



Solaris OS용 Sun Cluster 개념 안내서

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

부품 번호: 819-2062-10
2005년 8월, 개정판 A

Copyright 2005 Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 U.S.A. 모든 권리는 저작권자의 소유입니다.

이 제품 또는 문서는 저작권법의 보호를 받으며 그 사용과 복사, 배포 및 디컴파일을 제한하는 라이선스 하에서 배포됩니다. 본 제품 또는 설명서의 어떠한 부분도 Sun 및 Sun 소속 라이선스 부여자(있는 경우)의 사전 서면 승인 없이 어떠한 형태나 수단으로도 재생산할 수 없습니다. 글꼴 기술을 포함한 타사 소프트웨어에 대한 저작권 및 사용권은 Sun 공급업체에 있습니다.

제품 중에는 캘리포니아 대학에서 허가한 Berkeley BSD 시스템에서 파생된 부분이 포함되어 있을 수 있습니다. UNIX는 미국 및 다른 국가에서 X/Open Company, Ltd.를 통해 독점적으로 사용권이 부여되는 등록 상표입니다.

Sun, Sun Microsystems, Sun 로고, docs.sun.com, AnswerBook, AnswerBook2 Sun Cluster, SunPlex, Sun Enterprise, Sun Enterprise 10000, Sun Enterprise SyMON, Sun Management Center, Solaris, Solaris 볼륨 관리자, Sun StorEdge, Sun Fire, SPARCstation, OpenBoot 및 Solaris는 미국 및 다른 국가에서 Sun Microsystems, Inc.의 상표 또는 등록 상표입니다. 모든 SPARC 상표는 사용 허가를 받았으며 미국 및 다른 국가에서 SPARC International, Inc.의 상표 또는 등록 상표입니다. SPARC 상표를 사용하는 제품은 Sun Microsystems, Inc.가 개발한 구조를 기반으로 하고 있습니다. ORACLE, Netscape

Sun Microsystems, Inc.는 사용자 및 사용 허가자를 위해 OPEN LOOK 및 Sun™ GUI(그래픽 사용자 인터페이스)를 개발했습니다. Sun은 컴퓨터 업계를 위한 시각적 그래픽 사용자 인터페이스의 개념을 연구 개발한 Xerox사의 선구적인 노력을 높이 평가하고 있습니다. Sun은 Xerox와 Xerox Graphical User Interface에 대한 비독점적 사용권을 보유하고 있습니다. 이 사용권은 OPEN LOOK GUI를 구현하는 Sun의 정식 사용자에게도 적용되며 그렇지 않은 경우에는 Sun의 서면 사용권 계약을 준수해야 합니다.

미국 정부 권한 - 상용 소프트웨어. 정부 사용자는 Sun Microsystems, Inc. 표준 사용권 계약과 해당 FAR 규정과 보충 규정을 준수해야 합니다.

이 문서에서는 본문의 내용을 "있는 그대로" 제공하며, 법률을 위반하지 않는 범위 내에서 상품성, 특정 목적에 대한 적합성 또는 비침해에 대한 묵시적인 보증을 포함하여 모든 명시적 또는 묵시적 조건, 표현 및 보증을 배제합니다.



050816@12762



목차

머리말 7

1 소개 및 개요 13

Sun Cluster 시스템 소개 13

Sun Cluster 시스템에 대한 세 가지 관점 14

하드웨어 설치 및 서비스 관점 14

시스템 관리자 관점 15

응용 프로그램 개발자 관점 17

Sun Cluster 시스템 작업 18

2 하드웨어 서비스 제공업체를 위한 주요 개념 19

Sun Cluster 시스템 하드웨어 및 소프트웨어 구성 요소 19

클러스터 노드 20

클러스터 하드웨어 구성원용 소프트웨어 구성 요소 21

멀티 호스트 장치 22

Multi-Initiator SCSI 23

로컬 디스크 24

이동식 매체 24

클러스터 상호 연결 24

공용 네트워크 인터페이스 25

클라이언트 시스템 25

콘솔 액세스 장치 25

관리 콘솔 26

SPARC: SPARC용 Sun Cluster 토폴로지 27

SPARC: SPARC용 클러스터 쌍 토폴로지 27

SPARC: SPARC용 쌍+N 토폴로지 28

SPARC: SPARC용 N+1(스타) 토폴로지	29
SPARC: SPARC용 N*N(확장 가능) 토폴로지	30
x86: x86용 Sun Cluster 토폴로지	31
x86: x86용 클러스터 쌍 토폴로지	31

3 시스템 관리자 및 응용 프로그램 개발자를 위한 주요 개념 33

관리 인터페이스	33
클러스터 시간	34
고가용성 프레임워크	35
클러스터 구성원 모니터	36
페일패스트 기법	36
CCR(Cluster Configuration Repository)	37
전역 장치	37
장치 ID 및 DID 의사 드라이버	37
디스크 장치 그룹	38
디스크 장치 그룹 장애 복구	39
멀티 포트 디스크 장치 그룹	40
전역 이름 공간	42
로컬 및 전역 이름 공간 예제	42
클러스터 파일 시스템	43
클러스터 파일 시스템 사용	44
HAStoragePlus 자원 유형	44
Syncdir 마운트 옵션	45
디스크 경로 모니터링	45
DPM 개요	46
디스크 경로 모니터	47
쿼럼 및 쿼럼 장치	48
쿼럼 투표 수 정보	50
장애 차단 정보	50
장애 방지를 위한 페일패스트 기법	51
쿼럼 구성 정보	52
쿼럼 장치 요구 사항 준수	52
쿼럼 장치 모범 사례 준수	53
권장되는 쿼럼 구성	54
비전형적인 쿼럼 구성	56
바람직하지 않은 쿼럼 구성	56
데이터 서비스	58
데이터 서비스 메소드	60

페일오버 데이터 서비스	61
확장 가능 데이터 서비스	61
로드 균형 조정 정책	62
페일백 설정	64
데이터 서비스 결합 모니터	64
새로운 데이터 서비스 개발	65
확장 가능 서비스의 특성	65
Data Service API 및 Data Service Development Library API	66
데이터 서비스 트래픽을 위한 클러스터 상호 연결 사용	66
자원, 자원 그룹 및 자원 유형	67
Resource Group Manager (RGM)	68
자원 및 자원 그룹의 상태와 설정	68
자원 및 자원 그룹 등록 정보	70
데이터 서비스 프로젝트 구성	70
프로젝트 구성에 대한 요구 사항 결정	72
선행 프로세스 가상 메모리 제한 설정	73
페일오버 시나리오	74
공용 네트워크 어댑터 및 IP (Internet Protocol) Network Multipathing	79
SPARC: 동적 재구성 지원	80
SPARC: 동적 재구성 일반 설명	81
SPARC: CPU 장치에 대한 DR 클러스터링 참고 사항	81
SPARC: 메모리에 대한 DR 클러스터링 참고 사항	81
SPARC: 디스크 및 테이프 드라이브에 대한 DR 클러스터링 참고 사항	82
SPARC: 쿼럼 장치에 대한 DR 클러스터링 참고 사항	82
SPARC: 클러스터 상호 연결 인터페이스에 대한 DR 클러스터링 고려 사항	82
SPARC: 공용 네트워크 인터페이스에 대한 DR 클러스터링 참고 사항	83

4 FAQ 85

고가용성 FAQ	85
파일 시스템 FAQ	86
볼륨 관리 FAQ	87
데이터 서비스 FAQ	87
공용 네트워크 FAQ	88
클러스터 구성원 FAQ	89
클러스터 저장소 FAQ	90
클러스터 상호 연결 FAQ	90
클라이언트 시스템 FAQ	91
관리 콘솔 FAQ	91

단말기 집중 장치 및 시스템 서비스 프로세서 FAQ 92

색인 95

머리말

Solaris OS용 Sun™ Cluster 개념 안내서에는 SPARC® 및 x86 기반 시스템 모두의 SunPlex™ 시스템에 대한 개념 및 참조 정보가 있습니다.

주 - 이 문서에서 “x86”이라는 용어는 Intel 마이크로프로세서 칩 32비트 제품군을 말하며 AMD에서 만든 마이크로프로세서 칩과 호환 가능합니다.

SunPlex 시스템에는 Sun의 클러스터 솔루션을 구성하는 모든 하드웨어 및 소프트웨어 구성 요소가 포함되어 있습니다.

이 설명서는 Sun Cluster 소프트웨어에 대한 교육을 받은 전문 시스템 관리자를 대상으로 합니다. 이 문서는 계획이나 관측용 안내서가 아닙니다. 이 문서를 읽을 때는 이미 시스템 요구 사항을 결정하고 필요한 장비와 소프트웨어를 구입한 상태이어야 합니다.

이 설명서에서 설명하는 개념을 이해하려면 Solaris™ 운영 체제에 대한 지식과 Sun Cluster 시스템과 함께 사용되는 볼륨 관리자 소프트웨어에 대한 전문가 수준의 지식이 필요합니다.

주 - Sun Cluster 소프트웨어는 SPARC 및 x86의 두 가지 플랫폼에서 실행됩니다. 이 설명서의 정보는 특정 장, 절, 주, 머리글로 표시된 항목, 그림, 표 또는 예에서 언급된 경우를 제외하고는 두 플랫폼 모두와 관련됩니다.

활자체 규약

다음 표는 이 책에서 사용된 활자체 변경 사항에 대하여 설명합니다.

표 P-1 활자체 규약

서체 또는 기호	의미	예
AaBbCc123	명령, 파일 및 디렉토리의 이름, 그리고 컴퓨터 화면에 출력되는 내용입니다.	.login 파일을 편집하십시오. ls -a 명령을 사용하여 모든 파일을 나열하십시오. machine_name% you have mail.
AaBbCc123	컴퓨터 화면 상의 출력과는 달리 사용자가 직접 입력하는 사항입니다.	machine_name% su Password:
AaBbCc123	명령줄 자리 표시자: 실제 이름이나 값으로 대체됩니다.	파일을 삭제하려면 rm filename을 입력하십시오.
AaBbCc123	책 제목, 새로 나오는 용어, 강조 표시할 단어입니다.	사용자 설명서 의 6장을 읽으십시오. 패치 분석 을 수행하십시오. 파일을 저장하지 마십시오 . [일부 강조된 항목은 온라인에서 볼드로 표시됩니다.]

명령 예의 쉘 프롬프트

C 쉘, Bourne 쉘 및 Korn 쉘에 대한 기본 시스템 프롬프트 및 슈퍼유저 프롬프트는 다음 표와 같습니다.

표 P-2 쉘 프롬프트

셸	프롬프트
C 쉘 프롬프트	machine_name%
C 쉘 슈퍼유저 프롬프트	machine_name#

표 P-2 셸 프롬프트 (계속)

셸	프롬프트
Bourne 셸 및 Korn 셸 프롬프트	\$
Bourne 셸 및 Korn 셸 슈퍼유저 프롬프트	#

관련 문서

Sun Cluster 항목에 대한 정보는 다음 표에 나열된 설명서를 참조하십시오. 모든 Sun Cluster 설명서는 <http://docs.sun.com>에서 볼 수 있습니다.

주제	문서
개요	Solaris OS용 Sun Cluster 개요
개념	Solaris OS용 Sun Cluster 개념 안내서
하드웨어 설치 및 관리	Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS 개별 하드웨어 관리 설명서
소프트웨어 설치	Solaris OS용 Sun Cluster 소프트웨어 설치 안내서
데이터 서비스 설치 및 관리	Sun Cluster Data Services Planning and Administration Guide for Solaris OS 개별 데이터 서비스 설명서
데이터 서비스 개발	Solaris OS용 Sun Cluster 데이터 서비스 개발 안내서
시스템 관리	Solaris OS용 Sun Cluster 시스템 관리 안내서
오류 메시지	Sun Cluster Error Messages Guide for Solaris OS
명령 및 함수 참조	Sun Cluster Reference Manual for Solaris OS

Sun Cluster 전체 설명서 목록은 <http://docs.sun.com>에서 해당 Sun Cluster 소프트웨어 릴리스의 릴리스 노트를 참조하십시오.

Sun 문서 온라인 액세스

docs.sun.comSM 웹 사이트에서 Sun 기술 관련 문서를 온라인으로 이용할 수 있습니다. 다음 주소에서 docs.sun.com 아카이브를 살펴보고 특정 서적 제목이나 주제에 대해 검색할 수 있습니다. URL은 <http://docs.sun.com>입니다.

Sun 설명서 주문

Sun Microsystems에서는 제품 설명서를 인쇄물로 제공합니다. 문서 목록과 주문 방법은 <http://docs.sun.com>의 “인쇄본 문서를 구입하십시오”를 참조하십시오.

지원 받기

Sun Cluster 시스템 설치 및 사용에 문제가 있으면 서비스 담당자에게 문의하십시오. 문의할 때 다음 정보가 필요합니다.

- 이름 및 전자 메일 주소(있을 경우)
- 회사 이름, 주소 및 전화 번호
- 시스템 모델 및 일련 번호
- 운영 환경의 릴리스 번호(예: Solaris 9)
- Sun Cluster 소프트웨어의 릴리스 번호(예: 3.1 8/05)

다음 명령을 사용하여 서비스 담당자에게 제공할 시스템의 각 노드에 대한 정보를 수집합니다.

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>SPARC: prtdiag -v</code>	시스템 진단 정보를 표시합니다.
<code>scinstall -pv</code>	Sun Cluster 소프트웨어 릴리스 및 패키지 버전 정보를 표시합니다.

명령	기능
scstat	클러스터 상태에 대한 스냅샷을 제공합니다.
scconf -p	클러스터 구성 정보를 나열합니다.
scrgadm -p	설치된 자원, 자원 그룹 및 자원 유형에 대한 정보를 표시합니다.

/var/adm/messages 파일의 내용도 준비하십시오.

제품 교육

Sun Microsystems에서는 강의식 교육 과정과 자습식 교육 과정 등 다양한 교육 과정을 통하여 여러 Sun 기술 분야의 교육을 제공하고 있습니다. Sun에서 제공하는 교육 과정과 등록에 대한 자세한 내용을 보려면 <http://training.sun.com/>의 Sun Microsystems 교육을 방문하십시오.

소개 및 개요

Sun Cluster 시스템은 가용성이 높고 확장 가능한 서비스를 제공할 수 있도록 통합된 하드웨어 및 Sun Cluster 소프트웨어 솔루션입니다.

Solaris OS용 Sun Cluster 개념 안내서에서는 Sun Cluster 설명서의 주요 사용자에게 필요한 개념에 대해 설명합니다. 이 안내서의 대상은 다음과 같습니다.

- 클러스터 하드웨어를 설치하고 서비스를 제공하는 서비스 제공업체
- Sun Cluster 소프트웨어를 설치, 구성 및 관리하는 시스템 관리자
- Sun Cluster 제품에 현재 포함되지 않은 응용 프로그램의 페일오버 및 확장 가능 서비스를 개발하는 응용 프로그램 개발자

Sun Cluster 시스템에 대한 전체적인 개요를 보려면 전체 Sun Cluster 설명서 세트와 함께 이 설명서를 참조하십시오.

이 장에서는

- Sun Cluster 시스템을 소개하고 높은 수준의 개요를 제공합니다.
- Sun Cluster 사용 대상의 여러 관점에 대해 설명합니다.
- Sun Cluster 시스템을 사용하기 전에 알아야 할 주요 개념을 설명합니다.
- 절차와 관련 정보가 포함된 Sun Cluster 설명서와 주요 개념을 매핑합니다.
- 클러스터 관련 작업을 각 작업 수행 절차가 포함된 문서에 연결합니다.

Sun Cluster 시스템 소개

Sun Cluster 시스템은 Solaris 운영 체제를 클러스터 운영 체제로 확장합니다. 클러스터 또는 플렉스는 데이터베이스, 웹 서비스 및 파일 서비스를 포함한 네트워크 서비스나 응용 프로그램을 단일 클라이언트 환경으로 만들어주는 느슨한 결합의 컴퓨팅 노드 모음입니다.

각 클러스터 노드는 독립적으로 자체 프로세스를 실행하는 서버입니다. 이 프로세스는 다른 프로세스와 통신을 통해 사용자에게 응용 프로그램, 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, 페일오버 및 확장 가능한 서비스 지원, 모듈 단위로 확장할 수 있는 용량, 기존 하드웨어의 장애 복구 시스템에 비해 저렴한 항목 가격 등의 장점이 있습니다.

Sun Cluster 시스템의 목표는 다음과 같습니다.

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

장애 복구와 고가용성에 대한 자세한 내용은 **Solaris OS용 Sun Cluster 개요**의 “Sun Cluster를 사용하여 가용성이 높은 응용 프로그램 만들기”를 참조하십시오.

고가용성에 대한 질문과 대답은 **85 페이지 “고가용성 FAQ”**를 참조하십시오.

Sun Cluster 시스템에 대한 세 가지 관점

이 절에서는 Sun Cluster 시스템에 대한 서로 다른 세 가지 관점과 주요 개념 및 각 관점과 관련된 설명서에 대해 설명합니다. 이러한 관점을 대표하는 전문가 그룹은 다음과 같습니다.

- 하드웨어 설치 및 서비스 담당자
- 시스템 관리자
- 응용 프로그램 개발자

하드웨어 설치 및 서비스 관점

하드웨어 서비스 담당자의 경우에는 Sun Cluster 시스템을 서버, 네트워크 및 저장소가 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

추가 하드웨어 개념 정보

다음 절에는 위의 주요 개념과 관련된 자료가 있습니다.

- 20 페이지 "클러스터 노드"
- 22 페이지 "멀티 호스트 장치"
- 24 페이지 "로컬 디스크"
- 24 페이지 "클러스터 상호 연결"
- 25 페이지 "공용 네트워크 인터페이스"
- 25 페이지 "클라이언트 시스템"
- 26 페이지 "관리 콘솔"
- 25 페이지 "콘솔 액세스 장치"
- 27 페이지 "SPARC: SPARC용 클러스터 쌍 토폴로지"
- 29 페이지 "SPARC: SPARC용 N+1(스타) 토폴로지"

하드웨어 담당자를 위한 Sun Cluster 설명서

다음 Sun Cluster 설명서에는 하드웨어 서비스 개념과 관련된 절차 및 정보가 있습니다.

Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS

시스템 관리자 관점

시스템 관리자의 경우에는 Sun Cluster 시스템이 저장 장치를 공유하는 케이블로 연결된 서버(노드) 세트라고 생각할 수 있습니다. 시스템 관리자는 특수 작업을 수행하는 소프트웨어를 확인할 수 있습니다.

- 클러스터 노드 사이의 연결을 모니터링할 수 있도록 Solaris 소프트웨어에 통합된 특수 클러스터 소프트웨어
- 클러스터 노드에서 실행되는 사용자 응용 프로그램의 상태를 모니터링할 수 있는 특수 소프트웨어

- 디스크를 설정하고 관리하는 볼륨 관리 소프트웨어
- 직접 디스크에 연결되지는 않지만 모든 노드가 모든 저장 장치에 액세스할 수 있도록 하는 특수 클러스터 소프트웨어
- 노드에 로컬로 연결된 경우에도 모든 노드에 파일을 표시할 수 있도록 하는 특수 클러스터 소프트웨어

주요 개념 - 시스템 관리

시스템 관리자는 다음 개념과 프로세스를 이해해야 합니다.

- 하드웨어 및 소프트웨어 구성 요소 사이의 상호 작용
- 클러스터 설치 및 구성 방법에 대한 일반적인 흐름
 - Solaris 운영 체제 설치
 - Sun Cluster 소프트웨어 설치 및 구성
 - 볼륨 관리자 설치 및 구성
 - 클러스터에서 사용할 수 있도록 응용 프로그램 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

추가 시스템 관리자 개념 정보

다음 절에는 위의 주요 개념과 관련된 자료가 있습니다.

- 33 페이지 "관리 인터페이스"
- 34 페이지 "클러스터 시간"
- 35 페이지 "고가용성 프레임워크"
- 37 페이지 "전역 장치"
- 38 페이지 "디스크 장치 그룹"
- 42 페이지 "전역 이름 공간"
- 43 페이지 "클러스터 파일 시스템"
- 45 페이지 "디스크 경로 모니터링"
- 50 페이지 "장애 차단 정보"
- 58 페이지 "데이터 서비스"

시스템 관리자를 위한 Sun Cluster 설명서

다음 Sun Cluster 설명서에는 시스템 관리 개념과 관련된 절차와 정보가 있습니다.

- Solaris OS용 Sun Cluster 소프트웨어 설치 안내서
- Solaris OS용 Sun Cluster 시스템 관리 안내서
- Sun Cluster Error Messages Guide for Solaris OS

- Solaris OS용 Sun Cluster 3.1 8/05 릴리스 노트
- Sun Cluster 3.0-3.1 Release Notes Supplement

응용 프로그램 개발자 관점

Sun Cluster 시스템은 Oracle, NFS, DNS, Sun™ Java System Web Server, Apache 웹 서버(SPARC 기반 시스템) 및 Sun Java 시스템 디렉토리 서버 등의 응용 프로그램에서 사용할 **데이터 서비스**를 제공합니다. 별도 구입한 응용 프로그램이 Sun Cluster 소프트웨어의 제어를 통해 실행되도록 구성하면 데이터 서비스가 완성됩니다. Sun Cluster 소프트웨어에는 응용 프로그램을 시작, 중지, 모니터링하는 구성 파일과 관리 메소드가 있습니다. 새 페일오버나 확장 가능 서비스를 만들어야 하는 경우 Sun Cluster API(Application Programming Interface) 및 DSET API(Data Service Enabling Technologies API)를 사용하여 해당 응용 프로그램을 클러스터에서 데이터 서비스로 실행하는 데 필요한 구성 파일과 관리 메소드를 개발할 수 있습니다.

주요 개념 - 응용 프로그램 개발

응용 프로그램 개발자는 다음 사항을 이해해야 합니다.

- 사용하는 응용 프로그램이 페일오버 서비스나 확장 가능한 데이터 서비스로 실행될 수 있는지를 판단하는 응용 프로그램 특성
- Sun Cluster API, DSET API 및 “일반” 데이터 서비스개발자는 클러스터 환경에 맞게 응용 프로그램을 구성하기 위해 프로그램이나 스크립트를 작성할 때 사용할 가장 적합한 도구를 결정해야 합니다.

추가 응용 프로그램 개발자 개념 정보

다음 절에는 위의 주요 개념과 관련된 자료가 있습니다.

- 58 페이지 “데이터 서비스”
- 67 페이지 “자원, 자원 그룹 및 자원 유형”
- 4 장

응용 프로그램 개발자를 위한 Sun Cluster 설명서

다음 Sun Cluster 설명서에는 응용 프로그램 개발자 개념과 관련된 절차 및 정보가 있습니다.

- Solaris OS용 Sun Cluster 데이터 서비스 개발 안내서
- Sun Cluster Data Services Planning and Administration Guide for Solaris OS

Sun Cluster 시스템 작업

모든 Sun Cluster 시스템 작업에는 개념에 대한 약간의 배경 지식이 필요합니다. 다음 표에 높은 수준의 작업과 작업 단계를 설명하는 문서 목록이 있습니다. 이 책의 개념 절에서는 개념과 작업 사이의 매핑 관계를 설명합니다.

표 1-1 작업 맵: 문서에 사용자 작업 매핑

작업	지침
클러스터 하드웨어 설치	Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS
클러스터에 Solaris 소프트웨어 설치	Solaris OS용 Sun Cluster 소프트웨어 설치 안내서
SPARC: Sun™ Management Center 소프트웨어 설치	Solaris OS용 Sun Cluster 소프트웨어 설치 안내서
Sun Cluster 소프트웨어 설치 및 구성	Solaris OS용 Sun Cluster 소프트웨어 설치 안내서
볼륨 관리 소프트웨어 설치 및 구성	Solaris OS용 Sun Cluster 소프트웨어 설치 안내서 볼륨 관리 문서
Sun Cluster 데이터 서비스 설치 및 구성	Sun Cluster Data Services Planning and Administration Guide for Solaris OS
서비스 클러스터 하드웨어	Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS
Sun Cluster 소프트웨어 관리	Solaris OS용 Sun Cluster 시스템 관리 안내서
볼륨 관리 소프트웨어 관리	Solaris OS용 Sun Cluster 시스템 관리 안내서 및 사용자의 볼륨 관리 설명서
응용 프로그램 소프트웨어 관리	응용 프로그램 문서
문제 식별 및 권장하는 사용자 조치	Sun Cluster Error Messages Guide for Solaris OS
새 데이터 서비스 만들기	Solaris OS용 Sun Cluster 데이터 서비스 개발 안내서

하드웨어 서비스 제공업체를 위한 주요 개념

이 장에서는 Sun Cluster 시스템을 구성하는 하드웨어 구성 요소와 관련된 주요 개념에 대해 설명합니다. 주요 내용은 다음과 같습니다.

- 20 페이지 "클러스터 노드"
- 22 페이지 "멀티 호스트 장치"
- 24 페이지 "로컬 디스크"
- 24 페이지 "이동식 매체"
- 24 페이지 "클러스터 상호 연결"
- 25 페이지 "공용 네트워크 인터페이스"
- 25 페이지 "클라이언트 시스템"
- 25 페이지 "콘솔 액세스 장치"
- 26 페이지 "관리 콘솔"
- 27 페이지 "SPARC: SPARC용 Sun Cluster 토폴로지"
- 31 페이지 "x86: x86용 Sun Cluster 토폴로지"

Sun Cluster 시스템 하드웨어 및 소프트웨어 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
- 멀티 호스트 저장소(노드 사이의 공유 디스크)
- 이동식 매체(테이프 및 CD-ROM)

- 클러스터 상호 연결
- 공용 네트워크 인터페이스
- 클라이언트 시스템
- 관리 콘솔
- 콘솔 액세스 장치

Sun Cluster 시스템에서는 이 구성 요소를 결합하여 다양한 구성을 만들 수 있습니다. 다음 절에서는 이 구성 요소에 대해 설명합니다.

- 27 페이지 “SPARC: SPARC용 Sun Cluster 토폴로지”
- 31 페이지 “x86: x86용 Sun Cluster 토폴로지”

2-노드 클러스터 구성을 그림으로 설명한 예는 **Sun Cluster Overview for Solaris OS의 Solaris OS용 Sun Cluster 개요**의 “Sun Cluster 하드웨어 환경”를 참조하십시오.

클러스터 노드

클러스터 노드는 Solaris 운영 체제와 Sun Cluster 소프트웨어를 모두 실행하는 시스템입니다. 또한 클러스터 노드는 클러스터의 현재 구성원(클러스터 구성원)이거나 잠재적인 구성원입니다.

- SPARC: Sun Cluster 소프트웨어는 클러스터에서 한 개부터 16개까지의 노드를 지원합니다. 지원되는 노드 구성을 보려면 27 페이지 “SPARC: SPARC용 Sun Cluster 토폴로지”를 참조하십시오.
- x86: Sun Cluster 소프트웨어는 클러스터에서 두 개의 노드를 지원합니다. 지원되는 노드 구성을 보려면 31 페이지 “x86: x86용 Sun Cluster 토폴로지”를 참조하십시오.

클러스터 노드는 일반적으로 하나 이상의 멀티 호스트 장치에 연결됩니다. 멀티 호스트 장치에 연결되지 않은 노드는 클러스터 파일 시스템을 사용하여 멀티 호스트 장치에 액세스합니다. 예를 들어, 하나의 확장 가능한 서비스를 구성하면 멀티 호스트 장치에 직접 연결되지 않아도 노드가 서비스 요청에 응할 수 있습니다.

또한 병렬 데이터베이스 구성에 있는 노드는 모든 디스크를 동시에 액세스합니다.

- 디스크를 동시에 액세스하는 방법에 대한 자세한 내용은 22 페이지 “멀티 호스트 장치”를 참조하십시오.
- 병렬 데이터베이스 구성에 대한 자세한 내용은 27 페이지 “SPARC: SPARC용 클러스터 쌍 토폴로지” 및 31 페이지 “x86: x86용 클러스터 쌍 토폴로지”를 참조하십시오.

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

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

클러스터 구성원은 하나 이상의 물리적으로 독립된 네트워크를 통해 클러스터의 다른 노드와 통신합니다. 이렇게 물리적으로 독립된 네트워크 세트를 **클러스터 상호 연결**이라고 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. 페일오버가 발생할 수 있으므로 모든 노드는 백업 또는 보조 노드 역할을 하게 될 모든 노드의 작업 로드를 지원할 수 있을 만큼 충분한 여분의 용량을 확보해 두어야 합니다.

각 노드는 개별 루트(/) 파일 시스템을 부트합니다.

클러스터 하드웨어 구성원용 소프트웨어 구성 요소

어떤 노드가 클러스터 구성원 역할을 하려면 다음 소프트웨어가 설치되어 있어야 합니다.

