



Sun Cluster 3.0 12/01 개념

Sun Microsystems, Inc.
901 San Antonio Road
Palo Alto, CA 94303-4900
U.S.A. 650-960-1300

부품번호: 816-3338-10
2002년 1월, Revision A

Copyright 2001 Sun Microsystems, Inc. 901 San Antonio Road, Palo Alto, California 94303-4900 U.S.A. 모든 권리는 저작권자의 소유입니다. Sun Microsystems, Inc.는 이 문서에서 설명하는 제품에 구현된 기술에 대한 지적 재산권을 보유하고 있습니다. 특히, 이 지적 재산권에는 <http://www.sun.com/patents> 페이지에 있는 미국 특허권과 미국 및 다른 국가에서 등록하거나 출원한 특허권이 제한 없이 한 건 이상 포함될 수 있습니다.

이 문서 및 이 문서와 관련된 제품은 사용, 복사, 배포 및 역컴파일을 제한하는 사용권에 따라 배포됩니다. Sun과 권리 양도 사용자(있는 경우)의 사전 서면 승인이 없으면 이 제품이나 문서의 일부를 어떠한 수단이나 어떠한 형태로도 전제할 수 없습니다. 글꼴 기술을 포함한 협력 업체 소프트웨어에 대한 저작권 및 사용권은 Sun 공급업체에 있습니다.

제품 중에는 캘리포니아 대학에 사용권이 있는 Berkeley BSD 시스템에서 파생된 부분이 포함되어 있을 수 있습니다. UNIX는 미국 및 다른 나라에서 X/Open Company, Ltd를 통해 독점적으로 사용권이 부여되는 등록 상표입니다.

Sun, Sun Microsystems, Sun 로고, Java, Netra, Solaris, Sun StorEdge, iPlanet, Apache, Sun Cluster, Answerbook2, docs.sun.com, Solstice DiskSuite, Sun Enterprise, Sun Enterprise SyMON, Solaris JumpStart, JumpStart, Sun Management Center, OpenBoot, Sun Fire, SunPlex, SunSolve, SunSwift, 정확한 Java 로고, AnswerBook 로고, Netra 로고, Solaris 로고 및 iPlanet 로고는 미국 및 다른 나라에서 사용하는 Sun Microsystems, Inc.의 상표 또는 등록 상표입니다. 모든 SPARC 등록 상표는 미국 및 다른 나라에서 사용하는 SPARC International, Inc.의 등록 상표이므로 사용권을 부여 받아 사용해야 합니다. SPARC 상표가 있는 제품은 Sun Microsystems, Inc.에서 개발한 아키텍처를 기반으로 합니다.

ORACLE®은 Oracle Corporation의 등록 상표입니다. Netscape™는 미국 및 다른 나라에서 사용하는 Netscape Communications Corporation의 상표 또는 등록 상표입니다. Adobe® 로고는 Adobe Systems, Incorporated의 등록 상표입니다.

연방 정부 취득: 상용 소프트웨어—정부 사용자는 표준 사용권 조항 및 규정을 준수해야 합니다.

이 제품에는 Apache Software Foundation(<http://www.apache.org/>)에서 개발한 소프트웨어가 포함되어 있습니다.

이 문서에서는 본문의 내용을 “의미 그대로” 제공하며, 상품성, 특정 목적에 대한 적합성 또는 비위반과 같이 함축된 보증을 포함하여 모든 명시적 또는 함축적 조건, 표현 및 보증은 이 문서에 포함되어 있지 않습니다. 이 책임 부인 항목은 법률을 위반하지 않는 범위 내에서만 적용됩니다.

Copyright 2001 Sun Microsystems, Inc., 901 San Antonio Road, Palo Alto, CA 94303-4900 Etats-Unis. Tous droits réservés.

Sun Microsystems, Inc. a les droits de propriété intellectuels relatants à la technologie incorporée dans le produit qui est décrit dans ce document. En particulier, et sans la limitation, ces droits de propriété intellectuels peuvent inclure un ou plus des brevets américains énumérés à <http://www.sun.com/patents> et un ou les brevets plus supplémentaires ou les applications de brevet en attente dans les Etats - Unis et dans les autres pays.

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 des systèmes 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, Java, Netra, Solaris, Sun StorEdge, iPlanet, Apache, Sun Cluster, Answerbook2, docs.sun.com, Solstice DiskSuite, Sun Enterprise, Sun Enterprise SyMON, Solaris JumpStart, JumpStart, Sun Management Center, OpenBoot, Sun Fire, SunPlex, SunSolve, SunSwift, le logo 100% Pure Java, le logo AnswerBook, le logo Netra, le logo Solaris et le logo iPlanet 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® est une marque déposée registre de Oracle Corporation. Netscape™ est une marque de Netscape Communications Corporation aux Etats-Unis et dans d'autres pays. Le logo Adobe® est une marque déposée de Adobe Systems, Incorporated.

Ce produit inclut le logiciel développé par la base de Apache Software Foundation (<http://www.apache.org/>).

LA DOCUMENTATION EST FOURNIE "EN L'ETAT" ET TOUTES AUTRES CONDITIONS, DECLARATIONS ET GARANTIES EXPRESSES OU TACITES SONT FORMELLEMENT EXCLUES, DANS LA MESURE AUTORISEE PAR LA LOI APPLICABLE, Y COMPRIS NOTAMMENT TOUTE GARANTIE IMPLICITE RELATIVE A LA QUALITE MARCHANDE, A L'APTITUDE A UNE UTILISATION PARTICULIERE OU A L'ABSENCE DE CONTREFAÇON.



목차

	머리말	7
1.	소개 및 개요	11
	SunPlex 시스템 소개	11
	높은 가용성과 고장 허용	12
	SunPlex 시스템의 페일오버 및 확장성	12
	SunPlex 시스템에 대한 세 가지 관점	13
	하드웨어 설치 및 서비스 관점	13
	시스템 관리자 관점	14
	응용프로그램 프로그래머 관점	16
	SunPlex 시스템 작업	17
2.	주요 개념 - 하드웨어 서비스 제공업체	19
	SunPlex 시스템 하드웨어 구성 요소	19
	클러스터 노드	21
	멀티호스트 디스크	22
	로컬 디스크	24
	이동식 미디어	24
	클러스터 상호 연결	24
	공용 네트워크 인터페이스	25

	클라이언트 시스템	25
	콘솔 액세스 장치	26
	관리 콘솔	26
	Sun Cluster 토폴로지의 예	27
	클러스터 쌍 토폴로지	27
	Pair+M 토폴로지	28
	N+1(방사형) 토폴로지	29
3.	주요 개념 - 관리 및 응용프로그램 개발	31
	클러스터 관리 및 응용프로그램 개발	31
	관리 인터페이스	33
	클러스터 시간	33
	가용성이 높은 프레임워크	33
	글로벌 장치	36
	디스크 장치 그룹	37
	글로벌 이름 공간	39
	클러스터 파일 시스템	41
	정족수 및 정족수 장치	43
	블룸 관리자	47
	데이터 서비스	48
	새 데이터 서비스 개발	56
	데이터 서비스 트래픽을 위한 클러스터 상호 연결 사용	57
	자원, 자원 그룹 및 자원 유형	58
	PNM(Public Network Management) 및 NAFO(Network Adapter Failover)	61
	동적 재구성 지원	62
4.	자주 문의하는 질문(FAQ)	67
	높은 가용성 FAQ	67
	파일 시스템 FAQ	68

블룸 관리 FAQ	69
데이터 서비스 FAQ	69
공용 네트워크 FAQ	70
클러스터 구성원 FAQ	71
클러스터 기억 장치 FAQ	72
클러스터 상호 연결 FAQ	72
클라이언트 시스템 FAQ	73
관리 콘솔 FAQ	73
터미널 콘센트레이터 및 시스템 서비스 프로세서 FAQ	74
용어해설	77

머리말

Sun™ Cluster 3.0 12/01 개념에는 *SunPlex™* 시스템에 대한 개념 및 참조 정보가 있습니다. *SunPlex* 시스템에는 *Sun*의 클러스터 솔루션을 구성하는 모든 하드웨어 및 소프트웨어 구성 요소가 포함되어 있습니다.

이 문서는 *Sun Cluster* 소프트웨어에 대한 교육을 받은 전문 시스템 관리자를 위한 내용입니다. 이 문서는 계획이나 관측용 안내서가 아닙니다. 이 문서를 읽을 때는 이미 시스템 요구 사항을 결정하고 필요한 장비와 소프트웨어를 구입한 상태입니다.

이 책에서 설명하는 개념을 이해하려면 *Solaris™* 운영 환경에 대한 지식이 있어야 하고, *SunPlex* 시스템에서 사용하는 볼륨 관리자 소프트웨어에 익숙해야 합니다.

활자체 규약

글자체 또는 기호	의미	예
AaBbCc123	명령, 파일 및 디렉토리의 이름, 컴퓨터 화면상의 출력	.login 파일을 편집하십시오. 모든 파일을 표시하려면 <code>ls -a</code> 명령을 사용하십시오. % You have mail.
AaBbCc123	컴퓨터 화면상의 출력에 대하여 입력할 내용	% su Password:

글자체 또는 기호	의미	예
<i>AaBbCc123</i>	책 제목, 새 단어 및 용어, 강조 표시할 단어	<i>User's Guide</i> 의 제 6장을 읽으십시오. 이것을 class 옵션이라고 합니다. 이 작업을 수행하려면 반드시 슈퍼유저이어야 합니다.
	실제 이름이나 값으로 교체되는 명령행 변수	파일을 삭제하려면 <code>rm filename</code> 명령을 입력하십시오.

셸 프롬프트

셸	프롬프트
C 셸	<i>machine_name%</i>
C 셸 슈퍼유저	<i>machine_name#</i>
Bourne 셸 및 Korn 셸	\$
Bourne 셸 및 Korn 셸 슈퍼유저	#

관련 문서

내용	제목	문서 번호
설치	<i>Sun Cluster 3.0 12/01</i> 소프트웨어 설치 안내서	816-3344-10
하드웨어	<i>Sun Cluster 3.0 12/01 Hardware Guide</i>	816-2023

내용	제목	문서 번호
데이터 서비스	<i>Sun Cluster 3.0 12/01 Data Services Installation and Configuration Guide</i>	816-2024
API 개발	<i>Sun Cluster 3.0 12/01 Data Services Developer's Guide</i>	816-2025
관리	<i>Sun Cluster 3.0 12/01</i> 시스템 관리 안내서	816-3350-10
오류 메시지 및 문제 해결	<i>Sun Cluster 3.0 12/01 Error Messages Guide</i>	816-2028
릴리스 정보	<i>Sun Cluster 3.0 12/01</i> 릴리스 정보	816-3355-10

Sun 문서 주문

인터넷 전문 서점인 [Fatbrain.com](http://fatbrain.com)에서 Sun Microsystems 사의 정선된 제품 참고 문서를 공급합니다. 문서 목록과 주문 방법이 필요하면 다음 주소에 있는 [Fatbrain.com](http://fatbrain.com)의 Sun Documentation Center를 방문하십시오.

<http://www1.fatbrain.com/documentation/sun>

Sun 문서 온라인 액세스

docs.sun.comSM 웹 사이트에서 Sun 기술 문서에 액세스할 수 있습니다. 다음 주소에서 docs.sun.com 아카이브를 살펴보고 특정 서적 제목이나 내용에 대해 검색할 수 있습니다.

<http://docs.sun.com>

도움 받기

SunPlex 시스템 설치 및 사용에 문제가 있으면 서비스 담당자에게 문의하십시오. 문의할 때 다음과 같은 정보가 필요합니다.

- 이름 및 전자 우편 주소(있을 경우)
- 회사 이름, 주소 및 전화 번호
- 시스템 모델 및 일련 번호
- 운영 환경의 릴리스 번호(예: Solaris 8)
- Sun Cluster 소프트웨어의 릴리스 번호(예: Sun Cluster 3.0)

시스템에 있는 각 노드에 대한 정보를 서비스 제공업체에 제공하려면 다음 명령을 사용하여 정보를 수집하십시오.

명령	기능
<code>prtconf -v</code>	시스템 메모리의 크기를 표시하고 주변 장치에 대한 정보를 보고합니다.
<code>psrinfo -v</code>	프로세서에 대한 정보를 표시합니다.
<code>showrev --p</code>	설치된 패치를 알려줍니다.
<code>prtdiag -v</code>	시스템 진단 정보를 표시합니다.
<code>scinstall -pv</code>	Sun Cluster 소프트웨어 릴리스 및 패키지 버전 정보를 표시합니다.

`/var/adm/messages` 파일의 내용도 준비하십시오.

소개 및 개요

SunPlex 시스템은 가용성과 확장성이 높은 서비스를 제공할 수 있도록 통합된 하드웨어 및 소프트웨어 솔루션입니다.

Sun Cluster 3.0 12/01 개념에서는 SunPlex 설명서의 주요 사용자에게 필요한 개념을 설명합니다. 이 설명서의 대상은 다음과 같습니다.

- 클러스터 하드웨어를 설치하고 서비스를 제공하는 서비스 제공업체
- Sun Cluster 소프트웨어를 설치, 구성 및 관리하는 시스템 관리자
- 응용프로그램에 사용할 수 있도록 현재 Sun Cluster 제품에 포함되지 않은 페일오버 및 확장 가능 서비스를 개발하는 응용프로그램 개발자

이 설명서를 읽은 후에 나머지 SunPlex 설명서를 보면 SunPlex 시스템을 완전하게 이해할 수 있습니다.

이 장의 내용은 다음과 같습니다.

- SunPlex 시스템을 소개하고 높은 수준의 개요를 제공합니다.
- SunPlex 사용자의 몇 가지 관점에 대해 설명합니다.
- SunPlex 시스템을 사용하기 전에 알아야 할 주요 개념을 설명합니다.
- 절차와 관련 정보가 포함된 SunPlex 문서 및 주요 개념을 소개합니다.
- 클러스터 관련 작업을 수행하기 위한 각 작업 절차가 포함된 문서를 소개합니다.

SunPlex 시스템 소개

SunPlex 시스템은 Solaris 운영 환경을 클러스터 운영 체제로 확장합니다. 클러스터 또는 플렉스는 데이터베이스, 웹 서비스 및 파일 서비스를 포함한 네트워크 서비스나 응용프로그램을 단일 클라이언트 환경으로 만들어주는 느슨한 결합의 컴퓨팅 노드 모음입니다.

각 클러스터 노드는 독립적으로 자체 프로세스를 실행하는 서버입니다. 이 프로세스는 다른 프로세스와 통신을 통해 사용자에게 응용프로그램, 시스템 자원 및 데이터를 제공하고 네트워크 클라이언트에 하나로 표시되는 단일 시스템을 구성할 수 있습니다.

클러스터에는 기존의 단일 서버 시스템보다 좋은 여러 가지 장점이 있습니다. 즉, 페일 오버 및 확장 가능 서비스 지원, 모듈 단위로 확장할 수 있는 용량, 기존 하드웨어의 고장 허용 시스템에 비해 저렴한 항목 가격 등의 장점이 있습니다.

SunPlex 시스템의 목표는 다음과 같습니다.

- 소프트웨어 또는 하드웨어 결함으로 인한 시스템 작동 중지 시간을 줄이거나 제거합니다.
- 단일 서버 시스템이라면 작동이 중지될 수 있는 수준의 장애가 발생해도 일반 사용자가 데이터와 응용프로그램을 사용할 수 있도록 합니다.
- 클러스터에 노드를 추가하는 방법으로 프로세서를 추가하여 서비스를 확장하고 응용프로그램 처리량을 증가시킵니다.
- 전체 클러스터를 종료하지 않고도 유지 보수를 수행할 수 있도록 하여 높은 시스템 가용성을 제공합니다.

높은 가용성과 고장 허용

SunPlex 시스템은 가용성이 높은 시스템으로 설계되어 데이터 및 응용프로그램에 거의 연속적으로 액세스할 수 있습니다.

그러나 고장 허용 하드웨어 시스템은 데이터 및 응용프로그램에 대한 지속적인 액세스를 제공하지만 특수한 하드웨어를 사용해야 하기 때문에 비용이 많이 듭니다. 또한 소프트웨어 문제로 인한 고장은 일반적으로 고장 허용 시스템에서 해결할 수 없습니다.

SunPlex 시스템은 하드웨어와 소프트웨어 결함을 통해 높은 가용성을 제공합니다. 중복된 클러스터 상호 연결, 기억 장치 및 공용 네트워크를 사용하여 한 위치에서 고장이 발생해도 전체 시스템을 안전하게 보호합니다. 클러스터 소프트웨어는 클러스터 구성 노드의 상태를 계속 모니터링하고, 노드에 장애가 발생하면 클러스터에서 제외하여 데이터가 손상되지 않도록 보호합니다. 또한 클러스터는 서비스와 서비스에 사용되는 시스템 자원을 모니터링하여 장애가 발생하면 페일오버하거나 다시 시작합니다.

높은 가용성에 대한 질문과 대답은 “높은 가용성 FAQ”(67페이지)를 참조하십시오.

SunPlex 시스템의 페일오버 및 확장성

SunPlex 시스템을 사용하면 페일오버 서비스나 확장 가능 서비스를 구현할 수 있습니다. 일반적으로 페일오버 서비스는 높은 가용성(중복)만 제공하지만 확장 가능 서비스는 높은 가용성과 함께 높은 성능을 제공합니다. 단일 클러스터에서 페일오버 서비스와 확장 가능 서비스를 모두 지원할 수 있습니다.

페일오버 서비스

페일오버는 1차 노드에 장애가 발생할 경우에 자동으로 1차 노드의 서비스를 지정된 2차 노드로 전환하는 프로세스입니다. Sun Cluster 소프트웨어는 페일오버를 사용하여 높은 가용성을 제공합니다.

페일오버가 발생하면 클라이언트의 서비스 연결이 일시 중단될 수 있으며, 페일오버가 종료된 후에 다시 연결해야 할 수도 있습니다. 그러나 클라이언트에서는 서비스를 제공하는 물리적 서버를 알 수 없습니다.

확장 가능 서비스

페일오버는 중복성과 관련이 있는 반면, 확장성은 로드와 관계없이 일정한 응답 시간 또는 처리량을 제공합니다. 확장 가능 서비스를 사용하면 클러스터의 여러 노드가 동시에 응용프로그램을 실행할 수 있기 때문에 성능이 향상됩니다. 확장 가능한 구성에서는 클러스터의 각 노드가 데이터를 제공하고 클라이언트 요청을 처리할 수 있습니다.

페일오버 및 확장 가능 서비스에 대한 자세한 내용은 “데이터 서비스”(48페이지)를 참조하십시오.

SunPlex 시스템에 대한 세 가지 관점

이 단원에서는 SunPlex 시스템에 대한 서로 다른 세 가지 관점과 주요 개념 및 각 관점과 관련된 문서에 대하여 설명합니다. 각 관점은 다음과 같습니다.

- 하드웨어 설치 및 서비스 담당자
- 시스템 관리자
- 응용프로그램 프로그래머

하드웨어 설치 및 서비스 관점

하드웨어 서비스 담당자의 경우에는 SunPlex 시스템을 서버, 네트워크 및 기억 장치가 포함된 완성된 하드웨어 모음으로 생각할 수 있습니다. 이 구성 요소는 모든 구성 요소를 백업하여 단일 지점의 장애가 발생하지 않도록 케이블로 연결되어 있습니다.

주요 개념 – 하드웨어

하드웨어 서비스 담당자는 다음과 같은 클러스터 개념을 이해해야 합니다.

- 클러스터 하드웨어 구성 및 케이블 설치
- 설치 및 서비스 제공(추가, 제거, 교체)

- 네트워크 인터페이스 구성 요소(어댑터, 연결 장치, 케이블)
 - 디스크 인터페이스 카드
 - 디스크 배열
 - 디스크 드라이브
 - 관리 콘솔 및 콘솔 액세스 장치
- 관리 콘솔 및 콘솔 액세스 장치 설정

권장하는 하드웨어 개념 참조 단락

다음 단락에는 위의 주요 개념과 관련된 자료가 있습니다.

- “클러스터 노드”(21페이지)
- “멀티호스트 디스크”(22페이지)
- “로컬 디스크”(24페이지)
- “클러스터 상호 연결”(24페이지)
- “공용 네트워크 인터페이스”(25페이지)
- “클라이언트 시스템”(25페이지)
- “관리 콘솔”(26페이지)
- “콘솔 액세스 장치”(26페이지)
- “클러스터 쌍 토폴로지”(27페이지)
- “N+1(방사형) 토폴로지”(29페이지)

관련 SunPlex 문서

다음 SunPlex 문서에는 하드웨어 서비스 개념과 관련된 절차 및 정보가 있습니다.

- *Sun Cluster 3.0 12/01 Hardware Guide*

시스템 관리자 관점

시스템 관리자의 경우에는 SunPlex 시스템이 기억 장치를 공유하고 케이블로 연결된 서버(노드) 세트라고 생각할 수 있습니다. 시스템 관리자는 다음 사항을 이해해야 합니다.

- 클러스터 노드 사이의 연결을 모니터링할 수 있도록 Solaris 소프트웨어에 통합된 특수 클러스터 소프트웨어
- 클러스터 노드에서 실행되는 사용자 응용프로그램의 상태를 모니터링할 수 있는 특수 소프트웨어
- 디스크를 설정하고 관리하는 볼륨 관리 소프트웨어

- 모든 노드가 직접 디스크에 연결되지 않아도 모든 기억 장치에 액세스할 수 있도록 하는 특수 클러스터 소프트웨어
- 파일이 노드에 로컬로 연결된 것처럼 모든 노드에 파일이 표시되도록 하는 특수 클러스터 소프트웨어

주요 개념 – 시스템 관리

시스템 관리자는 다음 개념과 프로세스를 이해해야 합니다.

- 하드웨어 및 소프트웨어 구성 요소 사이의 상호 작용
- 다음 작업을 포함한 클러스터 설치 및 구성 방법에 대한 일반적인 흐름
 - Solaris 운영 환경 설치
 - Sun Cluster 소프트웨어 설치 및 구성
 - 볼륨 관리자 설치 및 구성
 - 클러스터에서 사용할 수 있도록 응용프로그램 소프트웨어 설치 및 구성
 - Sun Cluster 데이터 서비스 소프트웨어 설치 및 구성
- 클러스터 하드웨어 및 소프트웨어 구성 요소를 추가, 제거 및 교체하고 서비스를 제공하기 위한 클러스터 관리 절차
- 성능을 개선하기 위한 구성 변경

권장하는 시스템 관리자 개념 참조 단락

다음 단락에는 위의 주요 개념과 관련된 자료가 있습니다.

- “관리 인터페이스”(33페이지)
- “가용성이 높은 프레임워크”(33페이지)
- “글로벌 장치”(36페이지)
- “디스크 장치 그룹”(37페이지)
- “글로벌 이름 공간”(39페이지)
- “클러스터 파일 시스템”(41페이지)
- “정족수 및 정족수 장치”(43페이지)
- “볼륨 관리자”(47페이지)
- “데이터 서비스”(48페이지)
- “자원, 자원 그룹 및 자원 유형”(58페이지)
- “PNM(Public Network Management) 및 NAFO(Network Adapter Failover)”(61페이지)
- 4장

관련 SunPlex 문서 – 시스템 관리자

다음 SunPlex 문서에는 시스템 관리 개념과 관련된 절차와 정보가 있습니다.

- *Sun Cluster 3.0 12/01* 소프트웨어 설치 안내서
- *Sun Cluster 3.0 12/01* 시스템 관리 안내서
- *Sun Cluster 3.0 12/01 Error Messages Guide*

응용프로그램 프로그래머 관점

SunPlex 시스템은 Oracle, NFS, DNS, iPlanet™ Web Server, Apache Web Server, Netscape™ Directory Server 등과 같은 응용프로그램을 위한 데이터 서비스를 제공합니다. 별도 구입한 응용프로그램이 Sun Cluster 소프트웨어의 제어를 통해 실행되도록 구성하면 데이터 서비스가 완성됩니다. Sun Cluster 소프트웨어는 응용프로그램을 시작하고, 중지하고, 모니터링하는 구성 파일과 관리 방법을 제공합니다. 페일오버 서비스나 확장 가능 서비스를 새로 구성해야 하는 경우에 SunPlex API(Application Programming Interface) 및 DSET API(Data Service Enabling Technologies API)를 사용하면 클러스터에서 응용프로그램을 데이터 서비스로 실행할 수 있도록 필요한 구성 파일과 관리 방법을 개발할 수 있습니다.

주요 개념 – 응용프로그램 프로그래머

응용프로그램 프로그래머는 다음 사항을 이해해야 합니다.

- 사용하는 응용프로그램이 페일오버 서비스나 확장 가능 데이터 서비스로 실행될 수 있는지를 판단하는 응용프로그램 특성
- Sun Cluster API, DSET API 및 “범용” 데이터 서비스. 프로그래머가 클러스터 환경에 맞게 응용프로그램을 구성하려면 프로그램이나 스크립트를 작성할 때 가장 적합한 도구를 결정해야 합니다.

권장하는 응용프로그램 프로그래머 개념 참조 단락

다음 단락에는 위의 주요 개념과 관련된 자료가 있습니다.

- “데이터 서비스”(48페이지)
- “자원, 자원 그룹 및 자원 유형”(58페이지)
- 4장

관련 SunPlex 문서 – 응용프로그램 프로그래머

다음 SunPlex 문서에는 응용프로그램 프로그래머 개념과 관련된 절차 및 정보가 있습니다.

- *Sun Cluster 3.0 12/01 Data Services Developer's Guide*
- *Sun Cluster 3.0 12/01 Data Services Installation and Configuration Guide*

SunPlex 시스템 작업

모든 SunPlex 시스템 작업에 약간의 배경 개념이 필요합니다. 다음 표에 높은 수준의 작업과 작업 단계를 설명하는 문서 목록이 있습니다. 이 책의 개념 단원에서는 개념과 작업 사이의 매핑 관계를 설명합니다.

표1-1 작업 맵: 사용자의 작업을 위해 참조할 문서 소개

수행할 작업	참조할 문서
클러스터 하드웨어 설치	<i>Sun Cluster 3.0 12/01 Hardware Guide</i>
클러스터에 Solaris 소프트웨어 설치	<i>Sun Cluster 3.0 12/01</i> 소프트웨어 설치 안내서
Sun™ Management Center 소프트웨어 설치	<i>Sun Cluster 3.0 12/01</i> 소프트웨어 설치 안내서
Sun Cluster 소프트웨어 설치 및 구성	<i>Sun Cluster 3.0 12/01</i> 소프트웨어 설치 안내서
볼륨 관리 소프트웨어 설치 및 구성	<i>Sun Cluster 3.0 12/01</i> 소프트웨어 설치 안내서 볼륨 관리 문서
Sun Cluster 데이터 서비스 설치 및 구성	<i>Sun Cluster 3.0 12/01 Data Services Installation and Configuration Guide</i>
서비스 클러스터 하드웨어	<i>Sun Cluster 3.0 12/01 Hardware Guide</i>
Sun Cluster 소프트웨어 관리	<i>Sun Cluster 3.0 12/01</i> 시스템 관리 안내서
볼륨 관리 소프트웨어 관리	<i>Sun Cluster 3.0 12/01</i> 시스템 관리 안내서 및 볼륨 관리 문서
응용프로그램 소프트웨어 관리	응용프로그램 문서

표1-1 작업 맵: 사용자의 작업을 위해 참조할 문서 소개 계속

수행할 작업	참조할 문서
문제 식별 및 권장하는 사용자 조치	<i>Sun Cluster 3.0 12/01 Error Messages Guide</i>
새 데이터 서비스 구성	<i>Sun Cluster 3.0 12/01 Data Services Developer's Guide</i>

주요 개념 - 하드웨어 서비스 제공업체

이 단원에서는 SunPlex 시스템을 구성하는 하드웨어 구성 요소와 관련된 주요 개념에 대하여 설명합니다. 이 단원의 내용은 다음과 같습니다.

