Solaris OS용 Sun Cluster 개념 안내서

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

부품 번호: 819–0165–10
2004년 9월, 개정판 A

041217@10536

목차

머리말

*Solaris OS*용 *Sun™ Cluster* **개념 안내서**는 SPARC™ 및 x86 기반 시스템에서의 SunPlex™ 시스템에 대한 개념 및 참조 정보를 포함하고 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 문서는 Sun Cluster 소프트웨어에 대한 교육을 받은 전문 시스템 관리자를 위한 내용입니다. 이 문서는 계획이나 판촉용 안내서가 아닙니다. 이 문서를 읽을 때는 이미 시스템 요구 사항을 결정하고 필요한 장비와 소프트웨어를 구입한 상태이어야 합니다.

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

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	화면 상의 컴퓨터 출력과는 반대로 사용자가 직접 입력하는 사항입니다.	machine_name% **su** Password:
AaBbCc123	명령줄 자리 표시자: 실제 이름이나 값으로 대체됩니다.	파일을 삭제하려면 rm *filename*을 입력하십시오.
AaBbCc123	책 제목, 새로 나오는 용어, 강조 표시할 단어입니다.	**사용자 설명서**의 6장을 읽으십시오. 이를 *class* 옵션이라고 합니다. 파일을 저장하지 **마십시오**. (때때로 강조 표시는 온라인상에서 볼드로 표시됩니다.)

명령에 나오는 쉘 프롬프트의 예

다음 표에서는 C 쉘, Bourne 쉘 및 Korn 쉘에 대한 기본 시스템 프롬프트 및 수퍼유저 프롬프트를 보여줍니다.

표 P–2 쉘 프롬프트

쉘	프롬프트
C 쉘 프롬프트	machine_name%
C 쉘 수퍼유저 프롬프트	machine_name#
Bourne 쉘 및 Korn 쉘 프롬프트	$
Bourne 쉘 및 Korn 쉘 수퍼유저 프롬프트	#

관련 문서

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

주제	문서
개요	*Solaris OS*용 *Sun Cluster* **개요**
개념	*Solaris OS*용 *Sun Cluster* **개념 안내서**
하드웨어 설치 및 관리	*Sun Cluster 3.x Hardware Administration Manual for Solaris OS* 개별 하드웨어 관리 설명서
소프트웨어 설치	*Solaris OS*용 *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 OS*용 *Sun Cluster* **데이터 서비스 개발 안내서**
시스템 관리	*Solaris OS*용 *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 설명서 온라인 액세스

docs.sun.com℠ 웹 사이트에서 Sun 기술 관련 문서를 온라인으로 이용할 수 있습니다. docs.sun.com 아카이브를 찾아보거나 특정 책 제목 또는 주제를 검색할 수 있습니다. URL은 `http://docs.sun.com`입니다.

Sun 설명서 주문

Sun Microsystems에서는 제품 설명서를 인쇄물로 제공합니다. 설명서 목록 및 주문 방법은 `http://docs.sun.com`의 "인쇄본 문서를 구입하십시오"를 참조하십시오.

지원 받기

SunPlex 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 4/04)

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

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

1장

소개 및 개요

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

Solaris OS용 Sun Cluster 개념 안내서에서는 SunPlex 안내서의 주요 사용자에게 필요한 개념을 설명합니다. 이 안내서의 대상은 다음과 같습니다.

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

이 안내서를 읽은 후에 나머지 SunPlex 안내서를 보면 SunPlex 시스템을 완전히 이해할 수 있습니다.

이 장에서는 다음 내용에 대해 다룹니다.

- SunPlex 시스템을 소개하고 높은 수준의 개요를 제공합니다.
- SunPlex 사용자의 몇 가지 관점에 대해 설명합니다.
- SunPlex 시스템을 사용하기 전에 알아야 할 주요 개념을 설명합니다.
- 절차와 관련 정보가 포함된 SunPlex 문서와 주요 개념을 연결합니다.
- 클러스터 관련 작업을 각 작업 수행 절차가 포함된 문서에 연결합니다.

SunPlex 시스템 소개

SunPlex 시스템은 Solaris 운영 환경을 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.

SunPlex 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 복구 및 고가용성에 대한 자세한 내용은 *Solaris OS용 Sun Cluster* **개요**의 "Sun Cluster를 사용하여 응용 프로그램의 가용성 향상"을 참조하십시오.

고가용성에 대한 질문과 대답은 81 페이지 "고가용성 FAQ"를 참조하십시오.

SunPlex 시스템에 대한 세 가지 관점

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

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

하드웨어 설치 및 서비스 관점

하드웨어 서비스 담당자의 경우에는 SunPlex 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

권장하는 하드웨어 개념 참조 절

다음 절에는 위의 주요 개념과 관련된 자료가 있습니다.

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

관련 SunPlex 문서

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

Sun Cluster 3.x Hardware Administration Manual for Solaris OS

시스템 관리자 관점

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

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

권장하는 시스템 관리자 개념 참조 절

다음 절에는 위의 주요 개념과 관련된 자료가 있습니다.

- 31 페이지 “관리 인터페이스”
- 32 페이지 “클러스터 시간”
- 33 페이지 “고가용성 프레임워크”
- 35 페이지 “전역 장치”
- 36 페이지 “디스크 장치 그룹”
- 39 페이지 “전역 이름 공간”
- 40 페이지 “클러스터 파일 시스템”
- 42 페이지 “디스크 경로 모니터링”
- 47 페이지 “장애 차단 정보”
- 56 페이지 “데이터 서비스”

관련 SunPlex 문서 – 시스템 관리자

다음 SunPlex 문서에는 시스템 관리 개념과 관련된 절차와 정보가 있습니다.

- *Solaris OS*용 *Sun Cluster* **소프트웨어 설치 안내서**
- *Solaris OS*용 *Sun Cluster* **시스템 관리 안내서**
- *Sun Cluster Error Messages Guide for Solaris OS*
- *Solaris OS*용 *Sun Cluster 3.1 9/04* **릴리스 노트**
- *Sun Cluster 3.x Release Notes Supplement*

응용 프로그램 프로그래머 관점

SunPlex 시스템은 Oracle(SPARC 기반 시스템), NFS, DNS, Sun™ Java System Web Server(이전 명칭은 Sun Java System Web Server), Apache Web Server(SPARC 기반 시스템) 및 Sun Java System Directory Server(이전 명칭은 Sun Java System Directory Server)와 같은 응용 프로그램을 위해 **데이터 서비스**를 제공합니다. 별도 구입한 응용 프

로그램이 Sun Cluster 소프트웨어의 제어를 통해 실행되도록 구성하면 데이터 서비스가 완성됩니다. Sun Cluster 소프트웨어는 응용 프로그램을 시작, 중지 및 모니터하는 구성 파일과 관리 방법을 제공합니다. 새 페일오버 또는 확장 가능 서비스를 만들어야 하는 경우 SunPlex 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 방법을 개발할 수 있습니다.

주요 개념 – 응용 프로그램 프로그래머

응용 프로그램 프로그래머는 다음을 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 서비스. 프로그래머가 클러스터 환경에 맞게 응용 프로그램을 구성하려면 프로그램이나 스크립트를 작성할 때 가장 적합한 도구를 결정해야 합니다.

권장하는 응용 프로그램 프로그래머 개념 참조 절

다음 절에는 위의 주요 개념과 관련된 자료가 있습니다.

- 56 페이지 "데이터 서비스"
- 65 페이지 "자원, 자원 그룹 및 자원 유형"
- 4 장

관련 SunPlex 문서 – 응용 프로그램 프로그래머

다음 SunPlex 문서에는 응용 프로그램 프로그래머 개념과 관련된 절차 및 정보가 있습니다.

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

SunPlex 시스템 작업

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

표 1–1 작업 맵: 문서에 사용자 작업 매핑

수행할 작업	참조할 문서
클러스터 하드웨어 설치	*Sun Cluster 3.x Hardware Administration Manual for Solaris OS*
클러스터에 Solaris 소프트웨어 설치	*Solaris OS*용 *Sun Cluster* **소프트웨어 설치 안내서**
SPARC: Sun™ Management Center 소프트웨어 설치	*Solaris OS*용 *Sun Cluster* **소프트웨어 설치 안내서**
Sun Cluster 소프트웨어 설치 및 구성	*Solaris OS*용 *Sun Cluster* **소프트웨어 설치 안내서**
볼륨 관리 소프트웨어 설치 및 구성	*Solaris OS*용 *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.x Hardware Administration Manual for Solaris OS*
Sun Cluster 소프트웨어 관리	*Solaris OS*용 *Sun Cluster* **시스템 관리 안내서**
볼륨 관리 소프트웨어 관리	*Solaris OS*용 *Sun Cluster* **시스템 관리 안내서** 및 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 OS*용 *Sun Cluster* **데이터 서비스 개발 안내서**

2장

주요 개념 – 하드웨어 서비스 제공업체

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

SunPlex 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

SunPlex 시스템을 사용하면 25 페이지 “SPARC: Sun Cluster 토폴로지 예”에서 설명하는 다양한 구성 방법으로 이 구성 요소를 결합할 수 있습니다.

2 노드 클러스터 구성 예는 *Solaris OS용 Sun Cluster* **개요**의 “Sun Cluster 하드웨어 환경”을 참조하십시오.

클러스터 노드

클러스터 노드는 Solaris 운영 환경 및 Sun Cluster 소프트웨어를 모두 실행하는 시스템으로, 클러스터의 현재 구성원(**클러스터 구성원**)이거나 구성원이 될 수 있는 시스템입니다.

SPARC: Sun Cluster 소프트웨어를 사용하면 클러스터에 2 - 8개의 노드를 가질 수 있습니다. 지원되는 노드 구성은 25 페이지 “SPARC: Sun Cluster 토폴로지 예”를 참조하십시오.

x86: Sun Cluster 소프트웨어를 사용하면 클러스터에 2개의 노드를 가질 수 있습니다. 지원되는 노드 구성은 28 페이지 “x86: Sun Cluster 토폴로지 예”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 서비스 구성에서는 노드가 멀티 호스트 장치에 직접 연결되지 않아도 요청을 처리할 수 있습니다.

또한 병렬 데이터베이스의 노드는 모든 디스크에 동시에 액세스합니다. 병렬 데이터베이스 구성에 대한 자세한 내용은 20 페이지 “멀티 호스트 장치” 및 3 장을 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. 페일오버를 위해 모든 노드에 다른 노드의 워크로드를 백업하거나 보조 노드가 될 수 있을 만큼 충분한 용량이 있어야 합니다.

각 노드는 개별 루트(/) 파일 시스템을 부트합니다.

클러스터 하드웨어 구성원에 대한 소프트웨어 구성 요소

클러스터 구성원의 기능을 하려면 다음 소프트웨어가 설치되어야 합니다.

- Solaris 운영 환경
- 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 Manager와 같은 소프트웨어 볼륨 관리자가 필요하지 않을 수 있습니다.

- Solaris 운영 환경, Sun Cluster 및 볼륨 관리 소프트웨어 설치 방법에 대한 내용은 *Solaris OS***용** *Sun Cluster* **소프트웨어 설치 안내서**를 참조하십시오.
- 데이터 서비스의 설치 및 구성 방법에 대한 내용은 *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*를 참조하십시오.
- 이전의 소프트웨어 구성 요소에 대한 개념은 3 장을 참조하십시오.

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

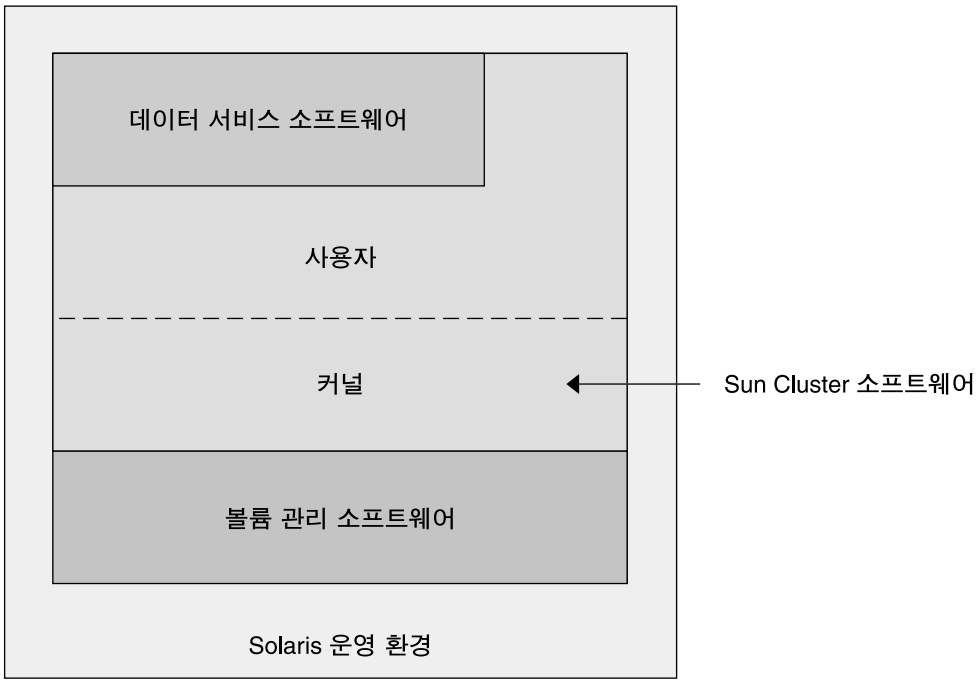


그림 2–1 높은 수준의 Sun Cluster 소프트웨어 구성 요소 관계

클러스터 구성원에 대한 질문과 대답은 4 장을 참조하십시오.

멀티 호스트 장치

한 번에 여러 노드에 연결될 수 있는 디스크는 멀티 호스트 장치입니다. Sun Cluster 환경에서 멀티 호스트 저장소를 사용하면 디스크 가용성을 높일 수 있습니다. Sun Cluster에서 2 노드 클러스터가 쿼럼을 설정하려면 멀티 호스트 저장소가 필요합니다. 3 노드 이상 클러스터에서는 멀티 호스트 저장소가 필요하지 않습니다.

멀티 호스트 장치에는 다음과 같은 기능이 있습니다.

- 노드 하나에 장애가 발생해도 계속 작동합니다.
- 응용 프로그램 데이터를 저장하고 응용 프로그램 바이너리와 구성 파일도 저장할 수 있습니다.
- 노드 장애로부터 보호합니다. 클라이언트 요청이 하나의 노드를 통해 데이터에 액세스할 때 노드에 장애가 발생하면 동일한 디스크에 직접 연결된 다른 노드를 사용하도록 요청이 전환됩니다.
- 멀티 호스트 디스크는 디스크를 "마스터"하는 기본 노드를 통하거나 로컬 경로를 통한 직접 동시 액세스에 의해 전역적으로 액세스됩니다. 현재는 Oracle Real Application Clusters만이 직접 동시 액세스를 사용합니다.

볼륨 관리자는 멀티 호스트 장치의 데이터 중복에 대한 미러된 구성 또는 RAID-5 구성에 제공됩니다. 현재 Sun Cluster는 SPARC 기반 클러스터에서만 사용 가능한 Solaris 볼륨 관리자™ 및 VERITAS Volume Manager를 볼륨 관리자로 지원하고 여러 하드웨어 RAID 플랫폼에서 RDAC RAID-5 하드웨어 컨트롤러를 지원합니다.

멀티 호스트 장치를 디스크 미러링 및 스트라이핑과 결합하면 노드 장애와 각 디스크 장애로부터 보호할 수 있습니다.

멀티 호스트 저장소에 대한 질문과 대답은 4 장을 참조하십시오.

Multi-Initiator SCSI

이 절의 내용은 SCSI 저장 장치에만 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 어댑터를 *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 종단 장치가 하나인 SCSI 장치나 차등 SCSI 장치로 구성된 경우에 이러한 구성을 Multi-initiator SCSI라고 합니다. 이 용어가 의미하는 것처럼 SCSI 버스에는 둘 이상의 SCSI initiator가 있습니다.

SCSI 스펙에서는 SCSI 버스에 있는 각 장치가 고유한 SCSI 주소를 갖도록 요구합니다 (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.

다음 도구 중 하나를 사용하여 `scsi-initiator-id` 등록 정보를 설정하는 방법으로 이 호스트 어댑터에 대해 선택된 SCSI 주소를 변경할 수 있습니다.

- `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 대하여 고유한 `scsi-initiator-id`를 설정하는 방법은 *Sun Cluster Hardware Collection*에서 각 디스크 인클로저에 대해 설명하는 장에 있습니다.

로컬 디스크

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

이러한 디스크의 파일 시스템을 전역 마운트 지점 아래에 놓아서 다른 노드가 사용할 수 있도록 만들 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 미러할 수 있지만, 노드 장애로부터 보호하지는 않습니다.

전역 장치에 대한 자세한 내용은 35 페이지 “전역 장치” 절을 참조하십시오.

이동식 매체

테이프 드라이브 및 CD-ROM 드라이브와 같은 이동식 매체가 클러스터에서 지원됩니다. 일반적으로, 클러스터링되지 않은 환경에서와 동일한 방법으로 이러한 장치를 설치, 구성하고 서비스를 제공할 수 있습니다. 이 장치는 Sun Cluster에 전역 장치로 구성되므로 클러스터의 모든 노드에서 각 장치에 액세스할 수 있습니다. 이동식 매체의 설치 및 구성에 대한 내용은 *Sun Cluster 3.x Hardware Administration Manual for Solaris OS*를 참조하십시오.

전역 장치에 대한 자세한 내용은 35 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 액세스할 수 있다고 가정합니다.

단일 장애 지점이 발생하지 않게 하려면 물리적으로 독립된 최소한 두 개 이상의 중복 네트워크를 통해 클러스터를 상호 연결하여 모든 노드를 연결해야 합니다. 두 노드 사이에 물리적으로 독립된 여러 개의 네트워크(2 - 6개)를 사용할 수 있습니다. 클러스터 상호 연결은어댑터, 연결 장치, 케이블 등의 세 가지 하드웨어로 구성됩니다.