- Solaris Operating System
- Sun Cluster 소프트웨어
- 데이터 서비스 응용 프로그램
- 볼륨 관리(Solaris 볼륨 관리자™ 또는 VERITAS Volume Manager)
하드웨어 RAID(Redundant Array of Independent Disks)를 사용하는 구성은 예외입니다. 이 구성에는 Solaris 볼륨 관리자 또는 VERITAS Volume Manager 등의 소프트웨어 볼륨 관리자가 필요하지 않을 수 있습니다.
- Solaris 운영 체제, Sun Cluster 및 볼륨 관리 소프트웨어를 설치하는 방법에 대한 자세한 내용은 **Solaris OS용 Sun Cluster 소프트웨어 설치 안내서**를 참조하십시오.
- 데이터 서비스를 설치하고 구성하는 방법에 대한 자세한 내용은 **Sun Cluster Data Services Planning and Administration Guide for Solaris OS**를 참조하십시오.
- 이전 소프트웨어의 구성 요소에 대한 개념 정보는 **3 장**을 참조하십시오.

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

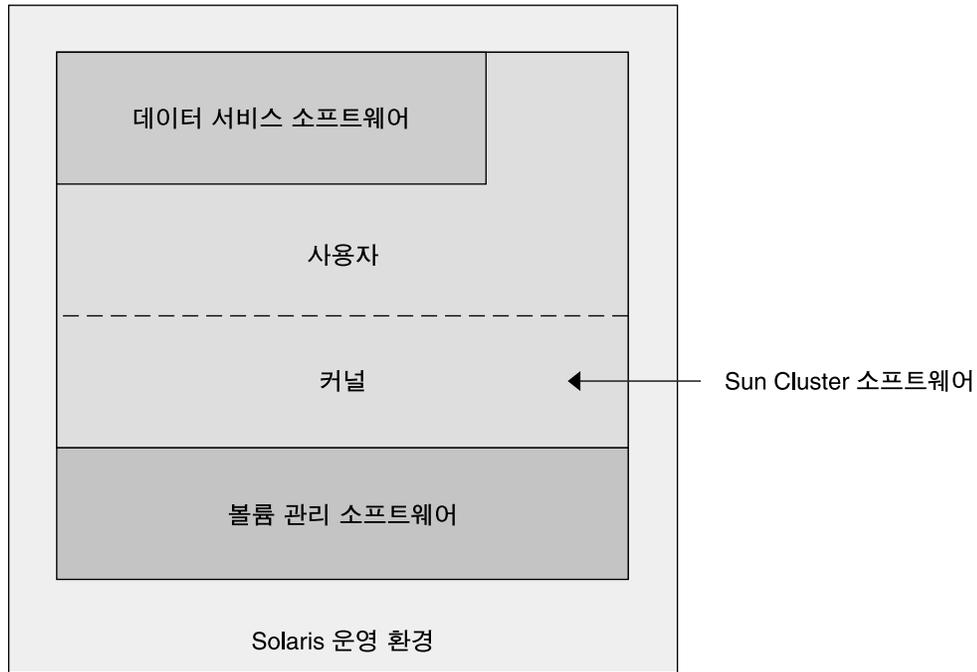


그림 2-1 높은 수준의 Sun Cluster 소프트웨어 구성 요소 관계

클러스터 구성원에 대한 질문과 대답은 4 장을 참조하십시오.

멀티 호스트 장치

한 번에 여러 노드에 연결될 수 있는 디스크는 멀티 호스트 장치입니다. Sun Cluster 환경에 멀티 호스트 저장소를 사용하면 디스크 가용성을 높일 수 있습니다. Sun Cluster 소프트웨어에서 2-노드 클러스터가 퀴럼을 설정하려면 멀티 호스트 저장소가 필요합니다. 3-노드 이상의 클러스터에는 퀴럼 장치가 필요하지 않습니다. 퀴럼에 대한 자세한 내용은 48 페이지 “퀴럼 및 퀴럼 장치”를 참조하십시오.

멀티 호스트 장치에는 다음과 같은 기능이 있습니다.

- 단일 노드 장애 복구.
- 응용 프로그램 데이터, 응용 프로그램 이진 및 구성 파일 저장 기능.
- 노드 장애로부터 보호. 클라이언트가 데이터를 요청한 노드에 장애가 생기면 해당 요청은 동일한 디스크에 직접 연결된 다른 노드를 사용하도록 전환됩니다.
- 디스크를 “마스터하는” 기본 노드를 통한 전역 액세스 또는 로컬 경로를 통한 직접 동시 액세스. 현재 직접 동시 액세스를 사용하는 응용 프로그램은 Oracle Real Application Clusters Guard뿐입니다.

볼륨 관리자는 멀티 호스트 장치의 데이터 중복에 대해 미러링 구성 또는 RAID-5 구성을 제공합니다. 현재 Sun Cluster는 SPARC 기반 클러스터에서만 사용할 수 있는 Solaris 볼륨 관리자와 VERITAS Volume Manager를 볼륨 관리자로 지원하며 여러 하드웨어 RAID 플랫폼에서 RDAC RAID-5 하드웨어 컨트롤러를 지원합니다.

멀티 호스트 장치에 디스크 미러링과 디스크 스트라이핑을 결합하면 노드 장애와 개별 디스크 장애로부터 보호할 수 있습니다.

멀티 호스트 저장소에 대한 질문과 대답은 4 장을 참조하십시오.

Multi-Initiator SCSI

이 절의 내용은 SCSI 저장소에만 적용되고 멀티 호스트 장치에 사용되는 광섬유 채널 저장소에는 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 *initiator*라고 합니다. 이 회로가 SCSI 버스에 대한 모든 버스 작업을 시작합니다. Sun 시스템에서 SCSI 호스트 어댑터의 기본 SCSI 주소는 7입니다.

클러스터 구성은 멀티 호스트 장치를 사용하여 여러 서버 노드 사이에서 저장소를 공유합니다. 클러스터 저장소가 단일 종결 SCSI 장치나 차동 SCSI 장치로 구성된 경우 이러한 구성을 Multi-initiator SCSI라고 합니다. 이 용어가 의미하는 것처럼 SCSI 버스에는 둘 이상의 SCSI initiator가 있습니다.

SCSI 사양에서는 SCSI 버스에 있는 각 장치에 고유한 SCSI 주소가 있어야 합니다. (호스트 어댑터 역시 SCSI 버스의 장치입니다). Multi-initiator 환경에서 기본 하드웨어 구성을 적용하면 모든 SCSI 호스트 어댑터의 기본값이 7이 되므로 충돌이 발생합니다.

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. 이후 사용되지 않을 주소의 예로는 새로운 드라이브를 빈 드라이브 슬롯에 설치하여 저장소를 추가할 경우가 있습니다.

대부분의 구성에서 보조 호스트 어댑터의 SCSI 주소로 6을 사용할 수 있습니다.

이 호스트 어댑터용으로 선택된 SCSI 주소를 변경하려면 다음 도구 중 하나를 사용하여 `scsi-initiator-id` 등록 정보를 설정합니다.

- `eeprom(1M)`
- SPARC 기반 시스템의 OpenBoot PROM
- x86 기반 시스템에서 BIOS 부트 이후 선택적으로 실행하는 SCSI 유틸리티

이 등록 정보를 노드에 대하여 전역으로 설정할 수도 있고 호스트 어댑터 단위로 설정할 수도 있습니다. 각 SCSI 호스트 어댑터에 대한 고유한 `scsi-initiator-id`를 설정하는 지침은 **Sun Cluster 3.0-3.1 With SCSI JBOD Storage Device Manual for Solaris OS**에 있습니다.

로컬 디스크

로컬 디스크는 단일 노드에만 연결되는 디스크입니다. 따라서 로컬 디스크는 노드 장애로부터 보호되지 않습니다(가용성이 높지 않음). 그러나 로컬 디스크를 포함한 모든 디스크는 전역 이름 공간에 포함되어 있어서 **전역 장치**로 구성됩니다. 그러므로 디스크 자체는 모든 클러스터 노드에서 볼 수 있습니다.

로컬 디스크의 파일 시스템을 전역 마운트 지점 아래에 놓아서 다른 노드에서 사용할 수 있도록 할 수 있습니다. 현재 이러한 전역 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 미리할 수 있지만, 노드 장애로부터 보호하지는 않습니다.

전역 장치에 대한 자세한 내용은 37 페이지 “전역 장치” 절을 참조하십시오.

이동식 매체

테이프 드라이브 및 CD-ROM 드라이브와 같은 이동식 매체가 클러스터에서 지원됩니다. 일반적으로 이런 장치는 클러스터되지 않은 환경과 동일한 방법으로 설치하고 구성하며 서비스를 제공합니다. 이런 장치는 Sun Cluster에 전역 장치로 구성되므로 클러스터의 모든 노드에서 각 장치에 액세스할 수 있습니다. 이동식 매체를 설치하고 구성하는 방법에 대한 자세한 내용은 **Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS**를 참조하십시오.

전역 장치에 대한 자세한 내용은 37 페이지 “전역 장치” 절을 참조하십시오.

클러스터 상호 연결

클러스터 상호 연결은 클러스터 노드 사이에 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 액세스한다고 가정합니다.

단일 실패 지점이 발생하지 않도록 물리적으로 독립된 둘 이상의 중복 네트워크 또는 경로를 통해 클러스터를 상호 연결하여 모든 노드를 연결해야 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 - 각 클러스터 노드에 상주하는 네트워크 인터페이스 카드. 이름은 qfe2와 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

저장소 인터페이스가 모두 포함되어 있습니다.

인터페이스가 여러 개인 네트워크 어댑터는 전체 어댑터에 장애가 발생할 경우 단일 장애 지점이 될 수 있습니다. 최대 가용성을 위해서 두 노드 사이의 경로가 단일 네트워크 어댑터에만 의존하지 않도록 클러스터를 계획하십시오.

- 연결 - 클러스터 노드의 외부에 있는 스위치. 연결 장치는 바로 전달 기능과 전환 기능을 수행하여 둘 이상의 노드를 연결할 수 있게 합니다. 2-노드 클러스터에서는 노드들이 각 노드에서 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 또는 어댑터와 연결 사이에 설치되는 물리적 연결.

클러스터 상호 연결에 대한 질문과 대답은 4 장을 참조하십시오.

공용 네트워크 인터페이스

클라이언트는 공용 네트워크 인터페이스를 통해 클러스터에 연결합니다. 각 네트워크 어댑터 카드는 카드에 여러 하드웨어 인터페이스가 있는지에 따라 하나 이상의 공용 네트워크에 연결할 수 있습니다. 여러 카드가 활성화되어 서로 페일오버 백업 역할을 하도록 구성된 여러 공용 네트워크 인터페이스 카드를 포함하도록 노드를 설정할 수 있습니다. 어댑터 중 하나가 실패하면 IP (Internet Protocol) Network Multipathing 소프트웨어가 호출되어 오류가 있는 인터페이스를 그룹에 있는 다른 어댑터로 페일오버합니다.

클러스터링에서 공용 네트워크 인터페이스를 위하여 특별히 하드웨어를 고려할 필요는 없습니다.

공용 네트워크에 대한 질문과 대답은 4 장을 참조하십시오.

클라이언트 시스템

클라이언트 시스템은 공용 네트워크를 통해 클러스터에 액세스하는 다른 서버나 워크스테이션을 포함합니다. 클라이언트측 프로그램은 클러스터에서 실행 중인 서버측 응용 프로그램이 제공하는 데이터나 다른 서비스를 사용합니다.

클라이언트 시스템은고가용성 시스템이 아닙니다. 클러스터의 데이터 및 응용 프로그램은고가용성입니다.

클라이언트 시스템에 대한 질문과 대답은 4 장을 참조하십시오.

콘솔 액세스 장치

모든 클러스터 노드에 대하여 콘솔 액세스가 있어야 합니다. 콘솔 액세스를 얻으려면 다음 장치 중 하나를 사용합니다.

- 클러스터 하드웨어와 함께 구입한 단말기 집중 장치

- Sun Enterprise E10000 서버의 SSP(System Service Processor)(SPARC 기반 클러스터용)
- Sun Fire™ 서버의 시스템 컨트롤러(SPARC 기반 클러스터용)
- 각 노드의 ttya를 액세스할 수 있는 다른 장치

Sun에서는 사용할 수 있는 단말기 집중 장치가 하나만 지원됩니다. 지원되는 Sun 단말기 집중 장치를 사용할 것인지는 사용자가 선택할 수 있습니다. 단말기 집중 장치를 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 E1000 서버에 대한 콘솔 액세스를 제공합니다. SSP는 Sun Enterprise E1000 서버를 지원하기 위해 구성된 이더넷 네트워크상의 시스템입니다. SSP는 Sun Enterprise E1000 서버의 관리 콘솔입니다. Sun Enterprise E10000 네트워크 콘솔 기능을 사용하면 네트워크 상의 어떤 워크스테이션에서도 호스트 콘솔 세션을 열 수 있습니다.

기타 콘솔 액세스 메소드에는 다른 단말기 집중 장치, 다른 노드로부터의 tip(1) 직렬 포트 액세스 및 단순 단말기 등이 포함됩니다. Sun™ 키보드 및 모니터를 사용하거나 하드웨어 서비스 제공업체에서 지원하는 다른 직렬 포트 장치를 사용할 수 있습니다.

관리 콘솔

관리 콘솔로 알려진 전용 UltraSPARC® 워크스테이션이나 Sun Fire V65x 서버를 사용하여 활성 클러스터를 관리할 수 있습니다. 일반적으로 관리 콘솔에는 CCP(Cluster Control Panel)와 Sun Management Center 제품의 Sun Cluster 모듈(SPARC 기반 클러스터 전용)과 같은 관리 도구 소프트웨어를 설치하여 실행합니다. CCP에서 cconsole을 사용하면 하나 이상의 노드 콘솔을 한 번에 연결할 수 있습니다. CCP 사용에 대한 자세한 내용은 **Solaris OS용 Sun Cluster 시스템 관리 안내서**의 1 장, “Sun Cluster 관리 소개”를 참조하십시오.

관리 콘솔은 클러스터 노드가 아닙니다. 공용 네트워크를 통해 또는 선택적으로 네트워크 기반 단말기 집중 장치를 사용하여 클러스터 노드에 원격 액세스할 때 관리 콘솔을 사용합니다. 클러스터가 Sun Enterprise E10000 플랫폼으로 구성된 경우에는 관리 콘솔에서 SSP로 로그인하고 netcon(1M) 명령을 사용하여 연결해야 합니다.

일반적으로, 노드를 모니터 없이 구성합니다. 그런 다음 관리 콘솔에서 telnet 세션을 통해 노드의 콘솔에 액세스합니다. 관리 콘솔은 단말기 집중 장치에 연결되고 단말기 집중 장치는 노드의 직렬 포트에 연결됩니다. Sun Enterprise E1000 서버는 SSP(System Service Processor)에서 연결합니다. 자세한 내용은 25 페이지 “콘솔 액세스 장치”를 참조하십시오.

Sun Cluster에는 전용 관리 콘솔이 필요하지 않지만 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 콘솔에 대한 질문과 대답은 4 장을 참조하십시오.

SPARC: SPARC용 Sun Cluster 토폴로지

토폴로지는 클러스터 노드를 Sun Cluster 환경에서 사용되는 저장소 플랫폼에 연결하는 연결 체계입니다. Sun Cluster 소프트웨어는 다음 지침을 준수하는 모든 토폴로지를 지원합니다.

- SPARC 기반 시스템으로 구성된 Sun Cluster 환경에서는 구현하는 저장소 구성에 상관없이 클러스터에서 최대 16개의 노드를 지원합니다.
- 공유 저장 장치는 저장 장치에서 지원하는 만큼의 노드에 연결될 수 있습니다.
- 공유 저장 장치를 클러스터의 모든 노드에 연결할 필요는 없지만 그러나 이러한 저장 장치는 최소 두 개의 노드에 연결해야 합니다.

Sun Cluster 소프트웨어에서는 특정 토폴로지를 사용하여 클러스터를 구성할 필요가 없습니다. 다음 토폴로지는 클러스터의 연결 체계를 설명하는 데 필요한 핵심 내용을 제공합니다. 이러한 토폴로지는 일반 연결 체계입니다.

- 클러스터 쌍
- 쌍+N
- N+1(스타)
- N*N(확장 가능)

다음 절에 각 토폴로지의 예를 보여주는 그림이 있습니다.

SPARC: SPARC용 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 사용하여 하나의 쌍에서 병렬 데이터베이스 응용 프로그램을 실행하고 다른 쌍에서 페일오버 또는 고가용성 응용 프로그램을 실행할 수 있습니다.

클러스터 파일 시스템을 사용하면 두 쌍의 구성을 가질 수도 있습니다. 모든 노드가 응용 프로그램 데이터를 저장하는 디스크에 직접 연결되어 있지 않더라도 세 개 이상의 노드에서 확장 가능 서비스나 병렬 데이터베이스를 실행할 수 있습니다.

다음 그림은 클러스터된 쌍 구성을 보여줍니다.

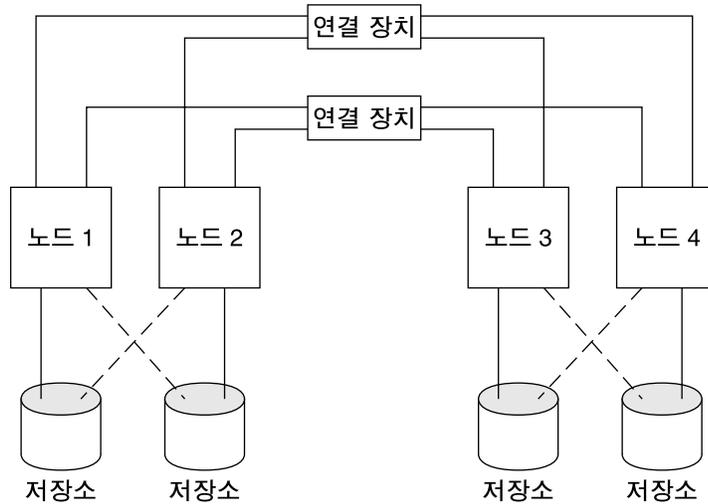


그림 2-2 SPARC: 클러스터 쌍 토폴로지

SPARC: SPARC용 쌍+N 토폴로지

쌍+N 토폴로지에는 공유 저장소 및 추가 노드 세트에 직접 연결된 노드의 쌍이 포함되어 있습니다. 이 추가 노드 세트는 자체적으로 직접 연결하지 않고 클러스터 상호 연결을 사용하여 공유 저장소에 액세스합니다.

다음 그림은 네 개의 노드 중 두 개의 노드(노드 3 및 노드 4)가 클러스터 상호 연결을 사용하여 저장소에 액세스하는 쌍+N 토폴로지의 예입니다. 이 구성은 공유 저장소에 직접 액세스하지 않은 추가 노드를 포함하도록 확장될 수 있습니다.

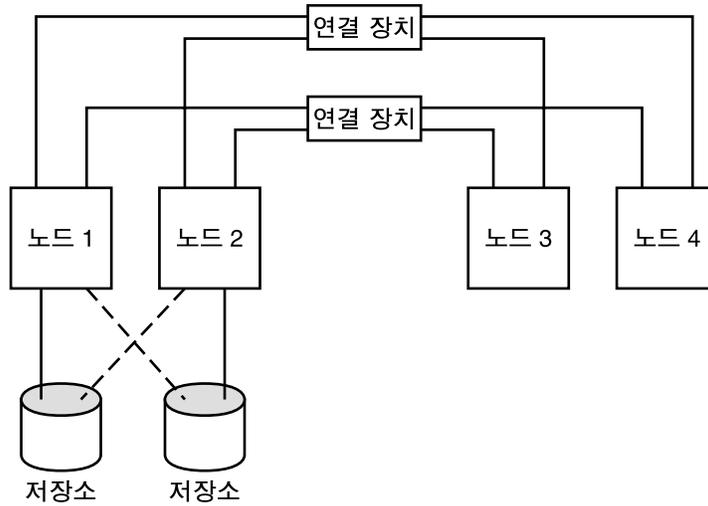


그림 2-3 쌍+N 토폴로지

SPARC: SPARC용 N+1(스타) 토폴로지

N+1 토폴로지에는 몇 개의 기본 노드와 하나의 보조 노드가 들어 있습니다. 기본 노드와 보조 노드를 동일하게 구성할 필요는 없습니다. 항상 기본 노드가 응용 프로그램 서비스를 제공합니다. 보조 노드는 기본 노드의 실패가 있을 때까지 비활동 상태일 필요는 없습니다.

보조 노드는 이러한 구성에서 모든 멀티 호스트 저장소에 물리적으로 연결된 유일한 노드입니다.

기본 노드에서 장애가 발생하면 Sun Cluster가 자원을 보조 노드로 페일오버합니다. 여기서 자원은 기본 노드로 다시 전환(자동 또는 수동으로)될 때까지 기능을 수행합니다.

보조 노드는 기본 노드 중 하나에 장애가 발생할 경우에 부하를 처리할 수 있을 만큼 충분한 CPU 용량이 있어야 합니다.

다음 그림은 N+1 구성의 예입니다.

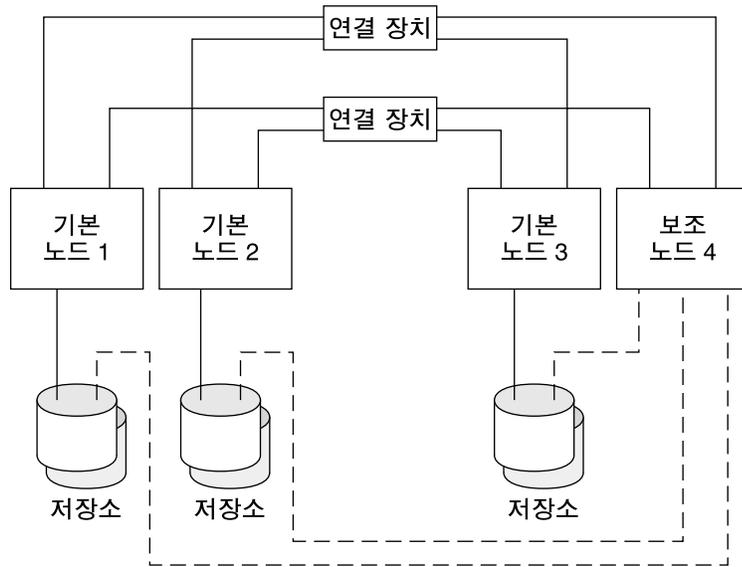


그림 2-4 SPARC: N+1 토폴로지

SPARC: SPARC용 N*N(확장 가능) 토폴로지

N*N 토폴로지에서는 클러스터의 모든 공유 저장 장치를 클러스터의 모든 노드에 연결할 수 있습니다. 이 토폴로지에서는고가용성 응용 프로그램을 서비스 성능의 저하 없이 한 노드에서 다른 노드로 페일오버할 수 있습니다. 페일오버가 발생하면 새 노드는 개별 상호 연결 대신 로컬 경로를 사용하여 저장 장치에 액세스할 수 있습니다.

다음 그림은 N*N 구성을 보여줍니다.

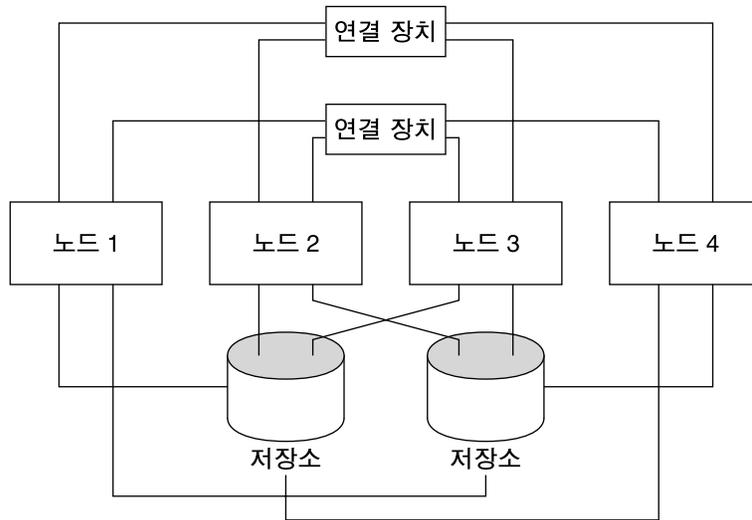


그림 2-5 SPARC: N*N 토폴로지

x86: x86용 Sun Cluster 토폴로지

토폴로지는 클러스터 노드를 클러스터에서 사용하는 저장소 플랫폼에 연결하는 연결 체계입니다. Sun Cluster는 다음 지침에서 설명하는 모든 토폴로지를 지원합니다.

- x86 기반 시스템으로 구성된 Sun Cluster는 한 클러스터에서 2개의 노드를 지원합니다.
- 공유 저장 장치는 두 노드에 모두 연결되어야 합니다.

Sun Cluster는 특정 토폴로지를 사용하여 클러스터를 구성할 필요가 없습니다. x86 기반 노드로 구성된 클러스터의 유일한 토폴로지인 다음의 클러스터 쌍 토폴로지는 클러스터 연결 체계를 설명하는 데 필요한 핵심 내용을 제공합니다. 이러한 토폴로지는 일반 연결 체계입니다.

다음 절에 토폴로지의 예를 보여주는 그림이 있습니다.

x86: x86용 클러스터 쌍 토폴로지

클러스터 쌍 토폴로지는 단일 클러스터 관리 프레임워크에서 작동하는 두 개의 노드입니다. 이 구성에서는 쌍 사이에서만 페일오버가 발생합니다. 그러나 모든 노드는 클러스터 상호 연결에 의해 연결되고 Sun Cluster 소프트웨어의 제어 하에 작동합니다. 이 토폴로지를 사용하여 해당 쌍에서 병렬 데이터베이스, 페일오버 또는 확장 가능 응용 프로그램을 실행할 수 있습니다.

다음 그림은 클러스터된 쌍 구성을 보여줍니다.

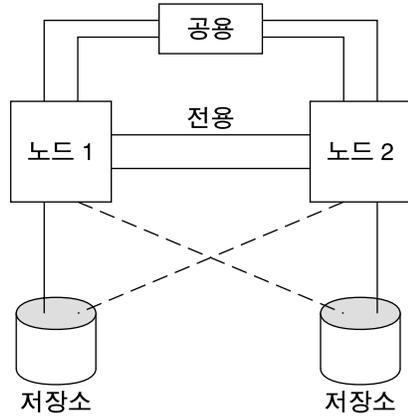


그림 2-6 x86: 클러스터 쌍 토폴로지

시스템 관리자와 응용 프로그램 개발자를 위한 주요 개념

이 장에서는 Sun Cluster 시스템의 소프트웨어 구성 요소와 관련된 주요 개념에 대해 설명합니다. 주요 내용은 다음과 같습니다.

- 33 페이지 “관리 인터페이스”
- 34 페이지 “클러스터 시간”
- 35 페이지 “고가용성 프레임워크”
- 37 페이지 “전역 장치”
- 38 페이지 “디스크 장치 그룹”
- 42 페이지 “전역 이름 공간”
- 43 페이지 “클러스터 파일 시스템”
- 45 페이지 “디스크 경로 모니터링”
- 48 페이지 “쿼럼 및 쿼럼 장치”
- 58 페이지 “데이터 서비스”
- 66 페이지 “데이터 서비스 트래픽을 위한 클러스터 상호 연결 사용”
- 67 페이지 “자원, 자원 그룹 및 자원 유형”
- 70 페이지 “데이터 서비스 프로젝트 구성”
- 79 페이지 “공용 네트워크 어댑터 및 IP (Internet Protocol) Network Multipathing”
- 80 페이지 “SPARC: 동적 재구성 지원”

이 정보는 Sun Cluster API 및 SDK를 사용하는 시스템 관리자와 응용 프로그램 개발자를 주요 대상으로 합니다. 클러스터 시스템 관리자는 클러스터 소프트웨어를 설치, 구성 및 관리하기 위한 준비 단계에서 이 정보를 사용할 수 있습니다. 응용 프로그램 개발자는 이 정보를 사용하여 작업할 클러스터 환경을 알 수 있습니다.

관리 인터페이스

여러 사용자 인터페이스 중에서 Sun Cluster 시스템을 설치, 구성 및 관리하는 방법을 선택할 수 있습니다. SunPlex Manager 그래픽 사용자 인터페이스(GUI) 또는 문서화된 명령줄 인터페이스를 통해 시스템 관리 작업을 수행할 수 있습니다. 명령줄

인터페이스의 맨 위에는 선택한 설치 및 구성 작업을 간소화하는 `scinstall` 및 `scsetup` 등의 유틸리티가 있습니다. 또한 Sun Cluster 시스템에는 Sun Management Center의 일부로 실행되어 특정 클러스터 작업에 GUI를 제공하는 모듈이 있습니다. 이 모듈은 SPARC 기반 클러스터에서만 사용할 수 있습니다. 관리 인터페이스에 대한 자세한 내용은 **Solaris OS용 Sun Cluster 시스템 관리 안내서**의 “관리 도구”를 참조하십시오.

클러스터 시간

클러스터에 있는 모든 노드 간의 시간은 동기화해야 합니다. 클러스터 작동에서 외부 시간 소스를 사용하여 클러스터 노드를 동기화할 것인지는 중요하지 않습니다. Sun Cluster 시스템은 NTP(Network Time Protocol)를 사용하여 노드 간에 클럭을 동기화합니다.

