



Sun Cluster: Guía de conceptos para el SO Solaris

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

Referencia: 819-2059
Agosto de 2005, Revisión A

Copyright 2005 Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 U.S.A. Reservados todos los derechos.

Este producto o documento está protegido por las leyes de propiedad intelectual y se distribuye bajo licencias que restringen su uso, copia, distribución y descompilación. No se puede reproducir ninguna parte de este producto o de este documento de ninguna forma ni por ningún medio sin la autorización previa por escrito de Sun y sus cedentes, si los hubiera. El software de otras empresas, incluida la tecnología de los tipos de letra, está protegido por la ley de copyright y con licencia de los distribuidores de Sun.

Algunas partes del producto pueden proceder de los sistemas Berkeley BSD, con licencia de la Universidad de California. UNIX es una marca registrada en los EE.UU. y otros países, bajo licencia exclusiva de X/Open Company, Ltd.

Sun, Sun Microsystems, el logotipo de Sun, docs.sun.com, AnswerBook, AnswerBook2 Sun Cluster, SunPlex, Sun Enterprise, Sun Enterprise 10000, Sun Enterprise SyMON, Sun Management Center, Solaris, Solaris Volume Manager, Sun StorEdge, Sun Fire, SPARCstation, OpenBoot y Solaris son marcas comerciales o marcas comerciales registradas de Sun Microsystems, Inc. en los EE.UU. y en otros países. Todas las marcas registradas SPARC se usan bajo licencia y son marcas comerciales o marcas registradas de SPARC International, Inc. en los EE.UU. y en otros países. Los productos con las marcas registradas de SPARC se basan en una arquitectura desarrollada por Sun Microsystems, Inc. ORACLE, Netscape

La interfaz gráfica de usuario OPEN LOOK y Sun™ fue desarrollada por Sun Microsystems, Inc. para sus usuarios y licenciatarios. Sun reconoce los esfuerzos pioneros de Xerox en la investigación y desarrollo del concepto de interfaces gráficas o visuales de usuario para la industria de la computación. Sun dispone de una licencia no exclusiva de Xerox para la interfaz gráfica de usuario de Xerox, que es extensiva a los licenciatarios de Sun que implementen la interfaz gráfica de usuario OPEN LOOK y que actúen conforme a los acuerdos de licencia por escrito de Sun.

Derechos gubernamentales de los EE.UU. – Software comercial. Los usuarios gubernamentales están sujetos al acuerdo de licencia estándar de Sun Microsystems, Inc. y a las disposiciones aplicables de la regulación FAR y sus suplementos.

ESTA DOCUMENTACIÓN SE PROPORCIONA "TAL CUAL". SE RENUNCIA A TODAS LAS CONDICIONES EXPRESAS O IMPLÍCITAS, REPRESENTACIONES Y GARANTÍAS, INCLUIDAS CUALQUIER GARANTÍA IMPLÍCITA DE COMERCIALIZACIÓN, ADECUACIÓN PARA UNA FINALIDAD DETERMINADA O DE NO CONTRAVENCIÓN, EXCEPTO EN AQUELLOS CASOS EN QUE DICHA RENUNCIA NO FUERA LEGALMENTE VÁLIDA.

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 Etats-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 Etats-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 Etats-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 REPOUDRE 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



Contenido

Prefacio 7

1 Introducción y visión general 13

Introducción al sistema Sun Cluster 14

Tres perspectivas sobre el sistema Sun Cluster 14

 Perspectiva del servicio y la instalación del hardware 15

 Perspectiva del administrador de sistemas 16

 Perspectiva de los desarrolladores de aplicaciones 17

Tareas de sistema de Sun Cluster 18

2 Conceptos clave para los proveedores de servicios de hardware 21

Componentes de software y hardware del sistema Sun Cluster 21

 Nodos del clúster 22

 Componentes de software para los miembros de hardware del clúster 23

 Dispositivos multisistema 24

 Iniciador múltiple SCSI 25

 Discos locales 26

 Medios extraíbles 26

 Interconexión de clúster 27

 Interfaces de red pública 28

 Sistemas cliente 28

 Dispositivos de acceso a la consola 28

 Consola administrativa 29

SPARC: Topologías de Sun Cluster para SPARC 30

 SPARC: Topología de par en clúster para SPARC 30

 SPARC: Topología de par+N para SPARC 31

SPARC: Topología N+1 (estrella) para SPARC	32
SPARC: Topología N*N (escalable) para SPARC	33
x86: Topologías de Sun Cluster para x86	34
x86: Topología de par en clúster para x86	35
3 Conceptos clave para los administradores de sistemas y los desarrolladores de aplicaciones	37
Interfaces administrativas	38
Hora del clúster	38
Estructura de alta disponibilidad	39
Cluster Membership Monitor	40
Mecanismo de recuperación rápida	41
Cluster Configuration Repository (CCR)	41
Dispositivos globales	42
ID de dispositivo y pseudocontrolador DID	42
Grupos de dispositivos de discos	43
Recuperación de fallos en grupos de dispositivos de disco	44
Grupos de dispositivos de disco multipuerto	45
Espacio de nombres global	47
Ejemplo de espacio de nombres local y global	48
Sistemas de archivos de clúster	48
Uso de los sistemas de archivos del clúster	49
Tipo de recurso de HASToragePlus	50
La opción de montaje Syncdir	50
Supervisión de la ruta de discos	51
Información general de DPM	51
Supervisión de las rutas del disco	53
Quórum y dispositivos del quórum	54
Acerca de los recuentos de votos del quórum	55
Acerca del aislamiento de fallos	56
Mecanismo de recuperación rápida para aislamiento de fallos	57
Acerca de las configuraciones del quórum	58
Cumplimiento de los requisitos de dispositivos de quórum	59
Mejores prácticas relacionadas con los dispositivos del quórum	59
Configuraciones de quórum recomendadas	60
Configuraciones de quórum atípicas	63
Configuraciones de quórum malas	63
Servicios de datos	65

Métodos de servicios de datos	67
Servicios de datos de recuperación de fallos	68
Servicios de datos escalables	68
Políticas de equilibrio de cargas	70
Configuración de retroceso	72
Supervisor de fallos del servicio de datos	72
Desarrollo de los nuevos servicios de datos	73
Características de los servicios escalables	73
API de servicio de datos y API de biblioteca de desarrollo de servicios de datos	74
Uso de interconexiones del clúster para el tráfico de servicios de datos	75
Tipos de recursos, grupos de recursos y recursos	76
Gestor de grupos de recursos (RGM)	77
Configuraciones y estados de los recursos y los grupos de recursos	77
Propiedades de los recursos y los grupos de recursos	79
Configuración de un proyecto de servicio de datos	79
Determinación de requisitos para la configuración de proyectos	81
Establecimiento de límites de memoria virtual por proceso	83
Casos de recuperación de fallos	83
Adaptadores de red públicos y Ruta múltiple de red de protocolo de Internet (IP)	89
SPARC: Compatibilidad con la reconfiguración dinámica	90
SPARC: Descripción general de la reconfiguración dinámica	91
SPARC: Consideraciones de clúster DR para los dispositivos CPU	91
SPARC: Consideraciones de clúster DR para la memoria	92
SPARC: Consideraciones de clúster DR para unidades de disco y cinta	92
SPARC: Consideraciones de clúster DR para los dispositivos del quórum	93
SPARC: Consideraciones de clúster DR para las interfaces interconectadas del clúster	93
SPARC: Consideraciones de clúster DR para las interfaces de red públicas	93
4 Preguntas más frecuentes	95
Preguntas más frecuentes sobre alta disponibilidad	95
FAQ sobre los sistemas de archivos	96
FAQ sobre gestión de volúmenes	97
FAQ sobre servicios de datos	98
FAQ sobre las redes públicas	99
FAQ sobre pertenencia al clúster	100

FAQ sobre almacenamiento en clúster	101
FAQ sobre interconexiones del clúster	101
FAQ sobre sistemas cliente	102
FAQ sobre la consola administrativa	102
FAQ sobre el concentrador de terminal y el procesador de servicio del sistema	103

Índice	107
---------------	------------

Prefacio

Sun™ Cluster Concepts Guide for Solaris OS contiene información de referencia sobre conceptos relacionados con el sistema SunPlex™ para sistemas basados en SPARC® y en x86.

Nota – En este documento el término x86” hace referencia a la familia de chips microprocesadores Intel de 32 bits y a los compatibles de AMD.

El sistema SuPlex incluye todos los componentes de hardware y software que componen la solución Sun Cluster.

Este documento está destinado a administradores de sistema con experiencia que tengan conocimientos sobre el software Sun Cluster. no se debe usar como guía de preventa o de planificación. Antes de leerlo, debe conocer su sistema y disponer del equipo y el software adecuados.

Para comprender los conceptos descritos en esta guía, debe tener conocimientos del sistema operativo Solaris™, además de experiencia con el software de gestión de volúmenes que se usa con el sistema Sun Cluster.

Nota – El software de Sun Cluster se ejecuta en dos plataformas: SPARC y x86. La información de este documento se aplica a ambas a menos que se especifique lo contrario de forma expresa en un capítulo, apartado, nota, elemento de lista, figura, tabla o ejemplo especiales.

Convenciones tipográficas

La tabla siguiente describe los cambios tipográficos utilizados en este manual.

TABLA P-1 Convenciones tipográficas

Tipo de letra o símbolo	Significado	Ejemplo
AaBbCc123	Los nombres de las órdenes, archivos, directorios y mensajes que aparecen en la pantalla del sistema	Edite el archivo <code>.login</code> . Utilice el comando <code>ls -a</code> para mostrar todos los archivos. <code>nombre_sistema% tiene correo.</code>
AaBbCc123	Lo que usted escribe, contrastado con la salida por la pantalla del sistema	<code>nombre_sistema% su</code> Contraseña:
<i>AaBbCc123</i>	Plantilla de la línea de órdenes: sustituir por un valor o nombre real	El comando necesario para eliminar un archivo es <code>rm filename</code> .
<i>AaBbCc123</i>	Títulos de los manuales, términos nuevos y palabras destacables	Consulte el capítulo 6 de la <i>Guía del usuario</i> . <i>Efectúe un análisis de modificaciones.</i> <i>No guarde el archivo.</i> [Observe que algunos elementos enfatizados, en línea, aparecen en negrita.]

Indicadores de los shells en los ejemplos de órdenes

La tabla siguiente muestra los indicadores predeterminados del sistema y los del superusuario para los shells C, Bourne y Korn.

TABLA P-2 Indicadores de intérprete de comandos

Shell	Preguntar
Indicador del shell C	machine_name%
Indicador de superusuario en el shell C	machine_name%
Indicador de los shells Bourne y Korn	\$
Indicador de superusuario en los shell Bourne y Korn	#

Documentación relacionada

Puede encontrar información sobre temas referentes a Sun Cluster en la documentación enumerada en la tabla siguiente. Toda la documentación sobre Sun Cluster está disponible en <http://docs.sun.com>.

Tema	Documentación
Visión general	<i>Sun Cluster para el sistema operativo Solaris: Visión general</i>
Conceptos	<i>Sun Cluster: Guía de conceptos para el SO Solaris</i>
Instalación y administración de hardware	<i>Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS</i> Guías de administración de hardware individuales
Instalación del software	<i>Software Sun Cluster: Guía de instalación para el sistema operativo Solaris</i>
Instalación y administración del servicio de datos	<i>Sun Cluster Data Services Planning and Administration Guide for Solaris OS</i> Guías de servicio de datos individuales
Desarrollo de los servicios de datos	<i>Sun Cluster: Guía del desarrollador de los servicios de datos del sistema operativo Solaris</i>
Administración de sistema	<i>Sun Cluster: Guía de administración del sistema para el SO Solaris</i>
Mensajes de error	<i>Sun Cluster Error Messages Guide for Solaris OS</i>
Referencias sobre las órdenes y las funciones	<i>Sun Cluster Reference Manual for Solaris OS</i>

Si desea ver una lista completa de la documentación sobre Sun Cluster, consulte las notas de la versión del software de Sun Cluster en <http://docs.sun.com>.

Acceso a la documentación de Sun en línea

La sede web docs.sun.comSM permite acceder a la documentación técnica de Sun en línea. Puede explorar el archivo docs.sun.com, buscar el título de un manual o un tema específicos. La URL es <http://docs.sun.com>.

Solicitud de documentación de Sun

Sun Microsystems ofrece una seleccionada documentación impresa sobre el producto. Si desea conocer una lista de documentos y cómo pedirlos, consulte "Adquirir documentación impresa" en <http://docs.sun.com>.

Obtención de ayuda

Si tiene problemas durante la instalación o el uso del sistema Sun Cluster, póngase en contacto con su proveedor de servicios y déle la información siguiente:

- Su nombre y dirección de correo electrónico (si estuviera disponible)
- El nombre, dirección y número de teléfono de su empresa
- Los modelos y números de serie de sus sistemas
- El número de versión del sistema operativo; por ejemplo Solaris 9
- El número de versión del software Sun Cluster (por ejemplo, 3.1 8/05)

Use las órdenes siguientes para reunir información sobre todos los nodos del sistema para el proveedor de servicio.

Comando	Función
<code>prtconf -v</code>	Muestra el tamaño de la memoria del sistema y ofrece información sobre los dispositivos periféricos
<code>psrinfo -v</code>	Muestra información acerca de los procesadores

Comando	Función
showrev -p	Indica las modificaciones instaladas
SPARC: prtdiag -v	Muestra información de diagnóstico del sistema
scinstall -pv	Muestra información sobre la versión y el paquete de Sun Cluster.
scstat	Ofrece una instantánea del estado del clúster.
scconf -p	Muestra información de configuración sobre el clúster
scrgadm -p	Muestra información sobre los recursos instalados, los grupos y tipos de recursos

Tenga también a mano el contenido del archivo `/var/adm/messages`.

Formación sobre el producto

Sun Microsystems ofrece cursos de formación sobre diversas tecnologías de Sun mediante una serie de cursos impartidos por instructores y cursos para que el alumno vaya a su ritmo. Para obtener información acerca de los cursos de formación que Sun pone a su disposición y para inscribirse en ellos, visite la Web de formación de Sun Microsystems en <http://training.sun.com/>.

Introducción y visión general

El sistema Sun Cluster es una solución integrada de hardware y software de Sun Cluster que se utiliza para crear servicios de alta disponibilidad y escalabilidad.

En Sun Cluster: Guía de conceptos para el SO Solaris se proporciona la información conceptual que necesitan los usuarios principales de la documentación de Sun Cluster, entre los que se incluyen:

- Proveedores de servicio que instalan y reparan hardware de clúster
- Administradores de sistema que instalan, configuran y administran el software Sun Cluster
- Desarrolladores de aplicaciones que desarrollan servicios de recuperación de fallos y servicios escalables para aplicaciones que no se incluyen actualmente en el producto Sun Cluster

Use este manual junto con la documentación de Sun Cluster para obtener una visión completa del sistema Sun Cluster.

En este capítulo

- Proporciona una introducción y una descripción general del sistema Sun Cluster
- Describe las distintas perspectivas de los usuarios sobre Sun Cluster
- Identifica los conceptos clave que debe conocer antes de trabajar con el sistema Sun Cluster
- Relaciona los conceptos clave con la documentación de Sun Cluster, que incluye procedimientos e información relacionada
- Correlaciona tareas relacionadas del clúster con la documentación que contiene procedimientos usados para realizar esas tareas

Introducción al sistema Sun Cluster

El sistema Sun Cluster convierte el sistema operativo Solaris en un sistema operativo en clúster. Un clúster, o plex, es una colección de nodos informáticos acoplados indirectamente que proporcionan una vista de cliente único de servicios de red o de aplicaciones, incluidos bases de datos, servicios de web y servicios de archivos.

Cada nodo del clúster es un servidor autónomo que ejecuta sus propios procesos. Éstos se comunican entre sí para formar lo que se parece (a un cliente de red) como un sólo sistema que proporciona de forma cooperativa aplicaciones, recursos de sistema y datos a usuarios.

Un clúster ofrece, respecto a los sistemas tradicionales de servidor único, varias ventajas que incluyen soporte para servicios de protección contra fallos y escalabilidad, capacidad para crecimiento modular y un precio básico económico en comparación a los sistemas de hardware tolerantes a fallos.

Los objetivos del sistema Sun Cluster son:

- Reducir o eliminar el tiempo de inactividad del sistema debido a los fallos de software o hardware
- Asegurar la disponibilidad de los datos y aplicaciones a los usuarios finales, cualquiera que sea el tipo de fallo que normalmente dejaría inactivo un sistema de servidor único
- Aumentar el rendimiento de las aplicaciones permitiendo escalar procesadores adicionales con solo agregar más nodos al clúster
- Proporcionar una mejor disponibilidad del sistema al permitirle realizar el mantenimiento sin tener que apagar todo el clúster

Para obtener más información acerca de la tolerancia ante fallos y de la alta disponibilidad, consulte “Aplicaciones con una gran disponibilidad con Sun Cluster” de *Sun Cluster para el sistema operativo Solaris: Visión general*.

Consulte “[Preguntas más frecuentes sobre alta disponibilidad](#)” en la página 95 para ver preguntas y respuestas sobre la alta disponibilidad.

Tres perspectivas sobre el sistema Sun Cluster

Esta sección describe tres perspectivas diferentes sobre el sistema Sun Cluster, así como los conceptos clave y la documentación relevante para cada una de ellas. Estas tres perspectivas son típicas de los siguientes profesionales:

- Personal de instalación y reparación de hardware
- Administradores del sistema
- Desarrolladores de aplicaciones

Perspectiva del servicio y la instalación del hardware

Para los profesionales que se dedican al servicio de hardware, el sistema Sun Cluster se asemeja a un conjunto de hardware listo para utilizar que incluye servidores, redes y dispositivos de almacenamiento. Estos componentes están todos interconectados de manera que todos tengan un recambio y no exista un punto débil para los fallos.

Conceptos clave – Hardware

Los profesionales del servicio de hardware deben comprender los siguientes conceptos relacionados con los clústeres.

- Configuraciones de hardware de clúster y cableado
- Instalación y reparación (agregar, eliminar, sustituir):
 - Componentes de la interfaz de red (adaptadores, uniones, cables)
 - Tarjetas interfaz de disco
 - Matrices de disco
 - Unidades de disco
 - La consola de administración y el dispositivo de acceso a la consola
- Configuración de la consola de administración y del dispositivo de acceso a la consola

Más información conceptual sobre el hardware

Los apartados siguientes contienen material relacionado con los conceptos clave anteriores:

- “Nodos del clúster ” en la página 22
- “Dispositivos multisistema ” en la página 24
- “Discos locales ” en la página 26
- “Interconexión de clúster ” en la página 27
- “Interfaces de red pública ” en la página 28
- “Sistemas cliente ” en la página 28
- “Consola administrativa ” en la página 29
- “Dispositivos de acceso a la consola ” en la página 28
- “SPARC: Topología de par en clúster para SPARC ” en la página 30
- “SPARC: Topología N+1 (estrella) para SPARC ” en la página 32

Documentación sobre Sun Cluster para profesionales del hardware

La siguiente documentación sobre Sun Cluster incluye procedimientos e información relacionada con los conceptos de servicio de hardware:

Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS

Perspectiva del administrador de sistemas

Para los administradores de sistemas, el sistema Sun Cluster es un conjunto de servidores (nodos) enlazados que comparten dispositivos de almacenamiento. Para el administrador de tareas, lo más importante es el software que realiza tareas específicas:

- Software para clústers especializado que se integra con el de Solaris para supervisar la conectividad entre los nodos del clúster
- Software especializado que supervisa el buen funcionamiento de los programas de aplicación del usuario que se ejecutan en sus nodos del clúster
- Software de gestión de volúmenes que configura y administra discos
- Software de clúster especializado que permite a todos los nodos acceder a todos los dispositivos de almacenamiento, incluso a aquellos no conectados directamente a discos
- Software de clúster especializado que permite a los archivos aparecer en todos los nodos como si estuvieran alojados localmente en él

Conceptos clave– Administración de sistemas

Es importante que los administradores de sistemas comprendan los conceptos y procesos siguientes:

- La interacción entre los componentes de hardware y software
- El flujo general sobre cómo instalar y configurar el clúster, que incluye:
 - Instalación del sistema operativo Solaris
 - Instalar y configurar Sun Cluster
 - Instalación y configuración de un gestor de volúmenes
 - Instalación y configuración del software de aplicación para que esté preparado para el clúster
 - Instalación y configuración del software de servicio de datos de Sun Cluster
- Procedimientos administrativos del clúster para agregar, eliminar, sustituir y reparar componentes de hardware y software del clúster
- Modificaciones de la configuración para mejorar el rendimiento

Más información conceptual sobre el administrador de sistemas

Los apartados siguientes contienen material relacionado con los conceptos clave anteriores:

- “Interfaces administrativas ” en la página 38
- “Hora del clúster ” en la página 38
- “Estructura de alta disponibilidad ” en la página 39
- “Dispositivos globales ” en la página 42
- “Grupos de dispositivos de discos ” en la página 43
- “Espacio de nombres global ” en la página 47
- “Sistemas de archivos de clúster ” en la página 48
- “Supervisión de la ruta de discos ” en la página 51
- “Acerca del aislamiento de fallos ” en la página 56
- “Servicios de datos ” en la página 65

Documentación sobre Sun Cluster para administradores de sistema

Los siguientes documentos sobre Sun Cluster incluyen procedimientos e información relacionada con los conceptos de administración de sistemas:

- *Software Sun Cluster: Guía de instalación para el sistema operativo Solaris*
- *Sun Cluster: Guía de administración del sistema para el SO Solaris*
- *Sun Cluster Error Messages Guide for Solaris OS*
- *Notas de la versión de Sun Cluster 3.1 8/05 para SO Solaris*
- *Sun Cluster 3.0-3.1 Release Notes Supplement*

Perspectiva de los desarrolladores de aplicaciones

El sistema Sun Cluster proporciona *servicios de datos* para aplicaciones como Oracle, NFS, DNS, Sun™ Java System Web Server, Apache Web Server (en sistemas basados en SPARC) y Sun Java System Directory Server. Los servicios de datos se crean al configurar aplicaciones listas para usar para que se ejecuten bajo el control del software Sun Cluster. El software Sun Cluster proporciona archivos de configuración y métodos de gestión para iniciar, detener y supervisar las aplicaciones. Si necesita crear un nuevo servicio escalable o de recuperación ante fallos, puede usar la interfaz de programación de aplicaciones (API) de Sun Cluster y la API de habilitación de servicios de datos (DSET API) para desarrollar los archivos de configuración necesarios y los métodos de gestión que faculden a su aplicación para ejecutarse como servicio de datos en el clúster.

Conceptos clave – Desarrollo de aplicaciones

Los desarrolladores de aplicaciones deben conocer los siguientes aspectos:

- Las características de la aplicación, para determinar si ésta puede ejecutarse como servicio de recuperación de fallos o de escalabilidad de datos.
- La API, la DSET API de Sun Cluster y el servicio de datos “genérico”. Los desarrolladores deben determinar qué herramienta es más adecuada para ellos a la hora de escribir programas o secuencias de comandos para configurar sus aplicaciones para el entorno de clústeres.

Más información conceptual sobre el desarrollador de aplicaciones

Los apartados siguientes contienen material relacionado con los conceptos clave anteriores:

- “Servicios de datos ” en la página 65
- “Tipos de recursos, grupos de recursos y recursos ” en la página 76
- Capítulo 4

Documentación sobre Sun Cluster para desarrolladores de aplicaciones

Los siguientes documentos sobre Sun Cluster incluyen procedimientos e información relacionada con los conceptos de desarrollador de aplicaciones:

- *Sun Cluster: Guía del desarrollador de los servicios de datos del sistema operativo Solaris*
- *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*

Tareas de sistema de Sun Cluster

Todas las tareas de sistema de Sun Cluster requieren algún tipo de conocimientos conceptuales. La tabla siguiente incluye una visión general de las tareas y la documentación que describe sus pasos. El apartado de conceptos de este manual describe la correlación entre los conceptos y estas tareas.

TABLA 1-1 Mapa de tareas: correspondencia entre las tareas del usuario y la documentación

Tarea	Instrucciones
Instalar el hardware del clúster	<i>Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS</i>

TABLA 1-1 Mapa de tareas: correspondencia entre las tareas del usuario y la documentación (Continuación)

Tarea	Instrucciones
Instalar el software de Solaris en el clúster	<i>Software Sun Cluster: Guía de instalación para el sistema operativo Solaris</i>
SPARC: Instalar el software Sun™ Management Center	<i>Software Sun Cluster: Guía de instalación para el sistema operativo Solaris</i>
Instalar y configurar el software Sun Cluster	<i>Software Sun Cluster: Guía de instalación para el sistema operativo Solaris</i>
Instalar y configurar el software de gestión de volúmenes	<i>Software Sun Cluster: Guía de instalación para el sistema operativo Solaris</i> La documentación del gestor de volúmenes
Instalar y configurar los servicios de datos de Sun Cluster	<i>Sun Cluster Data Services Planning and Administration Guide for Solaris OS</i>
Reparar el hardware del clúster	<i>Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS</i>
Administrar el software Sun Cluster	<i>Sun Cluster: Guía de administración del sistema para el SO Solaris</i>
Administrar el software de la gestión de volúmenes	<i>Sun Cluster: Guía de administración del sistema para el SO Solaris</i> y la documentación sobre gestión de volúmenes
Administrar el software de las aplicaciones	La documentación de su aplicación
Identificar los problemas y las acciones de usuario sugeridas	<i>Sun Cluster Error Messages Guide for Solaris OS</i>
Crear un servicio de datos nuevo	<i>Sun Cluster: Guía del desarrollador de los servicios de datos del sistema operativo Solaris</i>

Conceptos clave para los proveedores de servicios de hardware

Este capítulo describe los conceptos clave relacionados con los componentes de hardware de una configuración del sistema Sun Cluster. Los temas que se tratan son los siguientes:

- “Nodos del clúster ” en la página 22
- “Dispositivos multisistema ” en la página 24
- “Discos locales ” en la página 26
- “Medios extraíbles ” en la página 26
- “Interconexión de clúster ” en la página 27
- “Interfaces de red pública ” en la página 28
- “Sistemas cliente ” en la página 28
- “Dispositivos de acceso a la consola ” en la página 28
- “Consola administrativa ” en la página 29
- “SPARC: Topologías de Sun Cluster para SPARC ” en la página 30
- “x86: Topologías de Sun Cluster para x86 ” en la página 34

Componentes de software y hardware del sistema Sun Cluster

Esta información está dirigida principalmente a los proveedores de servicio de hardware a quienes ayuda a entender la relación entre los componentes de hardware para mejor instalar, configurar o reparar el hardware del clúster. Asimismo, a los administradores del sistema del clúster esta información les puede resultar útil para instalar, configurar y administrar el software del clúster.

Un clúster consta de varios componentes de hardware, entre los que se incluyen:

- Nodos de clúster con discos locales (no compartidos)
- Almacenamiento multisistema (discos compartidos entre nodos)

- Medios extraíbles (cintas y CD-ROM)
- Interconexión del clúster
- Interfaces de red públicas
- Sistemas cliente
- Consola de administración
- Dispositivos de acceso a la consola

El sistema Sun Cluster le permite combinar estos componentes en una serie de configuraciones. Las siguientes secciones describen estas configuraciones.

- [“SPARC: Topologías de Sun Cluster para SPARC ” en la página 30](#)
- [“x86: Topologías de Sun Cluster para x86 ” en la página 34](#)

Para ver una ilustración de una configuración de ejemplo de un clúster con dos nodos, consulte “Entorno de hardware de Sun Cluster” de *Sun Cluster para el sistema operativo Solaris: Visión general*.

