



# Sun Enterprise Server™ Alternate Pathing 2.3.1 用户指南

---

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

部件号: 806-5826-10  
2000年10月, 修订版 A

请将关于此文档的意见发送至: [docfeedback@sun.com](mailto:docfeedback@sun.com)

Copyright 2000 Sun Microsystems, Inc., 901 San Antonio Road, Palo Alto, California 94303-4900 U.S.A. 版权所有。

本产品或文档受版权保护，且按照限制其使用、复制、分发和反编译的许可证进行分发。未经 Sun 及其许可证颁发机构的书面授权，不得以任何方式、任何形式复制本产品或本文档的任何部分。第三方软件，包括字体技术，由 Sun 供应商提供许可和版权。

本产品的某些部分从 Berkeley BSD 系统派生而来，经 University of California 许可授权。UNIX 是在美国和其它国家注册的商标，经 X/Open Company, Ltd. 独家许可授权。下列声明适用于 Netscape Communicator™: (c) Copyright 1995 Netscape Communications Corporation. 版权所有。

Sun、Sun Microsystems、Sun 徽标、AnswerBook2、docs.sun.com、Sun Enterprise、OpenBoot 和 Solaris 是 Sun Microsystems, Inc. 在美国和其它国家的商标、注册商标或服务标记。所有的 SPARC 商标都按许可证使用，是 SPARC International, Inc. 在美国和其它国家的商标或注册商标。带有 SPARC 商标的产品均基于由 Sun Microsystems, Inc. 开发的体系结构。

OPEN LOOK 和 Sun™ 图形用户界面是 Sun Microsystems, Inc. 为其用户和许可证持有人开发的。Sun 承认 Xerox 在为计算机行业研究和开发可视或图形用户界面方面所作出的先行努力。Sun 以非独占方式从 Xerox 获得 Xerox 图形用户界面的许可证，该许可证涵盖实施 OPEN LOOK GUI 且遵守 Sun 的书面许可协议的许可证持有人。

本资料按“现有形式”提供，不承担明确或隐含的条件、陈述和保证，包括对特定目的的商业活动和适用性或非侵害性的任何隐含保证，除非这种不承担责任的声明是不合法的。

Sun Enterprise 10000 SSP 归属权:

本软件版权归 University of California 校务委员、Sun Microsystems, Inc. 和其它方面所有。下列条款适用于与本软件相关的所有文件，除非在单独的文件里明确声明。

著作人允许不限用途免费使用、复制、修改、分发和许可本软件及其文档资料，但必须在所有副本中保留现有的版权声明，并在所有分发品中注明上述版权声明。一经授权使用，不再需要书面协议、许可证或专利权使用费。本软件的修改版权归修改者所有。如果修改者在所适用的每个文件的首页上明确附加了新的条款，所做的修改不受此许可条款的制约。



请回收



Adobe PostScript

# 目录

---

前言	ix
本书结构	ix
阅读本书之前	x
使用 UNIX 命令	x
排印约定	xi
Shell 提示符	xi
订购 Sun 文档资料	xii
相关文档资料	xii
Web 上的 Sun 文档资料	xii
Sun 欢迎您提出宝贵意见	xiii
<b>1. 介绍 Alternate Pathing</b>	<b>1</b>
Alternate Pathing 的用途	1
Alternate Pathing 基本概念	4
物理路径	4
元磁盘	5
Sun StorEdge T3 磁盘上的路径优化	6
元网络	6
磁盘路径组	7
网络路径组	9

支持的软件版本	10
AP 配置示例	10
AP 和域	12
<b>2. Alternate Pathing 数据库</b>	<b>13</b>
管理数据库的副本	13
定位数据库以高效使用 RAS	14
创建和删除数据库	15
▼ 创建 AP 数据库的一个副本	15
▼ 删除 AP 数据库的副本	16
查看数据库信息	16
▼ 查看有关数据库副本的信息	17
查看路径组信息	17
▼ 查看未提交的磁盘项	18
▼ 查看已提交的磁盘项	18
▼ 查看未提交的网络项	19
▼ 查看已提交的网络项	19
<b>3. 使用元磁盘和磁盘路径组</b>	<b>21</b>
元磁盘的设备节点	21
自动切换元磁盘	22
磁盘可用性和性能折衷	24
磁盘镜像注意事项	25
使用磁盘路径组和元磁盘	29
▼ 创建磁盘路径组和元磁盘	29
▼ 从主路径切换到替代路径	32
▼ 切换回主路径	35
▼ 删除磁盘路径组和元磁盘	36
▼ 取消配置元磁盘	38

- ▼ 重新配置元磁盘 38
- 4. 使用 AP 引导设备 39
  - 将引导磁盘置于 AP 控制之下 39
    - ▼ 将引导磁盘置于 AP 控制之下 39
    - ▼ 使镜像引导磁盘具有替代路径 41
    - ▼ 使镜像引导磁盘脱离 AP 控制 42
    - ▼ 使引导磁盘脱离 AP 控制 42
  - AP 引导顺序 43
  - 使用单用户模式 43
- 5. 使用元网络路径组 45
  - 元网络接口 45
  - 使用文件路径组 46
    - ▼ 创建网络路径组和元网络 46
    - ▼ 切换网络路径组 50
    - ▼ 删除网络路径组和元网络 51
    - ▼ 取消配置元网络 51
    - ▼ 重新配置元网络 52
  - 使主网络接口具有替代路径 53
  - 为当前网络配置 AP 53
    - ▼ 为主网络创建网络路径组和元网络 54
    - ▼ 删除主网络的网络路径组和元网络 55
    - ▼ 取消配置主网络的元网络 56
    - ▼ 重新配置主网络的元网络 56
- 6. AP 和 DR 之间的交互作用 59
  - 一起使用 DR 和 AP 59
  - 维护正确的 AP 状态 61

