speicherbegrenzungen trägt zur Vermeidung des Ping-Pong-Szenarios bei und gibt dem Systemverwalter Zeit, die Parameter nach Bedarf zu korrigieren.
- Verwenden Sie die Speicherbegrenzung der Ressourcenverwaltung für den Lastausgleich.

Sie können zum Beispiel eine fehlerhaft arbeitende Anwendung mit Speicherbegrenzungen davon abhalten, zu viel Auslagerungsspeicher zu belegen.

Failover-Szenarien

Sie können die Verwaltungsparameter so konfigurieren, dass die Zuweisung in der Projektkonfiguration (`/etc/project`) im normalen Cluster-Betrieb und in Switchover- und Failover-Situationen funktioniert.

Die nachfolgenden Abschnitte sind als Beispiel gedachte Szenarien.

- In den ersten beiden Abschnitte, "Zwei-Knoten-Cluster mit zwei Anwendungen" und "Zwei-Knoten-Cluster mit drei Anwendungen," werden Failover-Szenarien für ganze Knoten beschrieben.
- Der Abschnitt "Failover der Ressourcengruppe" illustriert den Failover-Vorgang einer einzigen Anwendung.

In einer Sun Cluster-Umgebung konfigurieren Sie Anwendungen als Teil der Ressource. Anschließend konfigurieren Sie eine Ressource als Teil einer Ressourcengruppe (RG). Wenn ein Fehler auftritt, wechselt die Ressourcengruppe zusammen mit den zugeordneten Anwendungen auf einen anderen Knoten. In den nachstehenden Beispielen sind die Ressourcen nicht explizit angegeben. Angenommen, jede Ressource umfasst nur eine Anwendung.

Hinweis – Ein Failover erfolgt in der bevorzugten Reihenfolge, wie sie in der Knotenliste im RGM angegeben ist.

Für die nachstehenden Beispiele gelten folgende Beschränkungen:

- Anwendung 1 (Anw-1) ist in der Ressourcengruppe RG-1 konfiguriert.
- Anwendung 2 (Anw-2) ist in der Ressourcengruppe RG-2 konfiguriert.
- Anwendung 3 (Anw-3) ist in der Ressourcengruppe RG-3 konfiguriert.

Obwohl die Anzahl der zugeordneten Anteile gleich bleibt, ändert sich der jeder Anwendung zugewiesene Prozentsatz an CPU-Zeit nach dem Failover. Dieser Prozentsatz hängt von der Anzahl der auf dem Knoten ausgeführten Anwendungen sowie von der Anzahl an Anteilen ab, die jeder aktiven Anwendung zugeordnet sind.

In diesen Szenarien werden folgende Konfigurationen vorausgesetzt.

- Alle Anwendungen sind unter einem gemeinsamen Projekt konfiguriert.
- Jede Ressource enthält nur eine einzige Anwendung.
- Die Anwendungen sind die einzigen aktiven Prozesse auf den Knoten.
- Die Projektdatenbanken sind auf allen Knoten des Clusters gleich konfiguriert.

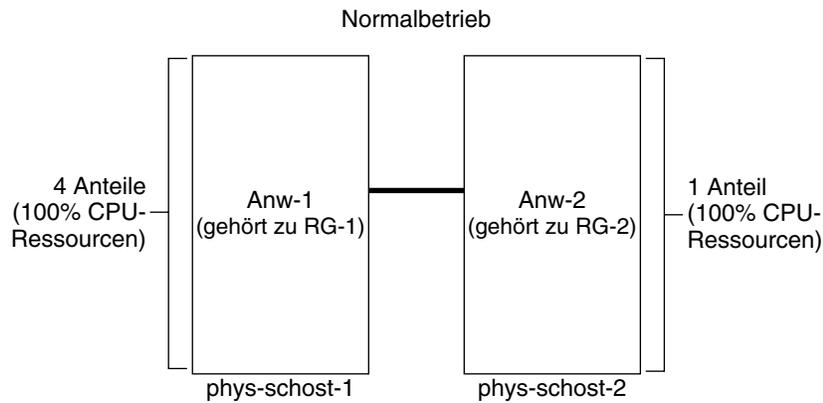
Zwei-Knoten-Cluster mit zwei Anwendungen

Auf einem Zwei-Cluster-Knoten können Sie zwei Anwendungen konfigurieren, um sicherzustellen, dass jeder reale Host (*phys-schost-1*, *phys-schost-2*) als Standardmaster für eine Anwendung fungiert. Jeder reale Host fungiert als Sekundärknoten für den anderen realen Host. Alle Projekte, die Anwendung 1 und Anwendung 2 zugeordnet sind, müssen in den Datenbankdateien der Projekte auf beiden Knoten vertreten sein. Im normalen Cluster-Betrieb wird jede Anwendung auf dem eigenen Standard-Master ausgeführt und erhält über die Verwaltungsfunktion die gesamte CPU-Zeit zugewiesen.

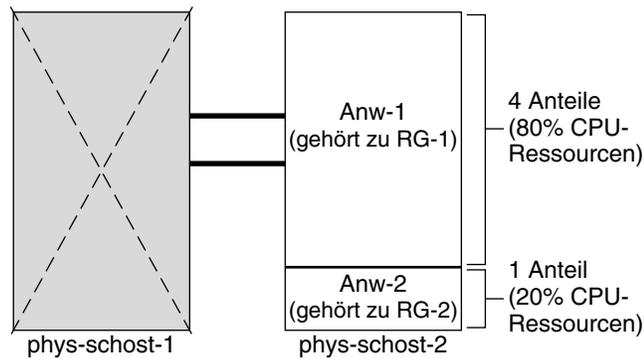
Nach einem Failover oder Switchover werden beide Anwendungen auf einem einzigen Knoten ausgeführt und erhalten die Anteile zugewiesen, die in der Konfigurationsdatei angegeben wurden. Dieser Eintrag in der `/etc/project`-Datei gibt zum Beispiel an, dass Anwendung 1 vier Anteile zugewiesen werden und Anwendung 2 ein Anteil zugewiesen wird.

```
Prj_1:100:project for App-1:root::project.cpu-shares=(privileged,4,none)
Prj_2:101:project for App-2:root::project.cpu-shares=(privileged,1,none)
```

Die nachstehende Abbildung illustriert den normalen und den Failover-Betrieb in dieser Konfiguration. Die Anzahl der zugewiesenen Anteile ändert sich nicht. Der Prozentsatz an CPU-Zeit, der jeder Anwendung zur Verfügung steht, kann sich jedoch ändern. Der Prozentsatz hängt von der Anzahl der Anteile ab, die jedem Prozess mit CPU-Zeit-Bedarf zugewiesen sind.



Failover-Betrieb: Ausfall von Knoten phys-schost-1



Zwei-Knoten-Cluster mit drei Anwendungen

Auf einem Zwei-Knoten-Cluster mit drei Anwendungen können Sie einen realen Host (*phys-schost-1*) als Standard-Master für eine Anwendung konfigurieren. Den zweiten realen Host (*phys-schost-2*) können Sie als Standard-Master für die restlichen beiden Anwendungen konfigurieren. Im folgenden Beispiel wird davon ausgegangen, dass die Projekt-Datenbankdatei auf jedem Knoten vorhanden ist. Die Projekt-Datenbankdatei wird bei einem Failover oder Switchover nicht geändert.

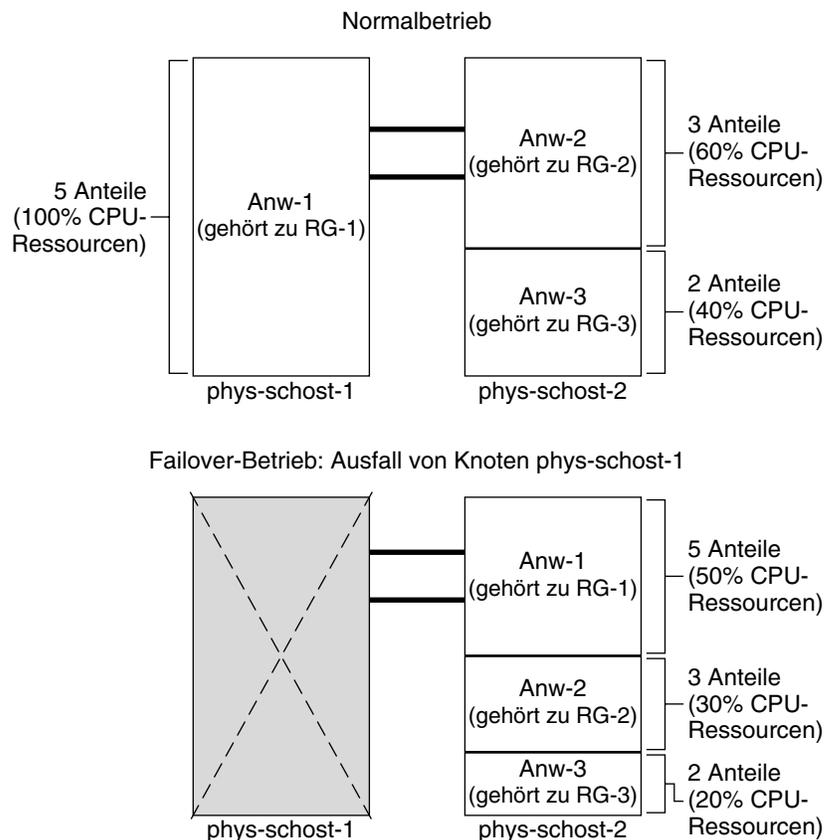
```
Prj_1:103:project for App_1:root::project.cpu-shares=(privileged,5,none)
Prj_2:104:project for App_2:root::project.cpu-shares=(privileged,3,none)
Prj_3:105:project for App_3:root::project.cpu-shares=(privileged,2,none)
```