Nodos del clúster

Un nodo de clúster es un equipo en el que se ejecuta el sistema operativo Solaris y el software Sun Cluster. Un nodo de clúster puede ser un miembro actual del clúster (un *miembro del clúster*) o bien un miembro potencial.

- SPARC: el software Sun Cluster admite hasta 16 nodos en un clúster. Consulte [“SPARC: Topologías de Sun Cluster para SPARC ” en la página 30](#) para conocer las configuraciones de nodos admitidas.
- x86: el software Sun Cluster admite dos nodos en un clúster. Consulte [“x86: Topologías de Sun Cluster para x86 ” en la página 34](#) para conocer las configuraciones de nodos admitidas.

Los nodos del clúster normalmente están conectados a uno o más dispositivos multisistema, aquéllos que no lo están usan el sistema de archivos del clúster para acceder a los dispositivos multisistema. Por ejemplo, una configuración de servicios escalables permite que los nodos atiendan peticiones sin estar directamente conectados a los dispositivos multisistema.

Además, los nodos en configuraciones de bases de datos paralelas comparten acceso simultáneo a todos los discos.

- Consulte [“Dispositivos multisistema ” en la página 24](#) para obtener información acerca del acceso simultáneo a discos.
- Consulte [“SPARC: Topología de par en clúster para SPARC ” en la página 30](#) y [“x86: Topología de par en clúster para x86 ” en la página 35](#) para obtener información acerca de las configuraciones de bases de datos paralelas.

Todos los nodos del clúster están agrupados bajo un nombre común (el nombre del clúster) que se utiliza para acceder a éste y gestionarlo.

Los adaptadores de redes públicas conectan los nodos a éstas para ofrecer acceso de cliente al clúster.

Los miembros del clúster se comunican con otros nodos del clúster a través de una o varias redes independientes. Este conjunto de redes independientes físicamente se conoce como *interconexión del clúster*.

Todos los nodos del clúster reciben información de cuándo un nodo se une o deja el clúster, conocen también los recursos que se están ejecutando localmente así como los que se están ejecutando en los otros nodos del clúster.

Los nodos del mismo clúster deben tener capacidades de procesamiento, memoria y E/S similares para permitir que la recuperación de fallos se produzca sin que haya una degradación importante en el rendimiento. Debido a la función de recuperación de fallos, cada nodo debe tener una capacidad de exceso suficiente para asimilar la carga de trabajo de todos los nodos para los que esté configurado como nodo secundario o de copia de respaldo.

Cada nodo ejecuta su propio sistema de archivos raíz individual (/).

Componentes de software para los miembros de hardware del clúster

Si desea trabajar como miembro de un clúster, un nodo debe tener el siguiente software instalado:

- Sistema operativo Solaris
- Software Sun Cluster
- Aplicación de servicio de datos
- Gestión de volúmenes (Gestor de volúmenes de Solaris™ o VERITAS Volume Manager)

Una excepción es la configuración que usa una matriz redundante de hardware de discos independientes (RAID). Esta configuración puede que no requiera un gestor de volúmenes de software como Gestor de volúmenes de Solaris o VERITAS Volume Manager.

- Consulte *Software Sun Cluster: Guía de instalación para el sistema operativo Solaris* para obtener información acerca de cómo instalar el sistema operativo Solaris, Sun Cluster y el software de gestión de volúmenes.
- Consulte *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* para obtener información acerca de cómo instalar y configurar los servicios de datos.
- Consulte el [Capítulo 3](#) para obtener información conceptual acerca de los componentes de software previos.

La siguiente figura proporciona una descripción general de los componentes de software que funcionan juntos para crear el entorno de software de Sun Cluster.

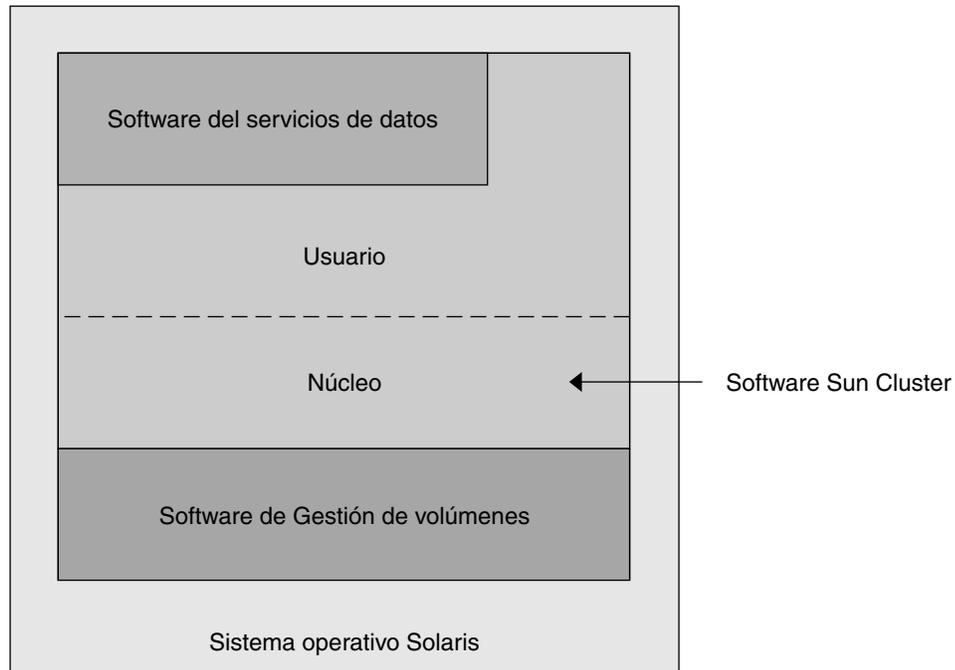


FIGURA 2-1 Relaciones generales entre los componentes del software Sun Cluster

Consulte el [Capítulo 4](#) para ver las preguntas y respuestas acerca de los miembros del clúster.

Dispositivos multisistema

Los discos que pueden conectarse a más de un nodo simultáneamente se denominan multisistema. En el entorno de Sun Cluster, el almacenamiento multisistema hace que los discos tengan una gran disponibilidad. El software Sun Cluster requiere almacenamiento multisistema para que los clústeres de dos nodos establezcan el quórum. Los clústeres que tengan más de dos nodos no requieren dispositivos del quórum. Para obtener más información acerca del quórum, consulte ["Quórum y dispositivos del quórum"](#) en la [página 54](#).

Los dispositivos multisistema poseen las características siguientes.

- Tolerancia a los fallos de un único nodo.
- Capacidad para almacenar los datos de las aplicaciones, los archivos binarios de las aplicaciones y los archivos de configuración.

- Protección frente a los fallos en los nodos. Si los clientes solicitan los datos a través de un nodo que falla, las solicitudes se desvían para usar otro nodo que tenga una conexión directa con los mismos discos.
- Acceso global a través de un nodo principal que “controle” los discos o acceso directo simultáneo a través de las rutas locales. La única aplicación que usa el acceso simultáneo directo actualmente es Oracle Real Application Clusters Guard.

Un gestor de volúmenes proporciona configuraciones duplicadas o RAID-5 para la redundancia de los dispositivos multisistema. Actualmente, Sun Cluster es compatible con Gestor de volúmenes de Solaris y con VERITAS Volume Manager, que está disponible para usarlo sólo en clústeres basados en SPARC, como gestores de volúmenes, y el controlador de hardware RDAC RAID-5 en varias plataformas RAID de hardware.

Al combinar los dispositivos multisistema con los discos de duplicación y la separación de datos, se consigue protección contra el fallo de los nodos y el de los discos individuales.

Consulte el [Capítulo 4](#) para ver las preguntas y respuestas acerca del almacenamiento multisistema.

Iniciador múltiple SCSI

Este apartado se aplica sólo a dispositivos de almacenamiento SCSI y no a los de fibra óptica que usan los dispositivos multisistema.

En un servidor autónomo, el nodo servidor controla las actividades del bus SCSI porque el circuito adaptador del sistema SCSI se conecta a ese servidor en un bus SCSI determinado. Este circuito adaptador del sistema SCSI se conoce con el nombre de *iniciador SCSI*. Este circuito inicia todas las actividades de este bus SCSI. La dirección SCSI predeterminada de los adaptadores en el sistema Sun es 7.

Las configuraciones del clúster comparten el almacenamiento entre varios nodos del servidor, usando los dispositivos multisistema. Cuando el almacenamiento del clúster consta de dispositivos SCSI diferenciales o de dispositivos con un único extremo, la configuración se denomina iniciador múltiple SCSI. Tal como implica esta terminología, existe más de un iniciador SCSI en el bus.

La especificación SCSI requiere que cada dispositivo de un bus SCSI tenga una única dirección SCSI. (El adaptador del sistema también es un dispositivo del bus SCSI.) La configuración predeterminada del hardware en un entorno de iniciador múltiple provoca un conflicto debido a que de forma predeterminada todos los adaptadores del sistema SCSI tienen el valor 7.

Para resolver este conflicto, se debe dejar uno de los adaptadores de cada bus SCSI con la dirección SCSI 7 y se debe configurar los otros adaptadores como direcciones SCSI no utilizadas. La planificación correcta dictamina que estas direcciones SCSI “no

utilizadas” incluyan tanto direcciones usadas actualmente como no utilizadas. Un ejemplo de direcciones no utilizadas en el futuro es la incorporación del almacenamiento instalando unidades nuevas en ranuras de unidad vacías.

En la mayoría de configuraciones, la dirección SCSI disponible para un segundo adaptador de sistema es 6.

Se pueden cambiar las direcciones SCSI seleccionadas de esos adaptadores del sistema mediante una de las siguientes herramientas de configuración de la propiedad `scsi-initiator-id`:

- `eeeprom(1M)`
- PROM de OpenBoot en un sistema basado en la plataforma SPARC
- La utilidad SCSI que se puede ejecutar opcionalmente después de que arranque la BIOS en un sistema basado en la plataforma x86

Se puede configurar esta propiedad tanto globalmente para un nodo como individualmente para cada adaptador. Las instrucciones para configurar un único `scsi-initiator-id` para cada adaptador de sistema SCSI se incluyen en *Sun Cluster 3.0-3.1 With SCSI JBOD Storage Device Manual for Solaris OS*.

Discos locales

Los discos locales son aquellos que sólo están conectados a un nodo individual. Los discos locales no están, en consecuencia, protegidos ante posibles fallos de los nodos (no están configurados para una alta disponibilidad). Sin embargo, todos los discos (también los locales), están incluidos en el espacio de nombre global y están configurados como *dispositivos globales*. Por lo tanto, los discos en sí son visibles desde todos los nodos del clúster.

Puede hacer que los sistemas de archivos de los discos locales estén disponibles para otros nodos. Para ello, debe colocarlos en un punto de montaje global. Si el nodo que tiene montado actualmente uno de estos sistemas de archivos globales falla, el resto de nodos pierde acceso a ese sistema de archivos. Usar un gestor de volúmenes permite duplicar estos discos de forma que un fallo no pueda provocar que los sistemas de archivos dejen de estar accesibles, aunque los gestores de volúmenes no protejan contra los fallos de los nodos.

Consulte la sección [“Dispositivos globales”](#) en la página 42 para obtener más información sobre los dispositivos globales.

Medios extraíbles

El clúster admite soportes extraíbles, como unidades de cinta y de CD-ROM. En general, estos dispositivos se instalan, configuran y se atienden del mismo modo que se haría en un entorno que no estuviera configurado en clúster. Estos servicios están

configurados como dispositivos globales en Sun Cluster, por lo que a cada dispositivo se puede acceder desde cualquier nodo del clúster. Consulte *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS* para obtener información acerca de la instalación y la configuración de los medios extraíbles.

Consulte la sección “Dispositivos globales ” en la página 42 para obtener más información sobre los dispositivos globales.

Interconexión de clúster

La *interconexión del clúster* es la configuración física de dispositivos que se utiliza para transferir comunicaciones privadas del clúster y de servicios de datos entre los nodos del clúster. Debido a que la interconexión se utiliza ampliamente para comunicaciones privadas del clúster, puede limitar el rendimiento.

La interconexión del clúster sólo puede conectar los distintos nodos del clúster. El modelo de seguridad de Sun Cluster asume que sólo los nodos del clúster tienen acceso físico a la interconexión del clúster.

Todos los nodos deben estar conectados por la interconexión del clúster a través de, al menos, dos redes independientes físicamente redundantes o rutas de acceso, para evitar que exista un único punto de fallo. Entre dos nodos cualesquiera puede tener varias redes independientes físicamente (de dos a seis).

La interconexión del clúster se compone de tres componentes del hardware: adaptadores, uniones y cables. La lista siguiente describe cada uno de estos componentes del hardware.

- **Adaptadores:** las tarjetas de interfaz de red que residen en cada nodo del clúster. Sus nombres se construyen a partir de un nombre de dispositivo seguido inmediatamente por un número de unidad física, por ejemplo: qfe2. Algunos adaptadores sólo tienen una conexión de red física, aunque otros como la tarjeta qfe tienen varias conexiones físicas. algunos también contienen interfaces de red y de almacenamiento.

Un adaptador de red con varias interfaces podría convertirse en un único punto de fallo si falla todo el adaptador. Para una máxima disponibilidad, se debe planificar el clúster de manera que la única ruta entre dos nodos no dependa de un único adaptador de red individual.

- **Uniones:** los conmutadores que residen fuera de los nodos del clúster. Las uniones llevan a cabo funciones de transporte y conmutación para permitir conectar más de dos nodos simultáneamente. En un clúster de dos nodos no hacen falta uniones porque los nodos pueden conectarse directamente entre sí con cables físicos redundantes conectados a adaptadores redundantes de cada nodo. Las configuraciones superiores a dos nodos en general requieren uniones.
- **Cables:** la conexión física que se instala entre dos adaptadores de red o entre un adaptador y una unión.

Consulte el [Capítulo 4](#) para ver las preguntas y respuestas acerca de las interconexiones del clúster.

Interfaces de red pública

Los clientes se conectan al clúster a través de interfaces de red pública. Todas las tarjetas adaptadoras de red pueden conectarse a una o más redes públicas, de acuerdo con las interfaces de hardware que la tarjeta tenga. Los nodos pueden configurarse para que incluyan múltiples tarjetas interfaz de red pública a fin de que varias estén activas y se utilicen en caso de recuperación de fallos como respaldo entre ellas. Si uno de los adaptadores falla, el software Ruta múltiple de red de protocolo de Internet (IP) se ejecuta para efectuar la recuperación ante fallos de la interfaz defectuosa en otro adaptador del grupo.

No hay consideraciones de hardware especiales relacionadas con la agrupación en clúster para las interfaces de red públicas.

Consulte el [Capítulo 4](#) para ver las preguntas y respuestas acerca de las redes públicas.

Sistemas cliente

Los sistemas cliente son las estaciones de trabajo u otros servidores que acceden al clúster a través de la red pública. Los programas cliente usan los datos u otros servicios proporcionados por las aplicaciones servidor que se ejecutan en el clúster.

Los sistemas cliente no son de alta disponibilidad. Los datos y aplicaciones del clúster son de alta disponibilidad.

Consulte el [Capítulo 4](#) para ver las preguntas y respuestas acerca de los sistemas cliente.

Dispositivos de acceso a la consola

Es necesario que disponga de acceso a la consola para todos los nodos del clúster. Para obtener acceso a la consola, utilice uno de los siguientes dispositivos:

- El concentrador de terminal que adquirió junto con el hardware del clúster
- El procesador de servicio del sistema (SSP) en los servidores Sun Enterprise E10000 (para clústeres basados en SPARC)
- El controlador del sistema en los servidores Sun Fire™ (también para clústeres basados en SPARC)
- Otro dispositivo que pueda acceder a `ttya` en cada nodo

Sun sólo dispone de un concentrador de terminal admitido y su uso es opcional. El concentrador de terminal permite el acceso a `/dev/console` en todos los nodos a través de una red TCP/IP. El resultado es el acceso a nivel de consola para todos los nodos desde una estación de trabajo remota desde cualquier lugar de la red.

El procesador de servicio del sistema (SSP) proporciona acceso a la consola para los servidores Sun Enterprise E1000. El SSP es un equipo de una red Ethernet que está configurado para admitir el servidor Sun Enterprise E1000. El SSP es la consola administrativa del servidor Sun Enterprise E1000. Gracias a la función Sun Enterprise E10000 Network Console, cualquier estación de trabajo de la red puede abrir una sesión de consola de sistema.

Otros métodos de acceso a la consola incluyen otros concentradores de terminal, acceso al puerto serie `tip` (1) desde otro nodo y terminales mudos. Puede usar teclados y monitores de Sun™ u otros dispositivos de puerto serie si el proveedor de servicio de hardware los admite.

Consola administrativa

Puede usar una estación de trabajo UltraSPARC® dedicada o un servidor Sun Fire V65x, conocida como *consola administrativa*, para administrar el clúster activo. Normalmente, se instala y ejecuta un software de herramienta administrativa, como el panel de control del clúster (CCP) y el módulo de Sun Cluster para el producto Sun Management Center (para utilizarlo sólo con clústeres basados en SPARC), en la consola administrativa. Si utiliza `cconsole` en CCP se le permitirá conectar a más de un nodo simultáneamente. Para obtener más información acerca del uso de CCP, consulte el Capítulo 1, “Introducción a la administración de Sun Cluster” de *Sun Cluster: Guía de administración del sistema para el SO Solaris*.

La consola de administración normalmente no es un nodo del clúster; se acostumbra a usar para obtener acceso remoto a los nodos del clúster, bien a través una red pública bien, opcionalmente, a través de un concentrador de terminales situado en una red. Si el clúster cuenta con una plataforma Sun Enterprise E10000, deberá iniciar sesión en SSP desde la consola administrativa y conectarse usando el comando `netcon (1M)`.

Normalmente los nodos se configuran sin monitor. Entonces se accede a la consola del nodo desde una sesión `telnet` desde la consola administrativa. La consola de administración está conectada a un concentrador de terminal y éste a su vez al puerto serie del nodo. En el caso de un servidor Sun Enterprise E1000, la conexión se realiza desde un procesador de servicio del sistema. Consulte “[Dispositivos de acceso a la consola](#)” en la [página 28](#) para obtener más información.

Sun Cluster no requiere una consola administrativa dedicada, pero si usa una, obtendrá los siguientes beneficios:

- Permite la gestión centralizada del clúster ya que agrupa herramientas de consola y gestión en la misma máquina
- Ofrece al proveedor de servicio de hardware una resolución de problemas potencialmente más rápida

Consulte el [Capítulo 4](#) para ver las preguntas y respuestas acerca de la consola administrativa.

SPARC: Topologías de Sun Cluster para SPARC

Una topología es el esquema de conexión que une los nodos del clúster con las plataformas de almacenamiento que se usan en un entorno de Sun Cluster. El software Sun Cluster es compatible con cualquier topología que cumpla las siguientes directrices.

- Un entorno de Sun Cluster que esté compuesto por sistemas basados en SPARC admite como máximo 16 nodos en un clúster, con independencia de la configuración de almacenamiento que se implemente.
- Un dispositivo de almacenamiento compartido puede conectarse a tantos nodos como se admitan.
- Los dispositivos de almacenamiento compartido no necesitan conectarse con todos los nodos del clúster. Sin embargo, estos dispositivos de almacenamiento deben conectarse a un mínimo de dos nodos.

El software Sun Cluster no requiere que se configure un clúster usando topologías específicas. Las topologías que se describen a continuación se incluyen para proporcionar el vocabulario que permita discutir un esquema de conexión del clúster. Estas topologías son esquemas de conexión típicos.

- Par en clúster
- Par+N
- N+1 (estrella)
- N*N (escalable)

Los apartados siguientes incluyen diagramas de ejemplo para cada topología.

SPARC: Topología de par en clúster para SPARC

Una topología de par en clúster consiste en dos o más nodos que funcionan en una misma estructura administrativa de clúster. En esta configuración sólo se produce recuperación de fallos entre pares. Sin embargo, todos los nodos están conectados mediante la interconexión del clúster y funcionan bajo el control del software de Sun Cluster. Esta topología se podría usar para ejecutar una aplicación de base de datos paralela en un par y una aplicación de recuperación de fallos o de escalabilidad en otro par.

Usando el sistema de archivos en clúster, podrá tener también una configuración de dos pares. Más de dos nodos pueden ejecutar un servicio escalable o una base de datos paralela, incluso aunque los nodos no estén directamente conectados a los discos que almacenan los datos de la aplicación.

La figura siguiente ilustra una configuración de par en clúster.

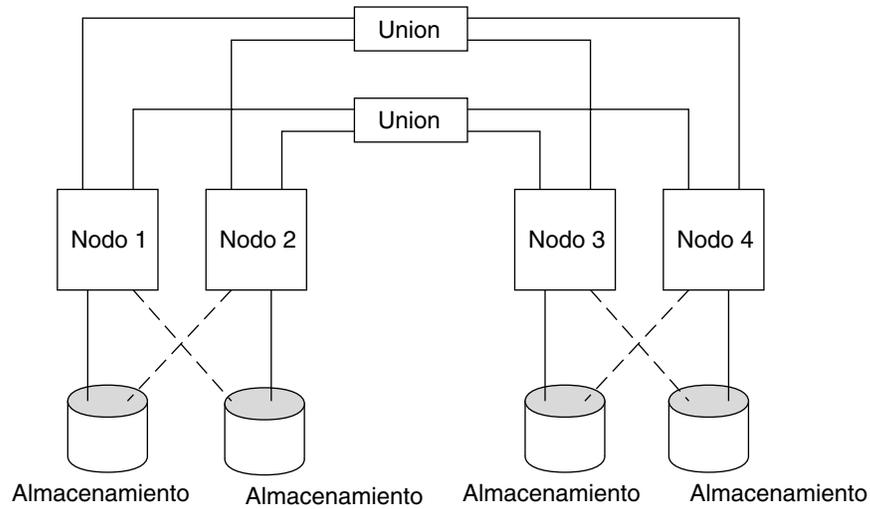


FIGURA 2-2 SPARC: Topología de par en clúster

SPARC: Topología de par+N para SPARC

La topología par+n incluye un par de nodos conectados directamente al almacenamiento compartido y un conjunto adicional de nodos que usa la interconexión del clúster para acceder al almacenamiento compartido; éstos no tienen conexión directa.

La figura siguiente muestra una topología par+n en que dos de los cuatro nodos (nodo 3 y 4) usan la interconexión del clúster para acceder al almacenamiento. Esta configuración puede ampliarse para que incluya nodos adicionales que no tengan acceso directo al almacenamiento compartido.

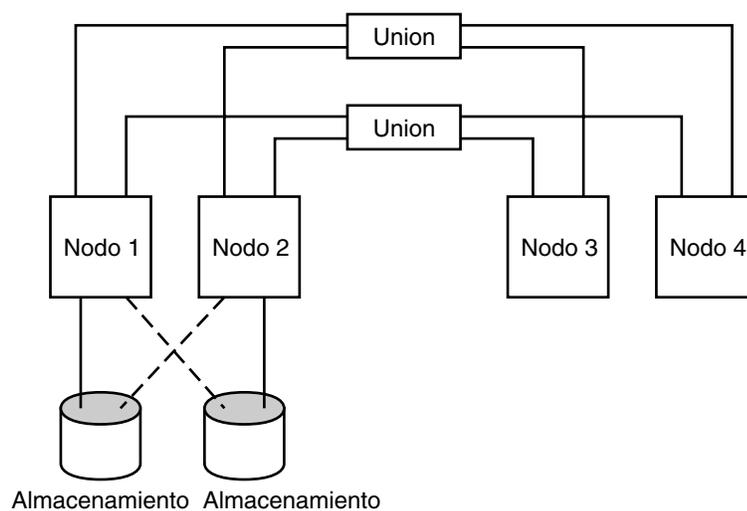


FIGURA 2-3 Topología de par+n

SPARC: Topología N+1 (estrella) para SPARC

Las topologías n+1 incluyen varios nodos primarios y uno secundario que no se deben configurar de forma idéntica. Los nodos primarios proporcionan de forma activa servicios de aplicación. El nodo secundario no debe estar desocupado mientras se espera que uno principal falle.

El nodo secundario es el único de la configuración que está conectado físicamente a todos las unidades de almacenamientos multisistema .

Si se produce un fallo en el nodo primario, Sun Cluster se recupera utilizando los recursos del nodo secundario, donde seguirán funcionando los recursos hasta que se vuelvan a pasar al nodo primario (ya sea de forma manual o automática).

El secundario siempre debe tener suficiente capacidad sobrante de CPU para manejar la carga si uno de los primarios falla.

La figura siguiente ilustra la configuración n+1.

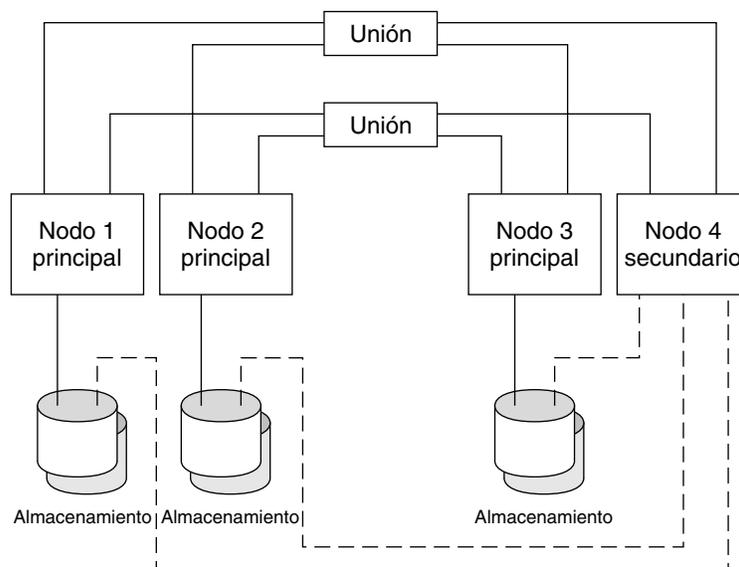


FIGURA 2-4 SPARC: Topología n+1

SPARC: Topología N*N (escalable) para SPARC

Una topología N*N activa todos los dispositivos de almacenamiento compartidos del clúster para conectar cada nodo del clúster. Esta topología activa aplicaciones de alta disponibilidad para que un nodo sustituya a otro en caso de fallo sin que el servicio se vea afectado. Cuando se produce un fallo, el nodo nuevo puede acceder al dispositivo de almacenamiento usando una ruta local en lugar de la interconexión privada.

La siguiente figura muestra una configuración n*n.

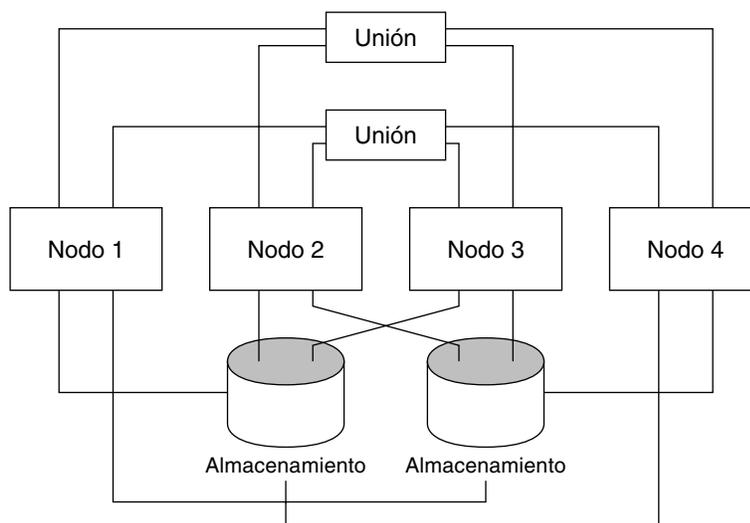


FIGURA 2-5 SPARC: Topología n*n

x86: Topologías de Sun Cluster para x86

Una topología es el esquema de conexión que une los nodos del clúster con las plataformas de almacenamiento que se usan en el clúster. Sun Cluster admite cualquier topología que cumpla las siguientes directrices.