- “클러스터 노드”(21페이지)
- “멀티호스트 디스크”(22페이지)
- “로컬 디스크”(24페이지)
- “이동식 미디어”(24페이지)
- “클러스터 상호 연결”(24페이지)
- “공용 네트워크 인터페이스”(25페이지)
- “클라이언트 시스템”(25페이지)
- “콘솔 액세스 장치”(26페이지)
- “관리 콘솔”(26페이지)
- “Sun Cluster 토폴로지의 예”(27페이지)

SunPlex 시스템 하드웨어 구성 요소

이 정보는 기본적으로 하드웨어 서비스 제공업체를 위한 내용입니다. 이 개념은 서비스 제공업체에서 클러스터 하드웨어를 설치, 구성하거나 서비스를 제공하기 전에 하드웨어 구성 요소 사이의 관계를 이해하는 데 도움이 됩니다. 클러스터 시스템 관리자는 클러스터 소프트웨어 설치, 구성 및 관리에 대한 배경 정보로 이 정보를 사용할 수 있습니다.

클러스터는 다음과 같은 몇 가지 하드웨어 구성 요소로 구성됩니다.

- 로컬 디스크(비공유)가 있는 클러스터 노드

- 멀티호스트 기억 장치(노드 사이의 공유 디스크)
- 이동식 미디어(테이프 및 CD-ROM)
- 클러스터 상호 연결
- 공용 네트워크 인터페이스
- 클라이언트 시스템
- 관리 콘솔
- 콘솔 액세스 장치

SunPlex 시스템을 사용하면 “Sun Cluster 토폴로지의 예”(27페이지)에서 설명하는 다양한 구성 방법으로 이 구성 요소를 결합할 수 있습니다.

다음 그림은 클러스터 구성 샘플입니다.

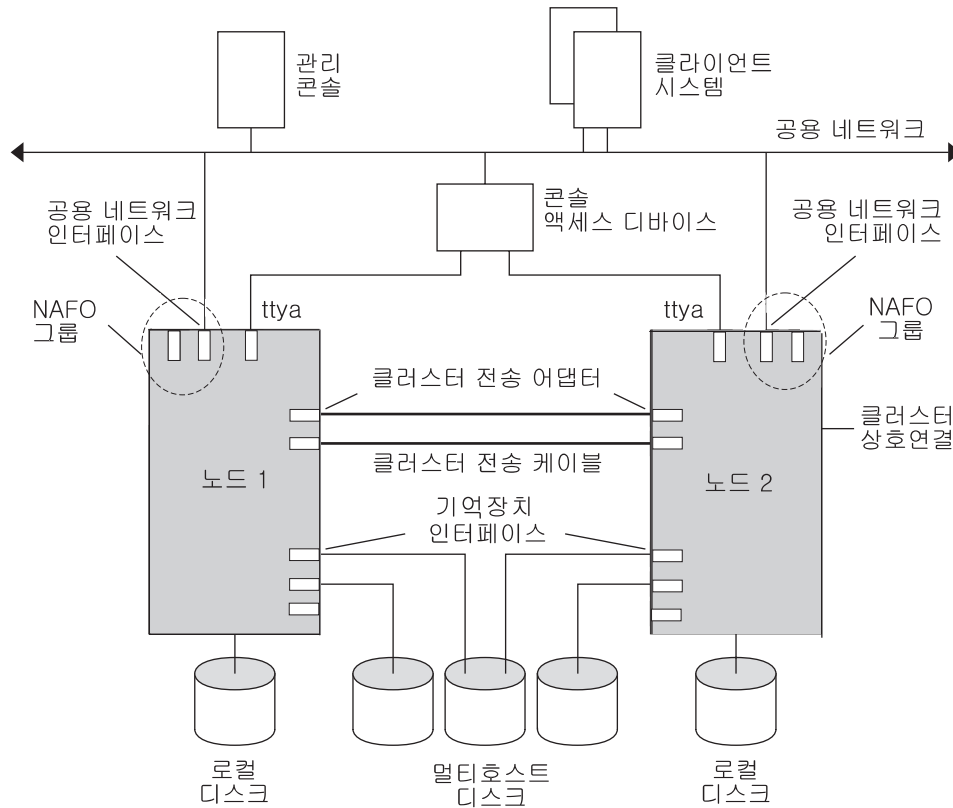


그림 2-1 2-노드 클러스터 구성 샘플

클러스터 노드

클러스터 노드는 **Solaris** 운영 환경 및 **Sun Cluster** 소프트웨어를 모두 실행하는 시스템으로, 클러스터의 현재 구성원(클러스터 구성원)이거나 구성원이 될 수 있는 시스템입니다. **Sun Cluster** 소프트웨어가 있으면 클러스터에 2 - 8개의 노드가 있을 수 있습니다. 지원되는 노드 구성은 “**Sun Cluster** 토폴로지의 예”(27페이지)를 참조하십시오.

클러스터 노드는 일반적으로 하나 이상의 멀티호스트 디스크에 연결됩니다. 멀티호스트 디스크에 연결되지 않은 노드는 멀티호스트 디스크에 액세스하기 위해 클러스터 파일 시스템을 사용합니다. 예를 들어, 확장 가능 서비스를 하나 구성하면 멀티호스트 디스크에 직접 연결되지 않아도 노드가 요청에 서비스를 제공할 수 있습니다.

병렬 데이터베이스 구성의 노드는 모든 디스크에 대한 동시 액세스를 공유합니다. 병렬 데이터베이스 구성에 대한 자세한 내용은 “멀티호스트 디스크”(22페이지) 및 3장을 참조하십시오.

클러스터의 모든 노드는 클러스터에 액세스하여 관리하기 위해 사용하는 공용 이름(클러스터 이름)으로 그룹화됩니다.

공용 네트워크 어댑터는 노드를 공용 네트워크에 연결하여 클러스터에 대한 클라이언트 액세스를 제공합니다.

클러스터 구성원은 물리적으로 독립된 하나 이상의 네트워크를 통해 클러스터의 다른 노드와 통신을 합니다. 이렇게 물리적으로 독립된 네트워크 세트를 클러스터 상호 연결이라고 합니다.

다른 노드가 클러스터에 결합되거나 클러스터에서 제거될 때 클러스터의 모든 노드가 이것을 인식합니다. 또한 클러스터의 모든 노드가 로컬로 실행되는 자원뿐 아니라 다른 클러스터 노드에서 실행되는 자원을 인식합니다.

성능이 크게 떨어지지 않고 페일오버가 발생하도록 하려면 동일한 클러스터의 노드가 모두 유사한 프로세싱, 메모리 및 I/O 기능을 사용해야 합니다. 페일오버를 위해 모든 노드에 다른 노드의 워크로드를 백업하거나 2차 노드가 될 수 있을 만큼 충분한 용량이 있어야 합니다.

각 노드는 개별 루트(/) 파일 시스템을 부트합니다.

클러스터 하드웨어 구성원을 위한 소프트웨어 구성 요소

클러스터 구성원의 기능을 하려면 다음 소프트웨어가 설치되어야 합니다.

- **Solaris** 운영 환경
- **Sun Cluster** 소프트웨어
- 데이터 서비스 응용프로그램
- 볼륨 관리자(**Solstice DiskSuite™** 또는 **VERITAS Volume Manager**)

독립된 디스크로 구성된 하드웨어 중복 배열(**RAID**)을 사용하는 **OPS(Oracle Parallel Server)** 구성에서는 예외입니다. 이러한 구성에서는 **Oracle** 데이터를 관리하기 위해 **Solstice DiskSuite** 또는 **VERITAS Volume Manager**와 같은 소프트웨어 볼륨 관리자가 필요없습니다.

Solaris 운영 환경, Sun Cluster 및 볼륨 관리 소프트웨어를 설치하는 방법은 *Sun Cluster 3.0 12/01* 소프트웨어 설치 안내서를 참조하십시오.

데이터 서비스 설치 및 구성 방법은 *Sun Cluster 3.0 12/01 Data Services Installation and Configuration Guide*를 참조하십시오.

이전의 소프트웨어 구성 요소에 대한 개념은 3장을 참조하십시오.

다음 그림은 Sun Cluster 소프트웨어 환경을 구성하기 위해 사용하는 높은 수준의 소프트웨어 구성 요소입니다.

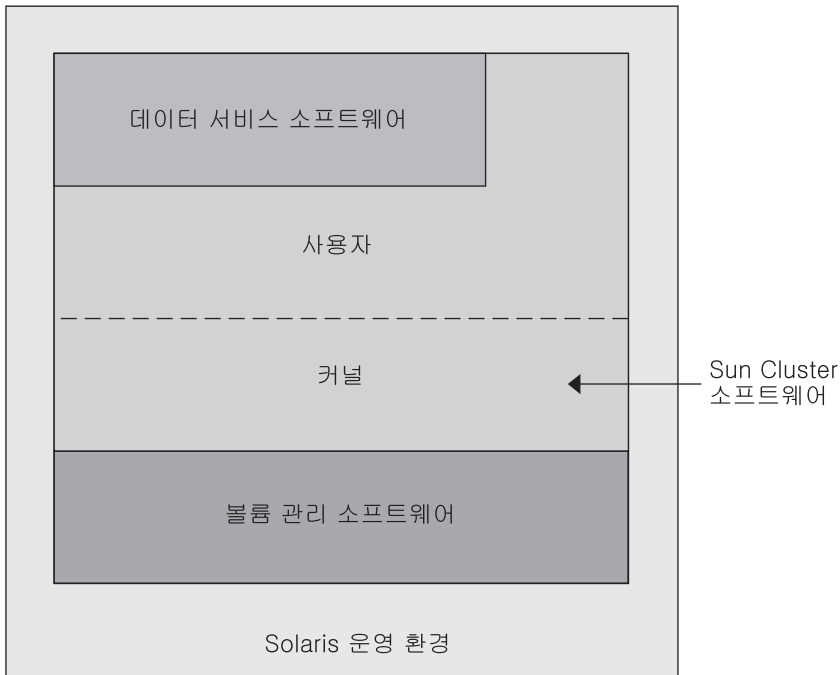


그림 2-2 높은 수준의 Sun Cluster 소프트웨어 구성 요소 관계
클러스터 구성원에 대한 질문과 대답은 4장을 참조하십시오.

멀티호스트 디스크

Sun Cluster에는 멀티호스트 기억 장치가 필요합니다. 멀티호스트 기억 장치는 동시에 둘 이상의 노드에 연결할 수 있는 디스크입니다. Sun Cluster 환경에 멀티호스트 기억 장치를 사용하면 디스크 가용성을 높일 수 있습니다.

멀티호스트 디스크에는 다음과 같은 기능이 있습니다.

- 노드 하나에 장애가 발생해도 계속 작동합니다.

- 응용프로그램 데이터를 저장하고 응용프로그램 바이너리와 구성 파일도 저장할 수 있습니다.
- 노드 장애로부터 보호합니다. 클라이언트 요청이 하나의 노드를 통해 데이터에 액세스할 때 노드에 장애가 발생하면 동일한 디스크에 직접 연결된 다른 노드를 사용하도록 요청이 스위치오버됩니다.
- 멀티호스트 디스크는 디스크를 “마스터”하는 1차 노드를 통하거나 로컬 경로를 통한 직접 동시 액세스에 의해 글로벌하게 액세스됩니다. 현재는 OPS만이 직접 동시 액세스를 사용합니다.

볼륨 관리자는 멀티호스트 디스크의 데이터 중복을 위해 이중화 또는 RAID-5 구성을 제공합니다. 현재 Sun Cluster에서는 Solstice DiskSuite™ 및 VERITAS Volume Manager를 볼륨 관리자로 사용하고 Sun StorEdge™ A3x00 기억 장치 유닛에 RDAC RAID-5 하드웨어 컨트롤러를 사용할 수 있습니다.

멀티호스트 디스크를 디스크 이중화 및 스트라이핑과 결합하면 노드 장애와 각 디스크 장애로부터 보호할 수 있습니다.

멀티호스트 기억 장치에 대한 질문과 대답은 4장을 참조하십시오.

Multi-Initiator SCSI

이 단원의 내용은 SCSI 기억 장치에만 적용되고 멀티호스트 디스크에서 사용하는 광섬유 채널 기억 장치에는 적용되지 않습니다.

독립형 서버에서는 서버 노드가 서버를 특정 SCSI 버스에 연결하는 SCSI 호스트 어댑터 회로를 사용하여 SCSI 버스 작동을 제어합니다. 이러한 SCSI 호스트 어댑터 회로를 SCSI initiator라고 합니다. 이 회로가 SCSI 버스에 대한 모든 버스 작업을 시작합니다. Sun 시스템에서 SCSI 호스트 어댑터의 기본 SCSI 주소는 7입니다.

클러스터 구성에서는 멀티호스트 디스크를 사용하여 여러 서버 노드 사이에서 기억 장치를 공유합니다. 클러스터 기억 장치가 종단 장치가 하나인 SCSI 장치나 차동 SCSI 장치로 구성된 경우에 이러한 구성을 multi-initiator SCSI라고 합니다. 이 용어가 의미하는 것처럼 SCSI 버스에는 둘 이상의 SCSI initiator가 있습니다.

SCSI 사양에 따라 SCSI 버스의 각 장치에 고유한 SCSI 주소가 있어야 합니다. 호스트 어댑터도 SCSI 버스의 장치입니다. Multi-initiator 환경의 기본 하드웨어 구성에서는 모든 SCSI 호스트 어댑터가 7로 기본 설정되기 때문에 충돌합니다.

이러한 충돌을 해결하려면 각 SCSI 버스에서 SCSI 호스트 어댑터 중 하나만 SCSI 주소를 7로 남겨 두고 나머지 호스트 어댑터는 사용하지 않는 SCSI 주소로 설정해야 합니다. 이후에 충돌하지 않도록 제대로 계획하려면 현재도 사용하지 않고 나중에도 사용하지 않을 주소를 “사용하지 않는” SCSI 주소로 생각해야 합니다. 나중에도 사용하지 않을 주소의 예로는 새로운 드라이브를 빈 드라이브 슬롯에 설치하여 기억 장치를 추가하는 경우가 있습니다. 대부분의 구성에서 2차 호스트 어댑터의 SCSI 주소로 6을 사용할 수 있습니다.

scsi-initiator-id OBP(Open Boot PROM) 등록 정보를 설정하여 이러한 호스트 어댑터에 대한 SCSI 주소를 변경할 수 있습니다. 이 등록 정보를 노드에 대하여 글로벌로 설정할 수도 있고 호스트 어댑터 단위로 설정할 수도 있습니다. 각 SCSI 호스트 어댑터에 대하여 고유한 scsi-initiator-id를 설정하는 방법은 *Sun Cluster 3.0 12/01 Hardware Guide*에서 각 디스크 인클로저에 대해 설명하는 단원에 있습니다.

로컬 디스크

로컬 디스크는 하나의 노드에만 연결된 디스크입니다. 따라서 노드 장애로부터 보호되지 않습니다(가용성이 높지 않음). 그러나 로컬 디스크를 포함한 모든 디스크가 글로벌 이름 공간에 포함되고 글로벌 장치로 구성됩니다. 따라서 디스크 자체는 모든 클러스터 노드에서 볼 수 있습니다.

이러한 디스크의 파일 시스템을 글로벌 마운트 포인트 아래에 지정하면 다른 노드가 사용하도록 만들 수 있습니다. 현재는 이러한 글로벌 파일 시스템 중 하나가 마운트된 노드에서 장애가 발생하면 모든 노드가 해당 파일 시스템에 액세스할 수 없게 됩니다. 볼륨 관리자를 사용하면 디스크 장애가 있어도 이러한 파일 시스템에 액세스할 수 있도록 디스크를 이중화할 수 있지만, 노드 장애로부터 보호하지는 않습니다.

글로벌 장치에 대한 자세한 내용은 “글로벌 장치”(36페이지) 단원을 참조하십시오.

이동식 미디어

테이프 드라이브 및 CD-ROM 드라이브와 같은 이동식 미디어가 클러스터에서 지원됩니다. 일반적으로 클러스터링되지 않은 환경에서와 동일한 방법으로 이러한 장치를 설치, 구성하고 서비스를 제공할 수 있습니다. 이 장치는 Sun Cluster에 글로벌 장치로 구성되므로 클러스터의 모든 노드에서 각 장치에 액세스할 수 있습니다. 이동식 미디어 설치와 구성에 대한 자세한 내용은 *Sun Cluster 3.0 12/01 Hardware Guide*를 참조하십시오.

글로벌 장치에 대한 자세한 내용은 “글로벌 장치”(36페이지) 단원을 참조하십시오.

클러스터 상호 연결

클러스터 상호 연결은 클러스터 노드 사이에서 클러스터 개인 통신과 데이터 서비스 통신을 전송하기 위해 사용하는 물리적 장치 구성입니다. 상호 연결은 클러스터 개인 통신에 광범위하게 사용되기 때문에 이 구성이 성능을 제한할 수 있습니다.

클러스터 노드만 클러스터 상호 연결에 연결될 수 있습니다. Sun Cluster 보안 모델에서는 클러스터 노드만 클러스터 상호 연결에 물리적으로 액세스할 수 있다고 가정합니다.

한 위치의 장애로 인해 작동이 중단되지 않도록 하려면 모든 노드가 클러스터 상호 연결에 의해 물리적으로 독립된 두 개 이상의 중복 네트워크나 경로를 통해 연결되어야 합니다. 두 노드 사이에 물리적으로 독립된 여러 개의 네트워크(2 - 6개)를 사용할 수

있습니다. 클러스터 상호 연결은 어댑터, 연결 장치, 케이블 등의 세 가지 하드웨어로 구성됩니다.

다음은 이러한 하드웨어 구성 요소 각각에 대한 설명입니다.

- 어댑터 - 노드에 있는 물리적 네트워크 카드. 이름은 **qfe2**와 같이 장치 이름 뒤에 물리적인 장치 번호를 붙여서 만듭니다. 일부 어댑터에는 단 하나의 물리적 네트워크 연결이 있지만 **qfe** 카드와 같은 다른 어댑터에는 여러 개의 물리적 연결이 있을 수 있습니다. 또한 일부 어댑터에는 네트워크 인터페이스와 기억 장치 인터페이스가 모두 포함되어 있습니다.

인터페이스가 여러 개인 네트워크 카드는 전체 카드에 장애가 발생할 경우 단일 장애 지점이 될 수 있습니다. 가용성을 최대한 높이려면 두 노드 사이의 경로가 단일 네트워크 카드에만 의존하지 않도록 클러스터를 계획하십시오.

- 연결 장치 - 클러스터 노드 외부에 있는 스위치. 연결 장치는 세 개 이상의 노드에 연결할 수 있도록 바로 전달 기능과 전환 기능을 수행합니다. 2-노드 클러스터에서는 노드가 각 노드의 중복 어댑터에 연결된 물리적 중복 케이블을 통해 서로 직접 연결될 수 있으므로 연결 장치가 필요없습니다. 일반적으로 세 개 이상의 노드로 된 구성에는 연결 장치가 필요합니다.
- 케이블 - 두 개의 네트워크 어댑터 사이 및 어댑터와 연결 장치 사이의 물리적 연결. 클러스터 상호 연결에 대한 질문과 대답은 4장을 참조하십시오.

공용 네트워크 인터페이스

클라이언트는 공용 네트워크 인터페이스를 통해 클러스터에 연결합니다. 각 네트워크 어댑터 카드는 카드에 있는 하드웨어 인터페이스 수에 따라 하나 이상의 공용 네트워크에 연결할 수 있습니다. 하나의 카드는 실제로 작동하고 다른 카드는 백업으로 작동하도록, 구성된 여러 개의 공용 네트워크 인터페이스 카드를 포함하여 노드를 설정할 수 있습니다. “PNM(Public Network Management)”이라는 Sun Cluster 소프트웨어 하위 시스템이 작동 중인 인터페이스를 모니터링합니다. 작동 중인 어댑터에 장애가 발생하면 인터페이스를 백업 어댑터 중 하나로 페일오버하기 위해 NAFO(Network Adapter Failover) 소프트웨어를 호출합니다.

클러스터링에서 공용 네트워크 인터페이스를 위하여 특별히 하드웨어를 고려할 필요는 없습니다.

공용 네트워크에 대한 질문과 대답은 4장을 참조하십시오.

클라이언트 시스템

클라이언트 시스템에는 공용 네트워크를 통해 클러스터에 액세스하는 다른 서버나 워크스테이션이 포함됩니다. 클라이언트측 프로그램은 클러스터에서 실행되는 서버측 응용프로그램에서 제공하는 데이터나 기타 서비스를 사용합니다.

클라이언트 시스템은 가용성이 높지 않습니다. 클러스터의 데이터 및 응용프로그램은 가용성이 높습니다.

클라이언트 시스템에 대한 질문과 대답은 4장을 참조하십시오.

콘솔 액세스 장치

모든 클러스터 노드에 대해 콘솔 액세스가 있어야 합니다. 콘솔에 액세스하려면 클러스터 하드웨어와 함께 구입한 터미널 콘센트레이터, Sun Enterprise E10000 서버의 SSP(System Service Processor), Sun Fire™ 서버의 시스템 컨트롤러 또는 각 노드의 ttya에 액세스할 수 있는 기타 장치를 사용하십시오.

Sun에서는 지원되는 터미널 콘센트레이터를 하나만 제공합니다. 지원되는 Sun 터미널 콘센트레이터를 사용할 것인지는 사용자가 선택할 수 있습니다. 터미널 콘센트레이터를 사용하면 TCP/IP 네트워크를 사용하여 각 노드에서 /dev/console에 액세스할 수 있습니다. 따라서 네트워크 상의 모든 원격 워크스테이션에서 콘솔 레벨로 각 노드에 액세스할 수 있습니다.

SSP(System Service Processor)는 Sun Enterprise E10000 서버에 대한 콘솔 액세스를 제공합니다. SSP는 Sun Enterprise E10000 서버를 지원하기 위해 구성된 이더넷 네트워크 상의 시스템입니다. SSP는 Sun Enterprise E10000 서버용 관리 콘솔입니다. Sun Enterprise E10000 네트워크 콘솔 기능을 사용하면 네트워크 상의 어떤 워크스테이션에서도 호스트 콘솔 세션을 열 수 있습니다.

다른 콘솔 액세스 방법에는 다른 터미널 콘센트레이터가 포함됩니다(다른 노드나 단순 터미널로부터의 tip(1) 직렬 포트 액세스). Sun™ 키보드 및 모니터를 사용하거나 하드웨어 서비스 제공업체에서 지원하는 다른 직렬 포트 장치를 사용할 수 있습니다.

관리 콘솔

관리 콘솔이라는 전용 SPARCstation™ 시스템을 사용하여 작동 중인 클러스터를 관리할 수 있습니다. 일반적으로 Sun Management Center 제품에는 CCP(Cluster Control Panel) 및 Sun Cluster 모듈과 같은 관리 도구 소프트웨어를 관리 콘솔에 설치하고 실행합니다. CCP에서 cconsole을 사용하면 동시에 둘 이상의 노드 콘솔에 연결할 수 있습니다. CCP 사용 방법은 Sun Cluster 3.0 12/01 시스템 관리 안내서의 내용을 참조하십시오.

관리 콘솔은 클러스터 노드가 아닙니다. 네트워크용 터미널 콘센트레이터를 통해 공용 또는 선택적 네트워크를 거쳐 클러스터 노드에 원격 액세스하기 위해 관리 콘솔을 사용합니다. 클러스터가 Sun Enterprise E10000 플랫폼으로 구성된 경우에는 관리 콘솔에서 netcon(1M) 명령을 사용하여 SSP(System Service Processor)에 로그인하고 연결할 수 있어야 합니다.

일반적으로 모니터 없이 노드를 구성합니다. 그러면 터미널 콘센트레이터에 연결되어 있는 관리 콘솔과 노드의 직렬 포트에 연결된 터미널 콘센트레이터에서 telnet 세션을 통해 노드 콘솔에 액세스합니다. Sun Enterprise E10000 서버의 경우에는 SSP에서 연결합니다. 자세한 내용은 “콘솔 액세스 장치”(26페이지)를 참조하십시오.

Sun Cluster에 전용 관리 콘솔이 반드시 필요한 것은 아니지만 전용 관리 콘솔을 사용하면 다음과 같은 이점이 있습니다.

- 동일한 시스템에서 콘솔과 관리 도구를 그룹화하여 중앙에서 클러스터를 관리할 수 있습니다.
- 하드웨어 서비스 제공업체에서 더욱 빠르게 문제를 분석할 수 있습니다.

관리 콘솔에 대한 질문과 대답은 4장을 참조하십시오.

Sun Cluster 토폴로지의 예

토폴로지는 클러스터에서 사용하는 기억 장치 플랫폼에 클러스터 노드를 연결하는 연결 구조입니다.

Sun Cluster는 다음과 같은 토폴로지를 지원합니다.

- 클러스터 쌍
- Pair+M
- N+1(방사형)

다음 단락에 각 토폴로지의 예를 보여주는 그림이 있습니다.

클러스터 쌍 토폴로지

클러스터 쌍 토폴로지는 단일 클러스터 관리 프레임워크에서 작동하는 두 개 이상의 노드 쌍입니다. 이 구성에서는 쌍 사이에서만 페일오버가 발생합니다. 그러나 모든 노드는 클러스터 상호 연결에 의해 연결되고 Sun Cluster 소프트웨어의 제어에 의해 작동합니다. 이 토폴로지를 사용하면 하나의 쌍에서 병렬 데이터베이스 응용프로그램을 실행하고 다른 쌍에서는 페일오버 응용프로그램이나 확장 가능 응용프로그램을 실행할 수 있습니다.

클러스터 파일 시스템을 사용하면, 모든 노드가 응용프로그램 데이터를 저장하는 디스크에 직접 연결되어 있지 않아도 세 개 이상의 노드가 확장 가능 서비스나 병렬 데이터베이스를 실행하는 두 쌍 구성도 사용할 수 있습니다.

다음 그림은 클러스터 쌍 토폴로지의 구성입니다.

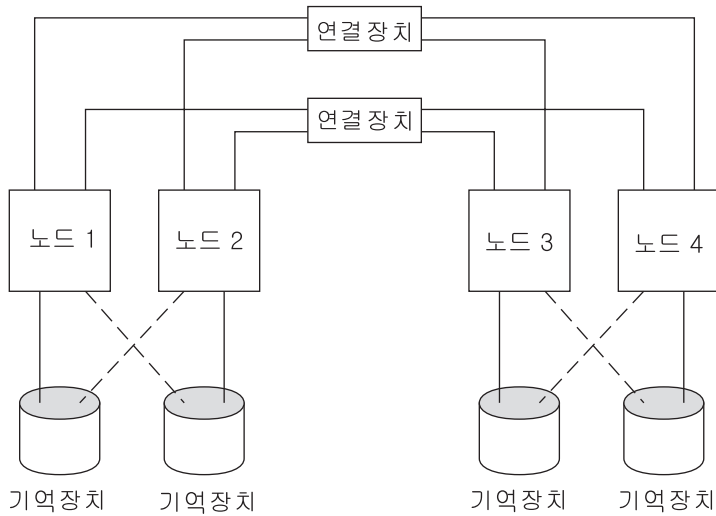


그림 2-3 클러스터 쌍 토폴로지

Pair+M 토폴로지

pair+M 토폴로지는 공유 기억 장치 및 추가 노드 세트에 직접 연결된 노드 쌍이 포함됩니다. 이 추가 노드 세트는 자체적으로 직접 연결하지 않고 클러스터 상호 연결을 사용하여 공유 기억 장치에 액세스합니다.