일반적으로 초의 끝자리 수 부분에 대해 시스템 시계를 변경해도 문제는 없습니다. 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 타임스탬프에 문제가 되거나 NTP 서비스에 혼란이 올 수 있습니다.

각 클러스터 노드에 Solaris 운영 체제를 설치할 때 해당 노드의 기본 시간과 날짜 설정을 변경할 수 있습니다. 일반적으로 출하시의 기본값을 승인할 수 있습니다.

`scinstall(1M)` 명령을 사용하여 Sun Cluster 소프트웨어를 설치하는 프로세스 중에 클러스터에 NTP를 구성하는 단계가 있습니다. Sun Cluster 소프트웨어는 모든 클러스터 노드 사이에 동등한 관계를 설정하는 템플릿 파일인 `ntp.cluster`를 제공합니다(설치된 클러스터 노드의 `/etc/inet/ntp.cluster` 참조). 한 개의 노드는 “선호하는” 노드로 지정됩니다. 노드들은 개인 호스트 이름으로 식별되고 시간 동기화가 클러스터 상호 연결을 거쳐 발생합니다. NTP에 대한 클러스터를 구성하는 방법에 대한 지침은 **Solaris OS용 Sun Cluster 소프트웨어 설치 안내서**의 2 장, “Sun Cluster 소프트웨어 설치 및 구성”를 참조하십시오.

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

일반적인 작동 하에서는 클러스터에서 시간을 조정할 필요가 없습니다. 그러나 Solaris 운영 체제를 설치할 때 시간이 잘못 설정되어 이를 변경하려는 경우 수행해야 할 절차는 **Solaris OS용 Sun Cluster 시스템 관리 안내서**의 7 장, “클러스터 관리”에 있습니다.

고가용성 프레임워크

Sun Cluster 시스템은 네트워크 인터페이스, 응용 프로그램, 파일 시스템 및 멀티 호스트 장치를 포함하여 사용자와 데이터 사이의 “경로”에 있는 모든 구성 요소의 가용성을 향상시킵니다. 일반적으로, 클러스터 구성 요소는 시스템에서 단일(소프트웨어 또는 하드웨어) 실패를 극복할 경우, 가용성이 높습니다.

다음 표에서는 Sun Cluster 구성 요소의 장애 종류(하드웨어 및 소프트웨어 모두)와 고가용성 프레임워크에 기본 제공되는 복구 종류에 대해 설명합니다.

표 3-1 Sun Cluster 장애 감지 및 복구 수준

장애가 발생한 클러스터 구성 요소	소프트웨어 복구	하드웨어 복구
데이터 서비스	HA API, HA 프레임워크	해당 없음
공용 네트워크 어댑터	IP (Internet Protocol) Network Multipathing	다중 공용 네트워크 어댑터 카드
클러스터 파일 시스템	기본 및 보조 복제본	멀티 호스트 장치
미러된 멀티 호스트 장치	볼륨 관리(Solaris 볼륨 관리자 및 SPARC 기반 클러스터 전용으로 사용 가능한 VERITAS Volume Manager)	하드웨어 RAID-5 (예: Sun StorEdge™ A3x00)
전역 장치	기본 및 보조 복제본	장치, 클러스터 전송 연결 장치에 대한 다중 경로
독립 네트워크	HA 전송 소프트웨어	여러 개인용 하드웨어 독립 네트워크
노드	CMM, 페일페스트 드라이버	다중 노드

Sun Cluster 소프트웨어의 고가용성 프레임워크는 노드 장애를 즉각적으로 감지하고 클러스터의 나머지 노드에 프레임워크 자원을 제공하는 새로운 서버를 만듭니다. 모든 프레임워크를 동시에 사용하지 못하는 경우는 없습니다. 손상된 노드의 영향을 받지 않는 프레임워크 자원은 복구가 진행 중인 동안에도 아무 제한없이 사용할 수 있습니다. 또한 실패한 노드의 프레임워크 자원은 복구되는 대로 사용할 수 있게 됩니다. 복구된 프레임워크 자원은 다른 모든 프레임워크 자원이 복구를 완료할 때까지 기다리지 않아도 됩니다.

대부분의 고가용성 프레임워크 자원은 해당 자원을 사용하는 응용 프로그램(데이터 서비스)으로 투명하게 복구됩니다. 프레임워크 자원 액세스의 시멘틱은 노드 실패에서 완전하게 보존됩니다. 응용 프로그램에서는 프레임워크 자원 서버가 다른 노드로 이동했음을 쉽게 감지할 수 없습니다. 한 노드에 장애가 발생하면 이 노드에 연결된 파일, 장치 및 디스크 볼륨을 사용하는 나머지 노드의 프로그램에서 이런 사실을 분명하게 알 수 있습니다. 다른 노드의 디스크에 대체 하드웨어 경로가 있는 경우에는 이 사실이 분명해집니다. 여러 노드에 대한 포트를 갖고 있는 멀티 호스트 장치를 사용하는 경우를 한 가지 예로 들 수 있습니다.

클러스터 구성원 모니터

데이터를 훼손하지 않고 안전하게 보존하려면, 모든 노드가 클러스터 구성원에서 일관되게 일치해야 합니다. 필요한 경우, CMM은 실패에 대한 응답에서 클러스터 서비스(응용 프로그램)의 클러스터 재구성에 통합됩니다.

CMM은 클러스터 전송 계층으로부터 다른 노드에 대한 연결 정보를 수신합니다. CMM은 클러스터 상호 연결을 사용하여 재구성 동안의 상태 정보를 교환합니다.

클러스터 구성원의 변경을 발견한 다음 CMM은 해당 클러스터에 대해 동기화된 구성을 수행합니다. 동기화된 구성에서는 클러스터 자원이 클러스터의 새 구성원을 기준으로 재배포될 수 있습니다.

이전의 Sun Cluster 소프트웨어 릴리스와는 달리, CMM은 전적으로 커널에서 실행됩니다.

클러스터가 여러 개의 별도 클러스터로 분할되지 않도록 보호하는 방법에 대한 자세한 내용은 50 페이지 "장애 차단 정보"를 참조하십시오.

페일패스트 기법

CMM이 노드에서 치명적인 문제를 확인하면 클러스터 프레임워크를 호출하여 강제로 해당 노드를 종료(패닉)시키고 클러스터 구성원에서 제거합니다. 이러한 작업을 수행하는 기법을 페일패스트라고 합니다. 페일패스트에 의해 노드가 종료되는 경우는 다음 두 가지가 있습니다.

- 노드가 클러스터를 벗어나서 쉘을 사용하지 않고 새 클러스터를 시작하려고 시도하면 공유 디스크에 액세스하지 못하도록 "보호"됩니다. 페일패스트의 용도에 대한 자세한 내용은 50 페이지 "장애 차단 정보"를 참조하십시오.
- 하나 이상의 클러스터 관련 데몬이 중단되면(clexecd, rpc.pmfd, rgmd 또는 rpc.ed) CMM에서 장애를 감지하여 노드가 강제 종료(패닉)됩니다.

클러스터 데몬이 중단되면 노드가 강제 종료(패닉)되고 다음과 유사한 메시지가 해당 노드의 콘솔에 표시됩니다.

```
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
```

패닉이 발생한 후에 노드가 재부트되어 클러스터로 다시 결합하려고 시도할 수 있습니다. 또한 클러스터가 SPARC 기반 시스템으로 구성된 경우에는 노드가 OpenBoot™ PROM(OBP) 프롬프트에 남아 있을 수 있습니다. 노드의 다음 작업은 auto-boot? 매개 변수의 설정에 따라 결정됩니다. OpenBoot PROM ok 프롬프트에서 eeprom(1M)을 사용하여 auto-boot?를 설정할 수 있습니다.

CCR(Cluster Configuration Repository)

CCR은 갱신 사항에 대해 2단계 완결 알고리즘을 사용합니다. 모든 클러스터 구성원이 성공적으로 업데이트되어야 하며 그렇지 않은 경우에는 업데이트 사항이 롤백됩니다. CCR은 클러스터 상호 연결을 사용하여 분배된 갱신 사항을 적용합니다.



주의 - CCR은 텍스트 파일로 구성되어 있지만 직접 CCR 파일을 편집하면 안됩니다. 각 파일에는 일관성이 유지되도록 체크섬 레코드가 포함됩니다. 수동으로 CCR 파일을 갱신하면 노드나 전체 클러스터의 기능이 정지될 수도 있습니다.

CCR은 쿼럼이 확립될 때만 클러스터가 실행되도록 하기 위해 CMM에 의존합니다. CCR은 클러스터에서 데이터 일관성을 확인해야 하는 책임을 갖고 있으므로 필요에 따라 복구를 수행하고 데이터를 갱신합니다.

전역 장치

Sun Cluster 시스템에서는 **전역 장치**를 사용하여 장치가 물리적으로 연결된 위치와 상관없이 어느 노드에서나 클러스터의 모든 장치로 액세스할 수 있는 클러스터 전반에 걸친고가용성 액세스를 제공합니다. 일반적으로 전역 장치에 액세스된 상태에서 노드에 장애가 발생하면 Sun Cluster 소프트웨어는 해당 장치에 대한 다른 경로를 자동으로 검색하여 해당 경로로 액세스를 재지정합니다. Sun Cluster 전역 장치에는 디스크, CD-ROM 및 테이프가 포함됩니다. 그러나 Sun Cluster 소프트웨어에서 지원하는 멀티 포트 전역 장치는 디스크뿐입니다. 따라서 CD-ROM과 테이프 장치는 현재고가용성 장치가 아닙니다. 각 서버의 로컬 디스크 역시 멀티 포트 상태가 아니므로고가용성 장치가 아닙니다.

클러스터는 클러스터 내의 각 디스크, CD-ROM 및 테이프 장치에 고유 ID를 자동으로 할당합니다. 이러한 할당은 클러스터의 어떤 노드에서도 각 장치에 일관되게 액세스할 수 있게 합니다. 전역 장치 이름 공간은 /dev/global 디렉토리에 저장됩니다. 자세한 내용은 42 페이지 “전역 이름 공간”을 참조하십시오.

멀티 포트 전역 장치는 장치에 대한 한 개 이상의 경로를 제공합니다. 멀티 호스트 디스크는 둘 이상의 노드가 호스트하는 디스크 장치 그룹의 일부이므로고가용성 멀티 호스트 디스크입니다.

장치 ID 및 DID 의사 드라이버

Sun Cluster 소프트웨어는 DID 의사 드라이버라고 알려진 구성을 통해 전역 장치를 관리합니다. 이 드라이버는 멀티 호스트 디스크, 테이프 드라이브 및 CD-ROM을 비롯하여 클러스터의 모든 장치에 고유 ID를 자동 할당할 때 사용합니다.

DID 의사 드라이버는 클러스터 전역 장치 액세스 기능의 필수 요소입니다. DID 드라이버는 클러스터의 모든 노드를 검사하여 고유한 디스크 장치 목록을 만든 다음 클러스터의 모든 노드에서 일관된 고유한 주 번호와 부 번호를 각 장치에 할당합니다. 전역 장치에 액세스할 때는 이전의 Solaris 장치 ID 대신 고유한 장치 ID를 사용합니다. 예를 들어 디스크의 경우에는 c0t0d0를 사용합니다.

이러한 접근 방법을 통해 디스크에 액세스하는 모든 응용 프로그램(볼륨 관리자 또는 원시 장치를 사용하는 응용 프로그램 등)은 클러스터 전체에 걸쳐 일관된 경로를 사용하게 됩니다. 이러한 일관성은 멀티 호스트 디스크에 대해 특히 유용합니다. 각 장치의 로컬 주 번호 및 부 번호는 노드마다 다를 수 있으므로 Solaris 장치 이름 지정 규칙도 변경될 수 있기 때문입니다. 예를 들어, Node1에서는 멀티 호스트 디스크를 c1t2d0으로 식별하고 Node2에서는 동일한 디스크를 완전히 다르게 c3t2d0으로 식별할 수 있습니다. DID 드라이버는 노드에서 대신 사용하는 d10 등의 전역 이름을 할당하므로 각 노드에 멀티 호스트 디스크에 대한 일관된 매핑을 제공합니다.

장치 ID는 scdidadm(1M) 및 scgdevs(1M)를 통해 업데이트 및 관리합니다. 자세한 내용은 다음의 매뉴얼 페이지를 참조하십시오.

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

디스크 장치 그룹

Sun Cluster 시스템에서 모든 멀티 호스트 장치는 Sun Cluster 소프트웨어로 제어되어야 합니다. 먼저 Solaris 볼륨 관리자 디스크 세트나 VERITAS Volume Manager 디스크 그룹(SPARC 기반 클러스터에서만 사용 가능) 같은 볼륨 관리자 디스크 그룹을 멀티 호스트 디스크에 만듭니다. 그런 다음에 볼륨 관리자 그룹을 **디스크 장치 그룹**으로 등록합니다. 디스크 장치 그룹은 전역 장치의 유형입니다. Sun Cluster 소프트웨어는 자동으로 클러스터의 각 디스크와 테이프 장치에 대한 원시 장치 그룹을 만듭니다. 그러나 사용자가 클러스터 장치 그룹을 전역 장치로 액세스할 때까지 이 클러스터 장치 그룹이 오프라인 상태를 유지합니다.

등록을 통해 특정 볼륨 관리자 디스크 그룹에 대한 경로가 있는 노드의 정보를 Sun Cluster 시스템에 제공합니다. 그러면 클러스터 전체에서 볼륨 관리자 디스크 그룹에 액세스할 수 있습니다. 둘 이상의 노드가 (마스터) 디스크 장치 그룹에 쓸 수 있으면 해당 디스크 장치 그룹에 저장된 데이터의 가용성이 높아집니다. 고가용성 디스크 장치 그룹을 사용하여 클러스터 파일 시스템을 포함시킬 수 있습니다.

주 - 디스크 장치 그룹은 자원 그룹의 영향을 받지 않습니다. 한 노드가 데이터 서비스에서 액세스하고 있는 디스크 그룹을 마스터하는 동안 다른 노드는 자원 그룹(데이터 서비스 프로세스 그룹을 지칭)을 마스터할 수 있습니다. 그러나 가장 좋은 방법은 특정 응용 프로그램 데이터를 저장하는 디스크 장치 그룹과 응용 프로그램 자원(응용 프로그램 데몬)이 들어있는 자원 그룹을 동일한 노드에 보관하는 것입니다. 디스크 장치 그룹과 자원 그룹 간의 연관에 대한 자세한 내용은 **Sun Cluster Data Services Planning and Administration Guide for Solaris OS**의 “Relationship Between Resource Groups and Disk Device Groups”를 참조하십시오.

노드가 디스크 장치 그룹을 사용하는 경우 볼륨 관리자 디스크 그룹은 기본 디스크에 다중 경로 지원을 제공하기 때문에 “전역”이 됩니다. 멀티 호스트 디스크에 물리적으로 연결된 각 클러스터 노드는 디스크 장치 그룹에 대한 경로를 제공합니다.

디스크 장치 그룹 장애 복구

디스크 엔클로저는 여러 개의 노드에 연결되어 있으므로 그 엔클로저에 있는 모든 디스크 장치 그룹은 현재 장치 그룹을 마스터하는 노드가 실패할 경우에 대체 경로를 통해 액세스할 수 있습니다. 장치 그룹을 마스터하는 노드의 실패는 복구 및 일관성 검사를 수행하는 데 시간이 소요되는 것을 제외하고는 장치 그룹에 대한 액세스에 영향을 주지 않습니다. 이 시간 동안, 모든 요청은 시스템이 장치 그룹을 사용가능하게 할 때까지 정체됩니다(응용 프로그램에서 알 수 있음).

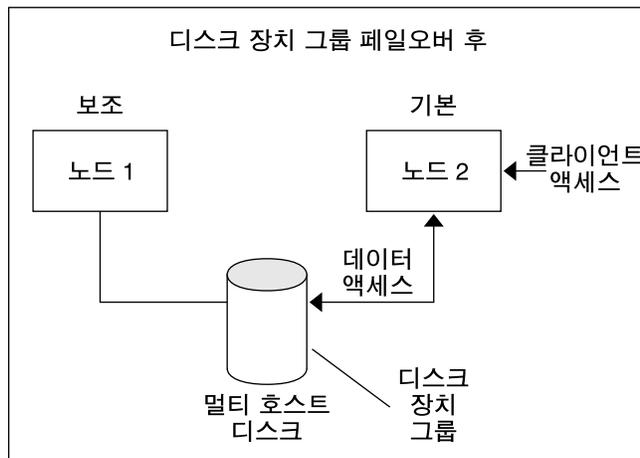
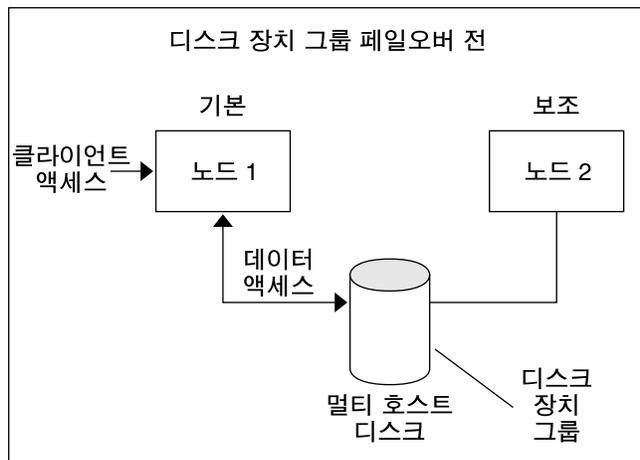


그림 3-1 페일오버 전후의 디스크 장치 그룹

멀티 포트 디스크 장치 그룹

이 절에서는 멀티 포트 디스크 구성에서 성능과 가용성의 균형을 유지할 수 있는 디스크 장치 그룹 등록 정보에 대해 설명합니다. Sun Cluster 소프트웨어는 다음과 같은 두 가지 등록 정보를 사용하여 멀티 포트 디스크 구성을 설정합니다. `preferenced` 및 `numsecondaries`. 페일오버가 발생할 경우에는 `preferenced` 등록 정보를 사용하여 노드에서 제어를 가정하는 순서를 제어할 수 있습니다. `numsecondaries` 등록 정보를 사용하여 장치 그룹에 대해 원하는 보조 노드 수를 설정합니다.

기본 노드에 장애가 발생한 경우 적절한 보조 노드를 기본 노드로 승격할 수 없으면 고가용성 서비스가 종료된 것으로 간주합니다. 서비스 페일오버가 발생하고 `preferenced` 등록 정보가 `true`이면 노드에서는 노드 목록에 지정된 순서에 따라 보조 노드를 선택합니다. 설정된 노드 목록은 노드가 기본적인 제어의 가정을 시도하거나 예비 노드에서 보조 노드로 전환하는 순서를 정의합니다. 장치 서비스의 기본 설정은 `scsetup(1M)` 유틸리티를 사용하여 동적으로 변경할 수 있습니다. 종속 서비스 공급자(예: 전역 파일 시스템)와 관련된 기본 설정은 장치 서비스의 기본 설정과 동일합니다.

보조 노드는 일반 작업 중에 기본 노드에 의해 검사점이 지정됩니다. 멀티 포트 디스크 구성에서 각 보조 노드의 검사점을 지정하면 클러스터 성능이 저하되고 메모리 오버헤드가 발생합니다. 예비 노드 지원을 구현하여 검사점 지정으로 발생한 성능 저하 및 메모리 오버헤드를 최소화합니다. 기본적으로 디스크 장치 그룹에는 한 개의 기본 노드와 한 개의 보조 노드가 있습니다. 사용 가능한 나머지 공급자 노드는 예비 노드가 됩니다. 페일오버가 발생하면 보조 노드가 기본 노드가 되고 노드 목록에서 우선 순위가 가장 높은 노드가 보조 노드가 됩니다.

원하는 보조 노드 수는 1부터 장치 그룹에 있는 작동 가능한 기본 공급자 노드 외의 공급자 노드 수 사이의 정수로 설정할 수 있습니다.

주 - Solaris 볼륨 관리자를 사용하는 경우 `numsecondaries` 등록 정보를 기본값 이외의 수로 설정하려면 디스크 장치 그룹을 만들어야 합니다.

장치 서비스를 위한 보조 노드의 기본 개수는 1입니다. 복제본 프레임워크에서 유지 관리되는 보조 공급자의 실제 수는 원하는 수로 지정할 수 있습니다. 단, 작동 가능한 기본 공급자 외의 공급자 수가 필요한 수보다 작지 않아야 합니다. `numsecondaries` 등록 정보를 변경하고 구성에 추가하거나 구성에서 삭제한 경우에는 노드 목록에서 다시 확인해야 합니다. 노드 목록과 원하는 보조 노드 수를 유지 관리하면 구성된 보조 노드 수와 프레임워크에서 허용하는 실제 노드 수 간의 충돌을 방지할 수 있습니다.

- (Solaris 볼륨 관리자) Solaris 볼륨 관리자 장치 그룹에서는 `metaset(1M)` 명령을 `preferenced` 및 `numsecondaries` 등록 정보 설정과 함께 사용하여 구성에 노드를 추가하고 제거합니다.
- (Veritas 볼륨 관리자) VxVM 디스크 장치 그룹에서는 `scconf(1M)` 명령을 `preferenced` 및 `numsecondaries` 등록 정보 설정과 함께 사용하여 구성에 노드를 추가하고 제거합니다.
- 디스크 장치 그룹의 등록 정보를 변경하는 절차에 대한 자세한 내용은 **Solaris OS용 Sun Cluster 시스템 관리 안내서**의 “클러스터 파일 시스템 관리 개요”를 참조하십시오.

전역 이름 공간

전역 장치를 활성화하는 Sun Cluster 소프트웨어 기법이 **전역 이름 공간**입니다. 전역 이름 공간에는 볼륨 관리자 이름 공간뿐 아니라 `/dev/global/` 계층이 포함됩니다. 전역 이름 공간은 멀티 호스트 디스크와 로컬 디스크(그리고 CD-ROM 및 테이프와 같은 다른 클러스터 장치) 둘 다를 반영하며, 멀티 호스트 디스크에 대한 여러 페일오버 경로를 제공합니다. 멀티 호스트 디스크에 물리적으로 연결된 각 노드는 클러스터의 모든 노드에 대해 저장소에 대한 경로를 제공합니다.

일반적으로 Solaris Volume Manager의 볼륨 관리자 이름 공간은 `/dev/md/diskset/dsk` (및 `rdsk`) 디렉토리에 있습니다. Veritas VxVM의 경우 볼륨 관리자 이름 공간은 `/dev/vx/dsk/disk-group` 및 `/dev/vx/rdsk/disk-group` 디렉토리에 있습니다. 이러한 이름 공간은 클러스터 전체에서 가져온 각각의 Solaris 볼륨 관리자 디스크 세트 및 VxVM 디스크 그룹의 디렉토리로 구성됩니다. 각 디렉토리에는 해당 디스크 세트 또는 디스크 그룹의 메타 장치나 볼륨에 필요한 장치 노드가 들어 있습니다.

Sun Cluster 시스템에서는 로컬 볼륨 관리자 이름 공간의 각 장치 노드가 `/global/.devices/node@nodeID` 파일 시스템의 장치 노드에 대한 심볼릭 링크로 교체됩니다. 여기서 `nodeID`는 클러스터의 노드를 나타내는 정수입니다. 또한 Sun Cluster 소프트웨어는 볼륨 관리자 장치를 해당 표준 위치에 심볼릭 링크로 계속 표시합니다. 전역 이름 공간과 표준 볼륨 관리자 이름 공간 둘 다 임의의 클러스터 노드에서 사용할 수 있습니다.

전역 이름 공간의 장점은 다음과 같습니다.

- 각 노드는 장치 관리 모델에서 약간 변경되었으나 독립적으로 남아 있습니다.
- 장치는 선택적으로 전역이 될 수 있습니다.
- 타사 링크 생성기를 계속 사용할 수 있습니다.
- 로컬 장치 이름이 제공되면 해당되는 전역 이름을 쉽게 매핑할 수 있습니다.

로컬 및 전역 이름 공간 예제

다음 표는 멀티 호스트 디스크 `c0t0d0s0`에 대한 전역 이름 공간과 로컬 이름 공간 사이의 매핑입니다.

표 3-2 로컬 및 전역 이름 공간 매핑

구성 요소 또는 경로	로컬 노드 이름 공간	전역 이름 공간
Solaris 논리 이름	<code>/dev/dsk/c0t0d0s0</code>	<code>/global/.devices/node@nodeID/dev/dsk/c0t0d0s0</code>
DID 이름	<code>/dev/did/dsk/d0s0</code>	<code>/global/.devices/node@nodeID/dev/did/dsk/d0s0</code>

표 3-2 로컬 및 전역 이름 공간 매핑 (계속)

구성 요소 또는 경로	로컬 노드 이름 공간	전역 이름 공간
Solaris 볼륨 관리자	/dev/md/diskset/dsk/d0	/global/.devices/node@nodeID/dev/md/diskset/dsk/d0
SPARC: VERITAS Volume Manager	/dev/vx/dsk/disk-group/v0	/global/.devices/node@nodeID/dev/vx/dsk/disk-group/v0

전역 이름 공간은 설치 시 자동으로 생성되며 재구성 부트를 실행할 때마다 갱신됩니다. scgdevs (1M) 명령을 실행하여 전역 이름 공간을 생성할 수도 있습니다.

클러스터 파일 시스템

클러스터 파일 시스템에는 다음과 같은 기능이 있습니다.

- 파일 액세스 위치가 투명합니다. 프로세스에서 시스템의 어느 위치에 있는 파일이라도 열 수 있습니다. 모든 노드의 프로세스는 동일한 경로 이름을 사용하여 파일을 찾을 수 있습니다.

주 - 클러스터 파일 시스템이 파일을 읽을 때 해당 파일에 대한 액세스 시간을 업데이트하지는 않습니다.

- 동기 프로토콜을 사용하여 파일이 동시에 여러 노드로부터 액세스될 경우에도 UNIX 파일 액세스 시멘틱을 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 권고 파일 잠금 기능을 제공합니다. 여러 클러스터 노드에서 실행되는 응용 프로그램은 클러스터 파일 시스템의 파일에 대하여 권고 파일 잠금 기능을 사용하여 데이터 액세스를 동기화할 수 있습니다. 클러스터에서 제거되는 노드와 잠금을 유지하는 동안 장애가 발생하는 응용 프로그램에서는 즉시 파일 잠금이 복구됩니다.
- 장애가 발생할 경우에도 데이터에 대한 액세스는 계속할 수 있습니다. 응용 프로그램은 디스크에 대한 경로가 계속 작동하면 실패하지 않습니다. 이러한 보장은 원시 디스크 액세스와 모든 파일 시스템 조작에 대해 유지됩니다.
- 클러스터 파일 시스템은 기본 파일 시스템 및 볼륨 관리 소프트웨어와 독립적으로 작동합니다. 클러스터 파일 시스템은 지원되는 디스크의 파일 시스템을 모두 전역으로 만듭니다.

파일 시스템을 전역 장치에 mount -g 명령을 사용하여 전역으로 또는 mount 명령을 사용하여 로컬로 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 서브세트에서 클러스터 파일 시스템을 마운트할 수 없습니다.

클러스터 파일 시스템은 별도로 구분되는 파일 시스템 형식이 아닙니다. 클라이언트는 기본 파일 시스템을 확인합니다(예: UFS).

클러스터 파일 시스템 사용

Sun Cluster 시스템에서는 모든 멀티 호스트 디스크가 디스크 장치 그룹에 포함되며 Solaris 볼륨 관리자 디스크 세트, VxVM 디스크 그룹 또는 소프트웨어 기반 볼륨 관리자의 제어를 받지 않는 개별 디스크 등이 디스크 장치 그룹이 될 수 있습니다.

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

파일 시스템을 마운트할 때 클러스터 파일 시스템을 마운트할 수 있습니다.

- **직접** – 명령행에서 mount 명령과 -g 또는 -o global 마운트 옵션을 사용하여 클러스터 파일 시스템을 마운트합니다. 예를 들면 다음과 같습니다.

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

- **자동** – 부트할 때 클러스터 파일 시스템을 마운트하려면 /etc/vfstab 파일에 global 마운트 옵션 항목을 만듭니다. 그런 다음 모든 노드의 /global 디렉토리 아래에 마운트 지점을 만듭니다. /global 디렉토리는 권장 위치이므로 반드시 따를 필요는 없습니다. 다음은 /etc/vfstab 파일에 포함된 클러스터 파일 시스템에 대한 행의 예입니다.

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

주 – Sun Cluster 소프트웨어에서 클러스터 파일 시스템에 대한 이름 지정 정책을 반드시 사용해야 하는 것은 아니지만 /global/disk-device-group과 같이 동일한 디렉토리에 모든 클러스터 파일 시스템의 마운트 지점을 만들면 관리가 쉬워집니다. 자세한 내용은 **Sun Cluster 3.1 9/04 Software Collection for Solaris OS(SPARC Platform Edition)** 및 **Solaris OS용 Sun Cluster 시스템 관리 안내서**를 참조하십시오.