- Sun Cluster, cuando se compone de sistemas basados en plataformas x86, admite dos nodos por clúster.
- Los dispositivos de almacenamiento compartido se deben conectar con ambos nodos.

Sun Cluster no obliga a configurar el clúster mediante topologías específicas. La siguiente topología de par en clúster, la única que admiten los clústers compuestos de nodos basados en plataformas x86, se describe con el único propósito de proporcionar una terminología que explique el esquema de conexiones del clúster. Esta topología es un esquema de conexión típico.

El apartado siguiente incluye una diagrama de ejemplo de la topología.

x86: Topología de par en clúster para x86

Una topología de par en clúster consiste en dos nodos que funcionan en una misma estructura administrativa de clúster. En esta configuración sólo se produce recuperación de fallos entre pares. Sin embargo, todos los nodos están conectados mediante la interconexión del clúster y funcionan bajo el control del software de Sun Cluster. Esta topología se podría usar para ejecutar una aplicación de base de datos paralela en un par y una aplicación de recuperación de fallos o de escalabilidad en otro par.

La figura siguiente ilustra una configuración de par en clúster.

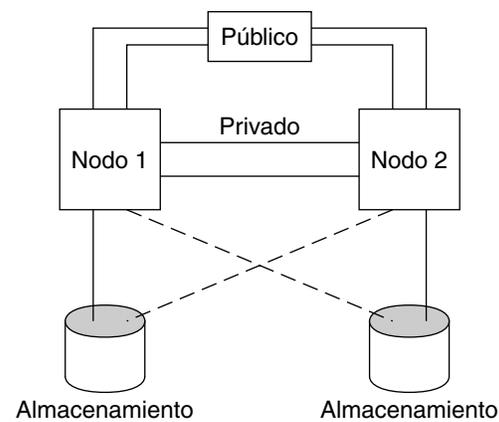


FIGURA 2-6 x86: Topología de par en clúster

Conceptos clave para los administradores de sistemas y los desarrolladores de aplicaciones

Este capítulo describe los conceptos clave relacionados con los componentes de software de un sistema Sun Cluster. Se tratan los siguientes temas:

- “Interfaces administrativas ” en la página 38
- “Hora del clúster ” en la página 38
- “Estructura de alta disponibilidad ” en la página 39
- “Dispositivos globales ” en la página 42
- “Grupos de dispositivos de discos ” en la página 43
- “Espacio de nombres global ” en la página 47
- “Sistemas de archivos de clúster ” en la página 48
- “Supervisión de la ruta de discos ” en la página 51
- “Quórum y dispositivos del quórum ” en la página 54
- “Servicios de datos ” en la página 65
- “Uso de interconexiones del clúster para el tráfico de servicios de datos ” en la página 75
- “Tipos de recursos, grupos de recursos y recursos ” en la página 76
- “Configuración de un proyecto de servicio de datos ” en la página 79
- “Adaptadores de red públicos y Ruta múltiple de red de protocolo de Internet (IP) ” en la página 89
- “SPARC: Compatibilidad con la reconfiguración dinámica ” en la página 90

Esta información está dirigida especialmente a los administradores de sistemas y a los desarrolladores de aplicaciones que usan el SDK y la API de Sun Cluster. Asimismo, esta información puede resultar útil a los administradores de sistemas del clúster para instalar, configurar y administrar el software del clúster. Los desarrolladores de aplicaciones pueden usar la información para entender el entorno del clúster en el que estarán trabajando.

Interfaces administrativas

Puede elegir la forma de instalar, configurar y administrar el sistema Sun Cluster desde varias interfaces de usuario, así como realizar tareas de administración del sistema a través de la interfaz gráfica de usuario (GUI) de SunPlex Manager o de la interfaz de línea de órdenes documentada. En la parte superior de la interfaz de línea de comandos hay algunas utilidades, como `scinstall` y `scsetup`, para simplificar las tareas de instalación y configuración. El sistema Sun Cluster también tiene un módulo que se ejecuta como parte de Sun Management Center e incluye una GUI para algunas tareas del clúster, y que está disponible sólo para usarlo en clústers basados en plataformas SPARC. Consulte “Herramientas de administración” de *Sun Cluster: Guía de administración del sistema para el SO Solaris* para obtener descripciones completas de las interfaces administrativas.

Hora del clúster

La hora en todos los nodos del clúster debe estar sincronizada. Que esta sincronización se realice con una fuente externa no es importante para el funcionamiento del clúster. El sistema Sun Cluster utiliza el protocolo Network Time Protocol (NTP) para sincronizar los relojes de los distintos nodos.

En general, una diferencia en el reloj del sistema de una fracción de segundo no causa problemas. Sin embargo, si ejecuta `date(1)`, `rdate(1M)` o `xntpdate(1M)` (ya sea de forma interactiva o con las secuencias de comandos `cron`) en un clúster activo, podrá forzar un cambio de hora mucho mayor que una fracción de segundo para sincronizar el reloj del sistema con un origen de hora, lo que podría causar problemas con la marca de hora de modificación de archivos o confundir al servicio NTP.

Al instalar el sistema operativo Solaris en cada nodo del clúster, tendrá la oportunidad de cambiar la hora predeterminada y la configuración de la fecha para el nodo. En general, puede aceptar el valor predeterminado sugerido.

Cuando se instala el software Sun Cluster usando `scinstall(1M)`, un paso del procedimiento consiste en configurar NTP para el clúster. El software Sun Cluster proporciona un archivo de plantilla, `ntp.cluster` (consulte `/etc/inet/ntp.cluster` en un nodo de clúster instalado), que establece una relación de igual a igual entre todos los nodos del clúster. Un nodo se establece como el “preferido”. Los nodos se identifican mediante sus nombres de hosts privados y la sincronización de hora se produce mediante la interconexión de los clústeres. Para obtener instrucciones acerca de cómo configurar un clúster para NTP, consulte el Capítulo 2, “Instalación y configuración del software Sun Cluster” de *Software Sun Cluster: Guía de instalación para el sistema operativo Solaris*.

También puede configurar uno o más servidores NTP externos al clúster y cambiar el archivo `ntp.conf` para reflejar esta configuración.

Durante el funcionamiento normal, nunca debería ser necesario ajustar la hora en el clúster. No obstante, si se estableció mal la hora al instalar el sistema operativo Solaris y desea modificarla, el procedimiento para hacerlo figura en el Capítulo 7, “Administración del clúster” de *Sun Cluster: Guía de administración del sistema para el SO Solaris*.

Estructura de alta disponibilidad

El sistema Sun Cluster hace que todos los componentes de la “ruta de acceso” entre los usuarios y los datos estén muy disponibles, incluidos elementos como las interfaces de red, las aplicaciones en sí, el sistema de archivos y los dispositivos multisistema. En general, un componente del clúster está muy disponible si sobrevive a cualquier fallo (de software o hardware) que se produzca en el sistema.

La siguiente tabla muestra los tipos de fallos de los componentes de Sun Cluster (tanto de hardware como de software) y los tipos de recuperaciones que se integran en la estructura de alta disponibilidad.

TABLA 3-1 Niveles de detección y recuperación de fallos de Sun Cluster

Componente del clúster fallido	Recuperación de software	Recuperación de hardware
Servicio de datos	API HA , estructura HA	N/A
Adaptador de red pública	Ruta múltiple de red de protocolo de Internet (IP)	Varias tarjetas adaptadoras de red pública
Sistema de archivos del clúster	Réplicas primaria y secundaria	Dispositivos multisistema
Dispositivo multisistema duplicado	Gestión de volúmenes (Gestor de volúmenes de Solaris y VERITAS Volume Manager, que está disponible sólo en clústeres basados en SPARC)	RAID-5 por hardware (por ejemplo: Sun StorEdge™ A3x00)
Dispositivo global	Réplicas primaria y secundaria	Varias rutas de acceso al dispositivo, uniones de transporte al clúster
Red privada	Software de transporte HA	Varias redes privadas independientes del hardware

TABLA 3-1 Niveles de detección y recuperación de fallos de Sun Cluster (Continuación)

Componente del clúster fallido	Recuperación de software	Recuperación de hardware
Nodo	CMM, controlador de recuperación rápida	Varios nodos

La estructura de alta disponibilidad del software de Sun Cluster detecta rápidamente cualquier fallo en los nodos y crea un nuevo servidor equivalente para los recursos de la estructura en otro nodo del clúster. En ningún momento se interrumpe completamente la disponibilidad de los recursos de la estructura. Los recursos de la estructura que no se han visto afectados por el fallo del nodo siguen estando disponibles durante la recuperación. Además, los recursos de la estructura del nodo fallido vuelven a estar disponibles en cuanto se recuperan. Un recurso de la estructura recuperado no tiene que esperar a que todos los demás recursos de la estructura completen su recuperación.

La mayoría de recursos de la estructura de alta disponibilidad se recuperan de manera transparente para las aplicaciones (servicios de datos) que utilizan el recurso. Las semánticas del acceso a los recursos de la estructura se conservan perfectamente entre los fallos de los nodos. Las aplicaciones no pueden detectar que el servidor del recurso de la estructura se ha movido a otro nodo. El error de un único nodo es completamente transparente para los programas de los nodos restantes que utilicen los archivos, los dispositivos y los volúmenes de discos adjuntos a este nodo. Esta transparencia es posible si hay una ruta de hardware alternativa a los discos desde otro nodo. Un ejemplo es el uso de dispositivos multisistema que tengan puertos con varios nodos.

Cluster Membership Monitor

Para asegurarse de que los datos permanezcan incorruptos, todos los nodos deben alcanzar un acuerdo uniforme sobre la pertenencia al clúster. Cuando es necesario, CMM coordina una reconfiguración de los servicios del clúster (aplicaciones) en respuesta a un fallo.

CMM recibe información sobre conectividad con otros nodos desde la capa de transporte del clúster. CMM usa la interconexión del clúster para intercambiar información de estado durante la reconfiguración.

Tras detectar un cambio en la pertenencia del clúster, CMM efectúa una configuración sincronizada de éste en la cual es posible que los recursos del clúster se redistribuyan, basándose en la nueva pertenencia del clúster.

A diferencia de las versiones de software anteriores de Sun Cluster, CMM se ejecuta por completo en el núcleo.

Consulte [“Acerca del aislamiento de fallos”](#) en la página 56 para obtener más información acerca de cómo se protege el clúster para impedir su división en numerosos clústeres separados.

Mecanismo de recuperación rápida

Si CMM detecta un problema grave en un nodo, lo notifica a la estructura del clúster para forzar el cierre (avisos graves) del nodo y hacer que deje de pertenecer al clúster. El mecanismo por el que esto se produce se llama *recuperación rápida*. Este mecanismo hace que un nodo se cierre de dos formas.

- Si un nodo abandona el clúster y después intenta crear uno nuevo sin tener quórum, queda “encerrado” y se le impide acceder a discos compartidos. Consulte “Acerca del aislamiento de fallos” en la página 56 para obtener información acerca de este uso del mecanismo de recuperación rápida.
- Si uno o varios de los daemons específicos del clúster sufre un fallo (clexecd, rpc.pmf, rgmd o rpc.ed), CMM detecta el fallo y el nodo emite avisos graves.

Cuando el fallo del daemon de un clúster provoca que un nodo envíe avisos graves, se muestra un mensaje similar al siguiente en la consola para dicho nodo.

```
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 fbf0
```

Después del aviso grave, es posible que el nodo rearranque e intente reunificar el clúster. Si el clúster está compuesto por sistemas basados en SPARC, puede que el nodo permanezca en el indicador OpenBoot™ PROM (OBP). La siguiente acción del nodo dependerá del valor del parámetro `auto-boot?`. Puede configurar `auto-boot?` con `eeprom(1M)`, en el indicador OpenBoot PROM `ok`.

Cluster Configuration Repository (CCR)

CCR usa un algoritmo de puesta al día de dos fases para las actualizaciones: Una actualización debe completarse satisfactoriamente en todos los miembros del clúster, de lo contrario, la actualización se deshace. CCR usa la interconexión del clúster para aplicar las actualizaciones distribuidas.



Caution – CCR se compone de archivos de texto que nunca se deben editar manualmente, ya que cada archivo contiene un registro de suma de control para asegurar la coherencia entre los nodos. La actualización manual de los archivos de CCR provocaría que un nodo o todo el clúster dejaran de funcionar.

CCR confía en CMM para garantizar que el clúster sólo se ejecute cuando se tenga el suficiente quórum y es responsable de verificar la uniformidad de los datos entre el clúster, efectuando recuperaciones según sea necesario y facilitando actualizaciones a los datos.

Dispositivos globales

El sistema Sun Cluster utiliza *dispositivos globales* para proporcionar a todos los dispositivos del clúster un acceso de alta disponibilidad en todo el clúster, desde cualquier nodo, con independencia del lugar al que esté acoplado físicamente el dispositivo. Por lo general, si un nodo falla a la hora de proporcionar acceso a un dispositivo global, el software Sun Cluster detecta automáticamente otra ruta al dispositivo y redirige el acceso a esa ruta. Los dispositivos globales de Sun Cluster pueden ser discos, CD-ROM y cintas. Sin embargo, los únicos dispositivos globales multipuerto que admite el software Sun Cluster son los discos. En consecuencia, los CD-ROM y los dispositivos de cinta no son actualmente dispositivos de alta disponibilidad. Los discos locales de cada servidor tampoco son multipuerto, por lo tanto no son dispositivos de alta disponibilidad.

El clúster asigna automáticamente ID exclusivos a cada disco, CD-ROM y unidad de cinta del clúster, lo que permite el acceso uniforme a todos los dispositivos desde cualquier nodo del clúster. El espacio de nombres de dispositivos global se conserva en el directorio `/dev/global`. Consulte ["Espacio de nombres global"](#) en la página 47 para obtener más información.

Los dispositivos globales multipuerto ofrecen más de una ruta de acceso al dispositivo. Dado que los discos multisistema forman parte de un grupo de dispositivos de discos que están alojados por más de un nodo, los discos multisistema tienen una alta disponibilidad.

ID de dispositivo y pseudocontrolador DID

El software Sun Cluster gestiona dispositivos globales a través de una estructura conocida como el pseudocontrolador del ID del dispositivo (DID). Este controlador se utiliza para asignar automáticamente identificadores exclusivos a todos los dispositivos del clúster, incluidos discos multisistema, unidades de cinta y CD-ROM.

También es una pieza integral de la función de acceso global a los dispositivos del clúster. El controlador DID prueba todos los nodos del clúster y genera una lista de dispositivos de disco únicos, asigna a cada dispositivo un número mayor y otro menor que son coherentes en todos los nodos del clúster. Para acceder a los dispositivos globales se utiliza el ID de dispositivo exclusivo en lugar de los ID de dispositivo tradicionales de Solaris como, por ejemplo, `c0t0d0` para un disco.

Este enfoque garantiza que cualquier aplicación que acceda a los discos (como, por ejemplo, el gestor de volúmenes o las aplicaciones que utilicen dispositivos sin formato) utilice una ruta coherente a través del clúster. Esta uniformidad es especialmente importante en discos multisistema, debido a que los números mayor y menor de cada dispositivo pueden variar entre distintos nodos, cambiando así

también las convenciones de asignación de nombres de dispositivos de Solaris. Por ejemplo, el nodo 1 podría identificar un disco multisistema como `c1t2d0` y el nodo 2 podría ver ese mismo disco de forma completamente distinta, como `c3t2d0`. El controlador DID asigna un nombre global, como `d10`, que los nodos usarán en su lugar, lo que generará una asignación uniforme con el disco multisistema.

Los ID se actualizan y administran mediante `sddidadm(1M)` y `scgdevs(1M)`. Consulte las siguientes páginas de comando `man` para obtener más información:

- `sddidadm(1M)`
- `scgdevs(1M)`

Grupos de dispositivos de discos

En el sistema Sun Cluster, todos los dispositivos multisistema deben estar bajo el control del software Sun Cluster. En primer lugar, se deben crear los grupos de discos del gestor de volúmenes —ya sea un conjunto de discos de Gestor de volúmenes de Solaris o un grupo de discos de VERITAS Volume Manager (disponible para usarlo sólo en clústeres basados en SPARC)— en los discos multisistema. A continuación, se registran los grupos de discos del gestor de discos como *grupos de dispositivos de discos* que forman un tipo de dispositivo global. Además, el software Sun Cluster crea automáticamente un grupo de dispositivos de bajo nivel para cada dispositivo de disco y de cinta del clúster. Sin embargo, estos grupos de dispositivos del clúster permanecen en estado fuera de línea hasta que se accede a ellos como dispositivos globales.

Este registro proporciona al sistema Sun Cluster información sobre qué nodos tienen una ruta a ciertos grupos de discos del gestor de volúmenes. En este punto, los grupos de discos del gestor de volúmenes se convierten en accesibles globalmente dentro del clúster. Si hay más de un nodo que pueda escribir (controlar) un grupo de dispositivos de disco, los datos almacenados en éste se consideran de alta disponibilidad. El grupo de dispositivos de disco de alta disponibilidad se puede utilizar para albergar los sistemas de archivos del clúster.

Nota – Los grupos de dispositivos de disco son independientes de los grupos de recursos. Un nodo puede actuar como maestro de un grupo de recursos (representando un grupo de procesos de servicios de datos), mientras que otro puede actuar de maestro de un grupo de discos a los que acceden los servicios de datos. Sin embargo, la mejor práctica consiste en mantener en el mismo nodo el grupo de dispositivos que almacena los datos de una aplicación en concreto y el grupo de recursos que contiene los recursos de la aplicación (el daemon de la aplicación). Consulte “Relationship Between Resource Groups and Disk Device Groups” de *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* para obtener más información acerca de la asociación entre los grupos de dispositivos de discos y los grupos de recursos.

Cuando un nodo utiliza un grupo de dispositivos de disco, el grupo de discos del gestor de volúmenes se convierte en “global” porque proporciona un soporte multirruta a los discos subyacentes. Cada nodo del clúster conectado físicamente a los discos multisistema proporciona una ruta de acceso al grupo de dispositivos de disco.

Recuperación de fallos en grupos de dispositivos de disco

Debido a que un alojamiento de disco está conectado a más de un nodo, todos los grupos de dispositivos de disco de ese alojamiento estarán disponibles a través de una ruta de acceso alternativa si falla el nodo que esté controlando en ese momento el grupo de dispositivos. El fallo del nodo que controla el grupo de dispositivos no afecta al acceso al grupo de dispositivos excepto por el tiempo que se tarda en realizar la recuperación y las comprobaciones de integridad. Durante ese tiempo, todas las peticiones se bloquean (de forma transparente para la aplicación) hasta que el sistema vuelve a hacer que el grupo de dispositivos esté disponible.

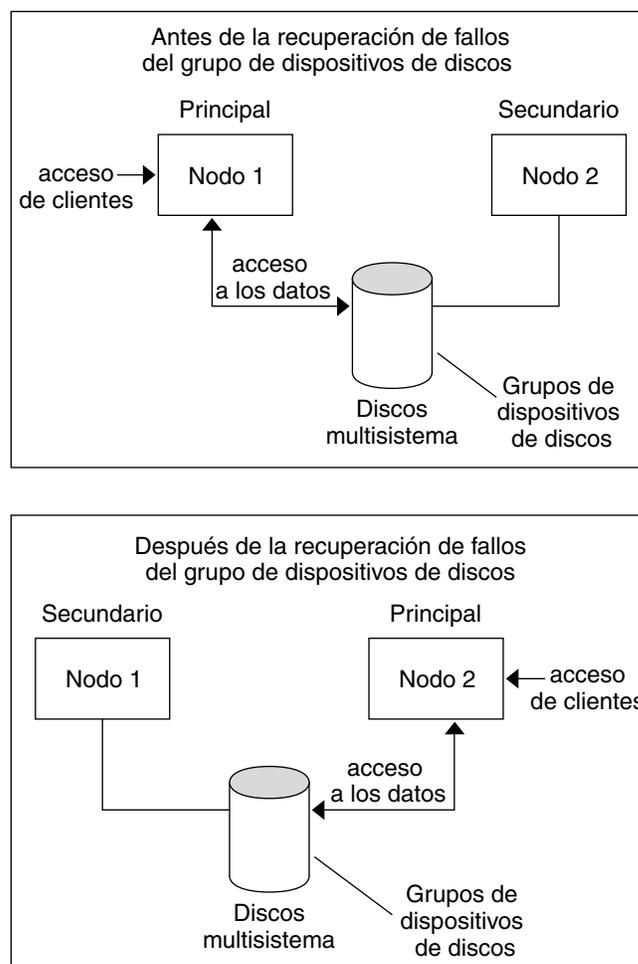


FIGURA 3-1 Grupo de dispositivos de disco antes y después de una recuperación de fallos

Grupos de dispositivos de disco multipuerto

Esta sección describe las propiedades de grupo de dispositivos de disco que le permiten equilibrar rendimiento y disponibilidad en una configuración de disco multipuerto. El software Sun Cluster proporciona dos propiedades que se usan para definir una configuración de disco multipuerto: `preferred` y `numsecondaries`. Puede controlar el orden en el que los nodos tratan de asumir el control si se produce una recuperación de fallos usando la propiedad `preferred`. La propiedad `numsecondaries` se utiliza para establecer un número deseado de nodos secundarios para un grupo de dispositivos.

Se considera que un servicio de alta disponibilidad está inactivo cuando el nodo primario falla y no hay nodos secundarios aptos para asumir las funciones del primario. Si se produce una recuperación de fallos y la propiedad `preferenced` tiene un valor `true`, entonces los nodos siguen el orden establecido en la lista de nodos para elegir cuál de ellos será el secundario. La lista de nodos establece el orden en el que los nodos intentarán asumir el control principal o dejarán de estar en reserva para ser secundarios. Puede cambiar dinámicamente la preferencia de un servicio de dispositivo mediante la utilidad `scsetup(1M)`. La preferencia asociada a los proveedores de servicios dependientes (por ejemplo, un sistema de archivos global) será idéntica a la preferencia del servicio de dispositivo.

El nodo primario comprueba los secundarios durante el funcionamiento normal. En una configuración de disco multipuerto, comprobar todos los nodos secundarios produce una degradación en el rendimiento del clúster y una sobrecarga en la memoria. La compatibilidad con el nodo de reserva se implementó para minimizar la degradación del rendimiento y la sobrecarga de la memoria que causaban las comprobaciones puntuales. De forma predeterminada, el grupo de dispositivos de disco tiene un nodo primario y otro secundario. El resto de nodos de proveedor disponibles aparecen como nodos de reserva. Si se produce una recuperación de fallos, el nodo secundario se convierte en primario y el nodo que tuviera una mayor prioridad en la lista, pasa a ser secundario.

Se puede definir el número que desee de nodos secundarios. Puede ser cualquier número entero entre uno y el número de nodos operativos de proveedor que no sean primarios del grupo de dispositivos.

Nota – Si está usando Gestor de volúmenes de Solaris, deberá crear el grupo de dispositivos de disco antes de definir la propiedad `numsecondaries` en un número distinto del predeterminado.

De manera predeterminada el número deseado de secundarios para servicios de dispositivos es de uno. El número real de proveedores secundarios que se mantiene en la estructura de réplica es el número que se indique, a menos que el número de proveedores operativos que no sean primarios sea inferior al número deseado. Debe modificar la propiedad `numsecondaries` y volver a comprobar la lista de nodos si está agregando o eliminando nodos de la configuración. Mantener la lista de nodos y el número deseado de nodos secundarios impide que se produzcan conflictos entre el número configurado de nodos secundarios y el número real que permite la estructura.

- (Gestor de volúmenes de Solaris) Use el comando `metaset(1M)` para los grupos de dispositivos de Gestor de volúmenes de Solaris junto con los ajustes de las propiedades `preferenced` y `numsecondaries`, para administrar la adición y la eliminación de nodos de la configuración.
- (Veritas Volume Manager) Use el comando `scconf(1M)` para los grupos de dispositivos VxVM, junto con los ajustes de las propiedades `preferenced` y `numsecondaries`, para administrar la adición y la eliminación de nodos de la

configuración.

- Consulte “Administración de sistemas de archivos del clúster: información general” de *Sun Cluster: Guía de administración del sistema para el SO Solaris* para obtener información sobre los procedimientos para cambiar las propiedades de los grupos de dispositivos.

Espacio de nombres global

El mecanismo de software de Sun Cluster que habilita los dispositivos globales se llama *espacio de nombres global*. El espacio de nombres incluye la jerarquía `/dev/global/`, así como los espacios de nombres del gestor de volúmenes. El espacio de nombres global representa los discos multisistema y locales (y cualquier otro dispositivo del clúster, como CD-ROM y cintas) e incluye varias rutas de acceso de recuperación de fallos a los discos multisistema. Cada nodo que está físicamente conectado a discos multisistema proporciona una ruta al almacenamiento para cualquier nodo del clúster.

Normalmente, para Solaris Volume Manager, los espacios de nombres del gestor de volúmenes están ubicados en los directorios `/dev/md/conjunto_discos/dsk` (y `rdsk`). Para Veritas VxVM, los espacios de nombres del gestor de volúmenes están ubicados en los directorios `/dev/vx/dsk/grupo-disco` y `/dev/vx/rdsk/grupo-disco`. Estos espacios de nombres están formados por directorios para cada conjunto de discos de Gestor de volúmenes de Solaris y cada grupo de discos de VxVM importados a través del clúster, respectivamente. Cada uno de estos directorios contiene un nodo de dispositivo para cada metadispositivo o volumen de dicho conjunto de discos o grupo de discos.

En el sistema Sun Cluster todos los nodos de dispositivos del espacio de nombres del gestor de volúmenes local se sustituyen por un enlace simbólico con un nodo de dispositivo del sistema de archivos `/global/.devices/node@IDnodo`, donde *IDnodo* es un número entero que representa los nodos del clúster. El software Sun Cluster continúa presentando los dispositivos del gestor de volúmenes como enlaces simbólicos y también en sus ubicaciones estándar. Tanto el espacio de nombres global como el estándar del gestor de volúmenes están disponibles desde cualquier nodo del clúster.

Entre las ventajas de los espacios de nombres, se incluyen las siguientes:

- Cada nodo permanece bastante independiente, con pocos cambios en el modelo de administración de dispositivos.
- Los dispositivos puede convertirse en globales de forma selectiva.
- Los generadores de enlaces de terceros siguen funcionando.
- Dado un nombre de dispositivo local, se ofrece una correlación simple para obtener un nombre global.

Ejemplo de espacio de nombres local y global

La tabla siguiente muestra las correlaciones entre los espacios de nombres local y global para un disco multisistema, `c0t0d0s0`.

TABLA 3-2 Asignación de espacio de nombres local y global

Componente o ruta	Espacio de nombres de nodo local	Espacio de nombres global
Nombre lógico de Solaris	<code>/dev/dsk/c0t0d0s0</code>	<code>/global/.devices/node@IDnodo/dev/dsk/c0t0d0s0</code>
Nombre DID	<code>/dev/did/dsk/d0s0</code>	<code>/global/.devices/node@IDnodo/dev/did/dsk/d0s0</code>
Gestor de volúmenes de Solaris	<code>/dev/md/conjunto_discos/dsk/d0</code>	<code>/global/.devices/node@IDnodo/dev/md/conjunto_discos/dsk/d0</code>
SPARC: VERITAS Volume Manager	<code>/dev/vx/dsk/grupo-discos/v0</code>	<code>/global/.devices/node@IDnodo/dev/vx/dsk/grupo-discos/v0</code>

El espacio de nombres global se genera automáticamente durante la instalación y se actualiza con cada arranque de reconfiguración; también se puede generar mediante la orden `scgdevs (1M)`.