다음 그림은 네 개의 노드 중 두 개의 노드(노드 3 및 노드 4)가 클러스터 상호 연결을 사용하여 기억 장치에 액세스하는 pair+M 토폴로지의 예입니다. 공유 기억 장치에 직접 액세스하지 않는 추가 노드를 포함하여 이 구성을 확장할 수 있습니다.

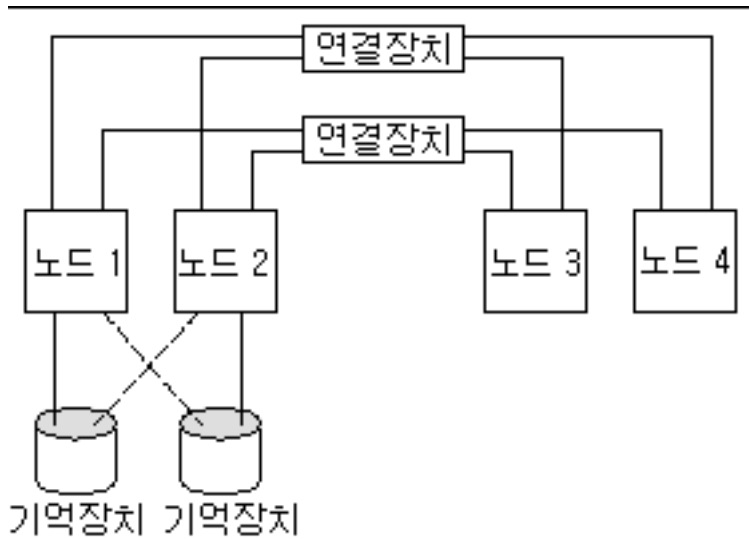


그림 2-4 Pair+M 토폴로지

N+1(방사형) 토폴로지

N+1 토폴로지에는 1차 노드 몇 개와 2차 노드 하나가 포함됩니다. 1차 노드와 2차 노드를 동일하게 구성할 필요는 없습니다. 항상 1차 노드가 응용프로그램 서비스를 제공합니다. 1차 노드에 장애가 발생할 때까지 대기하는 동안 2차 노드가 유휴 상태가 될 필요는 없습니다.

이 구성에서는 2차 노드만이 모든 멀티호스트 기억 장치에 물리적으로 연결됩니다.

1차 노드에서 장애가 발생하면 Sun Cluster가 자원을 2차 노드로 페일오버합니다. 그곳에서 자원은(자동 또는 수동으로) 다시 1차 노드로 전환될 때까지 기능을 수행합니다.

2차 노드는 1차 노드 중 하나에 장애가 발생할 경우에 부하를 처리할 수 있을 만큼 충분한 CPU 용량이 있어야 합니다.

다음 그림은 N+1 구성의 예입니다.

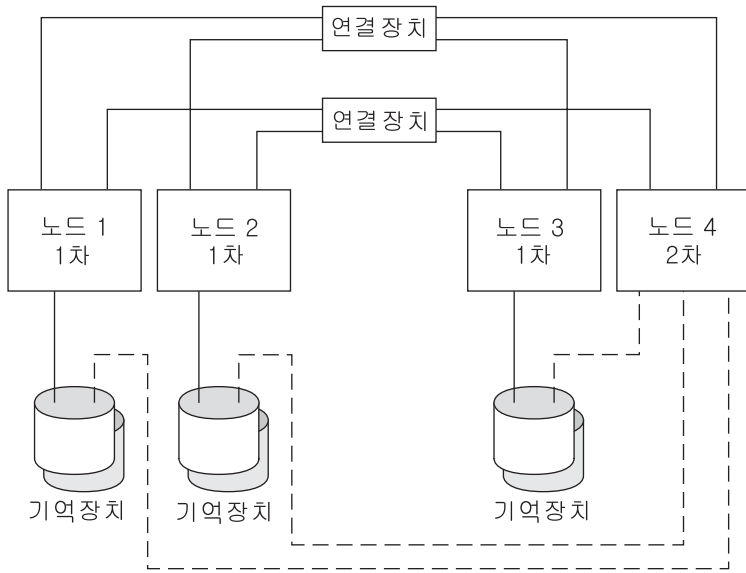


그림 2-5 N+1 토폴로지

주요 개념 - 관리 및 응용프로그램 개발

이 단원에서는 SunPlex 시스템의 소프트웨어 구성 요소와 관련된 주요 개념에 대해 설명합니다. 주요 내용은 다음과 같습니다.

- “관리 인터페이스”(33페이지)
- “클러스터 시간”(33페이지)
- “가용성이 높은 프레임워크”(33페이지)
- “글로벌 장치”(36페이지)
- “디스크 장치 그룹”(37페이지)
- “글로벌 이름 공간”(39페이지)
- “클러스터 파일 시스템”(41페이지)
- “정족수 및 정족수 장치”(43페이지)
- “블룸 관리자”(47페이지)
- “데이터 서비스”(48페이지)
- “새 데이터 서비스 개발”(56페이지)
- “자원, 자원 그룹 및 자원 유형”(58페이지)
- “PNM(Public Network Management) 및 NAFO(Network Adapter Failover)”(61페이지)
- “동적 재구성 지원”(62페이지)

클러스터 관리 및 응용프로그램 개발

이 정보는 주로 SunPlex API 및 SDK를 사용하는 시스템 관리자 및 응용프로그램 개발자를 위한 내용입니다. 클러스터 시스템 관리자는 클러스터 소프트웨어 설치, 구성 및

관리를 위한 배경 정보로 이 정보를 사용할 수 있습니다. 응용프로그램 개발자는 이 정보를 통해 작업할 클러스터의 환경을 이해할 수 있습니다.

다음은 클러스터 관리 개념을 클러스터 아키텍처에 매핑하는 방법을 높은 수준으로 설명하는 그림입니다.

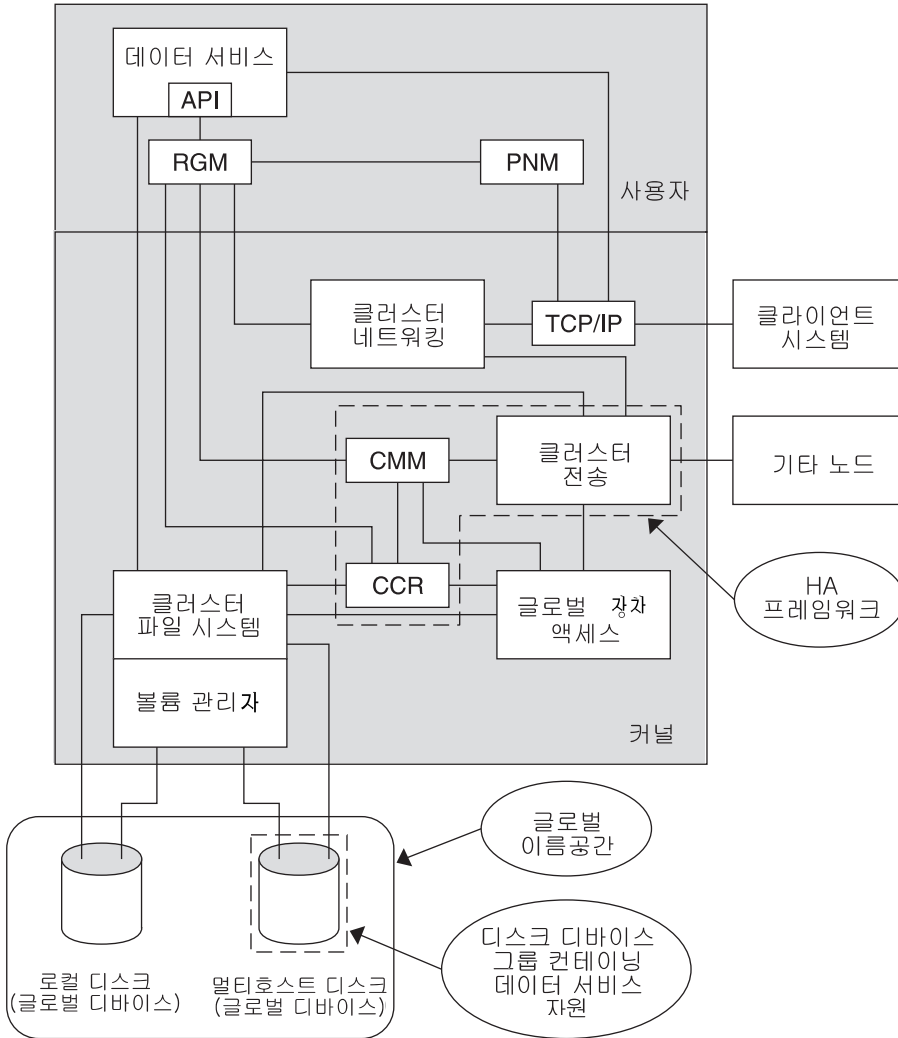


그림 3-1 Sun Cluster 소프트웨어 아키텍처

관리 인터페이스

여러 가지 사용자 인터페이스에서 **SunPlex** 시스템을 설치하고 구성하고 관리하는 방법을 선택할 수 있습니다. 문서화된 명령행 인터페이스를 통해 시스템 관리 작업을 수행할 수 있습니다. 명령행 인터페이스의 맨 위에는 선택한 구성 작업을 쉽게 할 수 있는 몇 가지 유틸리티가 있습니다. 또한 **SunPlex** 시스템에는 **Sun Management Center**의 일부로 실행되어 일부 클러스터 작업에 **GUI**를 제공하는 모듈이 있습니다. 관리 인터페이스에 대한 자세한 설명은 **Sun Cluster 3.0 12/01** 시스템 관리 안내서의 소개 단원을 참조하십시오.

클러스터 시간

클러스터에서는 모든 노드 사이의 시간이 동기화되어야 합니다. 클러스터 노드를 외부 시간 소스와 동기화할 것인지는 클러스터 작동에 중요하지 않습니다. **SunPlex** 시스템은 **NTP(Network Time Protocol)**를 사용하여 노드 사이의 시간을 동기화합니다.

일반적으로 시스템 시계에서 초의 끝자리 수 부분을 변경해도 문제가 발생하지는 않습니다. 그러나 작동하는 클러스터에서 `date(1)`, `rdate(1M)` 또는 `xntpdate(1M)` 명령을 실행하면(대화식으로 또는 `cron` 스크립트 내에서), 시스템 시계를 시간 소스와 동기화하기 위해 초의 끝자리 소수 부분보다 훨씬 큰 시간 값을 강제로 변경할 수 있습니다. 이러한 강제 변경으로 파일 수정 시간에 문제가 발생하거나 **NTP** 서비스에 혼란이 올 수 있습니다.

각 클러스터 노드에 **Solaris** 운영 환경을 설치할 때 노드에 대한 기본 시간 및 날짜 설정을 변경할 수 있습니다. 일반적으로 공장에서 설정된 기본값을 그대로 사용할 수 있습니다.

`scinstall(1M)` 명령을 사용하여 **Sun Cluster** 소프트웨어를 설치할 때 클러스터에 **NTP**를 구성하는 단계가 프로세스에 포함됩니다. **Sun Cluster** 소프트웨어에는 템플릿 파일 `ntp.cluster`가 있습니다(설치된 클러스터 노드에서 `/etc/inet/ntp.cluster` 참조). 이 파일은 노드 하나를 “기본” 노드로 사용하여 모든 클러스터 노드 사이에 피어 관계를 구성합니다. 노드는 독립된 호스트 이름으로 식별되고 클러스터 상호 연결 사이에 시간이 동기화됩니다. 클러스터에 **NTP**를 구성하는 방법은 **Sun Cluster 3.0 12/01** 소프트웨어 설치 안내서를 참조하십시오.

대신 클러스터 외부에 하나 이상의 **NTP** 서버를 설정하고 이 구성을 적용하여 `ntp.conf` 파일을 변경할 수도 있습니다.

정상적인 작동에서는 클러스터의 시간을 조정할 필요가 없습니다. 그러나 **Solaris** 운영 환경을 설치할 때 시간이 잘못 설정되어 변경하려는 경우에는 **Sun Cluster 3.0 12/01** 시스템 관리 안내서의 절차를 참조하십시오.

가용성이 높은 프레임워크

SunPlex 시스템은 네트워크 인터페이스, 응용프로그램, 파일 시스템 및 멀티호스트 디스크를 포함하여 사용자와 데이터 사이의 “경로”에 있는 모든 구성 요소의 가용성을

향상시킵니다. 일반적으로 시스템의 한 위치(소프트웨어 또는 하드웨어)에서 장애가 발생한 후에도 클러스터 구성 요소가 계속 작동하면 사용성이 높은 것입니다.

다음 표는 SunPlex 구성 요소에서 발생하는 하드웨어 및 소프트웨어 장애와 가용성이 높은 프레임워크에 구축된 복구 방법입니다.

표3-1 SunPlex 장애 탐지 및 복구 레벨

장애가 발생한 클러스터 구성 요소	소프트웨어 복구	하드웨어 복구
데이터 서비스	HA API, HA 프레임워크	없음
공용 네트워크 어댑터	NAFO(네트워크 어댑터 페일오버)	여러 개의 공용 네트워크 어댑터 카드
클러스터 파일 시스템	1차 및 2차 복제	멀티호스트 디스크
이중화된 멀티호스트 디스크	볼륨 관리자(Solstice DiskSuite 및 VERITAS Volume Manager)	하드웨어 RAID-5(예: Sun StorEdge A3x00)
글로벌 장치	1차 및 2차 복제	장치, 클러스터 전송 연결 장치에 대한 복수 경로
독립 네트워크	HA 전송 소프트웨어	여러 개의 개인용 하드웨어 독립 네트워크
노드	CMM, 페일패스트 드라이버	복수 노드

Sun Cluster 소프트웨어의 프레임워크는 가용성이 높기 때문에 노드 장애를 즉시 탐지하고 프레임워크 자원을 제공하는 새로운 서버를 남은 노드에 구성합니다. 모든 프레임워크를 동시에 사용하지 못하는 경우는 없습니다. 복구 중에도 충돌이 발생한 노드의 영향을 받지 않는 프레임워크 자원을 모두 사용할 수 있습니다. 또한 장애가 발생한 노드의 프레임워크 자원은 복구되는 대로 바로 사용할 수 있습니다. 다른 프레임워크 자원이 모두 복구될 때까지 복구된 프레임워크 자원이 대기하지 않아도 됩니다.

가용성이 높은 프레임워크 자원은 대부분 자원을 사용하는 응용프로그램(데이터 서비스)에 투명하게 복구됩니다. 장애 발생 전후에 프레임워크 자원 액세스에 대한 설정은 완전하게 보존됩니다. 프레임워크 자원 서버가 다른 노드로 이동된 것만 응용프로그램이 인식할 수 없습니다. 다른 노드에 디스크에 대한 다른 하드웨어 경로가 있기만 하면, 한 노드에 장애가 발생할 경우에 노드에 연결된 파일, 장치 및 디스크 볼륨을 사용하여 다른 노드의 프로그램에 완전히 투명하게 인식됩니다. 예를 들어, 여러 노드로 연결하는 포트가 있는 멀티호스트 디스크를 사용할 수 있습니다.

클러스터 멤버십 모니터

CMM(Cluster Membership Monitor)은 클러스터 구성원마다 하나씩 분산된 에이전트 세트입니다. 에이전트는 다음 작업을 수행하기 위해 클러스터 상호 연결을 통해 메시지를 교환합니다.

- 모든 노드(정족수)에 일관된 멤버십 화면 표시
- 멤버십이 변경될 경우에 등록된 콜백을 사용하여 동기화된 재구성 실행
- 클러스터 분할 처리(브레인 분할, 암네시아)
- 모든 클러스터 구성원 사이를 완전하게 연결

이전의 **Sun Cluster** 소프트웨어 릴리스와 달리 **CMM**은 완전히 커널에서 실행됩니다.

클러스터 멤버십

CMM의 기본 기능은 언제든지 전체 클러스터에서 클러스터에 포함된 노드 세트를 동일하게 인식하도록 하는 것입니다. 이러한 제약 조건을 클러스터 멤버십이라고 합니다.

CMM은 클러스터 구성원을 인식하여 데이터 무결성을 구현하기 위해 다음과 같은 기능을 수행합니다.

- 클러스터에서 노드 결합이나 제거와 같은 클러스터 멤버십 변경을 보고합니다.
- “잘못된” 노드를 클러스터에서 제거합니다.
- “잘못된” 노드가 복구될 때까지 클러스터에 연결되지 않도록 합니다.
- 클러스터가 하위 노드 세트로 분할되지 않도록 합니다.

클러스터가 여러 개의 별도 클러스터로 분할되지 않도록 보호하는 방법은 “정족수 및 정족수 장치”(43페이지)를 참조하십시오.

클러스터 멤버십 모니터 재구성

데이터가 손상되지 않도록 안전하게 보존하려면 모든 노드에서 클러스터 멤버십을 동일하게 인식해야 합니다. 장애가 발생할 경우에 필요하면 **CMM**이 클러스터 서비스(응용프로그램)의 클러스터 재구성을 조정합니다.

CMM은 클러스터 전송 계층으로부터 다른 노드에 대한 연결 정보를 받습니다. 재구성할 때 **CMM**은 클러스터 상호 연결을 사용하여 상태 정보를 교환합니다.

클러스터 멤버십 변경이 발견되면 **CMM**이 동기화된 클러스터 구성을 수행합니다. 이러한 경우에 새로운 클러스터 멤버십에 따라 클러스터 자원이 다시 분산될 수 있습니다.

파일패스트 메커니즘

CMM이 노드에서 치명적인 문제를 확인하면 클러스터 프레임워크를 호출하여 강제로 해당 노드를 중지시키고 클러스터 멤버십에서 제거합니다. 이러한 작업을 수행하

는 메커니즘을 페일패스트라고 합니다. 페일패스트 메커니즘은 다음 두 가지 방법으로 노드를 중지시킵니다.

- 노드가 클러스터에서 제거되어 노드 수가 부족한 상태로 새로운 클러스터를 시작하면 공유 디스크에 대한 액세스가 “금지”됩니다. 페일패스트 사용에 대한 자세한 내용은 (46페이지)“장애 방지”를 참조하십시오.
- 하나 이상의 클러스터 관련 데몬이 중단되면(clexecd, rpc.pmfd, rgmd 또는 rpc.ed) CMM에서 장애가 확인되어 노드가 종료됩니다. 클러스터 데몬이 중단되어 노드가 종료되면 해당 노드의 콘솔에 다음과 같은 메시지가 나타납니다.

```
panic[cpu0]/thread=40e60: Failfast: Aborting because "pmfd" died 35 seconds ago.  
409b8 cl_runtime: __OFZsc_syslog_msg_log_no_argsPviTCPcTB+48 (70f900, 30, 70df54, 407acc, 0)  
%10-7: 1006c80 000000a 000000a 10093bc 406d3c80 7110340 0000000 4001 fbf0
```

노드가 종료된 후에 다시 부트되어 클러스터에 다시 연결될 수도 있고 OpenBoot PROM (OBP) 프롬프트 상태로 있을 수도 있습니다. 취하는 조치는 OBP의 auto-boot? 매개 변수 설정에 따라 결정됩니다.

CCR(Cluster Configuration Repository)

CCR(Cluster Configuration Repository)은 클러스터 구성 및 상태에 대한 정보를 저장하기 위해 전체 클러스터에서 사용하는 독립 데이터베이스입니다. CCR은 분산 데이터베이스입니다. 각 노드에 전체 데이터베이스 사본이 유지됩니다. CCR을 사용하면 모든 노드에서 일관된 화면으로 “전체” 클러스터를 볼 수 있습니다. 데이터가 손상되지 않도록 하려면 각 노드에서 클러스터 자원의 현재 상태를 알아야 합니다.

CCR은 갱신을 위해 2단계 실행 알고리즘을 사용합니다. 즉, 모든 클러스터 구성원에 대해 성공적으로 갱신이 완료되어야 하고, 그렇지 않은 경우에는 롤백됩니다. CCR은 클러스터 상호 연결을 사용하여 분산된 갱신 사항을 적용합니다.



주의: CCR은 텍스트 파일로 구성되어 있지만 직접 CCR 파일을 편집하면 안됩니다. 각 파일에는 일관성이 유지되도록 체크섬 레코드가 포함됩니다. 직접 CCR 파일을 갱신하면 노드나 전체 클러스터의 작동이 중단될 수 있습니다.

CCR은 정족수가 구성된 경우에만 클러스터가 실행되도록 하기 위해 CMM을 사용합니다. CCR은 전체 클러스터에서 데이터 일관성을 확인하고, 필요할 경우에 복구 작업을 수행하고, 데이터 갱신을 지원하는 기능을 합니다.

글로벌 장치

SunPlex 시스템은 장치가 물리적으로 연결된 위치와 관계없이 어떤 노드에서나 클러스터의 모든 장치에 액세스할 수 있도록 글로벌 장치를 사용하여 전체 클러스터에 늘

은 가용성을 제공합니다. 일반적으로 글로벌 장치에 대한 액세스를 제공하는 동안 노드에 장애가 발생하면 **Sun Cluster** 소프트웨어가 자동으로 장치에 대한 다른 경로를 찾아서 해당 경로로 액세스를 전환합니다. **SunPlex** 글로벌 장치에는 디스크, **CD-ROM** 및 테이프가 포함됩니다. 그러나 멀티포트 글로벌 장치로는 디스크만 사용할 수 있습니다. 즉, **CD-ROM** 및 테이프 장치는 현재 가용성이 높은 장치가 아닙니다. 각 서버의 로컬 디스크도 멀티포트 상태가 아니므로 가용성이 높은 장치가 아닙니다.

클러스터는 자동으로 클러스터 내의 각 디스크, **CD-ROM** 및 테이프 장치에 고유한 **ID**를 할당합니다. 이러한 할당으로 인해 클러스터의 임의의 노드에서 각 장치에 일관되게 액세스할 수 있습니다. 글로벌 장치 이름 공간은 `/dev/global` 디렉토리에 저장됩니다. 자세한 내용은 “글로벌 이름 공간”(39페이지)을 참조하십시오.

멀티포트 글로벌 장치는 장치에 대한 경로를 하나 이상 제공합니다. 멀티호스트 디스크의 경우에는 디스크가 하나 이상의 노드에 의해 호스트되는 디스크 장치 그룹에 포함되기 때문에 멀티호스트 디스크의 가용성이 높아집니다.

DID(장치 ID)

Sun Cluster 소프트웨어는 **DID** 의사 드라이버라는 구성을 통해 글로벌 장치를 관리합니다. 이 드라이버는 멀티호스트 디스크, 테이프 드라이브 및 **CD-ROM**과 같은 클러스터의 모든 장치에 자동으로 고유 **ID**를 할당하는 데 사용됩니다.

DID(장치 ID) 의사 드라이버는 클러스터의 글로벌 장치 액세스 기능을 위해 중요한 부분입니다. **DID** 드라이버는 클러스터의 모든 노드를 검색하고 고유한 디스크 장치 목록을 만들어, 클러스터의 모든 노드에서 일관된 고유한 기본 번호와 하위 번호를 각 장치에 할당합니다. 글로벌 장치에 대한 액세스는 디스크에 `c0t0d0`을 사용했던 이전의 **Solaris** 장치 **ID** 대신 **DID** 드라이버에 의해 할당되는 고유 장치 **ID**를 사용하여 수행됩니다.

이러한 방법을 사용하면 디스크 장치를 이용하는 응용프로그램(원시 장치를 사용하는 볼륨 관리자나 응용프로그램)이 일관된 경로를 사용하여 장치에 액세스할 수 있습니다. 특히 멀티호스트 디스크의 경우에 이러한 일관성이 중요합니다. 그 이유는 각 장치의 로컬 기본 번호와 하위 번호가 노드마다 다를 수 있고, 이런 경우에 **Solaris** 장치 이름 지정 규칙도 달라질 수 있기 때문입니다. 예를 들어, **node1**에는 멀티호스트 디스크가 `c1t2d0`으로 표시되고 **node2**에는 동일한 디스크가 완전히 다르게 `c3t2d0`으로 표시될 수 있습니다. **DID** 드라이버는 노드에서 대신 사용하도록 **d10**과 같은 글로벌 이름을 할당하여 멀티호스트 디스크에 대한 일관된 매핑을 각 노드에 제공합니다.

사용자는 `scdidadm(1M)` 및 `scgdevs(1M)` 명령을 통해 장치 **ID**를 갱신하고 관리합니다. 자세한 내용은 각 설명서 페이지를 참조하십시오.

디스크 장치 그룹

SunPlex 시스템에서는 모든 멀티호스트 디스크가 **Sun Cluster** 소프트웨어에 의해 제어되어야 합니다. 먼저 멀티호스트 디스크에 볼륨 관리자 디스크 그룹(**Solstice DiskSuite** 디스크 세트 또는 **VERITAS Volume Manager** 디스크 그룹)을 만듭니다. 그런 다음에 볼륨 관리자 그룹을 디스크 장치 그룹으로 등록합니다. 디스크 장치 그룹은 글로벌 장

치의 유형입니다. **Sun Cluster** 소프트웨어는 자동으로 클러스터의 각 디스크와 테이프 장치에 대한 원시 장치 그룹을 만듭니다. 그러나 사용자가 클러스터 장치 그룹을 글로벌 장치로 액세스할 때까지 이 클러스터 장치 그룹이 오프라인 상태를 유지합니다.

등록을 하면 어느 노드가 어느 볼륨 관리자 디스크 그룹에 대한 경로를 갖는지에 대한 정보가 **SunPlex** 시스템에 제공됩니다. 그러면 클러스터 전체에서 볼륨 관리자 디스크 그룹에 액세스할 수 있습니다. 둘 이상의 노드가 디스크 장치 그룹에 쓸 수 있으면(마스터) 해당 디스크 장치 그룹에 저장된 데이터의 가용성이 높아집니다. 가용성이 높은 디스크 장치 그룹을 사용하면 클러스터 파일 시스템을 하우징할 수 있습니다.

주: 디스크 장치 그룹은 자원 그룹의 영향을 받지 않습니다. 따라서 노드 하나는 데이터 서비스에 의해 액세스되는 디스크 그룹을 마스터하고 다른 노드는 데이터 서비스 프로세스 그룹을 나타내는 자원 그룹을 마스터할 수 있습니다. 그러나 가장 좋은 방법은 특정 응용프로그램의 데이터를 저장하는 디스크 장치 그룹과 응용프로그램 자원(응용프로그램 데몬)이 포함된 자원 그룹을 동일한 노드에 유지하는 것입니다. 디스크 장치 그룹과 자원 그룹 사이의 관계에 대한 자세한 내용은 **Sun Cluster 3.0 12/01 Data Services Installation and Configuration Guide**의 개요 단원을 참조하십시오.