HASStoragePlus 자원 유형

HASStoragePlus 자원 유형은 UFS 및 VxFS 등의 비전역 파일 시스템 구성의 가용성을 높이도록 설계되었습니다. 로컬 파일 시스템을 Sun Cluster 환경에 통합하고 파일 시스템의 가용성을 높이려면 HASStoragePlus를 사용합니다. HASStoragePlus는 Sun Cluster가 로컬 파일 시스템을 페일오버할 수 있도록 검사, 마운트 및 강제 마운트 해제와 같은 추가 파일 시스템 기능을 제공합니다. 로컬 파일 시스템이 페일오버 기능을 사용하려면 유사 스위치오버 기능이 있는 전역 디스크 그룹에 있어야 합니다.

HASStoragePlus 자원 유형의 사용 방법에 대한 자세한 내용은 **Sun Cluster Data Services Planning and Administration Guide for Solaris OS**의 “Enabling Highly Available Local File Systems”를 참조하십시오.

HASStoragePlus는 자원과 자원이 종속된 디스크 장치 그룹의 시작을 동기화하는 데에도 사용됩니다. 자세한 내용은 67 페이지 “자원, 자원 그룹 및 자원 유형”을 참조하십시오.

Syncdir 마운트 옵션

syncdir 마운트 옵션은 UFS를 기본 파일 시스템으로 사용하는 클러스터 파일 시스템에서 사용할 수 있습니다. 그러나 syncdir을 지정하지 않으면 성능이 크게 향상됩니다. syncdir를 지정하면 쓰기는 POSIX 규격을 따르게 됩니다. syncdir를 지정하지 않으면 NFS 파일 시스템과 동일하게 작동합니다. 예를 들어 syncdir이 없으면 파일을 닫기 전까지 공간 부족 상태를 발견하지 못할 수도 있습니다. syncdir(및 POSIX 동작)이 있으면 쓰기 작업 동안 공간 부족 상태가 발견되었을 것입니다. syncdir를 지정하지 않아서 문제가 발생하는 경우는 거의 없습니다.

SPARC 기반 클러스터를 사용하는 경우 VxFS에는 UFS의 syncdir 마운트 옵션과 동일한 마운트 옵션이 없습니다. syncdir 마운트 옵션을 지정하지 않으면 VxFS가 UFS와 동일하게 작동합니다.

전역 장치 및 클러스터 파일 시스템에 대한 FAQ(자주 물어보는 질문)는 86 페이지 “파일 시스템 FAQ”를 참조하십시오.

디스크 경로 모니터링

현재 릴리스의 Sun Cluster 소프트웨어에서는 디스크 경로 모니터링(DPM)을 지원합니다. 이 절에서는 DPM, DPM 데몬, 디스크 경로 모니터링에 사용하는 관리 도구 등에 대한 개념 정보를 제공합니다. 디스크 경로 상태를 모니터, 모니터 해제 및 검사하는 방법에 대한 절차 정보는 **Solaris OS용 Sun Cluster 시스템 관리 안내서**를 참조하십시오.

주 - DPM은 Sun Cluster 3.1 10/03 소프트웨어보다 먼저 릴리스된 버전을 실행하는 노드에서는 지원되지 않습니다. 순환 업그레이드가 진행되는 동안에는 DPM 명령을 사용하지 마십시오. 모든 노드를 업그레이드한 후 DPM 명령을 사용하려면 노드가 온라인 상태여야 합니다.

DPM 개요

DPM은 보조 디스크 경로 가용성을 모니터링하여 페일오버 및 전환의 전체 안정성을 향상시킵니다. `scdpm` 명령을 사용하여 자원이 전환되기 전에 해당 자원이 사용하는 디스크 경로의 가용성을 확인합니다. `scdpm` 명령과 함께 제공되는 옵션을 사용하여 클러스터의 단일 노드 또는 모든 노드에 대한 디스크 경로를 모니터링할 수 있습니다. 명령줄 옵션에 대한 자세한 내용은 `scdpm(1M)` 설명서 페이지를 참조하십시오.

DPM 구성 요소는 `SUNWscu` 패키지에서 설치됩니다. `SUNWscu` 패키지는 표준 Sun Cluster 설치 절차에 따라 설치됩니다. 설치 인터페이스에 대한 자세한 내용은 `scinstall(1M)` 설명서 페이지를 참조하십시오. 다음 표에서는 DPM 구성 요소의 기본 설치 위치에 대해 설명합니다.

위치	구성 요소
데몬	<code>/usr/cluster/lib/sc/scdpmd</code>
명령줄 인터페이스	<code>/usr/cluster/bin/scdpm</code>
공유 라이브러리	<code>/user/cluster/lib/libscdpm.so</code>
데몬 상태 파일(런타임으로 작성됨)	<code>/var/run/cluster/scdpm.status</code>

멀티스레드 DPM 데몬이 각 노드에서 실행됩니다. DPM 데몬(`scdpmd`)은 노드가 부트될 때 `rc.d` 스크립트에 의해 시작됩니다. 문제가 발생하면 데몬이 `pmfd`에 의해 관리되고 자동으로 다시 시작됩니다. 다음 목록은 초기 시작 시 `scdpmd`가 수행되는 방식을 설명합니다.

주 - 시작 시에 각 디스크 경로의 상태는 알 수 없으므로 초기화됩니다.

1. DPM 데몬은 이전 상태 파일이나 CCR 데이터베이스에서 디스크 경로와 노드 이름 정보를 수집합니다. CCR에 대한 자세한 내용은 37 페이지 “CCR(Cluster Configuration Repository)”을 참조하십시오. DPM 데몬이 시작된 후 데몬이 지정된 파일 이름에서 모니터링되는 디스크의 목록을 읽게 할 수 있습니다.
2. DPM 데몬은 통신 인터페이스를 초기화하여 명령줄 인터페이스와 같이 데몬의 외부에 있는 구성 요소의 요청에 응답합니다.
3. DPM 데몬은 `scsi_inquiry` 명령을 사용하여 10분마다 모니터링되는 목록의 각 디스크 경로를 ping합니다. 통신 인터페이스가 수정 중인 항목의 내용에 액세스하지 못하도록 각 항목을 잠급니다.
4. DPM 데몬은 Sun Cluster Event Framework에게 알리고 UNIX `syslogd(1M)` 기법을 통해 경로의 새 상태를 기록합니다.

주 - 데몬과 관련된 모든 오류는 `pmfd(1M)`에 의해 보고됩니다. API의 모든 함수는 성공 시 0을 반환하고 실패 시 -1을 반환합니다.

DPM 데몬은 Sun StorEdge Traffic Manager, HDLM 및 PowerPath 등의 다중 경로 드라이버를 통해 볼 수 있는 논리적 경로의 가용성을 모니터링합니다. 다중 경로 드라이버는 DPM 데몬에서 개별적으로 오류를 발생하기 때문에 이 드라이버에 의해 관리되는 개별 물리 경로는 모니터링되지 않습니다.

디스크 경로 모니터

이 절에서는 클러스터에서 디스크 경로를 모니터링하는 두 가지 방법을 설명합니다. 첫 번째 방법은 `scdpm` 명령에 의해 제공됩니다. 이 명령을 사용하여 클러스터의 디스크 경로 상태를 모니터, 모니터 해제 또는 표시합니다. 이 명령은 오류가 있는 디스크의 목록을 인쇄하고 파일에서 디스크 경로를 모니터링할 때도 유용합니다.

클러스터의 디스크 경로를 모니터하는 두 번째 방법은 SunPlex Manager GUI (Graphical User Interface)에 의해 제공됩니다. SunPlex Manager는 클러스터에서 모니터되는 디스크 경로에 대한 토폴로지 뷰를 제공합니다. 이 뷰는 10분마다 업데이트되어 실패한 ping의 개수 정보를 제공합니다. SunPlex Manager GUI에 의해 제공되는 정보를 `scdpm(1M)` 명령과 함께 사용하여 디스크 경로를 관리합니다. SunPlex Manager에 대한 자세한 내용은 **Solaris OS용 Sun Cluster 시스템 관리 안내서**의 10 장, “GUI를 사용한 Sun Cluster 관리”를 참조하십시오.

scdpm 명령을 사용하여 디스크 경로 모니터

`scdpm(1M)` 명령은 다음 작업을 수행할 수 있는 DPM 관리 명령을 제공합니다.

- 새 디스크 경로 모니터
- 디스크 경로 모니터 해제
- CCR 데이터베이스에서 구성 데이터 다시 읽기
- 지정한 파일에서 모니터하거나 모니터 해제할 디스크 읽기
- 디스크 경로 또는 클러스터의 모든 디스크 경로에 대한 상태 보고
- 노드에서 액세스할 수 있는 모든 디스크 경로 인쇄

활성 노드에서 디스크 경로 인자와 함께 `scdpm(1M)` 명령을 실행하여 클러스터에서 DPM 관리 작업을 수행합니다. 디스크 경로 인자는 항상 노드 이름과 디스크 이름으로 구성됩니다. 노드 이름은 필수 항목이 아니며 노드 이름을 지정하지 않은 경우 기본적으로 `a11`로 설정됩니다. 다음 표에서는 디스크 경로에 대한 이름 지정 규약을 설명합니다.

주 - 전역 디스크 경로 이름은 전체 클러스터에 걸쳐 일관되므로 전역 디스크 경로 이름을 사용할 것을 권장합니다. UNIX 디스크 경로 이름은 전체 클러스터에 걸쳐 일관성이 없습니다. 한 디스크의 UNIX 디스크 경로는 클러스터 노드 간에 서로 다를 수 있습니다. 한 노드에서는 디스크 경로가 c1t0d0이고, 다른 노드에서는 c2t0d0이 될 수 있습니다. UNIX 디스크 경로 이름을 사용할 경우 DPM 명령을 실행하기 전에 `scdidadm -L` 명령을 사용하여 UNIX 디스크 경로 이름을 전역 디스크 경로 이름으로 매핑하십시오. `scdidadm(1M)` 설명서 페이지를 참조하십시오.

표 3-3 샘플 디스크 경로 이름

이름 유형	샘플 디스크 경로 이름	설명
전역 디스크 경로	<code>schost-1:/dev/did/dsk/d1</code>	<code>schost-1</code> 노드의 디스크 경로 <code>d1</code>
<code>all:d1</code>	클러스터에 있는 모든 노드의 <code>d1</code> 디스크 경로	
UNIX 디스크 경로	<code>schost-1:/dev/rdisk/c0t0d0s0</code>	<code>schost-1</code> 노드의 디스크 경로 <code>c0t0d0s0</code>
<code>schost-1:all</code>	<code>schost-1</code> 노드의 모든 디스크 경로	
모든 디스크 경로	<code>all:all</code>	클러스터에 있는 모든 노드의 모든 디스크 경로

SunPlex Manager를 사용하여 디스크 경로 모니터

SunPlex Manager를 사용하여 다음과 같은 기본 DPM 관리 작업을 수행할 수 있습니다.

- 디스크 경로 모니터
- 디스크 경로 모니터 해제
- 클러스터의 모든 디스크 경로 상태 보기

SunPlex Manager를 사용하여 디스크 경로 관리를 수행하는 절차는 SunPlex Manager 온라인 도움말을 참조하십시오.

쿼럼 및 쿼럼 장치

이 절에서는 다음 항목에 대해 설명합니다.

- 50 페이지 "쿼럼 투표 수 정보"
- 50 페이지 "장애 차단 정보"

- 52 페이지 “쿼럼 구성 정보”
- 52 페이지 “쿼럼 장치 요구 사항 준수”
- 53 페이지 “쿼럼 장치 모범 사례 준수”
- 54 페이지 “권장되는 쿼럼 구성”
- 56 페이지 “비전형적인 쿼럼 구성”
- 56 페이지 “바람직하지 않은 쿼럼 구성”

주 - Sun Cluster 소프트웨어에서 쿼럼 장치로 지원하는 특정 장치 목록은 Sun 서비스 공급자에게 문의하십시오.

클러스터 노드는 데이터와 자원을 공유하기 때문에 클러스터 한 개를 동시에 활성화되는 별도 분할 영역으로 분할하면 안 됩니다. 활성화된 분할 영역이 여러 개이면 데이터가 훼손될 수 있습니다. CMM(Cluster Membership Monitor)과 쿼럼 알고리즘은 클러스터 상호 연결이 분할되더라도 같은 클러스터의 한 인스턴스가 항상 운영되도록 보장합니다.

쿼럼과 CMM에 대한 소개는 **Solaris OS용 Sun Cluster 개요**의 “클러스터 멤버십”을 참조하십시오.

클러스터 분할 영역에서는 두 가지 유형의 문제가 발생합니다.

- 정보 분리
- 정보 유실

정보 분리는 노드 간의 클러스터 상호 연결이 끊어지고 하위 클러스터로 분할되는 경우에 발생합니다. 한 분할 영역의 노드는 다른 분할 영역의 노드와 통신할 수 없기 때문에 각 분할 영역은 자체를 유일한 분할 영역으로 “인식합니다.”

정보 유실은 클러스터가 종료된 후, 종료하기 전의 클러스터 구성 데이터를 사용하여 다시 시작할 경우에 발생합니다. 이 문제는 최근에 작동한 클러스터 분할 영역에 속하지 않은 노드에서 클러스터를 시작하면 발생할 수 있습니다.

Sun Cluster 소프트웨어는 다음 두 가지 방법으로 정보 분리 및 정보 유실을 방지합니다.

- 각 노드에 한 투표씩 할당
- 작동 클러스터에 대부분의 투표를 위임

대부분의 투표를 가진 분할 영역은 **쿼럼**을 얻어 작동할 수 있습니다. 이러한 다수 투표 체계는 한 클러스터에 세 개 이상의 노드가 구성된 경우 정보 분리와 정보 유실을 방지합니다. 그러나, 세 개 이상의 노드가 한 클러스터에 구성되어 있을 때는 노드 투표 수를 세는 것만으로 충분하지 않습니다. 노드가 두 개인 클러스터에서는 다수가 둘입니다. 그러한 2 노드 클러스터가 분할되는 경우 두 분할 영역 중 한 쪽에서 쿼럼을 얻으려면 외부 투표가 필요합니다. 필요한 외부 투표는 **쿼럼 장치**에서 제공합니다.

쿼럼 투표 수 정보

다음 정보를 확인하려면 `scstat -q` 명령을 사용합니다.

- 구성된 총 투표 수
- 현재 있는 투표 수
- 쿼럼을 위해 필요한 투표 수

이 명령에 대한 자세한 내용은 `scstat(1M)`을 참조하십시오.

두 노드와 쿼럼 장치는 쿼럼을 형성하기 위해 클러스터에 투표를 제공합니다.

노드는 노드의 상태에 따라 투표를 제공합니다.

- 노드는 부트하여 클러스터 구성원이 될 때 투표 수 1을 가집니다.
- 노드가 설치될 때의 투표 수는 0입니다.
- 시스템 관리자가 노드를 유지 보수 상태로 둘 경우 노드의 투표 수는 0입니다.

쿼럼 장치는 장치에 연결된 투표 수를 기준으로 투표를 제공합니다. 쿼럼 장치를 구성할 때 Sun Cluster 소프트웨어는 $N-1$ 의 투표 수를 쿼럼 장치에 할당합니다. 여기서 N 은 쿼럼 장치에 연결된 투표 수입니다. 예를 들어, 투표 수가 0이 아닌 두 노드에 연결된 쿼럼 장치는 쿼럼이 1입니다($2 - 1$).

쿼럼 장치가 투표를 제공하는 경우는 다음 중 하나의 조건이 충족될 때입니다.

- 쿼럼 장치가 현재 연결된 노드 중 적어도 한 개의 노드가 클러스터 구성원인 경우
- 쿼럼 장치가 현재 연결된 노드 중 적어도 한 개의 노드가 부트되고 있고, 그 노드가 쿼럼 장치를 소유한 이전 클러스터 분할 영역의 구성원이었던 경우

쿼럼 장치는 Solaris OS용 Sun Cluster 시스템 관리 안내서의 5 장, “쿼럼 관리”에 설명된 절차를 사용하여 클러스터 설치 중이나 이후에 구성할 수 있습니다.

장애 차단 정보

클러스터에서 가장 중요한 문제는 클러스터를 분할하는 장애(정보 분리)입니다. 정보 분리 현상이 발생하면 일부 노드에서는 통신을 할 수 없으므로 개별 노드나 일부 노드가 개별 클러스터나 하위 집합 클러스터를 형성하려고 할 수 있습니다. 각 하위 집합 또는 분할 영역은 멀티 호스트 장치를 단독으로 액세스하고 소유하고 있다고 “인식”할 수 있습니다. 여러 노드에서 디스크에 쓰기를 시도하면 데이터가 손상됩니다.

장애 차단 기능은 디스크에 대한 액세스를 물리적으로 막음으로써 멀티 호스트 장치에 대한 노드 액세스를 제한합니다. 노드가 클러스터에서 나갈 경우(실패하거나 분할되어), 실패 방지는 그 노드가 더 이상 디스크에 액세스할 수 없게 만듭니다. 현재 구성된 노드들만 디스크에 대한 액세스를 갖게 되므로, 데이터 무결성이 유지됩니다.

디스크 장치 서비스는 멀티 호스트 장치를 사용하는 서비스에 대한 페일오버 기능을 제공합니다. 현재 디스크 장치 그룹의 기본 노드(소유자) 역할을 하는 클러스터 구성원에 장애가 발생하거나 연결할 수 없으면 새 기본 노드가 선택됩니다. 새 기본 노드를 사용하면 잠시 중단되었던 디스크 장치 그룹에 대한 액세스를 계속 수행할 수

있습니다. 이 프로세스 중에 기존의 기본 노드에 있는 장치에 대한 액세스 권한을 제거한 다음 새 기본 노드를 시작해야 합니다. 그러나 구성원이 클러스터에서 이탈하여 도달할 수 없게 되면 클러스터는 기본 노드였던 장치를 해제하도록 해당 노드에 알릴 수 없습니다. 그러므로 남아있는 구성원이 실패한 구성원의 전역 장치를 제어하고 액세스할 수 있도록 하는 수단이 필요합니다.

Sun Cluster 시스템은 SCSI 디스크 예약 기능을 사용하여 실패 방지를 구현합니다. SCSI 예약 기능을 사용하면 장애가 발생한 노드가 멀티 호스트 디스크에 액세스하지 못하도록 멀티 호스트 장치가 해당 노드를 “차단”합니다.

SCSI-2 디스크 예약 기능은 디스크에 연결된 모든 노드에 대한 액세스 권한을 부여하는 예약 형식을 지원합니다(예약이 없는 경우). 또한 액세스를 단일 노드로 제한합니다(예약이 있는 노드).

클러스터 상호 연결을 통해 다른 노드가 더 이상 통신할 수 없다는 것을 발견한 클러스터 구성원은 장애 차단 절차를 시작하여 다른 노드가 공유 디스크에 액세스하지 못하도록 합니다. 장애 차단이 발생하면 차단된 노드는 중단(패닉)되고 “예약 충돌” 메시지가 콘솔에 나타납니다.

노드가 더 이상 클러스터 구성원이 아닌 경우에는, 이 노드와 다른 노드가 공유하는 모든 디스크에 대한 SCSI 예약이 트리거됩니다. 차단된 노드는 차단된 사실을 “인식”하지 못할 수도 있으며, 공유 디스크 중 하나에 액세스를 시도하면 예약이 감지되어 작동이 중단(패닉)됩니다.

장애 방지를 위한 페일패스트 기법

장애가 발생한 노드가 다시 부트되어 공유 저장소에 쓰지 못하도록 하기 위하여 클러스터 프레임워크에서 사용하는 기법을 **페일패스트**라고 합니다.

클러스터를 구성하는 노드는 쿼럼 디스크를 포함하여 액세스할 수 있는 디스크에 대하여 특정 `ioctl`, `MHIOCENFAILFAST`를 계속 사용할 수 있도록 합니다. `ioctl`은 디스크 드라이버에 대한 지시어입니다. `ioctl`은 다른 노드에서 예약된 디스크이기 때문에 디스크에 액세스할 수 없는 경우 노드가 자체적으로 중단(패닉)할 수 있는 기능을 제공합니다.

`MHIOCENFAILFAST ioctl`을 사용하면 드라이버는 노드가 디스크에 대해 실행하는 모든 읽기 및 쓰기 작업에서 반환되는 오류에서 `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 재부트하면 예약이 취소됩니다. `PGR (Persistent Group Reservation)`이 있는 SCSI-3 디스크의 경우에는 예약 정보가 디스크에 저장되어 노드를 재부트한 후에도 유지됩니다. 페일패스트 기법은 SCSI-2 디스크를 사용했는지 또는 SCSI-3 디스크를 사용했는지 여부에 관계없이 동일하게 작동합니다.

노드가 클러스터의 다른 노드와 연결이 끊어지고 퀴럼을 채울 수 있는 분할 영역에 포함되지 않은 경우에는 다른 노드에 의해 강제로 클러스터에서 제거됩니다. 분할 영역에 포함되고 퀴럼을 채울 수 있는 또 다른 노드가 공유 디스크를 예약할 수 있습니다. 퀴럼이 없는 노드에서 공유 디스크에 액세스를 시도하면 페일패스트 기법이 작동하여 예약 충돌이 발생하고 시스템이 중단(패닉)됩니다.

패닉 후에는 노드가 재부트되어 클러스터에 다시 연결될 수도 있고, 클러스터가 SPARC 기반 시스템으로 구성된 경우에는 OpenBoot™ PROM (OBP) 프롬프트가 표시될 수도 있습니다. auto-boot? 매개 변수 설정에 따라 수행할 작업이 결정됩니다. SPARC 기반 클러스터의 OpenBoot PROM ok 프롬프트에서 eeprom(1M)을 사용하여 auto-boot? 를 설정할 수 있습니다. 또한 x86 기반 클러스터에서는 이 매개 변수를 BIOS가 부트된 후에 선택적으로 실행하는 SCSI 유틸리티를 사용하여 설정할 수 있습니다.

퀴럼 구성 정보

다음 목록에는 퀴럼 구성에 대한 정보가 포함되어 있습니다.

- 퀴럼 장치는 사용자 데이터를 포함할 수 있습니다.
- N 퀴럼 장치가 1에서 N 까지의 노드 중 하나와 $N+1$ 노드에 각각 연결된 $N+1$ 구성에서는 1에서 N 까지의 모든 노드 또는 $N/2$ 노드 중 어느 하나가 중단되더라도 클러스터는 계속 유지됩니다. 이러한 가용성은 퀴럼 장치가 제대로 기능을 수행하고 있는 경우를 전제로 한 것입니다.
- 단일 퀴럼 장치가 모든 노드에 연결된 N -노드 구성에서는 $N-1$ 노드 중 어느 하나가 중단되더라도 클러스터는 계속 유지됩니다. 이러한 가용성은 퀴럼 장치가 제대로 기능을 수행하고 있는 경우를 전제로 한 것입니다.
- 하나의 퀴럼 장치가 모든 노드에 연결된 N 노드 구성에서는 모든 클러스터 노드를 사용할 수 있는 경우 퀴럼 장치에 장애가 발생해도 클러스터가 견딜 수 있습니다.

피해야 하는 퀴럼 구성에 대한 예는 56 페이지 “바람직하지 않은 퀴럼 구성”을 참조하십시오. 권장되는 퀴럼 구성에 대한 예는 54 페이지 “권장되는 퀴럼 구성”을 참조하십시오.

퀴럼 장치 요구 사항 준수

다음 요구 사항을 준수해야 합니다. 이 요구 사항을 무시하면 클러스터의 가용성이 손상될 수 있습니다.

- Sun Cluster 소프트웨어가 특정 장치를 퀴럼 장치로 지원하는지 확인합니다.

주 - Sun Cluster 소프트웨어에서 퀴럼 장치로 지원하는 특정 장치 목록은 Sun 서비스 공급자에게 문의하십시오.

Sun Cluster 소프트웨어는 다음 두 가지 유형의 퀴럼 장치를 지원합니다.

- SCSI-3 PGR 예약을 지원하는 멀티 호스트 공유 디스크
- SCSI-2 예약을 지원하는 이중 호스트 공유 디스크
- 2-노드 구성에서는 한 노드에 장애가 생기더라도 다른 노드는 확실하게 작동될 수 있도록 한 개 이상의 켜림 장치를 구성해야 합니다. [그림 3-2](#)를 참조하십시오.

피해야 하는 켜림 구성에 대한 예는 [56 페이지](#) “바람직하지 않은 켜림 구성”을 참조하십시오. 권장되는 켜림 구성에 대한 예는 [54 페이지](#) “권장되는 켜림 구성”을 참조하십시오.

켜림 장치 모범 사례 준수

다음 정보를 사용하여 토폴로지에 가장 적합한 켜림 구성을 평가합니다.

- 클러스터의 모든 노드에 연결될 수 있는 장치가 있습니까?
 - 있다면 그 장치를 켜림 장치로 구성합니다. 현재 구성이 가장 최적의 구성이므로 다른 켜림 장치를 구성할 필요가 **없습니다**.



주의 - 이 요구 사항을 무시하고 다른 켜림 장치를 추가하면 클러스터의 가용성이 감소됩니다.

- 없다면 이중 포트 장치를 구성합니다.
- 켜림 장치에서 제공한 총 투표 수는 반드시 노드에서 제공한 총 투표 수보다 적어야 합니다. 그렇지 않은 경우 모든 노드가 작동하고 있더라도 모든 디스크를 사용할 수 없으면 노드가 클러스터를 형성할 수 없습니다.

주 - 특정 환경에서는 요구 사항을 충족시키기 위해 전체 클러스터의 가용성을 줄여야 할 수도 있습니다. 그런 경우에는 이러한 최적의 구성을 무시할 수 있습니다. 그러나 최적의 구성을 따르지 않으면 전체 가용성은 감소됩니다. 예를 들어 [56 페이지](#) “비전형적인 켜림 구성”에 간략하게 설명된 구성에서는 클러스터의 가용성이 낮습니다. 이 구성에서는 켜림 투표 수가 노드 투표 수보다 많습니다. 클러스터에는 Nodes A와 Node B가 공유하는 저장소에 대한 액세스 권한을 상실하면 전체 클러스터에 장애가 발생하는 특성이 있습니다.

모범 사례에 대한 예외는 [56 페이지](#) “비전형적인 켜림 구성”을 참조하십시오.

- 저장 장치에 대한 액세스를 공유하는 모든 노드 쌍 사이에 켜림 장치를 지정합니다. 이 켜림 구성은 장애 차단 프로세스의 속도를 향상시킵니다. [54 페이지](#) “세 개 이상의 노드로 구성된 켜림”을 참조하십시오.
- 일반적으로 켜림 장치를 추가하여 총 클러스터 투표 수를 짝수로 만들면 총 클러스터 가용성이 감소됩니다.
- 켜림 장치에서는 노드 추가 또는 노드 제거 후 재구성 속도가 조금 느려집니다. 그러므로 필요 이상의 켜림 장치를 추가해서는 안 됩니다.

문제를 피하기는 퀴럼 구성에 대한 예는 56 페이지 “바람직하지 않은 퀴럼 구성”을 참조하십시오. 권장되는 퀴럼 구성에 대한 예는 54 페이지 “권장되는 퀴럼 구성”을 참조하십시오.

권장되는 퀴럼 구성

이 절에서는 권장되는 퀴럼 구성에 대한 예를 보여줍니다. 피해야 하는 퀴럼 구성에 대한 예는 56 페이지 “바람직하지 않은 퀴럼 구성”을 참조하십시오.

2-노드 구성 퀴럼

2 노드 클러스터를 형성하려면 두 개의 퀴럼 투표가 필요합니다. 두 개의 투표는 두 개의 클러스터 노드 또는 한 개의 노드와 한 개의 퀴럼 장치에서만 얻을 수 있습니다.

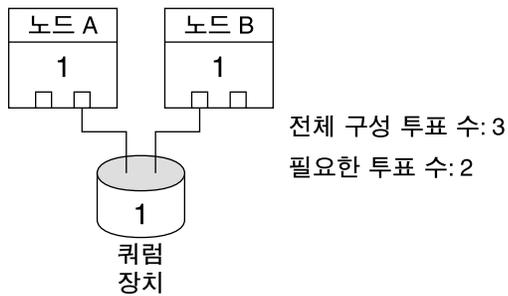
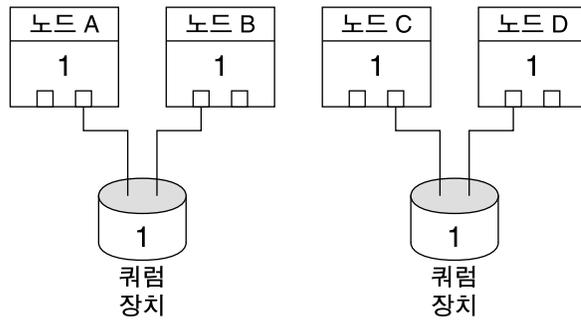


그림 3-2 2-노드 구성

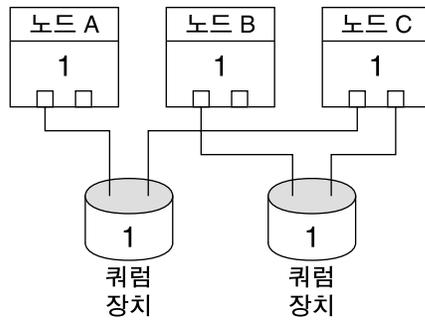
세 개 이상의 노드로 구성된 퀴럼

세 개 이상의 노드로 구성된 클러스터는 퀴럼 장치 없이 구성할 수 있습니다. 그러나 이렇게 구성하면 클러스터에서 대부분의 노드 없이 클러스터를 시작할 수 없습니다.