다음은 이러한 하드웨어 구성 요소 각각에 대한 설명입니다.

- 어댑터 – 각 클러스터 노드에 위치하는 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. 이 연결 장치는 바로 전달 기능과 전환 기능을 수행하여 두 개를 초과하는 노드를 함께 연결할 수 있게 합니다. 2 노드 클러스터에서는 노드들이 각 노드의 중복 어댑터에 연결된 물리적인 중복 케이블을 통해 서로 직접 연결될 수 있으므로, 연결 장치가 필요없습니다. 일반적으로 세 개 이상의 노드로 된 구성에는 연결 장치가 필요합니다.
- 케이블 – 두 네트워크 어댑터 사이 또는 어댑터와 연결 장치 사이의 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 Network Multipathing 소프트웨어가 호출되어 오류가 있는 인터페이스를 그룹 내의 다른 어댑터에 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 실행되는 서버측 응용 프로그램에서 제공하는 데이터나 다른 서비스를 사용합니다.

클라이언트 시스템은 고가용성 시스템이 아닙니다. 클러스터의 데이터 및 응용 프로그램은 고가용성입니다.

클라이언트 시스템에 대한 질문과 대답은 4 장을 참조하십시오.

콘솔 액세스 장치

모든 클러스터 노드에 대하여 콘솔 액세스가 있어야 합니다. 콘솔에 액세스하려면 클러스터 하드웨어와 함께 구입한 단말기 집중 장치, Sun Enterprise E10000™ 서버(SPARC 기반 클러스터용)의 SSP(System Service Processor), Sun Fire™ 서버(역시 SPARC 기반 클러스터용)의 시스템 컨트롤러 또는 각 노드의 `ttya`에 액세스할 수 있는 기타 장치를 사용하십시오.

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/console`에 액세스할 수 있습니다. 따라서 네트워크 상의 모든 원격 워크스테이션에서 콘솔 레벨로 각 노드에 액세스할 수 있습니다.

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

콘솔에 액세스하는 다른 방법으로는 다른 단말기 집중 장치, 다른 노드로부터의 `tip(1)` 직렬 포트 액세스, 단순 단말기 사용 등이 있습니다. Sun™ 키보드 및 모니터를 사용하거나 하드웨어 서비스 제공업체에서 지원하는 다른 직렬 포트 장치를 사용할 수 있습니다.

관리 콘솔

관리 콘솔이라고 부르는 전용 UltraSPARC® 워크스테이션이나 Sun Fire™ V65x 서버를 사용하여 활성 클러스터를 관리할 수 있습니다. 일반적으로 관리 콘솔에서는 Sun Management Center™ 제품(SPARC 기반 클러스터 전용)을 위한 Sun Cluster 모듈 및 CCP(Cluster Control Panel)와 같은 관리 도구 소프트웨어를 설치하고 실행합니다. CCP에서 `cconsole`을 사용하면 하나 이상의 노드 콘솔을 한 번에 연결할 수 있습니다. CCP 사용 방법은 *Sun Cluster System Administration Guide*의 내용을 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 SSP(System Service Processor)로 로그인하고 `netcon(1M)` 명령으로 연결할 수 있는 기능이 있어야 합니다.

일반적으로, 노드를 모니터 없이 구성합니다. 그러면 단말기 집중 장치에 연결되어 있는 관리 콘솔과 노드의 직렬 포트에 연결된 단말기 집중 장치에서 `telnet` 세션을 통해 노드 콘솔에 액세스합니다. (Sun Enterprise E10000 server의 경우에는 SSP에서 연결합니다.)자세한 내용은 23 페이지 "콘솔 액세스 장치"를 참조하십시오.

Sun Cluster에는 전용 관리 콘솔이 필요 없지만 전용 관리 콘솔을 사용하면 다음과 같은 이점이 있습니다.

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

관리 콘솔에 대한 질문과 대답은 4 장을 참조하십시오.

SPARC: Sun Cluster 토폴로지 예

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

- SPARC 기반 시스템으로 구성된 Sun Cluster는 구현한 저장 장치 구성과 상관 없이 한 클러스터에서 최대 8개의 노드를 지원합니다.
- 공유 저장 장치는 저장 장치에서 지원하는 만큼의 노드에 연결될 수 있습니다.
- 공유 저장 장치를 클러스터의 모든 노드에 연결할 필요는 없지만그러나 이러한 저장 장치는 최소 두 개의 노드에 연결해야 합니다.

Sun Cluster는 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: 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, 모든 노드가 응용 프로그램 데이터를 저장하는 디스크에 직접 연결되어 있지 않아도 세 개 이상의 노드가 확장 가능 서비스나 병렬 데이터베이스를 실행하는 두 쌍 구성도 사용할 수 있습니다.

다음 그림은 클러스터된 쌍 구성을 보여줍니다.

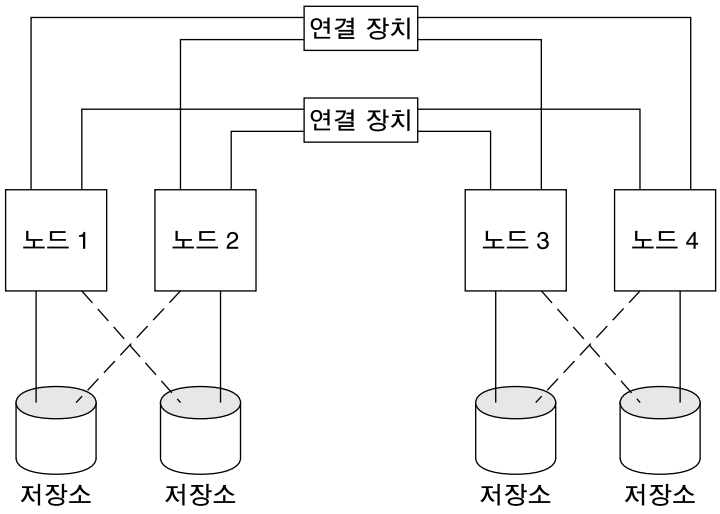


그림 2–2 SPARC: 클러스터 쌍 토폴로지

SPARC: 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.

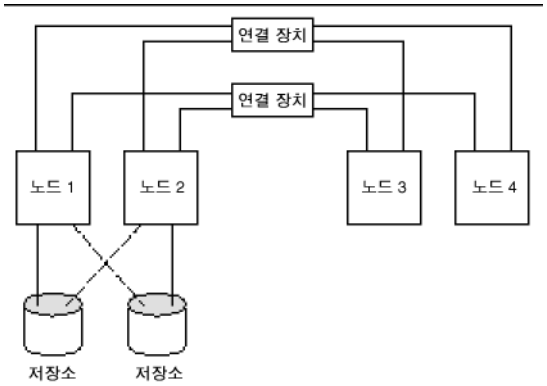


그림 2–3 SPARC: 쌍+N 토폴로지

SPARC: 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. 그곳에서 자원은 (자동 또는 수동으로) 다시 기본 노드로 전환될 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.

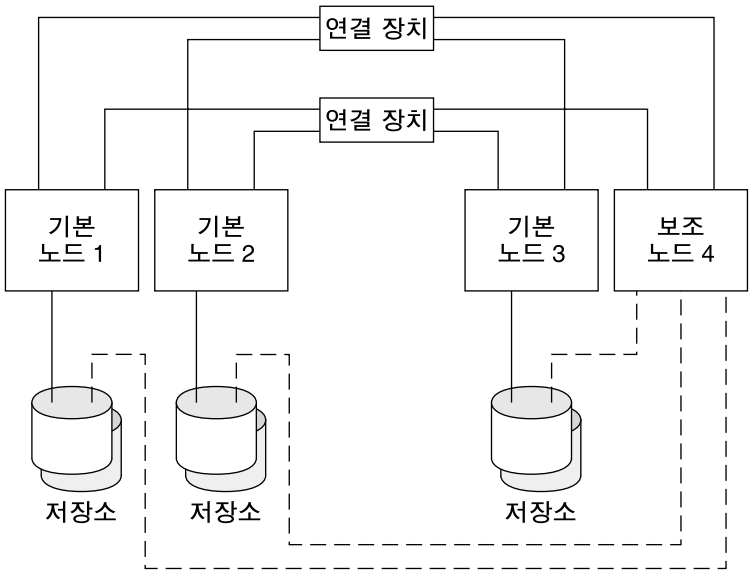


그림 2–4 SPARC: N+1 토폴로지

SPARC: N*N(확장 가능) 토폴로지

N*N 토폴로지를 사용하면 클러스터 내의 모든 공유 저장 장치가 해당 클러스터의 모든 노드에 연결할 수 있습니다. 가용성이 높은 응용 프로그램에서는 이 토폴로지를 사용하여 서비스 성능 감소 없이 한 노드에서 다른 노드로 페일오버할 수 있습니다. 페일오버가 발생하면 새 노드가 개별 상호 연결 대신 로컬 경로를 사용하여 저장 장치에 액세스할 수 있습니다.

다음 그림은 N*N 구성을 보여줍니다.

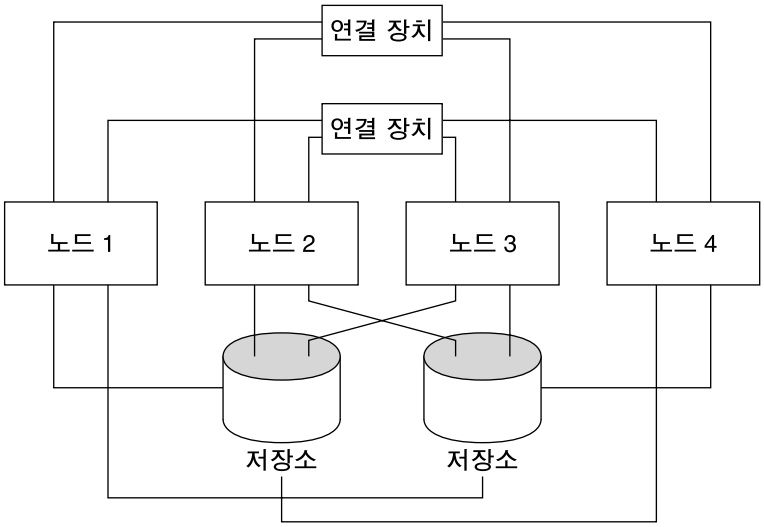


그림 2–5 SPARC: N*N 토폴로지

x86: Sun Cluster 토폴로지 예

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: 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 제어에 의해 작동합니다. 이 토폴로지를 사용하여 해당 쌍에서 병렬 데이터베이스, 페일오버 또는 확장 가능 응용 프로그램을 실행할 수 있습니다.

다음 그림은 클러스터된 쌍 구성을 보여줍니다.

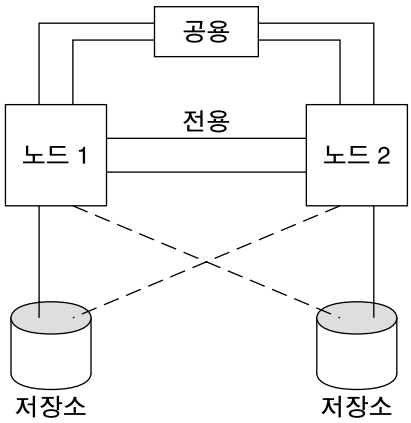


그림 2–6 x86: 클러스터 쌍 토폴로지

3장

주요 개념 – 관리 및 응용 프로그램 개발

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

이 정보는 주로 SunPlex API 및 SDK를 사용하는 시스템 관리자 및 응용 프로그램 개발자를 위한 내용입니다. 클러스터 시스템 관리자는 이 정보를 클러스터 소프트웨어의 설치, 구성 및 관리를 준비하는 데 사용할 수 있습니다. 응용 프로그램 개발자는 이 정보를 사용하여 작업할 클러스터 환경을 알 수 있습니다.

관리 인터페이스

여러 사용자 인터페이스에서 SunPlex 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 인터페이스 외에도 `scinstall` 및 `scsetup` 등의 유틸리티가 있으며 선택한 설치 및 구성 작업을 간소화합

니다. 또한 SunPlex 시스템에는 Sun Management Center의 일부로 실행되어 일부 클러스터 작업에 GUI를 제공하는 모듈이 있습니다. 이 모듈은 SPARC 기반 클러스터에서만 사용할 수 있습니다. 관리 인터페이스에 대한 자세한 설명은 *Solaris OS*용 *Sun Cluster* **시스템 관리 안내서**의 "관리 도구"를 참조하십시오.

클러스터 시간

클러스터에서 모든 노드들 사이의 시간은 동기화되어야 합니다. 클러스터 작동에서 외부 시간 소스를 사용하여 클러스터 노드를 동기화할 것인지는 중요하지 않습니다. SunPlex 시스템은 NTP(Network Time Protocol)를 사용하여 노드 사이의 시간을 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 환경을 설치할 경우, 노드에 대한 기본 시간 및 날짜 설정을 변경할 수 있는 기회가 주어집니다. 일반적으로 출하 시의 기본값을 승인할 수 있습니다.

`scinstall(1M)` 명령을 사용하여 Sun Cluster 소프트웨어를 설치할 때 클러스터에 NTP를 구성하는 단계가 프로세스에 포함됩니다. Sun Cluster 소프트웨어는 템플리트 파일 `ntp.cluster`가 있습니다(설치된 클러스터 노드에서 `/etc/inet/ntp.cluster` 참조). 이 파일은 노드 하나를 "기본" 노드로 사용하여 모든 클러스터 노드 사이에 피어 관계를 구성합니다. 노드는 독립된 호스트 이름으로 식별되고 클러스터 상호 연결 사이에 시간이 동기화됩니다. NTP를 위해 클러스터를 구성하는 방법은 *Solaris OS*용 *Sun Cluster* **소프트웨어 설치 안내서**의 "Sun Cluster 소프트웨어 설치 및 구성"을 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 환경을 설치할 때 시간을 잘못 설정하여 변경하려는 경우에는 *Solaris OS*용 *Sun Cluster* **시스템 관리 안내서**의 "클러스터 관리"에 있는 절차를 수행하십시오.

고가용성 프레임워크

SunPlex 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 SunPlex 구성 요소 장애(하드웨어 및 소프트웨어 모두)와 고가용성 프레임워크에 구축된 복구 작업을 보여줍니다.

표 3–1 SunPlex 장애 감지 및 복구 수준

장애가 발생한 클러스터 구성 요소	소프트웨어 복구	하드웨어 복구
데이터 서비스	HA API, HA 프레임워크	없음
공용 네트워크 어댑터	IP 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 전용인 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 프레임워크는 가용성이 높기 때문에 노드 장애를 즉시 탐지하고 남은 노드에 프레임워크 자원을 제공하는 새로운 서버를 구성합니다. 모든 프레임워크를 동시에 사용하지 못하는 경우는 없습니다. 손상된 노드의 영향을 받지 않는 프레임워크 자원은 복구가 진행 중인 동안에도 아무 제한 없이 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 자원을 사용하는 응용 프로그램(데이터 서비스)에 투명하게 복구됩니다. 프레임워크 자원 액세스의 시멘틱은 노드 실패에서 완전하게 보존됩니다. 응용 프로그램은 프레임워크 자원 서버가 다른 노드로 이동되었다는 것을 간단하게 알릴 수 없습니다. 단일 노드의 장애는 다른 노드의 디스크에 대한 대체 하드웨어 경로가 존재한다면, 이 노드에 접속된 파일, 장치 및 디스크 볼륨을 사용하는 나머지 노드에 있는 프로그램에서는 전혀 인식할 수 없습니다. 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 멤버쉽에서 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 멤버쉽에서의 변경을 발견하면 CMM은 클러스터의 동기화된 구성을 수행하며, 이 때 클러스터 자원은 클러스터의 새로운 멤버쉽을 기초로 재분배될 수 있습니다.

이전의 Sun Cluster 소프트웨어 릴리스와 달리, CMM은 완전히 커널에서 실행됩니다.

클러스터가 여러 개의 클러스터로 분할되지 않도록 보호하는 방법에 대한 자세한 내용은 47 페이지 "장애 차단 정보"를 참조하십시오.

페일패스트 기법

CMM이 노드에서 치명적인 문제를 확인하면 클러스터 프레임워크를 호출하여 강제로 해당 노드를 종료시키고 클러스터 구성원에서 제거합니다. 이러한 작업을 수행하는 기법을 **페일패스트**라고 합니다. 페일패스트 기법은 다음 두 가지 방법으로 노드를 중지시킵니다.