Im normalen Cluster-Betrieb sind der Anwendung 1 fünf Anteile auf dem entsprechenden Standard-Master *phys-schost-1* zugewiesen. Dies entspricht 100 Prozent der CPU-Zeit, weil es die einzige Anwendung ist, die CPU-Zeit auf diesem Knoten benötigt. Den Anwendungen 2 und 3 sind jeweils drei und zwei Anteile auf dem dazugehörigen Standard-Master *phys-schost-2* zugewiesen. Damit entfällt auf Anwendung 2 60 Prozent der CPU-Zeit und auf Anwendung 3 40 Prozent der CPU-Zeit im normalen Betrieb.

Wenn Anwendung 1 bei einem Failover oder Switchover auf *phys-schost-2* wechselt, bleiben die Anteile für alle drei Anwendungen gleich. Aber der Prozentsatz der CPU-Ressourcen wird anhand der Projekt-Datenbankdatei neu zugewiesen.

- Anwendung 1 erhält mit 5 Anteilen 50 Prozent der CPU.
- Anwendung 2 erhält mit 3 Anteilen 30 Prozent der CPU.
- Anwendung 3 erhält mit 2 Anteilen 20 Prozent der CPU.

Die nachstehende Abbildung illustriert den normalen und den Failover-Betrieb in dieser Konfiguration.



Failover der Ressourcengruppe

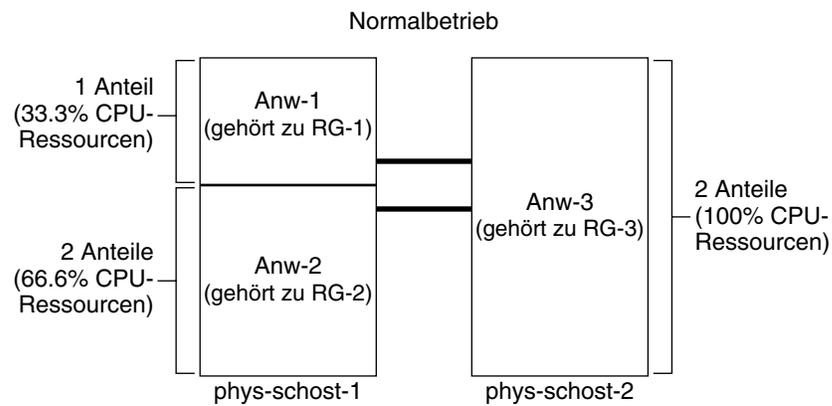
In einer Konfiguration, in der mehrere Ressourcengruppen denselben Standard-Master haben, kann eine Ressourcengruppe (und die zugeordneten Anwendungen) bei einem Failover oder Switchover auf einen Sekundärknoten wechseln. Dabei läuft der Standard-Master im Cluster weiter.

Hinweis – Die Anwendung, die den Failover-Vorgang durchführt, erhält dabei die Ressourcen auf dem Sekundärknoten zugewiesen, die in der Konfigurationsdatei angegeben sind. In diesem Beispiel haben die Projekt-Datenbankdateien auf Primär- und Sekundärknoten dieselbe Konfiguration.

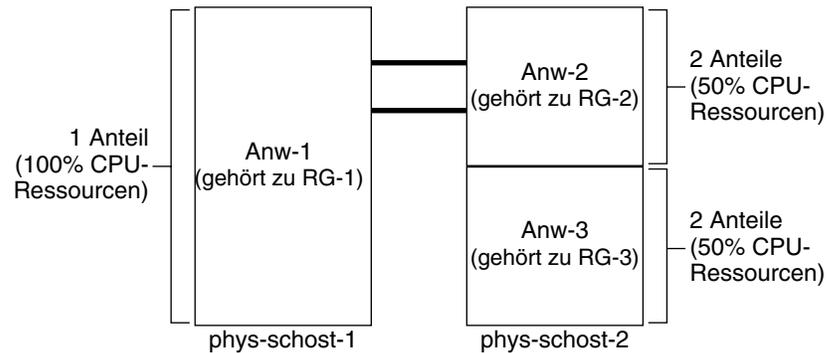
Diese Beispielskonfigurationsdatei gibt an, dass Anwendung 1 ein Anteil, Anwendung 2 zwei Anteile und Anwendung 3 zwei Anteile zugewiesen werden.

```
Prj_1:106:project for App_1:root::project.cpu-shares=(privileged,1,none)
Prj_2:107:project for App_2:root::project.cpu-shares=(privileged,2,none)
Prj_3:108:project for App_3:root::project.cpu-shares=(privileged,2,none)
```

Die nachstehende Abbildung illustriert den normalen und den Failover-Betrieb in dieser Konfiguration, wobei RG-2 mit Anwendung 2 zu *phys-schost-2* wechselt. Beachten Sie, dass sich die Anzahl der zugewiesenen Anteile nicht ändert. Der Prozentsatz an CPU-Zeit, der jeder Anwendung zur Verfügung steht, kann sich jedoch je nach Anzahl der Anteile ändern, die jeder Anwendung mit CPU-Zeit-Bedarf zugewiesen werden.



Failover-Betrieb: Failover von RG-2 auf phys-schost-2



Öffentliche Netzwerkadapter und Internet Protocol (IP) Network Multipathing

Clients führen über das öffentliche Netzwerk Datenanforderungen an den Cluster durch. Jeder Cluster-Knoten ist über ein Paar öffentlicher Netzwerkadapter mit mindestens einem öffentlichen Netzwerk verbunden.

Die Solaris IPMP-Software (Internet Protocol Network Multipathing) von Sun Cluster liefert den Basismechanismus für die Überwachung öffentlicher Netzwerkadapter und das Wechseln von IP-Adressen von einem Adapter auf einen anderen, wenn ein Fehler erkannt wird. Jeder Cluster-Knoten hat seine eigene Internet Protocol (IP) Network Multipathing-Konfiguration, die sich von der auf anderen Cluster-Knoten unterscheiden kann.

Öffentliche Netzwerkadapter sind als *IPMP-Gruppen* (Multipathing-Gruppen) organisiert. Jede Multipathing-Gruppe hat mindestens einen öffentlichen Netzwerkadapter. Jeder Adapter in einer Multipathing-Gruppe kann aktiv sein. Alternativ können Sie Standby-Schnittstellen konfigurieren, die bis zum Auftreten eines Failovers inaktiv bleiben.

Der `in.mpathd`-Multipathing-Dämon verwendet eine IP-Testadresse, um Fehler und Reparaturen zu erkennen. Erkennt der Multipathing-Dämon einen Fehler auf einem der Adapter, erfolgt ein Failover. Der gesamte Netzwerkzugriff geht vom fehlerhaften Adapter auf einen anderen funktionsfähigen Adapter in der Multipathing-Gruppe über. Daher bleibt die Konnektivität des öffentlichen Netzwerkes für den Knoten erhalten. Wenn eine Standby-Schnittstelle konfiguriert wurde, wählt der Dämon die Standby-Schnittstelle aus. Anderenfalls wählt der Dämon die Schnittstelle mit der niedrigsten Anzahl von IP-Adressen aus. Da das Failover auf Adapterschnittstellenebene erfolgt, sind Verbindungen auf höherer Ebene wie TCP mit Ausnahme einer kurzen Zeitspanne während des Failover-Vorgangs nicht betroffen. Wenn das Failover der IP-Adressen erfolgreich abgeschlossen ist, werden ARP-Übertragungen gesendet. Der Dämon erhält also die Konnektivität der Remote-Clients aufrecht.

Hinweis – Aufgrund der TCP-Wiederherstellungseigenschaften bei Stau kann an TCP-Endpunkten nach einem erfolgreichen Failover eine weitere Verzögerung entstehen. Einige Segmente sind möglicherweise beim Failover verloren gegangen, wodurch der Stausteuermechanismus im TCP aktiviert wird.