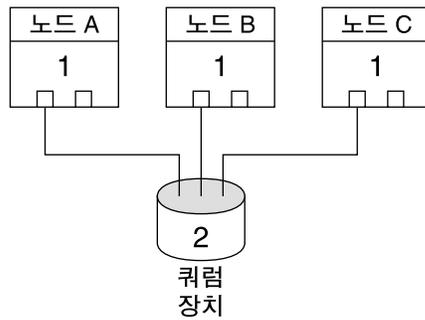
전체 구성 투표 수: 6
필요한 쿼럼 투표 수: 4

이러한 구성에서는 어느 쪽 쌍이든 작동하려면 각각의 쌍이 사용 가능해야 합니다.



전체 구성 투표 수: 5
필요한 쿼럼 투표 수: 3

이러한 구성에서는 보통 노드 A와 B에서 응용 프로그램을 실행하도록 구성하고 노드 C는 핫 스페어로 사용합니다.



전체 구성 투표 수: 5
필요한 쿼럼 투표 수: 3

이러한 구성에서는 하나 이상의 노드와 쿼럼 장치의 조합으로 하나의 클러스터를 형성할 수 있습니다.

비전형적인 켜림 구성

그림 3-3에서는 노드 A 및 노드 B에서 업무에 필수적인 응용 프로그램(예: Oracle 데이터베이스)을 실행하고 있다고 가정합니다. 노드 A와 노드 B를 사용할 수 없고 공유 데이터에 액세스할 수 없는 경우 전체 클러스터를 다운시키는 것을 원할 수 있습니다. 그렇지 않은 경우 이 구성은 고가용성을 제공하지 않기 때문에 최적의 구성이 아닙니다.

이러한 예외 사항과 관련된 모범 사례에 대한 자세한 내용은 53 페이지 “켜림 장치 모범 사례 준수”를 참조하십시오.

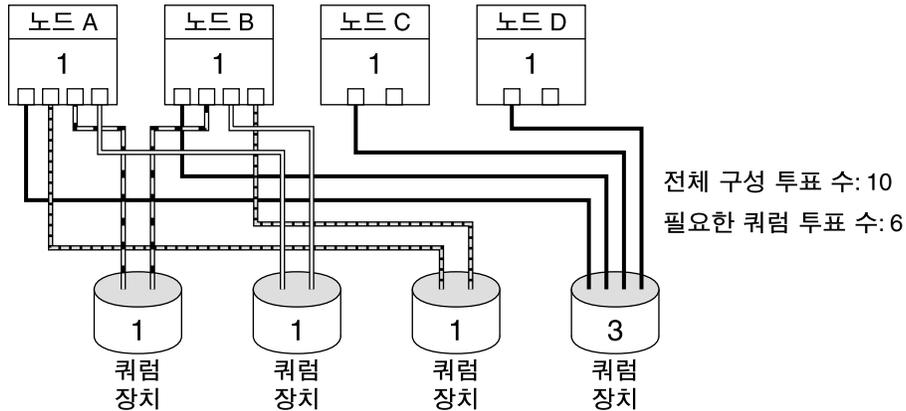
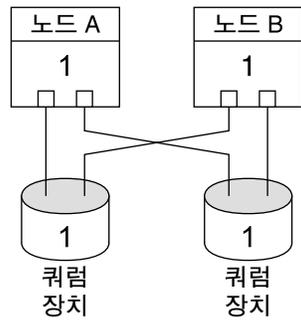


그림 3-3 비전형적 구성

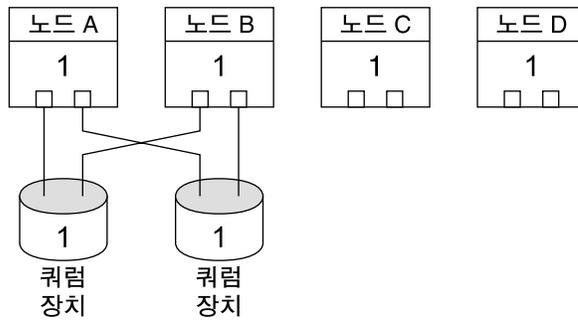
바람직하지 않은 켜림 구성

이 절에서는 피해야 하는 켜림 구성에 대한 예를 보여줍니다. 권장되는 켜림 구성에 대한 예는 54 페이지 “권장되는 켜림 구성”을 참조하십시오.



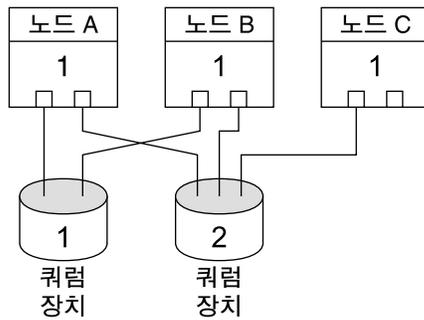
전체 구성 투표 수: 4
필요한 쿼럼 투표 수: 3

이러한 구성은 쿼럼 장치에서 제공한 총 투표 수가 노드에서 제공한 총 투표 수보다 반드시 적어야 한다는 쿼럼 장치 최적화 요건에 위배됩니다.



전체 구성 투표 수: 6
필요한 쿼럼 투표 수: 4

이러한 구성은 쿼럼 장치를 추가하여 총 클러스터 투표 수를 짝수로 만들면 안된다는 쿼럼 장치 최적화 요건에 위배됩니다. 이렇게 구성하면 가용성을 증가시킬 수 없습니다.



전체 구성 투표 수: 6
필요한 쿼럼 투표 수: 4

이러한 구성은 쿼럼 장치에서 제공한 총 투표 수가 노드에서 제공한 총 투표 수보다 반드시 적어야 한다는 쿼럼 장치 최적화 요건에 위배됩니다.

데이터 서비스

데이터 서비스는 Sun Java System Web Server 또는 Oracle과 같이 단일 서버가 아닌 클러스터에서 실행되도록 구성된 응용 프로그램을 가리키는 용어입니다. 데이터 서비스는 응용 프로그램, 특수 Sun Cluster 구성 파일 및 다음과 같은 응용 프로그램 작업을 제어하는 Sun Cluster 관리 메소드 등으로 구성됩니다.

- 시작
- 정지
- 모니터 및 수정 조치

데이터 서비스 유형에 대한 자세한 내용은 **Solaris OS용 Sun Cluster 개요**의 “데이터 서비스”를 참조하십시오.

그림 3-4에서는 단일 응용 프로그램 서버에서 실행되는 응용 프로그램(단일 서버 모델)과 클러스터에서 실행 중인 동일한 응용 프로그램(클러스터 서버 모델)을 비교합니다. 두 구성 간의 유일한 차이점은 클러스터 응용 프로그램이 보다 빠르게 실행되고 가용성이 더 높다는 점입니다.

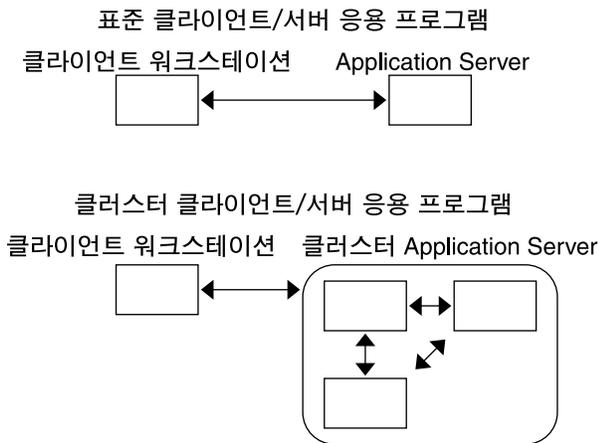


그림 3-4 표준 대 클러스터 클라이언트 서버 구성

단일 서버 모델에서는 특정의 공용 네트워크 인터페이스(호스트 이름)를 통해 서버에 액세스하도록 응용 프로그램을 구성합니다. 따라서 호스트 이름이 물리적 서버와 연결됩니다.

클러스터 서버 모델에서 공용 네트워크 인터페이스는 **논리 호스트 이름** 또는 **공유 주소**입니다. **네트워크 자원**이라는 용어는 논리 호스트 이름과 공유 주소를 모두 가리킬 때 사용합니다.

일부 데이터 서비스에서는 논리 호스트 이름이나 공유 주소를 네트워크 인터페이스로 지정해야 합니다. 논리 호스트 이름과 공유 주소는 서로 호환되지 않습니다. 그 외의 데이터 서비스에서는 논리 호스트 이름이나 공유 주소를 지정할 수 있습니다. 지정해야 하는 인터페이스 유형에 대한 자세한 내용은 각 데이터 서비스의 설치 및 구성을 참조하십시오.

네트워크 자원이 특정의 물리적 서버에 연결되어 있지 않습니다. 네트워크 자원은 물리적 서버 간에 마이그레이션할 수 있습니다.

네트워크 자원은 먼저 기본 노드에 연결됩니다. 기본 노드에 장애가 발생하면 네트워크 자원과 응용 프로그램 자원을 다른 클러스터 노드(보조 노드)로 페일오버합니다. 네트워크 자원이 페일오버하면 잠시 지연된 후에 응용 프로그램 자원이 보조 노드에서 계속 실행됩니다.

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

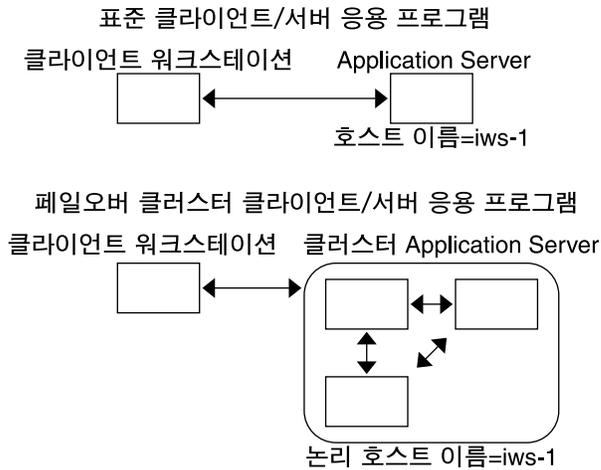


그림 3-5 고정 호스트 이름과 논리 호스트 이름 비교

공유 주소 역시 처음에는 한 개의 노드에 연결됩니다. 이 노드를 GIN(Global Interface Node)이라고 합니다. 공유 주소(일명 전역 인터페이스)는 클러스터에 대한 단일 네트워크 인터페이스로 사용됩니다.

논리 호스트 이름 모델과 확장 가능한 서비스 모델 간의 차이점은 확장 가능한 서비스 모델에는 각 노드에도 루프백 인터페이스에 활성화 상태로 구성된 공유 주소가 있다는 것입니다. 이러한 구성에서는 데이터 서비스의 여러 인스턴스를 여러 노드에서 동시에 활성화할 수 있습니다. "확장 가능 서비스"는 클러스터 노드를 추가하여 응용 프로그램에 CPU 기능을 추가하는 방법으로 성능을 확장할 수 있는 기능입니다.

GIF 노드에 장애가 발생하면 응용 프로그램 인스턴스를 실행하는 다른 노드에서 공유 주소를 시작하고 변경한 노드를 새 GIF 노드로 만들 수 있습니다. 아니면 이전에 응용 프로그램을 실행하지 않던 다른 클러스터 노드로 페일오버할 수 있습니다.

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

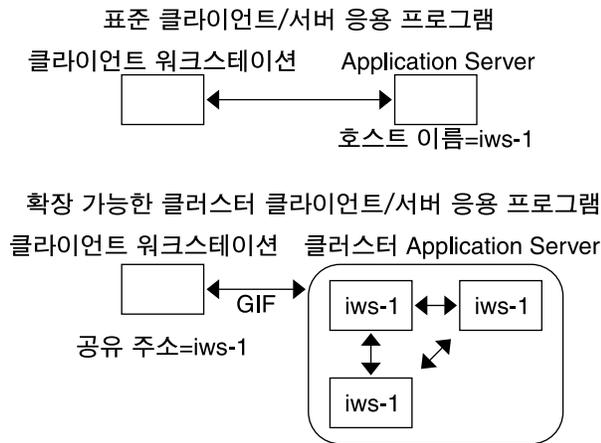


그림 3-6 고정 호스트 이름과 공유 주소 비교

데이터 서비스 메소드

Sun Cluster 소프트웨어는 일련의 서비스 관리 메소드를 제공합니다. 이 메소드는 Resource Group Manager(RGM)의 제어에 따라 실행되어 클러스터 노드의 응용 프로그램을 시작하고 중지하고 모니터링합니다. 이 메소드는 클러스터 프레임워크 소프트웨어, 멀티 호스트 장치와 함께 사용하여 응용 프로그램이 페일오버 또는 확장 가능 데이터 서비스가 되도록 합니다.

또한 RGM은 응용 프로그램의 인스턴스와 네트워크 자원(논리 호스트이름 및 공유 주소)과 같은 클러스터 내의 자원도 관리합니다.

Sun Cluster 소프트웨어에서 제공하는 메소드 외에 Sun Cluster 시스템에서도 API와 여러 가지 데이터 서비스 개발 도구를 제공합니다. 응용 프로그램 개발자는 이 도구를 사용하여 다른 응용 프로그램을 Sun Cluster 소프트웨어에서 가용성이 높은 데이터 서비스로 실행시키는 데이터 서비스 메소드를 개발할 수 있습니다.

페일오버 데이터 서비스

데이터 서비스가 실행되는 노드가 실패할 경우, 서비스는 사용자 간섭 없이 다른 작업 노드로 이주됩니다. 페일오버 서비스는 **페일오버 자원 그룹**을 사용합니다. 이 자원 그룹은 응용 프로그램 인스턴스 자원 및 네트워크 자원(**논리 호스트 이름**)을 위한 컨테이너입니다. 논리 호스트 이름은 하나의 노드에서 구성될 수 있는 IP 주소로, 나중에 기존 노드는 자동으로 구성이 중지되고 다른 노드에서 구성이 시작됩니다.

페일오버 데이터 서비스의 경우, 응용 프로그램 인스턴스는 단일 노드에서만 실행됩니다. 오류 모니터가 오류를 감지하면 동일한 노드에서 인스턴스를 다시 시작하거나 다른 노드에서 인스턴스를 시작하려고 시도합니다(페일오버). 데이터 서비스가 구성되는 방법에 따라 결과가 달라집니다.

확장 가능 데이터 서비스

확장 가능 데이터 서비스는 여러 노드에서 인스턴스가 사용되고 있을 가능성을 수반합니다. 확장 가능 서비스에서는 다음 두 가지의 자원 그룹을 사용합니다.

- **확장 가능 자원 그룹**에는 응용 프로그램 자원이 포함됩니다.
- **페일오버 자원 그룹**에는 확장 가능 서비스가 종속된 네트워크 자원(**공유 주소**)이 포함됩니다.

확장 가능 자원 그룹은 여러 노드에서 온라인 상태로 있을 수 있으므로 서비스의 여러 인스턴스가 한 번에 실행될 수 있습니다. 공유 주소를 호스트하는 페일오버 자원 그룹은 한 번에 한 노드에서만 온라인 상태입니다. 확장 가능 서비스를 호스트하는 모든 노드는 동일한 공유 주소를 사용하여 서비스를 호스트합니다.

서비스 요청은 단일 네트워크 인터페이스(전역 인터페이스)를 통해 클러스터에 보내집니다. 이러한 요청은 **로드 균형 조정 정책**에 따라 설정한 몇 가지 사전 정의 알고리즘 중 하나를 기준으로 각 노드에 분산됩니다. 클러스터는 로드 균형 조정 정책을 사용하여 몇몇 노드 사이의 서비스 부하 균형을 맞추는 로드 균형 조정 정책을 사용할 수 있습니다. 서로 다른 공유 주소를 호스트하는 노드에는 전역 인터페이스가 여러 개 있을 수 있습니다.

확장 가능 서비스의 경우 응용 프로그램 인스턴스는 여러 노드에서 동시에 실행됩니다. 전역 인터페이스를 호스트하는 노드가 실패할 경우, 전역 인터페이스는 다른 노드로 페일오버합니다. 실행 중인 응용 프로그램 인스턴스가 실패할 경우, 인스턴스는 동일한 노드에서 다시 시작하려고 시도합니다.

응용 프로그램 인스턴스는 동일한 노드에서 재시작될 수 없으며, 사용되지 않는 다른 노드는 서비스를 실행하도록 구성된 경우, 서비스는 사용되지 않는 노드로 페일오버됩니다. 그렇지 않은 경우에는 나머지 노드에서 서비스가 계속 실행되므로 서비스 처리 능력이 저하될 수 있습니다.

주 - 각 응용 프로그램 인스턴스에 대한 TCP 상태는 GIF 노드가 아니라 인스턴스가 있는 노드에 보존됩니다. 그러므로 GIF 노드의 실패는 연결에 영향을 주지 않습니다.

그림 3-7에서는 페일오버 자원 그룹 및 확장 가능 자원 그룹과, 확장 가능 서비스에 대한 이들 사이의 종속성에 대한 예를 보여줍니다. 이 예에는 세 가지 자원 그룹이 있습니다. 페일오버 자원 그룹에는 고가용성 DNS에 대한 응용 프로그램 자원과 고가용성 DNS 및 고가용성 Apache Web Server(SPARC 기반 클러스터 전용)에서 사용하는 응용 프로그램 자원이 포함됩니다. 확장 가능 자원 그룹에는 Apache Web Server의 응용 프로그램 인스턴스만 포함됩니다. 확장 가능 자원 그룹과 페일오버 자원 그룹 간에는 자원 그룹 종속성이 존재합니다(실선). 또한 모든 Apache 응용 프로그램 자원은 공유 주소인 네트워크 자원 schost-2에 종속됩니다(실선).

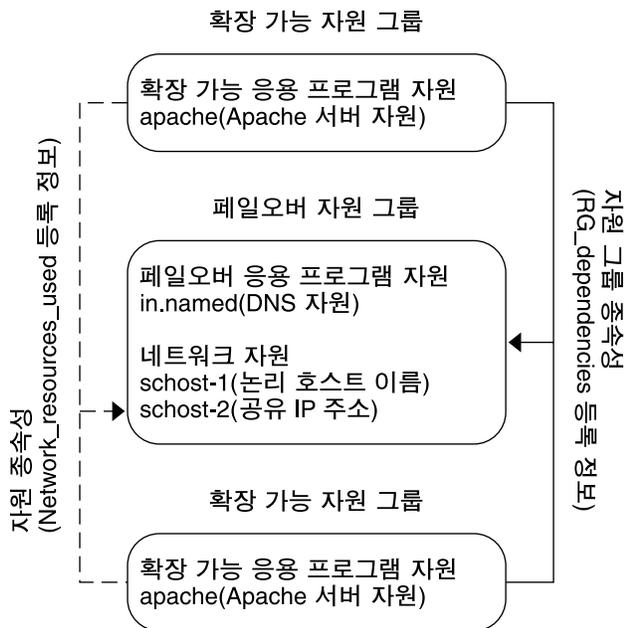


그림 3-7 SPARC: 페일오버 및 확장 가능 자원 그룹의 예

로드 균형 조정 정책

로드 균형 조정은 응답 시간과 처리량의 두 가지 측면 모두에서 확장 가능 서비스의 성능을 향상시킵니다. 확장 가능 데이터 서비스에는 두 가지 클래스가 있습니다.

- Pure
- Sticky

pure 서비스에서는 모든 인스턴스가 클라이언트의 요청에 응답할 수 있습니다. *sticky* 서비스에서는 클라이언트가 동일한 인스턴스에 요청을 보낼 수 있습니다. 그러한 요청은 다른 인스턴스에 보내지 않아도 됩니다.

pure 서비스는 가중된 로드 균형 조정 정책을 사용합니다. 이 로드 균형 조정 정책에서 클라이언트 요청은 기본적으로 클러스터의 서버 인스턴스에서 일정하게 분산됩니다. 예를 들어 3-노드 클러스터에서 각 노드의 가중치가 1이라고 가정합니다. 각 노드는 해당 서비스 대신 클라이언트의 요청 중 1/3을 처리합니다. 관리자는 `scrgadm (1M)` 명령 인터페이스 또는 SunPlex Manager GUI를 통해 언제든지 가중치를 변경할 수 있습니다.

Sticky 서비스에는 일반 *sticky*와 와일드카드 *sticky*라는 두 가지 종류가 있습니다. Sticky 서비스를 사용하면 여러 TCP 연결을 통해 동시 응용 프로그램 수준 세션에서 in-state 메모리(응용 프로그램 세션 상태)를 공유할 수 있습니다.

일반 *sticky* 서비스를 사용하면 클라이언트가 여러 개의 동시 TCP 연결 간에 상태를 공유할 수 있습니다. 클라이언트는 단일 포트를 수신하는 서버 인스턴스에 대해 “sticky”라고 합니다. 해당 서버 인스턴스가 활성 상태이고 액세스 가능하며, 서비스가 온라인 상태인 동안 로드 균형 조정 정책이 변경되지 않는 경우 클라이언트의 모든 요청은 동일한 서버 인스턴스로 가도록 보장됩니다.

예를 들어 클라이언트의 웹 브라우저는 세 개의 서로 다른 TCP 연결을 사용하여 포트 80의 공유 IP 주소로 연결됩니다. 그러나 이들 연결 간에 서비스에서 캐시된 세션 정보가 교환됩니다.

sticky 정책을 일반화하면 동일한 인스턴스에서 백그라운드로 세션 정보를 교환하는 여러 확장 가능 서비스까지 범위를 확장할 수 있습니다. 이러한 서비스가 세션 정보를 동일한 인스턴스에서 백그라운드로 교환할 때 클라이언트는 서로 다른 포트를 수신하는 동일한 노드의 여러 서버 인스턴스에 대해 “sticky”라고 합니다..

예를 들어, 전자 상거래 사이트에서 고객이 장바구니에 상품을 채울 때는 포트 80에서 HTTP를 사용하지만 고객이 장바구니의 상품 대금을 신용 카드로 지불하기 위해 보안 데이터를 보낼 때는 포트 443에서 SSL로 전환합니다.

와일드 카드 *sticky* 서비스는 동적으로 할당된 포트 번호를 사용하지만 클라이언트 요청은 여전히 동일한 노드로 전달합니다. 클라이언트는 IP 주소가 동일한 포트에 대해 “sticky 와일드 카드”입니다.

이 정책의 좋은 예는 수동 모드 FTP입니다. 예를 들어, 클라이언트가 포트 21의 FTP 서버에 연결되면 서버에서 해당 클라이언트를 동적 포트 범위에 있는 수신 포트 서버로 다시 연결하라는 지시를 내보냅니다. 이 IP 주소에 대한 모든 요청 동일한 노드로 전달됩니다. 서버에서 제어 정보를 통해 클라이언트에게 알립니다. .

이러한 각 *sticky* 정책마다 가중치가 적용된 로드 균형 조정 정책이 기본값으로 사용됩니다. 따라서, 클라이언트의 초기 요청은 로드 균형 조정기에서 지정한 인스턴스로 보내집니다. 클라이언트가 인스턴스를 실행 중인 노드에 대해 유사성을 설정한 이후에 들어온 요청은 조건에 따라 해당 인스턴스로 보내집니다. 노드에 액세스할 수 있어야 하며 로드 균형 조정 정책은 변경되지 않아야 합니다.

특정 로드 균형 조정 정책에 대한 추가 세부 정보는 다음과 같습니다.

- **Weighted.** 로드는 지정된 가중치에 따라 다양한 노드들 사이에 분배됩니다. 이 정책은 `Load_balancing_weights` 등록 정보에 `LB_WEIGHTED` 값을 사용하여 설정됩니다. 노드에 대한 가중치가 명백히 설정되지 않은 경우, 노드에 대한 가중치는 기본값인 1이 됩니다.
 가중치가 적용된 정책은 클라이언트로부터의 트래픽 중 특정 비율만큼만 특정 노드에 전달합니다. X가 가중치이고 A가 모든 활성 노드의 총 가중치라고 가정할 때, 총 새 연결 수의 약 X/A 가 활성 노드로 전달된다고 예상할 수 있습니다. 그러나 총 연결 수는 충분해야 합니다. 이 정책은 개별 요청을 다루지 않습니다.
 이 정책은 라운드 로빈(round robin)이 아닙니다. 라운드 로빈 정책은 클라이언트의 각 요청을 항상 서로 다른 노드로 이동시킵니다. 예를 들어 첫 번째 요청이 노드 1로 가면 두 번째 요청은 노드 2로 가는 방식으로 계속 진행됩니다.
- **Sticky.** 이 정책에서 포트 세트는 응용 프로그램 자원이 구성되는 시점에 알려집니다. 이 정책은 `Load_balancing_policy` 자원 등록 정보에 `LB_STICKY` 값을 사용하여 설정됩니다.
- **Sticky-wildcard.** 이 정책은 일반적인 “sticky” 정책의 슈퍼세트입니다. IP 주소로 식별하는 확장 가능 서비스에서는 서버가 포트를 할당하므로 사전에 포트를 알 수 없습니다. 포트는 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. 페일오버가 발생하면 기존의 보조 노드가 새 기본 노드가 됩니다. 페일백 설정은 기존의 기본 노드가 다시 온라인 상태가 될 때 발생할 작업을 지정합니다. 기존 기본 노드가 다시 기본 노드가 되는 페일백 옵션 또는 현재 기본 노드를 그대로 유지하는 옵션이 있습니다. `Failback` 자원 그룹 등록 정보 설정을 사용하여 원하는 옵션을 지정합니다.

자원 그룹을 호스트하는 기존 노드에서 장애가 발생하고 재부트되는 일이 반복될 때 페일백을 설정하면 자원 그룹의 가용성이 감소될 수 있습니다.

데이터 서비스 결함 모니터

각 `Sun Cluster` 데이터 서비스는 정기적으로 데이터 서비스를 규명하여 상태를 판별하는 오류 모니터를 제공합니다. 오류 모니터는 응용 프로그램 데몬이 실행 중인지 그리고 클라이언트 서비스가 제공되고 있는지 확인합니다. 프로브에 의해 반환된 정보를 기초로 데몬을 다시 시작하거나 페일오버를 야기하는 것과 같은 사전에 정의된 조치가 시작될 수 있습니다.

새로운 데이터 서비스 개발

Sun에서는 클러스터에서 여러 응용 프로그램이 페일오버 또는 확장 가능한 서비스로 작동하도록 할 수 있는 구성 파일과 관리 메소드 템플릿을 제공합니다. Sun에서 페일오버나 확장 가능 서비스로 실행할 응용 프로그램을 제공하지 않는 경우에는 다른 대안이 있습니다. Sun Cluster API 또는 DSET API를 사용하여 페일오버나 확장 가능 서비스로 실행할 응용 프로그램을 구성합니다. 그러나 모든 응용 프로그램이 확장 가능 서비스가 될 수 있는 것은 아닙니다.

확장 가능 서비스의 특성

응용 프로그램이 확장 가능 서비스가 될 수 있는지 여부는 일련의 기준에 따라 결정됩니다. 응용 프로그램이 확장 가능 서비스가 될 수 있는지를 결정하려면 **Solaris OS용 Sun Cluster 데이터 서비스 개발 안내서**의 “응용 프로그램 적합성 분석”를 참조하십시오. 일련의 기준을 요약하면 다음과 같습니다.

- 첫째, 해당 서비스는 하나 이상의 서버 인스턴스로 구성됩니다. 각 인스턴스는 클러스터의 서로 다른 노드에서 실행됩니다. 동일한 서비스에서 두 개 이상의 인스턴스가 동일한 노드에서 실행될 수는 없습니다.
- 둘째, 서비스에서 외부 논리 데이터 저장소를 제공하는 경우에는 주의해야 합니다. 여러 서버 인스턴스에서 이 저장소로 동시에 액세스할 때 업데이트 사항이 손실되거나 변경 중인 데이터를 읽지 않도록 동기화해야 합니다. 저장소를 메모리 내장 상태와 구별하기 위해 “외부”라는 용어를 사용합니다. “논리”라는 용어에는 저장소 자체가 복제될 수는 있어도 저장소는 단일 엔티티로 표시된다는 의미가 담겨 있습니다. 또한, 이 논리 데이터 저장소에는 서버 인스턴스가 저장소를 업데이트할 때마다 다른 인스턴스가 즉시 업데이트된 내용을 “볼” 수 있는 특성이 있습니다.