- 노드가 클러스터를 벗어나서 쿼럼을 사용하지 않고 새 클러스터를 시작하려고 시도하면 공유 디스크에 액세스하지 못하도록 "차단"됩니다. 페일패스트 사용에 대한 자세한 내용은 47 페이지 "장애 차단 정보"를 참조하십시오.
- 하나 이상의 클러스터 관련 데몬이 중단되면(`clexecd`, `rpc.pmfd`, `rgmd` 또는 `rpc.ed`) CMM에서 장애가 확인되어 노드가 종료됩니다. 클러스터 데몬의 종료로 노드가 패닉 상태가 되면 해당 노드의 콘솔에 다음과 비슷한 메시지가 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
```

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. OpenBoot PROM `ok` 프롬프트에서 `eeprom`(1M)을 사용하여 `auto-boot?`를 설정할 수 있습니다.

CCR(Cluster Configuration Repository)

CCR은 업데이트 사항에 대해 2단계 완결 알고리즘을 사용합니다. 즉, 모든 클러스터 구성원에 대해 성공적으로 업데이트가 완료되어야 하고, 그렇지 않은 경우에는 롤백됩니다. CCR은 클러스터 상호 연결을 사용하여 분배된 업데이트 사항을 적용합니다.

주의 – CCR은 텍스트 파일로 구성되어 있지만 직접 CCR 파일을 편집하면 안됩니다. 각 파일에는 일관성이 유지되도록 체크섬 레코드가 포함됩니다. 수동으로 CCR 파일을 업데이트하면 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 업데이트합니다.

전역 장치

SunPlex 시스템은 장치의 물리적 위치와 관계 없이 어떤 노드에서나 클러스터의 모든 장치에 액세스할 수 있도록 **전역 장치**를 사용하여 전체 클러스터에 고가용성을 제공합니다. 일반적으로 전역 장치에 대한 액세스를 제공하는 동안 노드에 장애가 발생하면 Sun Cluster 소프트웨어가 자동으로 장치에 대한 다른 경로를 찾아서 해당 경로로 액세스를 전환합니다. SunPlex 전역 장치에는 디스크, CD-ROM 및 테이프가 포함됩니다. 지원되는 멀티 포트 전역 장치는 디스크뿐입니다. 즉, CD-ROM 및 테이프 장치는 현재 고가용성 장치가 아닙니다. 각 서버의 로컬 디스크 역시 멀티 포트 상태가 아니므로 고가용성 장치가 아닙니다.

클러스터는 클러스터 내의 각 디스크, CD-ROM 및 테이프 장치에 고유 ID를 자동으로 할당합니다. 이러한 할당을 통해 클러스터의 모든 노드에서 각 장치에 일관되게 액세스할 수 있습니다. 전역 장치 이름 공간은 `/dev/global` 디렉토리에 저장됩니다. 자세한 내용은 39 페이지 “전역 이름 공간”을 참조하십시오.

멀티 포트 전역 장치는 장치에 대한 한 개 이상의 경로를 제공합니다. 멀티 호스트 디스크의 경우, 디스크는 여러 노드에 의해 호스팅된 디스크 장치 그룹의 일부분이므로 멀티 호스트 디스크는 가용성이 높아집니다.

DID(장치 ID)

Sun Cluster 소프트웨어는 DID 의사 드라이버라는 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(장치 ID) 의사 드라이버는 클러스터 전역 장치 액세스 기능의 필수적인 부분입니다. DID 드라이버는 클러스터의 모든 노드를 규명하고 고유한 디스크 장치 목록을 만들어, 클러스터의 모든 노드에서 일관되는 고유한 주 번호와 부 번호를 각각에 할당합니다. 전역 장치에 대한 액세스는 디스크에 `c0t0d0`을 사용했던 이전의 Solaris 장치 ID 대신 DID 드라이버에 의해 할당되는 고유 장치 ID를 사용하여 수행됩니다.

이러한 방법을 사용하면 디스크 장치를 이용하는 응용 프로그램(원래 장치를 사용하는 볼륨 관리자나 응용 프로그램)이 일관된 경로를 사용하여 장치에 액세스할 수 있게 합니다. 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, 노드 1에는 멀티 호스트 디스크가 c1t2d0으로 보이고 노드 2에는 동일한 디스크가 완전히 다르게 c3t2d0으로 보일 수 있습니다. DID 드라이버는 전역 이름을 할당하여, d10처럼 각 노드에 멀티 호스트 디스크에 대해 일관된 매핑을 제공합니다.

사용자는 scdidadm(1M) 및 scgdevs(1M) 명령을 통해 장치 ID를 업데이트하고 관리합니다. 자세한 내용은 다음 설명서 페이지를 참조하십시오.

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

디스크 장치 그룹

SunPlex 시스템에서는 모든 멀티 호스트 장치가 Sun Cluster 소프트웨어에 의해 제어되어야 합니다. 먼저 멀티 호스트 디스크에 볼륨 관리자 디스크 그룹인 Solaris 볼륨 관리자 디스크 세트 또는 VERITAS Volume Manager 디스크 그룹(SPARC 기반 클러스터에서만 사용 가능) 중 하나를 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 하면 어느 노드가 어느 볼륨 관리자 디스크 그룹에 대한 경로를 갖는지에 대한 정보를 SunPlex 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. 가용성이 높은 디스크 장치 그룹을 사용하면 클러스터 파일 시스템을 하우징할 수 있습니다.

주 – 디스크 장치 그룹은 자원 그룹의 영향을 받지 않습니다. 하나의 노드가 데이터 서비스에 의해 액세스되는 디스크 그룹을 마스터할 때, 다른 노드가 자원 그룹(데이터 서비스 프로세스 그룹을 나타내는)을 마스터할 수 있습니다. 그러나 가장 실용적인 것은 동일한 노드에서 응용 프로그램의 자원(응용 프로그램 디먼)을 포함하는 자원 그룹과 특수 응용 프로그램의 데이터를 저장하는 디스크 장치 그룹을 보존하는 것입니다. 디스크 장치 그룹과 자원 그룹 사이의 관계에 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.

디스크 장치 그룹을 사용하면 볼륨 관리자 디스크 그룹이 기반 디스크에 대한 복수 경로를 제공하기 때문에 볼륨 관리자 디스크 그룹이 "전역"이 됩니다. 물리적으로 멀티 호스트 디스크에 연결된 각 클러스터 노드는 디스크 장치 그룹에 대한 경로를 제공합니다.

디스크 장치 그룹 페일오버

디스크 인클로저는 여러 개의 노드에 연결되어 있으므로 그 인클로저에 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).

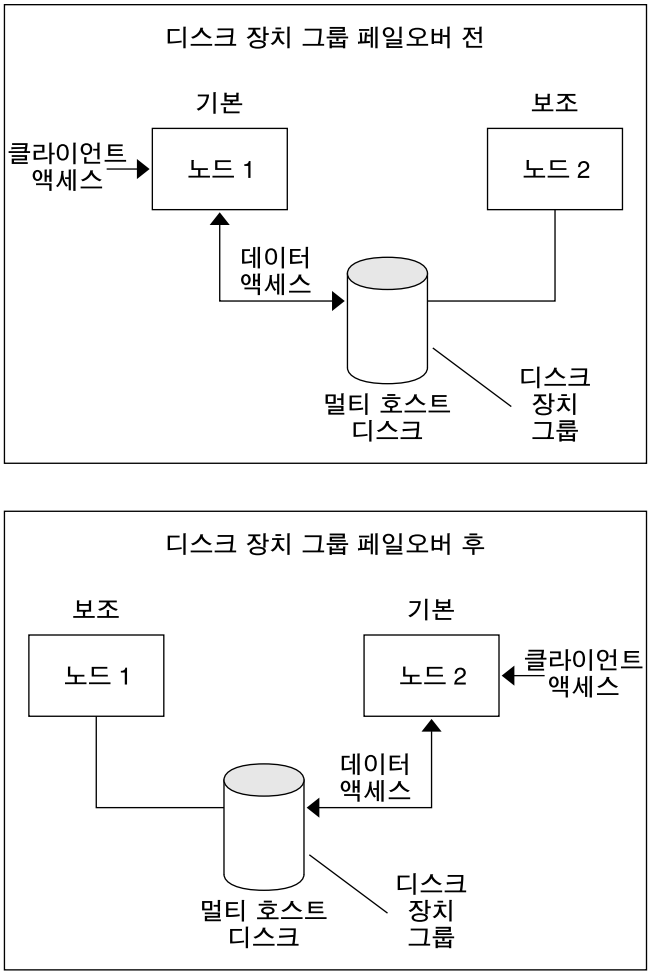


그림 3–1 디스크 장치 그룹 페일오버

멀티 포트 디스크 장치 그룹

이 절에서는 멀티 포트 디스크 구성에서 성능 및 가용성을 균형 조정할 수 있는 디스크 장치 그룹 등록 정보를 설명합니다. Sun Cluster 소프트웨어는 멀티 포트 디스크 구성을 설정하는 데 사용되는 두 가지 등록 정보preferenced 및 numsecondaries를 제공합니다. preferenced 등록 정보를 사용하면 페일오버가 발생할 경우에 노드에서 시도하는 제어 순서를 지정할 수 있습니다. numsecondaries 등록 정보를 사용하여 장치 그룹에 대해 원하는 보조 노드 수를 설정합니다.

기본 노드가 종료되었을 때 보조 노드를 기본 노드로 승격할 수 없는 경우 고가용성 서비스가 종료된 것으로 간주됩니다. 서비스 페일오버가 발생하고 preferenced 등록 정보가 true일 경우 노드는 노드 목록의 순서에 따라 보조 노드를 선택합니다. 설정된 노드 목록은 노드가 기본적으로 제어하려고 시도하거나 예비 노드에서 보조 노드로 전환하려고 시도하는 순서를 정의합니다. scsetup(1M) 유틸리티를 사용하여 장치 서비스의 기본 설정을 동적으로 변경할 수 있습니다. 종속 서비스 공급자(예: 전역 파일 시스템)와 연관된 기본 설정은 장치 서비스의 기본 설정이 됩니다.

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. 검사점 지정에 의해 발생되는 성능 저하 및 메모리 오버헤드를 최소화하기 위해 예비 노드 지원을 구현했습니다. 기본적으로 디스크 장치 그룹에는 기본 노드와 보조 노드가 하나씩 있습니다. 사용 가능한 나머지 공급자 노드는 예비 상태로 온라인됩니다. 페일오버가 발생하면 보조 노드가 기본 노드로 되고 노드 목록에서 우선 순위가 가장 높은 노드가 보조 노드가 됩니다.

보조 노드 수는 1부터 장치 그룹에서 기본을 제외한 공급자 노드 개수 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 숫자로 지정할 수 있습니다. 단, 작동 중인 기본이 아닌 공급자의 수가 해당 숫자보다 작지 않아야 합니다. 구성에서 노드를 추가하거나 제거할 경우 numsecondaries 등록 정보를 변경하고 노드 목록을 이중 검사할 수 있습니다. 노드 목록과 원하는 보조 노드 수를 유지하면 구성된 보조 노드의 수와 프레임워크에 허용된 실제 수가 충돌하지 않게 할 수 있습니다. Solaris 볼륨 관리자 장치 그룹에 대해서는 metaset(1M) 명령을 사용합니다. 또한 Veritas Volume Manager를 사용하는 경우 VxVM 장치 그룹용 scconf(1M) 명령을 preferenced 및 numsecondaries 등록 정보 설정과 함께 사용하여 구성에서의 노드 추가 및 제거를 관리합니다. 디스크 장치 그룹 등록 정보 변경의 절차 정보는 *Solaris OS*용 *Sun Cluster* **시스템 관리 안내서**의 "클러스터 파일 시스템 관리 개요"를 참조하십시오.

전역 이름 공간

전역 장치가 될 수 있도록 하는 Sun Cluster 소프트웨어 기법이 **전역 이름 공간**입니다. 전역 이름 공간에는 볼륨 관리자 이름 공간뿐 아니라 /dev/global/ 계층도 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. 물리적으로 멀티 호스트 디스크에 연결된 각 노드는 클러스터의 노드에 대한 저장소 경로를 제공합니다.

일반적으로 Solaris 볼륨 관리자의 경우 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. 이 이름 공간은 각각 클러스터를 통해 가져온 VxVM 디스크 그룹을 위한 디렉토리와 Solaris 볼륨 관리자 디스크 세트를 위한 디렉토리로 구성됩니다. 각각의 디렉토리는 해당 디스크 세트 또는 디스크 그룹의 볼륨이나 메타 장치 각각에 대해 장치 노드를 보관합니다.

SunPlex 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. Sun Cluster 소프트웨어는 계속해서 볼륨 관리자 장치를 표준 위치에 심볼릭 링크로 제공합니다. 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 예

다음 표는 멀티 호스트 디스크 c0t0d0s0에 대한 전역 이름 공간과 로컬 이름 공간 사이의 매핑입니다.

표 3–2 로컬 및 전역 이름 공간 매핑

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

표 3–2 로컬 및 전역 이름 공간 매핑 **(계속)**

구성 요소/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 업데이트됩니다. `scgdevs(1M)` 명령을 실행하여 전역 이름 공간을 생성할 수도 있습니다.

클러스터 파일 시스템

클러스터 파일 시스템에는 다음과 같은 기능이 있습니다.

- 파일 액세스 위치가 투명합니다. 프로세스는 위치에 관계 없이 시스템에 있는 파일을 열 수 있으므로 모든 노드의 프로세스들은 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 보증은 원래 디스크 액세스와 모든 파일 시스템 조작에 대해 유지됩니다.
- 클러스터 파일 시스템은 하부 파일 시스템 및 볼륨 관리 소프트웨어와 독립적으로 작동합니다. 클러스터 파일 시스템은 지원되는 디스크의 파일 시스템을 모두 전역으로 만듭니다.

`mount -g` 명령을 사용하여 전역으로 또는 `mount` 명령을 사용하여 로컬로 전역 장치에 파일 시스템을 마운트할 수 있습니다.

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

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

클러스터 파일 시스템은 별도로 구분되는 파일 시스템 형식이 아닙니다. 클라이언트는 기초가 되는 파일 시스템(예: UFS)을 보게 됩니다.

클러스터 파일 시스템 사용

SunPlex 시스템에서는 모든 멀티 호스트 디스크가 디스크 장치 그룹에 포함됩니다. Solaris 볼륨 관리자 디스크 세트, VxVM 디스크 그룹 또는 소프트웨어 기반의 볼륨 관리자에 의해 제어되지 않는 개별 디스크가 디스크 장치 그룹이 될 수 있습니다.

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

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

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

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

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

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

주 – Sun Cluster 소프트웨어에서 반드시 클러스터 파일 시스템에 이름 지정 정책을 사용해야 하는 것은 아니지만, /global/*disk-device-group*과 같이 동일한 디렉토리 아래에 모든 클러스터 파일 시스템에 대한 마운트 지점을 만들면 쉽게 관리할 수 있습니다. 자세한 내용은 *Sun Cluster 3.1 9/04 Software Collection for Solaris OS(SPARC Platform Edition)*와 *Solaris OS용 Sun Cluster* **시스템 관리 안내서**를 참조하십시오.

HAStoragePlus 자원 유형

HAStoragePlus 자원 유형은 UFS 및 VxFS와 같이 전역 파일 시스템이 아닌 구성의 가용성을 높이기 위해 설계되었습니다. 로컬 파일 시스템을 Sun Cluster 환경에 통합하고 파일 시스템의 가용성을 높이려면 HAStoragePlus 자원 유형을 사용하십시오.

HAStoragePlus는 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.

HAStoragePlus 자원 유형 사용 방법에 대한 내용은 *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*의 "Enabling Highly Available Local File Systems"를 참조하십시오.

또한 HAStoragePlus를 사용하면 자원과 자원이 사용하는 디스크 장치 그룹의 시작을 동기화할 수 있습니다. 자세한 내용은 65 페이지 "자원, 자원 그룹 및 자원 유형"을 참조하십시오.

Syncdir 마운트 옵션

syncdir 마운트 옵션은 UFS를 기반 파일 시스템으로 사용하는 클러스터 파일 시스템에 사용할 수 있습니다. 그러나 syncdir을 지정하지 않으면 성능이 크게 향상됩니다. syncdir을 지정한 경우, 쓰기의 POSIX 호환은 보증됩니다. 지정하지 않으면 NFS 파일 시스템과 동일하게 작동됩니다. 예를 들어, syncdir이 없으면 파일을 닫을 때까지 공간 부족 상태를 발견하지 못하는 경우가 있습니다. syncdir(및 POSIX 동작)이 있으면 쓰기 작업 동안 공간 부족 상태가 발견되었을 것입니다. syncdir을 지정하지 않아 문제가 생기는 경우는 드물기 때문에 syncdir을 지정하지 말고 성능 향상 이점을 얻는 것이 좋습니다.

SPARC 기반 클러스터를 사용하는 경우 Veritas VxFS에는 UFS의 syncdir 마운트 옵션에 해당하는 마운트 옵션이 없습니다. VxFS의 작동은 syncdir 마운트 옵션이 지정되지 않은 UFS의 경우와 동일합니다.

전역 장치와 클러스터 파일 시스템에 대한 FAQ는 82 페이지 "파일 시스템 FAQ"를 참조하십시오.

디스크 경로 모니터링

Sun Cluster 소프트웨어의 현재 릴리스는 DPM(Disk-Path Monitoring)을 지원합니다. 이 절에서는 DPM, DPM 데몬, 디스크 경로 모니터링에 사용하는 관리 도구 등에 대한 개념 정보를 제공합니다. 디스크 경로 상태 모니터링, 모니터링 해제 및 검사에 대한 개념은 *Solaris OS*용 *Sun Cluster* **시스템 관리 안내서**를 참조하십시오.

주 – Sun Cluster 3.1 4/04 소프트웨어 이전에 릴리스된 버전을 실행하는 노드에서는 DPM이 지원되지 않습니다. 순환 업그레이드가 진행되는 동안에는 DPM 명령을 사용하지 마십시오. 모든 노드를 업그레이드한 후 DPM 명령을 사용하려면 노드가 온라인 상태여야 합니다.

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 단일 노드 또는 클러스터의 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 위치를 설명합니다.

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