Multipathing-Gruppen bilden die Bausteine für logische Hostnamen und gemeinsam genutzte Adressen. Sie können auch unabhängig von logischen Hostnamen und gemeinsam genutzten Adressen Multipathing-Gruppen erstellen, um die Konnektivität der Cluster-Knoten mit dem öffentlichen Netzwerk zu überwachen. Dieselbe Multipathing-Gruppe kann auf einem Knoten eine beliebige Anzahl logischer Hostnamen oder gemeinsam genutzter Adressen hosten. Weitere Informationen zu logischen Hostnamen und gemeinsam genutzten Adressen finden Sie im *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

Hinweis – Der Internet Protocol (IP) Network Multipathing-Mechanismus ist auf das Erkennen und Verbergen von Adapterfehlern ausgelegt. Er dient nicht dazu, eine Wiederherstellung durchzuführen, nachdem ein Verwalter mit `ifconfig(1M)` eine der logischen (oder gemeinsam genutzten) IP-Adressen entfernt hat.. Die Sun Cluster-Software zeigt die logischen und gemeinsam genutzten IP-Adressen als Ressourcen an, die von RGM verwaltet werden. Ein Verwalter muss eine IP-Adresse mit `scrgadm(1M)` entfernen oder hinzufügen, um die Ressourcengruppe zu ändern, die die Ressource enthält.

Weitere Informationen zur Solaris-Implementierung von IP Network Multipathing finden Sie in der entsprechenden Dokumentation zu dem auf Ihrem Cluster installierten Solaris-Betriebssystem.

Version des Betriebssystems	Anweisungen
Betriebssystem Solaris 8	<i>IP Network Multipathing Administration Guide</i>
Betriebssystem Solaris 9	Kapitel 1, „IP Network Multipathing (Overview)“ in <i>IP Network Multipathing Administration Guide</i>
Betriebssystem Solaris 10	Teil VI, „IPMP“ in <i>System Administration Guide: IP Services</i>

SPARC: Unterstützung der dynamischen Rekonfiguration

Die Softwarefunktion zur Unterstützung der dynamischen Rekonfiguration (DR) in Sun Cluster 3.1 8/05 wird schrittweise umgesetzt. In diesem Abschnitt werden Konzepte und Erwägungen im Hinblick auf die Unterstützung der DR-Funktion durch Sun Cluster 3.1 8/05 beschrieben.

Alle für die DR-Funktion von Solaris dokumentierten Anforderungen, Verfahren und Einschränkungen gelten auch für die DR-Unterstützung von Sun Cluster (mit Ausnahme des Vorgangs zur Stilllegung der Betriebsumgebung). Lesen Sie daher die Dokumentation zur Solaris DR-Funktion, *bevor* Sie die DR-Funktion mit der Sun Cluster-Software verwenden. Lesen Sie insbesondere die Themen, die sich mit nicht vernetzten E/A-Geräten während eines DR-Trennungsvorgangs beschäftigen

Das *Sun Enterprise 10000 Dynamic Reconfiguration User Guide* und das *Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual* (aus den Dokumentationsreihen *Solaris 8 on Sun Hardware* oder *Solaris 9 on Sun Hardware*) können von <http://docs.sun.com> heruntergeladen werden.

SPARC: Allgemeine Beschreibung der dynamischen Rekonfiguration

Die DR-Funktion ermöglicht bestimmte Vorgänge wie zum Beispiel das Entfernen von Systemhardware bei laufenden Systemen. Die DR-Prozesse sind so ausgelegt, dass sie einen kontinuierlichen Systembetrieb sicherstellen, ohne das System anhalten oder die Cluster-Verfügbarkeit unterbrechen zu müssen.

DR arbeitet auf Board-Ebene. Deswegen hat ein DR-Vorgang Auswirkungen auf alle Komponenten eines Boards. Jedes Board kann mehrere Komponenten enthalten, einschließlich CPUs, Speicher und Peripherieschnittstellen für Plattenlaufwerke, Bandlaufwerke und Netzwerkverbindungen.

Das Entfernen eines Boards mit aktiven Komponenten führt zu Systemfehlern. Vor dem Entfernen eines Boards fragt das DR-Teilsystem andere Teilsysteme wie Sun Cluster ab, um festzustellen, ob die Komponenten des Boards genutzt werden. Wenn das DR-Teilsystem eine Board-Nutzung feststellt, wird der DR-Vorgang zur Board-Entfernung nicht ausgeführt. Ein DR-Vorgang zur Board-Entfernung ist insofern immer sicher, als das DR-Teilsystem Vorgänge an Boards mit aktiven Komponenten ablehnt.

Der DR-Vorgang zur Board-Hinzufügung ist ebenfalls immer sicher. CPUs und Speicher auf einem neu hinzugefügten Board werden automatisch vom System verfügbar gemacht. Der Systemverwalter muss den Cluster jedoch manuell konfigurieren, um Komponenten des neu hinzugefügten Boards aktiv nutzen zu können.

Hinweis – Das DR-Teilsystem hat mehrere Ebenen. Wenn ein Fehler auf einer unteren Ebene gemeldet wird, meldet auch die übergeordnete Ebene einen Fehler. Die untere Ebene meldet den spezifischen Fehler, die übergeordnete Ebene meldet jedoch nur **Unbekannter Fehler**. Diese Fehlermeldung können Sie gefahrlos ignorieren.

Die folgenden Abschnitte enthalten DR-spezifische Erwägungen zu den unterschiedlichen Gerätetypen.

SPARC: Erwägungen zur DR von CPU-Geräten im Cluster

Die Sun Cluster-Software lehnt einen DR-Vorgang zur Board-Entfernung nicht aufgrund vorhandener CPU-Geräte ab.

Wenn ein DR-Vorgang zur Board-Hinzufügung erfolgt, werden die CPU-Geräte auf dem hinzugefügten Board automatisch in den Systembetrieb integriert.

SPARC: Erwägungen zur DR von Speichern im Cluster

Für DR-Zwecke sind zwei Speichertypen zu berücksichtigen.

- Kernel-Speichergehäuse
- Nicht-Kernel-Speichergehäuse

Diese beiden Typen unterscheiden sich lediglich bei ihrer Nutzung. Die aktuelle Hardware ist bei beiden Typen die gleiche. Das Kernel-Speichergehäuse ist der vom Solaris-Betriebssystem genutzte Speicher. Die Sun Cluster-Software unterstützt die Board-Entfernung bei einem Board, das Kernel-Speichergehäuse enthält, nicht und lehnt jeden derartigen Vorgang ab. Wenn ein DR-Vorgang zur Board-Hinzufügung Speicher betrifft, bei dem es sich nicht um das Kernel-Speichergehäuse handelt, lehnt die Sun Cluster-Software den Vorgang nicht ab. Wenn ein den Speicher betreffender DR-Vorgang zur Board-Hinzufügung durchgeführt wird, wird der Speicher auf dem hinzugefügten Board automatisch in den Systembetrieb integriert.

SPARC: Erwägungen zur DR von Platten- und Bandlaufwerken im Cluster

Sun Cluster lehnt DR-Vorgänge zur Board-Entfernung auf aktiven Laufwerken des Primärknotens ab. DR-Vorgänge zur Board-Entfernung können auf nicht aktiven Laufwerken des Primärknotens und auf allen Laufwerken der Sekundärknoten durchgeführt werden. Der Cluster-Datenzugriff unterliegt durch den DR-Vorgang keinerlei Änderungen.

Hinweis – Sun Cluster lehnt DR-Vorgänge ab, die sich auf die Verfügbarkeit von Quorum-Geräten auswirken. Weitere Erwägungen zu Quorum-Geräten und dem Verfahren zur Ausführung von DR-Vorgängen bei Cluster-Geräten finden Sie unter [„SPARC: Erwägungen zur DR von Quorum-Geräten im Cluster“](#) auf Seite 94.

Unter „Dynamische Rekonfiguration von Quorum-Geräten“ in *Sun Cluster Handbuch Systemverwaltung für Solaris OS* erhalten Sie detaillierte Anweisungen zur Durchführung dieser Vorgänge.

SPARC: Erwägungen zur DR von Quorum-Geräten im Cluster

Wenn der DR-Vorgang zur Board-Entfernung ein Board betrifft, das eine Schnittstelle zu einem für das Quorum konfigurierten Gerät enthält, lehnt die Sun Cluster-Software den Vorgang ab. Außerdem identifiziert die Sun Cluster-Software das Quorum-Gerät, das von diesem Vorgang betroffen wäre. Sie müssen dieses Gerät als Quorum-Gerät deaktivieren, bevor Sie einen DR-Vorgang zur Board-Entfernung durchführen können.

In Kapitel 5, „Verwalten des Quorums“ in *Sun Cluster Handbuch Systemverwaltung für Solaris OS* erhalten Sie detaillierte Anweisungen zum Verwalten des Quorums.