Sun Cluster 시스템은 클러스터 파일 시스템과 전역 원시 분할 영역을 통해 이러한 외부 저장소를 제공합니다. 예를 들면, 서비스가 새로운 데이터를 외부 로그 파일에 기록하거나 기존 데이터를 제 위치에서 수정한다고 가정합니다. 이 서비스의 여러 인스턴스가 실행되는 경우 각 인스턴스는 이 외부 로그에 대한 액세스 권한을 가지고 있으므로 이 로그에 동시에 액세스할 수 있습니다. 각 인스턴스는 해당되는 액세스를 이 로그에 동기화해야 합니다. 그렇지 않으면, 인스턴스는 서로 간섭합니다. 서비스는 `fcntl(2)` 및 `lockf(3C)` 명령을 통한 일반 Solaris 파일 잠금을 사용하여 필요한 동기화를 수행합니다.

이 유형의 저장소에 대한 다른 예는 SPARC 기반 클러스터 또는 Oracle용고가용성 Oracle Real Application Clusters Guard 등의 백엔드 데이터베이스입니다. 이 백엔드 데이터베이스 서버 유형은 데이터베이스 쿼리나 업데이트 트랜잭션을 사용하여 내장된 동기화를 제공합니다. 따라서 여러 서버 인스턴스는 자체적으로 동기화를 구현할 필요가 없습니다.

Sun IMAP 서버는 확장 가능 서비스가 아닌 서비스의 예입니다. 이 서비스는 저장소를 갱신하지만, 그 저장소는 개인용이므로 여러 IMAP 인스턴스가 저장소에 기록하면 갱신 작업이 동기화되지 않았기 때문에 인스턴스들 간에 서로 덮어쓰게 됩니다. IMAP 서버는 동시 액세스를 동기화하기 위해 다시 작성해야 합니다.

- 마지막으로, 인스턴스에는 다른 인스턴스의 데이터에서 분리된 개별 데이터가 있을 수 있다는 점에 유의하십시오. 이러한 경우 데이터는 개별 데이터여서 해당 인스턴스만 조작할 수 있으므로 서비스에는 동기화된 동시 액세스가 필요하지 않습니다. 이 때에는 개별 데이터를 데이터 클러스터 파일 시스템에 저장하지 않도록 주의해야 합니다. 이 곳에 저장된 데이터는 전역으로 액세스될 수 있기 때문입니다.

Data Service API 및 Data Service Development Library API

Sun Cluster 시스템에서는 응용 프로그램의 가용성을 높이기 위해 다음과 같은 기능을 제공합니다.

- Sun Cluster 시스템의 일부로 제공되는 데이터 서비스
- Data Service API
- 데이터 서비스용 개발 라이브러리 API
- “일반” 데이터 서비스

Sun Cluster 시스템에서 제공하는 데이터 서비스를 설치하고 구성하는 방법에 대해서는 **Sun Cluster Data Services Planning and Administration Guide for Solaris OS**를 참조하십시오. **Sun Cluster 3.1 9/04 Software Collection for Solaris OS(SPARC Platform Edition)**에서는 Sun Cluster 프레임워크에서 다른 응용 프로그램의 가용성을 높이는 방법에 대해 설명합니다.

응용 프로그램 개발자는 Sun Cluster API를 사용하여 데이터 서비스 인스턴스를 시작하고 중지하는 스크립트와 오류 모니터를 개발할 수 있습니다. 이러한 도구를 사용하여 응용 프로그램은 페일오버 또는 확장 가능 데이터 서비스로 구현될 수 있습니다. Sun Cluster 시스템에서는 “일반적인” 데이터 서비스를 제공합니다. 이러한 일반적인 데이터 서비스는 응용 프로그램이 요구하는 시작과 중지 메소드를 신속하게 생성하여 데이터 서비스를 페일오버나 확장 가능 서비스로 구현하는 데 사용됩니다.

데이터 서비스 트래픽을 위한 클러스터 상호 연결 사용

클러스터에는 클러스터 상호 연결을 형성하는 노드 간 여러 네트워크 연결이 있어야 합니다. Sun Cluster 소프트웨어는 여러 개의 상호 연결을 사용하여 다음과 같은 목적을 달성합니다.

- 고가용성 확인
- 성능 향상

내부 트래픽(예: 파일 시스템 데이터 또는 확장 가능 서비스 데이터)의 경우에는 메시지가 사용 가능한 모든 상호 연결을 통해 라운드 로빈 방식으로 스트라이프됩니다. 클러스터 상호 연결은 노드 사이의 고가용 통신을 위해 응용 프로그램에도 사용

가능합니다. 예를 들어, 분산 응용 프로그램에는 통신을 필요로 하는 다른 노드에서 실행하는 구성 요소가 있을 수 있습니다. 공용 상호 연결이 아닌 클러스터 상호 연결을 사용하여, 이 연결은 각 링크에 대한 실패로부터 안전합니다.

노드 간 통신을 위해 클러스터 상호 연결을 사용하려면 응용 프로그램에서 Sun Cluster 설치 시 구성된 개인 호스트 이름을 사용해야 합니다. 예를 들어, node 1에 대한 개인 호스트 이름이 clusternode1-priv이면 클러스터 상호 연결을 통해 node 1과 통신할 때 이 이름을 사용합니다. 이 이름을 사용하여 열리는 TCP 소켓은 클러스터 상호 연결을 통해 라우트되며 네트워크 장애가 발생하더라도 투명하게 다시 라우트될 수 있습니다.

개인 호스트 이름은 Sun Cluster 설치 시 구성할 수 있으므로 클러스터 상호 연결에서는 그 당시에 선택한 이름을 사용합니다. 실제 이름을 결정하려면 `scha_cluster_get(3HA)` 명령을 `scha_privatelink_hostname_node` 인수와 함께 사용합니다.

응용 프로그램 통신과 내부 클러스터 통신은 모든 상호 연결을 통해 스트라이프됩니다. 응용 프로그램은 내부 클러스터 트래픽과 클러스터 상호 연결을 공유하므로 응용 프로그램이 사용할 수 있는 대역폭은 다른 클러스터 트래픽에서 사용하는 대역폭에 따라 달라집니다. 장애가 발생하면 내부 트래픽과 응용 프로그램 트래픽은 사용 가능한 모든 상호 연결을 통해 스트라이프됩니다.

각 노드에는 고정 `pernode` 주소도 할당됩니다. 이 `pernode` 주소는 `clprivnet` 드라이버로 연결됩니다. IP 주소는 다음 노드의 개인 호스트 이름에 매핑됩니다. `clusternode1-priv`. Sun Cluster 개인 네트워크 드라이버에 대한 자세한 내용은 `clprivnet(7)` 설명서 페이지를 참조하십시오.

응용 프로그램이 모든 지점에서 일관된 IP 주소를 필요로 하면 클라이언트와 서버측 모두에서 `pernode` 주소에 바인드하도록 응용 프로그램을 구성합니다. 이렇게 하면 모든 연결이 `pernode` 주소에서 시작되어 다시 돌아가는 것으로 보입니다.

자원, 자원 그룹 및 자원 유형

데이터 서비스는 여러 가지 유형의 자원을 사용합니다. Apache 웹 서버나 Sun Java System Web Server와 같은 응용 프로그램은 해당 응용 프로그램이 종속된 네트워크 주소(논리 호스트 이름 및 공유 주소)를 사용합니다. 응용 프로그램과 네트워크 자원이 RGM에 의해 관리되는 기본 단위를 구성합니다.

데이터 서비스는 자원 유형입니다. 예를 들어, Sun Cluster HA for Oracle는 `SUNW.oracle-server` 자원 유형이고 Sun Cluster HA for Apache는 `SUNW.apache` 자원 유형입니다.

자원은 전체 클러스터에서 정의된 자원 유형을 인스턴스화한 것입니다. 여러 자원 유형이 정의됩니다.

네트워크 자원은 `SUNW.LogicalHostname` 또는 `SUNW.SharedAddress` 자원 유형입니다. 이 두 가지 자원 유형은 Sun Cluster 소프트웨어에 의해 사전 등록됩니다.

HASStorage 및 HASStoragePlus 자원 유형은 자원이 종속된 디스크 장치 그룹과 자원의 시작을 동기화하는 데 사용됩니다. 이 자원 유형은 데이터 서비스를 시작하기 전에 클러스터 파일 시스템 마운트 지점, 전역 장치 및 장치 그룹 이름에 대한 경로를 사용할 수 있는지를 확인합니다. 자세한 내용은 *Data Services Installation and Configuration Guide*의 “Synchronizing the Startups Between Resource Groups and Disk Device Groups”를 참조하십시오. Sun Cluster 3.0 5/02 업데이트에서는 HASStoragePlus 자원 유형을 사용할 수 있고 로컬 파일 시스템의 가용성을 높여주는 다른 기능이 추가되었습니다. 이 기능에 대한 자세한 내용은 [44 페이지](#) “HASStoragePlus 자원 유형”을 참조하십시오.

RGM이 관리하는 자원은 **자원 그룹**이라는 그룹에 포함되므로 하나의 단위로 관리할 수 있습니다. 자원 그룹에 대하여 페일오버나 전환이 시작될 때 자원 그룹은 하나의 단위로 이동됩니다.

주 - 응용 프로그램 자원이 포함된 자원 그룹을 온라인으로 전환하면 해당 응용 프로그램이 시작됩니다. 데이터 서비스 시작 메소드는 응용 프로그램이 실행될 때까지 대기했다가 성공적으로 종료됩니다. 데이터 서비스 오류 모니터에서 데이터 서비스가 클라이언트에 서비스를 제공하는 것을 결정하는 것과 동일한 방법으로 응용 프로그램이 시작되어 실행되는 시기가 결정됩니다. 이 프로세스에 대한 자세한 내용은 **Sun Cluster Data Services Planning and Administration Guide for Solaris OS**를 참조하십시오.

Resource Group Manager (RGM)

RGM은 데이터 서비스(응용 프로그램)을 **자원 유형**으로 구현하여 자원으로 관리합니다. 이 구현은 Sun에서 제공되거나 개발자가 일반 데이터 서비스 템플릿, 데이터 서비스 개발 라이브러리(DSDL API) 또는 자원 관리 API (RM API)를 사용하여 작성합니다. 클러스터 관리자는 **자원 그룹**이라는 컨테이너에 자원을 만들고 관리합니다. RGM은 클러스터 구성원 변경에 대한 응답으로 선택된 노드에서 자원을 정지하였다가 시작합니다.

RGM은 **자원**과 **자원 그룹**을 대상으로 작업을 합니다. RGM 작업을 수행하면 자원과 자원 그룹의 상태가 온라인과 오프라인 상태로 전환됩니다. 자원 및 자원 그룹에 적용할 수 있는 상태 및 설정에 대한 자세한 내용은 [68 페이지](#) “자원 및 자원 그룹의 상태와 설정” 절을 참조하십시오.

RGM이 제어하는 환경에서 Solaris 프로젝트를 시작하는 방법에 대한 자세한 내용은 [70 페이지](#) “데이터 서비스 프로젝트 구성”을 참조하십시오.

자원 및 자원 그룹의 상태와 설정

관리자는 자원과 자원 그룹에 정적 설정을 적용합니다. 이 설정은 관리 작업을 통해서만 변경될 수 있습니다. RGM은 동적 “상태” 간에 자원 그룹을 전환합니다. 이 설정과 상태는 다음 목록에서 자세히 설명합니다.

- 관리 또는 관리 해제 - 이것은 자원 그룹에만 적용되는 클러스터 범위의 설정입니다. 자원 그룹은 RGM에 의해 관리됩니다. `scrgadm(1M)` 명령을 사용하면 RGM이 자원 그룹을 관리하거나 관리하지 않도록 할 수 있습니다. 이 설정은 클러스터를 재구성해도 변경되지 않습니다.

자원 그룹을 처음 만들 때는 관리되지 않습니다. 자원 그룹은 그룹에 속한 자원이 활성화되기 전에 관리되어야 합니다.

확장 가능 웹 서버와 같은 일부 데이터 서비스에서는 네트워크 자원을 시작하기 전과 중지한 후에 작업을 수행해야 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 메소드는 확장 가능 서비스의 네트워크 자원에 가장 많이 사용됩니다. 그러나 응용 프로그램이 수행하지 않은 초기화나 정리 작업에 이 메소드를 사용할 수 있습니다.

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

자원에 대한 일반 설정은 시스템에서 활성화되어 실행되는 것입니다.

모든 클러스터 노드에서 자원을 사용할 수 없게 하려면 자원을 비활성화합니다. 비활성화된 자원은 일반적인 용도로 사용할 수 없습니다.

- 온라인 또는 오프라인 - 이것은 자원 및 자원 그룹에 모두 적용되는 동적 상태입니다.

온라인 및 오프라인 상태는 스위치오버 또는 페일오버 중에 클러스터 재구성 단계를 통해 클러스터 전이로 변경됩니다. 이 상태는 관리 작업을 통해서도 변경할 수 있습니다. 자원이나 자원 그룹의 온라인 또는 오프라인 상태를 변경하려면 `scswitch(1M)` 명령을 사용합니다.

페일오버 자원이나 자원 그룹은 항상 한 노드에서는 온라인 상태가 될 수 있습니다. 확장 가능한 자원 또는 자원 그룹은 각 노드에서 온라인 상태일 수도 있고 오프라인 상태일 수도 있습니다. 스위치오버나 페일오버 과정에서 자원 그룹 및 이 그룹에 속한 자원은 한쪽 노드에서 오프라인이 되었다가 다른 노드에서 온라인화됩니다.

자원 그룹이 오프라인 상태이면 그룹의 모든 자원도 오프라인 상태입니다. 자원 그룹이 온라인 상태이면 그룹의 사용 가능한 모든 자원도 온라인 상태입니다.

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

Sun Cluster 데이터 서비스에 대해 자원 및 자원 그룹의 등록 정보 값을 구성할 수 있습니다. 표준 속성은 모든 데이터 서비스에 공통입니다. **Extension** 속성은 각 데이터 서비스에만 적용됩니다. 일부 표준 및 확장 등록 정보는 수정하지 않아도 되도록 기본 설정으로 구성됩니다. 다른 등록 정보들은 자원 작성 및 구성 프로세스의 일부로 설정해야 합니다. 각 데이터 서비스에 대한 설명서에서는 설정할 수 있는 자원 등록 정보와 설정 방법을 지정합니다.

표준 등록 정보는 보통 특정 데이터 서비스와 독립적인 자원 및 자원 그룹 등록 정보를 구성하는데 사용됩니다. 표준 등록 정보 집합에 대한 내용은 **Sun Cluster Data Services Planning and Administration Guide for Solaris OS**의 부록 A, "Standard Properties"를 참조하십시오.

RGM 확장 등록 정보는 응용 프로그램 바이너리 위치, 구성 파일 등과 같은 정보를 제공합니다. 사용하는 데이터 서비스를 구성하는 것처럼 확장 등록 정보를 수정할 수 있습니다. 확장 등록 정보에 대한 설명은 각 데이터 서비스 설명서를 참조하십시오.

데이터 서비스 프로젝트 구성

데이터 서비스는 RGM을 사용하여 온라인으로 전환할 때 Solaris 프로젝트 이름으로 시작하도록 구성할 수 있습니다. 구성은 RGM에 의해 관리되는 자원 또는 자원 그룹을 Solaris 프로젝트 ID와 연결합니다. 자원이나 자원 그룹을 프로젝트 ID로 매핑하면 Solaris 운영 체제에서 사용할 수 있는 정교한 제어 기능을 사용하여 클러스터 내의 작업 로드 및 사용량을 관리할 수 있습니다.

주 - 이러한 구성을 수행하려면 Solaris 9 이상에서 현재 릴리스의 Sun Cluster 소프트웨어를 실행해야 합니다.

Sun Cluster 환경에서 Solaris 관리 기능을 사용하면 다른 응용 프로그램과 노드를 공유할 때 가장 중요한 응용 프로그램에게 우선권을 부여할 수 있습니다. 통합된 서비스가 있거나 응용 프로그램이 페일오버된 경우에 여러 응용 프로그램이 하나의 노드를 공유할 수 있습니다. 여기에서 설명하는 관리 기능을 사용하면 우선 순위가 낮은 응용 프로그램이 CPU 시간과 같은 시스템 자원을 지나치게 소비하지 못하도록 하여 중요한 응용 프로그램의 가용성을 향상시킬 수 있습니다.

주 - 이 기능에 대한 Solaris 설명서에서는 CPU 시간, 프로세스, 작업 및 이와 유사한 구성 요소를 “자원”으로 설명합니다. 한편, Sun Cluster 설명서에서는 “자원”을 RGM의 제어를 받는 항목을 설명하는 용어로 사용합니다. 다음 절에서는 “자원”을 RGM의 제어를 받는 Sun Cluster 항목을 일컫는 용어로 사용합니다. “제공 항목”은 CPU 시간, 프로세스 및 작업을 일컫는 용어로 사용합니다.

이 절에서는 지정된 Solaris 9 project(4)에서 프로세스를 시작하도록 데이터 서비스를 구성하는 방법에 대한 개념을 설명합니다. 또한 Solaris 운영 체제가 제공하는 관리 기능을 사용하도록 설계된 여러 파일오버 시나리오와 제안에 대해서도 설명합니다.

관리 기능의 개념 및 절차에 대한 자세한 내용은 **System Administration Guide: Network Services**의 1 장, “Network Service (Overview)”를 참조하십시오.

클러스터에서 Solaris 관리 기능을 사용하도록 자원 및 자원 그룹을 구성하는 경우 다음과 같은 고급 프로세스를 사용합니다.

1. 응용 프로그램을 자원의 일부로 구성합니다.
2. 자원을 자원 그룹의 일부로 구성합니다.
3. 자원 그룹에서 자원을 사용합니다.
4. 자원 그룹을 관리 대상으로 만듭니다.
5. 자원 그룹에 대한 Solaris 프로젝트를 생성합니다.
6. 자원 그룹 이름을 단계 5에서 생성한 프로젝트에 연결하는 표준 등록 정보를 구성합니다.
7. 자원 그룹을 온라인 상태로 전환합니다.

표준 Resource_project_name 또는 RG_project_name 등록 정보를 구성하여 Solaris 프로젝트 ID를 자원 또는 자원 그룹에 연결하려면 scrgadm(1M) 명령에 -y 옵션을 사용합니다. 자원 또는 자원 그룹에 대한 등록 정보 값을 설정합니다. 등록 정보에 대한 정의는 **Sun Cluster Data Services Planning and Administration Guide for Solaris OS**의 부록 A, “Standard Properties”를 참조하십시오. 등록 정보에 대한 설명은 r_properties(5) 및 rg_properties(5)를 참조하십시오.

지정된 프로젝트 이름은 프로젝트 데이터베이스(/etc/project)에 있어야 하며 루트 사용자는 명명된 프로젝트의 구성원으로 구성해야 합니다. 프로젝트 이름 데이터베이스에 대한 개념 정보는 **System Administration Guide: Solaris Containers-Resource Management and Solaris Zones**의 2 장, “Projects and Tasks (Overview)”를 참조하십시오. 프로젝트 파일 구문에 대한 설명은 project(4)를 참조하십시오.

RGM은 자원 또는 자원 그룹을 온라인 상태로 전환할 때 관련 프로세스를 프로젝트 이름으로 시작합니다.

주 - 사용자는 언제든지 프로젝트를 사용하여 자원 또는 자원 그룹에 연결할 수 있습니다. 그러나 새 프로젝트 이름은 RGM을 사용하여 자원 및 자원 그룹을 오프라인 상태로 만든 다음 다시 온라인 상태로 만들기 전에는 적용되지 않습니다.

자원 및 자원 그룹을 프로젝트 이름으로 시작하면 다음과 같은 기능을 구성하여 클러스터 전체에서 시스템이 제공하는 항목을 관리할 수 있습니다.

- 확장 계정 - 작업 또는 프로세스를 기준으로 사용량을 기록할 수 있는 유연한 방법을 제공합니다. 확장 계정을 사용하면 사용 기록을 조사하여 이후의 작업 로드와 대용량 요구 사항을 평가할 수 있습니다.
- 컨트롤 - 시스템 제공 항목에 대한 제한을 위한 기법을 제공합니다. 프로세스, 작업 및 프로젝트가 지정된 시스템에서 제공하는 많은 양의 자원을 소비하지 못하도록 금지할 수 있습니다.
- FSS (Fair Share Scheduling) - 중요도를 기준으로 작업 로드 간에 사용 가능한 CPU 시간 할당을 제어하는 기능을 제공합니다. 작업 로드의 중요도는 각 작업 로드에서 할당되는 CPU 시간의 공유 횟수로 표현됩니다. 자세한 내용은 다음 설명서 페이지를 참조하십시오.
 - `dispadm(1M)`
 - `priocntl(1)`
 - `ps(1)`
 - `FSS(7)`
- 풀 - 응용 프로그램의 요구 사항에 따라 대화식 응용 프로그램에 분할 영역을 사용할 수 있는 기능을 제공합니다. 풀을 사용하면 서버를 분할하여 서로 다른 많은 소프트웨어 응용 프로그램을 지원할 수 있습니다. 풀을 사용하면 각 응용 프로그램에 대한 응답을 예측하기가 쉬워집니다.

프로젝트 구성에 대한 요구 사항 결정

Sun Cluster 환경에서 Solaris가 제공하는 컨트롤을 사용하도록 데이터 서비스를 구성하기 전에 스위치오버 또는 페일오버의 전 과정에 걸쳐 자원을 제어하고 추적하는 방법을 결정해야 합니다. 새 프로젝트를 구성하기 전에 클러스터 내의 종속성을 식별합니다. 예를 들어, 자원과 자원 그룹은 디스크 장치 그룹에 종속합니다.

`scrgadm(1M)`을 사용하여 구성된 `nodelist`, `failback`, `maximum primaries` 및 `desired primaries` 자원 그룹 등록 정보를 사용하여 자원 그룹의 노드 목록 우선 순위를 식별합니다.

- 자원 그룹과 디스크 장치 그룹 간의 노드 목록 종속성에 대한 간략한 설명은 **Sun Cluster Data Services Planning and Administration Guide for Solaris OS**의 “Relationship Between Resource Groups and Disk Device Groups”를 참조하십시오.
- 등록 정보에 대한 자세한 내용은 `rg_properties(5)`를 참조하십시오.

디스크 장치 그룹 노드 목록 우선 순위를 결정하려면 `scrgadm(1M)` 및 `scsetup(1M)`을 사용하여 구성된 `preferenced` 등록 정보와 `failback` 등록 정보를 사용합니다.

- `preferenced` 등록 정보에 대한 개념 정보는 40 페이지 “멀티 포트 디스크 장치 그룹”을 참조하십시오.
- 절차 정보는 **Solaris OS용 Sun Cluster 시스템 관리 안내서**의 “디스크 장치 그룹 관리”에 있는 “디스크 장치 등록 정보를 변경하는 방법”을 참조하십시오.
- 노드 구성과 페일오버 및 확장 가능 데이터 서비스의 동작에 대한 개념 정보는 19 페이지 “Sun Cluster 시스템 하드웨어 및 소프트웨어 구성 요소”를 참조하십시오.

모든 클러스터 노드를 동일하게 구성할 경우 기본 노드와 보조 노드에 동일한 사용 한계를 적용합니다. 프로젝트의 구성 매개 변수는 모든 노드의 구성 파일에 있는 응용 프로그램에서 모두 동일할 필요는 없습니다. 해당 응용 프로그램에 연결된 모든 프로젝트는 해당 응용 프로그램에 있는 모든 잠재적 마스터의 프로젝트 데이터베이스로 액세스할 수 있어야 합니다. 응용 프로그램 1은 *phys-schost-1*에 의해 마스터되지만 잠재적으로 *phys-schost-2* 또는 *phys-schost-3*으로 전환(스위치오버)되거나 페일오버될 수 있다고 가정합니다. 응용 프로그램 1에 연결된 프로젝트는 세 노드(*phys-schost-1*, *phys-schost-2* 및 *phys-schost-3*) 모두에서 액세스할 수 있어야 합니다.

주 - 프로젝트 데이터베이스 정보는 로컬 `/etc/project` 데이터베이스 파일이 되거나 NIS 맵 또는 LDAP 디렉토리 서비스에 저장될 수 있습니다.

Solaris 운영 체제에서는 사용 매개 변수를 유연하게 구성할 수 있으며 Sun Cluster에서 부과하는 제한 사항은 거의 없습니다. 구성 선택 항목은 사이트의 필요에 따라 다릅니다. 시스템을 구성하기 전에 다음 절의 일반 지침을 따르십시오.

선행 프로세스 가상 메모리 제한 설정

선행 프로세스를 기준으로 가상 메모리를 제한하도록 `process.max-address-space` 컨트롤을 설정합니다. `process.max-address-space` 값 설정에 대한 자세한 내용은 `rctladm(1M)`을 참조하십시오.

Sun Cluster 소프트웨어에서 관리 컨트롤을 사용하면 응용 프로그램의 불필요한 페일오버와 “핑퐁” 효과가 발생하지 않도록 메모리 제한을 적절하게 구성합니다. 일반적으로 다음과 같은 지침을 준수합니다.

- 메모리 제한을 너무 낮게 설정하지 마십시오.
응용 프로그램이 메모리 제한에 도달하면 페일오버될 수 있습니다. 이 지침은 가상 메모리 제한에 도달할 경우 예상치 않은 결과가 발생할 수 있는 데이터베이스 응용 프로그램에 특히 중요합니다.
- 기본 노드와 보조 노드에서 메모리 제한을 동일하게 설정하지 마십시오.
동일한 한계를 설정하면 응용 프로그램이 메모리 제한에 도달하여 동일한 메모리 제한을 갖는 보조 노드에 페일오버될 경우 핑퐁 효과가 발생할 수 있습니다. 보조 노드의 메모리 제한을 약간 더 높게 설정하십시오. 메모리 제한을 각기 다르게 설정하면 핑퐁 시나리오를 방지하여 시스템 관리자가 필요한 경우 매개 변수를

조정할 수 있는 시간을 제공합니다.

- 로드 균형 조정을 위해 자원 관리 메모리 제한을 사용하십시오.

예를 들어, 메모리 제한을 사용하여 잘못된 응용 프로그램이 과도한 스왑 공간을 차지하지 않도록 금지할 수 있습니다.

페일오버 시나리오

일반 클러스터 작업 중 및 스위치오버 또는 페일오버 상황에서 프로젝트 구성(/etc/project) 할당 작업이 수행되도록 관리 매개 변수를 구성할 수 있습니다.

다음 절은 시나리오 예입니다.

- 처음 두 절 “두 응용 프로그램이 있는 2-노드 클러스터” 및 “세 응용 프로그램이 있는 2-노드 클러스터”에서는 전체 노드에 대한 페일오버 시나리오를 보여줍니다.
- “자원 그룹 전용 페일오버” 절에서는 응용 프로그램에만 해당되는 페일오버 작업을 설명합니다.

Sun Cluster 환경에서 응용 프로그램을 자원의 일부로 구성합니다. 그런 다음 자원을 자원 그룹(RG)의 일부로 구성합니다. 장애가 발생하면 자원 그룹이 이와 연결된 응용 프로그램과 함께 다른 노드로 페일오버됩니다. 다음 예에서는 자원이 명시적으로 표시되지 않습니다. 각 자원에 응용 프로그램이 하나만 있는 것으로 가정합니다.

주 - 페일오버는 RGM에 설정된 기본 노드 목록 순서로 발생합니다.

다음 예에서는 이러한 제한 조건을 갖습니다.

- 응용 프로그램 1(App-1)은 자원 그룹 RG-1에 구성됩니다.
- 응용 프로그램 2 (App-2)는 자원 그룹 RG-2에 구성됩니다.
- 응용 프로그램 3(App-3)은 자원 그룹 RG-3에 구성됩니다.

할당된 공유 개수는 동일하게 유지되지만 각 응용 프로그램에 할당된 CPU 시간 비율은 페일오버 후에 변경됩니다. 이 비율은 노드에서 실행 중인 응용 프로그램의 수 및 각 활성 응용 프로그램에 할당된 공유 수에 따라 다릅니다.

이 시나리오에서는 다음과 같은 구성을 가정합니다.

- 모든 응용 프로그램이 공통 프로젝트 아래 구성됩니다.
- 각 자원에 응용 프로그램이 하나만 있습니다.
- 응용 프로그램이 노드에서 활성 상태인 유일한 프로세스입니다.
- 프로젝트 데이터베이스가 클러스터의 각 노드에 동일하게 구성됩니다.