디스크 장치 그룹을 사용하면 볼륨 관리자 디스크 그룹이 기반 디스크에 대한 복수 경로를 제공하기 때문에 볼륨 관리자 디스크 그룹이 “글로벌”이 됩니다. 물리적으로 멀티호스트 디스크에 연결된 각 클러스터 노드는 디스크 장치 그룹에 대한 경로를 제공합니다.

디스크 장치 그룹 페일오버

디스크 인클로저는 둘 이상의 노드에 연결되기 때문에 현재 장치 그룹을 마스터하는 노드에 장애가 발생해도 다른 경로를 통해 해당 인클로저의 모든 디스크 장치 그룹에 액세스할 수 있습니다. 장치 그룹을 마스터하는 노드에 장애가 발생해도 복구 및 일관성 검사를 수행하는 데 시간이 소요되는 것을 제외하고는 장치 그룹에 대한 액세스에 영향을 주지 않습니다. 이 시간 동안에는 다시 장치 그룹을 사용할 수 있게 될 때까지 모든 요청이 응용프로그램에 투명하게 차단됩니다.

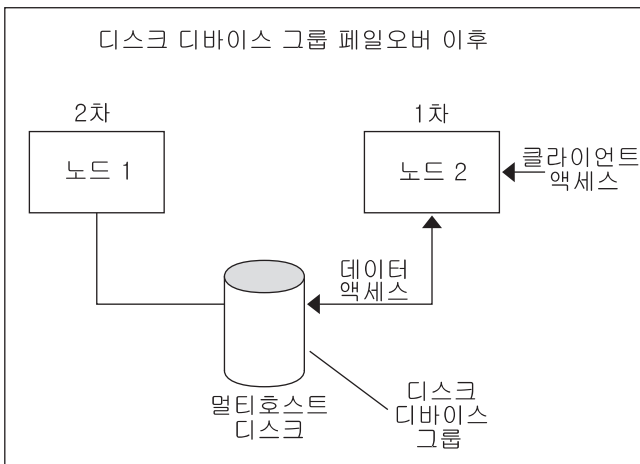
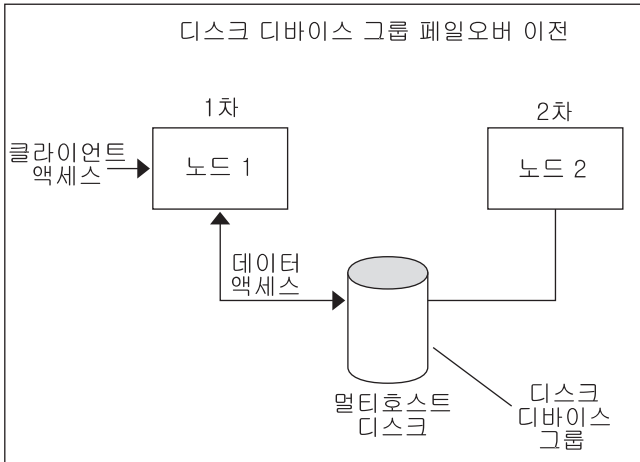


그림 3-2 디스크 장치 그룹 페일오버

글로벌 이름 공간

글로벌 장치를 활성화하는 **Sun Cluster** 소프트웨어 메커니즘이 글로벌 이름 공간입니다. 글로벌 이름 공간에는 볼륨 관리자 이름 공간뿐 아니라 `/dev/global/` 계층도 포함됩니다. 글로벌 이름 공간에는 멀티호스트 디스크와 로컬 디스크(CD-ROM 및 테이프와 같은 다른 클러스터 장치 포함)에 대한 정보가 모두 반영되어 멀티호스트 디스크에 대한 여러 개의 페일오버 경로를 제공합니다. 물리적으로 멀티호스트 디스크에 연결된 각 노드는 클러스터의 노드에 대한 기억 장치 경로를 제공합니다.

일반적으로 **Solstice DiskSuite**의 경우에는 볼륨 관리자 이름 공간이 `/dev/md/diskset/dsk`(및 `rdsk`) 디렉토리에 있고, **VxVM**의 경우에는 `/dev/vx/dsk/disk-group` 및 `/dev/vx/rdsk/disk-group` 디렉토리에 있습니다. 이 이름 공간은 각각 클러스터를

통해 가져온 VxVM 디스크 그룹을 위한 디렉토리와 Solstice DiskSuite 디스크 세트를 위한 디렉토리로 구성됩니다. 이 디렉토리는 해당 디스크 세트나 디스크 그룹에 있는 각 메타 장치나 볼륨에 대한 장치 노드를 하우징합니다.

SunPlex 시스템에서는 로컬 볼륨 관리자 이름 공간의 각 장치 노드가 /global/.devices/node@nodeID 파일 시스템의 장치 노드에 대한 기호 링크로 교체됩니다. 여기서 nodeID는 클러스터의 노드를 나타내는 정수입니다. Sun Cluster 소프트웨어는 표준 위치에도 계속 볼륨 관리자 장치를 기호 링크로 제공합니다. 글로벌 이름 공간과 표준 볼륨 관리자 이름 공간 둘 다 임의의 클러스터 노드에서 사용할 수 있습니다.

글로벌 이름 공간의 장점은 다음과 같습니다.

- 장치 관리 모델을 거의 변경하지 않고 각 노드를 독립적으로 유지합니다..
- 장치가 선택적으로 글로벌이 될 수 있습니다.
- 다른 회사의 링크 생성기를 계속 사용할 수 있습니다.
- 지정된 로컬 장치 이름에 대해 간단한 매핑으로 글로벌 이름을 만들 수 있습니다.

로컬 및 글로벌 이름 공간의 예

다음 표는 멀티호스트 디스크 c0t0d0s0에 대한 글로벌 이름 공간과 로컬 이름 공간 사이의 매핑입니다.

표3-2 로컬 및 글로벌 이름 공간 매핑

구성요소/경로	로컬 노드 이름 공간	글로벌 이름 공간
Solaris 논리 이름	/dev/dsk/c0t0d0s0	/global/.devices/node@ID/dev/dsk/c0t0d0s0
DID 이름	/dev/did/dsk/d0s0	/global/.devices/node@ID/dev/did/dsk/d0s0
Solstice DiskSuite	/dev/md/diskset/dsk/d0	/global/.devices/node@ID/dev/md/diskset/dsk/d0
VERITAS Volume Manager	/dev/vx/dsk/disk-group/v0	/global/.devices/node@ID/dev/vx/dsk/disk-group/v0

글로벌 이름 공간은 설치할 때 자동으로 만들어지고 재구성 후에 다시 부트할 때마다 갱신됩니다. scgdevs (1M) 명령을 실행하여 글로벌 이름 공간을 만들 수도 있습니다.

클러스터 파일 시스템

클러스터 파일 시스템은 한 노드의 커널과 디스크에 물리적으로 연결되어 있는 노드에서 실행되는 볼륨 관리자 및 기반 파일 시스템 사이의 프록시입니다.

클러스터 파일 시스템은 하나 이상의 노드에 물리적으로 연결되어 있는 글로벌 장치(디스크, 테이프, CD-ROM)에 의존합니다. 글로벌 장치는 노드가 기억 장치에 물리적으로 연결되었는지 여부에 관계없이 동일한 파일 이름(예: /dev/global/)을 통해 클러스터의 모든 노드에서 액세스할 수 있습니다. 정규 장치와 동일한 방법으로 글로벌 장치를 사용할 수 있습니다. 즉, `newfs` 또는 `mkfs` 명령을 사용하여 글로벌 장치에 파일 시스템을 만들 수 있습니다.

`mount -g` 명령을 사용하여 글로벌로 또는 `mount` 명령을 사용하여 로컬로 글로벌 장치에 파일 시스템을 마운트할 수 있습니다.

클러스터의 모든 노드에서 동일한 파일 이름(예: /global/foo)을 사용하여 클러스터 파일 시스템의 파일에 액세스할 수 있습니다.

클러스터 파일 시스템은 모든 클러스터 구성원에 마운트됩니다. 클러스터 구성원의 하위 세트에는 클러스터 파일 시스템을 마운트할 수 없습니다.

클러스터 파일 시스템은 별도로 구분되는 파일 시스템 유형이 아닙니다. 즉, 클라이언트에는 하부 파일 시스템(예: UFS)이 표시됩니다.

클러스터 파일 시스템 사용

SunPlex 시스템에서는 모든 멀티호스트 디스크가 디스크 장치 그룹에 포함됩니다. Solstice DiskSuite 디스크 세트, VxVM 디스크 그룹 또는 소프트웨어 기반의 볼륨 관리자에 의해 제어되지 않는 개별 디스크가 디스크 장치 그룹이 될 수 있습니다.

클러스터 파일 시스템의 가용성을 높이려면 하부 디스크 기억 장치를 둘 이상의 노드에 연결해야 합니다. 따라서 클러스터 파일 시스템에 만든 로컬 파일 시스템(노드의 로컬 디스크에 저장된 파일 시스템)은 가용성이 높지 않습니다.

일반적인 파일 시스템과 같이 두 가지 방법으로 클러스터 파일 시스템을 마운트할 수 있습니다.

- 직접—명령행에서 `mount` 명령에 `-g` 또는 `-o global` 마운트 옵션을 사용하여 클러스터 파일 시스템을 마운트합니다. 예를 들어 다음과 같은 명령을 실행합니다.

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

- 자동—부트할 때 클러스터 파일 시스템을 마운트하도록 /etc/vfstab 파일에 global 마운트 옵션이 있는 항목을 만듭니다. 그런 다음 모든 노드의 /global 디렉토리 아래에 마운트 포인트를 만듭니다. /global 디렉토리는 권장 위치이며 반드시 지정해야 하는 것은 아닙니다. 다음은 클러스터 파일 시스템을 마운트하기 위해 /etc/vfstab 파일에 포함되는 행의 예입니다.

```
/dev/md/oracle/dsk/d1 /dev/md/oracle/rdsk/d1 /global/oracle/data  
ufs 2 yes global,logging
```

주: Sun Cluster 소프트웨어에서 반드시 클러스터 파일 시스템에 이름 지정 정책을 사용해야 하는 것은 아니지만, `/global/disk-device-group`과 같이 동일한 디렉토리 아래에 모든 클러스터 파일 시스템에 대한 마운트 포인트를 만들면 쉽게 관리할 수 있습니다. 자세한 내용은 *Sun Cluster 3.0 12/01* 소프트웨어 설치 안내서 및 *Sun Cluster 3.0 12/01* 시스템 관리 안내서를 참조하십시오.

클러스터 파일 시스템 기능

클러스터 파일 시스템에는 다음과 같은 기능이 있습니다.

- 파일 액세스 위치가 투명합니다. 프로세스가 위치에 관계없이 시스템에 있는 파일을 열 수 있으므로 모든 노드의 프로세스가 동일한 경로 이름을 사용하여 파일을 찾을 수 있습니다.

주: 클러스터 파일 시스템이 파일을 읽을 때 해당 파일에 대한 액세스 시간을 갱신하지는 않습니다.

- 동기 프로토콜을 사용하여 파일이 동시에 여러 노드로부터 액세스될 경우에도 UNIX 파일 액세스 시맨틱을 보존합니다.
- 효율적으로 파일 데이터를 이동하기 위하여 **zero-copy** 벌크 I/O 이동과 함께 확장 캐싱이 사용됩니다.
- 클러스터 파일 시스템은 `fcntl(2)` 인터페이스를 사용하여 가용성이 높은 권장 파일 잠금 기능을 제공합니다. 여러 클러스터 노드에서 실행되는 응용프로그램은 클러스터 파일 시스템의 파일에 대한 권장 파일 잠금 기능을 사용하여 데이터 액세스를 동기화할 수 있습니다. 잠금이 유지되는 동안 클러스터에서 제거되는 노드와 장애가 발생하는 응용프로그램에서는 즉시 파일 잠금이 복원됩니다.
- 장애가 발생할 경우에도 데이터에 대한 액세스는 계속할 수 있습니다. 디스크에 대한 경로가 계속 작동하면 응용프로그램이 장애의 영향을 받지 않습니다. 이 기능은 원시 디스크 액세스와 모든 파일 시스템 작동에 적용됩니다.
- 클러스터 파일 시스템은 하부 파일 시스템 및 볼륨 관리 소프트웨어와 독립적으로 작동합니다. 클러스터 파일 시스템은 지원되는 디스크의 파일 시스템을 모두 글로벌로 만듭니다.

Syncdir 마운트 옵션

`syncdir` 마운트 옵션은 UFS를 기반 파일 시스템으로 사용하는 클러스터 파일 시스템에 사용할 수 있습니다. 그러나 `syncdir`을 지정하지 않으면 성능이 크게 향상됩니다. `syncdir` 옵션을 지정하면 쓰기 작업이 POSIX와 호환됩니다. 이 옵션을 지정하지 않으면 NFS 파일 시스템과 동일하게 작동합니다. 예를 들어, `syncdir`이 없으면 파일을 닫을 때까지 공간 부족 상태를 발견하지 못하는 경우가 있습니다. `syncdir`(및 POSIX 동작) 옵션이 있으면 쓰기 작업 중에 공간 부족 상태가 발견되었을 것입니다.

syncdir 옵션을 지정하지 않아서 문제가 발생하는 경우가 있지만 지정하지 않으면 성능이 향상되는 이점이 있기 때문에 지정하지 않는 것이 좋습니다.

글로벌 장치와 클러스터 파일 시스템에 대한 FAQ는 “파일 시스템 FAQ”(68페이지)를 참조하십시오.

정족수 및 정족수 장치

클러스터 노드는 데이터와 자원을 공유하기 때문에 클러스터가 동시에 작동하는 별도 파티션으로 분리되지 않습니다. CMM을 사용하면 클러스터 상호 연결이 분할된 경우에도 하나 이상의 클러스터가 항상 작동합니다.

클러스터 파티션에서는 브레인 분할 및 앰네시아 두 가지 유형의 문제가 발생합니다. 브레인 분할은 노드 사이의 클러스터 상호 연결이 끊어져서 클러스터가 하위 클러스터로 분할될 경우에 발생합니다. 그러면 각 하위 클러스터가 다른 파티션을 인식하지 못하고 자체 파티션만 있는 것으로 인식합니다. 이 문제는 클러스터 노드 사이의 통신 문제 때문에 발생합니다. 앰네시아는 클러스터가 종료된 후에 클러스터가 종료된 시간보다 오래된 클러스터 데이터를 사용하여 다시 시작할 경우에 발생합니다. 디스크에 여러 가지 프레임워크 데이터 버전이 저장되어 있을 경우에 최신 버전을 사용할 수 없으면 새로 구성된 클러스터가 시작될 때 이 문제가 발생할 수 있습니다.

각 노드에 한 표씩 부여하고 대부분의 표가 작동 클러스터에 참여하도록 지시하면 브레인 분할 및 앰네시아 문제를 방지할 수 있습니다. 다수 표가 있는 파티션은 정족수가 충족되기 때문에 작동할 수 있습니다. 클러스터에 노드가 두 개 이상이면 이러한 다수 표 메커니즘을 사용할 수 있습니다. 노드가 두 개인 클러스터에서는 두 개가 다수입니다. 이러한 클러스터가 분할되면 각 파티션이 정족수를 얻기 위해 외부 표가 필요합니다. 필요한 외부 표는 정족수 장치에서 제공합니다. 두 노드 사이에 공유하는 디스크를 정족수 장치로 사용할 수 있습니다. 정족수 장치로 사용하는 디스크에도 사용자 데이터를 저장할 수 있습니다.

표 3-3에서는 Sun Cluster 소프트웨어가 정족수를 사용하여 브레인 분할 및 앰네시아 문제를 해결하는 방법을 설명합니다.

표 3-3 클러스터 정족수와 브레인 분할 및 앰네시아 문제

파티션 유형	정족수 해결 방법
브레인 분할	다수 표가 있는 파티션(하위 클러스터)만 클러스터로 실행되도록 합니다. 이 경우에 다수 표를 갖는 파티션은 하나만 있을 수 있습니다.
앰네시아	클러스터가 부트될 때 최근 클러스터 멤버십의 구성원이었던 노드를 하나 이상 포함하도록 합니다. 그러면 최신 구성 데이터를 사용할 수 있습니다.

정족수 알고리즘은 동적으로 작동합니다. 즉, 클러스터 이벤트가 계산을 트리거하면 클러스터 수명 동안 계산 결과가 변경될 수 있습니다.

정족수 표 계산

클러스터 노드와 정족수 장치가 모두 투표하여 정족수를 채웁니다. 기본적으로 클러스터 노드가 부트되어 클러스터 구성원이 되면 클러스터에 정족수 투표 수 하나가 증가합니다. 노드를 설치하는 중이거나 관리자가 노드를 유지 보수 상태로 전환한 경우에는 노드의 투표 수가 0이 될 수도 있습니다.

정족수 장치의 정족수 투표 수는 장치에 연결된 노드 수에 따라 결정됩니다. 정족수 장치가 설정되면 최대 투표 수 $N-1$ 개를 받습니다. 여기서 N 은 정족수 장치에 대한 포트가 있고 투표 수가 0이 아닌 노드 수입니다. 예를 들어, 투표 수가 0이 아닌 두 개의 노드에 연결된 정족수 장치의 정족수 투표 수는 $1(2-1)$ 이 됩니다.

클러스터를 설치할 때 정족수 장치를 구성할 수도 있고 나중에 *Sun Cluster 3.0 12/01* 시스템 관리 안내서에서 설명하는 절차를 사용하여 정족수 장치를 구성할 수도 있습니다.

주: 현재 정족수 장치가 연결된 노드 중에서 하나 이상의 노드가 클러스터 구성원인 경우에만 정족수 장치가 투표 수에 포함됩니다. 또한, 클러스터 부트 중에는 현재 정족수 장치가 연결된 노드 중 하나 이상이 부트되고 있어야 하고 최근 부트된 클러스터가 종료될 때 노드가 이 클러스터의 구성원이 있어야만 정족수 장치가 투표 수에 포함됩니다.

정족수 구성

정족수 구성은 클러스터의 노드 수에 따라 달라집니다.

- 2-노드 클러스터 - 2-노드 클러스터를 구성하려면 두 개의 정족수 표가 필요합니다. 이러한 두 개의 투표는 두 개의 클러스터 노드 또는 하나의 노드와 정족수 장치로부터 제공됩니다. 그러나 한 노드에 장애가 발생할 경우에 다른 한 노드에서 계속 작동하도록 하려면 2-노드 클러스터에 정족수 장치가 구성되어 있어야 합니다.
- 노드가 세 개 이상인 클러스터 - 디스크 기억 장치 인클로저에 대한 액세스를 공유하는 모든 노드 쌍 사이에 정족수 장치를 지정해야 합니다. 예를 들어, 그림 3-3에 있는 클러스터와 비슷한 3-노드 클러스터가 있는 것으로 가정합니다. 이 그림에서 **nodeA**와 **nodeB**는 동일한 디스크 인클로저에 대한 액세스를 공유하고, **nodeB**와 **nodeC**는 다른 디스크 인클로저에 대한 액세스를 공유합니다. 그러면 정족수 투표가 모두 5개가 됩니다. 세 개는 노드에서 제공되고 두 개는 노드 사이에 공유하는 정족수 장치에서 제공됩니다. 클러스터를 구성하려면 대부분의 정족수 표가 필요합니다.

모든 노드 쌍 사이에 디스크 기억 장치 인클로저에 대한 액세스를 공유하는 정족수 장치를 지정할 필요도 없고 *Sun Cluster* 소프트웨어에 의해 강제로 지정되지도 않습니다. 그러나 $N+1$ 구성이 2-노드 클러스터로 축소된 후에 두 디스크 인클로저에 대한 액세스에도 실패하는 경우에는 필요한 정족수 투표를 제공할 수 있습니다. 모든 쌍 사이에 정족수 장치를 구성했으면 나머지 노드가 계속 클러스터로 작동할 수 있습니다.

이러한 구성의 예는 그림 3-3을 참조하십시오.

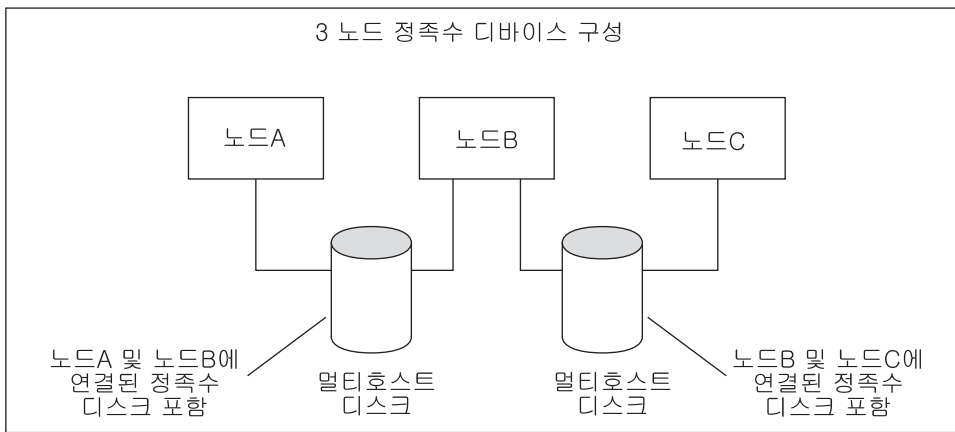
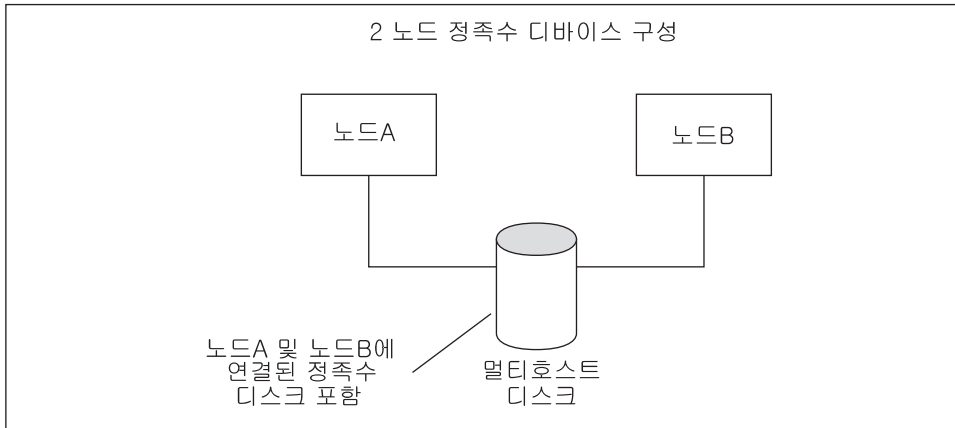


그림 3-3 정족수 장치 구성의 예

정족수 지침

정족수 장치를 설정할 때 다음 지침을 사용하십시오.

- 동일한 공유 디스크 기억 장치 인클로저에 연결된 모든 노드 사이에 정족수 장치를 설정하십시오. 노드에 장애가 발생할 경우에 다른 노드가 정족수를 유지하고 공유 인클로저의 디스크 장치 그룹을 마스터할 수 있도록 공유 인클로저 내의 디스크 하나를 정족수 장치로 추가하십시오.
- 두 개 이상의 노드에 정족수 장치를 연결해야 합니다.
- 정족수 장치는 이중 포트 정족수 장치로 사용되는 **SCSI-2** 또는 **SCSI-3** 디스크가 될 수 있습니다. 세 개 이상의 노드에 연결된 디스크는 정족수 장치로 사용되는지에 관계없이 **SCSI-3 PGR(Persistent Group Reservation)**을 지원해야 합니다. 자세한 내용은 *Sun Cluster 3.0 12/01* 소프트웨어 설치 안내서에서 계획 단원을 참조하십시오.
- 사용자 데이터가 포함된 디스크를 정족수 장치로 사용할 수 있습니다.

조언: 각 정족수 장치에 장애가 발생하지 않도록 보호하려면 노드 세트 사이에 두 개 이상의 정족수 장치를 구성하십시오. 다른 인클로저의 디스크를 사용하고, 각 노드 사이에 홀수 개의 정족수 장치를 구성하십시오.

장애 방지

클러스터에서 가장 중요한 문제는 클러스터를 분할하는 장애(브레인 분할)입니다. 이러한 문제가 발생하면, 모든 노드가 통신을 할 수 있는 것이 아니기 때문에 개별 노드나 노드 하위 세트가 개별 클러스터나 하위 세트 클러스터를 만들 수 있습니다. 각 하위 세트나 파티션이 멀티호스트 디스크에 대하여 단 하나의 액세스 및 소유권을 갖고 있는 것으로 인식할 수도 있습니다. 여러 노드가 디스크에 쓰려고 시도하면 데이터가 손상될 수 있습니다.

장애 방지는 물리적으로 디스크에 대한 액세스를 금지하여 멀티호스트 디스크에 대한 노드 액세스를 제한합니다. 노드에 장애가 발생하거나 분할되어 노드가 클러스터에서 제외되면 장애 방지 기능에 의해 해당 노드가 더 이상 디스크에 액세스할 수 없게 됩니다. 현재 구성원 노드만 디스크에 액세스할 수 있기 때문에 데이터 무결성이 유지됩니다.

디스크 장치 서비스는 멀티호스트 디스크를 사용하는 서비스에 페일오버 기능을 제공합니다. 현재 디스크 장치 그룹의 1차(소유자) 노드로 작동하는 클러스터 구성원에 장애가 발생하거나 연결이 안되면 잠깐 중단된 후에 디스크 장치 그룹에 대한 액세스를 계속할 수 있도록 새로운 1차 노드가 선택됩니다. 이 프로세스에서 새로운 1차 노드가 시작되려면 먼저 이전 1차 노드가 장치에 대한 액세스를 중지해야 합니다. 그러나 구성원이 클러스터에서 제외되어 연결되지 않으면 클러스터에서 장치에 대한 1차 소유권을 해제하도록 해당 노드에 알릴 수 없습니다. 따라서 작동하는 구성원이 장애가 발생한 구성원으로부터 제어 권한을 받아 글로벌 장치에 액세스할 수 있도록 하는 수단이 필요합니다.

SunPlex 시스템은 **SCSI** 디스크 예약 기능을 사용하여 장애 방지를 구현합니다. **SCSI** 예약 기능을 사용하면 장애가 발생한 노드가 멀티호스트 디스크로부터 “금지”되어 디스크에 액세스할 수 없습니다.

SCSI-2 디스크 예약은 디스크에 연결된 모든 노드에 액세스를 제공하거나(예약이 없을 경우) 단일 노드(예약된 노드)에만 액세스를 제공하는 방식의 예약을 지원합니다.

클러스터 구성원이 다른 노드가 더 이상 클러스터 상호 연결을 통해 통신할 수 없다는 것을 발견하면 다른 노드가 공유 디스크에 액세스하지 못하도록 장애 방지 절차를 시작합니다. 이러한 장애 방지가 발생하면 일반적으로 금지된 노드가 중단되고 콘솔에 “예약 충돌” 메시지가 표시됩니다.

예약 충돌이 발생하는 것은 노드가 더 이상 클러스터 구성원이 아니라는 것이 발견되면 이 노드와 다른 노드가 공유하는 모든 디스크에 **SCSI** 예약이 지정되기 때문입니다. 금지된 노드는 금지된 것을 인식하지 못할 수 있기 때문에 공유 디스크 중 하나에 액세스하려고 시도하면 예약을 발견하고 중단됩니다.

