



# Sun Cluster 簡介 (適用於 Solaris 作業系統)

---

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

文件號碼: 817-6914-10  
2004 年 4 月, 修訂版 A

Copyright 2004 Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 U.S.A. 版權所有

此產品或文件受著作權的保護，其使用、複製、分送以及反編譯，均受軟體使用權限制。未經 Sun 及其授權者 (如果有的話) 的書面授權，不得以任何形式或任何方法，重新製造本產品或文件的任何部分。至於協力廠商的軟體，包括字型技術，亦受著作權保護，並經過 Sun 供應商授權使用。

產品的某些部分可能源自 Berkeley BSD 系統，由加州大學授權。UNIX 是在美國和其他國家/地區的註冊商標，由 X/Open Company, Ltd. 獨家授權。

Sun、Sun Microsystems、Sun 徽標、docs.sun.com、AnswerBook、AnswerBook2、以及 Solaris 是 Sun Microsystems, Inc. 在美國和其他國家/地區的商標、註冊商標或服務標記。所有的 SPARC 商標都是在獲得授權的情況下使用，而且是 SPARC International, Inc. 在美國和其他國家/地區的商標或註冊商標。帶有 SPARC 商標的產品基於 Sun Microsystems, Inc. 開發的架構。

OPEN LOOK 和 Sun™ Graphical User Interface 是 Sun Microsystems, Inc. 為其使用者和授權者而開發的。Sun 認可 Xerox 研發電腦業之視覺化或圖形化使用者介面觀念的先驅貢獻。對於「Xerox 圖形使用者介面」，Sun 保有來自於 Xerox 的非獨家授權，這項授權的適用也涵蓋取得 Sun 的授權而使用 OPEN LOOK GUI、或者遵循 Sun 的書面授權合約的廠商。

美國政府的權利 – 商業軟體。政府使用者要遵守 Sun Microsystems, Inc. 的標準授權合約以及 FAR 及其補遺的適當條款。

文件以「現狀」提供，所有明示或暗示的條件、陳述與保證，包括對於適銷性、特定用途的適用性或非侵權行為的任何暗示性保證在內，均不負責，除非此免責聲明在法律上被認定為無效。



040521 @ 8606



# 目錄

---

前言 5

## 1 Sun Cluster 簡介 9

- 透過 Sun Cluster 使應用程式具有高度可用性 9
  - 可用性管理 10
  - 故障轉移服務和可延伸服務以及並列應用程式 10
  - IP 網路多重路徑 10
  - 儲存管理 11
- 監視故障 12
- 管理與配置工具 12
  - SunPlex Manager 13
  - 指令行介面 13
  - Sun Management Center 13
  - 基於角色存取控制 14

## 2 Sun Cluster 的主要概念 15

- 叢集節點 15
- 叢集互連 16
- 叢集成員關係 16
- 叢集配置儲存庫 17
- 故障監視器 17
  - 資料服務監視 17
  - 磁碟路徑監視 17
  - IP 多重路徑監視 18
- 法定裝置 18

|                         |           |
|-------------------------|-----------|
| 資料完整性                   | 18        |
| 故障隔離                    | 19        |
| 用於故障隔離的 Failfast 機制     | 19        |
| Devices                 | 20        |
| 全域裝置                    | 20        |
| 本機裝置                    | 21        |
| 磁碟裝置群組                  | 21        |
| 資料服務                    | 21        |
| 資源類型                    | 22        |
| 資源                      | 22        |
| 資源群組                    | 22        |
| 資料服務類型                  | 22        |
| <b>3 Sun Cluster 架構</b> | <b>25</b> |
| Sun Cluster 硬體環境        | 25        |
| Sun Cluster 軟體環境        | 26        |
| 叢集成員關係監視器               | 27        |
| 叢集配置儲存庫 (CCR)           | 28        |
| 叢集檔案系統                  | 28        |
| 可延伸資料服務                 | 29        |
| 平衡資料流量策略                | 30        |
| 多重主機磁碟儲存裝置              | 30        |
| 叢集互連                    | 31        |
| IP 網路多重路徑群組             | 31        |
| 公用網路介面                  | 32        |
| <br>                    |           |
| 索引                      | 33        |

# 前言

---

「*Sun™ Cluster* 簡介 (適用於 *Solaris* 作業系統)」透過說明 Sun Cluster 產品的用途及其實現此用途所用方法介紹 Sun Cluster 產品。本書還說明 Sun Cluster 的關鍵概念。本文件中的資訊可使您熟悉 Sun Cluster 的功能和功能性。

---

## 相關說明文件

有關 Sun Cluster 相關主題的資訊可從下表中列出的文件中獲得。所有 Sun Cluster 文件均可從 <http://docs.sun.com> 處獲得。

| 主題     | 說明文件   |
|--------|--|
| 概念     | 「 <i>Sun Cluster</i> 概念指南 (適用於 <i>Solaris</i> 作業系統)」   |
| 概述     | 「 <i>Sun Cluster</i> 簡介 (適用於 <i>Solaris</i> 作業系統)」   |
| 字彙表    | 「 <i>Sun Java Enterprise System 2003Q4 Glossary</i> 」  |
| 硬體管理   | 「 <i>Sun Cluster Hardware Administration Manual for Solaris OS</i> 」<br>個人硬體管理指南                 |
| 軟體安裝   | 「 <i>Sun Cluster</i> 軟體安裝指南 (適用於 <i>Solaris</i> 作業系統)」   |
| 資料服務管理 | 「 <i>Sun Cluster Data Service Planning and Administration Guide for Solaris OS</i> 」<br>個人資料服務指南 |
| 資料服務發展 | 「 <i>Sun Cluster</i> 資料服務開發者指南 (適用於 <i>Solaris</i> 作業系統)」  |
| 系統管理   | 「 <i>Sun Cluster System</i> 系統管理指南 (適用於 <i>Solaris</i> 作業系統)」                                    |

| 主題      | 說明文件   |
|---------|--|
| 錯誤訊息    | 「 <i>Sun Cluster Error Messages Guide for Solaris OS</i> 」 |
| 指令和功能參考 | 「 <i>Sun Cluster Reference Manual for Solaris OS</i> 」     |

如需 Sun Cluster 文件的完整清單，請參閱您所用 Sun Cluster 軟體發行版本的版本說明 (位於 <http://docs.sun.com>)。

---

## 線上存取 Sun 說明文件

[docs.sun.com](http://docs.sun.com)<sup>SM</sup> 網站可讓您存取 Sun 線上技術文件。您可以瀏覽 [docs.sun.com](http://docs.sun.com) 的歸檔檔案或搜尋特定書名或主題。此 URL 是 <http://docs.sun.com>。

---

## 訂購 Sun 說明文件

Sun Microsystems 提供列印的選取產品說明文件。如需文件清單以及如何訂購它們的有關資訊，請參閱位於 <http://docs.sun.com> 的「購買書面文件」。

---

## 取得說明

如果在安裝或使用 Sun Cluster 系統時遇到問題，請聯絡您的服務供應商並提供以下資訊：

- 您的姓名和電子郵件地址 (如果有的話)
- 您的公司名稱、地址和電話號碼
- 您系統的機型和序號
- 作業環境的版次編號 (例如，Solaris 9)
- Sun Cluster 軟體的版次編號 (例如，3.1 4/04)

使用下列指令收集您系統上每一個節點的相關資訊，提供給您的服務供應商：

| AdminTool 指令               | 函式                             |
|----------------------------|--------------------------------|
| <code>prtconf -v</code>    | 顯示系統記憶體的大小及報告周邊裝置的相關資訊         |
| <code>psrinfo -v</code>    | 顯示處理器的相關資訊                     |
| <code>showrev -p</code>    | 報告安裝了哪些修補程式                    |
| <code>prtdiag -v</code>    | 顯示系統診斷資訊                       |
| <code>scinstall -pv</code> | 顯示 Sun Cluster 軟體發行版本和套裝軟體版本資訊 |
| <code>scstat</code>        | 提供叢集狀態的快照                      |
| <code>scconf -p</code>     | 列示叢集配置資訊                       |
| <code>scrgadm -p</code>    | 顯示有關已安裝資源、資源群組與資源類型的資訊         |