두 응용 프로그램이 있는 2-노드 클러스터

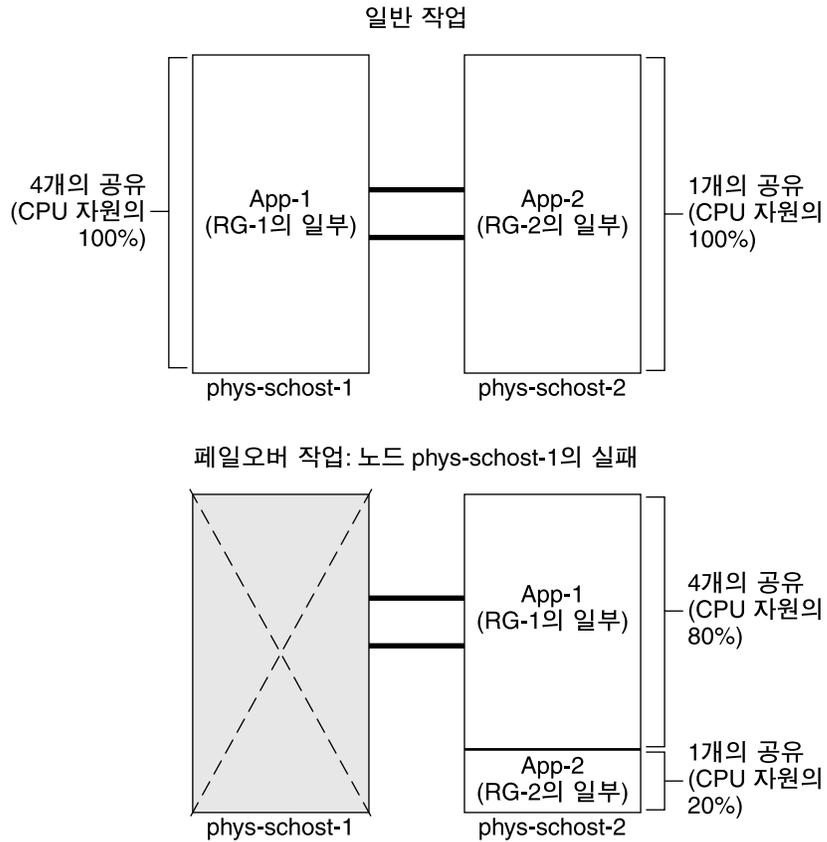
2-노드 클러스터에서 두 개의 응용 프로그램을 구성하여 물리적 호스트(*phys-schost-1*, *phys-schost-2*)가 각각 한 응용 프로그램의 기본 마스터 역할을 수행하도록 할 수 있습니다. 각 물리적 호스트는 다른 물리적 호스트에 대한 보조 노드 역할을 합니다.

응용 프로그램 1과 응용 프로그램 2에 연결된 모든 프로젝트가 두 개의 노드 모두에서 프로젝트 데이터베이스 파일에 표시되어야 합니다. 클러스터가 일반적으로 실행 중일 때는 각 응용 프로그램이 기본 마스터에서 실행되고, 이 때는 관리 기능에 의해 응용 프로그램에 모든 CPU 시간이 할당됩니다.

페일오버 또는 전환이 발생한 후에는 두 응용 프로그램 모두 단일 노드에서 실행되고 구성 파일에 지정된 대로 응용 프로그램에 공유가 할당됩니다. 예를 들어 /etc/project 파일의 해당 항목에서 응용 프로그램 1에 4개의 공유를, 응용 프로그램 2에 1개의 공유를 할당한다고 지정합니다.

```
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)
```

다음 다이어그램에서는 이 구성의 일반 작업과 페일오버 작업을 설명합니다. 할당되는 공유 수는 변경되지 않습니다. 그러나 각 응용 프로그램에서 사용할 수 있는 CPU 시간의 비율은 변경할 수 있습니다. 비율은 CPU 시간을 요구하는 각 프로세스에 할당된 공유 수에 따라 달라집니다.



세 응용 프로그램이 있는 2-노드 클러스터

세 개의 응용 프로그램이 있는 2-노드 클러스터에서 한 개의 물리적 호스트(*phys-schost-1*)를 한 응용 프로그램의 기본 마스터로 구성할 수 있습니다. 두 번째 물리적 호스트(*phys-schost-2*)는 나머지 두 응용 프로그램의 기본 마스터로 구성할 수 있습니다. 모든 노드에서 다음 예 프로젝트 데이터베이스 파일을 가정합니다. 페일오버 또는 전환이 발생할 때 프로젝트 데이터베이스 파일은 변경되지 않습니다.

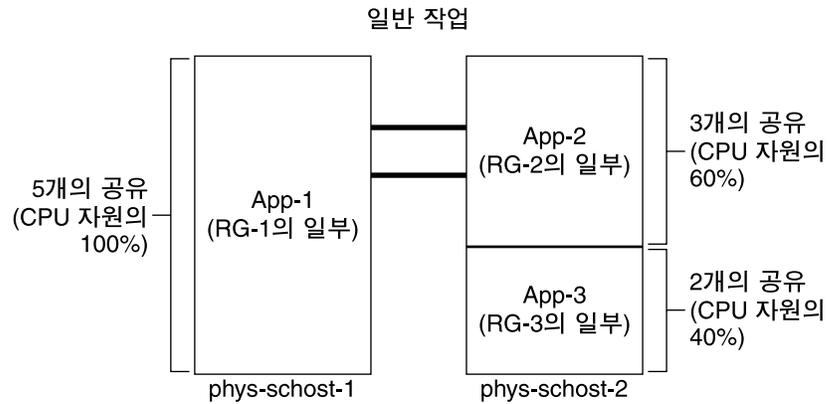
```
Prj_1:103:project for App_1:root::project.cpu-shares=(privileged,5,none)
Prj_2:104:project for App_2:root::project.cpu-shares=(privileged,3,none)
Prj_3:105:project for App_3:root::project.cpu-shares=(privileged,2,none)
```

클러스터가 정상적으로 실행 중인 경우 기본 마스터 *phys-schost-1*에서 응용 프로그램 1에 5개의 공유가 할당됩니다. 응용 프로그램 1이 이 노드에서 CPU 시간을 필요로 하는 유일한 응용 프로그램이기 때문에 이 수는 CPU 시간의 100%에 해당합니다. 응용 프로그램 2와 3에는 기본 마스터인 *phys-schost-2*에서 각각 3개와 2개의 공유가 할당됩니다. 일반 작업 중에 응용 프로그램 2는 CPU 시간의 60%를 받고 응용 프로그램 3은 40%를 받습니다.

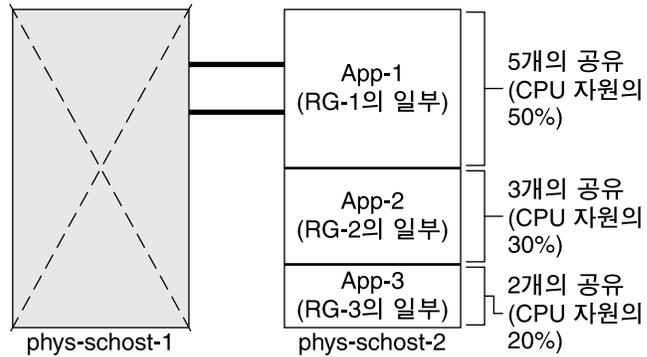
페일오버나 스위치오버가 발생하여 응용 프로그램 1이 *phys-schost-2*로 전환되더라도 세 응용 프로그램 모두에 대한 공유 수는 동일하게 유지됩니다. 그러나, CPU 자원 비율은 프로젝트 데이터베이스 파일에 따라 재할당됩니다.

- 5개의 공유를 갖는 응용 프로그램 1이 CPU의 50%를 받습니다.
- 3개의 공유를 갖는 응용 프로그램 2가 CPU의 30%를 받습니다.
- 2개의 공유를 갖는 응용 프로그램 3이 CPU의 20%를 받습니다.

다음 다이어그램에서는 이 구성의 일반 작업과 페일오버 작업을 설명합니다.



페일오버 작업: 노드 phys-schost-1의 실패



자원 그룹 전용 페일오버

여러 자원 그룹이 동일한 기본 마스터를 갖는 구성에서는 자원 그룹과 관련 응용 프로그램이 보조 노드로 페일오버되거나 전환될 수 있습니다. 반면에 기본 마스터는 클러스터에서 실행됩니다.

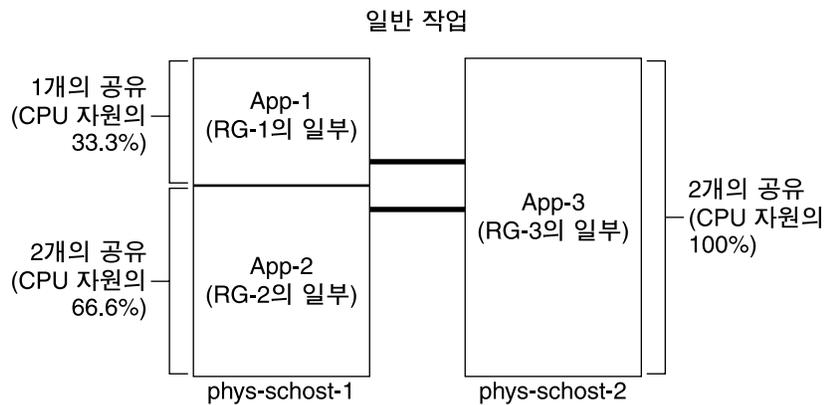
주 - 페일오버 중에 페일오버되는 응용 프로그램에는 보조 노드의 구성 파일에 지정된 대로 자원이 할당됩니다. 이 예에서 기본 및 보조 노드의 프로젝트 데이터베이스 파일 구조는 동일합니다.

예를 들어, 다음 샘플 구성 파일에서는 응용 프로그램 1에 1개의 공유, 응용 프로그램 2에 2개의 공유, 응용 프로그램 3에 2개의 공유가 할당되도록 지정합니다.

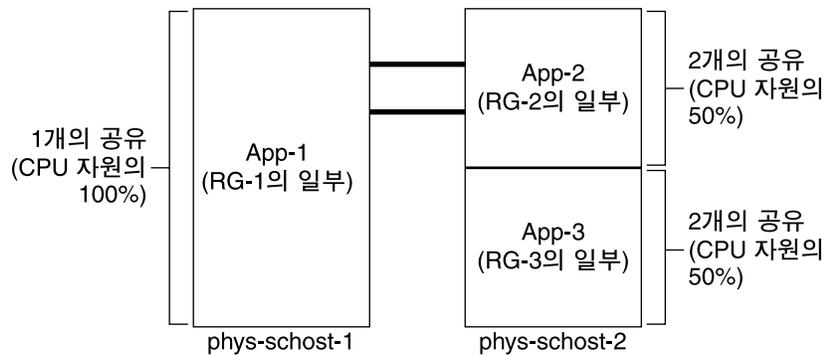
```
Prj_1:106:project for App_1:root::project.cpu-shares=(privileged,1,none)
Prj_2:107:project for App_2:root::project.cpu-shares=(privileged,2,none)
```

```
Prj_3:108:project for App_3:root::project.cpu-shares=(privileged,2,none)
```

다음 다이어그램에서는 이 구성의 일반 작업과 페일오버 작업에 대해 설명합니다. 여기서 응용 프로그램 2가 포함된 RG-2는 *phys-schost-2*로 페일오버됩니다. 할당되는 공유 수는 변경되지 않습니다. 그러나, 각 응용 프로그램에서 사용할 수 있는 CPU 시간 비율은 CPU 시간을 요구하는 각 응용 프로그램에 할당된 공유 수에 따라 변경될 수 있습니다.



페일오버 작업: RG-2가 *phys-schost-2*로 페일오버됩니다.



공용 네트워크 어댑터 및 IP (Internet Protocol) Network Multipathing

클라이언트는 공용 네트워크 인터페이스를 통해 클러스터에 데이터 요청을 합니다. 각 클러스터 노드는 공용 네트워크 어댑터 쌍을 통해 최소한 하나의 공용 네트워크에 연결됩니다.

Sun Cluster의 Solaris Internet Protocol(IP) Network Multipathing 소프트웨어는 공용 네트워크 어댑터를 모니터하고 오류가 감지된 경우 한 어댑터에서 다른 어댑터로 IP 주소를 페일오버하는 기본 기법을 제공합니다. 각 클러스터 노드에는 다른 클러스터 노드의 구성과 구별될 수 있는 자체 IP (Internet Protocol) Network Multipathing 구성이 있습니다.

공용 네트워크 어댑터는 *IP multipathing 그룹*(multipathing 그룹)으로 구성됩니다. 각 Multipathing 그룹에는 하나 이상의 공용 네트워크 어댑터가 있습니다. Multipathing 그룹의 각 어댑터는 활성화될 수 있습니다. 또한 페일오버가 발생하기 전에는 비활성 상태인 대기 인터페이스를 구성할 수 있습니다.

`in.mpathd multipathing` 데몬은 테스트 IP 주소를 사용하여 오류를 감지하고 복구합니다. `multipathing` 데몬이 어댑터 중 하나에서 오류를 감지하면 페일오버가 발생합니다. 모든 네트워크 액세스는 `multipathing` 그룹의 오류가 발생한 어댑터에서 정상적으로 작동하는 다른 어댑터로 페일오버됩니다. 따라서 데몬은 해당 노드에 대한 공용 네트워크 연결을 유지 관리합니다. 대기 인터페이스가 구성된 경우 데몬은 대기 인터페이스를 선택합니다. 그렇지 않으면 데몬은 가장 작은 수의 IP 주소를 가진 인터페이스를 선택합니다. 페일오버는 어댑터 인터페이스 수준에서 발생하므로 TCP와 같은 고급 연결은 페일오버 중에 일시적으로 지연되는 경우를 제외하고는 영향을 받지 않습니다. IP 주소의 페일오버가 성공적으로 수행되면 ARP 브로드캐스트가 전송됩니다. 따라서 데몬은 원격 클라이언트에 대한 연결을 유지 관리합니다.

주 - TCP의 정체 복구 특성으로 인해 성공적인 페일오버 후에도 TCP 끝점에서 추가 지연이 발생할 수 있습니다. TCP의 정체 제어 기법이 활성화되면 페일오버 중에 일부 세그먼트가 손실될 수 있습니다.

Multipathing 그룹은 논리 호스트 이름 및 공유 주소 자원에 대한 빌딩 블록을 제공합니다. 또한 사용자도 논리 호스트 이름과 공유 주소 자원의 Multipathing 그룹을 독립적으로 만들어서 클러스터 노드에 대한 공용 네트워크 연결을 모니터할 수 있습니다. 한 노드의 동일한 Multipathing 그룹이 여러 논리 호스트 이름이나 공유 주소 자원을 호스팅할 수 있습니다. 논리 호스트 이름과 공유 주소 자원에 대한 자세한 내용은 **Sun Cluster Data Services Planning and Administration Guide for Solaris OS**를 참조하십시오.

주 - IP (Internet Protocol) Network Multipathing 기법은 어댑터 장애를 감지하고 마스크하도록 설계되었습니다. 이 설계는 관리자가 `ifconfig(1M)`를 사용하여 논리(또는 공유) IP 주소 중 하나를 제거한 것을 복구하기 위한 것이 아닙니다. Sun Cluster 소프트웨어는 논리 및 공유 IP 주소를 RGM에서 관리하는 자원으로 인식합니다. 관리자가 IP 주소를 추가 또는 제거하는 올바른 방법은 `scrgadm(1M)`을 사용하여 해당 자원이 포함된 자원 그룹을 수정하는 것입니다.

IP Network Multipathing을 Solaris 에서 구현하는 방법에 대한 자세한 내용은 클러스터에 설치된 Solaris 운영 체제에 대한 해당 설명서를 참조하십시오.

운영 체제 릴리스	지침
Solaris 8 운영 체제	IP Network Multipathing Administration Guide
Solaris 9 운영 체제	IP Network Multipathing Administration Guide 의 1 장, "IP Network Multipathing (Overview)"
Solaris 10 Operating System	System Administration Guide: IP Services 의 파트 VI, "IPMP"

SPARC: 동적 재구성 지원

DR(동적 재구성) 소프트웨어 기능에 대한 Sun Cluster 3.1 8/05의 지원이 점진적으로 개발되고 있습니다. 이 절에서는 Sun Cluster 3.1 8/05의 DR 기능 지원에 대한 개념과 고려 사항에 대해 설명합니다.

Solaris DR 기능의 문서화된 요구 사항, 절차 및 제한 사항은 Sun Cluster DR 지원에도 모두 적용됩니다(운영 환경의 작동이 정지된 경우는 제외). 따라서 Sun Cluster 소프트웨어에서 DR 기능을 사용하기 전에 먼저 Solaris DR 기능에 대한 설명서를 검토합니다. 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** 또는 **Solaris 9 on Sun Hardware** 모음에 포함)은 모두 <http://docs.sun.com>에서 다운로드할 수 있습니다.

SPARC: 동적 재구성 일반 설명

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

DR은 보드 레벨로 작동합니다. 따라서 DR 작업은 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.

활성 구성 요소가 포함된 보드를 제거하면 시스템 오류가 발생합니다. 보드를 제거하기 전에 DR 하위 시스템은 Sun Cluster와 같은 다른 하위 시스템을 쿼리하여 보드의 구성 요소가 사용 중인지 확인합니다. 보드가 사용 중이면 DR 보드 제거 작업이 수행되지 않습니다. 따라서 DR 하위 시스템은 활성 구성 요소가 포함된 보드에 대해 작업을 거부하기 때문에 DR 보드 제거 작업을 실행해도 항상 안전합니다.

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

주 - DR 하위 시스템에는 여러 수준이 있습니다. 낮은 레벨에서 오류를 보고하면 상위 레벨도 오류를 보고합니다. 그러나 하위 수준에서 특정 오류를 보고하면 상위 수준에서 알 수 없는 오류를 보고합니다. 이 오류는 무시해도 됩니다.

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

SPARC: CPU 장치에 대한 DR 클러스터링 참고 사항

Sun Cluster 소프트웨어에는 CPU 장치가 있으므로 DR 보드 제거 작업을 거부하지 않습니다.

DR 보드 추가 작업이 성공하면 추가된 보드의 CPU 장치가 시스템 작업에 자동으로 통합됩니다.

SPARC: 메모리에 대한 DR 클러스터링 참고 사항

DR을 위해서는 두 가지의 메모리 유형을 고려해야 합니다.

- 커널 메모리 캐이지
- 커널이 아닌 메모리 캐이지

이 두 가지 유형은 용도만 다릅니다. 실제 하드웨어는 두 가지 유형이 동일합니다. 커널 메모리 캐이지는 Solaris 운영 체제에서 사용하는 메모리입니다. Sun Cluster 소프트웨어는 커널 메모리 캐이지가 있는 보드에서는 보드 제거 작업을 지원하지

않으며 이러한 작업은 모두 거부합니다. DR 보드 제거 작업이 커널 메모리 캐이지가 아닌 메모리와 관련되어 있으면 Sun Cluster 소프트웨어에서 작업을 거부하지 않습니다. 메모리에 속하는 DR 보드 추가 작업이 성공하면 추가된 보드의 메모리가 시스템 작업에 자동으로 통합됩니다.

SPARC: 디스크 및 테이프 드라이브에 대한 DR 클러스터링 참고 사항

Sun Cluster에서는 기본 노드에서 활성 드라이브에 대한 DR 보드 제거 작업을 거부할 수 없습니다. DR 보드 제거 작업은 기본 노드의 비활성 드라이브와 보조 노드의 모든 드라이브에서 수행될 수 있습니다. DR 작업이 끝나면 작업 이전과 마찬가지로 클러스터 데이터 액세스가 계속됩니다.

주 - Sun Cluster에서는 쉘 장치의 가용성에 영향을 주는 DR 작업을 할 수 없습니다. 쉘 장치와 쉘 장치에서 DR 작업을 수행하는 절차에 대한 고려 사항은 82 페이지 "SPARC: 쉘 장치에 대한 DR 클러스터링 참고 사항"을 참조하십시오.

이 작업을 수행하는 방법에 대한 자세한 지침은 Solaris OS용 Sun Cluster 시스템 관리 안내서의 "쉘 장치 동적 재구성"를 참조하십시오.

SPARC: 쉘 장치에 대한 DR 클러스터링 참고 사항

DR 보드 제거 작업이 쉘용으로 구성된 장치에 인터페이스가 있는 보드와 관련이 있으면 Sun Cluster 소프트웨어에서 이 작업을 거부합니다. 또한 Sun Cluster 소프트웨어는 해당 작업의 영향을 받는 쉘 장치도 식별합니다. DR 보드 제거 작업을 수행하기 전에 장치를 쉘 장치로 비활성화해야 합니다.

쉘을 관리하는 방법에 대한 자세한 내용은 Solaris OS용 Sun Cluster 시스템 관리 안내서의 5 장, "쉘 관리"를 참조하십시오.

SPARC: 클러스터 상호 연결 인터페이스에 대한 DR 클러스터링 고려 사항

DR 보드 제거 작업이 활성 클러스터 상호 연결 인터페이스가 있는 보드와 관련이 있으면 Sun Cluster 소프트웨어에서 작업을 거부합니다. 또한 Sun Cluster 소프트웨어는 해당 작업의 영향을 받는 인터페이스도 식별합니다. DR 작업에 성공하려면 먼저 Sun Cluster 관리 도구를 사용하여 활성 인터페이스를 비활성화해야 합니다.



주의 - Sun Cluster 소프트웨어에는 각 클러스터 노드에 작동 중인 다른 클러스터 노드 경로가 하나 이상 있어야 합니다. 다른 클러스터 노드에 대한 마지막 경로를 지원하는 독립 상호 연결 인터페이스를 비활성화하면 안됩니다.

이 작업을 수행하는 방법에 대한 자세한 지침은 **Solaris OS용 Sun Cluster 시스템 관리 안내서**의 “클러스터 상호 연결 관리”를 참조하십시오.

SPARC: 공용 네트워크 인터페이스에 대한 DR 클러스터링 참고 사항

DR 보드 제거 작업이 활성 공용 네트워크 인터페이스가 있는 보드와 관련이 있으면 Sun Cluster 소프트웨어에서 작업을 거부합니다. 또한 Sun Cluster 소프트웨어는 해당 작업의 영향을 받는 인터페이스도 식별합니다. 활성 네트워크 인터페이스가 있는 보드를 제거하기 전에 먼저 `if_mpadm(1M)` 명령을 사용하여 해당 인터페이스의 모든 트래픽을 `multipathing` 그룹의 작동 중인 다른 인터페이스로 전환합니다.



주의 - 비활성화된 네트워크 어댑터에서 DR 제거 작업을 수행하는 동안 나머지 네트워크 어댑터가 실패할 경우 가용성에 영향을 줍니다. DR 작업을 수행하는 동안 남은 어댑터를 페일오버할 수 없습니다.

공용 네트워크 인터페이스에서 DR 제거 작업을 수행하는 방법에 대한 자세한 내용은 **Solaris OS용 Sun Cluster 시스템 관리 안내서**의 “공용 네트워크 관리”를 참조하십시오.

FAQ

이 장은 Sun Cluster 시스템에 대하여 자주 문의하는 사항에 대한 응답으로 구성되어 있습니다. 응답은 주제별로 조직되어 있습니다.

고가용성 FAQ

질문: 고가용성 시스템이란 정확히 무엇입니까?

응답: Sun Cluster 시스템에서는 클러스터에서 응용 프로그램을 계속 실행하는 기능을 고가용성(HA)이라고 정의합니다. 장애가 발생하여 서버 시스템을 정상적으로 사용할 수 없을 때에도 응용 프로그램은 실행됩니다.

질문: 클러스터는 어떤 프로세스를 통해 높은 가용성을 제공합니까?

응답: 클러스터 프레임워크는 페일오버라고 하는 프로세스를 통해 고가용성 환경을 제공합니다. 페일오버는 장애가 발생한 노드로부터 작동 중인 다른 노드로 데이터 서비스 자원을 전환하기 위해 클러스터에서 수행하는 일련의 단계입니다.

질문: HA와 확장 가능 서비스간의 차이점은 무엇입니까?

응답: 고가용성 데이터 서비스에는 두 가지 유형이 있습니다.

- 페일오버
- 확장 가능

페일오버 데이터 서비스는 클러스터에서 한 번에 하나의 기본 노드에서만 응용 프로그램을 실행합니다. 다른 노드에서는 다른 응용 프로그램을 실행할 수 있지만, 각 응용 프로그램이 하나의 노드에서만 실행됩니다. 기본 노드에 장애가 발생하면 이 노드에서 실행 중인 응용 프로그램은 다른 노드로 페일오버되어 계속 실행됩니다.

확장 가능 서비스는 하나의 응용 프로그램을 여러 노드에 분산시켜서 하나의 논리 서비스를 작성합니다. 확장 가능 서비스는 실행되는 전체 클러스터에서 여러 노드와 프로세스를 조정합니다.

응용 프로그램마다 하나의 노드가 클러스터에 대한 물리적 인터페이스를 호스트합니다. 이러한 노드를 GIF(Global Interface) 노드라고 합니다. 여러 개의 GIF 노드가 클러스터에 존재할 수 있습니다. 각 GIF 노드는 확장 가능한 서비스에서 사용할 수 있는 하나 이상의 논리 인터페이스를 호스트합니다. 이러한 논리 인터페이스를 **전역 인터페이스**라고 합니다. 하나의 GIF 노드가 전역 인터페이스를 호스트하여 특정 응용 프로그램에 대한 모든 요청을 받고 Application Server를 실행하는 여러 노드로 이 요청을 전달합니다. GIF 노드에 장애가 발생하면 전역 인터페이스가 남아있는 노드로 페일오버합니다.

응용 프로그램을 실행 중인 노드에 장애가 발생하면 해당 응용 프로그램은 다른 노드에서 계속 실행되지만 성능은 다소 저하됩니다. 이 프로세스는 장애가 발생한 노드가 클러스터에 돌아갈 때까지 계속됩니다.

파일 시스템 FAQ

질문: 하나 이상의 클러스터 노드를고가용성 NFS 서버로 실행하고 다른 클러스터 노드는 클라이언트로 실행할 수 있습니까?

응답: 안됩니다. 루프백 마운트를 하면 안됩니다.

질문: Resource Group Manager의 제어를 받지 않는 응용 프로그램에 클러스터 파일 시스템을 사용할 수 있습니까?

응답: 예. 그러나 RGM의 제어가 없으면 응용 프로그램을 실행하고 있는 노드에 장애가 발생하면 직접 해당 응용 프로그램을 다시 시작해야 합니다.

질문: 모든 클러스터 파일 시스템의 /global 디렉토리에 마운트 지점이 있어야 합니까?

응답: 아닙니다. 그러나 /global과 같이 동일한 마운트 지점에 클러스터 파일 시스템을 두면 이러한 파일 시스템을 쉽게 구성하고 관리할 수 있습니다.

질문: 클러스터 파일 시스템을 사용하는 것과 NFS 파일 시스템을 내보내는 것이 어떤 차이가 있습니까?

응답: 여러 가지 차이점이 있습니다.

1. 클러스터 파일 시스템은 전역 장치를 지원합니다. NFS는 장치에 대한 원격 액세스를 지원하지 않습니다.
2. 클러스터 파일 시스템에는 전역 이름 공간이 있습니다. 하나의 마운트 명령만 필요합니다. NFS를 사용할 경우, 각 노드에서 파일 시스템을 마운트해야 합니다.
3. 클러스터 파일 시스템은 NFS를 수행하는 경우보다 많이 파일을 캐시합니다. 예를 들어 클러스터 파일 시스템은 읽기, 쓰기, 파일 잠금, 비동기 I/O를 위해 여러 노드에서 파일에 액세스할 때 파일을 캐시합니다.
4. 클러스터 파일 시스템은 향후에 원격 DMA 및 zero-copy 기능을 제공하는 고속 클러스터 상호 연결을 구축할 수 있도록 설계되었습니다.
5. 클러스터 파일 시스템에서 파일에 대한 등록 정보를 변경하면(예를 들어, chmod(1M) 명령 사용), 변경한 내용이 모든 노드에 즉시 적용됩니다. 내보낸 NFS 파일 시스템에서는 이 변경 내용을 적용하는 데 시간이 더 오래 걸릴 수 있습니다.

질문: 파일 시스템 /global/.devices/node@nodeID가 사용자의 클러스터 노드에 표시됩니다. 이 파일 시스템을 사용하여 가용성이 높은 전역 데이터를 저장할 수 있습니까?

응답: 이 파일 시스템은 전역 장치 이름 공간을 저장합니다. 이 파일 시스템은 일반적인 용도로 사용되지 않습니다. 이 파일 시스템은 전역이지만 전역 방식으로서는 액세스할 수 없습니다. 각 노드는 자체 전역 장치 이름 공간만 액세스합니다. 노드가 중단되면 다른 노드가 중단된 노드에 대한 이름 공간에 액세스할 수 있습니다. 이 파일 시스템은 가용성이 높지 않습니다. 전역 방식으로 액세스하거나 가용성이 높아야 하는 데이터를 저장할 경우에는 이 파일 시스템을 사용하면 안됩니다.

볼륨 관리 FAQ

질문: 모든 디스크 장치를 미러해야 합니까?

응답: 고가용성으로 간주되는 디스크 장치의 경우에는 미러해야 합니다. 그렇지 않으면, RAID-5 하드웨어를 사용하십시오. 모든 데이터 서비스는 고가용성 디스크 장치나 고가용성 디스크 장치에 마운트된 클러스터 파일 시스템을 사용해야 합니다. 이렇게 구성하면 하나의 디스크에 장애가 발생할 경우에도 안전합니다.

질문: 로컬 디스크(부트 디스크)에 하나의 볼륨 관리자를 사용하고 멀티 호스트 디스크에 다른 볼륨 관리자를 사용할 수 있습니까?