장애 방지를 위한 파일패스트 메커니즘

장애가 발생한 노드가 다시 부트되어 공유 기억 장치에 쓰지 못하도록 하기 위하여 클러스터 프레임워크에서 사용하는 메커니즘을 파일패스트라고 합니다.

클러스터를 구성하는 노드는 정족수 디스크를 포함하여 액세스할 수 있는 디스크에 대하여 특정 **ioctl**, **MHIOCENFAILFAST**를 계속 사용할 수 있도록 합니다. 이 **ioctl**은 디스크 드라이버에 대한 지시어이고, 디스크가 다른 노드에 예약되어 디스크에 액세스할 수 없을 경우에 노드가 종료될 수 있도록 합니다.

MHIOCENFAILFAST ioctl을 사용하면 노드가 디스크에 대해 실행하는 모든 읽기 및 쓰기에서 반환되는 오류로부터 드라이버가 **Reservation_Conflict** 오류 코드를 검사합니다. **ioctl**은 백그라운드에서 주기적으로 디스크에 테스트 작업을 실행하여 **Reservation_Conflict** 오류 코드를 검사합니다. **Reservation_Conflict** 오류 코드가 반환되면 포그라운드 및 백그라운드 제어 흐름 경로가 모두 중단됩니다.

SCSI-2 디스크의 경우에는 예약이 지속되지 않습니다. 즉, 노드를 다시 부트하면 예약이 취소됩니다. **PGR(Persistent Group Reservation)**이 있는 **SCSI-3** 디스크의 경우에는 예약 정보가 디스크에 저장되어 노드를 다시 부팅한 후에도 유지됩니다. 파일패스트 메커니즘은 **SCSI-2** 디스크를 사용하는 경우나 **SCSI-3** 디스크를 사용하는 경우에 모두 동일하게 작동합니다.

노드가 클러스터의 다른 노드와 연결이 끊어지고 정족수를 채울 수 있는 파티션에 포함되지 않은 경우에는 다른 노드에 의해 강제로 클러스터에서 제거됩니다. 정족수를 채울 수 있는 파티션에 포함된 다른 노드가 공유 디스크에 예약을 설정하고, 정족수가 채워지지 않은 노드가 예약된 공유 디스크에 액세스하려고 시도하면 파일패스트 메커니즘에 의해 예약 충돌이 발생하여 종료됩니다.

노드가 종료된 후에 다시 부팅되어 클러스터에 다시 연결될 수도 있고 **OpenBoot PROM (OBP)** 프롬프트 상태로 있을 수도 있습니다. 취하는 조치는 **OBP**의 **auto-boot?** 설정에 따라 결정됩니다.

볼륨 관리자

SunPlex 시스템은 미러 및 핫 스페어 디스크를 사용하여 데이터 가용성을 높이고 디스크 장애 및 교체를 처리하기 위해 볼륨 관리 소프트웨어를 사용합니다.

SunPlex 시스템 내부에는 볼륨 관리 구성 요소가 없지만 다음과 같은 볼륨 관리자를 사용합니다.

- Solstice DiskSuite
- VERITAS Volume Manager

클러스터의 볼륨 관리 소프트웨어는 다음과 같은 기능을 제공합니다.

- 노드 장애가 발생할 경우에 페일오버 처리
- 다른 노드로부터의 복수 경로 지원

■ 디스크 장치 그룹에 대한 투명한 원격 액세스

볼륨 관리 객체가 클러스터의 제어를 받으면 디스크 장치 그룹이 됩니다. 볼륨 관리자
에 대한 자세한 내용은 볼륨 관리자 소프트웨어 설명서를 참조하십시오.

주: 디스크 세트나 디스크 그룹을 계획할 때 클러스터 내에서 관련 디스크 장치 그룹
이 응용프로그램 자원(데이터)과 어떻게 연결되는지를 이해하는 것이 중요합니다. 이
문제에 대한 자세한 내용은 **Sun Cluster 3.0 12/01** 소프트웨어 설치 안내서 및 **Sun
Cluster 3.0 12/01 Data Services Installation and Configuration Guide**를 참조하십시오.

데이터 서비스

데이터 서비스는 Oracle이나 iPlanet Web Server와 같이 단일 서버가 아닌 클러스터에
서 실행되도록 구성된 다른 회사 응용 프로그램을 나타내는 용어입니다. 데이터 서비
스는 응용프로그램, Sun Cluster 구성 파일, 다음과 같은 응용프로그램 작업을 제어하
는 Sun Cluster 관리 메소드 등으로 구성됩니다.

- 시작
- 중지
- 모니터 및 수정 조치

그림 3-4에서는 단일 응용프로그램 서버에서 실행되는 응용프로그램(단일 서버 모델)
을 클러스터에서 실행되는 동일한 응용프로그램(클러스터 서버 모델)과 비교합니다.
사용자의 관점에서 보면 클러스터 응용프로그램이 더 빠르게 실행될 수 있고 가용성
이 높다는 것 외에는 두 구성 사이에 큰 차이가 없습니다.

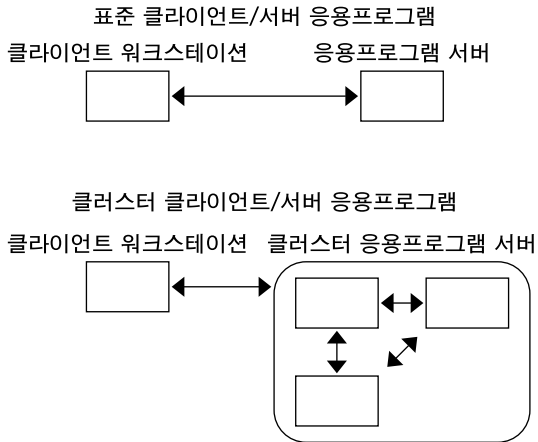


그림 3-4 표준 및 클러스터 클라이언트/서버 구성 비교

단일 서버 모델에서는 응용프로그램이 특정 공용 네트워크 인터페이스(호스트 이름)
를 통해 서버에 액세스하도록 구성합니다. 따라서 호스트 이름이 물리적 서버와 연결
됩니다.

클러스터 서버 모델에서는 공용 네트워크 인터페이스가 논리적 호스트 이름이나 공유 주소입니다. 네트워크 자원은 논리적 호스트 이름과 공유 자원을 모두 나타내는 용어입니다.

일부 데이터 서비스에서는 논리적 호스트 이름 또는 공유 주소를 네트워크 인터페이스로 지정해야 합니다. 두 가지를 동일하게 사용할 수 없습니다. 그 외의 데이터 서비스에서는 논리적 호스트 이름이나 공유 주소를 지정할 수 있습니다. 지정하는 인터페이스 종류에 대한 자세한 내용은 각 데이터 서비스에 대한 설치 및 구성을 참조하십시오.

네트워크 자원은 특정 물리적 서버에만 연결되는 것이 아니고 물리적 서버 사이를 이동할 수 있습니다.

네트워크 자원은 처음에 하나의 노드(1차)와 연결됩니다. 1차 노드에 장애가 발생하면 네트워크 자원과 응용프로그램 자원이 다른 클러스터 노드(2차)로 페일오버합니다. 네트워크 자원이 페일오버하면 잠시 지연된 후에 응용프로그램 자원이 2차 노드에서 계속 실행됩니다.

그림 3-5에서는 단일 서버 모델과 클러스터 서버 모델을 비교합니다. 클러스터 서버 모델에서는 네트워크 자원(이 예에서는 논리적 호스트 이름)이 둘 이상의 클러스터 노드 사이에 이동할 수 있습니다. 응용프로그램은 특정 서버와 연결된 호스트 이름 대신 논리적 호스트 이름을 사용하도록 구성됩니다.

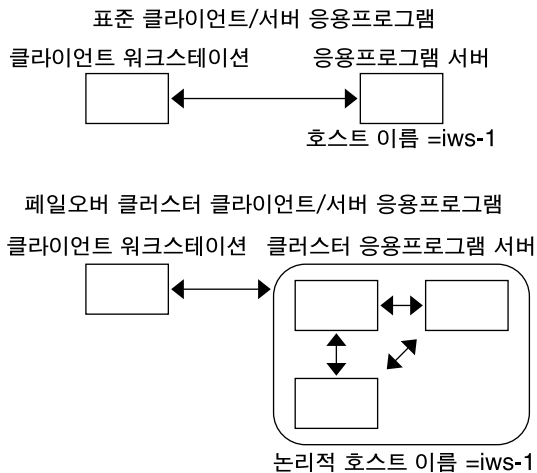


그림 3-5 고정 호스트 이름과 논리 호스트 이름 비교

공유 주소도 처음에는 하나의 노드와 연결됩니다. 이 노드를 **GIN(Global Interface Node)**이라고 합니다. 공유 주소는 클러스터에 대한 단일 네트워크 인터페이스로 사용됩니다. 이러한 인터페이스를 글로벌 인터페이스라고 합니다.

논리적 호스트 이름 모델과 확장 가능 서비스 모델 사이의 차이는 확장 가능 서비스 모델에서는 각 노드에도 루프백 인터페이스에 활성으로 구성된 공유 주소가 있다는 것입니다. 이러한 구성 때문에 여러 노드에 대하여 여러 데이터 서비스 인스턴스를 동시에 실행할 수 있습니다. “확장 가능 서비스”는 클러스터 노드를 추가하여 응용프로그램에 **CPU** 기능을 추가하는 방법으로 성능을 확장할 수 있는 기능입니다.

GIN에 장애가 발생하면 응용프로그램 인스턴스를 실행하는 다른 노드로 공유 주소를 변경하고 변경한 노드를 새 GIN으로 만들 수 있습니다. 아니면 이전에 응용프로그램을 실행하지 않던 다른 클러스터 노드로 페일오버할 수 있습니다.

그림 3-6에서는 단일 서버 구성과 확장 가능한 클러스터 서비스 구성을 비교합니다. 확장 가능 서비스 구성에서는 모든 노드에 대한 공유 주소가 있습니다. 페일오버 데이터 서비스에 논리적 호스트 이름을 사용하는 방법과 유사한 방법으로 특정 서버와 연결된 호스트 이름 대신 이 공유 주소를 사용하도록 응용프로그램이 구성됩니다.

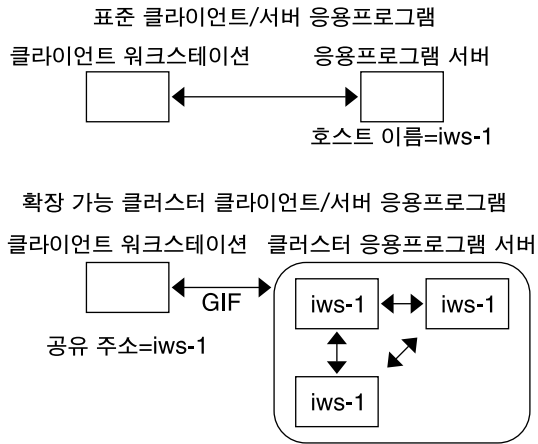


그림 3-6 고정 호스트 이름과 공유 주소 비교

데이터 서비스 방법

Sun Cluster 소프트웨어는 여러 가지 서비스 관리 메소드를 지원합니다. 이 메소드는 Resource Group Manager(RGM)의 제어에 따라 실행되어 클러스터 노드의 응용프로그램을 시작하고 중지하고 모니터링합니다. 클러스터 프레임워크 소프트웨어 및 멀티호스트 디스크와 함께 이 메소드를 사용하면 응용프로그램이 페일오버 또는 확장 가능 데이터 서비스가 될 수 있습니다.

또한 RGM은 응용프로그램의 인스턴스 및 네트워크 자원(논리 호스트이름 및 공유 주소)과 같은 클러스터 내의 자원도 관리합니다.

Sun Cluster 소프트웨어에서 제공하는 메소드 외에 SunPlex 시스템에서도 API와 여러 가지 데이터 서비스 개발 도구를 제공합니다. 이 도구를 사용하면 Sun Cluster 소프트웨어에서 다른 응용프로그램을 가용성이 높은 데이터 서비스로 실행할 수 있도록 응용프로그램 프로그래머가 필요한 데이터 서비스 메소드를 개발할 수 있습니다.

Resource Group Manager(RGM)

RGM은 데이터 서비스(응용프로그램)를 자원 유형으로 구현하여 자원으로 관리합니다. 이 구현은 Sun에서 제공되거나 개발자가 일반 데이터 서비스 템플릿, DSDL

API(데이터 서비스 개발 라이브러리 API) 또는 RMAPI(Resource Management API)를 사용하여 작성합니다. 클러스터 관리자는 *resource groups*라는 컨테이너에 자원을 만들어 관리합니다. RGM은 클러스터 멤버십 변경에 대한 응답으로 선택된 노드에서 자원을 정지하였다가 시작합니다.

RGM은 자원과 자원 그룹을 대상으로 작업을 합니다. RGM 작업을 수행하면 자원과 자원 그룹의 상태가 온라인 및 오프라인으로 전환됩니다. 자원과 자원 그룹에 적용할 수 있는 상태와 설정에 대한 자세한 설명은 (59페이지)“자원 및 자원 그룹의 상태와 설정” 단원을 참조하십시오.

페일오버 데이터 서비스

데이터 서비스가 실행되는 노드에 장애가 발생하면 사용자의 조작 없이 자동으로 서비스가 다른 작업 노드로 이동됩니다. 페일오버 서비스는 페일오버 자원 그룹을 사용합니다. 이 자원 그룹은 응용프로그램 인스턴스 자원 및 네트워크 자원(논리 호스트 이름)을 위한 컨테이너입니다. 논리 호스트 이름은 하나의 노드에서 구성될 수 있는 IP 주소로, 나중에 원래 노드에서 자동으로 구성이 중지되고 다른 노드에서 구성이 시작됩니다.

페일오버 데이터 서비스의 경우에는 응용프로그램 인스턴스가 단일 노드에서만 실행됩니다. 결함 모니터가 오류를 발견하면, 데이터 서비스가 구성된 방법에 따라 동일한 노드에서 인스턴스를 다시 시작하거나 다른 노드에서 인스턴스를 시작합니다(페일오버).

확장 가능 데이터 서비스

확장 가능 데이터 서비스에서는 여러 노드에서 인스턴스가 실행될 수 있습니다. 확장 가능 서비스는 확장 가능 자원 그룹과 페일오버 자원 그룹, 두 가지 자원 그룹을 사용합니다. 확장 가능 자원 그룹은 확장 가능 서비스에서 사용하는 네트워크 자원(공유 주소)를 포함하기 위해 응용프로그램 자원과 페일오버 자원 그룹을 포함합니다. 확장 가능 자원 그룹은 여러 노드에서 온라인 상태로 있을 수 있으므로 서비스의 여러 인스턴스가 동시에 실행될 수 있습니다. 공유 주소를 호스트하는 페일오버 자원 그룹은 한번에 한 노드에서만 온라인 상태가 될 수 있습니다. 확장 가능 서비스를 호스트하는 노드는 모두 동일한 공유 주소를 사용하여 서비스를 호스트합니다.

서비스 요청은 단일 네트워크 인터페이스(글로벌 인터페이스)를 통해 클러스터에 전달되고 로드 밸런싱 정책에 의해 설정된 사전 정의된 몇 가지 알고리즘 중 하나를 사용하여 노드에 분산됩니다. 클러스터는 로드 밸런싱 정책을 사용하여 몇몇 노드 사이의 서비스 부하 균형을 맞추는 로드 밸런싱 정책을 사용할 수 있습니다. 서로 다른 노드에 다른 공유 주소를 호스트하는 여러 개의 글로벌 인터페이스가 있을 수 있습니다.

확장 가능 서비스의 경우에는 응용프로그램 인스턴스를 몇 개의 노드에서 동시에 실행할 수 있습니다. 글로벌 인터페이스를 호스트하는 노드에 장애가 발생하면 글로벌 인터페이스가 다른 노드로 페일오버합니다. 응용프로그램 인스턴스 실행에 실패하면 인스턴스가 동일한 노드에서 다시 시작됩니다.

동일한 노드에서 응용프로그램 인스턴스를 다시 시작할 수 없고, 사용되지 않는 다른 노드가 서비스를 실행하도록 구성되어 있으면, 서비스가 사용되는 않는 노드로 페일

오버됩니다. 그렇지 않으면 계속 나머지 노드에서 실행되어 서비스 처리량이 감소할 수 있습니다.

주: 각 응용프로그램 인스턴스에 대한 TCP 상태는 글로벌 인터페이스 노드에 보존되지 않고 인스턴스가 있는 노드에 보존됩니다. 따라서 글로벌 인터페이스 노드에 장애가 발생해도 연결에는 영향을 주지 않습니다.

그림 3-7은 페일오버 자원 그룹과 확장 가능 자원 그룹 및 확장 가능 서비스를 위한 둘 사이의 관계에 대한 예입니다. 이 예에는 세 가지 자원 그룹이 있습니다. 페일오버 자원 그룹에는 가용성이 높은 DNS를 위한 응용프로그램 자원과 가용성이 높은 DNS 및 가용성이 높은 Apache Web Server에서 사용하는 네트워크 자원이 포함됩니다. 확장 가능 자원 그룹에는 Apache Web Server의 응용프로그램 인스턴스만 포함됩니다. 확장 가능 자원 그룹과 페일오버 자원 그룹 사이에는 자원 그룹 의존 관계가 있고(실선) Apache 응용프로그램 자원은 모두 공유 주소인 네트워크 자원 schost-2를 사용합니다(점선).

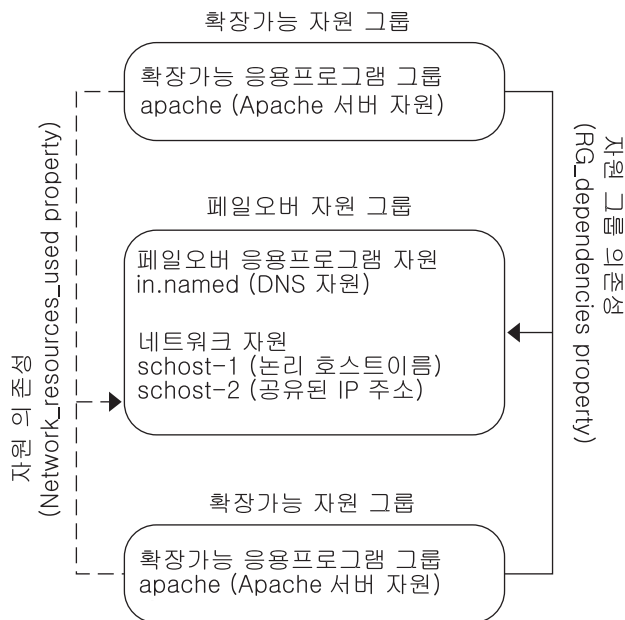


그림 3-7 페일오버 및 확장 가능 자원 그룹의 예

확장 가능 서비스 아키텍처

클러스터 네트워킹의 기본 목적은 데이터 서비스에 확장성을 제공하는 것입니다. 확장성은 서비스에 요청되는 로드가 증가할 경우 에 증가하는 워크로드에 맞게 새로운 노드를 클러스터에 추가하고 새로운 서버 인스턴스를 실행하여 데이터 서비스가 일정한

응답시간을 유지하는 기능을 의미합니다. 이러한 서비스를 확장 가능 데이터 서비스라고 합니다. 확장 가능 데이터 서비스의 좋은 예로 웹 서비스가 있습니다. 일반적으로 확장 가능 데이터 서비스는 각각 서로 다른 클러스터 노드에서 실행되는 여러 개의 인스턴스로 구성됩니다. 해당 서비스의 원격 클라이언트 관점에서는 이러한 모든 인스턴스가 하나의 서비스로 작동하면서 서비스의 기능을 구현합니다. 예를 들어, 서로 다른 노드에서 실행되는 여러 개의 httpd 데몬으로 확장 가능한 웹 서비스를 구성할 수 있습니다. 그러면 모든 httpd 데몬이 클라이언트 요청에 서비스를 제공할 수 있습니다. 요청에 서비스를 제공하는 데몬은 로드 밸런싱 정책에 의해 결정됩니다. 클라이언트에 응답이 전달될 때는 요청에 대하여 서비스를 제공하는 특정 데몬이 표시되는 것이 아니고 서비스로부터 전달되는 것처럼 표시되기 때문에 하나의 서비스로 표시됩니다.

확장 가능 서비스는 다음으로 구성됩니다.

- 확장 가능 서비스를 위한 네트워킹 하부 구조 지원
- 로드 밸런싱
- 네트워킹 및 데이터 서비스 지원(Resource Group Manager 사용)

다음 그림은 확장 가능 서비스의 아키텍처입니다.

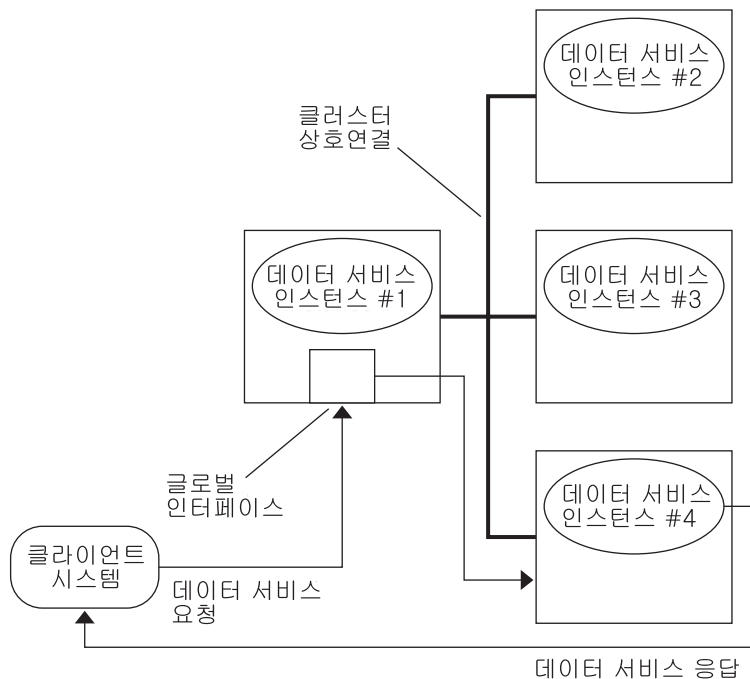


그림 3-8 확장 가능 서비스 아키텍처

글로벌 인터페이스를 호스트하지 않는 노드(프록시 노드)에는 자체 루프백 인터페이스에 대하여 호스트되는 공유 주소가 있습니다. 구성할 수 있는 로드 밸런싱 정책에 따라 글로벌 인터페이스에 전달되는 패킷이 다른 클러스터 노드에 분산됩니다. 구성할 수 있는 로드 밸런싱 정책은 다음 단락에서 설명합니다.

로드 밸런싱 정책

로드 밸런싱은 응답 시간과 처리량에서 확장 가능 서비스의 성능을 향상시킵니다.

확장 가능 데이터 서비스에는 *pure* 및 *sticky* 두 가지 서비스가 있습니다. *pure* 서비스는 포함된 모든 인스턴스가 클라이언트 요청에 응답할 수 있는 서비스입니다. *sticky* 서비스는 클라이언트마다 동일한 인스턴스로 요청을 보내는 서비스입니다. 이 요청은 다른 인스턴스에 전달되지 않습니다.

pure 서비스는 가중(*weighted*) 로드 밸런싱 정책을 사용합니다. 이 로드 밸런싱 정책에서는 기본적으로 클라이언트 요청이 클러스터의 서버 인스턴스에 고르게 분산됩니다. 예를 들어, 3-노드 클러스터에서 각 노드가 처리할 수 있는 로드 크기가 1이라고 가정합니다. 각 노드는 해당 서비스를 위해 모든 클라이언트로부터 받는 요청의 1/3에 해당하는 서비스를 제공합니다. 관리자는 언제든지 `scrgadm(1M)` 명령 인터페이스나 **SunPlex Manager GUI**를 통해 노드의 처리량을 변경할 수 있습니다.

sticky 서비스에는 일반 *sticky*와 와일드카드 *sticky* 두 가지 종류가 있습니다. **Sticky** 서비스에서는 여러 번의 TCP 연결에 걸쳐 진행되는 동시 응용프로그램 레벨 세션이 상태 메모리(응용 프로그램 세션 상태)를 공유할 수 있습니다.

일반 *sticky* 서비스에서는 클라이언트가 여러 개의 동시 TCP 연결 사이에 상태를 공유할 수 있습니다. 서버 인스턴스가 단일 포트를 통해 서비스 요청을 받을 경우에 클라이언트 상태를 “*sticky*”라고 합니다. 서비스가 온라인 상태인 동안 인스턴스가 계속 실행되어 액세스할 수 있고 로드 밸런싱 정책이 변경되지 않으면, 클라이언트가 모든 요청을 동일한 서버 인스턴스로 보낼 수 있습니다.

예를 들어, 클라이언트의 웹 브라우저는 서로 다른 세 개의 TCP 연결을 사용하여 공유 IP 주소의 포트 80에 연결하지만, 서비스에서는 연결 사이에 캐시된 세션 정보를 교환합니다.

sticky 정책을 일반화하면 동일한 인스턴스에서 백그라운드로 세션 정보를 교환하는 여러 개의 확장 가능 서비스로 확장됩니다. 이 인스턴스에서 백그라운드로 세션 정보를 교환하면, 동일한 노드에서 서로 다른 포트를 통해 서비스 요청을 받는 여러 서버 인스턴스에 대하여 클라이언트가 “*sticky*”하다고 합니다.

예를 들어, 전자 상거래 사이트의 고객이 포트 80에서 일반 HTTP를 사용하여 상품을 바구니에 채우지만, 바구니의 상품 대금을 신용 카드로 지불할 경우에는 보안 데이터를 보내기 위해 포트 443의 SSL로 전환합니다.

와일드 카드 *sticky* 서비스는 동적으로 할당되는 포트 번호를 사용하지만, 그래도 클라이언트의 요청이 동일한 노드에 전달될 것이라고 예상합니다. 이러한 경우에는 클라이언트가 동일한 IP 주소의 여러 포트에 대하여 “*sticky wildcard*” 상태입니다.

이러한 정책의 좋은 예로 수동 모드 FTP가 있습니다. 클라이언트가 FTP 서버의 포트 21로 연결하면, 서버가 동적 포트 범위에 있는 수신자 포트 서버에 다시 연결하도록 알립니다. 이 IP 주소에 대한 모든 요청은 서버가 제어 정보를 통해 클라이언트에 알리준 것과 동일한 노드로 전달됩니다.

이러한 *sticky* 정책 각각에 대해 기본적으로 가중(*weighted*) 로드 밸런싱 정책이 적용됩니다. 따라서 클라이언트의 초기 요청은 로드 밸런서가 지시하는 인스턴스로 전달됩니다. 클라이언트가 인스턴스를 실행하는 노드에 적응한 후에 계속 노드에 액세스할

수 있고 로드 밸런싱 정책이 변경되지만 않으면 이후의 요청이 노드에서 실행되는 인스턴스에 전달됩니다.

특정 로드 밸런싱 정책에 대한 자세한 내용은 아래 설명을 참조하십시오.