SPARC: Erwägungen zur DR von Cluster-Interconnect-Schnittstellen im Cluster

Wenn der DR-Vorgang zur Board-Entfernung ein Board betrifft, das zu einer aktiven Cluster-Interconnect-Schnittstelle gehört, lehnt die Sun Cluster-Software den Vorgang ab. Außerdem identifiziert die Sun Cluster-Software die Schnittstelle, die von dem Vorgang betroffen wäre. Sie müssen die aktive Schnittstelle mit einem Sun Cluster-Verwaltungstool deaktivieren, bevor der DR-Vorgang erfolgen kann.



Caution – Die Sun Cluster-Software erfordert, dass jeder Cluster-Knoten über mindestens einen funktionsfähigen Pfad zu jedem anderen Cluster-Knoten verfügt. Deaktivieren Sie keine privaten Interconnect-Schnittstellen, die den letzten Pfad zu einem Cluster-Knoten unterstützen.

Unter „Verwalten von Cluster-Interconnects“ in *Sun Cluster Handbuch Systemverwaltung für Solaris OS* erhalten Sie detaillierte Anweisungen zur Durchführung dieser Vorgänge.

SPARC: Erwägungen zur DR von öffentlichen Netzwerkschnittstellen im Cluster

Wenn der DR-Vorgang zur Board-Entfernung ein Board betrifft, das aktive öffentliche Netzwerkschnittstellen enthält, lehnt die Sun Cluster-Software den Vorgang ab. Außerdem identifiziert die Sun Cluster-Software die Schnittstelle, die von dem Vorgang betroffen wäre. Bevor ein Board mit einer aktiven öffentlichen Netzwerkschnittstelle entfernt werden kann, muss der gesamte Datenverkehr mithilfe des `if_mpadm(1M)`-Befehls von dieser Schnittstelle auf eine andere funktionsfähige Schnittstelle in der Multipathing-Gruppe umgeleitet werden.



Caution – Wenn der verbleibende Netzwerkadapter während des DR-Entfernungs Vorgangs für den deaktivierten Netzwerkadapter ausfällt, wird die Verfügbarkeit beeinträchtigt. Der verbleibende Adapter hat keine Möglichkeit, für die Dauer des DR-Vorgangs zu wechseln.

Unter „Verwalten des öffentlichen Netzwerks“ in *Sun Cluster Handbuch Systemverwaltung für Solaris OS* erhalten Sie detaillierte Anweisungen zur Durchführung eines DR-Entfernungs Vorgangs in einem öffentlichen Netzwerk.

Fragen und Antworten

Dieses Kapitel enthält die häufig gestellten Fragen zum Sun Cluster-System. Die Fragen sind nach Thema geordnet.

FAQs zur hohen Verfügbarkeit

Frage: Was genau ist ein hoch verfügbares System?

Antwort: Das Sun Cluster-System definiert hohe Verfügbarkeit (HA) als die Fähigkeit eines Clusters, eine Anwendung kontinuierlich auszuführen. Die Anwendung wird auch dann ausgeführt, wenn ein Fehler auftritt, nach dem ein Serversystem in der Regeln nicht mehr verfügbar ist.

Frage: Mit welchem Prozess sorgt der Cluster für Hochverfügbarkeit?

Antwort: Mit einem Prozess, der als Failover bezeichnet wird, sorgt das Cluster-Framework für eine hoch verfügbare Umgebung. Ein Failover ist eine Reihe von Schritten, mit denen der Cluster Datendienstressourcen von einem versagenden Knoten auf einen anderen funktionsfähigen Knoten im Cluster migriert.

Frage: Was ist der Unterschied zwischen einem Failover- und einem Scalable-Datendienst?

Antwort: Es gibt zwei Typen von hoch verfügbaren Datendiensten:

- Failover
- Scalable

Ein Failover-Datendienst führt eine Anwendung jeweils nur auf einem Primärknoten im Cluster aus. Andere Knoten können andere Anwendungen ausführen, aber jede Anwendung wird nur einmal auf einem einzigen Knoten ausgeführt. Wenn ein Primärknoten ausfällt, werden die auf dem ausgefallenen Knoten ausgeführten Anwendungen mit einem Failover auf einen anderen Knoten verschoben und dort weiter ausgeführt.

Ein Scalable-Dienst verteilt eine Anwendung auf mehrere Knoten, um einen einzigen logischen Dienst zu erstellen. Scalable-Dienste setzen die Anzahl der Knoten und Prozessoren auf dem ganzen Cluster, auf dem sie durchgeführt werden, wirksam ein.

Für jede Anwendung hostet ein Knoten die reale Schnittstelle zum Cluster. Dieser Knoten wird als globaler Schnittstellenknoten (GIF-Knoten) bezeichnet. Im Cluster können mehrere GIF-Knoten vorhanden sein. Jeder GIF-Knoten hostet eine oder mehr logische Schnittstellen, die von Scalable-Diensten genutzt werden können. Diese logischen Schnittstellen werden als *globale Schnittstellen* bezeichnet. Ein GIF-Knoten hostet eine globale Schnittstelle für alle Anforderungen einer bestimmten Anwendung und sendet sie an mehrere Knoten, auf denen der Anwendungsserver läuft. Wenn der GIF-Knoten ausfällt, wird die globale Schnittstelle auf einen auch weiterhin laufenden Knoten umgeleitet.

Wenn einer der Knoten ausfällt, auf denen die Anwendung ausgeführt wird, wird sie weiterhin auf den anderen Knoten mit leicht verminderter Leistung ausgeführt. Dieser Prozess wird fortgesetzt, bis der ausgefallene Knoten wieder zum Cluster hinzukommt.

FAQs zu Dateisystemen

Frage: Kann mindestens ein Cluster-Knoten als hoch verfügbarer NFS-Server mit anderen Cluster-Knoten als Client konfiguriert werden?

Antwort: Nein, schaffen Sie keine Schleifeneinhängung.

Frage: Kann ein Cluster-Dateisystem für Anwendungen verwendet werden, die nicht von Ressourcengruppen-Manager gesteuert werden?

Antwort: Ja. Allerdings müssen die Anwendungen ohne Steuerung durch den RGM nach dem Ausfall des Knotens, auf dem sie ausgeführt werden, manuell neu gestartet werden.

Frage: Müssen alle Cluster-Dateisysteme einen Einhängpunkt unter dem `/global`-Verzeichnis haben?

Antwort: Nein. Aber Cluster-Dateisysteme mit demselben Einhängpunkt (wie zum Beispiel `/global`) ermöglichen eine bessere Organisation und Verwaltung dieser Dateisysteme.

Frage: Welches sind die Unterschiede zwischen der Verwendung eines Cluster-Dateisystems und dem Exportieren von NFS-Dateisystemen?

Antwort: Es gibt mehrere Unterschiede:

1. Das Cluster-Dateisystem unterstützt globale Geräte. NFS unterstützt keinen Remote-Zugriff auf Geräte.
2. Das Cluster-Dateisystem hat einen globalen Namensraum. Nur ein einziger Einhängbefehl ist erforderlich. Bei NFS müssen Sie das Dateisystem auf jedem Knoten einhängen.

3. Das Cluster-Dateisystem speichert die Dateien in mehr Fällen als NFS im Cache. Das Cluster-Dateisystem speichert beispielsweise Dateien, wenn auf eine Datei von mehreren Knoten aus zum Lesen, Schreiben, für Dateisperren oder asynchronen E/A zugegriffen wird.
4. Das Cluster-Dateisystem ist auf die Nutzung zukünftiger schneller Cluster-Interconnects ausgelegt, die DMA-Funktionen und Funktionen ohne Zwischenspeicherung (Zero-Copy) im Fernzugriff zur Verfügung stellen.
5. Wenn Sie die Attribute für eine Datei (zum Beispiel mit `chmod (1M)`) in einem Cluster-Dateisystem ändern, wird diese Änderung sofort auf alle Knoten gespiegelt. Bei einem exportierten NFS-Dateisystem kann das wesentlich mehr Zeit in Anspruch nehmen.

Frage: Das Dateisystem `/global/.devices/node@nodeID` wird auf den Cluster-Knoten angezeigt. Kann dieses Dateisystem zum Speichern von Daten verwendet werden, die hoch verfügbar und global sein sollen?

Antwort: Diese Dateisysteme speichern den Namensraum für globale Geräte. Diese Dateisysteme sind nicht zum allgemeinen Gebrauch vorgesehen. Sie sind zwar global, aber der Zugriff auf sie erfolgt niemals global – jeder Knoten kann nur auf seinen eigenen Namensraum für globale Geräte zugreifen. Wenn ein Knoten ausgefallen ist, können andere Knoten nicht auf den Namensraum für den ausgefallenen Knoten zugreifen. Diese Dateisysteme sind nicht hoch verfügbar. Sie sollten nicht zum Speichern von Daten verwendet werden, die global zugänglich oder hoch verfügbar sein müssen.