응답: SPARC: 이 구성은 로컬 디스크를 관리하는 Solaris 볼륨 관리자 소프트웨어와 멀티 호스트 디스크를 관리하는 VERITAS Volume Manager에서 지원됩니다. 다른 조합은 지원되지 않습니다.

x86: 아니요, 이 구성은 지원되지 않습니다. x86 기반 클러스터에서는 Solaris 볼륨 관리자만 지원됩니다.

데이터 서비스 FAQ

질문: 어떤 Sun Cluster 데이터 서비스를 사용할 수 있습니까?

응답: 지원되는 데이터 서비스 목록은 **Solaris OS용 Sun Cluster 3.1 8/05 릴리스 노트**의 “지원 제품”에 있습니다.

질문: Sun Cluster 데이터 서비스에서 어떤 응용 프로그램 버전이 지원됩니까?

응답: 지원되는 응용 프로그램 버전 목록은 **Solaris OS용 Sun Cluster 3.1 8/05 릴리스 노트**의 “지원 제품”에 있습니다.

질문: 자체 데이터 서비스를 작성할 수 있습니까?

응답: 예. 자세한 내용은 **Solaris OS용 Sun Cluster 데이터 서비스 개발 안내서**의 11 장, “DSDL API 함수”를 참조하십시오.

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

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

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

응답: Sun Cluster HA for NFS의 경우를 제외하고, 설명서에서 페일오버 모드 자원 그룹의 LogicalHostname 자원을 사용할 것을 권장하는 모든 경우 SharedAddress 자원 또는 LogicalHostname 자원을 혼용하여 사용할 수 있습니다. SharedAddress 자원을 사용하면 약간의 추가 오버헤드가 발생합니다. 클러스터 네트워킹 소프트웨어가 SharedAddress에 대해서는 구성되지만 LogicalHostname에 대해서는 구성되지 않기 때문입니다.

확장 가능 및 페일오버 데이터 서비스를 모두 구성하여 클라이언트에서 동일한 호스트 이름을 사용하여 두 서비스에 모두 액세스할 수 있게 하려면 SharedAddress 자원을 사용하는 것이 좋습니다. 이 경우 SharedAddress 자원은 페일오버 응용 프로그램 자원과 함께 한 개의 자원 그룹에 포함됩니다. 확장 가능 서비스 자원은 별도의 자원 그룹에 포함되어 SharedAddress 자원을 사용하도록 구성됩니다. 이렇게 하면 확장 가능 서비스와 페일오버 서비스에서 모두 SharedAddress 자원에 구성된 동일한 호스트 이름/주소 집합을 사용할 수 있습니다.

공용 네트워크 FAQ

질문: 어떤 공용 네트워크 어댑터가 Sun Cluster 시스템을 지원합니까?

응답: 현재 Sun Cluster 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.

질문: 페일오버에서 MAC 주소의 역할은 무엇입니까?

응답: 페일오버가 발생할 경우, 새로운 ARP (Address Resolution Protocol) 패킷이 생성되어 전체에 브로드캐스팅됩니다. 이러한 ARP 패킷에는 새로운 MAC 주소(노드가 페일오버한 새로운 물리적 어댑터와 이전 IP 주소가 있습니다. 네트워크의 다른 시스템이 패킷 중 하나를 수신할 경우 그 시스템은 해당 ARP 캐시에서 이전 MAC-IP 매핑을 지우고 새 매핑 정보를 사용합니다.

질문: Sun Cluster 시스템이 local-mac-address?=true 설정을 지원합니까?

응답: 예. 실제로 IP Network Multipathing에서는 local-mac-address? 가 true로 설정되어야 합니다.

SPARC 기반 클러스터의 OpenBoot PROM ok 프롬프트에서 eeprom(1M)을 사용하여 local-mac-address?를 설정할 수 있습니다. MAC 주소는 x86 기반 클러스터에서 BIOS를 부트한 후에 선택적으로 실행하는 SCSI 유틸리티를 사용하여 설정할 수도 있습니다.

질문: IP (Internet Protocol) Network Multipathing에서 어댑터 간에 스위치오버를 수행하면 어느 정도 지연될 수 있습니까?

응답: 몇 분 동안 지연될 수 있습니다. IP (Internet Protocol) Network Multipathing 스위치오버를 수행할 때 Gratuitous ARP를 전송하기 때문입니다. 그러나 클라이언트와 클러스터를 연결하는 라우터에서 Gratuitous ARP를 사용할 지는 확실하지 않습니다. 따라서 라우터에서 이 IP 주소에 대한 ARP 캐시 항목의 시간이 만료될 때까지 이 항목은 오래된 MAC 주소를 사용할 수 있습니다.

질문: 네트워크 어댑터 오류가 얼마나 빨리 감지됩니까?

응답: 기본 장애 감지 시간은 10초입니다. 알고리즘에서 장애 감지 시간을 맞추려 하지만 실제 시간은 네트워크 로드 에 따라 달라집니다.

클러스터 구성원 FAQ

질문: 모든 클러스터 구성원이 동일한 루트 암호를 사용해야 합니까?

응답: 각 클러스터 구성원에서 동일한 루트 암호를 가질 필요는 없습니다. 그러나 모든 노드에서 동일한 루트 암호를 사용하면 쉽게 클러스터를 관리할 수 있습니다.

질문: 노드의 부트 순서가 중요합니까?

응답: 대부분의 경우 그렇지 않습니다. 그러나 정보 유실을 막으려면 부트 순서가 중요합니다. 예를 들어, 노드 2가 쿼럼 장치를 소유하고 있을 때 노드 1이 중단된 상태에서 사용자가 노드 2를 중단시키면 노드 1을 다시 실행하기 전에 노드 2를 먼저 실행해야 합니다. 이 순서대로 실행하면 오래된 클러스터 구성 정보를 가진 노드를 실수로 사용하지 않도록 방지할 수 있습니다. 정보 유실에 대한 자세한 내용은 50 페이지 “장애 차단 정보”를 참조하십시오.

질문: 클러스터 노드에서 로컬 디스크를 미러해야 합니까?

응답: 예. 이 미러링이 요구 사항은 아니지만 클러스터 노드의 디스크를 미러링하면 미러되지 않은 디스크 장애로 인해 노드가 중단되는 것을 방지할 수 있습니다. 클러스터 노드의 로컬 디스크를 미러하면 시스템 관리에 오버헤드가 부가됩니다.

질문: 클러스터 구성원을 백업하는 데는 어떤 문제가 있습니까?

응답: 하나의 클러스터에 대해 여러 가지 백업 방법을 사용할 수 있습니다. 한 가지 방법은 한 개의 노드를 테이프 드라이브나 라이브러리가 연결된 백업 노드로 사용하는 것입니다. 그리고 나서, 데이터를 백업하기 위해 클러스터 파일 시스템을 사용합니다. 이 노드를 공유 디스크에 연결하지는 마십시오.

데이터를 백업 및 복원하는 방법에 대한 추가 정보는 **Solaris OS용 Sun Cluster 시스템 관리 안내서**의 9 장, “클러스터 백업 및 복원”를 참조하십시오.

질문: 보조 노드로 사용될 수 있는 노드 상태는 언제입니까?

응답: Solaris 8 및 Solaris 9:

재부트 후 노드가 로그인 프롬프트를 표시하면 해당 노드가 보조 노드가 될 수 있는 상태입니다.

Solaris 10:

multi-user-server 시점이 실행 중이면 해당 노드가 보조 노드가 될 수 있는 상태입니다.

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

클러스터 저장소 FAQ

질문: 어떻게 멀티 호스트 저장소의 가용성을 높입니까?

응답: 멀티 호스트 저장소는 한 개의 디스크가 손실되더라도 미러링이나 하드웨어 기반 RAID-5 컨트롤러를 통해 계속 사용할 수 있기 때문에 가용성이 높습니다. 멀티 호스트 저장소는 여러 개의 연결을 갖고 있으므로 연결된 노드 중 하나가 손상되더라도 작동을 계속할 수 있습니다. 또한 호스트 버스 어댑터, 케이블 또는 디스크 컨트롤러의 오류에 대해 각 노드에서 연결된 저장소까지의 중복 경로가 허용됩니다.

클러스터 상호 연결 FAQ

질문: 어떤 클러스터 상호 연결이 Sun Cluster 시스템을 지원합니까?

응답: 현재 Sun Cluster 시스템에서는 다음 클러스터 상호 연결을 지원합니다.

- 모든 SPARC 기반 및 x86 기반 클러스터에 있는 이더넷(100BASE-T Fast Ethernet 및 1000BASE-SX Gb)
- 모든 SPARC 기반 및 x86 기반 클러스터에 있는 Infiniband
- SPARC 기반 클러스터 전용의 SCI

질문: “케이블”과 전송 “경로”의 차이점은 무엇입니까?

응답: 클러스터 전송 케이블은 전송 어댑터와 스위치를 사용하여 구성합니다. 케이블은 구성 요소끼리 연결하는 방식으로 어댑터와 스위치를 결합시킵니다. 클러스터 토폴로지 관리자는 사용 가능한 케이블을 통해 노드 사이에 종단 간 전송 경로를 구축합니다. 케이블이 직접 전송 경로에 매핑되지 않습니다.

케이블은 관리자에 의해 정적으로 “활성화”되고 “비활성화”됩니다. 케이블에는 “상태(state)”(활성 또는 비활성)는 있지만 “상태(status)”는 없습니다. 케이블이 비활성화되면 구성되지 않은 것으로 간주됩니다. 비활성화된 케이블은 전송 경로로 사용할 수 없습니다. 이 케이블은 검사를 하지 않으므로 상태를 알 수 없습니다. 케이블 상태를 확인하려면 `scconf -p` 명령을 사용합니다.

전송 경로는 클러스터 토폴로지 관리자에 의해 동적으로 구성됩니다. 전송 경로의 “상태(status)”는 토폴로지 관리자에 의해 결정됩니다. 경로에는 “온라인” 또는 “오프라인” 상태(status)가 있습니다. 전송 경로의 상태를 확인하려면 `scstat (1M)` 명령을 사용합니다.

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
```

이 네 개의 케이블로 두 개의 전송 경로를 만들 수 있습니다.

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

클라이언트 시스템 FAQ

질문: 클러스터에서 사용할 경우 특수 클라이언트 요구 사항이나 제한 사항을 고려해야 합니까?

응답: 클라이언트 시스템은 다른 서버에서와 마찬가지로 클러스터에 연결합니다. 어떤 경우에는 데이터 서비스 응용 프로그램에 따라, 클라이언트가 데이터 서비스 응용 프로그램에 연결할 수 있도록 클라이언트측 소프트웨어를 설치하거나 다른 구성 변경 사항을 수행해야 할 수도 있습니다. 클라이언트측 구성 요구 사항에 대한 자세한 내용은 **Sun Cluster Data Services Planning and Administration Guide for Solaris OS**의 1장, “Planning for Sun Cluster Data Services”를 참조하십시오.

관리 콘솔 FAQ

질문: Sun Cluster 시스템에 관리 콘솔이 필요합니까?

응답: 예.

질문: 관리 콘솔은 클러스터 전용이어야 합니까? 아니면, 다른 작업에도 사용할 수 있습니까?

응답: Sun Cluster 시스템에는 전용 관리 콘솔이 필요하지 않지만 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.

질문: 관리 콘솔을 클러스터에 “가까이”(예: 같은 방) 놓아야 합니까?

응답: 하드웨어 서비스 제공업체에 확인해 보십시오. 제공업체는 콘솔을 클러스터와 근접한 곳에 놓으라고 요청할 수 있습니다. 콘솔이 같은 방에 있어야 할 기술적인 이유는 없습니다.

질문: 거리 요구 사항이 모두 충족되면 관리 콘솔이 둘 이상의 클러스터를 서비스할 수 있습니까?

응답: 예. 하나의 관리 콘솔에서 여러 클러스터를 제어할 수 있습니다. 또한 클러스터 사이에서 하나의 단말기 집중 장치를 공유할 수도 있습니다.

단말기 집중 장치 및 시스템 서비스 프로세서 FAQ

질문: Sun Cluster 시스템에 단말기 집중 장치가 필요합니까?

응답: Sun Cluster 3.0으로 시작하는 모든 소프트웨어 릴리스는 단말기 집중 장치 없이 실행됩니다. 장애 방지를 위해 단말기 집중 장치가 필요했던 Sun Cluster 2.2 제품과는 달리, 이후 버전의 제품에서는 단말기 집중 장치가 반드시 필요하지는 않습니다.

질문: 대부분의 Sun Cluster 서버는 단말기 집중 장치를 사용하지만 Sun Enterprise E1000 서버는 사용하지 않습니다. 그 이유는 무엇입니까?

응답: 단말기 집중 장치는 실제로 대부분의 서버에서 직렬-이더넷 변환기로 사용되기 때문에 단말기 집중 장치의 콘솔 포트는 직렬 포트입니다. Sun Enterprise E1000 서버에는 직렬 콘솔이 없습니다. SSP (System Service Processor)는 이더넷이나 jtag 포트를 통한 콘솔입니다. Sun Enterprise E1000 서버에서는 항상 SSP를 콘솔로 사용합니다.

질문: 단말기 집중 장치를 사용할 경우 어떤 점이 좋습니까?

응답: 단말기 집중 장치를 사용하면 네트워크상의 모든 원격 워크스테이션에서 콘솔 수준으로 각 노드에 액세스할 수 있습니다. 노드가 SPARC 기반 노드의 OpenBoot PROM(OBP)에 있거나 x86 기반 노드의 부트 하위 시스템에 있을 때에도 이런 방식으로 액세스할 수 있습니다.

질문: Sun에서 지원하지 않는 단말기 집중 장치를 사용할 경우 알아야 할 사항은 무엇입니까?

응답: Sun에서 지원하는 단말기 집중 장치와 다른 콘솔 장치의 가장 큰 차이점은 Sun의 단말기 집중 장치에는 특수 펌웨어가 있다는 것입니다. 이 펌웨어는 부트할 때 단말기 집중 장치에서 콘솔로 중단 신호가 전송되지 않도록 합니다. 중단 신호나 중단 신호로 해석될 수 있는 신호를 콘솔로 전송할 수 있는 콘솔 장치가 있으면 중단 신호가 노드를 종료시킵니다.

질문: Sun에서 지원하는 단말기 집중 장치의 잠긴 포트를 재부트하지 않고 해제할 수 있습니까?

응답: 예. 다시 설정해야 하는 포트 번호를 확인하고 다음 명령을 입력하십시오.

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

Sun에서 지원하는 단말기 집중 장치를 구성하고 관리하는 방법에 대한 자세한 내용은 다음 설명서를 참조하십시오.

- Solaris OS용 Sun Cluster 시스템 관리 안내서의 “Sun Cluster 관리 개요”
- Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS의 2 장, “Installing and Configuring the Terminal Concentrator”

질문: 단말기 집중 장치 자체에 장애가 발생하면 어떻게 해야 합니까? 다른 집중 장치를 준비해야 합니까?

응답: 아니요. 단말기 집중 장치에 장애가 발생해도 클러스터의 가용성은 유지됩니다. 집중 장치가 다시 서비스를 제공할 때까지 노드 콘솔에 연결할 수 없게 됩니다.

질문: 단말기 집중 장치를 사용할 경우에 보안 문제는 없습니까?

응답: 일반적으로 단말기 집중 장치는 시스템 관리자가 사용하는 소규모 네트워크에 연결되며 다른 클라이언트 액세스에 사용되는 네트워크에는 연결되지 않습니다. 특수 네트워크에 대한 액세스를 제한하여 보안을 제어할 수 있습니다.

질문: SPARC: 테이프 또는 디스크 드라이브에서 동적 재구성을 어떻게 사용합니까?

응답: 다음 단계를 수행하십시오.

- 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. 드라이브가 현재 작동하는 장치 그룹에 속하면 82 페이지 “SPARC: 디스크 및 테이프 드라이브에 대한 DR 클러스터링 참고 사항”로 이동합니다.
- 드라이브가 기본 노드의 구성 요소인지 아니면 보조 노드의 구성 요소인지 확인하십시오. 드라이브가 보조 노드의 구성 요소이면 DR 제거 작업을 수행할 수 있습니다.
- 드라이브가 기본 노드의 구성 요소이면 DR 제거 작업을 수행하기 전에 기본 노드와 보조 노드를 전환해야 합니다.



주의 - 보조 노드에 대한 DR 작업을 수행할 때 현재 기본 노드에 장애가 발생하면 클러스터 가용성이 영향을 받습니다. 새로운 보조 노드가 제공될 때까지 기본 노드를 페일오버할 수 없습니다.

색인

A

API, 65-66
APIs, 68
auto-boot? 매개 변수, 36

C

CCP, 26
CCR, 37
CCR(Cluster Configuration Repository), 37
CD-ROM 드라이브, 24
clprivnet 드라이버, 67
Cluster Control Panel, 26
CMM, 36
 페일패스트 기법, 36
 참조 페일패스트
CPU 시간, 70-79

D

/dev/global/ 이름 공간, 42-43
DID, 37-38
DR, 참조 동적 재구성
DSDL API, 68

E

E10000, 참조 Sun Enterprise E10000

F

FAQ, 85-94
 고가용성, 85-86
 공용 네트워크, 88-89
 관리 콘솔, 91-92
 단말기 집중 장치, 92-94
 데이터 서비스, 87-88
 볼륨 관리, 87
 시스템 서비스 프로세서, 92-94
 클라이언트 시스템, 91
 클러스터 구성원, 89-90
 클러스터 상호 연결, 90-91
 클러스터 저장소, 90
 파일 시스템, 86-87
FAQ(자주 물어보는 질문), 참조 FAQ

G

/global 마운트 지점, 43-45, 86-87

H

HA, 참조 고가용성
HAStoragePlus, 44-45
HAStoragePlus, 67-70

I

ID
 노드, 42

ID (계속)

장치, 37-38
in.mpathd 데몬, 79
ioctl, 51-52
IP Network Multipathing, 79-80
 페일오버 시간, 88-89
IP 주소, 87-88
IPMP, 참조 IP Network Multipathing

L

local_mac_address, 88-89
LogicalHostname, 참조 논리 호스트 이름

M

MAC 주소, 88-89
Multi-initiator SCSI, 23
multipathing, 79-80

N

N+1(스타) 토폴로지, 29-30
N*N(확장 가능) 토폴로지, 30-31
Network Time Protocol, 34
NFS, 45
NTP, 34
numsecondaries 등록 정보, 40

O

Oracle Parallel Server, 참조 Oracle Real
 Application Clusters
Oracle Real Application Clusters, 65

P

pernode 주소, 66-67
Persistent Group Reservation, 51-52
PGR, 참조 Persistent Group Reservation
preferenced 등록 정보, 40
pure 서비스, 62

R

Resource_project_name 등록 정보, 72-73
RG_project_name 등록 정보, 72-73
RGM, 60, 67-70, 70-79
RMAPI, 68

S

scha_cluster_get 명령, 67
scha_privatelink_hostname_node
 인수, 67
SCSI
 Multi-initiator, 23
 Persistent Group Reservation, 51-52
 예약 충돌, 51-52
 장애 차단, 50-51
scsi-initiator-id 등록 정보, 23
SharedAddress, 참조 공유 주소
Solaris Resource Manager, 70-79
 가상 메모리 제한 구성, 73-74
 구성 요구 사항, 72-73
 페일오버 시나리오, 74-79
Solaris 볼륨 관리자, 멀티 호스트 장치, 23
Solaris 프로젝트, 70-79
SSP, 참조 시스템 서비스 프로세서
sticky 서비스, 62
Sun Cluster, 참조 클러스터
Sun Enterprise E10000, 92-94
 관리 콘솔, 26
Sun Management Center(SunMC), 33
SunPlex, 참조 클러스터
SunPlex Manager, 33
syncdir 마운트 옵션, 45

U

UFS, 45

V

VERITAS 볼륨 관리자, 멀티 호스트 장치, 23
VxFS, 45

개

개발자, 클러스터 응용 프로그램, 17
개인 네트워크, 20

경

경로, 전송, 90-91

고

고가용성
FAQ, 85-86
데이터 서비스, 35
프레임워크, 35-37

공

공용 네트워크, 참조 네트워크, 공용
공유 주소, 58
대 논리 호스트 이름, 87-88
전역 인터페이스 노드, 59
확장 가능 데이터 서비스, 61-62

관

관리, 클러스터, 33-83
관리 인터페이스, 33-34
관리 콘솔, 26-27
FAQ, 91-92

교

교육, 11

구

구성
가상 메모리 제한, 73-74
데이터 서비스, 70-79
병렬 데이터베이스, 20
저장소, 37
쿼럼, 52-53

구성 (계속)

클라이언트 서버, 58
구성원, 참조 클러스터, 구성원

그

그룹
디스크 장치
참조 디스크, 장치 그룹

기

기본 노드, 59
기본 소유권, 디스크 장치 그룹, 40-41

네

네트워크
개인, 20
공용, 25
FAQ, 88-89
IP Network Multipathing, 79-80
동적 재구성, 83
인터페이스, 88-89
공유 주소, 58
논리 호스트 이름, 58
로드 균형 조정, 62-64
어댑터, 25, 79-80
인터페이스, 25, 79-80
자원, 58, 67-70

노

노드, 20-21
nodeID, 42
기본, 40-41, 59
백업, 89-90
보조, 40-41, 59
부트 순서, 89-90
전역 인터페이스, 59

논

- 논리 호스트 이름, 58
- 대 공유 주소, 87-88
- 페일오버 데이터 서비스, 61

단

- 단말기 집중 장치, FAQ, 92-94
- 단일 서버 모델, 58

데

- 데이터, 저장, 86-87
- 데이터 서비스, 58-64
 - API, 65-66
 - FAQ, 87-88
 - 개발, 65-66
 - 고가용성, 35
 - 구성, 70-79
 - 라이브러리 API, 66
 - 메소드, 60
 - 오류 모니터, 64
 - 자원, 67-70
 - 자원 그룹, 67-70
 - 자원 유형, 67-70
 - 지원, 87-88
- 클러스터 상호 연결, 66-67
- 페일오버, 61
- 확장 가능, 61-62

동

- 동시 액세스, 20
- 동적 재구성, 80-83
 - CPU 장치, 81
 - 공용 네트워크, 83
 - 디스크, 82
 - 메모리, 81-82
 - 설명, 81
 - 쿼럼 장치, 82
- 클러스터 상호 연결, 82-83
- 테이프 드라이브, 82

드

- 드라이버, 장치 ID, 37-38

등

- 등록 정보
 - Resource_project_name, 72-73
 - RG_project_name, 72-73
 - 변경, 40-41
 - 자원, 70
 - 자원 그룹, 70

디

- 디스크
 - SCSI 장치, 23
 - 동적 재구성, 82
 - 로컬, 24, 37-38, 42-43
 - 미러링, 89-90
 - 볼륨 관리, 87
 - 멀티 호스트, 37-38, 38-41, 42-43
 - 장애 차단, 50-51
 - 장치 그룹, 38-41
 - 기본 소유권, 40-41
 - 멀티 포트, 40-41
 - 페일오버, 39-40
 - 전역 장치, 37-38, 42-43
- 디스크 경로 모니터링, 45-48

로

- 로드 균형 조정, 62-64
- 로컬 디스크, 24
- 로컬 이름 공간, 42-43
- 로컬 파일 시스템, 44-45

루

- 루트 비밀번호, 89-90

마

마운트

/global, 86-87
syncdir 사용, 45
전역 장치, 43-45
파일 시스템, 43-45

매

매체, 이동식, 24
매핑, 이름 공간, 42-43

멀

멀티 포트 디스크 장치 그룹, 40-41
멀티 호스트 장치, 22-23

메

메모리, 81-82

백

백업 노트, 89-90

병

병렬 데이터베이스, 20

보

보드 제거, 동적 재구성, 82
보조 노트, 59
보호, 36

복

복구

장애 감지, 35
페일백 설정, 64

볼

볼륨 관리

FAQ, 87
RAID-5, 87
Solaris 볼륨 관리자, 87
VERITAS 볼륨 관리자, 87
로컬 디스크, 87
멀티 호스트 디스크, 87
멀티 호스트 장치, 23
이름 공간, 42

부

부트 디스크, 참조 디스크, 로컬
부트 순서, 89-90

비

비밀번호, 루트, 89-90

서

서버 모델, 58

소

소프트웨어 구성 요소, 21-22

속

속성, 참조 등록 정보

시

시간, 노트 간, 34
시스템 서비스 프로세서, 25-26, 26
FAQ, 92-94

실

실패, 보호, 36

쌍

쌍+N 토폴로지, 28-29

어

어댑터, 참조 네트워크, 어댑터

업

업무에 필수적인 응용 프로그램, 56

에

에이전트, 참조 데이터 서비스

예

예약 충돌, 51-52

오

오류 모니터, 64

응

응용 프로그램, 참조 데이터 서비스

응용 프로그램 개발, 33-83

응용 프로그램 배포, 53

응용 프로그램 통신, 66-67

이

이동식 매체, 24

이름 공간, 42-43

인

인터페이스

참조 네트워크, 인터페이스
관리, 33-34

자

자원, 67-70

등록 정보, 70

상태, 68-69

설정, 68-69

자원 관리, 70-79

자원 그룹, 67-70

등록 정보, 70

상태, 68-69

설정, 68-69

페일오버, 61

확장 가능, 61-62

자원 그룹 관리자, 참조 RGM

자원 유형, 44-45, 67-70

장

장애

감지, 35

복구, 35

차단, 50-51

페일백, 64

장치

ID, 37-38

멀티 호스트, 22-23

전역, 37-38

쿼럼, 48-57

장치 그룹, 38-41

등록 정보 변경, 40-41

저

저장소, 22-23

FAQ, 90

SCSI, 23

동적 재구성, 82

전

전역

이름 공간, 37, 42-43

로컬 디스크, 24

인터페이스, 59

확장 가능 서비스, 61

장치, 37-38, 38-41

전역, 장치 (계속)
 로컬 디스크, 24
 마운트, 43-45
전역 인터페이스 노드, 59

정

정보 분리, 49, 50-51
정보 유실, 49

중

종료, 36

차

차단, 50-51

커

커널, 메모리, 81-82

케

케이블, 전송, 90-91

큰

콘솔
 관리, 25-26, 26-27
 FAQ, 91-92
 시스템 서비스 프로세서, 25-26
 액세스, 25-26

퀴

쿼럼, 48-57
 구성, 52
 권장되는 구성, 54-56
 모범 사례, 53-54
 비전형적인 구성, 56

쿼럼 (계속)

 요구 사항, 52-53
 잘못된 구성, 56-57
 장치, 48-57
 장치, 동적 재구성, 82
 투표 수, 50

클

클라이언트 서버 구성, 58
클라이언트 시스템, 25
 FAQ, 91
 제한 사항, 91
클러스터
 공용 네트워크, 25
 공용 네트워크 인터페이스, 58
 관리, 33-83
 구성, 37, 70-79
 구성원, 20, 36
 FAQ, 89-90
 재구성, 36
 노드, 20-21
 데이터 서비스, 58-64
 매체, 24
 목표, 13-14
 백업, 89-90
 보드 제거, 82
 부트 순서, 89-90
 비밀번호, 89-90
 상호 연결, 20, 24-25
 FAQ, 90-91
 데이터 서비스, 66-67
 동적 재구성, 82-83
 어댑터, 24
 연결, 25
 인터페이스, 24
 지원, 90-91
 케이블, 25
 서비스, 14-15
 설명, 13-14
 소프트웨어 구성 요소, 21-22
 시간, 34
 시스템 관리자 관점, 15-17
 응용 프로그램 개발, 33-83
 응용 프로그램 개발자 관점, 17
 작업 목록, 18
 장점, 13-14

클러스터 (계속)
저장소 FAQ, 90
토폴로지, 27-31, 31-32
파일 시스템, 43-45, 86-87
FAQ
참조 파일 시스템
HASToragePlus, 44-45
사용, 44
하드웨어, 14-15, 19-27
클러스터 구성원 모니터, 36
클러스터 서버 모델, 58
클러스터 쌍 토폴로지, 27-28, 31-32

테
테이프 드라이브, 24

토
토폴로지, 27-31, 31-32
N+1(스타), 29-30
N*N(확장 가능), 30-31
쌍+N, 28-29
클러스터 쌍, 27-28, 31-32

파
파일 시스템
FAQ, 86-87
NFS, 45, 86-87
syncdir, 45
UFS, 45
VxFS, 45
고가용성, 86-87
데이터 저장소, 86-87
로컬, 44-45
마운트, 43-45, 86-87
사용, 44
전역, 86-87
클러스터, 43-45, 86-87
클러스터 파일 시스템, 86-87
파일 잠금, 43

패
패닉, 36, 52

페
페일백, 64
페일오버
데이터 서비스, 61
디스크 장치 그룹, 39-40
시나리오, Solaris Resource Manager, 74-79
페일패스트, 36, 51-52

프
프레임워크, 고가용성, 35-37
프로젝트, 70-79

하
하드웨어, 14-15, 19-27, 80-83
참조 디스크
참조 저장소
동적 재구성, 80-83
클러스터 상호 연결 구성 요소, 24

호
호스트 이름, 58

확
확장 가능 데이터 서비스, 61-62