同時提供 `/var/adm/messages` 檔案的內容。

## 印刷排版慣例

下表說明本書在印刷上所作的變更。

表 P-1 印刷排版慣例

| 字體或符號                  | 涵義                   | 範例   |
|------------------------|----------------------|--|
| <code>AaBbCc123</code> | 指令、檔案和目錄的名稱以及電腦螢幕的輸出 | 編輯您的 <code>.login</code> 檔案。<br>請使用 <code>ls -a</code> ，來列出所有的檔案。<br><code>machine_name% you have mail.</code> |
| <b>AaBbCc123</b>       | 您鍵入的內容，與電腦螢幕上的輸出相對照  | <code>machine_name% su</code><br><code>password:</code>  |
| <code>AaBbCc123</code> | 指令行預留位置：用實際名稱或值取代    | 要刪除一個檔案，請鍵入 <code>rm filename</code> 。   |
| <code>AaBbCc123</code> | 書名、新專有名詞，以及要強調的專有名詞  | 請閱讀「使用者指南」中的第 6 章。<br>這些被稱為類別選項。<br>您必須是 <code>root</code> ，才能執行此動作。   |

---

## 指令範例中的 Shell 提示符號

以下表格列出使用於 C shell、Bourne shell 和 Korn shell 的預設系統提示符號以及超級使用者提示符號。

表 P-2 Shell 提示符號

| Shell                               | 提示符號          |
|-------------------------------------|---------------|
| C shell 提示符號                        | machine_name% |
| C shell 超級使用者提示符號                   | machine_name# |
| Bourne shell 和 Korn shell 提示符號      | \$            |
| Bourne shell 和 Korn shell 超級使用者提示符號 | #             |

# 第 1 章

---

## Sun Cluster 簡介

---

SunPlex 系統為一整合的硬體與 Sun Cluster 軟體解決方案，用於建立具有高度可用性的服務與可延伸服務。本章提供 Sun Cluster 功能的高階簡介。

本章包含下列小節：

- 第 9 頁的「透過 Sun Cluster 使應用程式具有高度可用性」
- 第 12 頁的「監視故障」
- 第 12 頁的「管理與配置工具」

---

## 透過 Sun Cluster 使應用程式具有高度可用性

叢集是作為單一連續可用的系統共同作業，以為使用者提供應用程式、系統資源和資料的兩個或多個系統 (或節點)。叢集中的每一個節點均是具有完整功能的獨立系統。但是，在叢集環境中，這些節點透過互連連線，並作為單一實體共同作業以提供增強的可用性和效能。

具有高度可用性的叢集透過使叢集在發生故障的整個過程中保持執行狀態，來提供對資料和應用程式的幾乎持續的存取，而這些故障通常會使單一伺服器系統當機。沒有任何單一故障 — 硬體、軟體或網路 — 可以使叢集失敗。相比之下，容錯性硬體系統雖提供對資料和應用程式的持續存取，但因為是特殊硬體，所以成本較高。容錯性系統通常無法處理軟體故障。

每個 Sun Cluster 系統是緊密連接的節點之集合，這些節點提供網路服務和應用程式的單一管理檢視。Sun Cluster 系統透過以下硬體和軟體的組合實現高度可用性：

- 冗餘磁碟系統提供儲存。通常，對這些磁碟系統進行鏡像，以使在磁碟或子系統失敗時作業不被中斷。磁碟系統的冗餘連線確保在伺服器、控制器或電纜出現故障時仍可存取資料。節點間的高速互連提供對資源的存取。還可將叢集中的所有節點連線至公用網路，以使多個網路上的用戶端可以存取該叢集。

- 可熱交換的冗餘元件 (例如電源供應器和散熱系統) 透過使系統在硬體發生故障後繼續作業改善可用性。可熱交換的元件提供在運行的系統中新增或移除硬體元件而不會使其當機的功能。
- Sun Cluster 軟體的高度可用性架構可迅速偵測到節點故障，並將應用程式或服務遷移至在同一環境中執行的另一節點。任何時候都不會出現所有應用程式均不可用的情況。不受故障節點影響的應用程式在恢復期間完全可用。此外，故障節點的應用程式一經恢復之後，便可使用。已恢復的應用程式無須等待所有其他應用程式完成其恢復。

## 可用性管理

如果系統內發生任何單一軟體或硬體故障，而應用程式仍不受其影響，則表示該應用程式具有高度可用性。由應用程式自身內部的錯誤或資料毀壞導致的故障除外。以下各項適用於具有高度可用性的應用程式：

- 恢復過程對於使用資源的應用程式是透明的。
- 資源存取完全不受節點故障的影響。
- 應用程式無法偵測到宿主節點是否已被移至另一節點。
- 單一節點的故障對使用連結至該節點的檔案、裝置和磁碟容體之其餘節點上的程式是完全透明的。

## 故障轉移服務和可延伸服務以及並列應用程式

故障轉移服務和可延伸服務以及並列應用程式可讓您使應用程式具有高度可用性並改善叢集上應用程式的效能。

故障轉移服務透過冗餘提供高度可用性。發生故障時，您可以將正在執行的應用程式配置為在同一節點上重新啟動，或移至叢集中另一節點，而無需使用者介入。

為提昇效能，可延伸服務平衡利用叢集中的多個節點以並行方式執行應用程式。在可延伸的配置中，叢集的每個節點均可提供資料並處理用戶端的要求。

並列資料庫可讓資料庫伺服器的多個實例執行以下作業：

- 加入叢集
- 同時處理對同一資料庫的不同查詢
- 在大型查詢中提供並列查詢功能

如需有關故障轉移服務和可延伸服務以及並列應用程式的更多資訊，請參閱第 22 頁的「資料服務類型」。

## IP 網路多重路徑

用戶端透過公用網路來將要求送至叢集。每個叢集節點均透過一個或多個公用網路配接卡連線至至少一個公用網路。

IP 網路多重路徑連結使伺服器有連線至同一子網路的多個網路埠。首先，IP 網路多重路徑連結軟體透過偵測網路配接卡的故障或修復提供從網路配接卡故障恢復的功能。然後，該軟體同時將網路位址在故障配接卡和替代配接卡之間切換。當多個網路配接卡可作業時，IP 網路多重路徑連結可透過在配接卡間分散外送的資料封包以提昇資料流量。

## 儲存管理

多重主機儲存裝置透過將磁碟連線至多個節點使磁碟具有高度可用性。多節點為存取資料提供多個路徑，當某個路徑失敗時，另一個路徑可以替代它。

多重主機磁碟可允許進行以下叢集處理：

- 容許單一節點發生故障。
- 集中應用程式資料、應用程式二進位檔及配置檔案。
- 免受故障節點的影響。如果用戶端要求是透過失敗的某個節點來存取資料，這些要求會切換至使用另一個直接連線相同磁碟的節點。
- 以全域方式透過「控制」磁碟的主要節點，或透過經由本地路徑的直接並行存取來提供存取。

## 容體管理支援

容體管理程式使您可以管理大量磁碟以及那些磁碟上的資料。容體管理程式透過提供以下功能，可以增加儲存容量和資料可用性：

- 磁碟機分條和鏈結
- 磁碟鏡像
- 磁碟機緊急備援
- 磁碟故障處理和磁碟更換

Sun Cluster 系統支援以下容體管理程式：

- Solaris Volume Manager
- VERITAS Volume Manager

## Sun StorEdge Traffic Manager

Sun StorEdge Traffic Manager 軟體從 Solaris 作業系統 8 核心 I/O 架構開始完全整合。Sun StorEdge Traffic Manager 軟體使您可以更有效地表示和管理裝置，這些裝置可透過 Solaris 作業環境之單一實例內的多個 I/O 控制器介面存取。Sun StorEdge Traffic Manager 架構啟動以下功能：

- 避免由於 I/O 控制器發生故障而受到 I/O 中斷的影響
- 在 I/O 控制器發生故障時自動切換至替代控制器
- 透過在多個 I/O 通道間平衡資料流量獲得提昇的 I/O 效能