Sistemas de archivos de clúster

El sistema de archivos del clúster dispone de las prestaciones siguientes:

- Las ubicaciones de los accesos de archivo son transparentes. Un proceso puede abrir un archivo que se encuentre en cualquier ubicación del sistema. Los procesos de todos los nodos pueden usar el mismo nombre de ruta para localizar un archivo.

Nota – Cuando el sistema de archivos del clúster lee archivos, no actualiza la hora de acceso en esos archivos.

- Se utilizan protocolos de coherencia para preservar la semántica de acceso a archivos UNIX aunque varios nodos estén accediendo al archivo al mismo tiempo.
- Para mover datos de archivos eficientemente se utiliza masivamente la antememoria y el movimiento de E/S en bloque sin copia.
- El sistema de archivos del clúster ofrece el bloqueo de archivos de aviso de alta disponibilidad a través de las interfaces `fcntl (2)`. Las aplicaciones que se ejecuten en varios nodos del clúster pueden sincronizar el acceso a los datos

mediante el bloqueo a los archivos de aviso en un archivo del sistema de archivos del clúster. Los bloqueos de archivo se recuperan inmediatamente desde los nodos que abandonan el clúster y las aplicaciones que fallan mientras se mantienen los bloqueos.

- El acceso continuo a los datos queda asegurado aunque se produzcan fallos. Las aplicaciones no se ven afectadas por fallos si sigue estando operativa una ruta de acceso a los discos. Esta garantía se mantiene para el acceso a discos de bajo nivel y todas las operaciones del sistema de archivos.
- Los sistemas de archivos del clúster son independientes del sistema de archivos subyacente y del software de gestión de volúmenes; convierten en global cualquier sistema de archivos en disco admitido.

Se puede montar un sistema de archivos en un dispositivo global con `mount -g` (globalmente) o con `mount` (localmente).

Los programas pueden acceder a un archivo en un sistema de archivos en clúster desde cualquier nodo del clúster a través del mismo nombre de archivo (por ejemplo, `/global/foo`).

Los sistemas de archivos del clúster se montan en todos los miembros del clúster, pero no puede montarse en un subconjunto de miembros del clúster.

Los sistemas de archivo clúster no son de tipo diferenciado. Los clientes verifican el sistema de archivos que subyace (por ejemplo, UFS).

Uso de los sistemas de archivos del clúster

En el sistema Sun Cluster todos los discos multisistema se sitúan en grupos de dispositivos de disco que pueden ser conjuntos de discos de Gestor de volúmenes de Solaris, grupos de discos de VxVM o discos individuales que no estén bajo el control de ningún gestor de volúmenes basado en software.

Para que un sistema de archivos del clúster sea de alta disponibilidad, el almacenamiento en disco subyacente debe estar conectado a más de un nodo. Por consiguiente, un sistema de archivos local (aquel que está almacenado en el disco local de un nodo) que se convierte en un sistema de archivos del clúster no es de alta disponibilidad.

Puede montar sistemas de archivos del clúster del mismo modo que montaría sistemas de archivos normales:

- **Manualmente** – Use el comando `mount` y las opciones de montaje `-g` o bien `-o global` para montar el sistema de archivos del clúster desde la línea de comandos, por ejemplo:

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

- **Automáticamente**– Cree una entrada en el archivo `/etc/vfstab` con una opción de montaje `global` para montar el sistema de archivos del clúster en el arranque. Después puede crear un punto de montaje en el directorio `/global` de todos los

nodos. El directorio `/global` es la ubicación que se recomienda, pero no es obligatorio usarlo. A continuación, figura una línea de ejemplo para un sistema de archivos del clúster desde un archivo `/etc/vfstab`:

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

Nota – Mientras el software Sun Cluster no impone ninguna directiva de asignación de nombre para los sistemas de archivos del clúster, puede facilitar la administración creando un punto de montaje para todos los sistemas de archivo del clúster en el mismo directorio como, por ejemplo, `/global/grupo-dispositivo-disco`. Consulte la *Sun Cluster 3.1 9/04 Software Collection for Solaris OS (SPARC Platform Edition)* y la *Sun Cluster: Guía de administración del sistema para el SO Solaris* para obtener más información.

Tipo de recurso de HASToragePlus

El tipo de recurso de HASToragePlus está designado para hacer configuraciones de sistema de archivo no globales, como UFS y VxFS, de alta disponibilidad. Use HASToragePlus para integrar su sistema de archivos local en el entorno de Sun Cluster y hacer que el sistema de archivos tenga una alta disponibilidad. HASToragePlus proporciona funciones de sistema de archivo adicionales como comprobaciones, montajes y desmontajes forzados que capacitan a Sun Cluster para recuperarse de los errores de los sistemas de archivos locales. Para la recuperación de fallos el sistema de archivos local debe residir en grupos de discos globales que tengan habilitados conmutadores de afinidad.

Consulte “Enabling Highly Available Local File Systems” de *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* para obtener información acerca de cómo usar el tipo de recurso de HASToragePlus.

HASToragePlus se puede usar también para sincronizar el inicio de los recursos y los grupos de dispositivos de disco de los que dependen los recursos. Para obtener más información, consulte “Tipos de recursos, grupos de recursos y recursos ” en la página 76.

La opción de montaje Syncdir

Puede usar la opción de montaje `syncdir` para los sistemas de archivos del clúster que usen UFS como el sistema de archivos subyacente. No obstante, el rendimiento mejora significativamente si no se especifica `syncdir`. Si especifica `syncdir`, se garantiza que lo que se escriba sea compatible con POSIX. Si no especifica `syncdir`, experimentará la misma conducta que con los sistemas de archivos NFS. Por ejemplo, sin `syncdir`, puede que no se detecte una situación de falta de espacio hasta que cierre un archivo. Con `syncdir` (y el comportamiento POSIX), la condición de falta de espacio se descubriría durante la operación de escritura. Los casos en los que se pueden producir problemas si no se especifica `syncdir` son poco habituales.

Si está utilizando un clúster basado en SPARC, VxFS no tiene ningún punto de montaje que sea equivalente a la opción de montaje `syncdir` para UFS. El comportamiento de VxFS es el mismo que para UFS cuando no se especifica la opción de montaje `syncdir`.

Consulte [“FAQ sobre los sistemas de archivos” en la página 96](#) para ver las preguntas frecuentes acerca de los dispositivos globales y los sistemas de archivos del clúster.

Supervisión de la ruta de discos

La versión actual del software Sun Cluster admite la supervisión de rutas de discos (DPM). Este apartado incluye información conceptual sobre DPM, el daemon DPM, y las herramientas de administración que se pueden usar para supervisar las rutas del disco. Consulte la *Sun Cluster: Guía de administración del sistema para el SO Solaris* para obtener información sobre los procedimientos para supervisar, dejar de supervisar y comprobar el estado de las rutas del disco.

Nota – DPM no se admite en los nodos que ejecutan versiones anteriores a Sun Cluster 3.1 10/03. No utilice órdenes de DPM durante una modernización. Una vez finalizada la modernización en todos los nodos, éstos deben estar en línea para poder utilizar las órdenes de DPM.

Información general de DPM

DPM mejora la disponibilidad general de la recuperación de fallos y la conmutación por supervisión de la disponibilidad de la ruta de acceso a disco secundaria. Utilice el comando `scdpm` para verificar la disponibilidad de la ruta del disco que utiliza un recurso antes de conmutarlo. Las opciones que se proporcionan con el comando `scdpm` le permiten supervisar las rutas de discos a un único nodo o a todos los nodos del clúster. Consulte la página de comando `man scdpm(1M)` para obtener más información acerca de las opciones de la línea de comandos.

Los componentes DPM se instalan a partir del paquete `SUNWscu`. El paquete `SUNWscu` se instala durante el procedimiento de instalación estándar de Sun Cluster. Consulte la página de comando `man scinstall(1M)` para obtener detalles sobre la interfaz de instalación. En la siguiente tabla se indica la ubicación predeterminada para la instalación de los componentes de DPM.

Ubicación	Componente
Daemon	/usr/cluster/lib/sc/scdpmd
Interfaz de línea de órdenes	/usr/cluster/bin/scdpm
Bibliotecas compartidas	/user/cluster/lib/libscdpm.so
Archivo de estado de daemon (creado en tiempo de ejecución)	/var/run/cluster/scdpm.status

En cada nodo se ejecuta un daemon DPM de subproceso múltiple. El daemon DPM (`scdpmd`) lo inicia la secuencia `rc.d` cuando arranca un nodo. Si surge algún problema, `pmfd` gestiona el daemon y lo reinicia automáticamente. La lista siguiente describe cómo funciona `scdpmd` en el arranque inicial.

Nota – En el arranque, el estado de cada ruta del disco se inicializa a UNKNOWN.

1. El daemon DPM recopila información sobre la ruta del disco y el nombre del nodo desde el archivo de estado anterior o desde la base de datos CCR. Consulte la “[Cluster Configuration Repository \(CCR\)](#)” en la página 41 para obtener más información acerca de CCR. Cuando se inicia el daemon DPM, éste puede forzarse para que lea la lista de discos supervisados a partir de un nombre de archivo especificado.
2. El daemon DPM inicializa la interfaz de comunicaciones para que responda a solicitudes de componentes que son externos al mismo, como la interfaz de línea de órdenes.
3. El daemon DPM realiza un ping a cada ruta del disco de la lista supervisada cada 10 minutos mediante órdenes de tipo `scsi_inquiry`. Todas las entradas están bloqueadas para evitar que la interfaz de comunicaciones acceda al contenido de una entrada que esté siendo modificada.
4. El daemon de DPM informa a la estructura sobre los eventos de Sun Cluster y registra el nuevo estado de la ruta mediante el mecanismo de UNIX `syslogd(1M)`.

Nota – `pmfd(1M)` registra todos los errores relacionados con el daemon. Todas las funciones de la API devuelven 0 cuando tienen éxito y -1 cuando se produce algún error.

El daemon DPM supervisa la disponibilidad de la ruta lógica que está visible a través de controladores multirruta como Gestor de tráfico Sun StorEdge, HDLM y PowerPath. Las rutas de acceso físicas individuales gestionadas por estos controladores no están supervisadas porque el controlador de ruta múltiple oculta los fallos individuales al daemon DPM.

Supervisión de las rutas del disco

Este apartado describe dos métodos para supervisar las rutas del disco en el clúster. El primer método lo ofrece el comando `scdpm`. Utilice esta orden para supervisar, dejar de supervisar o mostrar el estado de las rutas del disco del clúster; Este comando también es útil para imprimir la lista de discos con fallos y supervisar las rutas de disco desde un archivo.

El segundo método para supervisar las rutas de los discos en el clúster lo ofrece la interfaz gráfica de usuario (GUI) de SunPlex Manager. SunPlex Manager incluye una vista topográfica de las rutas del disco supervisadas del clúster. La vista se actualiza cada 10 minutos para incluir información sobre el número de pings que han fallado. Para administrar rutas del disco use la información que proporciona la GUI de SunPlex Manager junto con la orden `scdpm(1M)`. Consulte el Capítulo 10, “Administración de Sun Cluster con las interfaces gráficas de usuario” de *Sun Cluster: Guía de administración del sistema para el SO Solaris* para obtener información acerca de SunPlex Manager.

Uso de la orden `scdpm` para supervisar las rutas del disco

El comando `scdpm(1M)` proporciona a DPM comandos de administración que permiten realizar las tareas siguientes:

- Supervisar una ruta del disco nueva
- Dejar de supervisar una ruta del disco
- Volver a leer los datos de configuración de la base de datos CCR
- Leer los discos para supervisar o dejar de hacerlo a partir de un archivo especificado
- Informar del estado de una o todas las rutas de acceso de disco del clúster
- Imprimir todas las rutas de acceso de disco que son accesibles desde un nodo

Emita la orden `scdpm(1M)` con el argumento ruta-disco desde cualquier nodo para llevar a cabo tareas de administración de DPM en el clúster. Éste consta en todos los casos de un nombre de nodo y de un nombre de disco. El nombre del nodo no es obligatorio y su valor predeterminado es `all` si no se especifica otra cosa. La tabla siguiente describe las convenciones de asignación de nombres para la ruta del disco.

Nota – Se recomienda utilizar nombres de ruta del disco globales, ya que son coherentes dentro de todo el clúster. Los nombres de las rutas del disco UNIX no son coherentes en todo el clúster. La ruta del disco UNIX correspondiente a un disco determinado puede diferir en distintos nodos del clúster. La ruta podría ser `c1t0d0` en un nodo y `c2t0d0` en otro. Si utiliza nombres de ruta del disco UNIX, utilice la orden `scdidadm -L` para asignar el nombre UNIX al nombre global antes de ejecutar órdenes de DPM. Consulte la página de comando `man` de `scdidadm(1M)`.

TABLA 3-3 Ejemplos de nombres de rutas del disco

Tipo de nombre	Ejemplo de nombre de ruta de disco	Descripción
Ruta del disco global	<code>schost-1:/dev/did/dsk/d1</code>	Ruta del disco d1 en el nodo <code>schost-1</code>
<code>all:d1</code>	Ruta del disco d1 en todos los nodos del clúster	
Ruta del disco UNIX	<code>schost-1:/dev/rdisk/c0t0d0s0</code>	Ruta del disco <code>c0t0d0s0</code> en el nodo <code>schost-1</code>
<code>schost-1:all</code>	Todas las rutas del nodo <code>schost-1</code>	
Todas las rutas del disco	<code>all:all</code>	Todas las rutas del disco de todos los nodos del clúster

Uso de SunPlex Manager para supervisar las rutas del disco

SunPlex Manager permite realizar las siguientes tareas de administración DPM básicas:

- Supervisar una ruta del disco
- Dejar de supervisar una ruta del disco
- Ver el estado de todas las rutas del disco en el clúster

Consulte la ayuda en línea de SunPlex Manager para obtener información de procedimientos para realizar una administración de la ruta del disco utilizando SunPlex Manager.

Quórum y dispositivos del quórum

Esta sección se divide en los siguientes apartados:

- “Acerca de los recuentos de votos del quórum ” en la página 55
- “Acerca del aislamiento de fallos ” en la página 56
- “Acerca de las configuraciones del quórum ” en la página 58
- “Cumplimiento de los requisitos de dispositivos de quórum ” en la página 59
- “Mejores prácticas relacionadas con los dispositivos del quórum ” en la página 59
- “Configuraciones de quórum recomendadas ” en la página 60
- “Configuraciones de quórum atípicas ” en la página 63
- “Configuraciones de quórum malas ” en la página 63

Nota – Para obtener una lista de los dispositivos específicos que Sun Cluster admite como dispositivos de quórum, póngase en contacto con el proveedor de servicios Sun.

Debido a que los nodos del clúster comparten datos y recursos, un clúster nunca se debe dividir en particiones separadas que estén activas a la vez porque varias particiones activas pueden provocar que se dañen los datos. Cluster Membership Monitor (CMM) y el algoritmo de quórum garantizan que como mucho una instancia del mismo clúster está operativa cada vez, incluso si se particiona la interconexión del clúster.

Para obtener una introducción acerca del quórum y CMM, consulte “Pertenencia al clúster” de *Sun Cluster para el sistema operativo Solaris: Visión general*.

De las particiones de clústeres surgen dos tipos de problemas:

- Esquizofrenia
- Amnesia

La esquizofrenia se produce cuando se pierde la interconexión del clúster entre nodos y el clúster se particiona en clústeres secundarios. Cada partición “cree” que es la única existente porque los nodos de una partición no pueden comunicarse con el nodo de la otra partición.

La amnesia se produce cuando el clúster se reinicia después de un apagado teniendo datos de configuración del clúster más antiguos que los del momento del apagado. Este problema se da cuando se inicia el clúster en un nodo que no estaba en la última partición del clúster que estuvo en funcionamiento.

Sun Cluster evita la esquizofrenia y amnesia:

- Asignando un voto a cada nodo
- Controlando la mayoría de los votos de un clúster operativo

Una partición con la mayoría de votos tiene *quórum* y se le permite funcionar. Este mecanismo de votos de la mayoría evita la esquizofrenia y amnesia cuando hay más de dos nodos configurados en un clúster. Sin embargo, el recuento de los votos de los nodos por sí solo no es suficiente cuando hay más de dos nodos configurados en un clúster. En un clúster de dos nodos, la mayoría es dos. Si dicho clúster de dos nodos se particiona, es necesario un voto externo para que cualquiera de las particiones obtenga quórum. Este voto externo lo proporciona un *dispositivo del quórum*

Acerca de los recuentos de votos del quórum

Use la opción `-q` del comando `scstat` para determinar la siguiente información:

- Votos totales configurados
- Votos presentes actuales

- Votos necesarios para quórum

Para obtener más información acerca de este comando, consulte `scstat(1M)`.

Los nodos y los dispositivos de quórum aportan votos al clúster para formar quórum.

Un nodo aporta votos en función del estado del nodo:

- Un nodo tiene un recuento de votos de *uno* cuando arranca y es miembro del clúster.
- Un nodo tiene un recuento de voto de *cero* cuando el nodo se está instalando.
- Un nodo tiene un recuento de voto de *cero* cuando un administrador de sistema pone el nodo en estado de mantenimiento.

Los dispositivos del quórum aportan votos basados en el número de votos que están conectados al dispositivo. Cuando se configura un dispositivo del quórum, el software Sun Cluster asigna al dispositivo del quórum un recuento de votos de $N-1$, donde N es el número de votos conectados al dispositivo del quórum. Por ejemplo, un dispositivo del quórum conectado a dos nodos con recuentos distintos de cero, tiene un recuento del quórum igual a uno (dos menos uno).

Un dispositivo de quórum aporta votos si se cumple *una* de las siguientes condiciones:

- Al menos uno de los nodos a los que está conectado actualmente el dispositivo de quórum es miembro del clúster.
- Al menos uno de los nodos a los que el dispositivo de quórum está conectado está arrancando y dicho nodo era miembro de la última partición de clúster propietaria del dispositivo de quórum.

Los dispositivos del quórum se configuran durante la instalación del clúster o posteriormente utilizando los procedimientos descritos en el Capítulo 5, "Administración del quórum" de *Sun Cluster: Guía de administración del sistema para el SO Solaris*.

Acerca del aislamiento de fallos

Un problema fundamental de los clústers es un fallo que provoque en éstos una partición (denominada *esquizofrenia*). Cuando esto ocurre, no todos los nodos pueden comunicarse, por lo que algunos podrían intentar formar clústers individuales o subconjuntos que "creerían" que tienen permisos de acceso y de propiedad exclusivos respecto a los dispositivos multisistema. Cuando varios nodos intentan escribir en los discos, los datos se pueden dañar.

El aislamiento de fallos limita el acceso de los nodos a los dispositivos multisistema, evitando que físicamente se pueda acceder a ellos. Cuando un nodo abandona el clúster (falla o se particiona), el aislamiento de fallos se asegura de que el nodo ya no pueda acceder a los discos. Sólo los nodos miembros actuales tendrán acceso a los discos, conservándose así la integridad de los datos.

Los servicios de dispositivos de disco ofrecen prestaciones de recuperación de fallos para servicios que utilizan dispositivos multisistema. Cuando falla o no se puede localizar un miembro del clúster que actualmente funciona como el principal (propietario) del grupo de dispositivos de disco, se elige un nuevo miembro principal. El nuevo miembro principal habilita el acceso al grupo de dispositivos de disco para que se pueda continuar funcionando con la menor interrupción posible. Durante este proceso, el antiguo miembro principal debe perder sus derechos de acceso en favor de los dispositivos antes de que se pueda iniciar el nuevo miembro principal. Sin embargo, cuando un miembro se descuelga del clúster y deja de estar disponible, éste no puede informar al nodo que libere los dispositivos para los que era primario. Por tanto, se necesita un medio para permitir que los miembros supervivientes tomen control y accedan a los dispositivos globales de los miembros fallidos.

El sistema Sun Cluster usa reservas de disco SCSI para implementar el aislamiento de fallos, gracias a las cuales, los nodos fallidos se “aislan” de los dispositivos multisistema, evitando que accedan a estos discos.

Las reservas de disco SCSI-2 admiten una forma de reservas que garantizan el acceso a todos los nodos adjuntos al disco (cuando no hay ninguna reserva). De lo contrario, el acceso se restringe a un solo nodo (el que tiene la reserva).

Cuando un miembro del clúster detecta que otro nodo ya no se está comunicando a través de la interconexión del clúster, inicia un procedimiento de aislamiento de fallos para evitar que el otro nodo acceda a los discos compartidos. En este proceso, el nodo excluido emite avisos graves y aparece un mensaje de “conflicto de reserva” en la consola.

El descubrimiento de que el nodo ya no es un miembro del clúster desencadena una reserva SCSI en todos los discos que están compartidos entre este nodo y los demás. El nodo aislado puede que no sea “consciente” de que está siendo aislado y, si intenta acceder a uno de los discos compartidos, detecta la reserva y emite avisos graves.

Mecanismo de recuperación rápida para aislamiento de fallos

El mecanismo mediante el cual la estructura del clúster se asegura de que un nodo que ha fallado no pueda reiniciarse ni comenzar a escribir en un almacenamiento compartido se llama *recuperación rápida*.

Los nodos que son miembros del clúster habilitan permanentemente un `ioctl` específico, `MHIOCENFAILFAST`, para los discos a los que tienen acceso, incluidos los de quórum. Este `ioctl` es una directiva para el controlador del disco. El `ioctl` proporciona al nodo la capacidad de enviar mensajes graves en caso de que no pueda acceder al disco porque éste esté reservado por algún otro nodo.

El `MHIOCENFAILFAST ioctl` provoca que el controlador marque el error devuelto desde cada lectura y escritura que el nodo emita para el disco con el código de error `Reservation_Conflict`. En segundo plano y de forma periódica, el `ioctl` emite

operaciones de prueba para el disco para comprobar el código de error `Reservation_Conflict`. Las rutas de flujo de control en segundo y en primer plano emiten mensajes graves si se devuelve `Reservation_Conflict`.

En discos SCSI-2, las reservas no son persistentes, pues no resisten los rearranques de los nodos. En los discos SCSI-3 con reserva de grupo persistente (PGR), la información de reserva se almacena en el disco y permanece entre los rearranques de los nodos. El mecanismo de recuperación rápida funciona lo mismo, independientemente de que se tengan discos SCSI-2 o SCSI-3.

Si un nodo pierde conectividad con los otros nodos del clúster y no es parte de una partición que pueda conseguir el quórum, otro nodo lo expulsa del clúster. Otro nodo que forma parte de la partición que puede conseguir reservas de plazas del quórum en los discos compartidos. Cuando el nodo que no tiene quórum intenta acceder a los discos compartidos, recibe un conflicto de reserva y emite mensajes graves como resultado del mecanismo de recuperación rápida.

Después de la condición de aviso grave, el nodo podría rearrancar e intentar volver a unirse al clúster o, si éste se compone de sistemas basados en plataformas SPARC, permanecer en el indicador PROM (OBP) de OpenBoot™. La acción que se toma depende del valor del parámetro `auto-boot?`. Puede definir `auto-boot?` con `eeprom(1M)`, en el indicador OpenBoot PROM `ok` de un clúster basado en SPARC. Si lo desea, también puede configurar este parámetro con la utilidad `SCSI`, que puede ejecutar opcionalmente después de que arranque la BIOS en un clúster basado en x86.

Acerca de las configuraciones del quórum

La siguiente lista contiene hechos acerca de las configuraciones de quórum:

- Los dispositivos de quórum pueden contener datos de usuario.
- En una configuración $N+1$, donde N dispositivos del quórum se conectan a uno de 1 a través de N nodos y el nodo $N+1$, el clúster sobrevive al fallo de cualquiera de ellos 1 mediante N nodos o cualquiera de los $N/2$ nodos. Esta disponibilidad asume que el dispositivo de quórum está funcionando correctamente.
- En una configuración de N - nodos donde un único dispositivo del quórum se conecta a todos los nodos, el clúster puede sobrevivir al fallo de cualquiera de los $N-1$ nodos. Esta disponibilidad asume que el dispositivo de quórum está funcionando correctamente.
- En una configuración de N nodos donde un único dispositivo de quórum se conecta a todos los nodos, el clúster puede sobrevivir al fallo del dispositivo de quórum si todos los nodos del clúster están disponibles.

Para ver los tipos de configuraciones del quórum que se deben evitar, consulte [“Configuraciones de quórum malas ” en la página 63](#). Para ver los tipos de configuraciones del quórum recomendadas, consulte [“Configuraciones de quórum recomendadas ” en la página 60](#).

Cumplimiento de los requisitos de dispositivos de quórum

Debe cumplir los siguientes requisitos. Si hace caso omiso de estos requisitos, la disponibilidad del clúster puede verse comprometida.

- Asegúrese de que Sun Cluster admite el dispositivo específico como dispositivo de quórum.

Nota – Para obtener una lista de los dispositivos específicos que Sun Cluster admite como dispositivos de quórum, póngase en contacto con el proveedor de servicios Sun.

Sun Cluster admite dos tipos de dispositivos de quórum:

- Discos compartidos con varios sistemas que admiten reservas SCSI-3 PGR
- Discos compartidos de doble sistema que admiten reservas SCSI-2
- En una configuración de dos nodos, debe configurar al menos un dispositivo de quórum para asegurar que un único nodo puede continuar si el otro nodo falla. Consulte la [Figura 3-2](#).

Para ver los tipos de configuraciones del quórum que se deben evitar, consulte “[Configuraciones de quórum malas](#)” en la [página 63](#). Para ver tipos de configuraciones del quórum recomendadas, consulte “[Configuraciones de quórum recomendadas](#)” en la [página 60](#).

Mejores prácticas relacionadas con los dispositivos del quórum

Use la siguiente información para evaluar la mejor configuración de quórum para su topología:

- ¿Tiene un dispositivo capaz de estar conectado a todos los nodos del clúster?
 - En caso afirmativo, configure dicho dispositivo como el dispositivo de quórum. *No* necesita configurar otro dispositivo del quórum porque su configuración es óptima.



Precaución – Si ignora este requisito y añade otro dispositivo de quórum, el dispositivo de quórum adicional reduce la disponibilidad del clúster.

- En caso contrario, configure el dispositivo(s) de doble puerto.

- Asegúrese de que el número de votos aportados por los dispositivos de quórum es menor que el número de votos total aportados por los nodos. En caso contrario los nodos no pueden formar un clúster si todos los discos no están disponibles—, incluso si todos los nodos están funcionando.

Nota – En entornos concretos, puede que desee reducir la disponibilidad general del clúster para satisfacer sus necesidades. En estas situaciones, puede ignorar estas recomendaciones. Sin embargo, no cumplir estas recomendaciones reduce la disponibilidad general. Por ejemplo, en la configuración que se describe en [“Configuraciones de quórum atípicas ” en la página 63](#), el clúster está menos disponible: los votos del quórum superan los votos del nodo. El clúster tiene la propiedad de que si se pierde el acceso al almacenamiento compartido entre los nodos A y B, fallará el clúster entero.

Consulte [“Configuraciones de quórum atípicas ” en la página 63](#) para conocer las excepciones a estas mejores prácticas.

- Especifique un dispositivo de quórum entre cada par de nodos que compartan el acceso al dispositivo de almacenamiento. Esta configuración del quórum agiliza el proceso de aislamiento de fallos. Consulte [“Quórum en configuraciones de más de dos nodos ” en la página 61](#).
- En general, si la adición de un dispositivo de quórum iguala el número total de votos del clúster, la disponibilidad del clúster disminuye.
- Los dispositivos de quórum ralentizan ligeramente las reconfiguraciones después de que se una un nodo nuevo o se desactive uno antiguo. Por tanto, no agregue más dispositivos de quórum de los necesarios.