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 목록에서는 `scdpmd`가 초기 시작 단계에서 어떻게 작동하는지를 설명합니다.

주 – 시작 시에 각 디스크 경로의 상태는 알 수 없음으로 초기화됩니다.

1. DPM 데몬은 이전 상태 파일 또는 CCR 데이터베이스에서 디스크 경로 및 노드 이름 정보를 수집합니다. CCR에 대한 자세한 내용은 34 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 모니터되는 목록의 각 디스크 경로를 핑합니다. 통신 인터페이스가 수정 중인 항목의 내용에 액세스하지 못하도록 각 항목을 잠급니다.

4. DPM 데몬은 Sun Cluster Event Framework에 알림 메시지를 보내고 UNIX `syslogd`(1M) 기법을 통해 경로의 새 상태를 기록합니다.

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

DPM 데몬은 MPxIO, HDLM, PowerPath 등과 같은 다중 경로 드라이버를 통해 볼 수 있는 논리 경로의 가용성을 모니터합니다. 다중 경로 드라이버는 DPM 데몬에서 개별적으로 오류를 발생하기 때문에 이 드라이버에 의해 관리되는 개별 물리 경로는 모니터되지 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 모니터할 때에도 유용합니다.

클러스터의 디스크 경로를 모니터하는 두 번째 방법은 SunPlex Manager 그래픽 사용자 인터페이스(GUI)에 의해 제공됩니다. SunPlex Manager는 클러스터에서 모니터되는 디스크 경로에 대한 토폴로지 뷰를 제공합니다. 이 뷰는 10분마다 업데이트되어 실패한 핑의 개수 정보를 제공합니다. SunPlex Manager GUI에 의해 제공되는 정보를 `scdpm`(1M) 명령과 함께 사용하여 디스크 경로를 관리합니다. SunPlex Manager에 대한 자세한 내용은 *Solaris OS*용 *Sun Cluster* **시스템 관리 안내서**의 "그래픽 사용자 인터페이스를 통한 Sun Cluser 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 `all`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 `c1t0d0`이고, 다른 노드에서는 `c2t0d0`이 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	`schost-1:/dev/did/dsk/d1`	`schost-1` 노드의 디스크 경로 `d1`
	`all:d1`	클러스터에 있는 모든 노드의 `d1` 디스크 경로
UNIX 디스크 경로	`schost-1:/dev/rdsk/c0t0d0s0`	`schost-1` 노드의 디스크 경로 `c0t0d0s0`
	`schost-1:all`	`schost-1` 노드의 모든 디스크 경로
모든 디스크 경로	`all:all`	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.

- 47 페이지 “쿼럼 투표 수 정보”
- 47 페이지 “장애 차단 정보”

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 안됩니다. 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.

CMM에 대한 자세한 내용은 *Solaris OS***용** *Sun Cluster* **개요**의 “클러스터 멤버쉽”을 참조하십시오.

클러스터 분할 영역에서는 다음 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 영역에서는 다른 분할 영역이 있다는 것을 알지 못합니다.

정보 유실은 클러스터가 종료된 후, 종료하기 전의 클러스터 구성 데이터를 사용하여 다시 시작할 경우에 발생합니다. 이 문제는 마지막으로 기능을 수행하는 클러스터 분할 영역에 있지 않았던 노드에서 클러스터를 시작할 때 발생할 수 있습니다.

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

- 각 노드에 한 표씩 할당
- 작동 클러스터에 대부분의 표를 위임

대부분의 표를 가진 분할 영역은 **쿼럼**을 얻어 작동할 수 있습니다. 이러한 다수 표 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 것 만으로 충분하지 않습니다. 노드가 두 개인 클러스터에서는 다수가 둘입니다. 그러한 2 노드 클러스터가 분할되는 경우 두 분할 영역 중 한 쪽에서 쿼럼을 얻으려면 외부 표가 필요합니다. 필요한 외부 표는 **쿼럼 장치**에서 제공합니다.

쿼럼 투표 수 정보

다음 정보를 확인하려면 `scstat -q` 명령을 사용합니다.

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

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

두 노드와 쿼럼 장치는 쿼럼을 형성하기 위해 클러스터에 표를 제공합니다.

노드는 노드 상태에 따라 표를 제공합니다.

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

쿼럼 장치는 장치에 연결된 투표 수에 따라 표를 제공합니다. 쿼럼 장치를 구성하면 Sun Cluster 소프트웨어가 쿼럼 장치에 *N*-1개의 투표 수를 할당합니다. 여기서 *N*은 쿼럼 장치에 연결된 투표 수입니다. 예를 들어, 투표수가 0이 아닌 두 노드에 연결된 쿼럼 장치는 쿼럼 수가 1입니다(2 - 1).

쿼럼 장치가 표를 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 OS*용 *Sun Cluster* **시스템 관리 안내서**의 "쿼럼 관리"에 설명된 절차를 사용하여 클러스터를 설치하는 동안 또는 그 이후에 구성합니다.

장애 차단 정보

클러스터의 큰 문제는 클러스터가 분할되는(**정보 분리**) 문제입니다. 이 문제가 발생하면 일부 노드의 통신이 불가능하게 되어 개별 노드나 일부 노드가 개별 클러스터나 하위 클러스터를 형성하려고 시도할 가능성이 있습니다. 각 부분 또는 분할 영역에서는 멀티 호스트 장치에 대해 유일한 액세스 권한과 소유권을 가진 것으로 인식할 수 있습니다. 여러 노드가 디스크에 기록하려고 시도하면 데이터가 훼손될 수 있습니다.

장애 차단 기능은 디스크에 대한 액세스를 물리적으로 막음으로써 멀티 호스트 장치에 대한 노드 액세스를 제한합니다. 노드가 클러스터에서 나갈 경우(실패하거나 분할되어), 장애 차단 기능은 그 노드가 더이상 디스크에 액세스할 수 없도록 합니다. 현재 구성원 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(소유자) 노드로서 서비스를 제공하는 클러스터 구성원에 장애가 발생하거나 도달할 수 없게 되면, 새로운 기본 노드가 선택되어 부수적인 인터럽트만으로 계속해서 디스크 장치 그룹에 액세스할 수 있도록 합니다. 이 프로세

스 동안 새로운 기본 노드가 시작되려면 이전의 기본 노드가 장치에 대한 액세스를 포기해야만 합니다. 그러나 구성원이 클러스터에서 이탈하여 도달할 수 없게 되면 클러스터는 기본 노드였던 장치를 해제하도록 해당 노드에 알릴 수 없습니다. 그러므로 남아있는 구성원이 장애가 발생한 구성원으로부터 전역 장치를 제어하고 액세스할 수 있도록 하는 방법이 필요합니다.

SunPlex 시스템은 SCSI 디스크 예약 기능을 사용하여 장애 차단 기능을 실행합니다. SCSI 예약 기능을 사용하면 장애가 발생한 노드가 멀티 호스트 장치로부터 "차단되어" 디스크에 액세스할 수 없습니다.

SCSI-2 디스크 예약은 디스크에 접속된 모든 노드에 대한 액세스를 부여하거나(어떤 예약도 없을 경우) 단일 노드(예약이 있는 노드)에 대한 액세스로 제한하는 예약 양식을 지원합니다.

클러스터 상호 연결을 통해 다른 노드가 더 이상 통신할 수 없다는 것을 발견한 클러스터 구성원은 장애 차단 프로시저를 시작하여 다른 노드가 공유 디스크에 액세스하지 못하도록 합니다. 이러한 장애 차단 프로시저가 실행되면 액세스가 차단된 노드가 중단되고 콘솔에 "예약 충돌" 메시지가 표시됩니다.

예약 충돌은 특정 노드가 더 이상 클러스터 구성원이 아님을 발견한 후 이 노드와 다른 노드 사이에 공유되어 있는 모든 디스크에 대해 SCSI 예약이 적용되는 경우 발생합니다. 액세스가 차단된 노드는 차단되고 있음을 인식하지 못할 수 있으므로 공유 디스크 중 하나에 액세스를 시도하면 예약을 발견하게 되고 패닉 상태가 됩니다.

장애 차단을 위한 페일패스트 기법

장애가 발생한 노드가 재부트되어 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, MHIOCENFAILFAST를 계속 사용할 수 있도록 합니다. 이 ioctl은 디스크 드라이버에 대한 지시어이고, 디스크가 다른 노드에 예약되어 디스크에 액세스할 수 없을 경우에 노드가 종료될 수 있도록 합니다.

MHIOCENFAILFAST ioctl을 사용하면 노드가 디스크에 대해 실행하는 모든 읽기 및 쓰기에서 반환되는 오류에 대해 드라이버가 `Reservation_Conflict` 오류 코드를 검사합니다. ioctl은 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 사용하는 경우나 SCSI-3 디스크를 사용하는 경우에 모두 동일하게 작동합니다.

노드가 클러스터의 다른 노드와 연결이 끊어지고 쿼럼을 채울 수 있는 분할 영역에 포함되지 않은 경우에는 다른 노드에 의해 강제로 클러스터에서 제거됩니다. 쿼럼을 채울 수 있는 분할 영역에 포함된 다른 노드가 공유 디스크에 예약을 설정하고, 쿼럼이 채워지지 않은 노드가 예약된 공유 디스크에 액세스하려고 시도하면 페일패스트 기법에 의해 예약 충돌이 발생하여 종료됩니다.

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` 프롬프트에서는 `eeprom`(1M)을 사용하여 `auto-boot?`를 설정할 수 있습니다. x86 기반 클러스터에서는 BIOS 부트 이후 선택적으로 실행하는 SCSI 유틸리티를 사용하여 설정할 수 있습니다.

쿼럼 구성 정보

다음 목록에는 쿼럼 구성에 대한 정보가 포함되어 있습니다.

- 쿼럼 장치는 사용자 데이터를 포함할 수 있습니다.
- *N*개의 쿼럼 장치가 `1`에서 *N* 노드와 `N+1` 노드 중 하나에 각각 연결된 N+1 구성에서는 `1`부터 *N* 노드 전부 또는 *N*/ **2** 노드 중 어느 하나에 장애가 발생해도 클러스터가 견뎌냅니다. 이러한 가용성은 쿼럼 장치가 제대로 기능을 수행하고 있는 경우를 전제로 한 것입니다.
- 하나의 쿼럼 장치가 모든 노드에 연결된 *N* 노드 구성에서는 *N*- **1** 노드 중 하나에 장애가 발생하면 클러스터가 견딜 수 있습니다. 이러한 가용성은 쿼럼 장치가 제대로 기능을 수행하고 있는 경우를 전제로 한 것입니다.
- 하나의 쿼럼 장치가 모든 노드에 연결된 *N* 노드 구성에서는 모든 클러스터 노드를 사용할 수 있는 경우 쿼럼 장치에 장애가 발생해도 클러스터가 견딜 수 있습니다.

피해야 할 쿼럼 구성의 예는 55 페이지 “바람직하지 않은 쿼럼 구성”을 참조하십시오. 권장되는 쿼럼 구성의 예는 52 페이지 “권장되는 쿼럼 구성”을 참조하십시오.

쿼럼 장치 요구 사항 준수

다음 요구 사항을 준수해야 합니다. 그렇지 않으면 클러스터의 가용성이 손상될 수 있습니다.

- Sun Cluster 소프트웨어가 특정 장치를 쿼럼 장치로 지원하는지 확인합니다.

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

Sun Cluster 소프트웨어는 다음 두 가지 유형의 쿼럼 장치를 지원합니다.

- SCSI-3 PGR 예약을 지원하는 멀티 호스트 공유 디스크
- SCSI-2 예약을 지원하는 이중 호스트 공유 디스크

- 2 노드 구성에서는 최소한 한 쿼럼 장치는 한 노드에 장애가 발생해도 다른 노드가 계속할 수 있도록 구성해야 합니다. 그림 3–2를 참조하십시오.

피해야 할 쿼럼 구성 예는 55 페이지 "바람직하지 않은 쿼럼 구성"을 참조하십시오. 권장되는 쿼럼 구성 예는 52 페이지 "권장되는 쿼럼 구성"을 참조하십시오.

가장 적합한 쿼럼 장치 구성 준수

다음 정보를 사용하여 토폴로지에 가장 적합한 쿼럼 구성을 평가합니다.

- 클러스터의 모든 노드에 연결될 수 있는 장치가 있습니까?
  - 있다면 그 장치를 쿼럼 장치로 구성합니다. 현재의 구성이 최적의 구성이기 때문에 다른 쿼럼 장치를 구성할 필요가 **없습니다**.

주의 – 이 요구 사항을 무시하고 다른 쿼럼 장치를 추가하면 클러스터의 가용성이 감소됩니다.

  - 없다면 이중 포트 장치를 구성합니다.
- 쿼럼 장치에서 제공한 총 표 수는 반드시 노드에서 제공한 총 표 수보다 적어야 합니다. 그렇지 않을 경우 모든 노드가 기능을 수행하더라도 모든 디스크를 사용할 수 없는 경우에는 노드가 클러스터를 형성할 수 없습니다.

주 – 경우에 따라 특별한 환경에서는 필요에 맞도록 전체적인 클러스터 가용성을 낮추는 것이 바람직할 수 있습니다. 그런 경우에는 이러한 최적의 구성을 무시할 수 있습니다. 그러나 최적의 구성을 따르지 않으면 전체 가용성은 감소됩니다. 예를 들어, 54 페이지 "비전형적인 쿼럼 구성"에 설명된 구성에서는 클러스터의 가용성이 더 적습니다. 이 구성에서는 쿼럼 투표 수가 노드 투표 수보다 많습니다. 클러스터는 노드 A와 노드 B 간의 공유 저장소에 대한 액세스가 없어질 경우 클러스터 전체가 실패하는 특성을 가지고 있습니다.

최적의 구성이 아닌 예외는 54 페이지 "비전형적인 쿼럼 구성"을 참조하십시오.

- 저장 장치에 대한 액세스를 공유하는 모든 노드 쌍 사이에 쿼럼 장치를 지정합니다. 이렇게 쿼럼을 구성하면 장애 차단 프로세스의 속도가 빨라집니다. 53 페이지 "3 노드 이상 구성의 쿼럼"을 참조하십시오.
- 일반적으로 쿼럼 장치를 추가하여 총 클러스터 투표 수를 짝수로 만들면 총 클러스터 가용성이 감소됩니다.
- 쿼럼 장치에서는 노드 추가 또는 노드 제거 후 재구성 속도가 조금 느려집니다. 그러므로 필요 이상의 쿼럼 장치를 추가해서는 안됩니다.

피해야 할 쿼럼 구성 예는 55 페이지 “바람직하지 않은 쿼럼 구성”을 참조하십시오. 권장되는 쿼럼 구성예는 52 페이지 “권장되는 쿼럼 구성”을 참조하십시오.

권장되는 쿼럼 구성

피해야 할 쿼럼 구성 예는 55 페이지 “바람직하지 않은 쿼럼 구성”을 참조하십시오.

2 노드 구성의 쿼럼

2 노드 클러스터를 형성하려면 두 개의 쿼럼 표가 필요합니다. 이 두 개의 표는 두 개의 클러스터 노드에서 제공될 수도 있고 하나의 노드와 쿼럼 장치로부터 제공될 수도 있습니다.

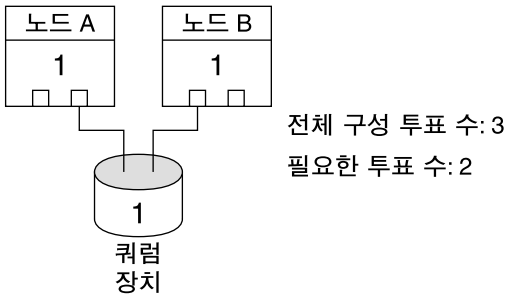
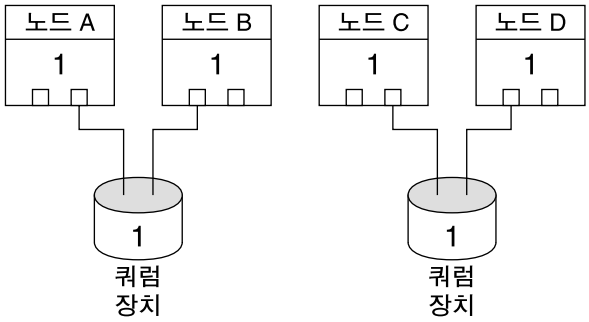


그림 3–2 2 노드 구성

3 노드 이상 구성의 쿼럼

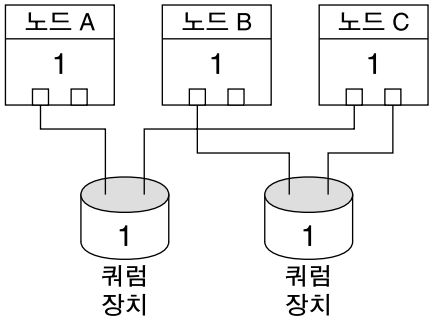
쿼럼 장치 없이 3 노드 이상의 클러스터를 구성할 수 있습니다. 그러나 그렇게 구성할 경우에는 클러스터에 대다수 노드가 없이는 클러스터를 시작할 수 없습니다.



전체 구성 투표 수: 6

필요한 쿼럼 투표 수: 4

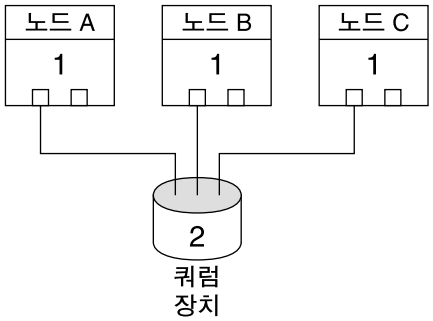
이러한 구성에서는 어느 쪽 쌍이든 작동하려면 각각의 쌍이 사용 가능해야 합니다.



전체 구성 투표 수: 5

필요한 쿼럼 투표 수: 3

이러한 구성에서는 보통 노드 A와 B에서 응용 프로그램을 실행하도록 구성하고 노드 C는 핫 스패어로 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 쿼럼 구성

그림 3–3은 노드 A와 노드 B에서 핵심 응용 프로그램(예: Oracle 데이터베이스)을 실행하고 있는 경우를 나타냅니다. 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.

이 예외와 관련된 최적의 구성에 대한 내용은 50 페이지 "가장 적합한 쿼럼 장치 구성 준수"를 참조하십시오.

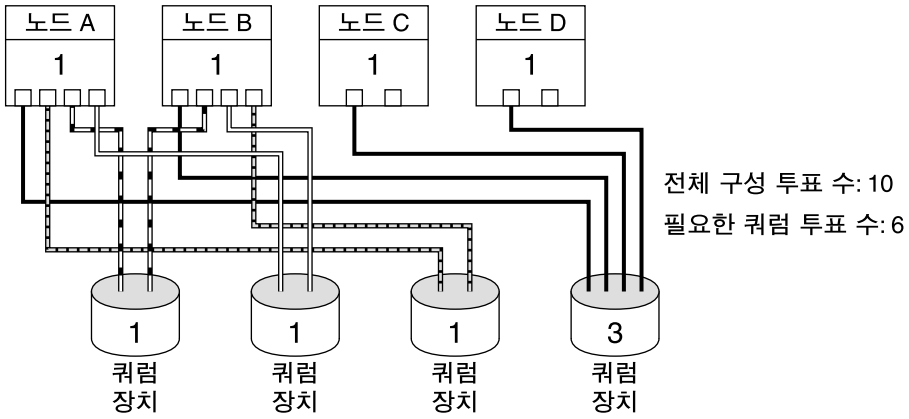
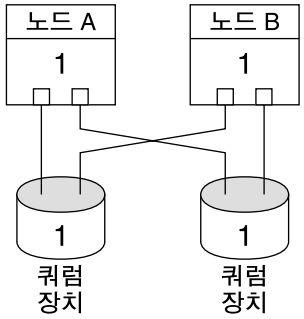


그림 3–3 비전형적 구성

바람직하지 않은 쿼럼 구성

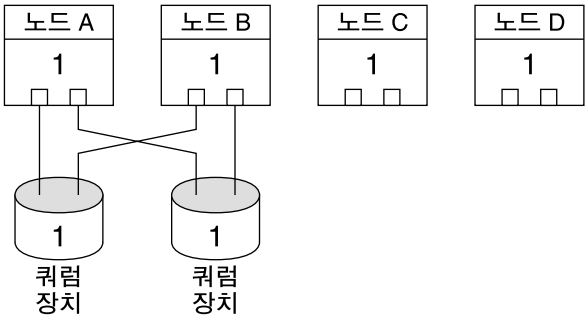
권장되는 쿼럼 구성 예는 52 페이지 "권장되는 쿼럼 구성"을 참조하십시오.



전체 구성 투표 수: 4

필요한 쿼럼 투표 수: 3

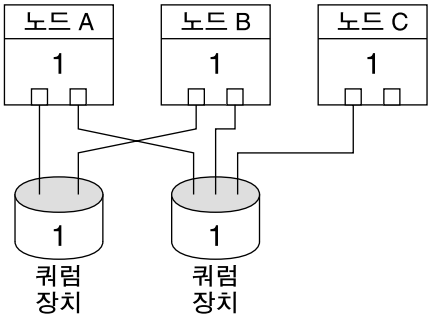
이러한 구성은 쿼럼 장치에서 제공한 총 투표 수가 반드시 노드에서 제공한 총 투표 수보다 적어야 한다는 쿼럼 장치 최적화 요건에 위배됩니다.



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: 5

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.

데이터 서비스

데이터 서비스라는 용어는 Sun Java System Web Server(이전 명칭은 Sun Java System Web Server) 또는 SPARC 기반 클러스터의 경우 Oracle처럼 단일 서버가 아닌 클러스터에서 실행되도록 구성된 타사 응용 프로그램을 의미합니다. 데이터 서비스는 응용 프로그램, Sun Cluster 구성 파일, 다음과 같은 응용 프로그램 작업을 제어하는 Sun Cluster 관리 메소드 등으로 구성됩니다.