## 硬體備用獨立式磁碟陣支援

Sun Cluster 系統支援使用硬體備用獨立式磁碟陣 (RAID) 和基於主機的軟體 RAID。硬體 RAID 使用儲存陣列的硬體冗餘或儲存系統的硬體冗餘，以確保獨立硬體故障不會影響資料可用性。如果您對各個獨立的儲存陣列進行鏡像，基於主機的軟體 RAID 可確保在整個儲存陣列處於離線狀態時，獨立硬體故障不會影響資料可用性。雖然您可以並行使用硬體 RAID 和基於主機的軟體 RAID，但您僅需要一個 RAID 解決方案以維持高度的資料可用性。

## 檔案系統支援

由於叢集系統的固有屬性之一為共用資源，因此叢集需要一個檔案系統，以滿足統一共用檔案的需要。Sun Cluster 檔案系統讓使用者或應用程式透過使用遠端或本機標準 UNIX API 存取叢集的任何節點上的任何檔案。如果將應用程式從一個節點移至另一節點，無需進行任何變更，該應用程式仍可存取相同檔案。無需進行任何變更，現有應用程式仍可完全利用叢集檔案系統。

---

## 監視故障

Sun Cluster 系統透過使用多重主機磁碟、多重路徑以及全域檔案系統使使用者和資料間的路徑具有高度可用性。Sun Cluster 系統監視以下各項的故障：

- 應用程式 – 大多數 Sun Cluster 資料服務提供一個故障監視器，該故障監視器定期探測資料服務以確定其運作狀態。故障監視器驗證應用程式常駐程式是否在執行以及用戶端是否得到服務。根據探測傳回的資訊，可以啟動預先定義的動作，例如重新啟動常駐程式或進行故障轉移。
- 磁碟路徑 – Sun Cluster 軟體支援磁碟路徑監視 (DPM)。DPM 透過報告輔助磁碟路徑的故障改善故障轉移和切換保護轉移的整體可靠性。
- 網際網路通訊協定 (IP) 多重路徑 – Sun Cluster 系統上的 Solaris IP 網路多重路徑軟體提供監視公用網路配接卡的基本機制。IP 多重路徑還可在偵測到故障時，使 IP 位址從一個配接卡故障轉移至另一配接卡。

---

## 管理與配置工具

您可以透過 SunPlex Manager GUI 或指令行介面 (CLI) 安裝、配置以及管理 Sun Cluster 系統。

Sun Cluster 系統還有一個作為 Sun Management Center 軟體的一部分執行的模組，該模組為某些叢集作業提供 GUI。

## SunPlex Manager

SunPlex Manager 是基於瀏覽器的工具，用以管理 Sun Cluster 系統。SunPlex Manager 軟體使管理員可以執行系統管理與監視、軟體安裝以及系統配置。

SunPlex Manager 軟體包括以下功能。

- 內建安全與授權機制
- 安全套接層 (SSL) 支援
- 基於角色存取控制 (RBAC)
- 可插接式驗證模組 (PAM)
- NAFO 與 IP 網路多重路徑群組管理工具
- 法定裝置、傳輸、共用儲存裝置以及資源群組管理
- 複雜錯誤檢查與私有互連的自動偵測

## 指令行介面

Sun Cluster 指令行介面是一組公用程式，您可以將其用於安裝和管理 Sun Cluster 系統以及管理 Sun Cluster 軟體的容體管理程式部分。

您可以透過 Sun Cluster CLI 執行以下 SunPlex 管理作業：

- 驗證 Sun Cluster 配置
- 安裝和配置 Sun Cluster 軟體
- 更新 Sun Cluster 配置
- 管理資源類型的註冊、資源群組的建立，以及資源群組內資源的啟動
- 變更資源群組與磁碟裝置群組的節點主次關係與狀態
- 使用基於角色存取控制 (RBAC) 來控制存取
- 關閉整個叢集

## Sun Management Center

Sun Cluster 系統還有一個作為 Sun Management Center 軟體的一部分執行的模組。Sun Management Center 軟體用作叢集管理和監視作業的基礎，並使系統管理員可以透過 GUI 或 CLI 執行以下作業：

- 配置遠端系統
- 監視效能
- 偵測與隔離硬體和軟體故障

Sun Management Center 軟體還可用作管理 Sun Cluster 伺服器內動態重新配置的介面。動態重新配置包括網域建立、動態板連結以及動態拆離。

## 基於角色存取控制

在傳統的 UNIX 系統中，root 使用者 (也稱為超級使用者) 的權力不受限制，它能夠讀取和寫入任何檔案、執行所有程式以及向任何程序傳送結束訊號。Solaris 基於角色存取控制 (RBAC) 可以替代具有無限權限的超級使用者模型。RBAC 使用最少權限的安全原則，即僅給予使用者執行他/她的工作所需的權限。

RBAC 使組織可以分佈超級使用者能力，並將它們組合至特殊使用者帳號或角色以指定給特定個人。這種分佈與組合可使安全策略多樣化。可在安全、網路、防火牆、備份以及系統作業之類的區域內為特殊用途的管理員設定帳號。

## 第 2 章

---

# Sun Cluster 的主要概念

---

本章說明與 Sun Cluster 系統的硬體和軟體元件相關的主要概念，這些是使用 Sun Cluster 系統之前需要理解的。

本章包含下列小節：

- 第 15 頁的「叢集節點」
- 第 16 頁的「叢集互連」
- 第 16 頁的「叢集成員關係」
- 第 17 頁的「叢集配置儲存庫」
- 第 17 頁的「故障監視器」
- 第 18 頁的「法定裝置」
- 第 20 頁的「Devices」
- 第 21 頁的「資料服務」

---

## 叢集節點

叢集節點是執行 Solaris 軟體與 Sun Cluster 軟體的機器。Sun Cluster 軟體可讓您在一個叢集中有兩個到八個節點。

叢集節點通常連結至一個或多個磁碟。未連結至磁碟的節點使用叢集檔案系統來存取多重主機磁碟。並列資料庫配置中的節點共用對一些或所有磁碟的並行存取。

當另一個節點加入或離開叢集時，叢集中的每個節點都會知道。此外，叢集中的每個節點也都知道正在本機執行的資源，以及在其他叢集節點上執行的資源。

相同叢集中的節點必須有類似的處理程序、記憶體和 I/O 能力，以便啓動故障轉移，而不至於大幅降低效能。由於可能發生故障轉移，因此每個節點都應有足夠的能力在某個節點發生故障時滿足服務層級協定的需要。

---

## 叢集互連

叢集互連是裝置的實體配置，這些裝置用於在叢集節點之間傳輸叢集私有通訊和資料服務通訊。

冗餘互連可在系統管理員隔離故障和修復通訊時，使作業經由其餘正常運作的互連繼續。Sun Cluster 軟體可偵測、修復並自動重新啓動經由已修復互連的通訊。

如需更多資訊，請參閱第 31 頁的「叢集互連」。

---

## 叢集成員關係

叢集成員關係監視器 (CMM) 是一組分散式代理程式，透過叢集互連交換訊息，以完成以下作業：

- 強制在所有節點 (法定數目) 上提供一致的成員關係檢視
- 驅動同步化重新配置以回應成員關係變更
- 處理叢集分割
- 透過將運作狀態不正常的節點留在叢集外直至其被修復，以確保所有叢集成員間的完整連線

CMM 的主要功能是建立叢集成員關係，這需要在隨時加入叢集的節點之間達成叢集範圍內的協議。CMM 將偵測每個節點上的主要叢集狀態變更，例如一個或多個節點間通訊中斷。CMM 依賴傳輸核心模組產生傳輸到叢集中其他節點的傳輸媒體所需的心跳訊號。當 CMM 在定義的逾時期間內未偵測到來自某個節點的心跳訊號時，則 CMM 將認為該節點已失敗，並且將啓動叢集重新配置以重新協商叢集成員關係。

為確定叢集成員關係並確保資料完整性，CMM 將執行以下作業：

- 記錄叢集成員關係的變更，例如節點加入或離開叢集
- 確保運作狀態不正常的節點離開叢集
- 確保運作狀態不正常的節點保持非作用中直至其被修復
- 防止叢集自行分割成節點子集