FAQs zur Datenträgerverwaltung

Frage: Müssen alle Plattengeräte gespiegelt werden?

Antwort: Damit ein Plattengerät als hoch verfügbar gilt, muss es gespiegelt werden oder RAID-5-Hardware verwenden. Alle Datendienste sollten entweder hoch verfügbare Plattengeräte oder Cluster-Dateisysteme verwenden, die auf hoch verfügbaren Plattengeräten eingehängt sind. Solche Konfigurationen tolerieren einzelne Plattenausfälle.

Frage: Kann ein Datenträger-Manager für die lokalen Platten (Boot-Platte) und ein anderer für die Multihostplatten verwendet werden?

Antwort: SPARC: Diese Konfiguration wird von der Solaris Volume Manager-Software zur Verwaltung der lokalen Platten und von VERITAS Volume Manager zur Verwaltung der Multihostplatten unterstützt. Keine andere Kombination wird unterstützt.

x86: Diese Konfiguration wird nicht unterstützt, da in x86-basierten Clustern nur Solaris Volume Manager unterstützt wird.

FAQs zu Datendiensten

Frage: Welche Sun Cluster-Datendienste sind verfügbar?

Antwort: Die Liste der unterstützten Datendienste finden Sie unter „Unterstützte Produkte“ in *Sun Cluster 3.1 8/05 Versionshinweise für Solaris OS*.

Frage: Welche Anwendungsversionen werden von den Sun Cluster-Datendiensten unterstützt?

Antwort: Die Liste der unterstützten Anwendungsversionen finden Sie unter „Unterstützte Produkte“ in *Sun Cluster 3.1 8/05 Versionshinweise für Solaris OS*.

Frage: Können eigene Datendienste geschrieben werden?

Antwort: Ja. Weitere Informationen finden Sie in Kapitel 11, „DSDL-API-Funktionen“ in *Sun Cluster Entwicklerhandbuch Datendienste für Solaris OS*.

Frage: Müssen beim Erstellen von Netzwerkressourcen numerische IP-Adressen oder Hostnamen angegeben werden?

Antwort: Die bevorzugte Methode zur Angabe von Netzwerkressourcen ist die Verwendung des UNIX-Hostnamens anstelle der numerischen IP-Adresse.

Frage: Was ist beim Erstellen von Netzwerkressourcen der Unterschied zwischen der Verwendung eines logischen Hostnamens (LogicalHostname-Ressource) und einer gemeinsam genutzten Adresse (SharedAddress-Ressource)?

Antwort: Wenn in der Dokumentation die Verwendung einer LogicalHostname-Ressource in einer Ressourcengruppe im Failover-Modus empfohlen wird, kann wahlweise eine SharedAddress-Ressource oder eine LogicalHostname-Ressource verwendet werden. Die einzige Ausnahme bildet Sun Cluster HA für NFS. Die Verwendung einer SharedAddress-Ressource sorgt für zusätzlichen Überlauf weil die Cluster-Netzwerksoftware für eine SharedAddress, aber nicht für einen LogicalHostname konfiguriert ist.

Der Vorteil einer SharedAddress wird deutlich, wenn Sie sowohl Scalable- als auch Failover-Datendienste konfigurieren und möchten, dass Clients mithilfe desselben Hostnamens auf beide Dienste zugreifen können. In diesem Fall ist bzw. sind die SharedAddress-Ressource(n) zusammen mit der Failover-Anwendungsressource in einer Ressourcengruppe enthalten. Die Scalable-Dienstressource ist in einer separaten Ressourcengruppe untergebracht und für die Verwendung der SharedAddress-Ressource konfiguriert. Sowohl Scalable- als auch Failover-Dienste können denselben Satz von Hostnamen/Adressen verwenden, der in der SharedAddress-Ressource konfiguriert ist.

FAQs zum öffentlichen Netzwerk

Frage: Welche öffentlichen Netzwerkadapter werden vom Sun Cluster-System unterstützt?

Antwort: Derzeit unterstützt das Sun Cluster-System öffentliche Netzwerkadapter für Ethernet (10/100BASE-T und 1000BASE-SX Gb). Zukünftig werden ggf. neue Schnittstellen unterstützt. Aktuelle Informationen erhalten Sie von Ihrem Sun-Vertreter.

Frage: Welche Rolle spielt die MAC-Adresse beim Failover?

Antwort: Bei einem Failover werden neue ARP-Pakete (Address Resolution Protocol-Pakete) generiert und an alle gesendet. Diese ARP-Pakete enthalten die neue MAC-Adresse (des neuen tatsächlichen Adapters, zu dem der Knoten gewechselt ist) und die alte IP-Adresse. Wenn ein anderer Rechner im Netzwerk eines dieser Pakete erhält, löscht er die alte MAC-IP-Zuordnung aus seinem ARP-Cache und verwendet die neue.

Frage: Unterstützt das Sun Cluster-System die Einstellung `local-mac-address? =true` ?

Antwort: Ja. Tatsächlich erfordert IP Network Multipathing, dass `local-mac-address?` auf `true` gesetzt sein *muss*.

Sie können `local-mac-address?` mit `eeprom(1M)` an der Eingabeaufforderung `ok` in einem SPARC-basierten Cluster einstellen. Alternativ können Sie die MAC-Adresse mit dem SCSI-Dienstprogramm einstellen, das Sie optional nach dem Starten der BIOS in einem x86-basierten Cluster ausführen.

Frage: Welche Verzögerung ist zu erwarten, wenn Internet Protocol (IP) Network Multipathing ein Switchover zwischen Adapters ausführt?

Antwort: Die Verzögerung kann mehrere Minuten betragen. Der Grund dafür ist, dass ein Internet Protocol (IP) Network Multipathing-Switchover das Senden eines freien ARP beinhaltet. Sie können jedoch nicht sicher sein, dass der Router zwischen Client und Cluster dieses freie ARP auch verwendet. Bis zum Ablauf der Gültigkeit des ARP-Cache-Eintrags für diese IP-Adresse auf dem Router wird für den Eintrag möglicherweise die veraltete MAC-Adresse verwendet.

Frage: Wie schnell werden Fehler an einem Netzwerkadapter erkannt?

Antwort: Die Standard-Fehlererkennungszeit beträgt 10 Sekunden. Der Algorithmus versucht, die Fehlererkennungszeit einzuhalten, doch die tatsächlich benötigte Zeit hängt von der Netzwerklast ab.

FAQs zu Cluster-Mitgliedern

Frage: Muss das Root-Passwort für alle Cluster-Mitglieder gleich sein?

Antwort: Sie sind nicht gezwungen, dasselbe Root-Passwort für jedes Cluster-Mitglied zu verwenden. Sie können jedoch die Cluster-Verwaltung vereinfachen, wenn Sie auf allen Knoten dasselbe Root-Passwort verwenden.

Frage: Ist es wichtig, in welcher Reihenfolge die Knoten gebootet werden?

Antwort: In den meisten Fällen nicht. Die Boot-Reihenfolge ist jedoch wichtig, um Amnesie zu verhindern. Wenn Knoten Zwei zum Beispiel der Eigentümer des Quorum-Geräts war, Knoten Eins nicht aktiv ist und Sie dann Knoten Zwei herunterfahren, müssen Sie Knoten Zwei vor Knoten Eins wieder hochfahren. Mit dieser Reihenfolge wird verhindert, dass Sie versehentlich einen Knoten mit veralteten Cluster-Konfigurationsinformationen starten. Details zu Amnesie finden Sie unter „[Informationen zum Fehlerschutz](#)“ auf Seite 57.

Frage: Müssen lokale Platten in einem Cluster-Knoten gespiegelt werden?

Antwort: Ja. Diese Spiegelung ist zwar keine Voraussetzung, aber das Spiegeln der Platten der Cluster-Knoten beugt einem Ausfall einer nicht gespiegelten Platte vor, der einen Knotenausfall nach sich zieht. Der Nachteil beim Spiegeln lokaler Platten eines Cluster-Knotens ist ein erhöhter Systemverwaltungs Aufwand.

Frage: Welche Sicherungslösungen gibt es für Cluster-Mitglieder?

Antwort: Sie können mehrere Sicherungsmethoden für einen Cluster verwenden. Eine Methode besteht darin, einen Knoten als Sicherungsknoten mit angeschlossenem Bandlaufwerk/Bibliothek zu verwenden. Dann sichern Sie die Daten mit dem Cluster-Dateisystem. Verbinden Sie diesen Knoten nicht mit den gemeinsam genutzten Platten.

Weitere Informationen zum Sichern und Wiederherstellen von Daten erhalten Sie in Kapitel 9, „Sichern und Wiederherstellen eines Clusters“ in *Sun Cluster Handbuch Systemverwaltung für Solaris OS*.