- 시작
- 중지
- 모니터 및 수정 조치
- 데이터 서비스 유형에 대한 자세한 내용은 *Solaris OS*용 *Sun Cluster* **개요**의 "데이터 서비스"를 참조하십시오.

그림 3–4에서는 단일 Application Server에서 실행되는 응용 프로그램(단일 서버 모델)을 클러스터에서 실행 중인 동일한 응용 프로그램(클러스터 서버 모델)과 비교합니다. 사용자의 관점에서 보면 클러스터 응용 프로그램이 더 빠르게 실행될 수 있고 가용성이 높다는 것 외에는 두 구성 사이에 큰 차이가 없습니다.

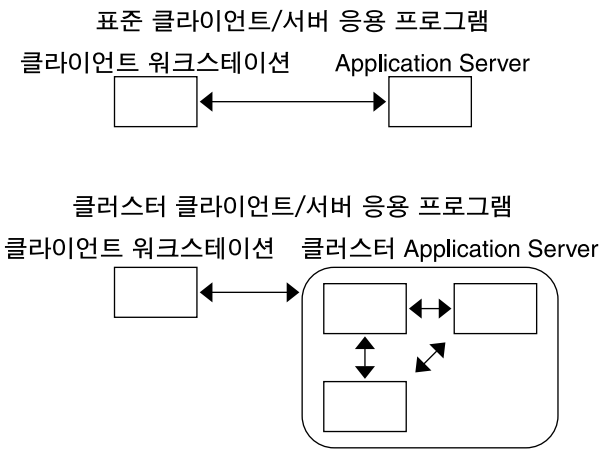


그림 3–4 표준 및 클러스터 클라이언트/서버 구성 비교

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

네트워크 자원은 특정 물리적 서버와 연결되지 않고 물리적 서버 사이에 마이그레이션할 수 있습니다.

네트워크 자원은 처음에 **기본** 노드에 연결됩니다. 기본 노드에 장애가 발생하면 네트워크 자원과 응용 프로그램 자원이 다른 클러스터 노드(보조)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.

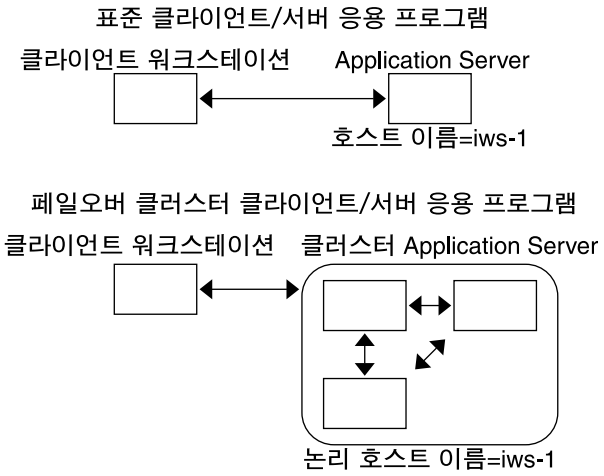


그림 3–5 고정 호스트 이름과 논리 호스트 이름 비교

또한 공유 주소는 처음에 한 노드와 연결됩니다. 이 노드를 GIN(Global Interface Node)이라고 합니다. 공유 주소는 클러스터에 대한 단일 네트워크 인터페이스로 사용됩니다. 이러한 인터페이스를 **전역 인터페이스**라고 합니다.

논리 호스트 이름 모델과 확장 가능한 서비스 모델 간의 차이점은 확장 가능한 서비스 모델의 경우 각 노드에도 루프백 인터페이스에 대해 활성으로 구성된 공유 주소가 있다는 것입니다. 이러한 구성 때문에 여러 노드에 대하여 여러 데이터 서비스 인스턴스를 동시에 실행할 수 있습니다. "확장 가능 서비스"는 클러스터 노드를 추가하여 응용 프로그램에 CPU 기능을 추가하는 방법으로 성능을 확장할 수 있는 기능입니다.

전역 인터페이스 노드에 장애가 발생하면 응용 프로그램 인스턴스를 실행하는 다른 노드로 공유 주소를 변경하고 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.

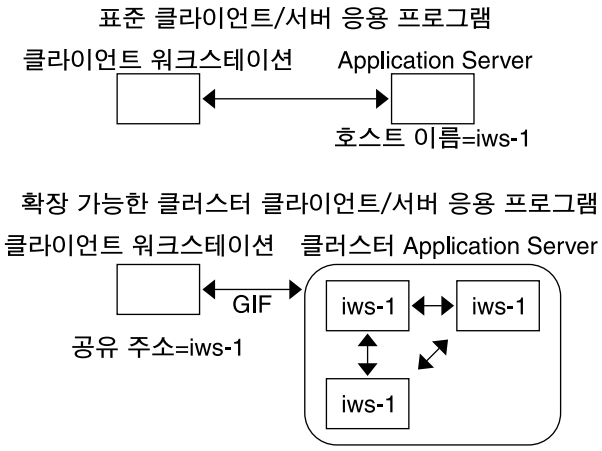


그림 3–6 고정 호스트 이름과 공유 주소 비교

데이터 서비스 메소드

Sun Cluster 소프트웨어는 일련의 서비스 관리 메소드를 제공합니다. 이 메소드는 RGM(Resource Group Manager)의 제어에 따라 실행되어 클러스터 노드의 응용 프로그램을 시작하고 중지하고 모니터합니다. 이 메소드는 클러스터 프레임워크 소프트웨어, 멀티호스트 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 SunPlex 시스템에서도 API와 여러 가지 데이터 서비스 개발 도구를 제공합니다. 이 도구를 사용하면 Sun Cluster 소프트웨어에서 다른 응용 프로그램을 가용성이 높은 데이터 서비스로 실행할 수 있도록 응용 프로그램 프로그래머가 필요한 데이터 서비스 메소드를 개발할 수 있습니다.

페일오버 데이터 서비스

데이터 서비스가 실행되는 노드가 실패할 경우, 서비스는 사용자 간섭 없이 다른 작업 노드로 마이그레이션됩니다. 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 원래 노드에서 자동으로 구성이 중지되고 다른 노드에서 구성이 시작됩니다.

페일오버 데이터 서비스의 경우, 응용 프로그램 인스턴스는 단일 노드에서만 실행됩니다. 오류 모니터가 오류를 발견하면, 데이터 서비스가 구성된 방법에 따라 동일한 노드에서 인스턴스를 재시작하려고 하거나 다른 노드에서 인스턴스를 시작하려고 합니다(페일오버).

확장 가능 데이터 서비스

확장 가능 데이터 서비스는 여러 노드에서 인스턴스가 사용되고 있을 가능성을 수반합니다. 확장 가능한 서비스는 두 가지 자원 그룹을 사용합니다. 응용 프로그램 자원을 포함하는 **확장 가능 자원 그룹** 그리고 확장 가능 서비스가 종속된 네트워크 자원(**공유 주소**)을 포함하는 페일오버 자원 그룹의 두 가지 자원 그룹을 사용합니다. 확장 가능 자원 그룹은 여러 노드에서 온라인 상태로 있을 수 있으므로 서비스의 여러 인스턴스가 한 번에 실행될 수 있습니다. 공유 주소를 호스팅하는 페일오버 자원 그룹은 한 번에 한 노드에서만 온라인 상태입니다. 확장 가능 서비스에 호스팅하는 모든 노드가 동일한 공유 주소를 사용하여 서비스를 호스팅합니다.

서비스 요청은 단일 네트워크 인터페이스(전역 인터페이스)를 통해 클러스터에 전달되고 **로드 균형 조정 정책**에 의해 설정된 사전 정의된 몇 가지 알고리즘 중 하나를 사용하여 노드에 분산됩니다. 클러스터는 로드 균형 조정 정책을 사용하여 몇몇 노드 사이의 서비스 부하 균형을 맞추는 로드 균형 조정 정책을 사용할 수 있습니다. 다른 공유 주소에 호스트하는 다른 노드에 여러 개의 GIF가 있을 수 있으므로 유의하십시오.

확장 가능 서비스의 경우 응용 프로그램 인스턴스는 여러 노드에서 동시에 실행됩니다. 전역 인터페이스를 호스트하는 노드가 실패할 경우, 전역 인터페이스는 다른 노드로 페일오버합니다. 응용 프로그램 인스턴스 실행이 실패하는 경우, 인스턴스는 동일한 노드에서 재시작을 시도합니다.

응용 프로그램 인스턴스는 동일한 노드에서 재시작될 수 없으며, 사용되지 않는 다른 노드는 서비스를 실행하도록 구성된 경우, 서비스는 사용되지 않는 노드로 페일오버됩니다. 그렇지 않으면, 나머지 노드에서 실행을 계속하여, 서비스 처리량이 줄어들 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 3–7에는 페일오버, 확장 가능 자원 그룹, 확장 가능 서비스를 위한 둘 사이의 관계 등에 대한 예가 있습니다. 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 사용되는 네트워크 자원이 포함됩니다. 확장 가능 자원 그룹에는 Apache Web Server의 응용 프로그램 인스턴스만 포함됩니다. 확장 가능 자원 그룹과 페일오버 자원 그룹 사이에는 자원 그룹 의존 관계가 있고(실선) Apache 응용 프로그램 자원은 모두 공유 주소인 네트워크 자원 `schost-2`를 사용합니다(점선).

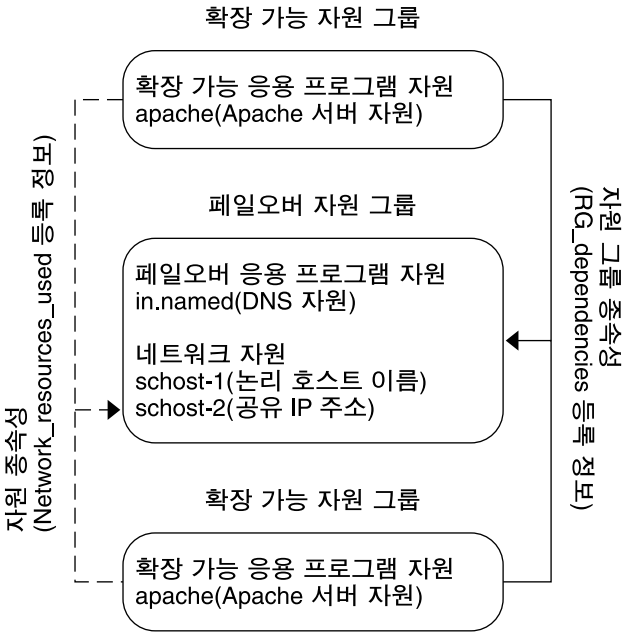


그림 **3–7** SPARC: 페일오버 및 확장 가능 자원 그룹의 예

로드 균형 조정 정책

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

확장 가능한 데이터 서비스에는 *pure* 및 *sticky*라는 두 가지 클래스가 있습니다. pure 서비스는 인스턴스가 클라이언트 요청에 응답할 수 있는 서비스입니다. sticky 서비스는 클라이언트가 같은 인스턴스에 요청을 보내는 서비스이며그러한 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 가정합시다. 각 노드는 해당 서비스 대신 클라이언트의 요청 중 1/3에 서비스를 제공합니다. 관리자는 언제든지 `scrgadm(1M)` 명령 인터페이스나 SunPlex Manager GUI를 통해 노드의 처리량을 변경할 수 있습니다.

sticky 서비스에는 **일반** *sticky*와 **와일드카드** *sticky* 등의 두 가지 종류가 있습니다. Sticky 서비스는 여러 TCP 연결을 거쳐 동시 응용 프로그램 레벨 세션이 in-state 메모리(응용 프로그램 세션 상태)를 공유할 수 있게 합니다.

보통 sticky 서비스는 클라이언트가 여러 개의 동시 TCP 연결 사이에 상태를 공유할 수 있게 합니다. 서버 인스턴스가 단일 포트에서 서비스 요청을 받을 경우에 클라이언트를 “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 브라우저가 세 개의 서로 다른 TCP 연결을 사용하여 포트 80에서 고유 IP 주소에 연결되지만, 그 연결들은 서비스에서 캐시된 세션 정보를 교환합니다.

sticky 정책을 일반화하면 여러 개의 확장 가능한 서비스가 동일한 인스턴스의 후면에서 세션 정보를 교환하는 것으로 확장할 수 있습니다. 이러한 서비스가 동일한 서비스의 백그라운드에서 세션 정보를 교환하면 동일한 노드의 서로 다른 포트에서 서비스 요청을 받는 여러 서버 인스턴스에 대하여 클라이언트가 “sticky”하다고 합니다.

예를 들어, e-commerce 사이트의 고객이 포트 80에서 일반 HTTP를 사용하여 상품들로 시장 바구니를 채우지만, 바구니의 상품 대금을 신용 카드 지불할 경우 보안 데이터를 보내기 위해 포트 443의 SSL로 전환합니다.

와일드 카드 sticky 서비스는 동적으로 할당된 포트 번호를 사용하지만, 여전히 클라이언트 요청이 같은 노드로 갈 것으로 예상합니다. 클라이언트는 IP 주소가 동일한 포트에 대해 “sticky wildcard”입니다.

이 정책의 좋은 예는 수동 모드 FTP입니다. 클라이언트는 포트 21의 FTP 서버에 연결되고, 동적 포트 범위의 청취자 포트 서버에 다시 연결하도록 서버에서 알립니다. IP 주소에 대한 모든 요청은 제어 정보를 통해 서버가 클라이언트를 알린 동일 노드로 전송됩니다.

이 sticky 정책 각각에 대해 가중된 로드 균형 조정 정책이 기본적으로 적용되므로 클라이언트의 초기 요청은 로드 균형 조정기에 의해 지시된 인스턴스에 보내집니다. 인스턴스가 실행되는 노드에 대한 유사성을 클라이언트가 확립하고 나면, 차후 요청은 액세스할 수 있는 경우 그 인스턴스로 보내집니다. 그리고 로드 균형 조정 정책은 변경되지 않습니다.

특정 로드 균형 조정 정책에 대한 추가 세부 사항이 아래에 설명되어 있습니다.

- 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. X=가중치이고 A=모든 활성 노드의 총 가중치라고 가정할 때 총 연결 수가 충분하다면, 활성 노드는 활성 노드에 보내질 총 새 연결 수의 약 X/A입니다. 이 정책은 개별 요청을 다루지 않습니다.

  이 정책은 라운드 로빈(round robin)이 아닙니다. 라운드 로빈 정책은 클라이언트의 각 요청이 항상 서로 다른 노드에 가도록 합니다. 예를 들어, 첫 번째 요청은 노드 1로, 두 번째 요청은 노드 2로 전달됩니다.

- 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 수퍼세트입니다. IP 주소에 의해 식별된 확장 가능 서비스의 경우, 포트는 서버에 의해 할당됩니다(미리 알려지지 않음). 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. 그러면 원래 보조 노드였던 노드가 새로운 기본 노드가 됩니다. 페일백 설정은 초기 기본 노드가 다시 온라인 상태가 될 때 수행되는 작업을 지정합니다. 초기 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.

특정 인스턴스에서, 예를 들어 자원 그룹을 호스트하는 원래 노드가 반복적으로 실패하고 재부트될 경우, 페일백을 설정하면 자원 그룹에 대한 가용성이 감소될 수 있습니다.

데이터 서비스 오류 모니터

각 SunPlex 데이터 서비스에는 주기적으로 데이터 서비스를 확인하여 안전 상태를 판단하는 오류 모니터가 있습니다. 오류 모니터는 응용 프로그램 데몬이 실행 중인지 그리고 클라이언트 서비스가 제공되고 있는지 확인합니다. 프로브에 의해 반환된 정보를 기초로, 디먼을 재시작하고 페일오버를 야기하는 것과 같은 사전에 정의된 조치가 초기화될 수 있습니다.

새 데이터 서비스 개발

Sun에서는 클러스터에서 여러 응용 프로그램이 페일오버 또는 확장 가능한 서비스로 작동하도록 할 수 있는 구성 파일과 관리 메소드 템플리트를 제공합니다. 고가용성 데이터 서비스로 실행할 응용 프로그램이 현재 Sun에서 제공되는 것이 아니면, API 또는 DSDL API을 사용하여 응용 프로그램을 취하고 이를 고가용성 데이터 서비스로 실행되도록 구성할 수 있습니다.

응용 프로그램이 페일오버 서비스가 될 수 있도록 할 것인지를 결정하는 기준이 있습니다. 자세한 기준은 응용 프로그램에 사용할 수 있는 API를 설명하는 SunPlex 문서에서 설명합니다.

여기서 사용하는 서비스가 확장 가능 데이터 서비스 구조의 장점을 취할 수 있는지 알 수 있도록 도와주는 일부 지침을 제시하기로 하겠습니다. 확장 가능 서비스에 대한 자세한 내용은 59 페이지 “확장 가능 데이터 서비스” 절을 참조하십시오.

다음 지침을 만족시키는 새로운 서비스는 확장 가능 서비스로 사용할 수 있습니다. 기존의 서비스가 이러한 지침을 정확하게 따르지 않으면, 서비스가 지침을 따르도록 부분을 재작성해야 할 수도 있습니다.

확장 가능 데이터 서비스는 다음 특징을 가집니다. 먼저 이러한 서비스는 하나 이상의 서버 **인스턴스**로 구성됩니다. 각 인스턴스는 클러스터의 서로 다른 노드에서 실행됩니다. 동일한 서비스에서 두 개 이상의 인스턴스가 동일한 노드에서 실행될 수는 없습니다.

두 번째, 서비스가 외부 논리 데이터 저장소를 제공할 경우, 여러 서버 인스턴스로부터 이 저장소로의 동시 액세스를 동기화하여, 업데이트 사항을 손실하거나 데이터가 변경되는 동안 데이터를 읽는 일이 발생하지 않도록 해야 합니다. 저장소에 저장된 상태를 메모리 내에 기억된 상태와 구별하기 위해 "외부"라고 표현하고, 저장소가 복제될 수는 있지만 단일 엔티티로 표시되기 때문에 "논리"라고 합니다. 게다가, 이 논리 데이터 저장소에는 서버 인스턴스가 저장소를 업데이트할 때마다 다른 인스턴스가 업데이트 사항을 즉시 볼 수 있도록 하는 등록 정보가 있습니다.

SunPlex 시스템은 클러스터 파일 시스템과 전역 원시 분할 영역을 통해 이러한 외부 저장 장치를 제공합니다. 예를 들면, 서비스가 새로운 데이터를 외부 로그 파일에 기록하거나 기존 데이터를 제 위치에서 수정한다고 가정합시다. 이 서비스의 여러 인스턴스가 실행될 경우, 각 인스턴스는 이 외부 로그에 대한 액세스를 가지므로 각각은 동시에 이 로그에 액세스할 수도 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 원하는 동기화를 수행할 수 있습니다.

이런 저장소 유형의 다른 예는 가용성이 큰 Oracle 또는 SPARC 기반 클러스터를 위한 Oracle Real Application Clusters와 같은 백엔드 데이터베이스입니다. 그러한 백엔드 데이터베이스 서버는 데이터베이스 조회와 업데이트 트랜잭션을 사용하여 내장된 동기화를 제공하므로, 여러 서버 인스턴스가 자체의 고유 동기화를 구현하지 않아도 됩니다.

현재 모습에서 확장 가능 서비스가 아닌 서비스의 예는 Sun의 IMAP 서버입니다. 이 서비스는 저장소를 업데이트하지만, 그 저장소는 개인용이므로 여러 IMAP 인스턴스가 저장소에 기록하면 업데이트 작업이 동기화되지 않았기 때문에 인스턴스들 간에 서로 덮어쓰게 됩니다. IMAP 서버는 동시 액세스를 동기화하기 위해 다시 작성해야 합니다.

마지막으로, 인스턴스에는 다른 인스턴스의 데이터에서 분리된 개인용 데이터가 있을 수 있다는 점에 유의하십시오. 그러한 경우, 데이터는 개인용이고, 해당되는 인스턴스만 이를 조작할 수 있으므로 서비스는 자체를 동시 액세스에 동기화하는데 관여하지 않아도 됩니다. 이 경우, 전역으로 액세스할 수 있게 될 수도 있으므로 클러스터 파일 시스템 아래에 개인 데이터를 저장하지 않도록 주의해야 합니다.

데이터 서비스 API 및 데이터 서비스 개발 라이브러리 API

SunPlex 시스템은 응용 프로그램의 가용성을 높이기 위하여 다음과 같은 기능을 제공합니다.

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

*Sun Cluster Data Services Planning and Administration Guide for Solaris OS*는 SunPlex 시스템과 함께 제공된 데이터 서비스를 설치하고 구성하는 방법을 설명합니다. *Sun Cluster 3.1 9/04 Software Collection for Solaris OS(SPARC Platform Edition)*은 Sun Cluster 프레임워크에서 다른 응용 프로그램을 사용하여 가용성을 높이는 방법을 설명합니다.

응용 프로그램 프로그래머가 Sun Cluster API를 사용하면 데이터 서비스 인스턴스를 시작하고 중지하는 스크립트와 오류 모니터를 개발할 수 있습니다. 이러한 도구를 사용하면, 응용 프로그램에 페일오버 또는 확장 가능 데이터 서비스 중 어느 것을 제공할 것인지 측정할 수 있습니다. 또한 SunPlex 시스템은 응용 프로그램을 페일오버 서비스나 확장 가능 서비스로 실행하기 위해 필요한 시작 및 중지 메소드를 신속하게 만들기 위해 사용할 수 있는 "일반" 데이터 서비스를 제공합니다.

데이터 서비스 트래픽에 클러스터 상호 연결 사용

클러스터에는 클러스터 상호 연결을 형성하는 노드 사이에 여러 네트워크 연결이 있어야 합니다. 클러스터 소프트웨어는 고가용성 및 성능 향상을 모두 실현하기 위해 다중 상호 연결을 사용합니다. 내부 트래픽(예를 들어, 파일 시스템 데이터 또는 확장 가능 서비스 데이터)의 경우, 메시지는 사용 가능한 모든 상호 연결을 통해 라운드 로빈 방식으로 스트립됩니다.