請參閱第 18 頁的「資料完整性」，以取得有關叢集如何保護自己以免於分割成多個獨立叢集的更多資訊。

---

## 叢集配置儲存庫

叢集配置儲存庫 (CCR) 是一個私有、叢集範圍的分散式資料庫，用來儲存專屬於叢集配置與狀態的資訊。為避免毀壞配置資料，每個節點均必須知道叢集資源的目前狀態。CCR 可確保所有節點均具有一致的叢集檢視。在發生錯誤或恢復情況或叢集的一般狀態變更時，CCR 將被更新。

CCR 結構包含以下類型的資訊：

- 叢集和節點名稱
- 叢集傳輸配置
- Solaris Volume Manager 磁碟集或 VERITAS 磁碟群組的名稱
- 可控制每個磁碟群組的節點清單
- 資料服務的可作業參數值
- 資料服務回呼方法的路徑
- DID 裝置配置
- 目前叢集狀態

---

## 故障監視器

Sun Cluster 系統透過監視應用程式本身、檔案系統以及網路介面，使使用者和資料間「路徑」上的所有元件具有高度可用性。

Sun Cluster 軟體可迅速偵測到節點故障，並為故障節點上的資源建立等效伺服器。Sun Cluster 軟體可確保未受故障節點影響的資源在恢復期間持續可用，並確保故障節點的資源一經恢復立即可用。

## 資料服務監視

每個 Sun Cluster 資料服務均提供一個故障監視器，可定期探測資料服務以確定其運作狀態。故障監視器可驗證應用程式常駐程式是否正在執行以及用戶端是否正在被服務。根據探測傳回的資訊，可以啟動預先定義的動作，例如重新啟動常駐程式或進行故障轉移。

## 磁碟路徑監視

Sun Cluster 軟體支援磁碟路徑監視 (DPM)。DPM 透過報告輔助磁碟路徑的故障來改善故障轉移和切換保護轉移的整體可靠性。您可以使用以下兩種方法之一監視磁碟路徑。第一種方法由 `scdpm` 指令提供。該指令可使您監視、取消監視或顯示叢集內磁碟路徑的狀態。請參閱 `scdpm(1M)` 線上援助頁，以取得有關指令行選項的詳細資訊。

監視叢集內磁碟路徑的第二種方法由 SunPlex Manager 圖形化使用者介面 (GUI) 提供。SunPlex Manager 提供受監視磁碟路徑的拓撲檢視。此檢視每 10 分鐘更新一次，以提供關於失敗偵測的數目。

## IP 多重路徑監視

每一個叢集節點均擁有自己的 IP 網路多重路徑連結配置，該配置可能與其他叢集節點上的此種配置不同。IP 網路多重路徑連結監視以下網路通訊故障：

- 網路配接卡的傳送和接收路徑已停止傳送資料封包。
- 網路配接卡的連結斷開。
- 交換器上的埠無法傳送和接收資料封包。
- 系統啟動時未出現群組中的實體介面。

---

## 法定裝置

法定裝置是由兩個或多個節點共用的磁碟，這些節點所投票數用於建立執行叢集的法定數目。只有具有可用的法定票數，叢集方能運作。法定裝置的使用時機，是在叢集劃分為個別的節點集，以便建立由哪一個節點集投票給新的叢集時。

叢集節點與法定裝置會投票以形成法定數目。依預設，當叢集節點啟動和成為叢集成員時，叢集節點會獲得一票的法定票數。當安裝節點或管理員將節點置於維護狀態時，節點可能會有零票。

法定裝置根據節點與裝置的連線數來獲得法定票數。當設定法定裝置時，它會獲得最大票數  $N-1$ ，其中， $N$  是由與法定裝置連線確定的票數。例如，與兩個有非零票數節點連線的法定裝置，擁有一票法定票數 (二減一)。

## 資料完整性

Sun Cluster 系統嘗試防止資料毀壞並確保資料完整性。由於叢集節點共用資料與資源，叢集嚴禁分割為同時處於作用中的獨立分割區。CMM 保證任何時候均只有一個叢集可以作業。

叢集分割區可能會導致兩種問題：**Split Brain** 與 **Amnesia**。**Split Brain** 發生於節點間的叢集互連遺失以及叢集分割為子叢集時，每個子叢集均相信自己是唯一的分割區。不知道有其他子叢集的子叢集可能導致共用資源衝突，例如網路位址重複和資料毀壞。

如果所有節點都認為叢集處於不穩定的狀態，就會發生 **Amnesia**。例如，有一個具有節點 A 和節點 B 的雙節點叢集。如果節點 A 失敗，將僅在節點 B 上更新 CCR 中的配置資料，而不在節點 A 上更新。如果稍後節點 B 失敗並且節點 A 重新啟動，則節點 A 將使用舊的 CCR 內容執行。這種狀態稱為 **Amnesia**，並可能導致使用舊的配置資訊執行叢集。

藉由投每個節點一票，並強制給予作業中的叢集多數票，便得以避免 Split Brain 與 Amnesia 的狀況發生。具有多數票的分割區有法定數目，並可以作業。如果叢集中有兩個以上的節點，則這種多數票機制會運作得極好。在兩個節點的叢集中，票數為兩票。如果這樣的叢集被分割，外部投票可以使分割區取得法定數目。此外投票由法定裝置提供。法定裝置可以是兩個節點間共用的任何磁碟。

表 2-1 說明 Sun Cluster 軟體如何使用法定數目避免 Split Brain 與 Amnesia 的狀況發生。

表 2-1 叢集法定數目以及 Split-Brain 與 Amnesia 問題

| 分割區類型       | 法定數目解決方案   |
|-------------|--|
| Split Brain | 只允許具有多數票的分割區 (子叢集) 作為叢集來執行 (僅可存在一個這樣具有多數票的分割區)。某個節點在法定數目的競爭中失敗後便會當機。 |
| Amnesia     | 保證在啟動叢集時，該叢集至少有一個節點是最新叢集成員關係中的成員之一 (因此具有最新的配置資料)。                    |

## 故障隔離

叢集的主要問題是導致叢集被分割的故障 (稱為 *Split Brain*)。發生此情形時，不是所有的節點均可通訊，所以個別節點或節點子集可能會嘗試形成個別或子集叢集。每個子集或分割區可能「相信」自己擁有對多重主機磁碟的唯一存取權和所有權。多個節點嘗試寫入磁碟可能導致毀壞資料。

故障隔離藉由防止存取磁碟，來限制節點存取多重主機磁碟。當節點離開叢集 (叢集失敗或被分割) 時，故障隔離可確保節點無法再存取磁碟。僅目前的成員可以存取磁碟，因此確保了資料完整性。

Sun Cluster 系統使用 SCSI 磁碟保留來實現故障隔離。使用 SCSI 保留，便可以將故障節點與多重主機磁碟相「隔離」，防止它們存取這些磁碟。

當叢集成員偵測到另一個節點已經不再經由叢集互連進行通訊，即會啟動故障隔離程序來防止故障節點存取共用磁碟。發生故障隔離時，隔離的節點便會當機，並且其主控台上會顯示「保留衝突」訊息。

## 用於故障隔離的 Failfast 機制

Failfast 機制會使故障節點當機，但它不阻止故障節點重新啟動。發生當機之後，該節點可能會重新啟動並嘗試重新加入叢集。

如果節點在叢集中失去與其他節點的連接，並且也不是可達法定容量的分割區，它會被其他節點強制從叢集中移除。另一可達法定容量之分割區部分的節點，將共用磁碟保留。作為 Failfast 機制的結果，無法定容量的節點將當機。

---

## Devices

全域檔案系統使整個叢集中所有檔案對於所有節點來說均是可存取與可視的。與此類似，Sun Cluster 軟體使叢集上的所有裝置在整個叢集中可存取與可視。即 I/O 子系統允許從任何節點存取叢集中的任何裝置，而不管裝置實體連結位置。該存取是指全域裝置存取。