Frage: Wann ist ein Knoten stabil genug, um als Sekundärknoten eingesetzt zu werden?

Antwort: Solaris 8 und Solaris 9:

Nach dem erneuten Booten ist ein Knoten stabil genug, um als Sekundärknoten eingesetzt zu werden, wenn der Knoten die Anmelde-Eingabeaufforderung anzeigt.

Solaris 10:

Ein Knoten ist stabil genug, um als sekundärer Knoten zu fungieren, wenn der Meilenstein `multi-user-server` ausgeführt wird.

```
# svcs -a | grep multi-user-server:default
```

FAQs zum Cluster-Speicher

Frage: Was macht den Multihost-Speicher hoch verfügbar?

Antwort: Multihost-Speicher ist hoch verfügbar, weil er den Verlust einer Einzelplatte dank der Spiegelung (oder dank hardwarebasierter RAID-5-Controller) übersteht. Ein Multihost-Speichergerät hat mehr als eine Hostverbindung. Deswegen widersteht es dem Verlust eines einzelnen Knotens, an den es angeschlossen ist. Zudem sorgen redundante Pfade von jedem Knoten zum angeschlossenen Speicher dafür, dass ein Versagen eines Hostbusadapters, Kabels oder Platten-Controllers toleriert wird.

FAQs zum Cluster-Interconnect

Frage: Welche Cluster-Interconnect werden vom Sun Cluster-System unterstützt?

Antwort: Derzeit unterstützt das Sun Cluster-System die folgenden Cluster-Interconnects:

- Ethernet (100BASE-T Fast Ethernet und 1000BASE-SX Gb) sowohl in SPARC-basierten als auch in x86-basierten Clustern
- Infiniband in SPARC-basierten und x86-basierten Clustern
- SCI nur in SPARC-basierten Clustern

Frage: Was ist der Unterschied zwischen einem "Kabel" und einem "Transportpfad" ?

Antwort: Cluster-Transportkabel werden unter Verwendung von Transportadaptern und Schaltern konfiguriert. Kabel verbinden Adapter und Schalter auf einer Komponente-zu-Komponenten-Basis. Der Cluster-Topologie-Manager setzt die verfügbaren Kabel ein, um durchgehende (von Ende zu Ende) Transportpfade zwischen den Knoten aufzubauen. Ein Kabel ist einem Transportpfad nicht direkt zugeordnet.

Kabel werden von einem Verwalter statisch "aktiviert" und "deaktiviert". Bei Kabeln gibt es einen "Zustand" (aktiv oder inaktiv), aber keinen "Status". Ein inaktives Kabel ist wie ein nicht konfiguriertes Kabel. Inaktive Kabel können nicht als Transportpfade eingesetzt werden. Diese Kabel werden nicht getestet und deshalb ist ihr Status nicht bekannt. Den Status des Kabels können Sie über den Befehl `scconf -p` abrufen.

Transportpfade werden dynamisch durch den Cluster-Topologie-Manager festgelegt. Der "Status" eines Transportpfades wird vom Topologie-Manager bestimmt. Ein Pfad kann den Status "online" oder "offline" haben. Den Status des Transportpfades können Sie mit dem Befehl `scstat (1M)` abrufen.

Betrachten Sie das folgende Beispiel eines Zwei-Knoten-Clusters mit vier Kabeln.

```
node1:adapter0      to switch1, port0
node1:adapter1      to switch2, port0
```

```
node2:adapter0    to switch1, port1
node2:adapter1    to switch2, port1
```

Zwei mögliche Transportpfade können aus diesen vier Kabeln gebildet werden.

```
node1:adapter0    to node2:adapter0
node2:adapter1    to node2:adapter1
```

FAQs zu Client-Systemen

Frage: Müssen bei der Verwendung eines Clusters besondere Client-Anforderungen oder Einschränkungen berücksichtigt werden?

Antwort: Client-Systeme stellen mit dem Cluster die Verbindung auf dieselbe Weise wie mit einem anderen Server her. Bei manchen Instanzen kann es je nach Datendienstanwendung erforderlich sein, clientseitige Software zu installieren oder sonstige Änderungen an der Konfiguration vorzunehmen, damit der Client eine Verbindung mit der Datendienstanwendung herstellen kann. Weitere Informationen zu clientseitigen Konfigurationsanforderungen finden Sie in Kapitel 1, „Planning for Sun Cluster Data Services“ in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

FAQs zur Verwaltungskonsole

Frage: Ist für das Sun Cluster-System eine Verwaltungskonsole erforderlich?

Antwort: Ja.

Frage: Muss die Verwaltungskonsole dem Cluster zugewiesen werden oder kann sie auch für andere Aufgaben eingesetzt werden?

Antwort: Das Sun Cluster-System benötigt keine dedizierte Verwaltungskonsole, die Verwendung einer solchen Konsole bietet jedoch folgende Vorteile:

- Sie ermöglicht eine zentralisierte Cluster-Verwaltung durch das Gruppieren von Konsolen- und Verwaltungstools auf demselben Rechner.
- Sie sorgt für eine potenziell schnellere Problemlösung durch Ihren Hardware-Dienstleister.

Frage: Muss sich die Verwaltungskonsole in der "Nähe" des Clusters befinden, zum Beispiel in demselben Raum?

Antwort: Fragen Sie bei Ihrem Hardware-Dienstleister nach. Der Dienstleister kann ggf. die räumliche Nähe der Konsole zum eigentlichen Cluster vorschreiben. Es gibt keine technischen Gründe, die eine Unterbringung der Konsole in demselben Raum erforderlich machen.

Frage: Können Verwaltungskonsolen mehr als einen Cluster verwalten, wenn ggf. vorhandene Abstandsvorgaben eingehalten werden?

Antwort: Ja. Sie können von einer einzigen Verwaltungskonsole aus mehrere Cluster steuern. Sie können auch einen einzigen Terminal-Konzentrator zwischen den Clustern teilen.

FAQs zu Terminal-Konzentrator und System Service Processor

Frage: Erfordert das Sun Cluster-System einen Terminal-Konzentrator?

Antwort: Softwareversionen ab Sun Cluster 3.0 benötigen keinen Terminal-Konzentrator. Anders als das Produkt Sun Cluster 2.2, das einen Terminal-Konzentrator als Fehlerschutz benötigte, hängen neuere Produkte nicht vom Terminal-Konzentrator ab.

Frage: Die meisten Sun Cluster-Server verwenden einen Terminal-Konzentrator, aber der Sun Enterprise E1000-Server nicht. Weshalb nicht?

Antwort: Der Terminal-Konzentrator ist für die meisten Server tatsächlich ein Seriell-zu-Ethernet-Konverter. Der Port der Konsole für den Terminal-Konzentrator ist ein serieller Port. Der Sun Enterprise E1000-Server verfügt nicht über eine serielle Konsole. Der System Service Processor (SSP) ist die Konsole, entweder über einen Ethernet- oder einen jtag-Port. Bei Sun Enterprise E1000-Servern verwenden Sie immer den SSP als Konsole.

Frage: Welches sind die Vorteile eines Terminal-Konzentrators?

Antwort: Mit einem Terminal-Konzentrator können Sie auf Konsolenebene von einer Remote-Workstation an einer beliebigen Stelle im Netzwerk auf jeden Knoten zugreifen. Dieser Zugriff ist selbst dann möglich, wenn sich der Knoten am OpenBoot PROM (OBP) in einem SPARC-basierten Knoten oder einem Boot-Teilsystem in einem x86-basierten Knoten befindet.

Frage: Welches sind die für die Qualifizierung eines nicht von Sun unterstützten Terminal-Konzentrators erforderlichen Kenntnisse?

Antwort: Der Hauptunterschied zwischen dem von Sun unterstützten Terminal-Konzentrator und anderen Konsolengeräten besteht darin, dass der Terminal-Konzentrator von Sun mit spezieller Firmware ausgestattet ist. Diese Firmware hindert den Terminal-Konzentrator daran, einen Abbruch an die bootende Konsole zu senden. Der Knoten wird durch ein Konsolengerät heruntergefahren, wenn das Gerät einen Abbruch oder ein von der Konsole als Abbruch interpretiertes Signal senden kann.

Frage: Kann ein gesperrter Port am von Sun unterstützten Terminal-Konzentrator freigeschaltet werden, ohne ihn neu zu booten?

Antwort: Ja. Beachten Sie, dass die Port-Nummer neu eingestellt werden muss, und geben Sie folgende Befehle ein:

```
telnet tc
Enter Annex port name or number: cli
annex: su -
annex# admin
admin : reset port_number
admin : quit
annex# hangup
#
```

Weitere Informationen zur Konfiguration und Verwaltung des von Sun unterstützten Terminal-Konzentrators erhalten Sie in den folgenden Handbüchern:

- „Überblick über die Verwaltung von Sun Cluster“ in *Sun Cluster Handbuch Systemverwaltung für Solaris OS*
- Kapitel 2, „Installing and Configuring the Terminal Concentrator“ in *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*

Frage: Was geschieht, wenn der Terminal-Konzentrator selbst ausfällt? Ist ein zweiter als Standby-Gerät erforderlich?