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.

노드간 통신을 위해 클러스터 상호 연결을 사용하려면, 응용 프로그램은 클러스터 설치 시 구성된 개인 호스트 이름을 사용해야 합니다. 예를 들어, 노드 1의 개인 호스트 이름이 `clusternode1-priv`인 경우, 클러스터 상호 연결을 통해 노드 1로 통신할 때 이 이름을 사용하십시오. 이 이름을 사용하여 열린 TCP 소켓은 클러스터 상호 연결을 통해 라우트되며 네트워크 오류가 발생하더라도 투명하게 다시 라우트될 수 있습니다.

개인 호스트 이름이 설치 중에 구성될 수 있기 때문에, 클러스터 상호 연결은 이 시점에 선택된 모든 이름을 사용할 수 있습니다. 실제 이름은 `scha_cluster_get`(3HA)에서 `scha_privatelink_hostname_node` 인자를 사용하여 얻을 수 있습니다.

응용 프로그램 수준에서 클러스터 상호 연결을 사용할 경우에는 각 노드 쌍 사이에 하나의 상호 연결이 사용되지만 서로 다른 노드 쌍에는 별도의 상호 연결이 사용될 수 있습니다. 예를 들어, 세 개의 SPARC 기반 노드에서 실행하고 클러스터 상호 연결을 통해 통신하는 응용 프로그램의 경우,인터페이스 `qfe1`에서 노드 1과 노드 3 사이의 통신이 진행되는 동안 인터페이스 `hme0`에서 노드 1과 노드 2 사이의 통신이 수행될 수 있습니다. 즉, 두 노드간 응용 프로그램 통신은 단일 상호 연결로 제한되는 반면, 내부 클러스터 통신은 모든 상호 연결을 통해 스트립됩니다.

응용 프로그램이 내부 클러스터 트래픽과 상호 연결을 공유하므로, 응용 프로그램에 사용 가능한 대역폭은 다른 클러스터 트래픽에 사용되는 대역폭에 따라 다릅니다. 장애가 발생할 경우에 내부 트래픽은 나머지 상호 연결을 통해 라운드 로빈될 수 있고, 장애가 발생한 상호 연결의 응용 프로그램 연결은 작동하는 상호 연결로 전환될 수 있습니다.

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

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

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

응용 프로그램에서 모든 경우에 대해 일관된 IP 주소를 필요로 하는 경우, 클라이언트와 서버측 모두에서 pernode 주소를 바인드하여 모든 연결이 pernode 주소를 경유하는 것으로 보일 수 있도록 응용 프로그램을 구성하십시오.

자원, 자원 그룹 및 자원 유형

데이터 서비스는 여러 가지 유형의 **자원**을 사용합니다. Sun Java System Web Server(이전 명칭은 Sun Java System Web Server) 또는 Apache Web Server와 같은 응용 프로그램은 자신이 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 자원 유형 `SUNW.oracle-server`이고 Sun Cluster HA for Apache는 자원 유형 `SUNW.apache`입니다.

주 – 자원 유형 `SUNW.oracle-server`는 SPARC 기반 클러스터에서만 사용됩니다.

자원은 전체 클러스터에서 정의된 **자원 유형**을 인스턴스화한 것입니다. 몇 가지 자원 유형이 정의되어 있습니다.

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

SUNW.HAStorage 및 HAStoragePlus 자원 유형은 자원이 사용하는 디스크 장치 그룹과 자원의 시작을 동기화하는 데 사용됩니다. 이 자원 유형은 데이터 서비스를 시작하기 전에 클러스터 파일 시스템 마운트 지점의 경로, 전역 장치 및 장치 그룹 이름을 사용할 수 있는지 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 HAStoragePlus 자원 유형을 사용할 수 있고 다른 기능이 추가되어 로컬 파일 시스템의 가용성을 높일 수 있습니다. 이 기능에 대한 자세한 내용은 41 페이지 "HAStoragePlus 자원 유형"을 참조하십시오.)

RGM에서 관리하는 자원은 하나의 단위로 관리할 수 있도록 **자원 그룹**이라는 그룹에 포함됩니다. 자원 그룹에 대하여 페일오버나 전환이 시작될 때 자원 그룹은 하나의 단위로 이동됩니다.

주 – 응용 프로그램 자원이 포함된 자원 그룹을 온라인으로 전환하면 응용 프로그램이 시작됩니다. 데이터 서비스 시작 메소드는 응용 프로그램이 시작되어 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.

RGM(Resource Group Manager)

RGM은 데이터 서비스(응용 프로그램)를 **자원 유형**으로 구현하여 자원으로 관리합니다. 이 구현은 Sun에서 제공되거나 개발자가 일반 데이터 서비스 템플리트, 데이터 서비스 개발 라이브러리 API(DSDL API) 또는 자원 관리 API(RMAPI)를 사용하여 작성합니다. 클러스터 관리자는 *resource groups*라는 컨테이너에 자원을 만들어 관리합니다. RGM은 클러스터 멤버쉽 변경에 대한 응답으로 선택된 노드에서 자원을 정지하였다가 시작합니다.

RGM은 **자원**과 **자원 그룹**을 대상으로 작업을 합니다. RGM 작업을 수행하면 자원과 자원 그룹의 상태가 온라인 및 오프라인으로 전환됩니다. 자원 및 자원 그룹에 적용 가능한 상태와 설정에 대한 자세한 설명은 66 페이지 "자원 및 자원 그룹의 상태와 설정" 절을 참조하십시오. RGM의 제어에 따라 자원 관리 프로젝트를 시작하는 방법은 65 페이지 "자원, 자원 그룹 및 자원 유형"을 참조하십시오.

자원 및 자원 그룹의 상태와 설정

관리자는 자원과 자원 그룹에 정적 설정을 적용합니다. 이 설정은 관리 작업을 통해서만 변경될 수 있습니다. RGM은 자원 그룹의 동적 "상태"를 변경합니다. 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 관리 해제하도록 할 수 있습니다. 이 설정은 클러스터를 재구성해도 변경되지 않습니다.

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

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

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

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

INIT 및 FINI 메소드의 가장 일반적인 용도는 확장 가능한 서비스를 위해 네트워크 자원에 사용되지만 응용 프로그램에 의해 수행되지 않는 초기화 또는 삭제 작업에도 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 사용하지 못하도록 하려면 자원을 비활성화하면 됩니다. 비활성화된 자원은 일반적인 용도로 사용할 수 없습니다.

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

  이러한 상태는 스위치오버 또는 페일오버 중 클러스터 재구성 단계가 진행됨에 따라 변경됩니다. 이 설정은 관리 작업을 통해서만 변경될 수 있습니다. `scswitch(1M)`는 자원이나 자원 그룹의 온라인 또는 오프라인 상태를 바꿀 때 사용할 수 있습니다.

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

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

자원 및 자원 그룹 등록 정보

SunPlex 데이터 서비스를 위해 자원과 자원 그룹에 대한 등록 정보 값을 구성할 수 있습니다. 표준 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 정보에 대한 내용은 *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*의 "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

RGM을 사용하여 온라인 상태로 가져올 때 Solaris 프로젝트 이름으로 시작하도록 데이터 서비스를 구성할 수 있습니다. 구성은 RGM에 의해 관리되는 자원 또는 자원 그룹을 Solaris 프로젝트 ID와 연결합니다. 자원 또는 자원 그룹을 프로젝트 ID로 매핑하면 Solaris 환경에서 사용할 수 있는 세부 제어 기능을 사용하여 클러스터 내에서 작업 로드 및 소비를 관리할 수 있습니다.

주 – 이 구성은 Solaris 9에서 Sun Cluster 소프트웨어의 현재 릴리스를 실행하는 경우에만 수행할 수 있습니다.

클러스터 환경에서 Solaris 관리 기능을 사용하면 노드를 다른 응용 프로그램과 공유할 때 가장 중요한 응용 프로그램에 우선 순위가 부여되게 할 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 다른 응용 프로그램이 CPU 시간과 같은 시스템 자원을 과소비하지 못하게 하여 중요 응용 프로그램의 가용성을 향상시킬 수 있습니다.

주 – 이 기능에 대한 Solaris 설명서에서는 CPU 시간, 프로세스, 작업 및 유사 구성 요소를 '자원'으로 설명합니다. 반면에 Sun Cluster 설명서에서는 '자원'이 RGM의 제어를 받는 항목을 설명하는 용어로 사용됩니다. 다음 절에서는 RGM의 제어를 받는 Sun Cluster 항목을 나타내는 용어로 '자원'을 사용하고, CPU 시간, 프로세스 및 작업을 나타내는 용어로 '공급 항목'을 사용합니다.

이 절에서는 지정된 Solaris 9 `project`(4)에서 프로세스를 시작하도록 데이터 서비스를 구성하는 방법에 대한 개념을 설명합니다. 또한, Solaris 환경에서 제공되는 관리 기능의 사용 계획을 위한 여러 페일오버 시나리오 및 권장 사항을 설명합니다. 관리 기능에 대한 자세한 개념과 절차는 *Solaris 9 System Administrator Collection*의 *System Administration Guide: Resource Management and Network Services*를 참조하십시오.

클러스터에서 Solaris 관리 기능을 사용하도록 자원 및 자원 그룹을 구성할 경우 다음과 같은 고급 프로세스를 사용하는 것이 좋습니다.

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 정보 정의는 *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*의 "Standard Properties"를 참조하십시오. 등록 정보에 대한 설명은 `r_properties`(5) 및 `rg_properties`(5)를 참조하십시오.

지정된 프로젝트 이름이 프로젝트 데이터베이스(`/etc/project`)에 존재해야 하며, 루트 사용자는 명명된 프로젝트의 구성원으로 구성되어야 합니다. 프로젝트 이름 데이터베이스에 대한 개념은 *Solaris 9 System Administrator Collection*에서 *System Administration Guide: Resource Management and Network Services*의 "Projects and Tasks"를 참조하십시오. 프로젝트 파일 구문에 대한 자세한 내용은 `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 자원 또는 자원 그룹이 오프라인 상태로 전환된 다음 RGM을 사용하여 온라인 상태로 다시 전환할 때까지는 새 프로젝트 이름이 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 소비를 기록할 수 있는 유연한 방법을 제공합니다. 확장 계정을 사용하면 사용 기록을 조사하여 이후의 작업 로드에 대한 용량 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 로드에 할당되는 CPU 시간의 공유 횟수로 표현됩니다. FSS를 기본 스케줄러로 설정하는 명령줄 설명은 `dispadmin`(1M)을 참조하십시오. 자세한 내용은 `priocntl`(1), `ps`(1) 및 `FSS`(7)을 참조하십시오.

- 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 전환 또는 페일오버를 통해 자원을 제어 및 추적할 방법을 결정해야 합니다. 새 프로젝트를 구성하기 전에 클러스터 내에서 종속성을 식별하는 것이 좋습니다. 예를 들어, 자원과 자원 그룹은 디스크 장치 그룹에 종속합니다. `scrgadm` (1M)과 함께 구성된 `nodelist`, `failback` , `maximum_primaries` 및 `desired_primaries` 자원 그룹 등록 정보를 사용하여 자원 그룹에 대한 노드 목록 우선 순위를 식별합니다. 자원 그룹과 디스크 장치 그룹 사이의 노드 목록 종속성에 대한 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.

`scrgadm`(1M) 및 `scsetup`(1M)과 함께 구성된 `preferenced` 및 `failback` 등록 정보를 사용하여 디스크 장치 그룹 노드 목록 등록 정보를 확인합니다. 절차 정보는*Solaris OS*용 *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 개념은 17 페이지 "SunPlex 시스템 하드웨어 및 소프트웨어 구성 요소"를 참조하십시오.

모든 클러스터 노드를 동일하게 구성할 경우 기본 노드와 보조 노드에 동일한 사용 한계를 적용합니다. 모든 노드에서 구성 파일에 있는 모든 응용 프로그램에 대해 동일한 프로젝트 구성 매개 변수를 적용할 필요는 없습니다. 적어도 응용 프로그램의 모든 잠정적 마스터에 있는 프로젝트 데이터베이스에서는 해당 응용 프로그램과 연결된 모든 프로젝트에 액세스할 수 있어야 합니다. 응용 프로그램 1이 *phys-schost-1*에 의해 마스터로 지정되지만 *phys-schost-2* 또는 *phys-schost-3*으로 전환되거나 페일오버될 수 있다고 가정합니다. 응용 프로그램 1에 연결된 프로젝트를 세 노드(*phys-schost-1*, *phys-schost-2*, *phys-schost-3*) 모두에서 액세스할 수 있어야 합니다.

주 – 프로젝트 데이터베이스 정보는 로컬 `/etc/project` 데이터베이스 파일에 저장될 수도 있고 NIS 맵이나 LDAP 디렉토리 서비스에 저장될 수도 있습니다.

Solaris 환경에서는 사용 매개 변수를 유연하게 구성할 수 있지만 Sun Cluster에서는 몇 가지 제한이 부과됩니다. 구성 선택 항목은 사이트의 필요에 따라 다릅니다. 시스템을 구성하기 전에 다음 절의 일반 지침을 따르십시오.

선행 프로세스 가상 메모리 한계 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 Cluster에서 관리 컨트롤을 사용할 경우 불필요한 응용 프로그램 페일오버 및 응용 프로그램 "핑퐁" 효과를 금지하도록 메모리 한계를 적절하게 구성합니다. 일반적으로 다음과 같습니다.

- 메모리 한계를 너무 낮게 설정하지 마십시오.

  응용 프로그램이 메모리 한계에 도달하면 페일오버될 수 있습니다. 이 지침은 가상 메모리 한계에 도달할 경우 예상치 않은 결과가 발생할 수 있는 데이터베이스 응용 프로그램에 특히 중요합니다.

- 기본 노드와 보조 노드에서 메모리 한계를 동일하게 설정하지 마십시오.

  동일한 한계를 설정하면 응용 프로그램이 메모리 한계에 도달하여 동일한 메모리 한계를 갖는 보조 노드에 페일오버될 경우 핑퐁 효과가 발생할 수 있습니다. 보조 노드의 메모리 한계를 약간 더 높게 설정하십시오. 메모리 한계를 각기 다르게 설정하면 핑퐁 시나리오를 방지하여 시스템 관리자가 필요한 경우 매개 변수를 조정할 수 있는 시간을 제공합니다.

- 로드 균형 조정을 위해 자원 관리 메모리 한계를 사용하십시오.

  예를 들어, 메모리 한계를 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 2노드 클러스터" 및 "세 응용 프로그램이 있는 2노드 클러스터"에서는 전체 노드에 대한 페일오버 시나리오를 보여줍니다.
- "자원 그룹 전용 페일오버" 절에서는 응용 프로그램에만 해당되는 페일오버 작업을 설명합니다.

클러스터 환경에서 응용 프로그램은 자원의 일부로 구성되고 자원은 자원 그룹(RG)의 일부로 구성됩니다. 페일오버가 발생하면 자원 그룹은 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 발생됩니다.

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.

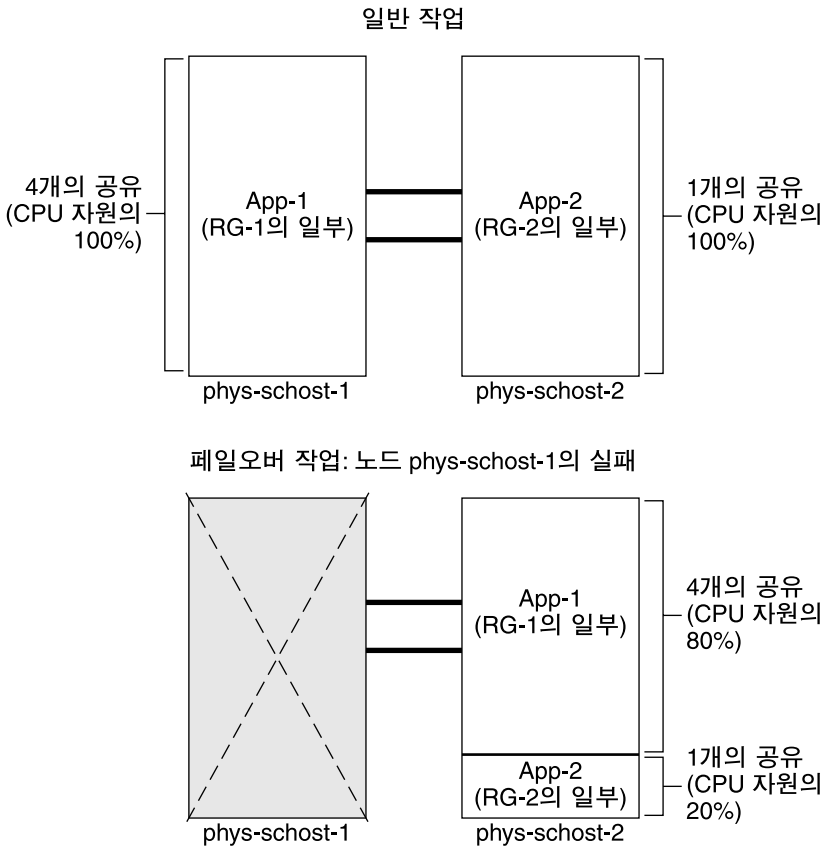
두 응용 프로그램이 있는 2노드 클러스터

2노드 클러스터에서 두 응용 프로그램을 구성하여 각 물리적 호스트(*phys-schost-1*, *phys-schost-2*)가 한 응용 프로그램에 대한 기본 마스터 역할을 하는지 확인할 수 있습니다. 각 물리적 호스트는 다른 물리적 호스트에 대한 보조 노드 역할을 합니다. 응용 프로그램 1 및 응용 프로그램 2에 연결된 모든 프로젝트가 두 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 이 항목에 응용 프로그램 1에 4개의 공유가 할당되고 응용 프로그램 2에 1개의 공유가 할당된다고 지정한 경우는 다음과 같습니다.