### 全域裝置

Sun Cluster 系統使用全域裝置來提供從任何節點對叢集中任何裝置的叢集範圍、具有高度可用性的存取。通常，如果節點在提供對全域裝置的存取時失敗，Sun Cluster 軟體會切換至該裝置的其他路徑，並將存取重新導向至該路徑。針對全域裝置，重新導向很簡單，因為對裝置使用了同一名稱，而不管路徑為何。對遠端裝置的存取與對使用同一名稱的本機裝置的存取方法相同。此外，存取叢集上全域裝置的 API 與用於存取本機裝置的 API 相同。

Sun Cluster 全域裝置包含磁碟、CD-ROM 與磁帶。但是，磁碟是唯一受支援的多埠式全域裝置。該有限支援表示 CD-ROM 和磁帶裝置目前不是高可用性裝置。每部伺服器上的本機磁碟亦不是多埠式，因此不是高可用性裝置。

叢集可以為叢集中的每個磁碟、CD-ROM 和磁帶裝置指定唯一的 ID。這種指定可讓叢集中的任何節點對各個裝置進行一致存取。

### 裝置 ID

Sun Cluster 軟體藉由稱為裝置 ID (DID) 驅動程式的建構來管理全域裝置。此驅動程式用於將唯一的 ID 自動指定給叢集中的每個裝置，包括多重主機磁碟、磁帶機和 CD-ROM。

DID 驅動程式是叢集中全域裝置存取功能的主要部分。DID 驅動程式將探測叢集的所有節點，並建立唯一磁碟裝置清單。DID 驅動程式還將為每個裝置指定唯一的主要和次要編號，該編號在叢集的所有節點上一致。全域裝置的存取透過由 DID 驅動程式指定的唯一 DID 來執行，而不是透過傳統的 Solaris DID 來執行。

這種方法可確保存取磁碟的任何應用程式 (例如 Solaris Volume Manager 或 Sun Java System Directory Server) 在叢集中使用一致的路徑。這種一致性對於多重主機磁碟尤其重要，因為每個裝置的本機主要和次要編號會隨著節點的不同而改變。這些編號也可以變更 Solaris 裝置命名慣例。

## 本機裝置

Sun Cluster 還可管理本機裝置。僅可於正在執行服務並與叢集有實體連線的節點上存取這些裝置。本機裝置的效能高於全域裝置，因為本機裝置無須在多個節點上同時複製狀態資訊。除非裝置可由多個節點共用，否則裝置的網域故障將取消對裝置的存取。

## 磁碟裝置群組

磁碟裝置群組使容體管理程式磁碟群組變為「全域」，因為它為基礎磁碟提供多重路徑和多重主機支援。實體連接到多重主機磁碟的每一個叢集節點，均提供了一個磁碟裝置群組的路徑。

在 Sun Cluster 系統中，多重主機磁碟可藉由註冊為磁碟裝置群組受控於 Sun Cluster 軟體。這種註冊為 Sun Cluster 系統提供有關各個節點具有到達哪些容體管理程式磁碟群組的路徑之資訊。Sun Cluster 軟體可為叢集中的每一個磁碟與磁帶裝置建立原始磁碟裝置群組。這些叢集裝置群組保持離線狀態，直至您透過掛載全域檔案系統或存取原始資料庫檔案將它們作為全域裝置存取。

---

## 資料服務

資料服務是軟體和配置檔案的組合，它可使應用程式無需修改即可以 Sun Cluster 配置執行。以 Sun Cluster 配置執行時，應用程式作為受控於群組管理程式 (Resource Group Manager, RGM) 的資源執行。資料服務使您可以將應用程式 (例如 Sun Java System Web Server 或 Oracle 資料庫) 配置為在叢集而非單一伺服器上執行。

資料服務軟體實現了 Sun Cluster 管理方法，以對應用程式執行以下作業：

- 啟動應用程式
- 停止應用程式
- 監視應用程式中的故障並從這些故障中恢復

資料服務的配置檔案定義資源的屬性，對於 RGM，該資源即代表應用程式。

RGM 控制叢集中故障轉移服務與可延伸資料服務的處理。RGM 負責啟動和停止叢集所選取節點上的資料服務，以回應叢集成員關係變更。RGM 使資料服務應用程式可以利用叢集架構。

RGM 將資料服務作為資源控制。這些實現由 Sun 提供或由使用通用資料服務範本、資料服務開發程式庫 API (DSDL API) 或資源管理 API (RMAPI) 的開發者建立。叢集管理員建立並管理稱為資源群組的容器中的資源。RGM 和管理員動作使資源和資源群組在線上狀態與離線狀態之間切換。

## 資源類型

資源類型是屬性的集合，這些屬性向叢集描述應用程式。該集合包括有關如何在叢集的節點上啟動、停止以及監視應用程式之資訊。資源類型還包括應用程式特定屬性，這些屬性需要定義以在叢集中使用應用程式。Sun Cluster 資料服務有數個預先定義的資源類型。例如，Sun Cluster HA for Oracle 屬於資源類型 SUNW.oracle-server；而Sun Cluster HA for Apache 屬於資源類型 SUNW.apache。

## 資源

資源是在叢集範圍內定義的資源類型的實例。資源類型可使應用程式的多個實例安裝在叢集中。當您初始化資源時，RGM 將為應用程式特定屬性指定值，並且該資源將繼承資源類型層級上的所有屬性。

資料服務可以利用數種類型的資源。應用程式 (例如 Apache Web Server 或 Sun Java System Web Server) 使用它們所依賴的網路位址 (邏輯主機名稱與共用位址)。應用程式和網路資源形成受 RGM 管理的基本單位。

## 資源群組

受 RGM 管理的資源被放入資源群組中以便作為單位管理。資源群組是一組相關或相互依賴的資源。例如，源自 SUNW.LogicalHostname 資源類型的資源可能會與源自 Oracle 資料庫資源類型的資源被放入同一資源群組中。如果在資源群組上啟動了故障轉移或切換保護轉移，則資源群組會作為一個單位來遷移。

## 資料服務類型

資料服務使應用程式具有高度可用性，並且可延伸服務可以防止重要應用程式在叢集內發生任何單一故障後中斷。

配置資料服務時，您必須將資料服務配置為以下類型的資料服務之一：

- 故障轉移資料服務
- 可延伸資料服務
- 並列資料服務

## 故障轉移資料服務

故障轉移是叢集自動將應用程式從出現故障的主要節點轉移至指定的冗餘次要節點的處理程序。故障轉移應用程式具有以下特性：

- 能夠僅在叢集中的一個節點上執行

- 不為叢集所知
- 依賴叢集架構以取得高度可用性

如果錯誤監視器偵測到錯誤，則會嘗試於同一節點重新啓動實例，或於其他節點啓動實例 (故障轉移)，視資料服務的配置方式而定。故障轉移服務使用故障轉移資源群組，它是應用程式實例資源與網路資源 (邏輯主機名稱) 的容器。邏輯主機名稱是 IP 位址，可以在某個節點配置上線，稍後自動在原始節點配置下線，並在其他節點配置上線。

用戶端的服務可能有短暫中斷，並且可能需要在完成故障轉移之後重新連線。但是，用戶端並不會知道到提供服務之實體伺服器中的變更。

## 可延伸資料服務

可延伸資料服務使應用程式實例可以在多個節點上同時執行。可延伸服務使用兩種資源群組。可延伸資源群組包含應用程式資源，故障轉移資源群組包含可延伸服務所依賴的網路資源 (共用位址)。可延伸資源群組可以在多個節點上處於線上狀態，因此多個服務實例可同時執行。放置共用位址的故障轉移資源群組一次只在一個節點上啓動成爲線上。宿主可延伸服務的所有節點均使用相同的共用位址宿主服務。

叢集透過單一網路介面 (全域介面) 接收服務要求。基於由平衡資料流量策略設定的數個預先定義的演譯法之一，這些要求將分配到多個節點。叢集可以使用平衡資料流量策略，來均衡各個節點之間的服务負載。

## 並列應用程式

Sun Cluster 系統透過使用並列資料庫，提供在叢集的所有節點上共用並列執行的應用程式之環境。Sun Cluster Support for Oracle Parallel Server/Real Application Clusters 是一組套裝軟體，安裝後可以使 Oracle Parallel Server/Real Application Clusters 在 Sun Cluster 節點上執行。這種資料服務還使 Sun Cluster Support for Oracle Parallel Server/Real Application Clusters 可透過 Sun Cluster 指令來管理。