Antwort: Nein. Bei einem Versagen des Terminal-Konzentrators geht keine Cluster-Verfügbarkeit verloren. Sie können erst dann wieder eine Verbindung mit den Knotenkonsolen herstellen, wenn der Konzentrator wieder arbeitet.

Frage: Wenn sieht es bei einem Terminal-Konzentrator in puncto Sicherheit aus?

Antwort: Im Allgemeinen ist der Terminal-Konzentrator mit einem kleinen Netzwerk verbunden, das von Systemverwaltern verwendet wird, und nicht mit einem Netzwerk, das für sonstigen Client-Zugriff eingesetzt wird. Sie können die Sicherheit durch einen beschränkten Zugriff auf dieses Netzwerk steuern.

Frage: SPARC: Wie wird die dynamische Rekonfiguration für ein Band- oder Plattenlaufwerk eingesetzt?

Antwort: Führen Sie folgende Schritte durch:

- Stellen Sie fest, ob das Platten- oder Bandlaufwerk zu einer aktiven Gerätegruppe gehört. Wenn das Laufwerk zu keiner aktiven Gerätegruppe gehört, können Sie den DR-Entfernungsvorgang ausführen.
- Wenn der DR-Vorgang zur Board-Entfernung ein aktives Platten- oder Bandlaufwerk betrifft, lehnt das System den Vorgang ab und identifiziert die betroffenen Laufwerke. Wenn das Laufwerk zu einer aktiven Gerätegruppe gehört, erhalten Sie weitere Informationen unter [„SPARC: Erwägungen zur DR von Platten- und Bandlaufwerken im Cluster “](#) auf Seite 93.
- Stellen Sie fest, ob das Laufwerk eine Komponente des Primär- oder Sekundärknotens ist. Wenn das Laufwerk eine Komponente des Sekundärknotens ist, können Sie den DR-Entfernungsvorgang ausführen.

- Wenn das Laufwerk eine Komponente des Primärknotens ist, müssen Sie vom Primär- auf den Sekundärknoten umschalten, bevor Sie den DR-Entfernungsvorgang für das Gerät durchführen können.



Caution – Wenn der aktuelle Primärknoten ausfällt, während Sie den DR-Vorgang auf einem Sekundärknoten ausführen, wirkt sich das auf die Cluster-Verfügbarkeit aus. Der Primärknoten kann kein Failover auf einen anderen Knoten ausführen, bis ein neuer Sekundärknoten bereitgestellt wird.

Index

A

Adapter, *Siehe* Netzwerk, Adapter
Agenten, *Siehe* Datendienste
Amnesie, 55
Anwendung, *Siehe* Datendienste
Anwendungsentwicklung, 37-95
Anwendungskommunikation, 75-76
Anwendungsverteilung, 61
APIs, 73-75, 77
Attribute, *Siehe* Eigenschaften
auto-boot? parameter, 41

B

Bandlaufwerk, 26
Board-Entfernung, Dynamische
 Rekonfiguration, 93
Boot-Platte, *Siehe* Platten, lokal
Boot-Reihenfolge, 102

C

CCP, 29
CCR, 41-42
CD-ROM-Laufwerk, 26
Client/Server-Konfiguration, 65
Client-Systeme, 28
 Einschränkungen, 104
 FAQs, 104
clprivnet-Treiber, 76

Cluster

Anwendungsentwickler, 17-18
Anwendungsentwicklung, 37-95
Aufgabenliste, 18-19
Beschreibung, 14
Board-Entfernung, 93
Boot-Reihenfolge, 102
Dateisystem, 48-51, 98-99
 FAQs
 Siehe auch Dateisystem
 HAStoragePlus, 50
 Verwenden, 49-50
Datendienste, 65-73
Datenträger, 26
FAQs zum Speicher, 103
Hardware, 15-16, 21-29
Interconnect, 23, 27
 Adapter, 27
 Datendienste, 75-76
 Dynamische Rekonfiguration, 94
 FAQs, 103-104
 Kabel, 27
 Schnittstellen, 27
 Unterstützung, 103-104
 Verbindungspunkte, 27
Knoten, 22-23
Konfiguration, 79-89
Konfiguration, 41-42
Mitglieder, 22, 40-41
 FAQs, 102
 Rekonfiguration, 40-41
Öffentliche Netzwerkschnittstelle, 65
öffentliches Netzwerk, 27-28

- Cluster (Fortsetzung)
 - Passwort, 102
 - Service, 15-16
 - Sicherung, 102
 - Softwarekomponenten, 23-24
 - Systemverwaltung, 16-17
 - Topologien, 30-34, 34-35
 - Verwaltung, 37-95
 - Vorteile, 14
 - Zeit, 38-39
 - Ziele, 14
- Cluster Control Panel, 29
- Cluster-Konfigurations-Repository, 41-42
- Cluster-Mitglieder-Monitor, 40-41
- CMM, 40-41
 - Failfast-Mechanismus, 40-41
 - Siehe auch* Failfast
- CPU-Zeit, 79-89

D

- Dateisperrung, 48
- Dateisystem
 - Cluster, 48-51, 98-99
 - Cluster-Dateisystem, 98-99
 - Datenspeicherung, 98-99
 - Einhängen, 48-51, 98-99
 - FAQs, 98-99
 - Global, 98-99
 - Hohe Verfügbarkeit, 98-99
 - lokal, 50
 - NFS, 50-51, 98-99
 - syncdir, 50-51
 - UFS, 50-51
 - VxFS, 50-51
- Dateisysteme, Verwenden, 49-50
- Daten, Speichern, 98-99
- Datendienste, 65-73
 - APIs, 73-75
 - Bibliotheks-API, 74-75
 - Cluster-Interconnect, 75-76
 - Entwickeln, 73-75
 - Failover, 68
 - FAQs, 100
 - Fehler-Monitor, 72-73
 - Hoch verfügbar, 40
 - Konfiguration, 79-89

- Datendienste (Fortsetzung)
 - Methoden, 68
 - Ressourcen, 76-79
 - Ressourcengruppen, 76-79
 - Ressourcentypen, 76-79
 - Scalable, 68-70
 - Unterstützt, 100
- Datenträger, wechselbar, 26
- Datenträgerverwaltung
 - FAQs, 99
 - Lokale Platten, 99
 - Multihost-Geräte, 25
 - Multihostplatten, 99
 - Namensraum, 47
 - RAID-5, 99
 - Solaris Volume Manager, 99
 - VERITAS Volume Manager, 99
 - /dev/global/ Namensraum, 47-48
- DID, 42-43
- DR, *Siehe* Dynamische Rekonfiguration
- DSDL API, 77
- Dynamische Rekonfiguration, 91-95
 - Bandlaufwerke, 93
 - Beschreibung, 92
 - Cluster-Interconnect, 94
 - CPU-Geräte, 93
 - Öffentliches Netzwerk, 94-95
 - Platten, 93
 - Quorum-Geräte, 94
 - Speicher, 93

E

- E10000, *Siehe* Sun Enterprise E10000
- Eigenschaften
 - Ändern, 45-47
 - resource_project_name, 82-83
 - Ressourcen, 79
 - Ressourcengruppen, 79
 - RG_project_name, 82-83
- Einhängen
 - Dateisysteme, 48-51
 - /Global, 98-99
 - Globale Geräte, 48-51
 - mit syncdir, 50-51
- Einzelservermodell, 65
- Entwickler, Cluster-Anwendungen, 17-18

F

- Failback, 72
- Failfast, 41, 58
- Failover
 - Datendienste, 68
 - Plattengerätegruppen, 44-45
 - Szenarien, Solaris
 - Ressourcen-Manager, 84-89
- FAQs, 97-107
 - Client-Systeme, 104
 - Cluster-Interconnect, 103-104
 - Cluster-Mitglieder, 102
 - Cluster-Speicher, 103
 - Dateisysteme, 98-99
 - Datendienste, 100
 - Datenträgerverwaltung, 99
 - Hohe Verfügbarkeit, 97-98
 - Öffentliches Netzwerk, 101
 - System Service Processor, 105-107
 - Terminal-Konzentrator, 105-107
 - Verwaltungskonsole, 104-105
- Fehler
 - Erkennung, 39
 - Failback, 72
 - Schutz, 57-58
 - Schützen, 41
 - Wiederherstellung, 39
- Fehler-Monitor, 72-73
- Fragen und Antworten Asked Questions, *Siehe* FAQs
- Framework, Hochverfügbarkeit, 39-42