- **가중(Weighted)**. 지정된 가중치에 따라 로드 가 여러 노드에 분산됩니다. 이 정책은 `Load_balancing_weights` 등록 정보에 `LB_WEIGHTED` 값을 사용하여 설정됩니다. 노드에 대한 가중치가 명확하게 설정되지 않으면 해당 노드의 가중치는 기본값인 `1`이 됩니다.

이 정책은 라운드 로빈(**round robin**) 방식이 아닙니다. 라운드 로빈 정책에서는 클라이언트의 각 요청이 항상 서로 다른 노드에 전달됩니다. 즉, 첫 번째 요청은 노드 `1`로, 두 번째 요청은 노드 `2`로 전달됩니다. 가중 정책을 사용하면 클라이언트로부터 받은 트래픽 중 일정한 비율이 특정 노드에 전달됩니다. 이 정책에서는 개별 요청의 주소를 지정하지 않습니다.

- **Sticky**. 이 정책에서는 응용프로그램 자원이 구성될 때 포트 세트가 알려집니다. 이 정책은 `Load_balancing_policy` 자원 등록 정보에 `LB_STICKY` 값을 사용하여 설정됩니다.
- **Sticky-wildcard**. 이 정책은 일반적인 “sticky” 정책의 상위 세트입니다. IP 주소에 의해 식별되는 확장 가능 서비스의 경우에는 포트가 서버에 의해 할당되고, 사전에 알려지지 않습니다. 포트가 변경될 수도 있습니다. 이 정책은 `Load_balancing_policy` 자원 등록 정보에 `LB_STICKY_WILD` 값을 사용하여 설정됩니다.

페일백 설정

자원 그룹은 한 노드에서 다른 노드로 페일오버됩니다. 그러면 초기의 2차 노드가 새로운 1차 노드가 됩니다. 페일백 설정은 초기 1차 노드가 다시 온라인 상태가 될 때 수행되는 작업을 지정합니다. 초기 1차 노드가 다시 1차 노드가 되는 페일백 옵션 또는 현재 1차 노드를 그대로 유지하는 옵션이 있습니다. `Failback` 자원 그룹 등록 정보 설정을 사용하여 원하는 옵션을 지정합니다.

일부 인스턴스에서는 자원 그룹을 호스트하는 원래 노드에 장애가 발생하여 다시 부트하는 과정을 반복하면 페일백 설정으로 인해 자원 그룹에 대한 가용성이 떨어질 수 있습니다.

데이터 서비스 결함 모니터

각 **SunPlex** 데이터 서비스에는 주기적으로 데이터 서비스를 확인하여 안전 상태를 판단하는 결함 모니터가 있습니다. 결함 모니터는 응용프로그램 데몬이 실행되고 있는지 그리고 클라이언트에 서비스가 제공되고 있는지를 확인합니다. 프로브에서 반환되는 정보에 따라 데몬 재시작이나 페일오버 실행과 같이 사전 정의된 작업이 시작될 수 있습니다.

새 데이터 서비스 개발

Sun에서는 클러스터에서 여러 응용프로그램을 페일오버 또는 확장 가능 서비스로 작동시킬 수 있도록 구성 파일과 관리 메소드 템플릿을 제공합니다. 페일오버 또는 확장 가능 서비스로 실행할 응용프로그램이 현재 Sun에서 제공하는 응용프로그램이 아닌 경우에는 API나 DSDL API를 사용하여 응용프로그램이 페일오버 또는 확장 가능 서비스로 실행되도록 구성할 수 있습니다.

응용프로그램이 페일오버 서비스가 될 수 있는지를 판단하는 몇 가지 기준이 있습니다. 자세한 기준은 응용프로그램에 사용할 수 있는 API를 설명하는 SunPlex 문서에서 설명합니다.

여기서는 현재 제공하는 서비스에서 확장 가능 서비스의 아키텍처를 사용할 수 있는지 확인할 수 있도록 몇 가지 지침을 제공합니다. 확장 가능 서비스에 대한 자세한 내용은 (51페이지)“확장 가능 데이터 서비스” 단원을 참조하십시오.

새로운 서비스가 다음 조건을 충족시키면 확장 가능 서비스를 사용할 수 있습니다. 기존의 서비스가 이 조건을 정확하게 따르지 않으면 서비스가 조건을 따르도록 일부분을 다시 구성해야 할 수도 있습니다.

확장 가능 데이터 서비스에는 다음과 같은 특징이 있습니다. 첫째, 확장 가능 데이터 서비스는 하나 이상의 서버 인스턴스로 구성 됩니다. 각 인스턴스는 서로 다른 클러스터 노드에서 실행됩니다. 동일한 노드에서 동일한 서비스의 인스턴스를 두 개 이상 실행할 수 없습니다.

둘째, 서비스가 외부 논리 데이터 저장소를 제공하면, 변경될 때마다 데이터를 읽거나 갱신 사항을 읽지 않도록, 여러 서버 인스턴스로부터 이 저장소에 대한 동시 액세스를 동기화해야 합니다. 저장소에 저장된 상태를 메모리 내에 기억된 상태와 구별하기 위해 “외부”소가 복제될 수는 있지만 단일 엔티티로 표시되기 때문에 “논리”라고 합니다. 또한, 이 논리 데이터 저장소에는 서버 인스턴스가 저장소를 갱신할 때마다 갱신 사항이 즉시 다른 인스턴스에 표시되는 속성이 있습니다.

SunPlex 시스템은 클러스터 파일 시스템과 글로벌 원시 파티션을 통해 이러한 외부 기억 장치를 제공합니다. 예를 들어, 서비스가 새로운 데이터를 외부 로그 파일에 기록하거나 기존 데이터를 수정한다고 가정합니다. 이 서비스의 인스턴스가 여러 개 실행되면 각 인스턴스가 이 외부 로그에 대한 액세스 권한을 갖고 동시에 이 로그에 액세스할 수 있습니다. 따라서 각 인스턴스가 이 로그에 대한 액세스를 동기화해야 합니다. 그렇지 않으면 각 인스턴스가 서로 간섭을 합니다. 서비스는 `fcntl(2)` 및 `lockf(3C)` 명령을 통한 일반 Solaris 파일 잠금을 사용하여 원하는 동기화를 수행할 수 있습니다.

이러한 저장소의 다른 예로는 가용성이 높은 Oracle이나 Oracle Parallel Server와 같은 백엔드 데이터베이스가 있습니다. 이러한 백엔드 데이터베이스 서버는 데이터베이스 조회 및 갱신 트랜잭션을 사용하여 내장된 동기화 기능을 제공하므로, 여러 서턴스가 자체의 고유 동기화를 구현하지 않아도 됩니다.

현재 단계에서 확장 가능 서비스가 아닌 서비스의 예로는 Sun의 IMAP 서버가 있습니다. 이 서비스는 저장소를 갱신하지만, 그 저장소는 개인용이므로 여러 IMAP 인스턴스가 저장소에 기록할 경우에는 갱신 사항이 동기화되지 않기 때문에 서로 겹쳐씁니다. IMAP 서버는 동시 액세스를 동기화하도록 다시 구성해야 합니다.

끝으로, 인스턴스에 다른 인스턴스의 데이터와 분리된 개인용 데이터가 있을 수 있습니다. 이러한 경우에는 데이터가 개인용이므로 해당 인스턴스만 이 데이터를 처리할 수 있기 때문에 서비스에서 동시 액세스를 동기화할 필요가 없습니다. 이 경우에는 나중에 개인용 데이터에 대한 글로벌 액세스가 가능해질 수도 있기 때문에 이 개인용 데이터를 클러스터 파일 시스템에 저장하지 않도록 해야 합니다.

데이터 서비스 API 및 데이터 서비스 개발 라이브러리 API

SunPlex 시스템은 응용프로그램의 가용성을 높이기 위하여 다음과 같은 기능을 제공합니다.

- SunPlex 시스템의 일부로 제공되는 데이터 서비스
- Data Service API
- Data Service Development Library API
- “일반” 데이터 서비스

*Sun Cluster 3.0 12/01 Data Services Installation and Configuration Guide*에서는 SunPlex 시스템에 제공되는 데이터 서비스를 설치하고 구성하는 방법을 설명합니다. *Sun Cluster 3.0 12/01 Data Services Developer's Guide*에서는 Sun Cluster 프레임워크에서 다른 응용프로그램의 가용성을 높이는 방법을 설명합니다.

응용프로그램 프로그래머가 Sun Cluster API를 사용하면 데이터 서비스 인스턴스를 시작하고 중지하는 스크립트와 결함 모니터를 개발할 수 있습니다. 이러한 도구를 사용하면 응용프로그램을 페일오버 또는 확장 가능 데이터 서비스로 구현할 수 있습니다. 또한 SunPlex 시스템은 응용프로그램을 페일오버 서비스나 확장 가능 서비스로 실행하기 위해 필요한 시작 및 중지 메소드를 신속하게 만들기 위해 사용할 수 있는 “일반” 데이터 서비스를 제공합니다.

데이터 서비스 트래픽을 위한 클러스터 상호 연결 사용

클러스터에는 클러스터 상호 연결을 구성하는 여러 개의 네트워크 연결이 노드 사이에 있어야 합니다. 클러스터 소프트웨어는 가용성을 높이고 성능을 향상시키기 위해 복수 상호 연결을 사용합니다. 내부 트래픽(예: 파일 시스템 데이터 또는 확장 가능 서비스 데이터)의 경우에는 메시지가 사용할 수 있는 모든 상호 연결을 통해 라운드 로빈 방식으로 스트라이프됩니다.

노드 사이에서 수행되는 통신의 가용성을 높이기 위하여 응용프로그램이 클러스터 상호 연결을 사용할 수도 있습니다. 예를 들어, 분산 응용프로그램이 통신이 필요한 서로 다른 노드에서 구성 요소를 실행할 수 있습니다. 공용 전송 기능 대신 클러스터 상호 연결을 사용하면 링크 하나에 장애가 발생해도 이러한 연결이 끊어지지 않습니다.

노드 사이의 통신을 위해 클러스터 상호 연결을 사용하려면 응용프로그램이 클러스터가 설치될 때 구성된 개인용 호스트 이름을 사용해야 합니다. 예를 들어, 노드 1에 대한 개인용 호스트 이름이 `clusternode1-priv`이면, 노드 1에 대한 클러스터 상호 연결을 통해 통신을 할 때 이 이름을 사용하십시오. 이 이름을 사용하여 열리는 TCP 소

켓은 클러스터 상호 연결을 통한 경로를 사용하고 네트워크 장애가 발생할 경우에는 투명하게 경로를 다시 지정할 수 있습니다.

설치 중에 개인용 호스트 이름이 구성될 수 있기 때문에, 설치할 때 선택한 이름을 클러스터 상호 연결이 사용할 수 있습니다. 실제 이름은 `scha_cluster_get(3HA)` 명령에 `scha_privatelink_hostname_node` 인수를 사용하여 얻을 수 있습니다.

응용프로그램 레벨에서 클러스터 상호 연결을 사용할 경우에는 각 노드 쌍 사이에 하나의 상호 연결이 사용되지만 서로 다른 노드 쌍에는 별도의 상호 연결이 사용될 수 있습니다. 예를 들어, 세 개의 노드에서 실행하고 클러스터 상호 연결을 통해 통신하는 응용프로그램을 고려하십시오. 인터페이스 `qfe1`에서 노드 1과 노드 3 사이의 통신이 진행되는 동안 인터페이스 `hme0`에서 노드 1과 노드 2 사이의 통신이 수행될 수 있습니다. 즉, 두 노드 사이의 응용프로그램 통신은 단일상호 연결로 제한되는 반면, 내부 클러스터 통신은 모든 상호 연결에 스트라이프됩니다.

응용프로그램이 내부 클러스터 트래픽과 상호 연결을 공유하므로, 응용프로그램에 사용할 수 있는 대역폭은 다른 클러스터 트래픽에 사용되는 대역폭에 따라 달라집니다. 장애가 발생할 경우에 내부 트래픽은 나머지 상호 연결로 라운드 로빈될 수 있고, 장애가 발생한 상호 연결의 응용프로그램 연결은 작동하는 상호 연결로 전환될 수 있습니다.

두 가지 유형의 주소가 클러스터 상호 연결을 지원하고, 개인용 호스트 이름에 대하여 `gethostbyname(3N)` 명령을 실행하면 일반적으로 두 개의 IP 주소가 반환됩니다. 첫 번째 주소는 논리 *pairwise* 주소라고 하고 두 번째 주소는 논리 *pernode* 주소라고 합니다.

각 노드 쌍에 별도의 논리 *pairwise* 주소가 할당됩니다. 이렇게 작은 논리 네트워크는 연결에 대한 페일오버를 지원합니다. 각 노드에 고정 *pernode* 주소도 할당됩니다. 즉, `clusternode1-priv`에 대한 논리 *pairwise* 주소는 노드마다 다른 반면, `clusternode1-priv`에 대한 논리 *pernode* 주소는 각 노드가 동일합니다. 그러나 노드 자체에는 *pairwise* 주소가 없기 때문에 노드 1에 대하여 `gethostbyname(clusternode1-priv)` 명령을 수행하면 논리 *pernode* 주소가 반환됩니다.

클러스터 상호 연결을 통해 연결을 받은 다음 보안을 위해 IP 주소를 확인하는 응용프로그램은 첫 번째 IP 주소뿐 아니라 `gethostbyname` 명령에서 반환되는 모든 IP 주소에 대하여 확인해야 합니다.

모든 위치의 응용프로그램에 일관된 IP 주소가 필요하다면, 모든 연결이 *pernode* 주소에서 오고 가는 것처럼 표시되도록 클라이언트와 서버 양쪽의 응용프로그램을 *pernode* 주소에 바인드하여 구성하십시오.

자원, 자원 그룹 및 자원 유형

데이터 서비스는 여러 가지 유형의 자원을 사용합니다. 즉, Apache 웹 서버나 iPlanet 웹 서버와 같은 응용프로그램은 응용프로그램을 위해 네트워크 주소(논리 호스트 이름 및 공유 주소)를 사용합니다. 응용프로그램과 네트워크 자원이 RGM에 의해 관리되는 기본 단위를 구성합니다.

데이터 서비스는 자원 유형입니다. 예를 들어, **Sun Cluster HA for Oracle**은 자원 유형 `SUNW.oracle`이고 **Sun Cluster HA for Apache**는 자원 유형 `SUNW.apache`입니다.

자원은 전체 클러스터에서 정의된 자원 유형의 인스턴스입니다. 몇가지의 정의된 자원 유형이 있습니다.

네트워크 자원은 `SUNW.LogicalHostname` 또는 `SUNW.SharedAddress` 자원 유형 중 하나입니다. 이 두 가지 자원 유형은 **Sun Cluster** 소프트웨어에 의해 사전에 등록됩니다.

`SUNW.HAStorage` 자원 유형은 자원이 사용하는 디스크 장치 그룹과 자원의 시작을 동기화하는 데 사용됩니다. 이 자원 유형은 데이터 서비스를 시작하기 전에 클러스터 파일 시스템 마운트 포인트의 경로, 글로벌 장치 및 장치 그룹 이름을 사용할 수 있는지 확인합니다.

RGM에서 관리하는 자원은 하나의 단위로 관리할 수 있도록 자원 그룹이라는 그룹에 포함됩니다. 자원 그룹은 페일오버나 스위치오버가 자원 그룹에서 초기화된 경우에 하나의 단위로 이주됩니다.

주: 응용프로그램 자원이 포함된 자원 그룹을 온라인으로 전환하면 응용프로그램이 시작됩니다. 데이터 서비스 시작 메소드는 응용프로그램이 시작되어 실행될 때까지 대기했다가 성공적으로 종료됩니다. 데이터 서비스 결합 모니터에서 데이터 서비스가 클라이언트에 서비스를 제공하는 것을 결정하는 것과 동일한 방법으로 응용프로그램이 시작되어 실행되는 시기가 결정됩니다. 이 프로세스에 대한 자세한 내용은 *Sun Cluster 3.0 12/01 Data Services Installation and Configuration Guide*의 내용을 참조하십시오.

자원 및 자원 그룹의 상태와 설정

관리자는 자원과 자원 그룹에 정적 설정을 적용합니다. 이 설정은 관리 작업을 통해서만 변경될 수 있습니다. **RGM**은 동적인 “상태” 사이에서 자원 그룹을 전환합니다. 이러한 설정과 상태는 다음 목록에서 설명합니다.

- **Managed** 또는 **unmanaged** - 이것은 자원 그룹에만 적용되는 클러스터 범위의 설정입니다. 자원 그룹은 **RGM**에 의해 관리됩니다. `scrgadm(1M)` 명령을 사용하면 **RGM**이 자원 그룹을 관리하거나 관리하지 않도록 할 수 있습니다. 이 설정은 클러스터를 재구성해도 변경되지 않습니다.

자원 그룹을 처음 만들 때는 관리되지 않습니다. 그룹에 있는 자원이 활성화되기 전에 관리되도록 해야 합니다.

확장 가능 웹 서버와 같은 일부 데이터 서비스에서는 네트워크 자원을 시작하기 전과 중지한 후에 작업을 해야 합니다. 이 작업은 시작(**INIT**) 및 종료(**FINI**) 데이터 서비스 메소드에 의해 수행됩니다. **INIT** 메소드는 자원이 있는 자원 그룹이 관리되는 상태인 경우에만 실행됩니다.

자원 그룹이 관리되지 않는 상태에서 관리되는 상태로 변경되면 그룹에 대하여 등록된 **INIT** 메소드가 그룹의 자원에 대하여 실행됩니다.

자원 그룹이 관리되는 상태에서 관리되지 않는 상태로 변경되면 등록된 **FINI** 메소드가 호출되어 삭제를 수행합니다.

INIT 및 FINI 메소드의 가장 일반적인 용도는 확장 가능한 서비스를 위해 네트워크 자원에 사용되지만 응용프로그램에 의해 수행되지 않는 초기화 또는 삭제 작업에도 사용할 수 있습니다.

- **Enabled** 또는 **disabled** - 이것은 자원에 적용되는 클러스터 범위의 설정입니다. `scrgadm(1M)` 명령을 사용하면 자원을 활성화하거나 비활성화할 수 있습니다. 이 설정은 클러스터를 재구성해도 변경되지 않습니다.

자원에 대한 정상 설정은 시스템에서 활성화되어 실행되는 것입니다.

모든 클러스터 노드에서 자원을 사용하지 못하도록 하려면 자원을 비활성화하면 됩니다. 비활성화된 자원은 일반적인 용도로 사용할 수 없습니다.

- **Online** 또는 **offline** - 이것은 자원 및 자원 그룹에 모두 적용되는 동적 상태입니다. 이 상태는 스위치오버 또는 페일오버 중에 클러스터 재구성 단계를 통해 클러스터 트랜지션으로 변경됩니다. 이 설정은 관리 작업을 통해서만 변경될 수 있습니다. `scswitch(1M)` 명령을 사용하면 자원이나 자원 그룹의 상태를 온라인 또는 오프라인으로 변경할 수 있습니다.

페일오버 자원이나 자원 그룹은 항상 한 노드에서만 온라인 상태가 될 수 있습니다. 확장 가능 자원이나 자원 그룹은 일부 노드에서는 온라인 상태이고 다른 노드에서는 오프라인 상태일 수 있습니다. 스위치오버나 페일오버 중에는 자원 그룹과 포함된 자원이 한 노드에서는 오프라인 상태가 되고 다른 노드에서는 온라인 상태가 됩니다.

자원 그룹이 오프라인이면 모든 자원이 오프라인 상태가 됩니다. 자원 그룹이 온라인이면 모든 자원이 온라인 상태가 됩니다.

자원 그룹에는 여러 자원이 포함될 수 있고, 자원 사이에는 의존 관계가 있습니다. 이러한 의존성을 위해서는 자원이 특정 순서로 온라인 및 오프라인 상태가 되어야 합니다. 자원을 온라인 및 오프라인 상태로 변경하는 메소드의 실행 시간은 자원마다 다를 수 있습니다. 자원의 의존성과 시작 및 중지 시간의 차이 때문에 클러스터 재구성 중에 단일 자원 그룹 내에 있는 자원의 온라인 및 오프라인 상태가 서로 다를 수 있습니다.

자원 및 자원 그룹 등록 정보

SunPlex 데이터 서비스를 위해 자원과 자원 그룹에 대한 등록 정보 값을 구성할 수 있습니다. 표준 등록 정보는 모든 데이터 서비스에 공통입니다. 확장 등록 정보는 각 데이터 서비스에만 적용됩니다. 일부 표준 및 확장 등록 정보는 기본 설정으로 구성되며 수정하지 않아도 됩니다. 다른 등록 정보는 자원을 만들고 구성하는 프로세스에서 설정해야 합니다. 설정할 수 있는 자원 등록 정보와 설정 방법은 각 데이터 서비스에 대한 문서에서 설명합니다.

표준 등록 정보는 보통 특정 데이터 서비스와 독립적인 자원 및 자원 그룹 등록 정보를 구성하는 데 사용됩니다. 표준 등록 정보 세트는 *Sun Cluster 3.0 12/01 Data Services Installation and Configuration Guide*의 부록에서 설명합니다.

확장 등록 정보는 응용프로그램 바이너리 및 구성 파일의 위치와 같은 정보를 제공합니다. 데이터 서비스를 구성할 때 확장 등록정보를 수정할 수 있습니다. 확장 등록 정

보 세트는 *Sun Cluster 3.0 12/01 Data Services Installation and Configuration Guide*에서 데이터 서비스에 대하여 설명하는 각 장에서 설명합니다.

PNM(Public Network Management) 및 NAFO(Network Adapter Failover)

클라이언트는 공용 네트워크 인터페이스를 통해 클러스터에 데이터를 요청합니다. 각 클러스터 노드는 공용 네트워크 어댑터를 통해 하나 이상의 공용 네트워크에 연결됩니다.

Sun Cluster PNM(Public Network Management) 소프트웨어는 공용 네트워크 어댑터를 모니터링하고 결함이 발견될 경우에 한 어댑터에서 다른 어댑터로 IP 주소를 페일오버하기 위한 기본 메커니즘을 제공합니다. 클러스터 노드마다 고유한 PNM 구성이 있기 때문에 다른 클러스터 노드에서는 구성이 다를 수 있습니다.

여러 개의 공용 네트워크 어댑터가 NAFO 그룹(네트워크 어댑터 페일오버 그룹)을 구성합니다. 각 NAFO 그룹에는 하나 이상의 공용 네트워크 어댑터가 있습니다. 각 NAFO 그룹에서 동시에 하나의 어댑터만 작동할 수 있고, 동일한 그룹에 속한 다른 어댑터는 PNM 데몬이 작동하는 어댑터에 대하여 결함을 발견하여 어댑터 페일오버가 발생할 경우에 사용되는 백업 어댑터 기능을 합니다. 페일오버는 작동하는 어댑터와 연결된 IP 주소를 백업 어댑터로 전환하여 노드에 대한 공용 네트워크 연결을 유지합니다. 페일오버는 어댑터 인터페이스 레벨에서 발생하기 때문에 TCP와 같은 고급 연결은 페일오버의 영향을 받지 않고 짧은 지연만 발생합니다.

주: 페일오버 중에 일부 세그먼트가 손실될 수 있기 때문에 TCP의 혼잡 복구 특성으로 인해 TCP에서 혼잡 제어 메커니즘을 작동하기 위해 성공적인 페일오버 후에 TCP 엔드포인트에 추가 지연이 발생할 수 있습니다.

NAFO 그룹은 논리 호스트 이름 및 공유 주소 자원에 대한 빌딩 블록을 제공합니다. `scrgadm(1M)` 명령은 필요할 경우에 자동으로 NAFO 그룹을 만듭니다. 또한 사용자도 논리 호스트 이름과 공유 주소 자원의 NAFO 그룹을 독립적으로 만들어서 클러스터 노드에 대한 공용 네트워크 연결을 모니터링할 수 있습니다. 한 노드의 동일한 NAFO 그룹이 여러 논리 호스트 이름이나 공유 주소 자원을 호스팅할 수 있습니다. 논리 호스트 이름 및 공유 자원에 대한 자세한 내용은 *Sun Cluster 3.0 12/01 Data Services Installation and Configuration Guide*를 참조하십시오.

주: NAFO 메커니즘은 어댑터 장애를 발견하여 마스킹하도록 설계되었습니다. 이것은 관리자가 `ifconfig(1M)` 명령으로 논리(또는 공유) IP 주소 중 하나를 제거하여 복구하기 위해 설계된 것이 아닙니다. Sun Cluster 소프트웨어는 논리 및 공유 IP 주소를 RGM에 의해 관리되는 자원으로 인식합니다. 관리자가 정확하게 IP 주소를 추가하거나 제거하려면 `scrgadm(1M)` 명령을 사용하여 자원이 포함된 자원 그룹을 수정해야 합니다.

PNM 결함 발견 및 페일오버 프로세스

PNM은 어댑터를 통과하는 정상적인 네트워크 트래픽 때문에 정상 어댑터의 패킷 카운터가 변경된다고 가정하고 작동 어댑터의 패킷 카운터를 주기적으로 확인합니다. 패킷 카운터가 일정 시간 동안 변경되지 않으면 PNM이 ping 시퀀스가 되어 트래픽을 강제로 현재 작동하는 어댑터로 보냅니다. PNM은 각 시퀀스의 끝에서 패킷 카운터의 변경을 확인하고, ping 시퀀스가 여러 번 반복된 후에도 패킷 카운터가 변경되지 않으면 어댑터 결함을 선언합니다. 이러한 이벤트가 발생하면 사용할 수 있는 백업 어댑터로 페일오버가 트리거됩니다.

입력 패킷과 출력 패킷의 카운터 모두 또는 두 패킷 카운터 중 하나가 일정 시간 동안 변경되지 않으면 ping 시퀀스가 시작되도록 PNM이 입력 패킷과 출력 패킷의 카운터를 모니터합니다.

핑 시퀀스는 ALL ROUTER 멀티캐스트 주소(224.0.0.2), ALL HOST 멀티캐스트 주소(224.0.0.1) 및 로컬 서브넷 브로드캐스트 주소에 대한 핑으로 구성됩니다.

핑은 최소 비용 우선(least-costly-first) 방식으로 구성되므로 비용이 적게 드는 핑이 성공하면 비용이 많이 드는 핑은 실행되지 않습니다. 또한 핑은 어댑터에 트래픽을 발생시키는 수단으로만 사용됩니다. 핑의 종료 상태는 어댑터가 작동하는지 아니면 결함이 있는지를 판단하는 데는 도움이 되지 않습니다.

이 알고리즘에는 네 가지의 조정가능한 매개변수가 있습니다. inactive_time, ping_timeout, repeat_test 및 slow_network가 그 매개변수입니다. 이 매개변수를 사용하면 장애 감지 속도와 정확도를 상대적으로 조정할 수 있습니다. 매개변수에 대한 세부사항과 변경 방법에 대해서는 Sun Cluster 3.0 12/01 시스템 관리 안내서에서 공용 네트워크 매개변수를 변경하는 절차를 참조하십시오.

NAFO 그룹의 작동 어댑터에서 결함이 발견된 후에 백업 어댑터를 사용할 수 없으면 그룹이 DOWN으로 선언되고 백업 어댑터에 대한 테스트가 계속됩니다. 그렇지 않고 백업 어댑터를 사용할 수 있으면 백업 어댑터로 페일오버가 발생합니다. 결함이 있는 작동어댑터가 중지되어 제거된 동안에는 논리 주소 및 관련 플러그가 백업 어댑터로 "이동"됩니다.