並列應用程式已被配置爲在叢集環境中執行，以便該應用程式可由兩個或多個節點同時控制。在 Oracle Parallel Server/Real Application Clusters 環境中，多個 Oracle 實例合作提供對相同共用資料庫的存取。Oracle 用戶端可以使用任一實例存取該資料庫。因此，如果一個或多個實例失敗，用戶端可連線至其餘正常運作的實例並繼續存取資料庫。



## 第 3 章

---

# Sun Cluster 架構

---

Sun Cluster 架構允許將一組系統作為單一大型系統進行部署、管理和檢視。

本章包含下列小節：

- 第 25 頁的「Sun Cluster 硬體環境」
- 第 26 頁的「Sun Cluster 軟體環境」
- 第 29 頁的「可延伸資料服務」
- 第 30 頁的「多重主機磁碟儲存裝置」
- 第 31 頁的「叢集互連」
- 第 31 頁的「IP 網路多重路徑群組」

---

## Sun Cluster 硬體環境

叢集由以下硬體元件組成：

- 具有本機磁碟 (未共用) 的叢集節點，提供叢集的主要運算平台。
- 多重主機儲存裝置，提供在節點間共用的磁碟。
- 配置為全域裝置的可移除媒體，例如磁帶和 CD-ROM。
- 叢集互連，為節點間通訊提供通道。
- 公用網路介面，可讓用戶端系統使用網路介面存取叢集上的資料服務。

圖 3-1 說明這些硬體元件如何配合工作。

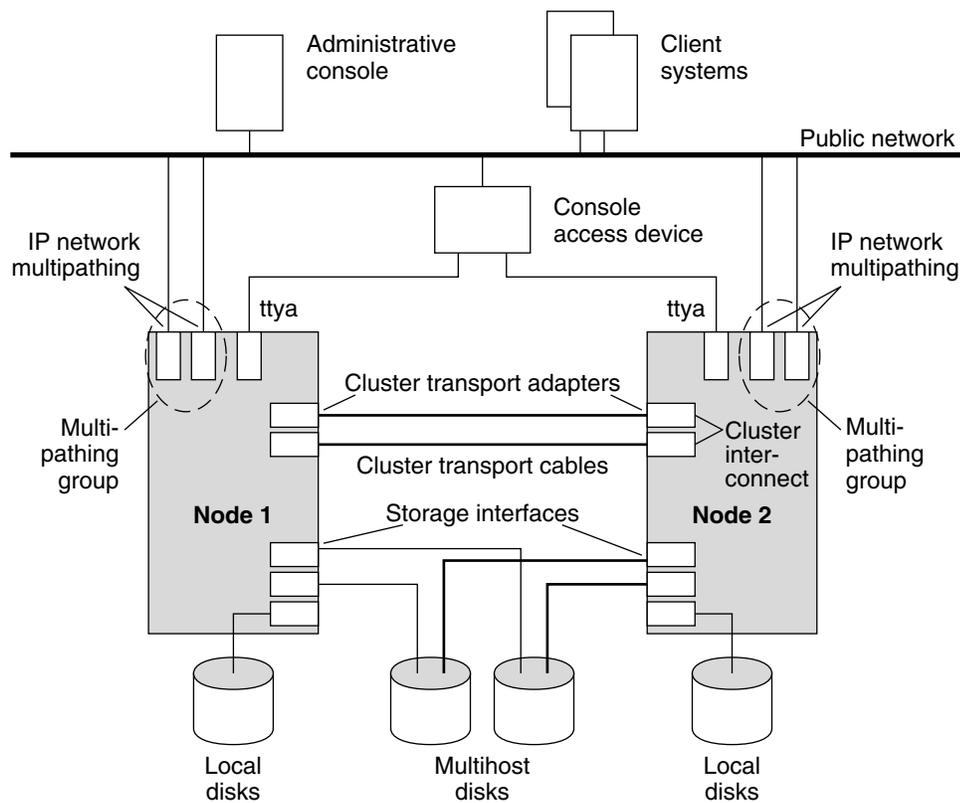


圖 3-1 Sun Cluster 硬體元件

## Sun Cluster 軟體環境

若要作為叢集成員運作，節點必須安裝有以下軟體：

- Solaris 軟體
- Sun Cluster 軟體
- 資料服務應用程式
- 容體管理 (Solaris™ Volume Manager 或 VERITAS Volume Manager)
  - 一個例外的情形是使用方塊中的容體管理的配置。該配置可能不需要軟體容體管理程式。

圖 3-2 顯示軟體元件的高階檢視，這些軟體元件共同作業以建立 Sun Cluster 軟體環境。

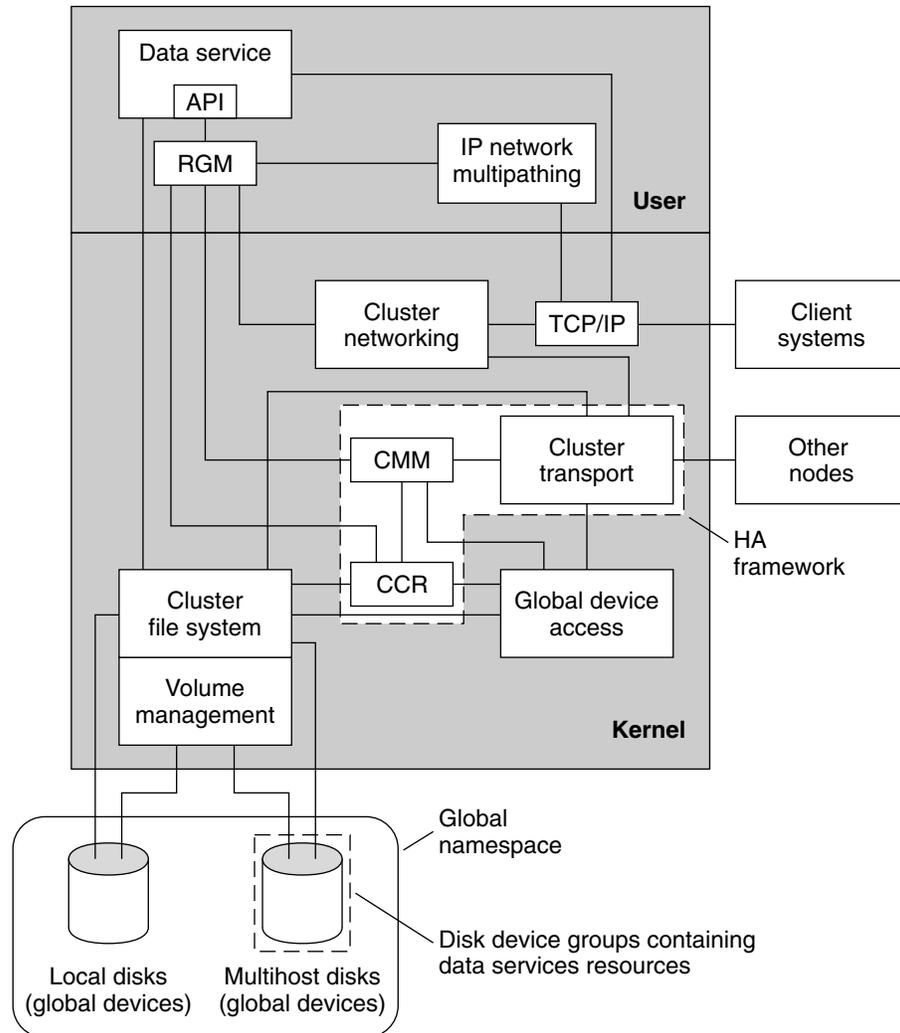


圖 3-2 Sun Cluster 軟體架構

## 叢集成員關係監視器

為了確保資料免於毀壞，所有節點必須針對叢集成員關係達成一致協議。必要時，CMM 會協調叢集服務的叢集重新配置以回應故障。

CMM 從叢集傳輸層接收有關連接到其他節點的資訊。在重新配置期間，CMM 使用叢集交互連接來交換狀態資訊。

偵測到叢集成員關係變更後，CMM 會執行叢集的同步化配置。在此配置中，叢集資源可能會根據新的叢集成員關係重新分配。

CMM 完全在核心中執行。

## 叢集配置儲存庫 (CCR)

CCR 依賴 CMM 來保證叢集只有在到達法定數目時才能執行。CCR 負責驗證整個叢集的資料一致性，必要時執行復原，以及促使資料的更新。

## 叢集檔案系統

叢集檔案系統是以下兩者間的代理：

- 某個節點上的核心程式與基礎檔案系統
- 在與磁碟實體連線的節點上執行的容體管理程式

叢集檔案系統依賴全域裝置 (磁碟、磁帶、CD-ROM)。全域裝置可以從叢集中的任何節點，使用相同檔案名稱 (例如 /dev/global/) 來存取。該節點無需有與儲存裝置的實體連線。您可以像使用一般裝置一樣使用全域裝置，亦即，您可以使用 `newfs` 或 `mkfs` 在全域裝置上建立檔案系統。