G

- Geclusterte-Paare-Topologie, 30-31, 34-35
- Geclustertes Servermodell, 65
- Gemeinsam genutzte Adresse, 65
 - Globaler Schnittstellenknoten, 66
 - Scalable-Datendienste, 68-70
 - Versus logischer Hostname, 100
- Gerät
 - Global, 42-43
 - ID, 42-43
- Geräte
 - Multihost, 24-25
 - Quorum, 55-64
- Gerätegruppe, 43-47

- Gerätegruppe (Fortsetzung)
 - Ändern von Eigenschaften, 45-47
- Gleichzeitiger Zugriff, 22
- Global
 - Gerät, 42-43, 43-47
 - Einhängen, 48-51
 - Lokale Platten, 26
 - Namensraum, 42, 47-48
 - Lokale Platten, 26
 - Schnittstelle, 66
 - Scalable-Dienste, 69
 - /global Einhängpunkt, 48-51
 - /Global Einhängpunkt, 98-99
- Globaler Schnittstellenknoten, 66
- Gruppen
 - Plattengerät
 - Siehe* Platten, Gerätegruppen

H

- HA, *Siehe* Hochverfügbarkeit
- Hardware, 15-16, 21-29, 91-95
 - Siehe auch* Platten
 - Siehe auch* Speicher
 - Cluster-Interconnect-Komponenten, 27
 - Dynamische Rekonfiguration, 91-95
- HAStoragePlus, 50, 76-79
- Herunterfahren, 41
- Hoch verfügbar, Datendienste, 40
- Hochverfügbarkeit, Framework, 39-42
- Hohe Verfügbarkeit, FAQs, 97-98
- Hostname, 65

I

- ID
 - Gerät, 42-43
 - Knoten, 47
- in.mpathd-Dämon, 90
- ioctl, 58
- IP-Adresse, 100
- IP Network Multipathing, 89-91
 - Failover-Zeit, 101
- IPMP, *Siehe* IP Network Multipathing

K

- Kabel, Transport, 103-104
- Kernel, Speicher, 93
- Knoten, 22-23
 - Boot-Reihenfolge, 102
 - Globale Schnittstelle, 66
 - Knoten-ID, 47
 - Primärknoten, 45-47, 66
 - Sekundärknoten, 45-47, 66
 - Sicherung, 102
- Konfiguration
 - Client/Server, 65
 - Datendienste, 79-89
 - parallele Datenbank, 22
 - Repository, 41-42
 - Virtuelle Speicherbegrenzungen, 83-84
- Konfigurationen, Quorum, 59-60
- Konsole
 - System Service Processor, 28-29
 - Verwaltung, 28-29, 29
 - FAQs, 104-105
 - Zugriff, 28-29

L

- Lastausgleich, 70-72
- local_mac_address, 101
- LogicalHostname, *Siehe* Logischer Hostname
- Logischer Hostname, 65
 - Failover-Datendienste, 68
 - Versus gemeinsam genutzte Adresse, 100
- lokale Namensräume, 48
- Lokale Platten, 26
- lokales Dateisystem, 50

M

- MAC-Adresse, 101
- Mitgliedschaft, *Siehe* Cluster, Mitgliedschaft
- Multi-Initiator-SCSI, 25-26
- Multihost-Gerät, 24-25
- Multipathing, 89-91
- Multiport-Plattengerätegruppen, 45-47

N

- N+1 (star)-Topologie, 32-33
- N*N (scalable)-Topologie, 33-34
- Namensräume, 47-48, 48
- Network Time Protocol, 38-39
- Netzwerk
 - Adapter, 27-28, 89-91
 - Gemeinsam genutzte Adresse, 65
 - Lastausgleich, 70-72
 - Logischer Hostname, 65
 - öffentlich, 27-28
 - Öffentlich
 - Dynamische Rekonfiguration, 94-95
 - FAQs, 101
 - IP Network Multipathing, 89-91
 - Schnittstellen, 101
 - privat, 23
 - Ressourcen, 65, 76-79
 - Schnittstellen, 27-28, 89-91
- NFS, 50-51
- NTP, 38-39
- numsecondaries property, 45

O

- Öffentliches Netzwerk, *Siehe* Netzwerk, öffentlich
- Oracle Parallel Server, *Siehe* Oracle Real Application Cluster
- Oracle Real Application Cluster, 74

P

- pair+N-Topologie, 31-32
- Panik, 41, 58
- Parallele Datenbank-Konfigurationen, 22
- Passwort, Root, 102
- Persistent Group Reservation, 58
- Pfad, Transport, 103-104
- PGR, *Siehe* Persistent Group Reservation
- Platten
 - Dynamische Rekonfiguration, 93
 - Fehlerschutz, 57-58
 - Gerätegruppen, 43-47
 - Failover, 44-45
 - Multiport, 45-47

Platten, Gerätegruppen (Fortsetzung)
 Primäre Eigentümerschaft, 45-47
 Globale Geräte, 42-43, 47-48
 lokal, 26, 47-48
 Lokal
 Datenträgerverwaltung, 99
 Spiegelung, 102
 Lokall, 42-43
 Multihost, 42-43, 43-47, 47-48
 SCSI-Geräte, 25-26
Plattenpfadüberwachung, 51-54
preferred property, 45
Primäre Eigentümerschaft,
 Plattengerätegruppen, 45-47
Primärknoten, 66
Privates Netzwerk, 23
Pro-Knoten-Adresse, 75-76
Projekte, 79-89

Q

Quorum, 55-64
 Anforderungen, 59-60
 atypische Konfigurationen, 63
 Empfehlungen, 60-61
 Empfohlene Konfigurationen, 61-63
 Gerät, dynamische Rekonfiguration, 94
 Geräte, 55-64
 Konfigurationen, 59
 Stimmenanzahl, 56-57
 unzulässige Konfigurationen, 63-64

R

recovery, Failback-Einstellungen, 72
Reiner Dienst, 70
Reservierungskonflikt, 58
Resource_project_name Eigenschaft, 82-83
Ressourcen, 76-79
 Eigenschaften, 79
 Zustände, 77-79
Ressourcengruppen, 76-79
 Eigenschaften, 79
 Einstellungen, 77-79
 Failover, 68
 Scalable, 68-70

Ressourcengruppen (Fortsetzung)
 Zustände, 77-79
Ressourcengruppen-Manager, *Siehe* RGM
Ressourcentypen, 50, 76-79
Ressourcenverwaltung, 79-89
Ressources, Einstellungen, 77-79
RG_project_name Eigenschaft, 82-83
RGM, 68, 76-79, 79-89
RMAPI, 77
Root-Passwort, 102

S

Scalable-Datendienste, 68-70
scha_cluster_get -Befehl, 75
scha_privatelink_hostname_node-Argument, 75
Schnittstellen
 Siehe Netzwerk, Schnittstellen
 Verwaltung, 38
Schulung, 11
Schutz, 57-58
Schützen, 41
SCSI
 Fehlerschutz, 57-58
 Multi-Initiator, 25-26
 Persistent Group Reservation, 58
 Reservierungskonflikt, 58
 SCSI-Initiator-Id, Eigenschaft, 26
Sekundärknoten, 66
Servermodelle, 65
SharedAddress, *Siehe* Gemeinsam genutzte
 Adresse
Sicherungsknoten, 102
Softwarekomponenten, 23-24
Solaris-Projekte, 79-89
Solaris Ressourcen-Manager, 79-89
 Failover-Szenarien, 84-89
 Konfigurationsanforderungen, 82-83
 Konfigurieren virtueller
 Speicherbegrenzungen, 83-84
Solaris Volume Manager, Multihost-Geräte, 25
Speicher, 24-25, 93
 Dynamische Rekonfiguration, 93
 FAQs, 103
 SCSI, 25-26
Split Brain, 55, 57-58
SSP, *Siehe* System Service Processor

Sticky-Dienst, 70
Sun Cluster, *Siehe* Cluster
Sun Enterprise E10000, 105-107
 Verwaltungskonsole, 29
Sun Management Center (SunMC), 38
SunPlex, *Siehe* Cluster
SunPlex Manager, 38
syncdir-Einhängeoption, 50-51
System Service Processor, 28-29, 29
 FAQs, 105-107

Z

Zeit, Zwischen Knoten, 38-39
Zuordnung, Namensräume, 48

T

Terminal-Konzentrator, FAQs, 105-107
Topologien, 30-34, 34-35
 Geclustertes Paar, 30-31, 34-35
 N+1 (star), 32-33
 N*N (scalable), 33-34
 pair+N, 31-32
Treiber, Geräte- ID, 42-43

U

UFS, 50-51

V

VERITAS Volume Manager,
 Multihost-Geräte, 25
Verwaltung, Cluster, 37-95
Verwaltung Schnittstellen, 38
Verwaltungskonsole, 29
 FAQs, 104-105
VxFS, 50-51

W

Wechseldatenträger, 26
Wichtige Anwendungen, 63
Wiederherstellung, Fehler detection, 39