Para ver los tipos de configuraciones del quórum que se deben evitar, consulte [“Configuraciones de quórum malas ” en la página 63](#). Para ver tipos de configuraciones del quórum recomendadas, consulte [“Configuraciones de quórum recomendadas ” en la página 60](#).

Configuraciones de quórum recomendadas

Esta sección contiene ejemplos de configuraciones recomendadas del quórum. Para ver los tipos de configuraciones del quórum que se deben evitar, consulte [“Configuraciones de quórum malas ” en la página 63](#).

Quórum en configuraciones de dos nodos

Son necesarios dos votos de quórum para que se forme un clúster de dos nodos. Estos dos votos pueden proceder de los dos nodos del clúster o de un nodo y un dispositivo del quórum.

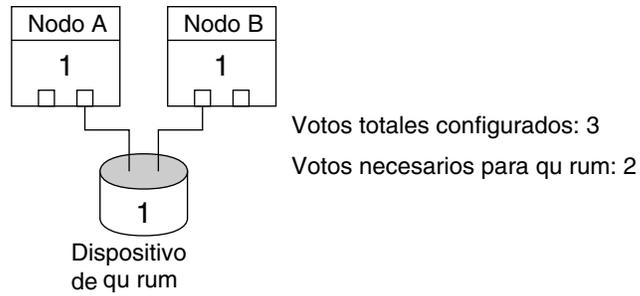
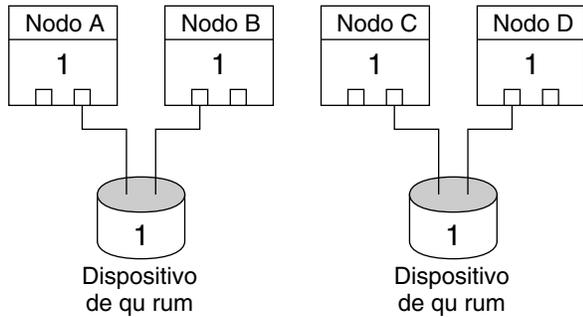


FIGURA 3-2 Configuración de dos nodos

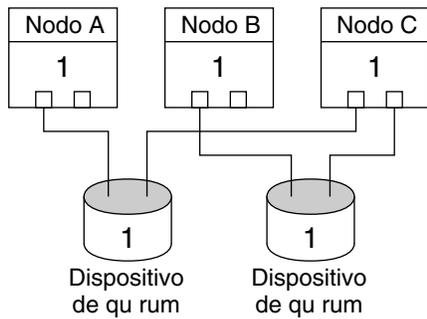
Quórum en configuraciones de más de dos nodos

Puede configurar un clúster con más de dos nodos sin dispositivo del quórum. No obstante, si lo hace así, no podrá iniciar el clúster sin una mayoría de nodos en el clúster.



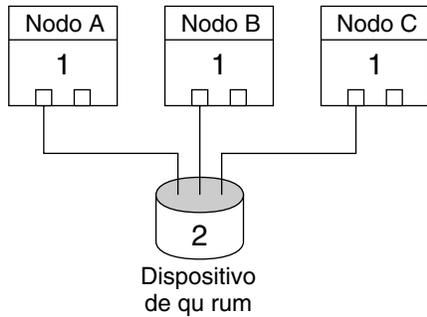
Votos totales configurados: 6
 Votos necesarios para quorum: 4

En esta configuración, cada par debe estar disponible para cada par para sobrevivir.



Votos totales configurados: 5
 Votos necesarios para quorum: 3

En esta configuración, las aplicaciones se configuran normalmente para ejecutarse en el Nodo A y el Nodo B y utilizar el Nodo C como respaldo.



Votos totales configurados: 5
 Votos necesarios para quorum: 3

En esta configuración, la combinación de uno o más nodos y el dispositivo de quorum pueden formar un cluster.

Configuraciones de quórum atípicas

En la [Figura 3-3](#) se da por hecho que se están ejecutando aplicaciones con un papel fundamental (una base de datos de Oracle, por ejemplo) en el nodo A y en el B. Si el Nodo A y el Nodo B no están disponibles y no se puede acceder a los datos compartidos, es posible que prefiera que todo el clúster se apague. En caso contrario, esta configuración no es óptima porque no proporciona una alta disponibilidad.

Para obtener información acerca de las mejores prácticas relacionadas con esta excepción, consulte [“Mejores prácticas relacionadas con los dispositivos del quórum”](#) en la [página 59](#).

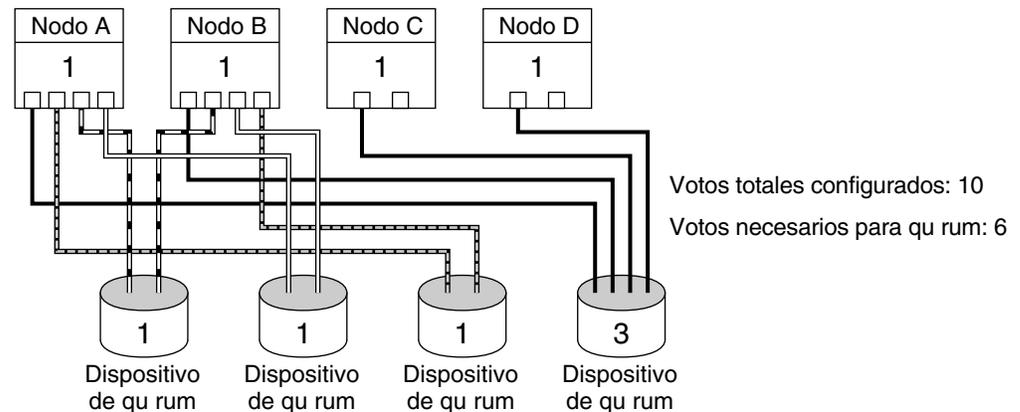
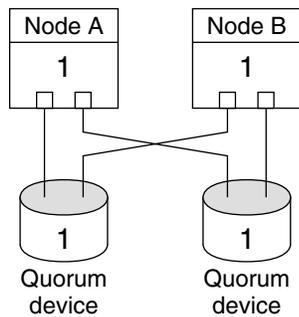


FIGURA 3-3 Configuración atípica

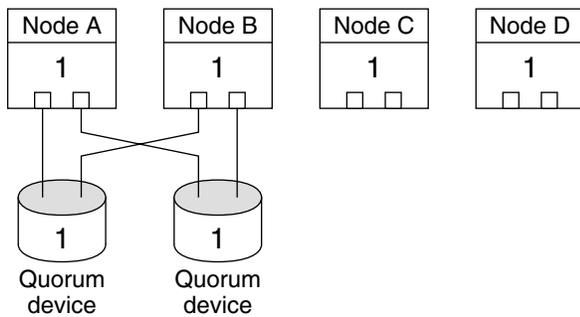
Configuraciones de quórum malas

Esta sección contiene ejemplos de configuraciones que se deben evitar. Para ver tipos de configuraciones del quórum recomendadas, consulte [“Configuraciones de quórum recomendadas”](#) en la [página 60](#).



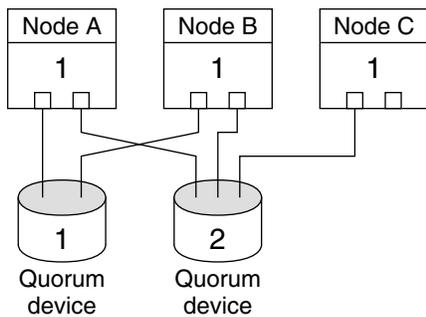
Total configured votes: 4
Votes required for quorum: 3

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.

Servicios de datos

El término *servicio de datos* describe una aplicación, como por ejemplo Oracle o Sun Java System Web Server, que se ha configurado para ejecutarse en un clúster en lugar de en un servidor individual. Un servicio de datos se compone de una aplicación, archivos de configuración especializados de Sun Cluster y métodos de gestión de Sun Cluster que controlan las acciones siguientes de la aplicación.

- Inicio
- Detener
- Supervisar y tomar acciones correctoras

Para obtener información acerca de los tipos de servicios de datos, consulte “Servicios de datos” de *Sun Cluster para el sistema operativo Solaris: Visión general*.

En la [Figura 3-4](#) se compara una aplicación que se ejecuta en un único servidor de aplicaciones (modelo de servidor único) con la misma aplicación pero ejecutándose en un clúster (modelo de servidor en clúster). La única diferencia entre las dos configuraciones es que la aplicación en clúster puede que se ejecute más rápidamente y tenga una mayor disponibilidad.

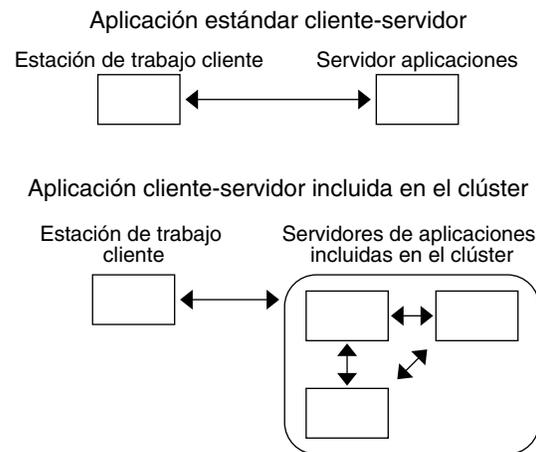


FIGURA 3-4 Configuración estándar y configuración cliente-servidor

En el modelo de servidor único, la aplicación se configura para que acceda al servidor a través de una interfaz de red pública concreta (un nombre de sistema). El nombre de sistema está asociado con este servidor físico.

En el modelo de servidor en clúster, la interfaz de red pública es un *nombre de sistema lógico* o una *dirección compartida*. El término *recursos de red* se utiliza para referirse a los nombres de sistema lógicos y para las direcciones compartidas.

Algunos servicios de datos requieren que se especifiquen los nombres de sistema lógicos o las direcciones compartidas como interfaces de red. Los nombres de sistema lógicos y las direcciones compartidas no se pueden intercambiar; otros, en cambio, permiten especificar nombres de sistema lógicos o direcciones compartidas. Consulte la instalación y la configuración para cada servicio de datos para obtener detalles acerca del tipo de interfaz que debe especificar.

Un recurso de red no está asociado a un servidor físico específico. Un recurso de red puede migrar entre servidores físicos.

Un recurso de red está asociado inicialmente a un nodo, el *primario*. Si el primario falla, el recurso de red y el recurso de aplicación se recuperan del fallo usando un nodo de clúster diferente (secundario). Cuando el recurso de red se recupera del problema, después de un breve intervalo, el recurso de aplicación continúa ejecutándose en el secundario.

En la **Figura 3-5** se compara un modelo de servidor único con un modelo de servidor en clúster. Tenga en cuenta que en el modelo de servidor basado en clúster un recurso de red (nombre de sistema lógico, en este ejemplo) puede trasladarse entre dos o más nodos del clúster. La aplicación está configurada para usar este nombre de sistema lógico en lugar de uno asociado a un servidor en particular.

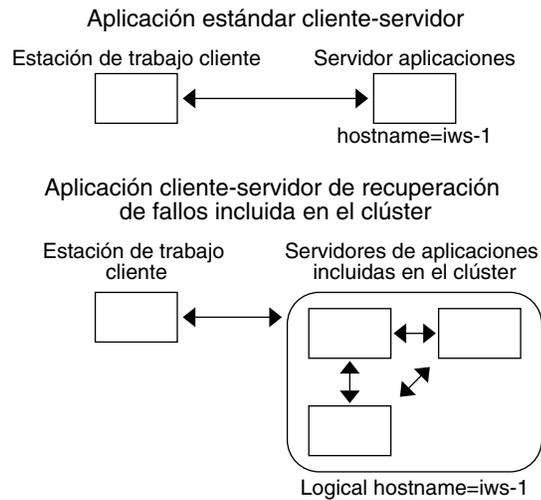


FIGURA 3-5 Comparación entre los nombres de sistema fijo y lógico

Una dirección de red también está asociada inicialmente a otro nodo. Este nodo se denomina nodo de interfaz global. Una dirección compartida (conocida como *interfaz global*) se utiliza como interfaz de red única para el clúster.

La diferencia entre el modelo de nombre de sistema lógico y el modelo de servicio escalable es que en este último, cada nodo tiene la dirección de red activamente configurada en su interfaz de bucle inverso. Esta configuración permite que haya

múltiples instancias de un servicio de datos activas en varios nodos simultáneamente. El término “servicio escalable” significa que puede agregarse más potencia de CPU a la aplicación agregando nodos del clúster adicionales con lo que el rendimiento se multiplicará.

Si el nodo de interfaz global falla, la dirección compartida puede iniciarse en otro nodo que también esté ejecutando una instancia de la aplicación (convirtiendo así al otro nodo en el nuevo nodo de interfaz global). O, la dirección compartida puede trasladarse a otro nodo del clúster que no estuviera ejecutando la aplicación previamente.

En la [Figura 3-6](#) se compara una configuración de servidor único con una configuración de servicio escalable basada en clúster. Tenga en cuenta que, en la configuración de servicio escalable, la dirección compartida está presente en todos los nodos. De forma análoga a como se usan los nombres de sistema para los servicios de datos de recuperación de fallos, la aplicación se configura para usar esta dirección compartida en lugar de un nombre de sistema asociado a un servidor determinado.

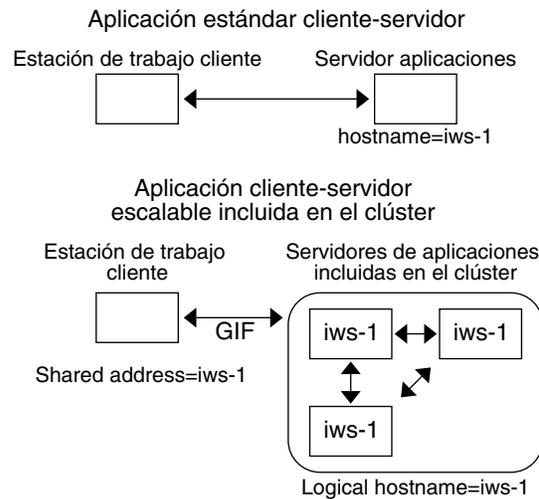


FIGURA 3-6 Comparación entre los nombres de sistema fijos y las direcciones compartidas

Métodos de servicios de datos

El software Sun Cluster proporciona un conjunto de métodos de gestión de servicios. Estos métodos se ejecutan bajo el control de Gestor de grupos de recursos (RGM), que los utiliza para iniciar, detener y supervisar la aplicación en los nodos del clúster. Estos métodos, junto con el software de la estructura del clúster y los dispositivos multisistema, permiten a las aplicaciones convertirse en servicios de datos a prueba de fallos o escalables.

RGM también gestiona los recursos del clúster, incluidas las instancias de una aplicación y de los recursos de red (nombres de sistema lógicos y direcciones compartidas).

Además de los métodos proporcionados mediante el software de Sun Cluster, el sistema Sun Cluster también proporciona una API y varias herramientas de desarrollo de servicio de datos. Estas herramientas permiten a los desarrolladores de aplicaciones desarrollar los métodos de servicios de datos necesarios para que otras aplicaciones se puedan ejecutar como servicios de datos de alta disponibilidad con el software Sun Cluster.

Servicios de datos de recuperación de fallos

Si el nodo en el que se está ejecutando el servicio de datos (el primario) falla, el servicio migra a otro nodo en funcionamiento sin intervención del usuario. Los servicios de recuperación de fallos usan un *grupo de recursos de recuperación de fallos*, que es un contenedor para los recursos de instancias de aplicaciones y recursos de la red (*nombres de sistemas lógicos*). Los nombres de sistemas lógicos son direcciones IP que pueden configurarse como activas en un nodo y, posteriormente, configurarse automáticamente como inactivas en el nodo original y activarse en otro nodo.

En los servicios de datos a prueba de fallos, las instancias de las aplicaciones sólo se ejecutan en un nodo individual. Si el supervisor de fallos detecta un error, intenta reiniciar la instancia en el mismo nodo o bien intenta iniciarla en otro nodo (el de recuperación de fallos). El resultado depende de la forma en que se haya configurado el servicio de datos.

Servicios de datos escalables

Los servicios de datos escalables tienen la capacidad de activar instancias en varios nodos; Los servicios escalables utilizan los dos siguientes grupos de recursos:

- Un *grupo de recursos escalables* contiene los recursos de aplicación.
- Un grupo de recursos de recuperación de fallos contiene los recursos de red (*direcciones compartidas*) de los que depende el servicio escalable.

El grupo de recursos escalable puede estar en línea en varios nodos, de forma que se puedan ejecutar varias instancias del servicio simultáneamente. El grupo de recurso a prueba de fallos que aloja la dirección compartida está disponible en un solo nodo cada vez. Todos los nodos que alojan un servicio escalable usan la misma dirección compartida para alojar el servicio.

Las solicitudes de servicio entran en el clúster a través de una interfaz de red única (la interfaz global). Estas solicitudes se distribuyen a los nodos, basándose en uno de los algoritmos predefinidos especificados por la *directiva de equilibrado de cargas* que el clúster puede usar para equilibrar la carga del servicio entre varios nodos. Puede haber varias interfaces globales en distintos nodos que alojen otras direcciones compartidas.

En los servicios escalables, las instancias de las aplicaciones se ejecutan en varios nodos simultáneamente. Si el nodo que aloja la interfaz global falla, ésta se traspassa a otro nodo. Si falla una instancia de aplicación que se está ejecutando, la instancia intenta reiniciarse en el mismo nodo.

Si no se puede reiniciar una instancia de una aplicación en el mismo nodo, y se configura otro nodo que no esté en uso para ejecutar el servicio, éste se transfiere a este nodo. De lo contrario, el servicio continúa ejecutándose en los nodos restantes, lo que posiblemente provocará una degradación del rendimiento del servicio.

Nota – El estado TCP para cada instancia de aplicación se conserva en el nodo de la instancia, no en el de la interfaz global. Por tanto el fallo en el nodo de interfaz global no afecta a la conexión.

La [Figura 3-7](#) muestra un ejemplo de grupo de recursos escalables y de recuperación de fallos y las dependencias que existen entre ellos en cuanto a los servicios escalables. Este ejemplo muestra tres grupos de recursos. El grupo de recursos de recuperación de fallos contiene recursos de aplicación que usan los DNS de alta disponibilidad y recursos de red que usan éstos y el servidor Apache Web Server de alta disponibilidad (sólo en clústeres basados en SPARC). Los grupos de recursos escalables sólo contienen instancias de la aplicación de Apache Web Server. Tenga en cuenta que las dependencias de los grupos de recursos existen entre los grupos de recursos de recuperación de fallos y los escalables (líneas sólidas). Adicionalmente, todos los recursos de aplicación de Apache dependen del recurso de red `schost-2`, que es una dirección compartida (líneas de puntos).

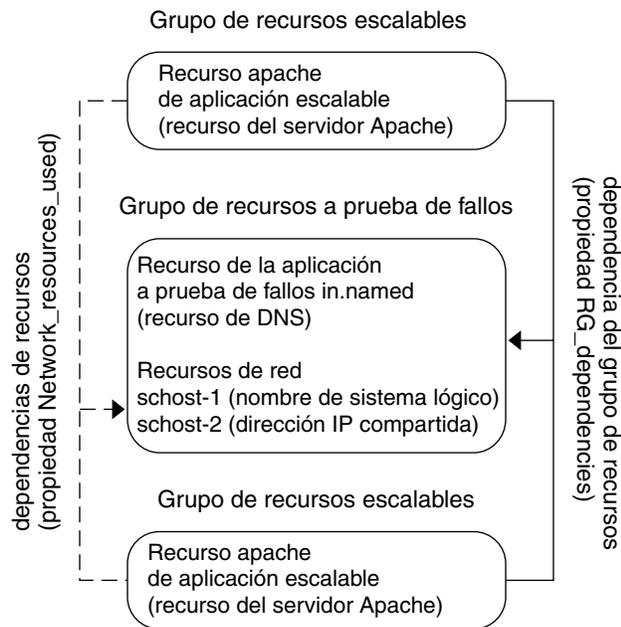


FIGURA 3-7 SPARC: Ejemplo de los grupos de recursos a prueba de fallos y escalables

Políticas de equilibrio de cargas

El equilibrio de cargas mejora el rendimiento del servicio escalable, tanto en tiempo de respuesta como en rendimiento. Hay dos clases de servicios de datos escalables.

- Puro
- Adosado

En un servicio *puro*, todas las instancias pueden responder a las solicitudes del cliente. En un servicio *adosado*, un cliente puede enviar solicitudes a la misma instancia. Estas peticiones no se redirigen a otras instancias.

Un servicio *puro* usa una política de equilibrio de cargas ponderada bajo la cual, predeterminadamente, las peticiones de los clientes se distribuyen de manera uniforme entre las instancias del servidor en el clúster. Por ejemplo, en un clúster de tres nodos, suponga que cada nodo tiene un peso de 1. Cada nodo atenderá 1/3 de las solicitudes procedentes de cualquier cliente en nombre de dicho servicio. El administrador puede cambiar el peso en cualquier momento usando la interfaz de comando `scrgadm (1M)` o la GUI de SunPlex Manager.

Un servicio *adosado* tiene dos perspectivas: *adosado normal* y *adosado con comodines*. Los servicios *adosados* permiten sesiones simultáneas de aplicación a través de varias conexiones TCP para que compartan la memoria en estado (estado de la sesión de aplicación).

Los normales permiten a un cliente compartir el estado entre varias conexiones TCP simultáneas. Se considera que el cliente es “adosado” con respecto a la instancia del servidor que recibe en un único puerto. Al cliente se le garantiza que todas sus solicitudes van a ir a la misma instancia del servidor, siempre que ésta permanezca activa y accesible y que la directiva de equilibrio de cargas no cambie mientras el servicio esté en línea.

Por ejemplo, un navegador Web en el cliente conecta con una dirección IP compartida en el puerto 80 usando tres conexiones TCP diferentes. Sin embargo, las conexiones intercambian con el servicio información de sesión almacenada en caché.

Una generalización de una directiva adosada amplía a varios servicios escalables dicha información de sesión de intercambio en segundo plano y en la misma instancia. Cuando estos servicios intercambian información de sesión en segundo plano y en la misma instancia, se considera que el cliente está “adosado” a varias instancias de servidor en el mismo nodo recibiendo información a través de distintos puertos. .

Por ejemplo, un cliente de un sitio de comercio electrónico llena su carro de la compra con artículos utilizando HTTP en el puerto 80. El cliente entonces cambia a SSL en el puerto 443 para enviar los datos de forma segura para pagar con tarjeta de crédito los artículos del carro.

Los servicios adosados *comodín* usan números de puerto asignados dinámicamente, pero siguen esperando que las peticiones de clientes vayan al mismo nodo. El cliente está “adosado con comodín” a los puertos que tienen la misma dirección IP.

Un buen ejemplo de esta política es el FTP de modalidad pasiva. Por ejemplo, un cliente se conecta a un servidor FTP mediante el puerto 21. El servidor entonces da instrucciones al cliente para que se vuelva a conectar a un servidor con puerto de escucha que esté en el intervalo de puertos dinámicos. Todos los requisitos para esta dirección IP se reenvían al mismo nodo mediante el que el servidor informó al cliente a través de la información de control .

Para cada una de estas directivas adosadas, la directiva de equilibrio de carga en función del peso está activada de forma predeterminada. Por lo tanto, una solicitud inicial del cliente se dirige a la instancia que indica el equilibrador de carga. Después de que el cliente establezca una afinidad para el nodo en el que se está ejecutando la instancia, las solicitudes futuras estarán dirigidas condicionalmente a dicha instancia. El nodo debe estar accesible y la directiva de equilibrado de carga no se debe haber modificado.

Los detalles adicionales de las directivas específicas de equilibrado de carga son las siguientes.

- **Ponderada:** la carga se distribuye entre varios nodos según valores de peso especificados. Esta directiva se define usando el valor `LB_WEIGHTED` para la propiedad `Load_balancing_weights`. Si no se establece explícitamente el peso de un nodo, éste toma de forma predeterminada el valor uno.

La política de ponderación redirige cierta parte del tráfico de los clientes a un nodo determinado. Si consideramos que X =peso y A =el peso total de todos los nodos activos, un nodo activo recibirá aproximadamente X/A del total de conexiones nuevas que se dirijan al nodo activo. Sin embargo, el número total de conexiones debe ser lo suficientemente grande. Esta política no trata de peticiones individuales.

Tenga en cuenta que esta política no es de tipo giratoria. Una política de este tipo provocará que cada solicitud de un cliente vaya a un nodo distinto. Por ejemplo, la primera solicitud iría al nodo 1, la segunda al nodo 2, etc.

- Adosada: en esta norma el conjunto de puertos se da a conocer en el momento en el que se configuran los recursos de la aplicación. Esta directiva se define usando el valor `LB_STICKY` para la propiedad `Load_balancing_policy`.
- Adosada-comodín: esta política tiene menos limitaciones que la “adosada” normal. En el caso de un servicio escalable que esté identificado mediante la dirección IP, los puertos los asigna el servidor (y no se conocen de antemano). Los puertos podrían cambiar. Esta directiva se define usando el valor `LB_STICKY_WILD` para la propiedad `Load_balancing_policy`.

Configuración de retroceso

Los grupos de recurso se cubren entre nodos. Cuando se produce una recuperación de fallos, el nodo secundario original pasa a ser el nuevo primario. La configuración de retroceso especifica las acciones que se producirán cuando el nodo principal vuelva a estar en línea. Las opciones son hacer que el primario original se convierta de nuevo en primario (retroceso) o permitir que el que haya tomado su papel continúe con él. Las opciones se especifican usando la configuración de la propiedad del grupo de recursos `Failback`.

Si falla el nodo original que alberga el grupo de recursos y se reinicia varias veces, la configuración de un retroceso podría reducir la disponibilidad del grupo de recursos.

Supervisor de fallos del servicio de datos

Cada servicio de datos de Sun Cluster proporciona un supervisor de fallos que explora periódicamente el servicio de datos para determinar su buen estado. Un supervisor de fallos comprueba que los daemons de las aplicaciones estén funcionando y que se esté dando servicio a los clientes. De acuerdo con la información que devuelven las comprobaciones, se pueden llevar a cabo acciones predefinidas, como reiniciar daemons o producir una recuperación de fallos.

Desarrollo de los nuevos servicios de datos

Sun proporciona archivos de configuración y plantillas de métodos de gestión que permiten hacer funcionar varias aplicaciones como servicios a prueba de fallos o escalables dentro de un clúster. Si Sun no le ofrece la aplicación que necesita ejecutar como servicio escalable o de recuperación de fallos, tiene una alternativa. Use una API de Sun Cluster o la API DSET para configurar la aplicación para que se ejecute como servicio escalable o de recuperación de fallos. Sin embargo, no todas las aplicaciones pueden convertirse en servicios escalables.

Características de los servicios escalables

Hay una serie de criterios que determinan si una aplicación puede convertirse o no en servicio escalable. Para determinar si su aplicación puede convertirse en servicio escalable, consulte “Análisis de la validez de la aplicación” de *Sun Cluster: Guía del desarrollador de los servicios de datos del sistema operativo Solaris*. Este conjunto de criterios se resume a continuación.

- Primero, el servicio se compone de una o varias *instancias* de servidor, cada una de las cuales se ejecuta en un nodo distinto del clúster. Dos o más instancias del mismo servicio no pueden ejecutarse en el mismo nodo.
- Segundo, si el servicio proporciona un almacén de datos lógico externo, deberá actuar con precaución. El acceso simultáneo a este almacén desde varias instancias de servidor debe estar sincronizado para evitar que se pierdan actualizaciones o datos de lectura mientras se estén modificando. Observe que se usa “externo” para distinguir este almacén del estado en memoria. El término “lógico” indica que el almacén aparece como una única entidad, aunque puede que esté replicado. Además, este almacén de datos lógico tiene la propiedad de que cada vez que una instancia de servidor actualiza el almacén, las demás instancias pueden “ver” la actualización de forma inmediata.