叢集檔案系統具備下述功能：

- 檔案存取位置是透明的。程序可以開啓位於系統中任何位置的檔案。此外，所有節點上的程序可以使用相同路徑名稱找到檔案。

---

**注意** – 當叢集檔案系統讀取檔案時，並不會更新這些檔案上的存取時間。

---

- 使用一致的通訊協定來保持 UNIX 檔案存取語意，即使檔案是從多個節點並行地被存取。
- 廣泛的快取與 zero-copy bulk I/O 移動配合使用，使檔案資料的移動更有效率。
- 叢集檔案系統藉由使用 `fcntl(2)` 介面來提供高度可用的建議檔案鎖定功能。藉由使用叢集檔案系統檔案上的建議檔案鎖定功能，在多個叢集節點上執行的應用程式可以同步化資料的存取。檔案鎖可立即由離開叢集的節點，以及維持鎖定時故障的應用程式加以回復。
- 即使發生故障時，仍可確保資料的持續存取。只要磁碟的路徑仍然是作業中，應用程式不會受到故障的影響。這項保證適用於原始磁碟存取和所有的檔案系統作業。
- 叢集檔案系統獨立於基礎檔案系統及容體管理軟體。叢集檔案系統可使任何受支援的磁碟檔案系統成爲全域檔案系統。

## 可延伸資料服務

叢集網路的主要目標是為資料服務提供可延伸性。可延伸性意味著，當服務的負載增加時，資料服務可以在一定時間內回應此增加的工作負荷量，以使新的節點可以加入叢集並且新的伺服器實例可以執行。可延伸資料服務的典型範例是網際網路服務。通常，可延伸資料服務是由許多實例所組成，每一個實例執行於叢集的不同節點上。這些實例一起對於該服務的遠端用戶端來說為單一服務，並實現該服務的功能。在不同節點上執行多個 httpd 常駐程式的可延伸 Web 服務，可以讓任何常駐程式服務用戶端要求。為要求提供服務的常駐程式所依據的便是**平衡資料流量策略**。對用戶端的回覆看上去是來自服務，而不是服務該要求的特定常駐程式，因此保留了單一服務的外觀。

下圖說明了可延伸服務的架構。

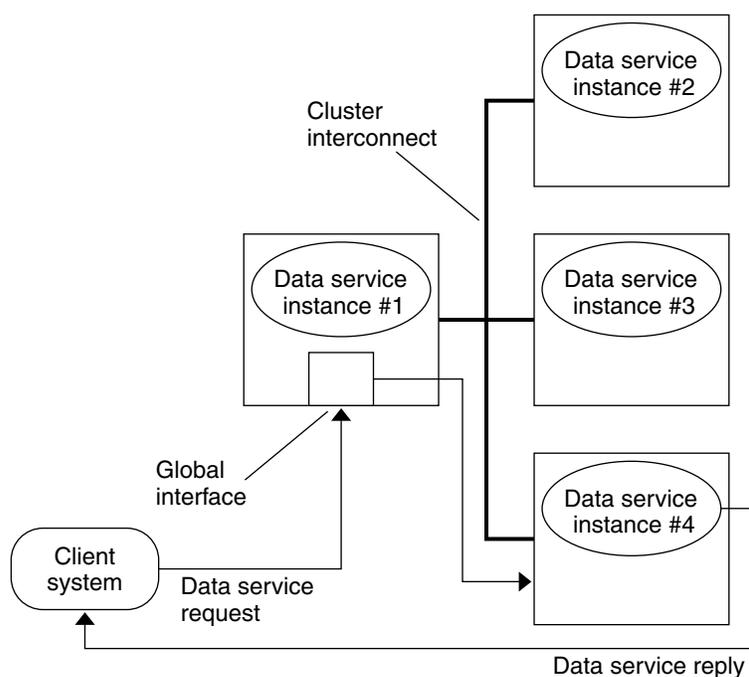


圖 3-3 可延伸資料服務的架構

沒有放置整體介面的節點 (代理節點) 將共用位址放在其回送介面上。進入全域介面的資料封包會根據可配置平衡資料流量策略分配至其他叢集節點。可能的平衡資料流量策略說明如後。

## 平衡資料流量策略

平衡資料流量可以在回應時間及產量上增進可延伸服務的效能。

可延伸資料服務的類別有兩種：*Pure* 與 *Sticky*。*Pure* 服務是任何實例均可回應用戶端要求的服務。*Sticky* 服務使叢集可以平衡節點要求的資料流量。那些要求不會重新導向至其他實例。

*Pure* 服務使用加權平衡資料流量策略。在此平衡資料流量策略下，依預設用戶端要求會平均地分配給叢集中的伺服器實例。例如，在每個節點權重均為 1 的三節點叢集中，每個節點代表該服務分別為來自任何用戶端的要求提供三分之一的服務。透過 `scrgadm (1M)` 指令介面或 `SunPlex Manager GUI` 可隨時變更權重。

*Sticky* 服務有兩種類型：*Ordinary Sticky* 與 *Wildcard Sticky*。*Sticky* 服務允許在多個 TCP 連接上並行處理應用程式層次階段作業，以共用 `in-state` 記憶體 (應用程式階段作業狀態)。

*Ordinary Sticky* 服務允許用戶端共用多個並行 TCP 連接之間的狀態。用戶端被稱為「sticky」是因為該伺服器實例偵聽單一埠。只要該實例維持啟動與可存取的状态，且當此服務處於線上狀態時，平衡資料流量策略未曾改變，即可保證用戶端的所有要求均會到達相同的伺服器實例。

*Wildcard Sticky* 服務使用動態指定的通訊埠編號，但是仍然希望用戶端要求會到達相同的節點。用戶端在同一 IP 位址的埠上呈現「*Sticky Wildcard*」。

---

## 多重主機磁碟儲存裝置

`Sun Cluster` 軟體透過利用多重主機磁碟儲存裝置 (一次可連線至多個節點) 使磁碟具有高度可用性。容體管理軟體可用於將這些磁碟分佈至由叢集節點控制的共用儲存裝置。然後，配置這些磁碟以在發生故障時移至其他節點。在 `Sun Cluster` 系統中使用多重主機磁碟可提供多種優勢，包括：

- 對檔案系統的全域存取
- 對檔案系統和資料的多個存取路徑
- 容許單一節點故障

---

## 叢集互連

必須使用叢集互連 (透過至少兩個實體獨立的冗餘網路或路徑) 來連線所有節點，以避免單一故障點。冗餘需要兩個互連，但最多可使用六個互連分散流量，以避免通訊阻塞並改善冗餘功能和可延伸性。Sun Cluster 互連使用 Fast Ethernet、Gigabit-Ethernet、Sun Fire Link 或 Scalable Coherent Interface (SCI, IEEE 1596-1992)，以進行高效能且叢集私有的通訊。

在叢集環境中，用於節點間通訊的高速、低延時互連及協定是必要的。Sun Cluster 系統中的 SCI 互連經由典型網路介面卡 (NIC) 提供改善的效能。Sun Cluster 將遠端共用記憶體 (RSM™) 介面用於整個 Sun Fire Link 網路中節點間的通訊。RSM 是一個 Sun 訊息傳送介面，對於遠端記憶體作業效率很高。