```
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 시간 비율은 CPU 시간을 요구하는 각 프로세스에 할당된 공유 개수에 따라 변경될 수 있습니다.



세 응용 프로그램이 있는 2노드 클러스터

세 응용 프로그램이 있는 2노드 클러스터에서는 물리적 호스트(*phys-schost-1*) 하나를 한 응용 프로그램의 기본 마스터로 구성하고 두 번째 물리적 호스트(*phys-schost-2*)를 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.

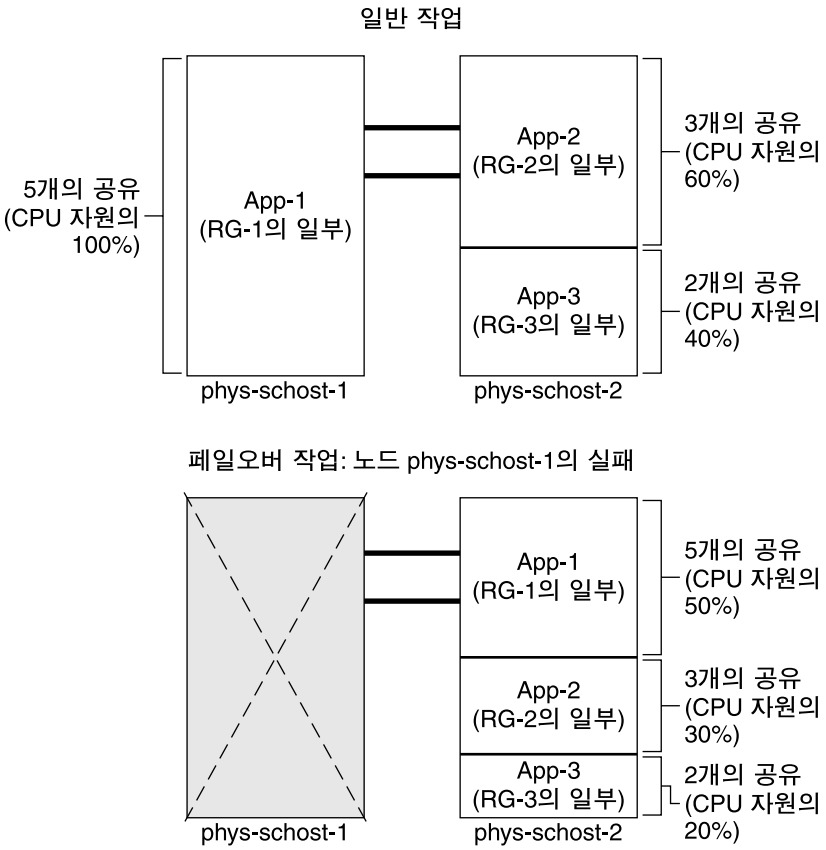
```
Prj_1:103:project for App-1:root::project.cpu-shares=(privileged,5,none)
Prj_2:104:project for App_2:root::project.cpu-shares=(privileged,3,none)
Prj_3:105:project for App_3:root::project.cpu-shares=(privileged,2,none)
```

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 3은 해당 기본 마스터 *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.

페일오버 또는 스위치오버가 발생하고 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.



자원 그룹 전용 페일오버

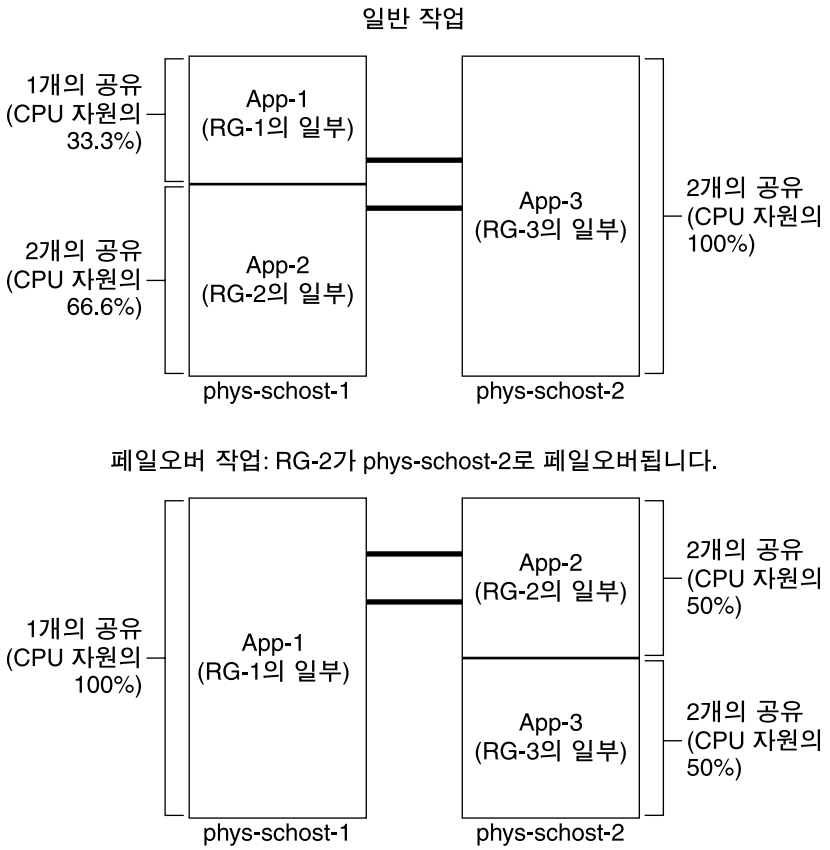
여러 자원 그룹이 동일한 기본 마스터를 갖는 구성에서는 자원 그룹과 관련 응용 프로그램이 보조 노드로 페일오버되거나 전환될 수 있습니다. 반면에 기본 마스터는 클러스터에서 실행됩니다.

주 – 페일오버 중에 페일오버되는 응용 프로그램에는 보조 노드의 구성 파일에 지정된 대로 자원이 할당됩니다. 이 예에서 기본 노드와 보조 노드의 프로젝트 데이터베이스 파일은 동일한 구성을 갖습니다.

예를 들어, 다음 샘플 구성 파일에서는 응용 프로그램 1에 1개의 공유가, 응용 프로그램 2에 2개의 공유가, 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 작업을 설명합니다. 여기서 응용 프로그램 2를 포함하는 RG-2는 *phys-schost-2*에 페일오버됩니다. 할당되는 공유 수는 변경되지 않습니다. 그러나, 각 응용 프로그램에서 사용할 수 있는 CPU 시간 비율은 CPU 시간을 요구하는 각 응용 프로그램에 할당된 공유 개수에 따라 변경될 수 있습니다.



공용 네트워크 어댑터 및 IP 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 IP(Internet Protocol) Network Multipathing 소프트웨어는 공용 네트워크 어댑터를 모니터하고 오류가 감지될 경우 IP 주소를 다른 어댑터로 페일오버하는 기본 기법을 제공합니다. 각 클러스터 노드에는 다른 클러스터 노드에서는 다를 수도 있는 자체 IP 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 있으며, 페일오버가 발생할 때까지는 비활성 상태로 유지되는 대기 인터페이스를 구성할 수 있습니다. `in.mpathd` multipathing 데몬은 테스트 IP 주소를 사용하여 오류 및 복구를 감지합니다. multipathing 데몬이 어댑터 중 하나에서 오류를 감지하면 페일오버가 발생됩니다. 네트워크에 액세스할 때마다 오류가 있는 어댑터가 multipathing 그룹에 있는 일반적으로 작동하는 다른 어댑터로 페일오버되기 때문에 노드에 대한 공용 네트워크 연결이 항상 유지됩니다. 대기 인터페이스를 구성한 경우 데몬은 대기 인터페이스를 선택합니다. 그렇지 않으면, `in.mpathd`가 가장 작은 IP 주소 번호를 갖는 인터페이스를 선택합니다. 페일오버는 어댑터 인터페이스 레벨에서 발생하므로 TCP와 같은 고급 연결은 페일오버 동안의 간단한 임시 지연을 제외하고는 영향을 받지 않습니다. IP 주소 페일오버가 성공적으로 완료되면 ARP 브로드캐스트가 전송됩니다. 따라서 원격 클라이언트에 대한 연결이 유지됩니다.

주 – 일부 세그먼트가 페일오버 동안 유실되면 TCP에서 정체 제어 기법이 활성화되므로 TCP의 정체 복구 특성상 성공적인 페일오버 후 TCP 종점에서 추가 지연이 발생할 수도 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 Network Multipathing 기법의 설계는 어댑터 장애를 발견하고 마스크하기 위한 것입니다. 관리자가 `ifconfig(1M)` 명령으로 논리(또는 공유) IP 주소 중 하나를 제거한 것으로부터 복구하기 위해 설계된 것이 아닙니다. Sun Cluster 소프트웨어는 논리 및 공유 IP 주소를 RGM에 의해 관리되는 자원으로 인식합니다. 관리자가 IP 주소를 추가하거나 제거하는 정확한 방법은 `scrgadm(1M)` 명령을 사용하여 자원이 포함된 자원 그룹을 수정하는 것입니다.

IP Network Multipathing의 Solaris 구현에 대한 자세한 내용은 클러스터에 설치된 Solaris 운영 환경의 해당 설명서를 참조하십시오.

운영 환경 릴리스	참고 항목
Solaris 8 운영 환경	*IP Network Multipathing Administration Guide*
Solaris 9 운영 환경	*System Administration Guide: IP Services*의 "IP Network Multipathing Topics"

SPARC: 동적 재구성 지원

DR(동적 재구성) 소프트웨어 기능에 대한 Sun Cluster 3.1 4/04의 지원이 한 단계씩 개발되고 있습니다. 이 절에서는 Sun Cluster 3.1 4/04의 DR 기능 지원에 대한 개념과 참고 사항을 설명합니다.

Solaris DR 기능에 대하여 문서화된 요구 사항, 절차 및 제한은 모두 Sun Cluster DR 지원에 적용됩니다(운영 체제 작동 정지 제외). 따라서 Sun Cluster 소프트웨어에서 DR 기능을 사용하려면 **먼저** Solaris DR 기능에 대한 설명서를 참조하십시오. 특히 DR 연결 종료 작업 중에 비네트워크 IO 장치에 영향을 주는 문제를 확인해야 합니다. *Sun Enterprise 10000 Dynamic Reconfiguration User Guide* 및 *Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual*(*Solaris 8 on Sun Hardware* 또는 *Solaris 9 on Sun Hardware* 모음 중에 포함)은 `http://docs.sun.com`에서 다운로드할 수 있습니다.