El sistema de Sun Cluster proporciona un almacén externo a través de su sistema de archivos en clúster y sus particiones sin formato globales. Como ejemplo, supongamos que un servicio escribe datos nuevos a un archivo de registro cronológico externo o modifica los datos que haya. Cuando se ejecutan varias instancias de este servicio, cada instancia tiene acceso a su registro externo y cada una de ellas puede acceder simultáneamente a este registro. Cada instancia debe sincronizar su acceso a este registro o de lo contrario se interferirán entre ellas. El servicio puede usar normalmente el bloqueo de archivos de Solaris mediante `fcntl(2)` y `lockf(3C)` para lograr la sincronización necesaria.

Otro ejemplo de este tipo de almacén es una base de datos de fondo (back-end), como pueda ser una instalación de Oracle Real Application Clusters Guard de alta disponibilidad para clústeres basados en SPARC o en Oracle. Este tipo de

servidores de bases de datos de fondo (back-end) proporcionan sincronización integrada utilizando una consulta de base de datos o transacciones de actualización. En consecuencia, varias instancias de servidor no necesitan implementar su propia sincronización.

El servidor Sun IMAP es un ejemplo de servicio que no es escalable. El servicio actualiza un almacenamiento, pero es privado y cuando varias instancias de IMAP escriben en él, se sobrescriben porque las actualizaciones no están sincronizadas. El servidor IMAP debe volver a escribirse para que se sincronice el acceso simultáneo.

- Por último, tenga en cuenta que las instancias pueden tener datos privados que no tengan nada que ver con los datos de otras instancias. En estos casos, el servicio no necesita un acceso simultáneo sincronizado porque los datos son privados y sólo puede gestionarlos la propia instancia. En estas circunstancias, deberá tener cuidado de no almacenar estos datos privados en el sistema de archivos en clúster porque estarían accesibles globalmente.

API de servicio de datos y API de biblioteca de desarrollo de servicios de datos

El sistema Sun Cluster proporciona las siguientes funciones para que las aplicaciones tengan una alta disponibilidad:

- Servicios de datos que se suministran como parte del sistema Sun Cluster
- Una API de servicio de datos
- Una API de biblioteca de desarrollo para los servicios de datos
- Un servicio de datos “genérico”

En *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* se describe la forma de instalar y configurar los servicios de datos que se suministran con el sistema Sun Cluster. En *Sun Cluster 3.1 9/04 Software Collection for Solaris OS (SPARC Platform Edition)* se describe cómo se gestionan otras aplicaciones para que tengan una alta disponibilidad en la estructura de Sun Cluster.

La API de Sun Cluster permite que los desarrolladores de aplicaciones puedan desarrollar supervisores de fallos y secuencias de comandos para iniciar y detener las instancias de servicios de datos. Con estas herramientas, una aplicación se puede implementar para que funcione como servicio de datos escalable o de recuperación de fallos. El sistema Sun Cluster proporciona un servicio de datos “genérico”. Use este servicio de datos genérico para generar rápidamente los métodos de inicio y detención que requiera una aplicación, así como implementar los servicios de datos como servicios escalables o de recuperación de fallos.

Uso de interconexiones del clúster para el tráfico de servicios de datos

Un clúster debe tener varias conexiones de red entre los nodos que formen la interconexión del clúster. El software Sun Cluster utiliza varias interconexiones para lograr los siguientes objetivos:

- Garantizar una alta disponibilidad
- Mejorar el rendimiento

Para el tráfico interno (como datos de sistema de archivos o datos de servicios escalables), los mensajes se desvían a través de las interconexiones disponibles de forma giratoria. La interconexión del clúster también está disponible para aplicaciones, para comunicaciones de alta disponibilidad entre nodos. Por ejemplo, una aplicación distribuida podría tener componentes que se ejecutaran en distintos nodos y que necesitasen comunicarse. Al usar la interconexión del clúster en lugar del transporte público, éstas conexiones pueden soportar el fallo de un enlace individual.

Para usar la interconexión del clúster para la comunicación entre nodos, una aplicación necesita los nombres de sistema privados configurados cuando se instaló Sun Cluster. Por ejemplo, si el nombre de sistema privado para el nodo 1 es `clusternode1-priv`, use ese nombre para comunicarse mediante la interconexión del clúster con el nodo 1. Los zócalos TCP abiertos usando este nombre se dirigen por la interconexión del clúster y pueden redirigirse de forma transparente en caso de fallo de la red.

Dado que es posible configurar nombres de sistema privados durante la instalación de Sun Cluster, la interconexión del clúster utiliza cualquier nombre que se elija en ese momento. Para determinar cuál es el nombre real, use el comando `scha_cluster_get(3HA)` con el argumento `scha_privatelink_hostname_node`.

Tanto las comunicaciones de las aplicaciones como las comunicaciones internas del clúster se distribuyen a través de las interconexiones. Como las aplicaciones comparten la interconexión del clúster con el tráfico interno de éste, el ancho de banda disponible para las aplicaciones depende del ancho de banda que utilice el resto del tráfico del clúster. Si se produce un fallo, el tráfico interno y el de la aplicación se distribuyen por las interconexiones disponibles.

A cada nodo se le asigna también una dirección por nodo fija. Esta dirección está incluida en el controlador `clprivnet`. La dirección IP hace referencia al nombre de sistema privado del nodo: `clusternode1-priv`. Para obtener información acerca del controlador de red privado de Sun Cluster, consulte la página de comando `man clprivnet(7)`.

Si la aplicación requiere direcciones IP coherentes en todos los puntos, configure la aplicación para enlazar a la dirección por nodo tanto en el cliente como en el servidor. Todas las conexiones parece entonces que se originan en la dirección por nodo y que vuelven a ella.

Tipos de recursos, grupos de recursos y recursos

Los servicios de datos usan varios tipos de *recursos*: Las aplicaciones como Apache Web Server o Sun Java System Web Server usan direcciones de red (nombres lógicos del sistema y direcciones compartidas) de las que dependen las aplicaciones. Los recursos de aplicación y red forman una unidad básica que gestiona RGM.

Los servicios de datos son tipos de recursos. Por ejemplo, Sun Cluster HA para Oracle es el tipo de recurso `SUNW.oracle-server` y Sun Cluster HA for Apache es `SUNW.apache`.

Un recurso es una concreción de *tipo de recurso* que está definida a nivel del clúster. Se definen varios tipos de recursos.

Los recursos de red pueden ser tipos de recursos `SUNW.LogicalHostname` o `SUNW.SharedAddress`. Estos dos tipos de recursos los ha registrado previamente el software Sun Cluster.

Los tipos de recursos `HASStorage` y `HASStoragePlus` se utilizan para sincronizar el inicio de los recursos y los grupos de dispositivos de disco de los que dependen los recursos. Estos tipos de recursos garantizan que antes de que se inicie un servicio de datos, estén disponibles las rutas a los puntos de montaje, los dispositivos globales y los nombres de grupos de dispositivos del sistema de archivos del clúster. Para obtener más información, consulte "Synchronizing the Startups Between Resource Groups and Disk Device Groups" en la *Data Services Installation and Configuration Guide*. El tipo de recurso `HASStoragePlus` comenzó a estar disponible en Sun Cluster 3.0 5/02 y se le ha agregado otra función que hace que los sistemas de archivos locales tengan una alta disponibilidad. Para obtener más información acerca de esta función, consulte "Tipo de recurso de `HASStoragePlus` " en la página 50.

Los recursos administrados mediante RGM se ubican en grupos llamados *grupos de recursos*, por lo que se pueden administrar como si fueran una unidad. El grupo de recursos migra como unidad si se inicia una recuperación de fallos o un cambio.

Nota – Cuando se pone en línea un grupo de recursos que contiene recursos de aplicación, se inicia la aplicación en cuestión. El método de inicio del servicio de datos espera a que la aplicación esté funcionando para cerrarse correctamente. Determinar cuándo la aplicación está completamente en marcha sigue el mismo proceso que el supervisor de fallos utiliza para saber que un servicio de datos está ofreciendo servicio a clientes. Consulte la *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* para obtener más información sobre el proceso.

Gestor de grupos de recursos (RGM)

RGM controla los servicios de datos (aplicaciones) como recursos, que las implementaciones de *tipos de recursos* gestionan. Éstas las proporciona Sun o las crea un desarrollador con una plantilla genérica de servicios de datos, una API de la biblioteca de desarrollo de servicios de datos (API DSDL) o con una API de gestión de recursos (RMAPI). El administrador del clúster crea y gestiona los recursos en contenedores que reciben el nombre de *grupos de recursos*. RGM para e inicia los grupos de recursos en nodos seleccionados como respuesta a cambios en la pertenencia al clúster.

RGM actúa en *recursos* y *grupos de recursos*. Las acciones de RGM provocan que los recursos y los grupos de recursos pasen de estar en línea a estar fuera de línea. Encontrará una descripción completa de los estados y las configuraciones se pueden aplicar a los recursos y a los grupos de recursos en la sección “[Configuraciones y estados de los recursos y los grupos de recursos](#)” en la página 77.

Consulte “[Configuración de un proyecto de servicio de datos](#)” en la página 79 para obtener información acerca de cómo iniciar proyectos de Solaris bajo el control de RGM.

Configuraciones y estados de los recursos y los grupos de recursos

Los administradores aplican a recursos y grupos de recursos configuraciones estáticas que sólo pueden cambiarse con acciones administrativas. RGM asigna a los recursos “estados” dinámicos. Estos valores y estados se describen en la lista siguiente.

- Gestionados o no gestionados: son valores que afectan a todo el clúster y sólo se aplican a grupos de recursos. RGM gestiona los grupos de recursos. El comando `scrgadm (1M)` se puede usar para que RGM gestione o deje de gestionar un grupo de recursos. Estos valores no cambian con la reconfiguración del clúster.

Cuando se crea un grupo de recursos por primera vez, no se gestiona. Un grupo de recursos debe estar administrado antes de activar los recursos colocados en el grupo.

En algunos servicios de datos, por ejemplo, un servidor Web escalable, el trabajo se debe hacer antes de iniciar los recursos de red y después de detenerlos. Este trabajo se hace con los métodos de servicio de datos de inicialización (INIT) y finalización (FINI). Los métodos INIT sólo se ejecutan si el grupo de recursos en el que éstos residen está en estado gestionable.

Cuando un grupo de recursos cambia de no gestionado a gestionado, los métodos INIT registrados para el grupo se ejecutan en todos los recursos.

Cuando un grupo de recursos cambia de gestionado a no gestionado, todos los métodos FINI registrados se ejecutan para realizar una limpieza.

Los métodos INIT y FINI se utilizan habitualmente para los recursos de red para servicios escalables. Sin embargo, estos métodos se pueden utilizar para cualquier tarea de inicio o de limpieza que no ejecute la aplicación.

- **Habilitados o inhabilitados:** son los valores a nivel del clúster que se aplican a los recursos. El comando `scrgadm (1M)` se puede usar para habilitar o inhabilitar los recursos. Estos valores no cambian con la reconfiguración del clúster.

El valor normal para un recurso es que esté habilitado y funcionando activamente en el sistema.

Si desea que un recurso no esté disponible en ningún nodo del clúster, deshabilite el recurso. De esta manera dejará de estar disponible para uso general.

- **En línea o fuera de línea:** son estados dinámicos y se aplican a recursos y grupos de recursos.

El estado cambia de en línea a fuera de línea a medida que el clúster avanza por los distintos pasos de reconfiguración durante los procesos de recuperación de fallos o conmutación por fallos. El estado también puede cambiar mediante las acciones administrativas. Use el comando `scswitch (1M)` para cambiar del estado en línea al estado fuera de línea de un recurso o grupo de recursos.

Un recurso o un grupo de recursos a prueba de fallos sólo pueden estar en línea en un nodo simultáneamente. Un recurso o un grupo de recursos escalables pueden estar en línea en algunos nodos y fuera de línea en otros. Durante una conmutación o una recuperación de fallos, los grupos de recursos y los recursos que contienen se ponen en fuera de línea en un nodo y después se vuelven a poner en línea en otro distinto.

Si un grupo de recursos está fuera de línea, entonces todos sus recursos estarán también fuera de línea. Si un grupo de recursos está en línea, entonces todos sus recursos habilitados estarán también en línea.

Los grupos de recursos pueden contener varios recursos, pudiendo haber dependencias entre ellos que requieren que los recursos se pongan en línea y fuera de línea en un orden determinado. Los métodos usados para poner los recursos en línea y fuera de línea pueden ocupar tiempos distintos en cada uno de ellos.

Debido a las dependencias de recursos y a las diferencias de tiempo de inicio y finalización, los recursos de un mismo grupo de recursos pueden tener estados de puesta en línea y fuera de línea distintos durante una reconfiguración del clúster.

Propiedades de los recursos y los grupos de recursos

Puede configurar valores de propiedades para los recursos y los grupos de recursos de sus servicios de datos de Sun Cluster. Las propiedades estándar son comunes a todos los servicios de datos. Las propiedades de extensión son específicas de cada servicio de datos. Algunas propiedades estándar y de extensión se configuran con valores predeterminados para que no se tengan que modificar. Otras necesitan configurarse como parte del proceso de creación y configuración de recursos. La documentación de cada servicio de datos especifica qué propiedades de recurso pueden establecerse y cómo hacerlo.

Las propiedades estándar se usan para configurar propiedades de recursos y grupos de recursos que normalmente son independientes de todos los servicios de datos. Para conocer el conjunto de propiedades estándar, consulte el Apéndice A, "Standard Properties" de *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

Las propiedades de extensión de RGM ofrecen información como la ubicación de los binarios de la aplicación y los archivos de configuración. Las propiedades de extensión se modifican a medida que se configuran los servicios de datos. El conjunto de propiedades de extensión se describe en la guía específica para servicios de datos.

Configuración de un proyecto de servicio de datos

Los servicios de datos se pueden configurar para que se inicien con un nombre de proyecto de Solaris cuando RGM los ponga en línea. La configuración asocia un recurso o un grupo de recursos gestionados por RGM con un OD de proyecto de Solaris. La asignación de un recurso o un grupo de recursos a un ID de proyecto le permite utilizar controles sofisticados que están disponibles en el sistema operativo Solaris para gestionar las cargas de trabajo y el consumo en el clúster.

Nota – Podrá llevar a cabo esta configuración sólo si está utilizando la versión actual del software Sun Cluster con Solaris 9 como mínimo.

El uso de la función de gestión de Solaris en un entorno de Sun Cluster le permite garantizar que las aplicaciones más importantes tengan prioridad a la hora de compartir un nodo con otras aplicaciones. Las aplicaciones podrían compartir un nodo si hay servicios consolidados o porque se haya producido una recuperación de fallos.

El uso de las funciones de gestión descritas en este documento puede mejorar la disponibilidad de una aplicación que sea básica impidiendo que otras aplicaciones con una prioridad inferior consuman en exceso suministros del sistema como, por ejemplo, el tiempo de la CPU.

Nota – La documentación de Solaris sobre esta función asigna el término “recursos” al tiempo de la CPU, los procesos, las tareas y otros componentes similares. Mientras que en la documentación de Sun Cluster, se utiliza el término “recursos” para describir entidades que están bajo el control de RGM. En la siguiente sección se utilizará el término “recurso” para referirse a las entidades de Sun Cluster que están bajo el control de RGM. En esta sección se utiliza el término “suministros” para referirse al tiempo de la CPU, los procesos y las tareas.

Esta sección ofrece una descripción conceptual de la configuración de los servicios de datos para ejecutar procesos en un `project(4)` especificado de Solaris 9. En esta sección también se describen varias situaciones de recuperación de fallos, así como sugerencias para planificar el uso de la función de gestión que proporciona el sistema operativo Solaris.

Para obtener documentación sobre los conceptos y los procedimientos relacionados con la función de gestión, consulte el Capítulo 1, “Network Service (Overview)” de *System Administration Guide: Network Services*.

Cuando vaya a configurar recursos y grupos de recursos para usar la función de gestión de Solaris en un clúster, use los siguientes procedimientos generales:

1. Configure las aplicaciones como parte del recurso.
2. Configure los recursos como parte del grupo de recursos.
3. Habilite los recursos en el grupo de recursos.
4. Haga que el grupo de recursos esté administrado.
5. Cree un proyecto de Solaris para el grupo de recursos.
6. Configure propiedades estándar para asociar el nombre del grupo de recursos al proyecto creado en el paso 5.
7. Ponga en línea el grupo de recursos.

Para configurar las propiedades estándar `Resource_project_name` o `RG_project_name` para asociar el ID de proyecto de Solaris al recurso o grupo de recursos, use la opción `--y` con el comando `scrgadm(1M)`. Configure los valores de la propiedad al recurso o grupo de recursos. Consulte el Apéndice A, “Standard Properties” de *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* para ver las definiciones de las propiedades. Consulte `r_properties(5)` y `rg_properties(5)` para ver las descripciones de las propiedades.

El nombre de proyecto especificado debe existir en la base de datos de proyectos (`/etc/project`) y el superusuario debe estar configurado como miembro de dicho proyecto. Consulte el Capítulo 2, “Projects and Tasks (Overview)” de *System*

Administration Guide: Solaris Containers-Resource Management and Solaris Zones para obtener información conceptual sobre la base de datos de nombres de proyectos. Consulte `project(4)` para conocer la sintaxis del archivo del proyecto.

Cuando RGM pone en línea el recurso o grupo de recursos, ejecuta los procesos relacionados que hay bajo el nombre del proyecto.

Nota – Los usuarios pueden asociar recursos o grupos de recursos con proyectos en cualquier momento. Sin embargo, el nombre del nuevo proyecto no entra en vigor hasta que el recurso o el grupo de recursos se ponga fuera de línea y luego en línea de nuevo usando RGM.

Ejecutar recursos y grupos de recursos bajo el nombre del proyecto permite configurar las funciones siguientes para gestionar suministros de sistema en todo el clúster.

- Contabilidad ampliada: ofrece una forma flexible de registrar el consumo según las tareas o los procesos. La contabilidad ampliada permite examinar el uso histórico y evaluar las necesidades de capacidad para cargas de trabajo futuras.
- Controles: proporcionan un mecanismo para ajustarse a los suministros del sistema. Puede evitarse que los procesos, tareas y proyectos consuman grandes cantidades de suministros de sistema especificados.
- Programación de partición justa (FSS): ofrece la posibilidad de controlar la ubicación de tiempo de CPU disponible entre cargas de trabajo según su importancia. Ésta se mide por el número de particiones de tiempo de CPU que se asigna a cada carga de trabajo. Consulte las siguientes páginas de comando man para obtener más información:
 - `dispadmin(1M)`
 - `priocntl(1)`
 - `ps(1)`
 - `FSS(7)`
- Agrupaciones: ofrecen la posibilidad de usar particiones para aplicaciones interactivas según los requisitos de éstas. La agrupaciones pueden usarse para particionar un servidor que admite distintas aplicaciones de software. El uso de agrupaciones produce una respuesta más predecible para cada aplicación.

Determinación de requisitos para la configuración de proyectos

Antes de configurar los servicios de datos para que utilicen los controles proporcionados por Solaris en un entorno de Sun Cluster, debe decidir cómo desea controlar y supervisar los recursos en los procesos de recuperación de fallos y conmutación. Identifique las dependencias del clúster antes de configurar un nuevo proyecto. Por ejemplo, los recursos y grupos de recursos dependen de grupos de dispositivos de disco.

Use las propiedades de grupo de recursos `nodelist`, `failback`, `maximum primaries` y `desired primaries` que están configuradas con `scrgadm(1M)` para identificar las prioridades de la lista de nodos para el grupo de recursos.

- Para obtener una breve información sobre las dependencias existentes entre los grupos de recursos y los grupos de dispositivos de discos, consulte “Relationship Between Resource Groups and Disk Device Groups” de *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.
- Para ver una descripción detallada de las propiedades, consulte `rg_properties(5)`.

Use las propiedades `preferenced property` y `failback` que están configuradas con `scrgadm(1M)` y `scsetup(1M)` para determinar las prioridades de la lista de nodos del grupo de dispositivos.

- Para obtener información conceptual acerca de la propiedad `preferenced`, consulte “Grupos de dispositivos de disco multipuerto ” en la página 45.
- Para obtener información sobre los procedimientos, consulte “How To Change Disk Device Properties” en la “Administración de grupos de dispositivos de disco” de *Sun Cluster: Guía de administración del sistema para el SO Solaris*.
- Para obtener información conceptual acerca de la configuración del nodo y del comportamiento de los servicios de datos escalables y de recuperación de fallos, consulte la “Componentes de software y hardware del sistema Sun Cluster ” en la página 21.

Si se configura todos los nodos del clúster de forma idéntica, los límites de uso se hacen cumplir de forma idéntica en los nodos primarios y secundarios. Los parámetros de configuración de los proyectos no deben ser idénticos para todas las aplicaciones en los archivos de configuración de todos los nodos. Todos los proyectos que estén asociados a la aplicación deben, como mínimo, estar accesibles para la base de datos en todos los posibles maestros de esta aplicación. Suponga que *phys-schost-1* actúa como maestro de la Aplicación 1, pero ésta podría conmutarse o efectuar la recuperación de fallos con *phys-schost-2* o *phys-schost-3*. El proyecto que esté asociado a la aplicación 1 debe estar accesible en los tres nodos (*phys-schost-1*, *phys-schost-2* y *phys-schost-3*).

Nota – La información de la base de datos del proyecto puede ser un archivo de base de datos local `/etc/project` o puede almacenarse en la asignación NIS o en el servicio de directorio LDAP.

El sistema operativo Solaris hace posible una configuración flexible de los parámetros de uso y Sun Cluster impone pocas restricciones. Las opciones de configuración dependen de las necesidades de la sede. Considere las directrices generales de los apartados siguientes antes de configurar los sistemas.

Establecimiento de límites de memoria virtual por proceso

Configure el control `process.max-address-space` para limitar la memoria virtual a cada proceso. Consulte `rctladm(1M)` para obtener información detallada acerca de la configuración del valor de `process.max-address-space`.

Cuando use los controles de gestión con el software Sun Cluster, configure los límites de memoria de forma adecuada para impedir las recuperaciones de fallos innecesarias y el efecto “ping-pong” de las aplicaciones. En general, deberá seguir las siguientes directrices.

- No configure unos límites de memoria demasiado bajos.
Cuando una aplicación llegue a su límite de memoria, podría producirse una recuperación de fallos. Esta directriz es especialmente importante para aplicaciones de base de datos, ya que al alcanzar un límite de memoria virtual se pueden producir consecuencias inesperadas.
- No configure límites de memoria idénticos en nodos primarios y secundarios.
Límites idénticos podrían causar un efecto ping-pong cuando una aplicación alcance su límite de memoria y se recupere el error en un nodo secundario con un límite de memoria idéntico. Configure el límite de memoria un poco por encima en el nodo secundario. La diferencia en los límites de memoria ayuda a evitar la situación ping-pong y da al administrador del sistema un periodo de tiempo en el que ajustar los parámetros como sea necesario.
- Use los límites de memoria de la gestión de recursos para el equilibrio de cargas.
Por ejemplo, puede usar límites de memoria para evitar que una aplicación que no funcione bien consuma demasiado espacio de intercambio en disco.

Casos de recuperación de fallos

Puede configurar parámetros de gestión para que la asignación en la configuración del proyecto (`/etc/project`) funcione normalmente en el clúster y en las situaciones de conmutación o recuperación de fallos.

Los apartados siguientes son casos de ejemplo.

- Los dos primeros apartados, “Clúster de dos nodos con dos aplicaciones” y “Clúster de dos nodos con tres aplicaciones”, muestran escenarios de recuperación de fallos para nodos completos.
- El apartado “Recuperación de fallos sólo de grupos de recursos” ilustra el funcionamiento de la recuperación de fallos sólo para una aplicación.

En un entorno de Sun Cluster las aplicaciones se configuran como parte de un recurso. A continuación, debe configurar el recurso como parte de un grupo de recursos (RG). Cuando se produce un fallo, el grupo de recursos, junto con sus aplicaciones asociadas, se recupera del fallo con otro nodo. En los ejemplos siguientes los recursos no se muestran específicamente. Se presupone que cada recurso sólo tiene una aplicación.

Nota – La recuperación de fallos se produce en el orden de lista de nodos especificado en la RGM.

Los ejemplos siguientes tienen estas limitaciones:

- La aplicación 1 (App-1) está configurada en el grupo de recursos RG-1.
- La aplicación 2 (App-2) está configurada en el grupo de recursos RG-2.
- La aplicación 3 (App-3) está configurada en el grupo de recursos RG-3.

Aunque los números de particiones asignadas siguen siendo los mismos, el porcentaje de tiempo de CPU asignado a cada aplicación cambia después de la recuperación de fallos. Este porcentaje depende del número de aplicaciones que se están ejecutando en el nodo y del número de particiones que se han asignado a cada aplicación activa.

En estos casos, se asumen las configuraciones siguientes.

- Todas las aplicaciones están configuradas bajo un proyecto común.
- Cada recurso dispone de una única aplicación.
- Las aplicaciones son los únicos procesos activos de los nodos.
- Las bases de datos de los proyectos están configuradas igual en todos los nodos del clúster.

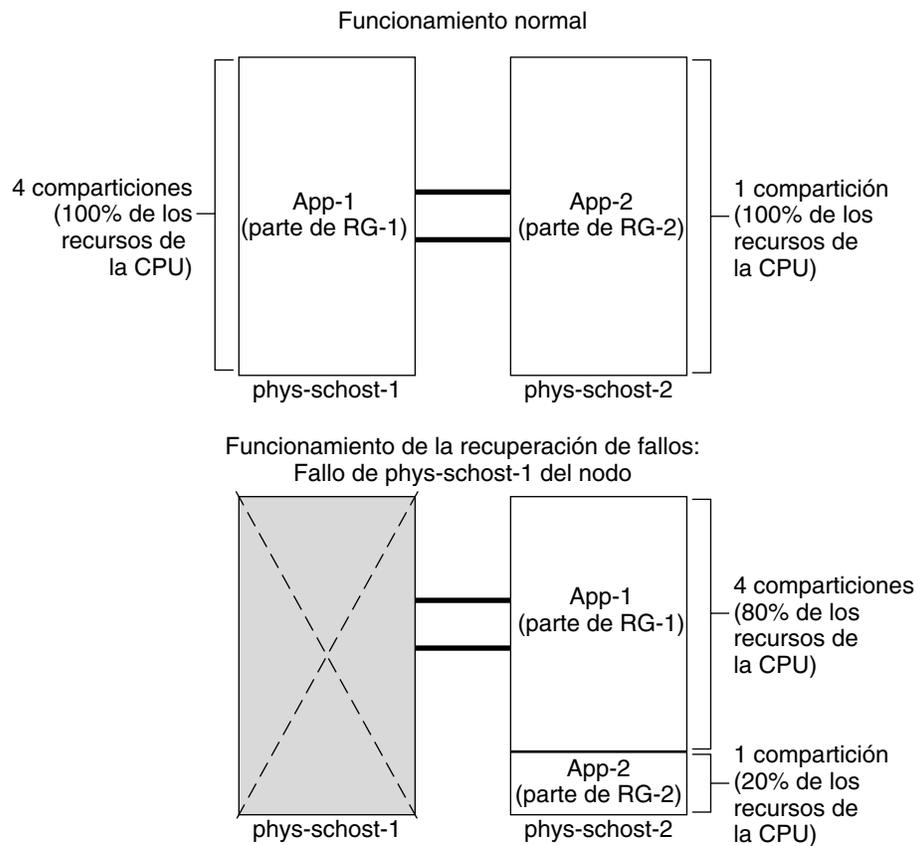
Clúster de dos nodos con dos aplicaciones

Puede configurar dos aplicaciones en un clúster de dos nodos para asegurarse de que cada sistema físico (*phys-schost-1*, *phys-schost-2*) actúe como el maestro predeterminado de una aplicación. Cada sistema físico actúa como el nodo secundario del otro. Todos los proyectos que estén asociados a la aplicación 1 y 2 deben estar representados en los archivos de la base de datos en ambos nodos. Cuando el clúster se está ejecutando normalmente, todas las aplicaciones se ejecutan en su principal predeterminado, donde la función de gestión le asigna todo el tiempo de CPU.