IP 주소 페일오버가 성공적으로 완료되면 ARP 브로드캐스트가 전송됩니다. 따라서 원격 클라이언트에 대한 연결이 유지됩니다.

동적 재구성 지원

DR(동적 재구성) 소프트웨어 기능에 대한 Sun Cluster 3.0의 지원이 한 단계씩 개발되고 있습니다. 이 단원에서는 Sun Cluster 3.0 12/01의 DR 기능 지원에 대한 개념과 참고 사항을 설명합니다.

Solaris 8 DR 기능에 대하여 문서화된 요구 사항, 절차 및 제한은 모두 Sun Cluster DR 지원에 적용됩니다(운영 체제 작동 정지 제외). 따라서 Sun Cluster 소프트웨어에서 DR 기능을 사용하기 전에 Solaris 8 DR 기능에 대한 문서를 참조하십시오. 특히 DR 연결 종료 작업 중에 비네트워크 IO 장치에 영향을 주는 문제를 확인해야 합니다. Sun Enterprise 10000 Dynamic Reconfiguration User Guide 및 Sun Enterprise 10000

Dynamic Reconfiguration Reference Manual(Solaris 8 on Sun Hardware 모음 중에 포함)은 <http://docs.sun.com>에서 다운로드할 수 있습니다.

동적 재구성에 대한 일반적인 설명

DR 기능을 사용하면 실행하는 시스템에서 시스템 하드웨어 제거와 같은 작업을 할 수 있습니다. DR 프로세스는 시스템을 중단하거나 클러스터 가용성을 방해하지 않고 연속적으로 시스템을 작동할 수 있도록 설계되었습니다.

DR은 보드 레벨로 작동합니다. 따라서 DR 작동이 보드의 모든 구성 요소에 영향을 줍니다. 각 보드에는 CPU, 메모리를 비롯하여 디스크 드라이브, 테이프 드라이브 및 네트워크 연결을 위한 주변 장치 인터페이스를 포함한 여러 구성 요소가 포함될 수 있습니다.

보드를 제거하면 보드에 있는 구성 요소를 사용하는 시스템의 기능이 중단됩니다. 보드를 제거하기 전에 DR 하위 시스템이 보드의 구성 요소가 사용되고 있는지 확인합니다. 사용하고 있는 장치를 제거하면 시스템에 장애가 발생할 수 있습니다. DR 하위 시스템이 장치를 사용하고 있는 것을 확인하면 하위 시스템이 DR 보드 제거 작업을 거부합니다. 따라서 항상 안전하게 DR 보드 제거 작업을 수행할 수 있습니다.

DR 보드 추가 작업도 항상 안전합니다. 새로 추가되는 보드의 CPU와 메모리는 시스템에 의해 자동으로 서비스에 포함됩니다. 그러나 새로 추가되는 보드의 다른 구성 요소를 바로 사용하려면 시스템 관리자가 직접 클러스터를 구성해야 합니다.

주: DR 하위 시스템에는 여러 레벨이 있습니다. 낮은 레벨에서 오류를 보고하면 상위 레벨도 오류를 보고합니다. 그러나 하위 레벨이 특정 오류를 보고하면 상위 레벨이 "Unknown 오류"를 보고합니다. 시스템 관리자는 상위 레벨에 의해 보고되는 "Unknown 오류"를 무시해야 합니다.

다음 단원에서는 서로 다른 장치 유형에 대한 DR 참고 사항을 설명합니다.

CPU 장치에 대한 DR 클러스터링 참고 사항

DR 보드 제거 작업이 보드에 있는 CPU에 영향을 줄 경우에는 DR 하위 시스템이 작업을 허용하고 이 CPU를 사용하는 노드를 자동으로 중지시킵니다.

DR 보드 추가 작업이 추가되는 보드의 CPU에 영향을 줄 경우에는 DR 하위 시스템이 이 CPU를 사용하여 자동으로 노드를 시작합니다.

메모리에 대한 DR 클러스터링 참고 사항

DR의 목적을 위해 두 가지 유형의 메모리를 고려해야 합니다. 이 두 가지 유형은 용도만 다릅니다. 실제 하드웨어는 두 가지 유형이 동일합니다.

운영 체제에 사용되는 메모리를 커널 메모리 케이지라고 합니다. Sun Cluster 소프트웨어는 커널 메모리 케이지가 포함된 보드에 대해서는 보드 제거 작업을 지원하지 않

고 이러한 작업은 모두 거부합니다. DR 보드 제거 작업이 커널 메모리 캐이지가 아닌 메모리에 영향을 줄 경우에는 DR 하위 시스템이 작업을 허용하고 해당 메모리를 사용하는 노드를 자동으로 중단시킵니다.

DR 보드 추가 작업이 메모리에 영향을 줄 경우에는 DR 하위 시스템이 새 메모리를 사용하여 자동으로 노드를 시작합니다.

디스크 및 테이프 드라이브에 대한 DR 클러스터링 참고 사항

1차 노드에서 현재 작동하는 드라이브에 대한 DR 제거 작업은 허용되지 않습니다. 1차 노드에서 현재 작동하지 않는 드라이브와 2차 노드의 드라이브에 대한 DR 제거 작업만 수행할 수 있습니다. DR 작업 전후에 모두 클러스터 데이터 액세스는 계속됩니다.

주: 정족수 장치의 가용성에 영향을 주는 DR 작업은 허용되지 않습니다. 정족수 장치에 대한 참고 사항과 정족수 장치에 대하여 DR 작업을 수행하기 위한 절차는 (65페이지)“정족수 장치에 대한 DR 클러스터링 참고 사항”을 참조하십시오.

다음 단계에서는 디스크나 테이프 드라이브에 대하여 DR 제거 작업을 수행하는 절차를 간략하게 설명합니다. 이 작업을 수행하기 위한 자세한 방법은 *Sun Cluster 3.0 U1 System Administration Guide*를 참조하십시오.

1. 디스크나 테이프 드라이브가 현재 작동하는 장치 그룹에 포함되었는지 확인하십시오.
 - 드라이브가 현재 작동하는 장치 그룹에 포함되지 않았으면 드라이브에 대하여 DR 제거 작업을 수행할 수 있습니다.
 - DR 보드 제거 작업이 현재 작동하는 디스크나 테이프 드라이브에 영향을 줄 경우에는 시스템이 작업을 거부하고 작업의 영향을 받을 드라이브를 식별합니다. 장치가 현재 작동하는 장치 그룹에 포함되었으면 (64페이지)단계 2로 이동하십시오.
2. 드라이브가 1차 노드의 구성 요소인지 아니면 2차 노드의 구성 요소인지 확인하십시오.
 - 드라이브가 2차 노드의 구성 요소이면 DR 제거 작업을 수행할 수 있습니다.
 - 드라이브가 1차 노드의 구성 요소이면 DR 제거 작업을 수행하기 전에 1차 노드와 2차 노드를 전환해야 합니다.



주의: 2차 노드에 대한 DR 작업을 수행할 때 현재 1차 노드에 장애가 발생하면 클러스터 가용성이 영향을 받습니다. 새로운 2차 노드가 제공될 때까지 1차 노드를 페일오버할 수 없습니다.

정족수 장치에 대한 DR 클러스터링 참고 사항

현재 정족수 장치로 구성된 장치에 대해서는 DR 제거 작업을 수행할 수 없습니다. DR 보드 제거 작업이 정족수 장치에 영향을 줄 경우에는 시스템이 작업을 거부하고 작업의 영향을 받을 정족수 장치를 식별합니다. DR 제거 작업을 수행하려면 정족수 장치의 기능을 비활성화해야 합니다.

다음 단계에서는 정족수 장치에 대하여 DR 제거 작업을 수행하기 위한 절차를 간략하게 설명합니다. 이 작업을 수행하기 위한 자세한 방법은 *Sun Cluster 3.0 U1 System Administration Guide*를 참조하십시오.

1. DR 작업을 수행하는 장치가 아닌 다른 장치를 정족수 장치로 활성화하십시오.
2. DR 작업을 수행하는 장치의 정족수 장치 기능을 해제하십시오.
3. 장치에 대하여 DR 제거 작업을 수행하십시오.

독립 상호 연결 인터페이스에 대한 DR 클러스터링 참고 사항

현재 작동하는 독립 상호 연결 인터페이스에 대해서는 DR 작업을 수행할 수 없습니다. DR 보드 제거 작업이 현재 작동하는 독립 상호 연결 인터페이스에 영향을 줄 경우에는 시스템이 작업을 거부하고 작업의 영향을 받을 인터페이스를 식별합니다. 현재 작동하는 인터페이스를 제거하려면 먼저 비활성화해야 합니다(아래 주의 사항 참조). 인터페이스가 독립 상호 연결로 교체될 경우, 상태가 동일하게 유지되기 때문에 추가로 Sun Cluster를 재구성하는 단계가 필요없습니다.

다음 단계에서는 독립 상호 연결 인터페이스에 대하여 DR 제거 작업을 수행하기 위한 절차를 간략하게 설명합니다. 이 작업을 수행하기 위한 자세한 방법은 *Sun Cluster 3.0 U1 System Administration Guide*를 참조하십시오.



주의: Sun Cluster에서는 각 클러스터 노드에 다른 모든 클러스터 노드에 대한 경로가 하나 이상 작동하고 있어야 합니다. 다른 클러스터 노드에 대한 마지막 경로를 지원하는 독립 상호 연결 인터페이스를 비활성화하면 안 됩니다.

1. DR 작업을 수행하는 상호 연결 인터페이스가 포함된 전송 케이블을 비활성화하면 안 됩니다.
2. 물리적인 독립 상호 연결 인터페이스에 대하여 DR 제거 작업을 수행하십시오.

공용 네트워크 인터페이스에 대한 DR 클러스터링 참고 사항

현재 작동하지 않는 공용 네트워크 인터페이스에 대하여 DR 제거 작업을 수행할 수 있습니다. DR 보드 제거 작업이 현재 작동하는 공용 네트워크 인터페이스에 영향을 줄 경우에는 시스템이 작업을 거부하고 작업의 영향을 받을 인터페이스를 식별합니다. 먼저 현재 작동하는 공용 네트워크 인터페이스가 NAFO 그룹에 있는 현재 작동하는 어댑터 인스턴스 상태에서 제거되어야 합니다.



주의: 비활성화된 네트워크 어댑터에 대하여 **DR** 제거 작업을 수행하는 동안 현재 작동하는 네트워크 어댑터에 장애가 발생하면 가용성에 영향을 받습니다. **DR** 작업을 수행하는 동안에는 현재 작동하는 어댑터를 페일오버할 수 없습니다.

다음 단계에서는 공용 네트워크 인터페이스에 대하여 **DR** 제거 작업을 수행하기 위한 절차를 간략하게 설명합니다. 이 작업을 수행하기 위한 자세한 방법은 *Sun Cluster 3.0 U1 System Administration Guide*를 참조하십시오.

1. 현재 작동하는 어댑터가 **NAFO** 그룹에서 제거될 수 있도록 백업 어댑터로 전환하십시오.
2. **NAFO** 그룹에서 어댑터를 제거하십시오.
3. 공용 네트워크 인터페이스에 대하여 **DR** 작업을 수행하십시오.

자주 묻는 질문(FAQ)

이 장은 SunPlex 시스템에 대하여 자주 묻는 질문에 대한 응답으로 구성되어 있습니다. 응답은 주제별로 구성되어 있습니다.

높은 가용성 FAQ

- 가용성이 높은 시스템이란 정확히 무엇입니까?

SunPlex 시스템은 서버 시스템을 정상적으로 사용할 수 없는 장애가 발생할 경우에도 응용프로그램을 계속 실행하는 클러스터의 기능을 높은 가용성(HA)이라고 정의합니다.

- 클러스터는 어떤 프로세스를 통해 높은 가용성을 제공합니까?

클러스터 프레임워크는 페일오버라는 프로세스를 통해 가용성이 높은 환경을 제공합니다. 페일오버는 클러스터에서 장애가 발생한 노드로부터 작동 중인 다른 노드로 데이터 서비스 자원을 전환하기 위해 클러스터에서 수행하는 일련의 단계입니다.

- HA와 확장 가능 서비스 사이의 차이점은 무엇입니까?

데이터 서비스에는 페일오버와 확장 가능 두 가지의 주요 기능이 있습니다.

HA 서비스는 클러스터에서 한 번에 하나의 1차 노드에서만 응용프로그램을 실행합니다. 다른 노드에서는 다른 응용프로그램을 실행할 수 있지만, 각 응용프로그램이 하나의 노드에서만 실행됩니다. 1차 노드에 장애가 발생하면 장애가 발생한 노드에서 실행하는 응용프로그램이 다른 노드로 페일오버되어 계속 실행됩니다.

확장 가능 서비스는 하나의 응용프로그램을 여러 노드에 분산시켜 하나의 논리 서비스를 제공합니다. 확장 가능 서비스는 실행되는 전체 클러스터에서 노드와 프로세서 수를 조정합니다.

응용프로그램마다 하나의 노드가 클러스터에 대한 물리적 인터페이스를 호스트합니다. 이러한 노드를 GIN(Global Interface Node)이라고 합니다. 클러스터에 여러

개의 GIN이 있을 수 있습니다. 각 GIN은 확장 가능 서비스에서 사용할 수 있는 하나 이상의 논리적 인터페이스를 호스트합니다. 이러한 논리적 인터페이스를 글로벌 인터페이스라고 합니다. 하나의 GIN이 글로벌 인터페이스를 호스트하여 특정 응용프로그램에 대한 모든 요청을 받고 응용프로그램 서버를 실행하는 여러 노드로 이 요청을 전달합니다. GIN에 장애가 발생하면 정상 작동하는 노드로 글로벌 인터페이스가 페일오버됩니다.

응용프로그램을 실행하는 노드에 장애가 발생하면 장애가 발생한 노드가 클러스터에 복귀될 때까지 응용프로그램이 다른 노드에서 계속 실행되고, 이 경우에는 약간 성능이 떨어집니다.

파일 시스템 FAQ

- 하나 이상의 클러스터 노드를 가용성이 높은 NFS 서버로 실행하고 다른 클러스터 노드는 클라이언트로 실행할 수 있습니까?
안됩니다. 루프백 마운트를 하면 안됩니다.
- **Resource Group Manager** 제어를 받지 않는 응용프로그램에 클러스터 파일 시스템을 사용할 수 있습니까?
예. 그러나 RGM 제어가 없으면 응용프로그램을 실행하고 있는 노드에 장애가 발생할 경우에 직접 응용프로그램을 다시 시작해야 합니다.
- 모든 클러스터 파일 시스템의 마운트 포인트가 **/global** 디렉토리에 있어야 합니까?
아닙니다. 그러나 /global과 같이 동일한 마운트 포인트에 클러스터 파일 시스템을 지정하면 이러한 파일 시스템을 쉽게 구성하고 관리할 수 있습니다.
- 클러스터 파일 시스템을 사용하는 것과 NFS 파일 시스템을 내보내는 것이 어떤 차이가 있습니까?
몇 가지 차이점이 있습니다.
 1. 클러스터 파일 시스템은 글로벌 장치를 지원합니다. NFS는 장치에 대한 원격 액세스를 지원하지 않습니다.
 2. 클러스터 파일 시스템에는 글로벌 이름 공간이 있습니다. 하나의 마운트 명령만 필요합니다. NFS를 사용할 경우에는 각 노드에 파일 시스템을 마운트해야 합니다.
 3. 클러스터 파일 시스템에서는 NFS를 사용할 때보다 파일을 캐시하는 경우가 많습니다. 예를 들어, 읽기, 쓰기, 파일 잠금 및 비동기 I/O를 위해 여러 노드에서 파일에 액세스합니다.
 4. 클러스터 파일 시스템은 하나의 서버에 장애가 발생할 경우에 완벽한 페일오버 기능을 제공합니다. NFS는 여러 서버를 지원하지만 읽기 전용 서버에서만 페일오버가 가능합니다.
 5. 클러스터 파일 시스템은 향후에 원격 DMA 및 zero-copy 기능을 제공하는 고속 클러스터 상호 연결을 구축할 수 있도록 설계되었습니다.

6. 클러스터 파일 시스템에서 파일에 대한 등록 정보를 변경하면(예를 들어, `chmod(1M)` 명령 사용), 변경한 내용이 모든 노드에 즉시 적용됩니다. 내보낸 NFS 파일 시스템에서 이 작업을 수행하려면 훨씬 많은 시간이 소요될 수 있습니다.

- 파일 시스템 `/global/devices/<node>@<node ID>`가 내 클러스터 노드에 표시됩니다. 이 파일 시스템을 사용하여 가용성이 높은 글로벌 데이터를 저장할 수 있습니까?

이 파일 시스템은 글로벌 장치 이름 공간을 저장합니다. 이것은 일반적인 용도에 사용하는 파일 시스템이 아닙니다. 데이터는 글로벌이지만 글로벌 방식으로 액세스할 수 없습니다. 각 노드가 자체 글로벌 장치 이름 공간만 액세스합니다. 노드가 중단되면 다른 노드가 중단된 노드에 대한 이름 공간에 액세스할 수 없습니다. 이 파일 시스템은 가용성이 높지 않습니다. 글로벌 방식으로 액세스하거나 가용성이 높아야 하는 데이터를 저장할 경우에는 이 파일 시스템을 사용하면 안 됩니다.

볼륨 관리 FAQ

- 모든 디스크 장치를 이중화해야 합니까?

가용성이 높아야 하는 디스크 장치는 이중화해야 합니다. 그렇지 않으면 RAID-5 하드웨어를 사용하십시오. 모든 데이터 서비스가 가용성이 높은 디스크 장치나 가용성이 높은 디스크 장치에 마운트된 클러스터 파일 시스템을 사용해야 합니다. 이렇게 구성하면 하나의 디스크에 장애가 발생할 경우에도 안전합니다.

- 로컬 디스크(부트 디스크)에 하나의 볼륨 관리자를 사용하고 멀티호스트 디스크에 다른 볼륨 관리자를 사용할 수 있습니까?

Solstice DiskSuite 소프트웨어로 로컬 디스크를 관리하고 VERITAS Volume Manager로 멀티호스트 디스크를 관리할 경우에 이구성이 지원됩니다. 다른 조합은 지원되지 않습니다.

데이터 서비스 FAQ

- 어떤 SunPlex 데이터 서비스를 사용할 수 있습니까?

지원되는 데이터 서비스 목록은 *Sun Cluster 3.0 12/01* 릴리스 정보를 참조하십시오.

- SunPlex 데이터 서비스에서 어떤 응용프로그램 버전이 지원됩니까?

지원되는 응용프로그램 버전 목록은 *Sun Cluster 3.0 12/01* 릴리스 정보를 참조하십시오.

- 자체 데이터 서비스를 구성할 수 있습니까?

예. 자세한 내용은 Data Service Development Library API와 함께 제공되는 *Sun Cluster 3.0 12/01 Data Services Developer's Guide* 및 Data Service Enabling Technologies 문서를 참조하십시오.

- 네트워크 자원을 제공할 때 숫자 IP 주소나 호스트 이름을 지정해야 합니까?

네트워크 자원을 지정하는 데는 숫자 IP 주소를 사용하는 것보다 UNIX 호스트 이름을 사용하는 것이 좋습니다.

- 네트워크 자원을 제공할 때 논리 호스트 이름(LogicalHostname 자원)을 사용하는 것과 공유 주소(SharedAddress 자원)를 사용하는 것이 어떤 차이가 있습니까?

Sun Cluster HA for NFS의 경우가 아니면 문서가 페일오버 모드 자원 그룹의 LogicalHostname 자원을 사용하기 위해 호출할 때마다 SharedAddress 자원 또는 LogicalHostname 자원 중 한 가지를 사용할 수 있습니다. SharedAddress 자원을 사용하면 클러스터 네트워킹 소프트웨어가 SharedAddress에 구성되고 LogicalHostname에는 구성되지 않기 때문에 약간의 추가 오버헤드가 발생합니다.

확장 가능 및 페일오버 데이터 서비스를 모두 구성하여 클라이언트가 동일한 호스트 이름을 사용하여 두 서비스에 모두 액세스할 수 있도록 하려는 경우에 SharedAddress를 사용하면 좋습니다. 이 경우에 SharedAddress 자원은 페일 오버 응용프로그램 자원과 함께 자원 그룹에 포함되지만, 확장 가능 서비스 자원은 별도의 자원 그룹에 포함되어 SharedAddress를 사용하도록 구성됩니다. 그러면 확장 가능 서비스와 페일오버 서비스가 모두 SharedAddress 자원에 구성된 동일한 호스트 이름/주소 세트를 사용할 수 있습니다.

공용 네트워크 FAQ

- 어떤 공용 네트워크 어댑터가 SunPlex 시스템을 지원합니까?

현재는 SunPlex 시스템이 이더넷(10/100BASE-T 및 1000BASE-SX Gb) 공용 네트워크 어댑터를 지원합니다. 이후에 새로운 인터페이스가 지원될 수 있으므로 최신 정보는 Sun 영업 담당자에게 문의하십시오.

- 페일오버에서 MAC 주소는 어떤 기능을 합니까?

페일오버가 발생하면 새로운 ARP(Address Resolution Protocol) 패킷이 생성되어 전체에 브로드캐스트됩니다. 이 ARP 패킷에는 노드가 페일오버된 새로운 물리적 어댑터의 새 MAC 주소와 이전 IP 주소가 포함됩니다. 네트워크상의 다른 시스템이 이 패킷 중 하나를 받으면 ARP 캐시에서 이전의 MAC-IP 매핑을 제거하고 새 MAC-IP를 사용합니다.

- SunPlex 시스템이 OBP(OpenBoot™ PROM)에서 호스트 어댑터에 대한 local-mac-address?=true 설정을 지원합니까?

아니오. 이 변수는 지원되지 않습니다.

- NAFO가 현재 작동하는 어댑터와 백업 어댑터 사이에 스위치오버를 수행할 때 시간이 얼마나 지연됩니까?

몇 분 동안 지연될 수 있습니다. 이것은 **NAFO** 스위치오버가 수행될 때 **ARP**를 외부로 전송하기 때문입니다. 그러나 클라이언트와 클러스터 사이의 라우터가 반드시 **ARP**를 사용하는 것은 아닙니다. 따라서 라우터에서 이 **IP** 주소에 대한 **ARP** 캐시 항목의 시간이 만료될 때까지 **MAC** 주소를 사용할 수 있습니다. 양쪽 **NAFO** 어댑터가 이더넷 스위치에 연결되어 지연이 발생할 수도 있습니다. **NAFO** 스위치오버가 수행되면 **NAFO** 어댑터 중 하나는 제거되고 두 번째 어댑터는 인식됩니다. 그러면 이더넷 스위치가 현재 포트를 비활성화하고 다른 포트를 활성화해야 합니다. 약간의 시간이 걸릴 수도 있습니다. 이더넷에서도 스위치와 새로 활성화되는 어댑터 사이에 속도 협상이 진행되어 시간이 걸립니다. 최종적으로 스위치오버가 완료되면 **NAFO**가 새로 활성화된 어댑터가 정상인지 몇 가지 검사를 수행하여 모든 것이 제대로 작동하는지 확인합니다.

클러스터 구성원 FAQ

- 모든 클러스터 구성원이 동일한 루트 암호를 사용해야 합니까?

각 클러스터 구성원에 동일한 루트 암호를 사용할 필요는 없습니다. 그러나 모든 노드에서 동일한 루트 암호를 사용하면 쉽게 클러스터를 관리할 수 있습니다.

- 노드의 부트 순서가 중요합니까?

대부분의 경우에 그렇지 않습니다. 그러나 앰네시아(자세한 내용은 “정족수 및 정족수 장치”(43페이지) 참조)를 방지하기 위해서는 부트 순서가 중요합니다. 예를 들어, 노드 2가 정족수 장치를 소유하고 있을 때 노드 1이 중단된 상태에서 사용자가 노드 2를 중단시키면 노드 1을 다시 실행하기 전에 노드 2를 먼저 실행해야 합니다. 그러면 시간이 지난 클러스터 구성 정보를 사용하여 노드를 실행하는 사고를 방지할 수 있습니다.

- 클러스터 노드에서 로컬 디스크를 이중화해야 합니까?

예. 이러한 이중화가 반드시 필요한 것은 아니지만, 클러스터 노드 디스크를 이중화하면 디스크를 이중화하지 않을 경우에 발생할 수 있는 노드 중단을 방지할 수 있습니다. 클러스터 노드의 로컬 디스크를 이중화하면 시스템 관리를 위해 오버헤드가 증가하는 단점이 있습니다.

- 클러스터 구성원을 백업하는 데는 어떤 문제가 있습니까?

하나의 클러스터에 대해 여러 가지 백업 방법을 사용할 수 있습니다. 한 가지 방법은 노드 하나에 테이프 드라이브/라이브러리를 연결하여 백업 노드로 사용하는 것입니다. 그런 다음 클러스터 파일 시스템을 사용하여 데이터를 백업합니다. 이 노드를 공유 디스크에 연결하지는 마십시오.

백업 및 복원 절차에 대한 자세한 내용은 *Sun Cluster 3.0 12/01* 시스템 관리 안내서를 참조하십시오.

클러스터 기억 장치 FAQ

- 멀티호스트 기억 장치의 가용성이 높은 이유는 무엇입니까?

멀티호스트 기억 장치는 하나의 디스크에 장애가 발생한 후에도 이중화나 하드웨어 기반의 RAID-5 컨트롤러를 통해 계속 사용할 수 있기 때문에 가용성이 높습니다. 멀티호스트 기억 장치에는 호스트가 두 개 이상 연결되어 있기 때문에 연결된 노드 하나에 장애가 발생해도 안전하게 서비스를 제공할 수 있습니다.

클러스터 상호 연결 FAQ

- SunPlex 시스템은 어떤 클러스터 상호 연결을 지원합니까?

현재는 SunPlex 시스템이 이더넷(100BASE-T Fast Ethernet 및 1000BASE-SX Gb) 클러스터 상호 연결을 지원합니다.

- “케이블”과 전송 “경로” 사이에는 어떤 차이가 있습니까?

클러스터 전송 케이블은 전송 어댑터와 스위치를 사용하여 구성됩니다. 케이블은 구성 요소끼리 연결하는 방식으로 어댑터와 스위치를 결합합니다. 클러스터 토폴로지 관리자는 사용 가능한 케이블을 통해 노드 사이에 종단 간 전송 경로를 구축합니다. 케이블이 직접 전송 경로에 매핑되지는 않습니다.

케이블은 관리자에 의해 정적으로 “활성화”되고 “비활성화”됩니다. 케이블은 활성화 또는 비활성화 “상태(state)”가 있지만 항상 존재하는 “상태(status)”는 아닙니다. 케이블이 비활성화되면 구성되지 않은 것과 같습니다. 비활성화된 케이블은 전송 경로로 사용할 수 없습니다. 비활성화된 경로는 인식되지 않기 때문에 상태를 확인할 수 없습니다. 케이블의 상태는 `scconf -p` 명령을 사용하여 볼 수 있습니다.

전송 경로는 클러스터 토폴로지 관리자에 의해 동적으로 구성됩니다. 전송 경로의 “상태(status)”는 토폴로지 관리자에 의해 결정됩니다. 경로의 상태는 “온라인”이나 “오프라인”이 될 수 있습니다. 전송 경로의 상태는 `scstat (1M)` 명령을 사용하여 볼 수 있습니다.