SPARC: 동적 재구성 일반 설명

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

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

활성 구성 요소를 포함하는 보드를 제거하면 시스템 오류가 발생합니다. 보드를 제거하기 전에 DR 하위 시스템은 Sun Cluster와 같은 다른 하위 시스템을 쿼리하여 보드의 구성 요소가 사용 중인지를 확인합니다. 보드가 사용 중이면 DR 보드 제거 작업이 수행되지 않습니다. 즉, DR 하위 시스템은 활성 구성 요소가 포함된 보드에 대해서는 보드 제거 작업을 거부하기 때문에 DR 보드 제거 작업 명령을 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. 하위 수준에서 오류를 보고하면 상위 수준도 오류를 보고합니다. 그러나, 하위 수준에서 특정 오류를 보고할 때 상위 수준에서는 "알 수 없는 오류"로 보고합니다. 시스템 관리자는 상위 수준에서 보고되는 "알 수 없는 오류"를 무시해야 합니다.

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

SPARC: CPU 장치에 대한 DR 클러스터링 고려 사항

Sun Cluster 소프트웨어는 CPU 장치의 존재로 인해 DR 보드 제거 작업을 거부하지 않습니다.

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

SPARC: 메모리에 대한 DR 클러스터링 고려 사항

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

운영 체제에 사용되는 메모리를 커널 메모리 케이지라고 합니다. Sun Cluster 소프트웨어는 커널 메모리 케이지가 포함된 보드에 대해서는 보드 제거 작업을 지원하지 않고 이러한 작업은 모두 거부합니다. DR 보드 제거 작업이 커널 메모리 케이지가 아닌 메모리에 속할 경우 Sun Cluster는 해당 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 고려 사항

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

주 – Sun Cluster에서는 쿼럼 장치의 가용성에 영향을 주는 DR 작업을 할 수 없습니다. 쿼럼 장치에 대한 참고 사항과 쿼럼 장치에 대하여 DR 작업을 수행하기 위한 절차는 79 페이지 “SPARC: 쿼럼 장치에 대한 DR 클러스터링 고려 사항”을 참조하십시오.

이러한 작업을 수행하는 방법에 대한 자세한 내용은 *Solaris OS*용 *Sun Cluster* **시스템 관리 안내서**의 “작업 맵: 쿼럼 장치 동적 재구성”을 참조하십시오.

SPARC: 쿼럼 장치에 대한 DR 클러스터링 고려 사항

DR 보드 제거 작업이 쿼럼에 대해 구성된 장치의 인터페이스를 포함하는 보드에 속할 경우 Sun Cluster는 해당 작업을 거부하고 작업의 영향을 받는 쿼럼 장치를 식별합니다. DR 보드 제거 작업을 수행하기 전에 장치를 쿼럼 장치로 비활성화해야 합니다.

이러한 작업을 수행하는 방법에 대한 자세한 내용은 *Solaris OS*용 *Sun Cluster* **시스템 관리 안내서**의 "작업 맵: 쿼럼 장치 동적 재구성"을 참조하십시오.

SPARC: 클러스터 상호 연결 인터페이스에 대한 DR 클러스터링 고려 사항

DR 보드 제거 작업이 활성 클러스터 상호 연결 인터페이스를 포함하는 보드에 속할 경우 Sun Cluster는 해당 작업을 거부하고 작업의 영향을 받는 인터페이스를 식별합니다. DR 작업을 성공하려면 Sun Cluster 관리 도구를 사용하여 활성 인터페이스를 비활성화해야 합니다(아래 주의 참조).

이러한 작업을 수행하는 방법에 대한 자세한 내용은 *Solaris OS*용 *Sun Cluster* **시스템 관리 안내서**의 "클러스터 상호 연결 관리"를 참조하십시오.

주의 – Sun Cluster에서는 각 클러스터 노드에서 다른 모든 클러스터 노드에 대하여 하나 이상의 경로가 작동하고 있어야 합니다. 다른 클러스터 노드에 대한 마지막 경로를 지원하는 독립 상호 연결 인터페이스를 비활성화하면 안됩니다.

SPARC: 공용 네트워크 인터페이스에 대한 DR 클러스터링 고려 사항

DR 보드 제거 작업이 활성 공용 네트워크 인터페이스를 포함하는 보드에 속할 경우 Sun Cluster는 해당 작업을 거부하고 작업의 영향을 받는 인터페이스를 식별합니다. 활성 네트워크 인터페이스가 있는 보드를 제거하기 전에 `if_mpadm`(1M) 명령을 사용하여 해당 인터페이스의 모든 트래픽을 정상적으로 작동하는 Multipathing 그룹의 다른 인터페이스로 전환해야 합니다.

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 OS*용 *Sun Cluster* **시스템 관리 안내서**의 "공용 네트워크 관리"를 참조하십시오.

4장

질문과 대답

INDEXTERM-343

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

고가용성 FAQ

- **고가용성 시스템이란 정확히 무엇입니까?**

  SunPlex 시스템은 서버 시스템을 정상적으로 사용할 수 없는 장애가 발생할 경우에도 응용 프로그램을 계속 실행하는 클러스터의 기능을 고가용성(HA)이라고 정의합니다.

- **클러스터는 어떤 프로세스를 통해 고가용성을 제공합니까?**

  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.

- **페일오버와 확장 가능 데이터 서비스 간의 차이점은 무엇입니까?**

  데이터 서비스에는 페일오버와 확장 가능 두 가지의 주요 기능이 있습니다.

  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 노드가 실패할 경우, 실패한 노드에서 실행되는 응용 프로그램은 다른 노드로 페일오버하여 실행을 계속합니다.

  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. 클러스터에는 여러 개의 GIF 노드가 있을 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 실행하는 노드에 장애가 발생하면 장애가 발생한 노드가 클러스터에 복귀될 때까지 응용 프로그램이 다른 노드에서 계속 실행되고, 이 경우에는 약간 성능이 떨어집니다.

파일 시스템 FAQ

- **하나 이상의 클러스터 노드를 가용성이 높은 NFS 서버로 실행하고 다른 클러스터 노드는 클라이언트로 실행할 수 있습니까?**

  안 됩니다. 루프백 마운트를 하면 안 됩니다.

- **Resource Group Manager의 제어를 받지 않는 응용 프로그램에 클러스터 파일 시스템을 사용할 수 있습니까?**

  예. 그러나 RGM 제어가 없으면 응용 프로그램을 실행하고 있는 노드에 장애가 발생할 경우에 직접 응용 프로그램을 다시 시작해야 합니다.

- **모든 클러스터 파일 시스템에서 /global 디렉토리에 마운트 지점이 있어야 합니까?**

  아닙니다. 그러나 `/global`과 같이 동일한 마운트 지점에 클러스터 파일 시스템을 두면 이러한 파일 시스템을 쉽게 구성하고 관리할 수 있습니다.

- **클러스터 파일 시스템을 사용하는 것과 NFS 파일 시스템을 내보내는 것이 어떤 차이가 있습니까?**

  몇 가지 차이점이 있습니다.

  1. 클러스터 파일 시스템은 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, 읽기, 쓰기, 파일 잠금 및 비동기 I/O를 위해 여러 노드에서 파일에 액세스합니다.
  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 이를 수행하는 데 더 많은 시간이 소요될 수 있습니다.

- **파일 시스템 /global/.devices/node@<nodeID>가 클러스터 노드에 나타납니다. 이 파일 시스템을 사용하여 가용성이 높은 전역 데이터를 저장할 수 있습니까?**

  이 파일 시스템은 전역 장치 이름 공간을 저장합니다. 이것은 일반적인 용도에 사용하는 파일 시스템이 아닙니다. 데이터는 전역이지만 전역 방식으로 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 안 됩니다.

볼륨 관리 FAQ

- **모든 디스크 장치를 미러해야 합니까?**

  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.

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

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

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

데이터 서비스 FAQ

- **사용할 수 있는 SunPlex 데이터 서비스는 무엇입니까?**

  지원되는 데이터 서비스 목록은 *Sun Cluster 3.1 9/04 Release Notes for Solaris OS*의 "Supported Products"를 참조하십시오.

- **SunPlex 데이터 서비스에서 지원되는 응용 프로그램 버전은 무엇입니까?**

  지원되는 응용 프로그램 버전 목록은 *Solaris OS***용** *Sun Cluster 3.1 9/04* **릴리스 노트**의 "지원 제품"을 참조하십시오.

- **자체 데이터 서비스를 작성할 수 있습니까?**

  예. 자세한 내용은 *Solaris OS***용** *Sun Cluster* **데이터 서비스 개발 안내서**의 "데이터 서비스 개발 라이브러리 참조"를 참조하십시오.

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

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

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

  Sun Cluster HA for NFS의 경우가 아니면 문서가 페일오버 모드 자원 그룹의 `LogicalHostname` 자원을 사용하기 위해 호출할 때마다 `SharedAddress` 자원 또는 `LogicalHostname` 자원 중 한 가지를 사용할 수 있습니다. `SharedAddress` 자원을 사용하면 클러스터 네트워킹 소프트웨어가 `SharedAddress`에 대해 구성되며, `LogicalHostname`에 대해 구성되지 않으므로 일부 추가 오버헤드가 발생합니다.

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

공용 네트워크 FAQ

- **어떤 공용 네트워크 어댑터가 SunPlex 시스템을 지원합니까?**

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

- **페일오버에서 MAC 주소의 역할은 무엇입니까?**

  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.

- **SunPlex 시스템은 `local-mac-address?=true` 설정을 지원합니까**?

  예. 실제로 IP Network Multipathing에서는 `local-mac-address?`가 `true`로 설정되어야 합니다.

  SPARC 기반 클러스터의 OpenBoot PROM `ok` 프롬프트에서 `eeprom`(1M)을 사용하여 `local-mac-address?`를 설정할 수 있습니다. x86 기반 클러스터에서는 BIOS 부트 이후 선택적으로 실행하는 SCSI 유틸리티를 사용하여 설정할 수 있습니다.

- **IP Network Multipathing에서 어댑터 스위치오버를 수행할 때 어느 정도 지연될 수 있습니까?**

  몇 분 동안 지연될 수 있습니다. 이것은 IP Network Multipathing 전환이 수행될 때 ARP를 외부로 전송하기 때문입니다. 그러나 클라이언트와 클러스터 사이의 라우터가 반드시 ARP를 사용하는 것은 아닙니다. 따라서 라우터에서 이 IP 주소에 대한

ARP 캐시 항목의 시간이 만료될 때까지 이전의 MAC 주소를 사용할 수 있습니다.

- **네트워크 어댑터 오류가 얼마나 빨리 감지됩니까?**

  기본 오류 감지 시간은 10초입니다. 알고리즘에서 오류 감지 시간을 맞추려 하지만 실제 시간은 네트워크 로드에 따라 달라집니다.

클러스터 구성원 FAQ

- **모든 클러스터 구성원이 동일한 루트 암호를 사용해야 합니까?**

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

- **노드의 부트 순서가 중요합니까?**

  대부분의 경우에는 그렇지 않습니다. 그러나 정보 유실을 방지하기 위해서는 부트 순서가 중요합니다(정보 유실에 대한 내용은 47 페이지 “장애 차단 정보” 참조). 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. 그러면, 클러스터 구성 정보 날짜가 지난 노드를 가져오는 일이 없어집니다.

- **클러스터 노드에서 로컬 디스크를 미러해야 합니까?**

  예. 이러한 미러링이 반드시 필요한 것은 아니지만, 클러스터 노드 디스크를 미러하면 디스크를 미러하지 않을 경우에 발생할 수 있는 노드 중단을 방지할 수 있습니다. 클러스터 노드의 로컬 디스크를 미러하면 시스템 관리에 오버헤드가 부가됩니다.

- **클러스터 구성원을 백업하는 데는 어떤 문제가 있습니까?**

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

  데이터 백업 및 복원 방법에 대한 추가 정보는 *Solaris OS*용 *Sun Cluster* **시스템 관리 안내서**의 “클러스터 백업 및 복원”을 참조하십시오.

- **보조 노드로 사용될 수 있는 노드 상태는 언제입니까?**

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

클러스터 저장소 FAQ

- **어떻게 멀티 호스트 저장소의 가용성을 높입니까?**

  멀티 호스트 저장소는 하나의 디스크에 장애가 발생한 후에도 미러링이나 하드웨어 기반 RAID-5 컨트롤러를 통해 계속 사용할 수 있기 때문에 가용성이 높습니다. 멀티 호스트 저장 장치는 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

- **SunPlex 시스템에서 지원하는 클러스터 상호 연결은 무엇입니까?**

  현재 SunPlex 시스템은 SPARC 기반 및 x86 기반 클러스터 모두에서 이더넷(100BASE-T Fast Ethernet 및 1000BASE-SX Gb) 클러스터 상호 연결을 지원합니다. SunPlex 시스템은 SPARC 기반 클러스터에서만 SCI 네트워크 인터페이스 클러스터 상호 연결을 지원합니다.

- **"케이블"과 전송 "경로" 사이에는 어떤 차이가 있습니까?**

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

  케이블은 관리자에 의해 정적으로 "활성화"되고 "비활성화"됩니다. 케이블의 상태는 활성화 또는 비활성화된 "정적인 상태(state)"를 말하는 것이지, "동적인 상태(status)"를 말하는 것이 아닙니다. 케이블이 비활성화되면 구성되지 않은 것으로 간주됩니다. 비활성화된 케이블은 전송 경로로 사용할 수 없습니다. 비활성화된 케이블은 검사 대상이 아니므로 상태를 알 수 없습니다. 케이블의 상태는 `scconf -p` 명령을 사용하여 볼 수 있습니다.

  전송 경로는 클러스터 토폴로지 관리자에 의해 동적으로 구성됩니다. 전송 경로의 "상태(status)"는 토폴로지 관리자에 의해 결정됩니다. 경로의 상태는 "온라인" 또는 "오프라인"일 수 있습니다. 전송 경로의 상태는 `scstat(1M)`를 사용하여 확인할 수 있습니다.

  다음은 케이블이 4개인 2 노드 클러스터의 예입니다.

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

  네 개의 케이블로 두 개의 전송 경로를 만들 수 있습니다.

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

클라이언트 시스템 FAQ

- **클러스터에서 사용할 경우 특수 클라이언트 요구 사항이나 제한 사항을 고려해야 합니까?**

  클라이언트 시스템은 다른 서버에서처럼 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*의 해당 장을 참조하십시오.

관리 콘솔 FAQ

- **SunPlex 시스템에 관리 콘솔이 필요합니까?**

  예.

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

  SunPlex 시스템에는 전용 관리 콘솔이 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 콘솔이 클러스터에 "가까이"(예: 같은 방) 있어야 합니까?**

  하드웨어 서비스 제공업체에 확인해 보십시오. 제공업체에서 클러스터 자체에 근접하게 콘솔이 위치되도록 요구할 수도 있습니다. 콘솔이 같은 방에 위치되어야 하는 기술적인 이유는 없습니다.

- **거리 요구 사항이 일단 충족되면 관리 콘솔이 둘 이상의 클러스터에 서비스를 제공할 수 있습니까?**

  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

- **SunPlex 시스템에 단말기 집중 장치가 필요합니까?**

  Sun Cluster 3.0으로 시작하는 모든 소프트웨어 버전은 단말기 집중 장치 없이 실행됩니다. 장애를 방지하기 위해 단말기 집중 장치가 필요했던 Sun Cluster 2.2 제품과 달리, 이후 버전에서는 단말기 집중 장치가 반드시 필요하지는 않습니다.

- **대부분의 SunPlex 서버에서 단말기 집중 장치를 사용하는데 Sun Enterprise E10000 server는 이것을 사용하지 않습니다. 그 이유는 무엇입니까?**

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

- **단말기 집중 장치를 사용할 경우 어떤 점이 좋습니까?**

  단말기 집중 장치를 사용하면 SPARC 기반 노드가 OpenBoot PROM(OBP)에 있거나 x86 기반 노드가 부트 하위 시스템에 있을 경우를 비롯하여 네트워크의 모든 원격 워크스테이션에서 각 노드에 콘솔 수준으로 액세스할 수 있습니다.

- **Sun에서 지원하지 않는 단말기 집중 장치를 사용할 경우, 어떤 사항을 알아야 합니까?**

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

- **Sun에서 지원하는 단말기 집중 장치의 잠긴 포트를 재부트하지 않고 해제할 수 있습니까?**

  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 지원되는 단말기 집중 장치의 구성 및 관리 방법에 대한 자세한 내용은 다음 설명서를 참조하십시오.

  - *Solaris OS*용 *Sun Cluster* **시스템 관리 안내서**의 "Sun Cluster 관리 개요"
  - *Sun Cluster 3.x Hardware Administration Manual for Solaris OS*의 "Installing and Configuring the Terminal Concentrator"

- **단말기 집중 장치 자체가 실패할 경우에는 어떻습니까? 다른 단말기 집중 장치를 준비해야 합니까?**

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

- **단말기 집중 장치를 사용할 경우에 보안 문제는 없습니까?**

  일반적으로, 단말기 집중 장치는 시스템 관리자가 사용되는 소규모 네트워크에 접속되며, 다른 클라이언트 액세스에 사용되는 네트워크에는 접속되지 않습니다. 특수 네트워크에 대한 액세스를 제한하여 보안을 제어할 수 있습니다.

- **SPARC: 테이프 또는 디스크 드라이브에서 동적 재구성을 어떻게 사용합니까?**

  - 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 포함되었으면 79 페이지 "SPARC: 디스크 및 테이프 드라이브에 대한 DR 클러스터링 고려 사항"으로 이동하십시오.
  - 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