Después de una recuperación de fallos o cambio, las dos aplicaciones se ejecutan en un único nodo en que se les asigna particiones, como se especifica en el archivo de configuración. Por ejemplo, esta entrada del archivo `/etc/project` especifica que la aplicación 1 tiene asignadas 4 particiones y que la aplicación 2 tiene asignada 1 partición.

```
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)
```

El diagrama siguiente ilustra las operaciones normales y de recuperación de fallos de esta configuración. El número de particiones que se asignen no cambia. No obstante, el porcentaje de tiempo de la CPU disponible para cada aplicación puede modificarse. El porcentaje depende del número de particiones que estén asignadas a cada proceso que requiera tiempo de la CPU.



Clúster de dos nodos con tres aplicaciones

En un clúster de dos nodos con tres aplicaciones, puede configurar un sistema físico (*phys-schost-1*) como el maestro predeterminado de una aplicación. El segundo sistema físico (*phys-schost-2*) se puede configurar como el maestro predeterminado para las dos aplicaciones restantes. Consideremos el siguiente archivo de base de datos de proyectos de ejemplo que está en todos los nodos. El archivo de base de datos de proyectos no cambia cuando se produce una recuperación de fallos o un cambio.

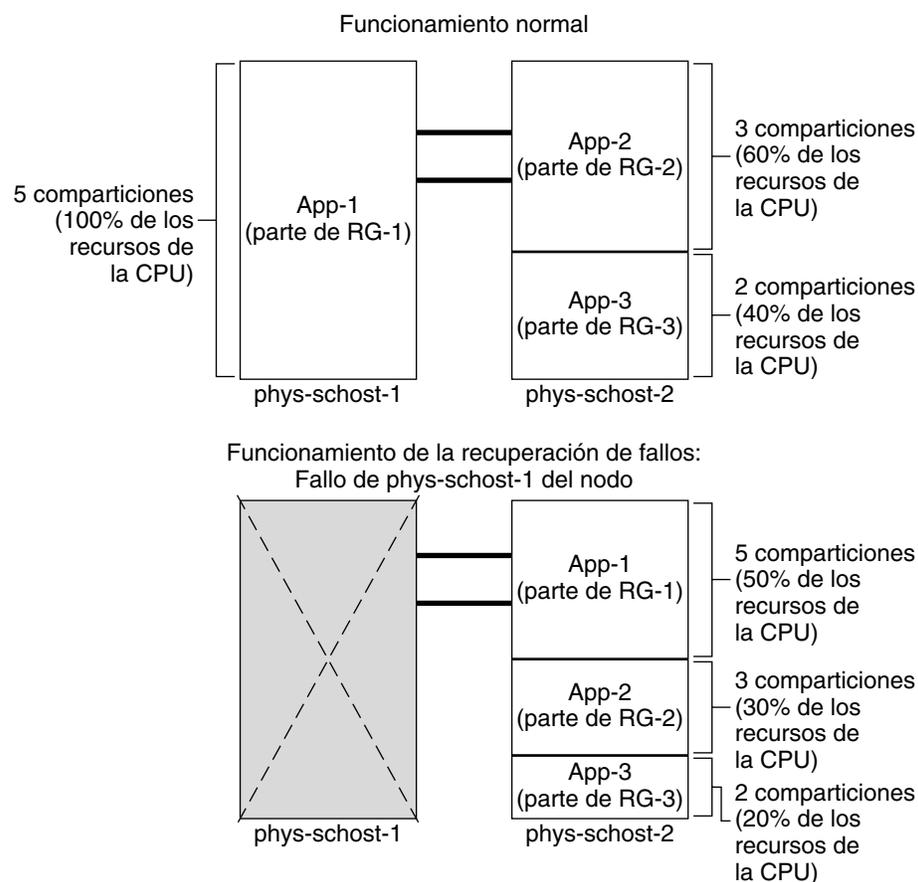
```
Prj_1:103:project for App_1:root::project.cpu-shares=(privileged,5,none)
Prj_2:104:project for App_2:root::project.cpu-shares=(privileged,3,none)
Prj_3:105:project for App_3:root::project.cpu-shares=(privileged,2,none)
```

Cuando el clúster se está ejecutando normalmente, a la aplicación 1 se le asignan 5 particiones en su maestro predeterminado, *phys-schost-1*. Este número es equivalente al 100 por cien del tiempo de CPU porque es la única aplicación que lo demanda en ese nodo. Las aplicaciones 2 y 3 tienen asignadas 3 y 2 particiones respectivamente en el maestro predeterminado, *phys-schost-2*. La aplicación 2 recibiría el 60 por ciento del tiempo de CPU y la aplicación 3 recibiría el 40 por ciento durante el funcionamiento normal.

Si se produce una recuperación de fallos o una conmutación y la aplicación se conmuta con *phys-schost-2*, las particiones de las tres aplicaciones permanecen iguales. Sin embargo, los porcentajes de recursos de CPU se reasignan según el archivo de base de datos de proyectos.

- La aplicación 1, con 5 particiones, recibe el 50 por ciento de CPU.
- La aplicación 2, con tres particiones, recibe el 30 por ciento de CPU.
- La aplicación 3, con 2 particiones, recibe el 20 por ciento de CPU.

El diagrama siguiente ilustra el funcionamiento normal y las operaciones de recuperación de fallos de esta configuración.



Recuperación de fallos sólo de grupos de recursos

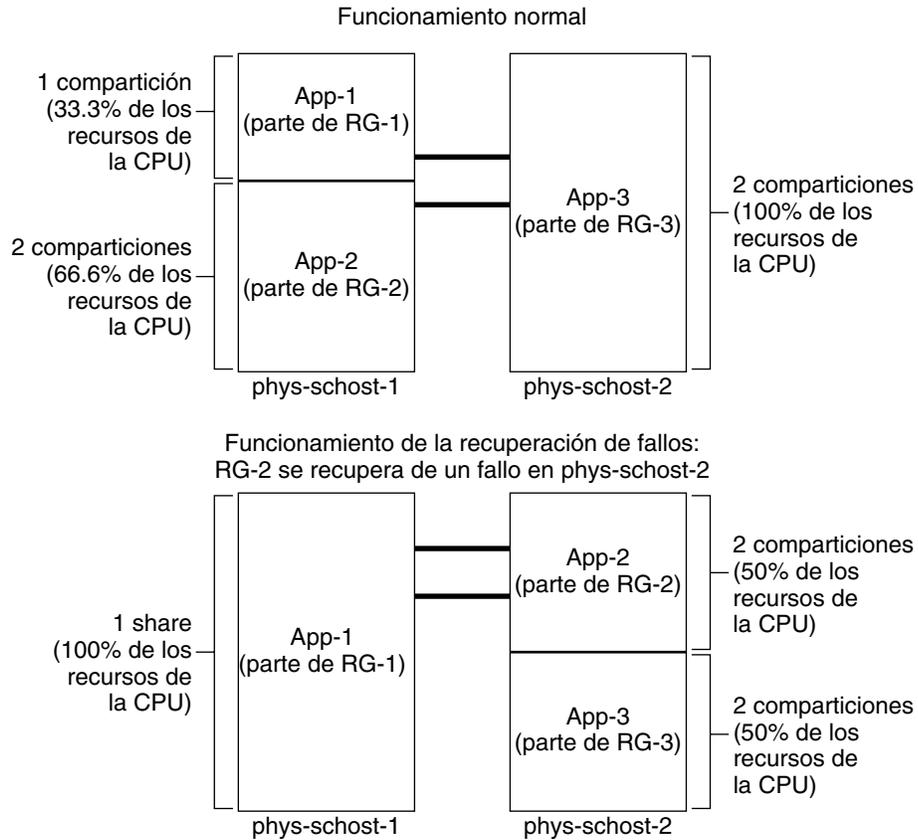
En una configuración en la que varios grupos de recursos tienen el mismo maestro predeterminado, un grupo de recursos (y sus aplicaciones asociadas) pueden entrar en recuperación de fallos o cambiarse a un nodo secundario. Mientras tanto el maestro predeterminado está ejecutándose en el clúster.

Nota – Durante la recuperación de fallos a la aplicación que falla se le asignan recursos, como los que se especifican en el archivo de configuración en el nodo secundario. En este ejemplo, los archivos de base de datos del proyecto de los nodos primario y secundario tienen la misma configuración.

Por ejemplo, este archivo de configuración de ejemplo especifica que a la aplicación 1 se le asigna 1 partición, a la aplicación 2 se le asignan 2 y a la aplicación 3 se le asignan otras 2.

```
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)
```

El siguiente diagrama ilustra el funcionamiento normal y de recuperación de fallos de esta configuración, donde RG-2, que contiene la aplicación 2, falla e intenta recuperarse con *phys-schost-2*. Tenga en cuenta que el número de particiones asignadas no cambia. Sin embargo, el porcentaje de tiempo de la CPU disponible para cada aplicación se puede cambiar, en función del número de particiones que estén asignadas a cada aplicación que requiera tiempo de la CPU.



Adaptadores de red públicos y Ruta múltiple de red de protocolo de Internet (IP)

Los clientes hacen peticiones de datos al clúster a través de la red pública. Cada nodo del clúster está conectado como mínimo a una red pública a través de un par de adaptadores.

El software de ruta múltiple de red de protocolo de Internet (IP) de Solaris en Sun Cluster proporciona el mecanismo básico para supervisar los adaptadores de red públicos y realizar la recuperación de fallos en direcciones IP desde un adaptador a otro cuando se produce un fallo. Todos los nodos del clúster tienen su propia configuración de Ruta múltiple de red de protocolo de Internet (IP), que puede ser distinta de la configuración de otros nodos del clúster.

Los adaptadores de red públicos están organizados en *grupos de ruta IP múltiple* (grupos de ruta múltiple). Cada grupo de ruta múltiple tiene uno o más adaptadores de red pública. Cada adaptador de un grupo de ruta múltiple puede estar activo. Alternativamente, puede configurar interfaces de modo de espera que estén inactivas hasta que se produzca una recuperación de fallos.

El daemon de ruta múltiple `in.mpathd` usa una dirección IP de prueba para detectar los fallos y realizar las reparaciones. Si detecta un fallo en uno de los adaptadores, se produce una recuperación de fallos. Un acceso de red se recupera del fallo desde el adaptador que ha fallado a otro adaptador funcional en el grupo de ruta múltiple. En consecuencia, el daemon mantiene la conectividad de red pública para el nodo. Si configuró una interfaz de modo de espera, el daemon elegirá dicha interfaz. De lo contrario, el daemon elegirá la interfaz con el menor número de direcciones IP. Dado que la recuperación de fallos se produce en el nivel de interfaz del adaptador, las conexiones de un nivel superior, como las TCP, no se ven afectadas, excepto por un breve retraso durante la recuperación de fallos. Cuando la recuperación de fallos de las direcciones IP se completa correctamente, se envían las difusiones ARP. En consecuencia, el daemon mantiene la conectividad con los clientes remotos.

Nota – Debido a la función de recuperación del colapso de TCP, los puntos finales TCP pueden experimentar otros retrasos después de que la recuperación del fallo haya sido correcta. Puede que algunos segmentos se hayan perdido durante la recuperación de fallos, lo que activaría el mecanismo de control de colapsos en TCP.

Los grupos de ruta múltiple ofrecen bloques de construcción para nombres de sistema lógicos y recursos de dirección compartida. También se pueden crear grupos de ruta múltiple independientemente de los nombres de sistema lógicos y los recursos de

dirección compartida para supervisar la conectividad de red pública de los nodos del clúster. El mismo grupo de ruta múltiple de un nodo puede alojar cualquier número de nombres de sistema lógicos o recursos de dirección compartida. Si desea obtener más información sobre sistemas lógicos y recursos de direcciones compartidos, consulte la *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

Nota – El diseño del mecanismo de Ruta múltiple de red de protocolo de Internet (IP) está enfocado a detectar y solucionar fallos del adaptador. Su diseño no está pensado para recuperarse tras el uso del comando `ifconfig(1M)` por parte de un administrador para eliminar una de las direcciones IP lógicas o compartidas. El software Sun Cluster considera las direcciones IP compartidas y lógicas como recursos administrados por RGM. El modo correcto para que un administrador agregue o elimine una dirección IP es usar el comando `scrpadm(1M)` para modificar el grupo de recursos que contiene el recurso.

Para obtener más información acerca de la implementación de Solaris de varias rutas de red IP, consulte la documentación pertinente del sistema operativo Solaris que esté instalado en su clúster.

Versión del sistema operativo	Instrucciones
Sistema operativo Solaris 8	<i>IP Network Multipathing Administration Guide</i>
Sistema operativo Solaris 9	Capítulo 1, "IP Network Multipathing (Overview)" de <i>IP Network Multipathing Administration Guide</i>
Sistema operativo Solaris 10	Parte VI, "IPMP" de <i>System Administration Guide: IP Services</i>

SPARC: Compatibilidad con la reconfiguración dinámica

El soporte de Sun Cluster 3.1 8/05 para la función de software de reconfiguración dinámica (DR) se está desarrollando en fases incrementales. Este apartado describe los conceptos y consideraciones para el soporte de Sun Cluster 3.1 8/05 de la función DR.

Todos los requisitos, las restricciones y los procedimientos que están documentados para la función DR de Solaris se aplican también al soporte de DR de Sun Cluster (excepto por la operación de quiescencia del entorno operativo). En consecuencia, revise la documentación de la función DR de Solaris *antes* de usar dicha función con el software Sun Cluster. Debe prestar especial atención a los problemas que afectan a los dispositivos de E/S que no están en red durante una operación de desacople de DR.

Los manuales *Sun Enterprise 10000 Dynamic Reconfiguration User Guide* y *Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual* (de las colecciones *Solaris 8 on Sun Hardware* o *Solaris 9 on Sun Hardware*) están disponibles para descargarlos de <http://docs.sun.com>.

SPARC: Descripción general de la reconfiguración dinámica

La función DR habilita operaciones tales como la extracción de hardware del sistema en sistemas que estén en ejecución. Los procesos DR están diseñados para asegurar el funcionamiento continuo del sistema sin necesidad de pararlo o interrumpir la disponibilidad del clúster.

DR funciona a nivel de placa. Por lo tanto, la operación DR afecta a todos los componentes de la placa. Cada placa puede contener varios componentes, como CPU, memorias e interfaces periféricas para unidades de disco, unidades de cinta y conexiones en red.

La extracción de una placa que contenga componentes activos podría provocar errores en el sistema. Antes de extraer una placa, el subsistema DR consulta otros subsistemas, como Sun Cluster, para determinar si los componentes de la placa se están utilizando. Si el subsistema DR encuentra que la placa se está usando, la operación de extraer la placa por DR no se lleva a cabo. Por lo tanto, siempre es más seguro llevar a cabo una operación de extracción de placa por DR porque el subsistema DR rechaza operaciones en las placas que contienen componentes activos.

La operación de adición de la placa DR también es siempre segura. El sistema pone en funcionamiento automáticamente las CPU y la memoria de una placa recién añadida. Sin embargo, el administrador del sistema deberá configurar manualmente el clúster para que use de forma activa componentes que estén en la placa recién agregada.

Nota – El subsistema DR tiene varios niveles. Si un nivel inferior informa de un error, el superior también informa del mismo error. No obstante, cuando el nivel inferior informa sobre un error específico, el nivel superior indica que se ha producido un error desconocido. Este error se puede omitir sin problema.

Los apartados siguientes describen consideraciones de DR para los distintos tipos de dispositivos.

SPARC: Consideraciones de clúster DR para los dispositivos CPU

El software Sun Cluster no rechaza una operación de extracción de placa por DR debido a la presencia de los dispositivos CPU.

Cuando una operación de añadir placa con DR tenga éxito, los dispositivos de CPU de la placa añadida se incorporarán automáticamente al funcionamiento del sistema.

SPARC: Consideraciones de clúster DR para la memoria

Para los propósitos de DR, se deben considerar dos tipos de memoria.

- paquete de memoria del núcleo
- paquete de memoria que no es del núcleo

que difieren sólo en su uso. El hardware real es el mismo para ambos tipos. El paquete de memoria del núcleo es la memoria utilizada por el sistema operativo Solaris. El software Sun Cluster no admite las operaciones de extracción de placas que contengan paquetes de memoria del núcleo y rechaza las operaciones de este tipo. Cuando una operación de extracción de placa por DR pertenece a una memoria distinta de la del paquete del núcleo, el software Sun Cluster no rechaza la operación. Cuando una operación de añadir placa con DR que pertenezca a memoria tenga éxito, la memoria de la placa añadida se incorporará automáticamente al funcionamiento del sistema.

SPARC: Consideraciones de clúster DR para unidades de disco y cinta

Sun Cluster rechaza las operaciones DR de extraer-placa en las unidades activas del nodo principal. Las operaciones de extracción de placa por DR se pueden llevar a cabo en unidades inactivas en el nodo primario o en cualquier unidad del nodo secundario. Después de la operación de DR, el acceso a los datos del clúster prosigue de la misma forma.

Nota – Sun Cluster rechaza las operaciones DR que afecten a la disponibilidad de los dispositivos del quórum. Para ver las consideraciones acerca de los dispositivos del quórum y los procedimientos para realizar operaciones de DR en ellas, consulte [“SPARC: Consideraciones de clúster DR para los dispositivos del quórum ”](#) en la página 93.

Consulte “Reconfiguración dinámica con los dispositivos del quórum” de *Sun Cluster: Guía de administración del sistema para el SO Solaris* para obtener instrucciones detalladas acerca de cómo llevar a cabo estas acciones.

SPARC: Consideraciones de clúster DR para los dispositivos del quórum

Si la operación de extracción de placa por DR pertenece a una placa que contenga una interfaz para un dispositivo configurado para el quórum, el software Sun Cluster rechazará la operación. El software Sun Cluster también identifica el dispositivo del quórum que se vería afectado por la operación. Es necesario inhabilitar el dispositivo como del quórum antes de poder efectuar una operación de extracción de placa con DR.

Consulte el Capítulo 5, “Administración del quórum” de *Sun Cluster: Guía de administración del sistema para el SO Solaris* para obtener instrucciones detalladas acerca de la forma de administrar el quórum.

SPARC: Consideraciones de clúster DR para las interfaces interconectadas del clúster

Si la operación de extracción de la placa por DR pertenece a una placa que contiene una interfaz de interconexión del clúster activa, el software Sun Cluster rechaza la operación. El software Sun Cluster también identifica la interfaz que se vería afectada por la operación. Debe usar una herramienta administrativa de Sun Cluster para deshabilitar la interfaz activa para que la operación de DR tenga éxito.



Caution – El software Sun Cluster requiere que cada nodo del clúster tenga como mínimo una ruta operativa para los demás nodos del clúster. No inhabilite una interfaz de interconexión privada en el caso de que represente la última ruta a cualquiera de los nodos del clúster.

Consulte “Administración de las interconexiones del clúster” de *Sun Cluster: Guía de administración del sistema para el SO Solaris* para obtener instrucciones detalladas acerca de cómo realizar estas acciones.

SPARC: Consideraciones de clúster DR para las interfaces de red públicas

Si la operación de extracción de la placa por DR pertenece a una placa que contenga una interfaz de red pública activa, el software Sun Cluster rechazará la operación. El software Sun Cluster también identifica la interfaz que se vería afectada por la operación. Antes de eliminar una placa que tenga una interfaz de red presente, conmute todo el tráfico de dicha interfaz a otra interfaz funcional del grupo con varias rutas usando el comando `if_mpadm(1M)`.



Caution – Si el resto de adaptadores de red fallan mientras se está efectuando una supresión de DR en el adaptador de red inhabilitado, la disponibilidad se verá afectada. El adaptador restante no tiene a quién transferir el control durante la operación de DR.

Consulte “Administración de la red pública” de *Sun Cluster: Guía de administración del sistema para el SO Solaris* para obtener instrucciones detalladas acerca de cómo realizar una operación de extracción por DR en una interfaz de red pública.

Preguntas más frecuentes

Este capítulo incluye respuestas a las preguntas más frecuentes acerca del sistema Sun Cluster. Las preguntas están organizadas por temas.

Preguntas más frecuentes sobre alta disponibilidad

Pregunta: ¿Qué es exactamente un sistema de alta disponibilidad?

Respuesta: En el sistema Sun Cluster, se llama alta disponibilidad (HA) a la capacidad de un clúster para mantener en funcionamiento una aplicación. La aplicación sigue ejecutándose incluso cuando se produce un fallo que, normalmente, dejaría un sistema de servidor fuera de servicio.

Pregunta: ¿Cuál es el proceso por el que el clúster proporciona alta disponibilidad?

Respuesta: A través de un proceso conocido como recuperación de fallos, la estructura del clúster proporciona un entorno de alta disponibilidad. Recuperación de fallos es una serie de pasos que realiza el clúster para migrar recursos de servicios de datos de un nodo fallido a otro operativo dentro del clúster.

Pregunta: ¿Cuál es la diferencia entre un servicio de datos a prueba de fallos y otro escalable?

Respuesta: Hay dos tipos de servicios de datos de alta disponibilidad:

- recuperación tras error
- escalable

Un servicio de datos a prueba de fallos ejecuta una aplicación sólo en un nodo primario del clúster cada vez. Los demás nodos pueden ejecutar otras aplicaciones, pero cada una se ejecuta sólo en un nodo. Si un nodo principal falla, las aplicaciones que se ejecutan en el nodo que ha fallado se recuperan del error y continúan ejecutándose en otro nodo.

Un servicio escalable reparte una aplicación entre varios nodos para crear un servicio lógico único que aprovecha el número de nodos y procesadores de todo el clúster en el que se ejecutan.

Para cada aplicación un nodo aloja la interfaz física para el clúster. Este nodo se denomina interfaz global (GIF). En un clúster pueden existir varios nodos GIF. Cada uno de ellos aloja una o más interfaces lógicas que pueden usar los servicios escalables y que reciben el nombre de *interfaces globales*. Un nodo GIF aloja una interfaz global para todas las solicitudes hacia una aplicación en particular y las despacha a los distintos nodos en los que se esté ejecutando el servidor de la aplicación. Si el nodo GIF falla, la interfaz global se traslada a un nodo superviviente.

Si falla alguno de los nodos en los que se ejecuta la aplicación, ésta continúa ejecutándose en otros nodos con una leve pérdida de rendimiento. Este proceso continúa hasta que el nodo que ha fallado vuelve al clúster.

FAQ sobre los sistemas de archivos

Pregunta: ¿Puedo ejecutar uno o varios de los nodos de un clúster como servidores NFS de alta disponibilidad con otros nodos de clúster como clientes?

Respuesta: No, no haga un montaje de bucle cerrado.

Pregunta: ¿Puedo usar un sistema de archivos en clúster para aplicaciones que no estén bajo el control de Gestor de grupos de recursos?

Respuesta: Sí. Sin embargo, sin el control de RGM, es necesario reiniciar manualmente las aplicaciones después del fallo del nodo en que se están ejecutando.

Pregunta: ¿Deben tener todos los sistemas de archivos en clúster un punto de montaje en el directorio `/global`?

Respuesta: No. Sin embargo, si se sitúan los sistemas de archivos clúster bajo el mismo punto de montaje, como `/global`, se consigue una mejor organización y gestión de estos sistemas de archivos.

Pregunta: ¿Cuáles son las diferencias entre usar el sistema de archivos del clúster y exportar sistemas de archivos NFS?

Respuesta: Hay varias diferencias:

1. El sistema de archivos del clúster admite dispositivos globales. NFS no admite acceso remoto a los dispositivos.

2. El sistema de archivos del clúster dispone de un espacio de nombres global. Sólo es necesaria una orden de montaje. Con NFS, debe montarse el sistema de archivos en cada uno de los nodos.
3. El sistema de archivos del clúster almacena en antememoria los archivos en más casos que NFS. Por ejemplo, el sistema de archivos en clúster almacena en caché los archivos cuando se accede a éstos desde varios nodos para efectuar tareas de lectura, escritura, bloqueo de archivos o E/S asíncronas.
4. El sistema de archivos del clúster está creado para aprovechar futuras interconexiones rápidas de clúster que ofrezcan DMA remotas y funciones de copia cero.
5. Si cambia los atributos de un archivo (usando `chmod (1M)`, por ejemplo) en un sistema de archivos en clúster, el cambio se refleja inmediatamente en todos los nodos. Con un sistema de archivos NFS exportado, este cambio puede requerir mucho más tiempo.

Pregunta: El sistema de archivos `/global/.devices/node@ID nodo` aparece en mis nodos del clúster. ¿Puedo usar este sistema de archivos para almacenar datos que deseo que estén altamente disponibles y sean globales?

Respuesta: Estos sistemas de archivos almacenan los espacios de nombre de dispositivo global. Estos sistemas de archivos no están destinados a un uso general. Aunque son globales, a ellos nunca se accede de manera global: cada nodo accede sólo a su propio espacio de nombre de dispositivo global. Si un nodo está caído, el resto no puede acceder al espacio de nombres del nodo en cuestión. Estos sistemas de archivos no son de alta disponibilidad. No deberían usarse para almacenar datos que hayan de estar globalmente accesibles o altamente disponibles.

FAQ sobre gestión de volúmenes

Pregunta: ¿Necesito duplicar todos los dispositivos de disco?

Respuesta: Para que un dispositivo de disco se considere de alta disponibilidad, debe estar duplicado o usar hardware RAID-5. Todos los servicios de datos deben usar dispositivos de disco de alta disponibilidad o sistemas de archivos del clúster montados en dispositivos de disco de alta disponibilidad. Estas configuraciones pueden tolerar fallos de disco puntuales.

Pregunta: ¿Puedo usar un gestor de volúmenes para los discos locales (disco de arranque) y otro distinto para los discos multisistema?

Respuesta: SPARC: esta configuración se admite si el software Gestor de volúmenes de Solaris administra los discos locales y VERITAS Volume Manager administra los discos multisistema. No se admite ninguna otra combinación.

x86: no, esta configuración no se admite, puesto que Gestor de volúmenes de Solaris sólo se admite en los clústeres basados en la plataforma x86.

FAQ sobre servicios de datos

Pregunta: ¿Qué servicios de datos de Sun Cluster están disponibles?

Respuesta: La lista de servicios de datos compatibles se incluye en “Productos admitidos” de *Notas de la versión de Sun Cluster 3.1 8/05 para SO Solaris*.

Pregunta: ¿Qué versiones de aplicaciones son compatibles con los servicios de datos de Sun Cluster?

Respuesta: La lista de versiones de aplicaciones compatibles se incluye en “Productos admitidos” de *Notas de la versión de Sun Cluster 3.1 8/05 para SO Solaris*.

Pregunta: ¿Puedo escribir mi propio servicio de datos?

Respuesta: Sí. Consulte el Capítulo 11, “Funciones de la API de DSDL” de *Sun Cluster: Guía del desarrollador de los servicios de datos del sistema operativo Solaris* para obtener más información.

Pregunta: Al crear recursos de red, ¿debo especificar direcciones IP numéricas o nombres de sistema?

Respuesta: El método preferido para especificar recursos de red es usar el nombre de sistema de UNIX en lugar de la dirección IP numérica.