다음은 케이블이 네 개인 2-노드 클러스터의 예입니다.

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

네 개의 케이블로 다음과 같이 두 개의 전송 경로를 만들 수 있습니다.


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

클라이언트 시스템 FAQ

- 클러스터에서 사용할 경우 특별히 고려해야 할 클라이언트 요구 사항이나 제한 사항이 있습니까?

클라이언트 시스템도 다른 서버처럼 클러스터에 연결합니다. 어떤 경우에는 데이터 서비스 응용프로그램에 따라, 클라이언트가 데이터 서비스 응용프로그램에 연결할 수 있도록 클라이언트측 소프트웨어를 설치하거나 다른 구성 변경사항을 수행해야 할 수도 있습니다. 클라이언트측 구성 요구사항에 대한 자세한 내용은 *Sun Cluster 3.0 12/01 Data Services Installation and Configuration Guide*에서 해당 단원을 참조하십시오.

관리 콘솔 FAQ

- SunPlex 시스템에 관리 콘솔이 필요합니까?

예.

- 관리 콘솔은 클러스터 전용이어야 합니까? 아니면 다른 작업에도 사용할 수 있습니까?

SunPlex 시스템에는 전용 관리 콘솔이 필요 없지만, 전용 관리 콘솔을 사용하면 다음과 같은 이점이 있습니다.

- 동일한 시스템에서 콘솔과 관리 도구를 그룹화하여 중앙에서 클러스터를 관리할 수 있습니다.
 - 하드웨어 서비스 제공업체에서 더욱 신속하게 문제를 분석할 수 있습니다.
 - 관리 콘솔이 클러스터에 “가까이”(예를 들어, 같은 방) 있어야 합니까?
- 하드웨어 서비스 제공업체에 문의해 보십시오. 제공업체가 콘솔을 클러스터 가까이 설치하도록 요구할 수도 있습니다. 그러나 기술적으로는 콘솔이 같은 방에 있어야 할 이유가 없습니다.
- 거리 요구사항이 우선적으로 충족되었을 경우에 관리 콘솔이 둘 이상의 클러스터에 서비스를 제공할 수 있습니까?

예. 하나의 관리 콘솔에서 여러 클러스터를 제어할 수 있습니다. 또한 클러스터 사이에서 하나의 터미널 콘센트레이터를 공유할 수도 있습니다.

터미널 콘센트레이터 및 시스템 서비스 프로세서 FAQ

■ SunPlex 시스템에 터미널 콘센트레이터가 필요합니까?

Sun Cluster 3.0으로 시작하는 모든 소프트웨어 버전은 터미널 콘센트레이터 없이 실행됩니다. 장애를 방지하기 위해 터미널 콘센트레이터가 필요했던 Sun Cluster 2.2 제품과 달리, 이후 버전에서는 터미널 콘센트레이터가 반드시 필요하지는 않습니다.

■ 대부분의 SunPlex 서버에서 터미널 콘센트레이터를 사용하는데 E10000은 이것을 사용하지 않습니다. 그 이유는 무엇입니까?

터미널 콘센트레이터는 실제로 대부분의 서버에서 직렬-이더넷 변환기로 사용되기 때문에 터미널 콘센트레이터의 콘솔 포트는 직렬 포트입니다. 그러나 Sun Enterprise E10000 server에는 직렬 콘솔이 없습니다. SSP(System Service Processor)는 이더넷이나 jtag 포트를 통한 콘솔입니다. Sun Enterprise E10000 server에서는 항상 SSP를 콘솔로 사용합니다.

■ 터미널 콘센트레이터를 사용할 경우 어떤 점이 좋습니까?

터미널 콘센트레이터를 사용하면 노드가 OBP(OpenBoot PROM)에 있는 경우를 포함하여 네트워크의 어느 위치에 있는 원격 워크스테이션에서나 콘솔 레벨로 각 노드에 액세스할 수 있습니다.

■ Sun에서 지원하지 않는 터미널 콘센트레이터를 사용할 경우, 어떤 사항을 알아야 합니까?

Sun에서 지원하는 터미널 콘센트레이터와 다른 콘솔 장치의 가장 큰 차이점은 Sun 터미널 콘센트레이터에는 부트할 때 콘솔로 중단 신호가 전송되지 않도록 하는 특수 펌웨어가 있다는 것입니다. 중단 신호나 중단 신호로 해석될 수 있는 신호를 콘솔로 전송할 수 있는 콘솔 장치가 있으면 이 장치가 노드를 종료시킵니다.

■ 다시 부트하지 않고서도 Sun에서 지원하는 터미널 콘센트레이터의 잠긴 포트를 해제할 수 있습니까?

예. 다시 설정해야 하는 포트 번호를 확인하고 다음 작업을 수행하십시오.

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

Sun에서 지원하는 터미널 콘센트레이터를 구성하고 관리하는 방법은 Sun Cluster 3.0 12/01 시스템 관리 안내서를 참조하십시오.

- 터미널 콘센트레이터 자체에 장애가 발생하면 어떻게 해야 합니까? 다른 콘센트레이터를 준비해야 합니까?

아니오. 터미널 콘센트레이터에 장애가 발생해도 클러스터의 가용성은 유지됩니다. 콘센트레이터가 다시 서비스를 제공할 때까지 노드 콘솔에 연결할 수 없게 됩니다.

- 터미널 콘센트레이터를 사용할 경우에 보안 문제는 없습니까?

일반적으로 터미널 콘센트레이터는 다른 클라이언트가 액세스하는 네트워크에는 연결되지 않고 시스템 관리자가 사용하는 소규모 네트워크에 연결됩니다. 이 네트워크에 대한 액세스를 제한하면 보안 문제를 관리할 수 있습니다.

용어해설

이 용어집은 SunPlex 3.0 문서에서 사용됩니다.

A

관리 콘솔
앰네시아

클러스터 관리 소프트웨어를 실행하기 위해 사용하는 워크스테이션. 잘못된 CCR(Cluster Configuration Data)로 시스템이 종료된 후에 클러스터가 다시 시작되는 상태. 예를 들어, 노드 1만 작동하는 2-노드 클러스터의 경우에 노드 1에서 클러스터 구성이 변경되면 노드 2의 CCR이 잘못된 상태가 됩니다. 클러스터가 종료된 후 노드 2에서 다시 시작되면 노드 2의 잘못된 CCR로 인해 앰네시아 상태가 발생합니다.

자동 페일백

1차 노드에 장애가 발생하여 나중에 클러스터 구성원으로 다시 시작된 후에 자원 그룹이나 장치 그룹을 다시 1차 노드로 전환하는 프로세스.

B

백업 그룹

“NAFO 그룹”을 참조하십시오.

C

체크포인트

1차 노드와 2차 노드 사이에 소프트웨어 상태를 동기화하기 위해 1차 노드에서 2차 노드로 보내는 정보. “1차” 및 “2차”를 참조하십시오. 클러스터 파일 시스템을 공유하고 함께 구성되어 페일오버, 병렬 또는 확장 가능 자원을 실행하도록 상호 연결된 둘 이상의 노드 또는 도메인.

클러스터

CCR(Cluster Configuration Repository)

Sun Cluster 소프트웨어에서 클러스터 구성 정보를 지속적으로 저장하기 위해 사용하는 가용성이 높은 복제 데이터 저장소.

클러스터 파일 시스템	전체 클러스터에서 기존 로컬 파일 시스템에 대하여 가용성이 높은 액세스를 제공하는 클러스터 서비스.
클러스터 상호 연결	케이블, 클러스터 전송 연결 장치 및 클러스터 전송 어댑터를 포함하는 하드웨어 네트워킹 기반 구조. Sun Cluster 및 데이터 서비스 소프트웨어는 클러스터 내에서 이 기반 구조를 사용하여 통신을 합니다.
클러스터 구성원	현재 구현된 클러스터에서 작동하는 구성원. 이 구성원은 다른 클러스터 구성원과 자원을 공유하고 다른 클러스터 구성원 및 클러스터의 클라이언트에 서비스를 제공할 수 있습니다. “클러스터 노드”를 참조하십시오.
CMM(Cluster Membership Monitor)	클러스터 구성원에 대하여 일관된 기록을 유지하는 소프트웨어. 다른 클러스터 소프트웨어에서 가용성이 높은 서비스를 찾을 때 이 구성원 정보를 사용합니다. CCM 을 사용하면 클러스터의 구성원에 포함되지 않은 시스템이 데이터를 손상시키고 손상되거나 일관성이 없는 데이터를 클라이언트에 전송할 수 없습니다.
클러스터 노드	클러스터 구성원이 되도록 구성된 노드. 클러스터 노드가 현재 구성원일 수도 있고 아닐 수도 있습니다. “클러스터 구성원”을 참조하십시오.
클러스터 전송 어댑터	노드에 있으면서 노드를 클러스터 상호 연결에 연결하는 네트워크 어댑터. “클러스터 상호 연결”을 참조하십시오.
클러스터 전송 케이블	엔드포인트에 연결하는 네트워크 연결. 클러스터 전송 어댑터와 클러스터 전송 연결 장치 사이 또는 두 개의 클러스터 전송 어댑터 사이의 연결. “클러스터 상호 연결”을 참조하십시오.
클러스터 전송 연결 장치	클러스터 상호 연결의 일부로 사용되는 하드웨어 스위치. “클러스터 상호 연결”을 참조하십시오.
배열	동일한 노드에 있음을 나타내는 등록 정보. 이 개념은 성능 개선을 위해 클러스터를 구성할 때 사용됩니다.

D

데이터 서비스	RGM(Resource Group Manager) 제어에 의해 가용성이 높은 자원으로 실행되도록 설계된 응용프로그램.
기본 마스터	페일오버 자원 유형이 온라인 상태로 전환되는 기본 클러스터 구성원.
장치 그룹	디스크처럼 클러스터 HA 구성의 다른 노드로부터 마스터될 수 있는 사용자 정의된 장치 자원 그룹. 이 그룹에는 디스크, Solstice

DiskSuite 디스크 세트 및 VERITAS Volume Manager 디스크 그룹과 같은 장치 자원이 포함될 수 있습니다.

장치 ID	<p>Solaris를 통해 사용할 수 있는 장치 식별 메커니즘. 장치 ID에 대한 설명은 <code>devid_get(3DEVID)</code> 설명서 페이지를 참조하십시오.</p> <p>Sun Cluster DID 드라이버는 장치 ID를 사용하여 서로 다른 클러스터 노드에 있는 Solaris 논리적 이름 사이의 관계를 결정합니다. DID 드라이버는 각 장치에서 장치 ID를 확인합니다. 확인한 장치 ID가 클러스터의 다른 위치에 있는 장치와 일치하면 두 장치에 동일한 DID 이름이 할당됩니다. 전에 클러스터에 표시되었던 장치가 아니면 새로운 DID 이름이 할당됩니다. “Solaris 논리 이름”과 “DID 드라이버”를 참조하십시오.</p>
DID 드라이버	<p>전체 클러스터에서 일관된 장치 이름 공간을 제공하기 위해 Sun Cluster 소프트웨어에서 구현하여 사용하는 드라이버. “DID 이름”을 참조하십시오.</p>
DID 이름	<p>SunPlex 시스템에서 글로벌 장치를 식별하는 데 사용됩니다. 이것은 Solaris 논리 이름과 일대일 또는 일대다 관계를 갖는 클러스터 식별자입니다. 형식은 <code>dXsY</code>입니다. 여기서 <code>X</code>는 정수이고 <code>Y</code>는 슬라이스 이름입니다. “Solaris 논리 이름”을 참조하십시오.</p>
디스크 장치 그룹	<p>“장치 그룹”을 참조하십시오.</p>
DLM(Distributed Lock Manager)	<p>공유 디스크 Oracle Parallel Server(OPS) 환경에서 사용하는 잠금 소프트웨어. DLM을 사용하면 다른 노드에서 실행되는 Oracle 프로세스가 데이터베이스 액세스를 동기화할 수 있습니다. DLM은 가용성을 높이기 위해 설계되었습니다. 프로세스나 노드에 장애가 발생해도 나머지 노드를 종료하고 다시 시작할 필요가 없습니다. 장애를 복구하기 위해 신속하게 DLM이 재구성됩니다.</p>
디스크 세트	<p>“장치 그룹”을 참조하십시오.</p>
디스크 그룹	<p>“장치 그룹”을 참조하십시오.</p>

E

엔드포인트 이벤트	<p>클러스터 전송 어댑터나 클러스터 전송 연결 장치의 물리적 포트, 관리 대상 오브젝트에 대한 상태, 소유권, 위험도 또는 설명의 변경.</p>
-----------	---

F

페일백 페일패스트	“자동 페일백”을 참조하십시오. 잘못된 작동으로 손상이 발생하기 전에 클러스터에서 결함이 있는 노드를 순차적으로 종료하고 제거하는 작업.
페일오버	장애가 발생한 후에 현재 1차 노드에서 새로운 1차 노드로 자원 그룹이나 장치 그룹을 자동 전환하는 작업.
페일오버 자원	한 번에 한 노드에서만 올바르게 마스터될 수 있는 자원. “단일 인스턴스 자원”과 “확장 가능 자원”을 참조하십시오.
결함 모니터	데이터 서비스의 여러 부분을 확인하고 조치를 취하는 데 사용되는 프로그램과 결함 데몬. “자원 모니터”를 참조하십시오.

G

일반 자원 유형	데이터 서비스의 템플릿. 일반 자원 유형을 사용하면 간단한 응용 프로그램을 페일오버 데이터 서비스(한 노드에서 정지하고, 다른 노드에서 시작)로 만들 수 있습니다. 이 유형에서는 SunPlex API를 사용하여 프로그래밍을 할 필요가 없습니다.
일반 자원	일반 자원 유형에 포함되어 자원 그룹 관리자의 제어를 받는 응용 프로그램 데몬과 하위 프로세스.
글로벌 장치	디스크, CD-ROM 및 테이프와 같은 모든 클러스터 구성원에게 액세스할 수 있는 장치.
글로벌 장치 이름 공간	전체 클러스터에서 사용하는 글로벌 장치에 대한 논리적 이름이 포함된 이름 공간. Solaris 환경에서 로컬 장치는 /dev/dsk, /dev/rdisk 및 /dev/rmt 디렉토리에 정의됩니다. 글로벌 장치 이름 공간은 /dev/global/dsk, /dev/global/rdisk 및 /dev/global/rmt 디렉토리에 글로벌 장치를 정의합니다.
글로벌 인터페이스	공유 주소를 물리적으로 호스트하는 글로벌 네트워크 인터페이스. “공유 주소”를 참조하십시오.
글로벌 인터페이스 노드	글로벌 인터페이스를 호스트하는 노드.
글로벌 자원	Sun Cluster 소프트웨어의 커널 레벨에서 제공되는 가용성이 높은 자원. 글로벌 자원으로는 디스크(HA 장치 그룹), 클러스터 파일 시스템 및 글로벌 네트워킹이 있습니다.

H

HA 데이터 서비스
하트비트

“데이터 서비스”를 참조하십시오.
사용할 수 있는 모든 클러스터 상호 연결 전송 경로를 통해 전송되는 주기적인 메시지. 지정된 간격 및 재시도 횟수 후에 하트비트가 없으면 전송 통신의 내부 페일오버가 다른 경로로 트리거될 수 있습니다. 클러스터 구성원에 대한 모든 경로에 장애가 발생하면 **CMM**이 클러스터 정족수를 다시 평가합니다.

I

인스턴스

“자원 호출”을 참조하십시오.

L

로드 밸런싱

확장 가능 서비스에만 적용됩니다. 클라이언트 요청에 시간제 방식으로 서비스가 제공되도록 클러스터의 노드 사이에 응용프로그램 로드를 분산시키는 프로세스. 자세한 내용은 (51페이지)“확장 가능 데이터 서비스”을 참조하십시오.

로드 밸런싱 정책

확장 가능 서비스에만 적용됩니다. 노드 사이에 응용프로그램 요청 로드를 분산시키는 기본 방식. 자세한 내용은 (51페이지)“확장 가능 데이터 서비스”을 참조하십시오.

로컬 디스크

지정된 클러스터 노드에만 물리적으로 연결된 디스크.

논리 호스트

응용프로그램, 응용프로그램 데이터가 있는 디스크 세트나 디스크 그룹, 클러스터에 액세스하는 데 사용되는 네트워크 주소 등을 포함한 **Sun Cluster 2.0**(최소) 개념. **SunPlex** 시스템에서는 더 이상 이 개념을 사용하지 않습니다. **SunPlex** 시스템에서 이 개념을 구현하는 방법은 “디스크 장치 그룹”(37페이지) 및 “자원, 자원 그룹 및 자원 유형”(58페이지)을 참조하십시오.

논리 호스트 이름 자원

네트워크 주소를 나타내는 논리 호스트 이름의 모음이 포함된 자원. 논리 호스트 이름 자원은 한 번에 하나의 노드에 의해서만 마스터될 수 있습니다. “논리 호스트”를 참조하십시오.

논리 네트워크 인터페이스

인터넷 아키텍처에서는 하나의 호스트가 하나 이상의 **IP** 주소를 가질 수 있습니다. **Sun Cluster** 소프트웨어는 논리적 네트워크 인터페이스를 추가로 구성하여 여러 개의 논리적 네트워크 인터페이스와 하나의 물리적 네트워크 인터페이스 사이의 매핑을 구성합니다. 각 논리 네트워크 인터페이스에는 하나의 **IP** 주소가 있습니다. 이러한 매핑으로 하나의 물리적 네트워크 인터페이스가 여러 **IP** 주소에 응답할 수 있습니다. 또한 이러한 매핑을 사용하면 테이크오버 또는 스

위치오버 이벤트에서 추가 하드웨어 인터페이스 없이 클러스터 구성원 사이에 IP 주소가 이동할 수 있습니다.

M

마스터 메타 장치 상태 데이터베이스 복제(복제)	“1차”를 참조하십시오. 모든 메타 장치의 구성과 상태 및 오류 조건을 기록하여 디스크에 저장한 데이터베이스. 이 정보는 Solstice DiskSuite 디스크 세트의 올바른 작동을 위해 중요하므로 복제됩니다.
멀티홈 호스트	둘 이상의 공용 네트워크에 있는 호스트.
멀티호스트 디스크	물리적으로 여러 노드에 연결되어 있는 디스크.

N

NAFO(Network Adapter Failover) 그룹 네트워크 주소 자원	어댑터 장애가 발생할 경우에 서로 백업하도록 동일한 노드와 동일한 서브넷에 구성된 하나 이상의 네트워크 어댑터 세트. “네트워크 자원”을 참조하십시오.
네트워크 자원	하나 이상의 논리 호스트 이름이나 공유 주소를 포함하는 자원. “논리 호스트 이름 자원”과 “공유 주소 자원”을 참조하십시오.
노드	Sun Enterprise E10000 server 에서 SunPlex 시스템에 포함될 수 있는 물리적 시스템 또는 도메인. “호스트”라고도 합니다.
비클러스터 모드	-x 부트 옵션을 사용하여 클러스터 구성원을 부트할 때 얻을 수 있는 상태. 이 상태에서는 노드가 더 이상 클러스터 구성원이 아니지만 계속 클러스터 노드로 남아 있습니다. “클러스터 구성원”과 “클러스터 노드”를 참조하십시오.

P

병렬 자원 유형	클러스터 환경에서 실행되어 여러 노드(두 개 이상)에서 동시에 마스터될 수 있도록 설계된 병렬 데이터베이스와 같은 자원 유형.
병렬 서비스 인스턴스	개별 노드에서 실행되는 병렬 자원 유형의 인스턴스.
예비 마스터	“예비 1차”를 참조하십시오.
예비 1차	1차 노드에 장애가 발생할 경우에 페일오버 자원 유형을 마스터할 수 있는 클러스터 구성원. “기본 마스터”를 참조하십시오.

1차	자원 그룹이나 장치 그룹이 현재 온라인 상태인 노드. 즉, 1차 노드는 현재 자원과 관련된 서비스를 구현하거나 호스트하는 노드입니다. “2차”를 참조하십시오.
1차 호스트 이름	1차 공용 네트워크의 노드 이름. 이 이름은 항상 /etc/nodename에 지정된 노드 이름입니다. “2차 호스트 이름”을 참조하십시오.
개인용 호스트 이름	클러스터 상호 연결을 통해 노드와 통신을 하는 데 사용되는 호스트 이름 별명.
PNM(Public Network Management)	단일 네트워크 어댑터나 케이블 장애로 노드 가용성이 떨어지는 것을 방지하기 위해 결합 모니터링 및 페일오버를 사용하는 소프트웨어. PNM 페일오버는 네트워크 어댑터 페일오버 그룹이라고 하는 네트워크 어댑터 세트를 사용하여 클러스터 노드와 공용 네트워크 사이의 중복 연결을 제공합니다. 자원의 가용성을 높이기 위해 결합 모니터링 기능과 페일오버 기능이 함께 작동합니다. “네트워크 어댑터 페일오버 그룹”을 참조하십시오.

Q

정족수 장치	클러스터 실행에 필요한 정족수를 설정하기 위해 사용되는 투표에 포함되는 두 개 이상의 노드에서 공유하는 디스크. 투표 정족수가 충족될 경우에만 클러스터가 작동합니다. 정족수 장치는 클러스터가 별도의 노드 세트로 분할되어 새 클러스터를 구성하는 노드 세트를 설정할 때 사용됩니다.
--------	---

R

자원	자원 유형의 인스턴스. 동일한 유형의 자원이 많이 있을 수 있습니다. 각 자원은 많은 기반 응용프로그램의 인스턴스가 클러스터에서 실행될 수 있도록 고유한 이름과 등록 정보 값 세트를 갖습니다.
자원 그룹	RGM에 의해 하나의 단위로 관리되는 자원 모음. RGM에 의해 관리되는 각 자원은 자원 그룹에 구성되어야 합니다. 일반적으로 서로 관련되어 의존하는 자원이 그룹화됩니다.
RGM(Resource Group Manager)	선택된 클러스터 노드에서 자동으로 클러스터 자원을 시작하고 정지시켜 클러스터 자원의 가용성과 확장성을 높이는 데 사용하는 소프트웨어. RGM은 하드웨어 또는 소프트웨어 장애가 발생하거나 다시 부트할 때 사전에 구성된 정책에 따라 작동합니다.
자원 그룹 상태	지정된 노드에서 자원 그룹의 상태.

자원 호출	노드에서 실행되는 자원 유형의 인스턴스. 노드에서 시작된 자원을 나타내는 추상적 개념.
RMAPI(Resource Management API)	클러스터 환경에서 응용프로그램의 가용성을 높이는 SunPlex 시스템 내의 응용프로그램 프로그래밍 인터페이스.
자원 모니터	자원 유형 구현에서 자원이 올바르게 실행되는지 그리고 어떻게 수행하는지 판단하기 위해 주기적으로 결함 확인을 실행하는 옵션 부분.
자원 상태(state)	지정된 노드에서 Resource Group Manager 자원의 상태.
자원 상태(status)	결함 모니터에서 보고되는 자원의 상태.
자원 유형	데이터 서비스, LogicalHostname 또는 SharedAddress 클러스터 객체에 부여되는 고유한 이름. 데이터 서비스 자원 유형은 페일 오버 유형이 될 수도 있고 확장 가능 유형이 될 수도 있습니다. “데이터 서비스,” “페일오버 자원” 및 “확장 가능 자원”을 참조하십시오.
자원 유형 등록정보	RGM에 의해 자원 유형의 일부로 저장되고, 지정된 유형의 자원을 설명하고 관리하기 위해 사용되는 키-값 쌍.

S

SCI(Scalable Coherent Interface) 확장 가능 자원	클러스터 상호 연결로 사용되는 고속 상호 연결 하드웨어. 원격 클라이언트에 단일 서비스 형태로 서비스를 제공하기 위해 클러스터 상호 연결을 사용하는, 여러 노드에서 실행되는(각 노드에 하나의 인스턴스) 자원.
확장 가능 서비스	여러 노드에서 동시에 실행되도록 구현된 데이터 서비스.
2차	1차 노드에 장애가 발생할 경우에 디스크 장치 그룹과 자원 그룹을 마스터할 수 있는 클러스터 구성원. “1차”를 참조하십시오.
2차 호스트 이름	2차 공용 네트워크에서 노드에 액세스하기 위해 사용하는 이름. “1차 호스트 이름”을 참조하십시오.
공유 주소 자원	해당 노드에서 서비스를 확장할 수 있도록 클러스터 내의 노드에서 실행되는 모든 확장 가능 서비스에 의해 바인드될 수 있는 네트워크 주소. 하나의 클러스터에 여러 개의 공유 주소가 있을 수 있고, 하나의 서비스가 여러 공유 주소에 바인드될 수 있습니다.
단일 인스턴스 자원	전체 클러스터에서 하나의 자원만 사용할 수 있는 자원.

Solaris 논리 이름	Solaris 장치를 관리하기 위해 일반적으로 사용하는 이름. 디스크의 경우에는 일반적으로 /dev/rdisk/c0t2d0s2와 같은 이름을 사용합니다. 이러한 Solaris 논리 장치 이름마다 기반이 되는 Solaris 물리적 장치 이름이 있습니다. “ DID 이름”과 “ Solaris 물리적 이름”을 참조하십시오.
Solaris 물리적 이름	Solaris 에서 장치 드라이버에 의해 장치에 부여되는 이름. 이 이름은 Solaris 시스템의 /devices 트리 아래에 경로로 표시됩니다. 예를 들어, SCSI 디스크는 일반적으로 다음과 같은 Solaris 물리적 이름을 갖습니다./devices/sbus@1f,0/SUNW,fas@e,8800000/sd@6,0:c,raw “ Solaris 논리 이름”을 참조하십시오.
Solstice DiskSuite	SunPlex 시스템에서 사용하는 볼륨 관리자. “볼륨 관리자”를 참조하십시오.
브레인 분할	하나의 클러스터가 여러 파티션으로 분리되어 각 파티션이 다른 파티션의 존재를 인식하지 못하는 상태.
Sun Cluster (소프트웨어)	SunPlex 시스템에 포함된 소프트웨어 부분. SunPlex 를 참조하십시오.
SunPlex	가용성과 확장성이 높은 서비스를 제공하기 위해 사용하는 통합된 하드웨어 및 Sun Cluster 소프트웨어 시스템.
스위치백	“페일백”을 참조하십시오.
스위치오버	자원 그룹이나 장치 그룹을 클러스터의 한 마스터(노드)에서 다른 마스터(자원 그룹이 여러 1차 노드에 대해 구성된 경우에는 여러 마스터)로 순차적으로 옮기는 작업. 스위치오버는 관리자가 scswitch(1M) 명령을 사용하여 시작합니다.
SSP(System Service Processor)	Enterprise 10000 구성에서 주로 클러스터 구성원과 통신을 하기 위해 클러스터 외부에 사용되는 장치.

T

테이크오버	“페일오버”를 참조하십시오.
터미널 콘센트레이터	Enterprise 10000 이외의 구성에서 주로 클러스터 구성원과 통신을 하기 위해 사용되는 클러스터 외부 장치.

V

VERITAS Volume Manager

볼륨 관리자

SunPlex 시스템에서 사용하는 볼륨 관리자. “볼륨 관리자”를 참조하십시오.

디스크 스트라이핑, 연결, 이중화, 메타 장치나 볼륨의 동적 증가 등을 통해 데이터 신뢰성을 제공하는 소프트웨어 제품.