叢集互連由以下硬體元件組成：

- **配接卡** – 位於每個叢集節點的網路介面卡。具有多重介面的網路配接卡在整個配接卡出現故障時會成為單一故障點。
- **接點** – 位於叢集節點之外的交換器。接點執行傳輸和交換功能，讓您連線兩個以上的節點。在雙節點叢集中，您不需要接點，因為透過冗餘實體電纜可直接將節點相互連線。那些冗餘電纜均連線至每個節點上的冗餘配接卡。多於兩個節點的配置需要接點。
- **電纜** – 位於兩個網路配接卡之間或配接卡與接點之間的實體連線。

圖 3-4 顯示如何連線三個元件。

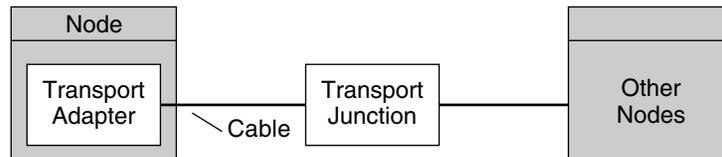


圖 3-4 叢集交互連接

---

## IP 網路多重路徑群組

可將公用網路配接卡組織到 IP 多重路徑群組 (多重路徑群組) 中。每個多重路徑群組均有一個或多個公用網路配接卡。多重路徑群組中的每個配接卡都可以處於作用中，或者您可以配置備用介面 (除非發生故障轉移，否則處於非作用中)。

多重路徑群組可提供邏輯主機名稱與共用位址資源的基礎。節點上的相同多重路徑群組可以擁有任意數目的邏輯主機名稱或共用位址資源。若要監視叢集節點的公用網路連線，您可建立多重路徑。

如需有關邏輯主機名稱和共用位址資源的更多資訊，請參閱「*Sun Cluster Data Services Planning and Administration Guide for Solaris OS*」。

## 公用網路介面

用戶端透過公用網路介面連接至叢集。每一個網路配接卡可以連接至一或多個公用網路，這要根據配接卡是否有多重硬體介面而定。您可以設定節點來包含已配置的多重公用網路介面卡，以便多個配接卡均處於作用中，並且彼此作為故障轉移的備份。如果一個配接卡失敗，系統將呼叫 Sun Cluster 中的 Solaris 網際網路通訊協定 (IP) 網路多重路徑軟體，以將失敗的介面故障轉移至群組中的其他配接卡。

# 索引

---

## A

amnesia, 18

## F

Failfast, 19

## I

ID, 裝置, 20

IP 網路多重路徑連結, 10, 18, 31

IPMP

參閱IP 網路多重路徑連結

## O

Oracle Parallel Server/Real Application Clusters, 11

## S

scdpm 指令, 17

SCSI, 19

Solaris 容體管理程式, 11

split-brain, 18, 19

Sun Cluster Support for Oracle Parallel Server/Real Application Clusters, 23

Sun Management Center, 13

Sun StorEdge Traffic Manager, 11

SunPlex Manager, 13, 17

## V

VERITAS Volume Manager (VxVM), 11

## 工

工具, 12

## 互

互連, 參閱叢集, 互連

## 介

介面, 18, 31, 32

## 元

元件

軟體, 26

硬體, 25

## 公

公用網路, 參閱網路, 公用

## 冗

### 冗餘

- 硬體, 12
- 磁碟系統, 9

## 分

- 分割, 叢集, 18

## 代

- 代理程式, 參閱資料服務

## 可

- 可用性管理, 10
- 可延伸
  - 服務, 10
  - 資料服務, 23
  - 架構, 29
  - 資源群組, 23
- 可延伸性, 參閱可延伸

## 平

- 平衡資料流量
  - 描述, 29
  - 策略, 30

## 本

- 本機裝置, 21

## 全

- 全域名稱空間, 20
- 全域裝置
  - 掛載, 28
  - 描述, 20
- 磁碟裝置群組, 21

## 共

- 共用位址, 可延伸資料服務, 23
- 共用磁碟群組, 23

## 多

- 多重主機儲存裝置, 11
- 多重路徑, 10, 12, 18, 31

## 存

- 存取控制, 14

## 成

- 成員關係, 15, 16, 27

## 投

- 投票計數, 法定, 18

## 並

- 並列
  - 資料庫, 10, 15
  - 應用程式, 10, 23

## 服

- 服務, 參閱資料服務

## 法

- 法定, 18

## 保

- 保留衝突, 19

## 恢

恢復, 9, 10

## 指

指令行介面 (CLI), 13

## 挂

掛載, 28

## 故

### 故障

偵測, 12  
硬體和軟體, 12  
隔離, 19

### 故障轉移

由 Oracle Parallel Server/Real Application  
Clusters 軟體提供, 23  
服務, 10  
透明, 10  
資料服務, 22

## 容

容錯性, 9  
容體管理, 11, 30

## 配

配接卡, 參閱網路, 配接卡

### 配置

工具, 12  
並列資料庫, 15  
儲存庫, 17, 28

## 高

高度可用性, 9

## 基

基於角色存取控制 (RBAC), 14

## 軟

### 軟體

元件, 26  
故障, 12  
高度可用性, 9  
基於主機的, 12  
備用獨立式磁碟陣 (RAID), 12

## 通

通訊管理程式, 11

## 備

備用獨立式磁碟陣 (RAID), 12

## 硬

### 硬體

Sun StorEdge Traffic Manager, 11  
故障, 12  
高度可用性, 9  
備用獨立式磁碟陣 (RAID), 12  
環境, 25  
叢集互連, 31  
叢集節點, 15

## 當

當機, 19

## 節

節點, 15

## 裝

### 裝置

- ID (DID), 20
- 本機, 21
- 全域, 20
- 法定, 18
- 群組, 21

## 資

資料完整性, 18

### 資料服務

#### 可延伸

- pure, 30
- sticky, 30
- 架構, 29
- 資源, 23

並列, 23

定義, 21

故障監視, 12

故障轉移, 22

資源, 22

資源群組, 22

資源類型, 22

類型, 22

資料服務開發程式庫 API (DSDL API), 21

資料庫, 10

### 資源

共用的, 12

定義, 22

恢復, 10

#### 群組

故障轉移, 22

描述, 22

類型, 22

資源群組管理程式 (RGM)

功能性, 21

資源群組, 與, 22

資源管理 API (RMAPI), 21

## 隔

隔離, 19

## 監

### 監視

故障, 12

磁碟路徑, 17

網路介面, 18

## 磁

### 磁碟

本機, 20

全域裝置, 20

多重主機, 11, 20, 21, 30

法定, 18

故障隔離, 19

裝置群組, 21

管理, 11

鏡像, 11, 12

磁碟路徑監視 (DPM), 17

## 管

管理, 工具, 12

## 網

### 網路

介面, 10, 31

#### 公用

IP 網路多重路徑連結, 10, 18, 31

描述, 32

監視, 12

平衡資料流量, 29, 30

配接卡, 10, 18, 32

網際網路通訊協定 (IP), 23

## 儲

### 儲存

多重主機, 30

陣列, 12

管理, 11

儲存庫, 17, 28

儲存裝置, 多重主機, 11

## 應

應用程式  
請參閱資料服務  
並列, 10, 23  
容錯性, 9  
高度可用性, 9  
監視, 12

## 邏

邏輯主機名稱, 故障轉移資料服務, 22

## 檔

檔案系統  
掛載, 28  
叢集, 12, 28  
檔案鎖定, 28

## 環

環境  
軟體, 26  
硬體, 25

## 叢

叢集  
互連, 16, 31  
公用網路, 32  
分割, 18  
成員, 15, 27  
成員關係, 16  
配置, 17, 28  
通訊, 16  
節點, 15  
檔案系統, 12, 28  
叢集成員關係監視器 (CMM), 16, 27  
叢集配置儲存庫 (CCR), 17, 28

## 驅

驅動程式, 參閱裝置, ID (DID)