Pregunta: Al crear recursos de red, ¿cuál es la diferencia entre usar un nombre de sistema lógico (un recurso LogicalHostname) o una dirección compartida (un recurso SharedAddress)?

Respuesta: Excepto en el caso de Sun Cluster HA para NFS, cada vez que la documentación recomienda el uso de un recurso LogicalHostname en un grupo de recursos en modo Failover, se podrá utilizar un recurso SharedAddress o bien uno LogicalHostname. El uso de un recurso SharedAddress lleva consigo varias sobrecargas adicionales porque el software de red del clúster está configurado para SharedAddress, no para LogicalHostname.

La ventaja de usar un recurso SharedAddress queda de manifiesto cuando se configuran servicios de datos escalables y de recuperación de fallos y se desea que los clientes puedan acceder a ambos servicios utilizando el mismo nombre de sistema. En este caso, los recursos SharedAddress y el recurso de aplicación de recuperación de fallos están incluidos en un grupo de recursos. El recurso de servicio escalable está en un grupo de recursos separado y está configurado para usar el recurso SharedAddress. Los servicios escalables y de recuperación de fallos pueden usar el mismo conjunto de nombres de sistemas y direcciones que está configurado en el recurso SharedAddress .

FAQ sobre las redes públicas

Pregunta: ¿Qué adaptadores de red pública admite el sistema Sun Cluster?

Respuesta: Actualmente, el sistema Sun Cluster es compatible con los adaptadores de red pública Ethernet (10/100BASE-T y 1000BASE-SX Gb). Debido a que en el futuro podrían admitirse interfaces nuevas, compruebe con el representante de ventas de Sun cuál es la información más actual.

Pregunta: ¿Cuál es el papel de la dirección MAC en la recuperación de fallos?

Respuesta: Cuando se produce una recuperación de fallos, se generan nuevos paquetes de protocolo de resolución de direcciones (ARP) y se envían a todo el mundo. Estos paquetes ARP contienen la nueva dirección MAC (del nuevo adaptador físico al que se ha transferido el nodo) y la dirección IP antigua. Cuando otra máquina de la red recibe uno de estos paquetes, borra la correlación MAC-IP antigua de la antememoria ARP y usa la nueva.

Pregunta: ¿Admite el sistema Sun Cluster la configuración `local-mac-address? =true` ?

Respuesta: Sí. En realidad, la Ruta múltiple de red IP requiere que el valor `local-mac-address?` *esté* configurado como `true`.

Puede definir `local-mac-address?` con `eeprom(1M)`, en el indicador OpenBoot PROM `ok` de un clúster basado en SPARC. También puede definir una dirección MAC con la utilidad SCSI que se ejecute opcionalmente después de que arranque la BIOS en un clúster basado en x86.

Pregunta: ¿Cuánto retraso cabe esperar cuando Ruta múltiple de red de protocolo de Internet (IP) realiza una conmutación entre los adaptadores?

Respuesta: El retraso podría ser de varios minutos. El motivo es que cuando se realiza una conmutación en Ruta múltiple de red de protocolo de Internet (IP), la operación envía ARP innecesarios. Sin embargo, no se puede asegurar que el router entre el cliente y el clúster vayan a usar ARP innecesarios. Por ello, hasta que venza la entrada de caché de ARP para esta dirección IP en el router, la entrada podrá usar la dirección MAC antigua.

Pregunta: ¿Con qué velocidad se detecta el fallo de un adaptador de red?

Respuesta: El tiempo de detección de fallos predeterminado es de 10 segundos. Este algoritmo intenta cumplir el tiempo de detección de fallos predeterminado, pero el real depende de la carga de la red.

FAQ sobre pertenencia al clúster

Pregunta: ¿Tienen que tener la misma contraseña de superusuario todos los miembros del clúster?

Respuesta: No es necesario que el superusuario comparta la misma contraseña en todos los miembros del clúster. Sin embargo, esta práctica simplifica mucho la administración del clúster.

Pregunta: ¿Es importante el orden en el que arrancan los nodos?

Respuesta: En la mayoría de los casos no. Sin embargo, el orden de arranque es importante para evitar la amnesia. Por ejemplo, si el nodo dos era el propietario del dispositivo del quórum y el nodo uno estaba caído, y después se desactiva el nodo dos, hay que volver a activar éste antes que aquél. Este orden impide que se active de forma accidental un nodo con información de configuración del clúster desfasada. Consulte "[Acerca del aislamiento de fallos](#)" en la [página 56](#) para obtener detalles acerca de la amnesia.

Pregunta: ¿Es necesario realizar una duplicación de los discos locales en un nodo del clúster?

Respuesta: Sí. Aunque la duplicación no es un requisito, duplicar los discos del nodo del clúster impide que un fallo de disco no duplicado deje inactivo el nodo. El inconveniente de realizar una duplicación de los discos locales de un nodo del clúster es que complica la administración del sistema.

Pregunta: ¿Cuáles son los inconvenientes de realizar copias de respaldo de los miembros del clúster?

Respuesta: En un clúster se pueden usar varios métodos de copia de respaldo. Un método consiste en dejar un nodo para que sirva de respaldo con una unidad de cinta o una biblioteca adjunta. A continuación se debe usar el sistema de archivos del clúster para realizar la copia de seguridad de los datos. No conecte este nodo a los discos compartidos.

Consulte el Capítulo 9, "Copia de seguridad y restauración de un clúster" de *Sun Cluster: Guía de administración del sistema para el SO Solaris* para obtener información adicional acerca de cómo hacer copias de respaldo y restaurar los datos.

Pregunta: ¿Cuándo está un nodo en el suficiente buen estado como para usarlo como secundario?

Respuesta: Solaris 8 y Solaris 9:

Después de un arranque, un nodo está en el suficiente buen estado como para ser secundario cuando éste muestra el indicador de inicio de sesión.

Solaris 10:

Un nodo está en buen estado para ser un nodo secundario si el hito `multi-user-server` se está ejecutando.

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

FAQ sobre almacenamiento en clúster

Pregunta: ¿Qué hace que el almacenamiento multisistema sea de alta disponibilidad?

Respuesta: El almacenamiento multisistema es de alta disponibilidad porque puede sobrevivir a la pérdida de un único disco gracias a la duplicación (o gracias a los controladores RAID-5 basados en hardware). Como los dispositivos de almacenamiento multisistema tienen más de una conexión de sistema, también pueden resistir la pérdida de un nodo al cual estén conectados. Además, las rutas redundantes desde cada nodo al almacenamiento conectado ofrecen tolerancia para el fallo de un adaptador del bus del sistema, de un cable o del controlador de disco.

FAQ sobre interconexiones del clúster

Pregunta: ¿Qué interconexiones del clúster admite el sistema Sun Cluster?

Respuesta: Actualmente, el sistema Sun Cluster admite las siguientes interconexiones del clúster:

- Ethernet (100BASE-T Fast Ethernet y 1000BASE-SX Gb) en clústeres basados en SPARC y en x86
- Infiniband en clústeres basados en SPARC y en x86
- SCI sólo en clústeres basados en SPARC

Pregunta: ¿Cuál es la diferencia entre un “cable” y una “ruta” de transporte?

Respuesta: Los cables de transporte del clúster se configuran usando adaptadores y conmutadores de transporte. Los cables unen adaptadores y conmutadores según cada componente. El gestor de la topología del clúster usa los cables disponibles para crear las rutas de transporte entre los extremos de los nodos. Un cable no se correlaciona directamente con una ruta de transporte.

Los cables son “habilitados” e “inhabilitados” por el administrador estáticamente. Los cables tienen un “estado” (habilitado o inhabilitado), pero no un “estatus”. Si se inhabilita un cable, es como si estuviera sin configurar. Los cables que están inhabilitados no pueden usarse como rutas de transporte. Estos cables no están probados y, en consecuencia, se desconoce su estado. Puede obtener el estado de un cable usando el comando `scconf -p`.

El gestor de topología del clúster establece dinámicamente las rutas de transporte. El “estatus” de una ruta de transporte viene determinado por el gestor de topologías. Una ruta puede tener un estado de “en línea” o “fuera de línea.” Puede obtener el estatus de una ruta de transporte usando el comando `scstat (1M)`.

Considere el ejemplo siguiente de clúster de dos nodos con cuatro cables.

```
node1:adapter0      to switch1, port0
node1:adapter1      to switch2, port0
node2:adapter0      to switch1, port1
node2:adapter1      to switch2, port1
```

A partir de estos cuatro cables, se pueden formar dos rutas de transporte posibles.

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

FAQ sobre sistemas cliente

Pregunta: ¿Es necesario considerar alguna necesidad o restricción de cliente para usarla con un clúster?

Respuesta: Los sistemas cliente se conectan a un clúster del mismo modo que lo harían a cualquier otro servidor. En algunos casos, dependiendo de la aplicación del servicio de datos, es posible que necesite instalar software de lado del cliente o llevar a cabo otros cambios de configuración para que el cliente pueda conectarse a la aplicación del servicio de datos. Consulte el Capítulo 1, “Planning for Sun Cluster Data Services” de *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* para obtener más información acerca de los requisitos de configuración del cliente.

FAQ sobre la consola administrativa

Pregunta: ¿El sistema Sun Cluster requiere una consola administrativa?

Respuesta: Sí.

Pregunta: ¿Tiene que tratarse de una consola de administración exclusiva para el clúster? ¿O puede usarse para otras tareas?

Respuesta: El sistema Sun Cluster no requiere una consola administrativa dedicada, pero si usa una, obtendrá los siguientes beneficios:

- Permite la gestión centralizada del clúster ya que agrupa herramientas de consola y gestión en la misma máquina
- Ofrece al proveedor de servicio de hardware una resolución de problemas potencialmente mas rápida

Pregunta: ¿Debe colocarse la consola administrativa “cerca” del clúster (por ejemplo, en la misma habitación)?

Respuesta: Compruébelo con su proveedor de servicio de hardware. El proveedor puede requerir que la consola esté ubicada cerca del clúster. No existe ninguna razón técnica por la que la consola deba estar situada en la misma habitación.

Pregunta: ¿Puede atender una consola administrativa a más de un clúster, si se cumplen primero otros requisitos de distancia?

Respuesta: Sí. Se pueden controlar varios clústeres desde una misma consola de administración. También puede compartirse un concentrador de terminal simple entre varios clústeres.

FAQ sobre el concentrador de terminal y el procesador de servicio del sistema

Pregunta: ¿Requiere el sistema Sun Cluster un concentrador de terminal?

Respuesta: Ninguna versión de software a partir de la Sun Cluster 3.0 requiere un concentrador de terminal para ejecutarse. En cambio, los productos Sun Cluster 2.2 sí requieren un concentrador de terminal para aislamiento de fallos. Los productos posteriores no lo necesitan.

Pregunta: La mayoría de los servidores Sun Cluster utilizan un concentrador de terminal, pero el servidor Sun Enterprise E1000 no lo usa. ¿Por qué no?

Respuesta: El concentrador de terminal es en la práctica un conversor serie para Ethernet para la mayoría de los servidores. El puerto de la consola del concentrador de terminal es un puerto serie. El servidor Sun Enterprise E1000 no tiene una consola serie. El procesador de servicio del sistema (SSP) es la consola, a través de Ethernet o del puerto `jtag`. Para el servidor Sun Enterprise E1000 se utiliza siempre SSP para las consolas.

Pregunta: ¿Cuáles son las ventajas de usar un concentrador de terminal?

Respuesta: El uso de un concentrador de terminal proporciona acceso de consola a cada nodo desde cualquier estación de trabajo de la red. Este acceso se proporciona incluso cuando el nodo está en OpenBoot PROM (OBP) en un nodo basado en SPARC o un subsistema de arranque en un nodo basado en x86.

Pregunta: Si uso un concentrador de terminal que no sea compatible con Sun, ¿qué necesito saber para saber si sirve el que quiero usar?

Respuesta: La principal diferencia entre el concentrador de terminal que es compatible con Sun y otros dispositivos de consola es que el concentrador de terminal de Sun tiene un firmware especial. Este firmware impide que el concentrador de terminal envíe una parada a la consola cuando arranca. Si dispone de un dispositivo de consola que puede enviar una parada, o una señal que la consola pueda interpretar como una parada, dicha parada desconectará el nodo.

Pregunta: ¿Puedo liberar un puerto bloqueado en el concentrador de terminal compatible con Sun sin reiniciarlo?

Respuesta: Sí. Anote el número de puerto que necesita restaurarse y escriba las órdenes siguientes:

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

Consulte los siguientes manuales para obtener información acerca de cómo configurar y administrar el concentrador de terminal compatible con Sun.

- “Información general sobre la administración de Sun Cluster” de *Sun Cluster: Guía de administración del sistema para el SO Solaris*
- Capítulo 2, “Installing and Configuring the Terminal Concentrator” de *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*

Pregunta: ¿Qué ocurre si falla el propio concentrador de terminal? ¿Es necesario tener otro de recambio?

Respuesta: No. No se pierde ninguna disponibilidad del clúster si el concentrador de terminal falla. Se pierde la posibilidad de conectarse a las consolas del nodo hasta que el concentrador vuelva a estar operativo.

Pregunta: Si se usa un concentrador de terminal, ¿qué ocurre con la seguridad?

Respuesta: Normalmente, el concentrador de terminal está adjunto a una pequeña red que usan los administradores de sistema, no a una red que se utilice para el acceso de otros clientes. La seguridad se puede controlar limitando el acceso a esa red en particular.

Pregunta: SPARC: ¿Cómo se utiliza la reconfiguración dinámica con una cinta o unidad de disco?

Respuesta: Siga estos pasos:

- Determine si el disco o la unidad de cinta forma parte de un grupo de dispositivos activo. Si la unidad no forma parte de un grupo de dispositivos activos, puede llevar a cabo la operación de extracción DR en él.

- Si la operación de extracción de placa DR pudiera afectar a un disco o unidad de cinta activos, el sistema rechazará la operación e identificará las unidades que se verían afectadas por la operación. Si la unidad forma parte de un grupo de dispositivos activos, acceda a [“SPARC: Consideraciones de clúster DR para unidades de disco y cinta ” en la página 92.](#)
- Determine si la unidad es un componente del nodo primario o del secundario. Si la unidad es un componente del nodo secundario, puede llevar a cabo la operación de extracción DR en el mismo.
- Si la unidad es un componente del nodo primario, es necesario intercambiar los nodos primario y secundario antes de realizar la operación de extracción DR en el dispositivo.



Caution – Si el nodo principal falla durante una operación de DR en un nodo secundario, la disponibilidad del clúster queda afectada. El nodo primario no tiene dónde transferirse hasta que se proporcione un nodo secundario nuevo.

Índice

A

acceso simultáneo, 22
adaptadores, *Ver* red, adaptadores
administración, clúster, 37-94
agentes, *Ver* servicios de datos
almacenamiento, 24-25
 FAQ, 101
 reconfiguración dinámica, 92
 SCSI, 25-26
alta disponibilidad
 estructura, 39-41
 FAQ, 95-96
altamente disponibles, servicios de datos, 40
amnesia, 55
API, 73-74, 77
API DSDL, 77
aplicación, *Ver* servicios de datos
aplicaciones con un papel fundamental, 63
argumento
 scha_privatelink_hostname_node, 75
asignación, espacio de nombres, 48
atributos, *Ver* propiedades
aviso grave, 41, 58

B

bloqueo de archivo, 48

C

cable, transporte, 101-102

CCP, 29
CCR, 41
cierre, 41
clúster
 administración, 37-94
 componentes de software, 23-24
 configuración, 41, 79-89
 contraseña, 100-101
 copia de respaldo, 100-101
 desarrollo de aplicaciones, 37-94
 descripción, 14
 eliminación de placa, 92
 FAQ sobre almacenamiento en clúster, 101
 hardware, 15-16, 21-30
 hora, 38-39
 interconexión, 23, 27-28
 adaptadores, 27
 cables, 27
 compatible, 101-102
 FAQ, 101-102
 interfaces, 27
 reconfiguración dinámica, 93
 servicios de datos, 75-76
 uniones, 27
 interfaz de red pública, 65
 lista de tareas, 18-19
 medios, 26-27
 miembros, 22, 40
 FAQ, 100-101
 reconfiguración, 40
 nodos, 22-23
 objetivos, 14
 orden de arranque, 100-101

- clúster (Continuación)
 - perspectiva de los desarrolladores de aplicaciones, 17-18
 - perspectiva del administrador de sistemas, 16-17
 - red pública, 28
 - servicio, 15-16
 - servicios de datos, 65-72
 - sistema de archivos, 48-51, 96-97
 - FAQ
 - Ver también* sistema de archivos
 - HASStoragePlus, 50
 - uso, 49-50
 - topologías, 30-34, 34-35
 - ventajas, 14
- Cluster Configuration Repository, 41
- Cluster Membership Monitor, 40
- CMM, 40
 - mecanismo de recuperación rápida, 40
 - Ver también* recuperación rápida
- comando `scha_cluster_get`, 75
- componentes de software, 23-24
- comunicación de las aplicaciones, 75-76
- concentrador de terminal, FAQ, 103-105
- configuración
 - base de datos paralela, 22
 - cliente-servidor, 65
 - repositorio, 41
 - servicios de datos, 79-89
- configuración cliente-servidor, 65
- configuraciones, quórum, 59
- configuraciones de bases de datos paralelas, 22
- configurar, límites de memoria virtual, 83
- conflicto de reserva, 57-58
- consola
 - acceso, 28-29
 - administrativa, 29-30
 - FAQ, 102-103
 - administrativo, 28-29
 - procesador de servicio del sistema, 28-29
- consola administrativa, 29-30
 - FAQ, 102-103
- contraseña, root, 100-101
- contraseña root, 100-101
- controlador, ID del controlador, 42-43
- controlador `clprivnet`, 75

D

- `daemon in.mpathd`, 89
- datos, almacenamiento, 96-97
- desarrollador, aplicaciones en clúster, 17-18
- desarrollo de aplicaciones, 37-94
 - `/dev/global/` espacio de nombres, 47-48
- DID, 42-43
- dirección compartida, 65
 - frente a nombre de sistema lógico, 98
 - nodo de interfaz global, 66
 - servicios de datos escalables, 68-70
- dirección IP, 98
- dirección MAC, 99
- dirección por nodo, 75-76
- disco de arranque, *Ver* discos, locales
- discos
 - aislamiento de fallos, 56-57
 - dispositivos globales, 42-43, 47-48
 - dispositivos SCSI, 25-26
 - grupos de dispositivos, 43-47
 - multipuerto, 45-47
 - propiedad principal, 45-47
 - recuperación de fallos, 44-45
 - local
 - gestión de volúmenes, 97
 - locales, 26, 42-43, 47-48
 - duplicación, 100-101
 - multisistema, 42-43, 43-47, 47-48
 - reconfiguración dinámica, 92
- discos locales, 26
- dispositivo
 - global, 42-43
 - ID, 42-43
- dispositivo multisistema, 24-25
- dispositivos
 - multisistema, 24-25
 - quórum, 54-64
- Distribución de aplicaciones, 60
- DR, *Ver* reconfiguración dinámica

E

- E10000, *Ver* Sun Enterprise E10000
- eliminación de placa, reconfiguración
 - dinámica, 92
- encierro, 41
- equilibrio de cargas, 70-72

espacio de nombres, 48
espacio de nombres local, 48
espacios de nombres, 47-48
esquizofrenia, 55, 56-57
estructura, alta disponibilidad, 39-41

F

fallo, 56-57
 aislamiento, 56-57
 detección, 39
 encierro, 41
 recuperación, 39
 retroceso, 72
FAQ, 95-105
 almacenamiento en clúster, 101
 alta disponibilidad, 95-96
 concentrador de terminal, 103-105
 consola administrativa, 102-103
 gestión de volúmenes, 97
 interconexiones del clúster, 101-102
 miembros del clúster, 100-101
 procesador de servicio del sistema, 103-105
 red pública, 99
 servicios de datos, 98
 sistemas cliente, 102
 sistemas de archivos, 96-97
formación, 11

G

gestión de recursos, 79-89
gestión de volúmenes
 discos locales, 97
 discos multisistema, 97
 dispositivos multisistema, 25
 espacio de nombres, 47
 FAQ, 97
 RAID-5, 97
 Solaris Volume Manager, 97
 VERITAS Volume Manager, 97
gestor de grupos de recursos, *Ver* RGM
global
 dispositivo, 42-43, 43-47
 discos locales, 26
 montaje, 48-51

global (Continuación)
 espacio de nombres, 42, 47-48
 discos locales, 26
 interfaz, 66
 servicios escalables, 69
/global punto de montaje, 48-51, 96-97
grupo de dispositivos, 43-47
 cambio de propiedades, 45-47
grupos
 dispositivo de disco
 Ver discos, grupos de dispositivos
grupos de dispositivos de disco
 multipuerto, 45-47
grupos de recursos, 76-79
 configuraciones, 77-78
 escalable, 68-70
 estados, 77-78
 propiedades, 79
 recuperación de fallos, 68

H

HA, *Ver* alta disponibilidad
hardware, 15-16, 21-30, 90-94
 Ver también almacenamiento
 Ver también discos
 componentes de interconexiones del
 clúster, 27
 reconfiguración dinámica, 90-94
HASToragePlus, 50, 76-79
hora, entre nodos, 38-39
hora de la CPU, 79-89

I

ID
 dispositivo, 42-43
 nodo, 47
iniciador múltiple SCSI, 25-26
interfaces
 Ver red, interfaces
 administrativas, 38
interfaces administrativas, 38
ioctl, 57-58
IPMP, *Ver* Ruta múltiple de red IP

L

local_mac_address, 99
LogicalHostname, *Ver* nombre de sistema lógico

M

medios, extraíbles, 26-27
medios extraíbles, 26-27
memoria, 92
miembro, *Ver* clúster, miembros
modelo de servidor en clúster, 65
modelo de servidor único, 65
modelos de servidor, 65
montaje
 con syncdir, 50-51
 dispositivos globales, 48-51
 /global, 96-97
 sistemas de archivos, 48-51

N

Network Time Protocol, 38-39
NFS, 50-51
nodo de interfaz global, 66
nodo de respaldo, 100-101
nodo primario, 66
nodo secundario, 66
nodos, 22-23
 copia de respaldo, 100-101
 IDnodo, 47
 interfaz global, 66
 orden de arranque, 100-101
 primario, 45-47, 66
 secundario, 45-47, 66
nombre de sistema, 65
nombre de sistema lógico, 65
 frente a dirección compartida, 98
 servicios de datos de recuperación de fallos, 68
NTP, 38-39
núcleo, memoria, 92

O

opción de montaje syncdir, 50-51

Oracle Parallel Server, *Ver* Oracle Real Application Clusters
Oracle Real Application Clusters, 73
orden de arranque, 100-101

P

panel de control del clúster, 29
parámetro auto-boot?, 41
PGR, *Ver* Reserva de grupo persistente
Preguntas más frecuentes, *Ver* FAQ
procesador de servicio del sistema, 28-29, 29
 FAQ, 103-105
propiedad numsecondaries, 45
propiedad preferenced, 45
propiedad principal, grupos de dispositivos de disco, 45-47
propiedad Resource_project_name, 81-82
propiedad RG_project_name, 81-82
propiedad scsi-initiator-id, 26
propiedades
 cambio, 45-47
 grupos de recursos, 79
 recursos, 79
 Resource_project_name, 81-82
 RG_project_name, 81-82
proyectos, 79-89
proyectos de Solaris, 79-89

Q

quórum, 54-64
 configuraciones, 58, 59
 configuraciones atípicas, 63
 configuraciones erróneas, 63-64
Quórum, configuraciones recomendadas, 60-63
quórum
 dispositivo, reconfiguración dinámica, 93
 dispositivos, 54-64
 mejores prácticas, 59-60
 recuento de votos, 55-56
 requisitos, 59

R

reconfiguración dinámica, 90-94

- reconfiguración dinámica (Continuación)
 - descripción, 91
 - discos, 92
 - dispositivos de CPU, 91-92
 - interconexión del clúster, 93
 - memoria, 92
 - quórum devices, 93
 - red pública, 93-94
 - unidades de cinta, 92
- recuperación
 - configuración de retroceso, 72
 - detección de fallos, 39
- recuperación de fallos
 - casos, Solaris Resource Manager, 83-89
 - grupos de dispositivos de disco, 44-45
 - servicios de datos, 68
- recuperación rápida, 41, 57-58
- recursos, 76-79
 - configuraciones, 77-78
 - estados, 77-78
 - propiedades, 79
- red
 - adaptadores, 28, 89-90
 - dirección compartida, 65
 - equilibrio de cargas, 70-72
 - interfaces, 28, 89-90
 - nombre de sistema lógico, 65
 - privada, 23
 - pública, 28
 - FAQ, 99
 - interfaces, 99
 - reconfiguración dinámica, 93-94
 - Ruta múltiple de red IP, 89-90
- recursos, 65, 76-79
- red privada, 23
- red pública, *Ver* red, pública
- Reserva de grupo persistente, 57-58
- retroceso, 72
- RGM, 67, 76-79, 79-89
- RMAPI, 77
- ruta, transporte, 101-102
- ruta múltiple, 89-90
- Ruta múltiple de red IP, 89-90
 - tiempo de recuperación de fallos, 99

S

- SCSI
 - aislamiento de fallos, 56-57
 - conflicto de reserva, 57-58
 - iniciador múltiple, 25-26
 - Reserva de grupo persistente, 57-58
- servicio adosado, 70
- servicio de datos
 - métodos, 67-68
 - supervisor de fallos, 72
- servicio puro, 70
- servicios de datos, 65-72
 - altamente disponibles, 40
 - API, 73-74
 - API de biblioteca, 74
 - compatibles, 98
 - configuración, 79-89
 - desarrollo, 73-74
 - escalable, 68-70
 - FAQ, 98
 - grupos de recursos, 76-79
 - inteconexión del clúster, 75-76
 - recuperación de fallos, 68
 - recursos, 76-79
 - tipos de recursos, 76-79
- servicios de datos escalables, 68-70
- SharedAddress, *Ver* dirección compartida
- sistema de archivos, sistema de archivos en clúster, 96-97
- sistema de archivos
 - almacenamiento de datos, 96-97
 - alta disponibilidad, 96-97
 - clúster, 48-51, 96-97
 - FAQ, 96-97
 - global, 96-97
 - local, 50
 - montaje, 48-51, 96-97
 - NFS, 50-51, 96-97
 - syncdir, 50-51
 - UFS, 50-51
 - VxFS, 50-51
- sistema de archivos local, 50
- sistemas cliente, 28
 - FAQ, 102
 - restricciones, 102
- sistemas de archivos, uso, 49-50
- Solaris Resource Manager, 79-89
 - casos de recuperación de fallos, 83-89

- Solaris Resource Manager (Continuación)
 - configurar límites de memoria virtual, 83
 - requisitos de configuración, 81-82
- Solaris Volume Manager, dispositivos multisistema, 25
- SSP, *Ver* procesador de servicio del sistema
- Sun Cluster, *Ver* clúster
- Sun Enterprise E10000, 103-105
 - consola administrativa, 29
- Sun Management Center (SunMC), 38
- SunPlex, *Ver* clúster
- SunPlex Manager, 38
- supervisión de la ruta de discos, 51-54
- supervisor de fallos, 72

T

- tipos de recursos, 50, 76-79
- topología de par en clúster, 30-31, 35
- topología de par+N, 31-32
- topología N+1 (estrella), 32-33
- topología N*N (escalable), 33-34
- topologías, 30-34, 34-35
 - N+1 (estrella), 32-33
 - N*N (escalable), 33-34
 - par en clúster, 30-31, 35
 - par+N, 31-32

U

- UFS, 50-51
- unidad de CD-ROM, 26-27
- unidad de cinta, 26-27

V

- VERITAS Volume Manager, dispositivos multisistema, 25
- VxFS, 50-51