A.	AP 组件	63
B.	AP 手册页	65
C.	驱动程序层	67
	词汇表	69
	索引	71

# 图形

---

- 图 1-1 有替代路径的 I/O 设备 2
- 图 1-2 在 I/O 控制器出现故障后切换路径 3
- 图 1-3 为 DR 分离操作切换路径 4
- 图 1-4 物理路径 5
- 图 1-5 元磁盘示例 6
- 图 1-6 元网络 7
- 图 1-7 磁盘路径组切换 8
- 图 1-8 网络路径组 9
- 图 1-9 典型的 AP 配置 10
- 图 1-10 AP 和磁盘镜像 11
- 图 3-1 系统板和磁盘控制器 25
- 图 3-2 系统板和控制器 26
- 图 3-3 镜像卷示例 1 26
- 图 3-4 镜像卷示例 2 27
- 图 3-5 镜像卷示例 3 28
- 图 C-1 AP 磁盘驱动程序层 67
- 图 C-2 AP 网络驱动程序层 68





# 前言

---

《*Sun Enterprise Server Alternate Pathing 2.3.1 用户指南*》介绍了 Sun Enterprise™ 服务器系列产品中的 Alternate pathing (AP) 组件。某些 AP 功能只适用于 Sun Enterprise 10000 服务器，且已在本手册中注明。

---

## 本书结构

本指南包括以下章节：

第一章介绍 AP。

第二章讲述 AP 数据库操作。

第三章介绍元磁盘和磁盘路径组，并说明其使用方法。

第四章介绍无人职守系统的引导问题。

第五章介绍元网络和网络路径组，并说明其使用方法。

第六章介绍 Dynamic Reconfiguration (DR) 和 AP 是如何一同工作的。

附录 A 提供所有 AP 命令的列表。

附录 B 概述 AP 的基本体系结构。

附录 C 概述 AP 的基本驱动程序。

---

## 阅读本书之前

本手册适用于在 UNIX® 系统，尤其是在基于 Solaris™ 操作环境的系统方面有实践经验的 Sun Enterprise 系统管理员。如果不具备这些知识，请首先阅读随本系统一起提供的 Solaris 用户和系统管理员 AnswerBook2™ 文档资料，并考虑进行 UNIX 系统管理培训。

---

## 使用 UNIX 命令

本文档不包括有关基本 UNIX 命令和过程（例如关闭系统、引导系统和配置设备）的信息。

有关此类信息，请参考以下的一本或多本资料：

- 适用于 Solaris 软件环境（尤其是那些涉及 Solaris 系统管理的软件环境）的 AnswerBook 联机文档资料
- 与系统一起收到的其它软件文档资料

---

## 排印约定

表 P-1 排印约定

字体或符号	含义	示例
AaBbCc123	命令、文件和目录的名称； 计算机屏幕输出。	编辑 .login 文件。 使用 <code>ls -a</code> 列出所有文件。 % You have mail.
<b>AaBbCc123</b>	键入的内容，与计算机屏幕输出相区别。	% <b>su</b> Password:
<i>AaBbCc123</i>	书名、新词或术语以及要强调的词。 命令行变量；将用实际名称或值来替代。	请阅读 《 <i>用户指南</i> 》的第六章。 这些称为类选项。 要执行该操作， <i>必须是 root 用户</i> 。 要删除文件，键入 <code>rm 文件名</code> 。

---

## Shell 提示符

表 P-2 Shell 提示符

Shell	提示符
C shell	<i>machine_name</i> %
C shell 超级用户	<i>machine_name</i> #
Bourne shell 和 Korn shell	\$
Bourne shell 和 Korn shell 超级用户	#

---

## 订购 Sun 文档资料

Fatbrain.com, 一个 Internet 专业书店, 储备了有关 Sun Microsystems, Inc. 的精选产品文档资料。

有关这些文档的列表及其订购方法, 请访问 Fatbrain.com 上的 Sun Documentation Center, 网址为:

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

---

## 相关文档资料

表 P-3 相关文档资料

用途	书名
安装	《Solaris 8 10/00 Sun 硬件平台指南》
参考 (手册页)	《Sun Enterprise Server Alternate Pathing 2.3.1 Reference Manual》
发行说明	《Sun 硬件的 Solaris 8 10/00 发行说明补充资料》
其它文档资料	《Sun Enterprise 10000 Dynamic Reconfiguration User Guide》 《Sun Enterprise 6x00, 5x00, 4x00, 3x00 Dynamic Reconfiguration User Guide》 《Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual》

---

## Web 上的 Sun 文档资料

可通过 docs.sun.com<sup>SM</sup> 网站访问 Web 上的 Sun 技术文档资料。可以在如下网址上浏览 docs.sun.com 存档资料或搜索特定的书名或主题:

<http://docs.sun.com>

---

## Sun 欢迎您提出宝贵意见

我们致力于提高文档资料的质量，欢迎您提出宝贵意见和建议。您可以将意见通过电子邮件发送给我们，地址如下：

`docfeedback@sun.com`

请在电子邮件的主题行中包含文档的部件号。



# 介绍 Alternate Pathing

---

本章介绍了 Alternate Pathing 的基本用途并概述了 Alternate Pathing 的概念和术语。

---

## Alternate Pathing 的用途

Alternate pathing (AP) 支持 I/O 控制器的高可用性—这些控制器是位于系统板上的硬件组件，可使 Sun Enterprise 服务器与 I/O 设备（如磁盘和网络）进行通讯。有了 AP 后，每台 I/O 设备都可连接两个 I/O 控制器。

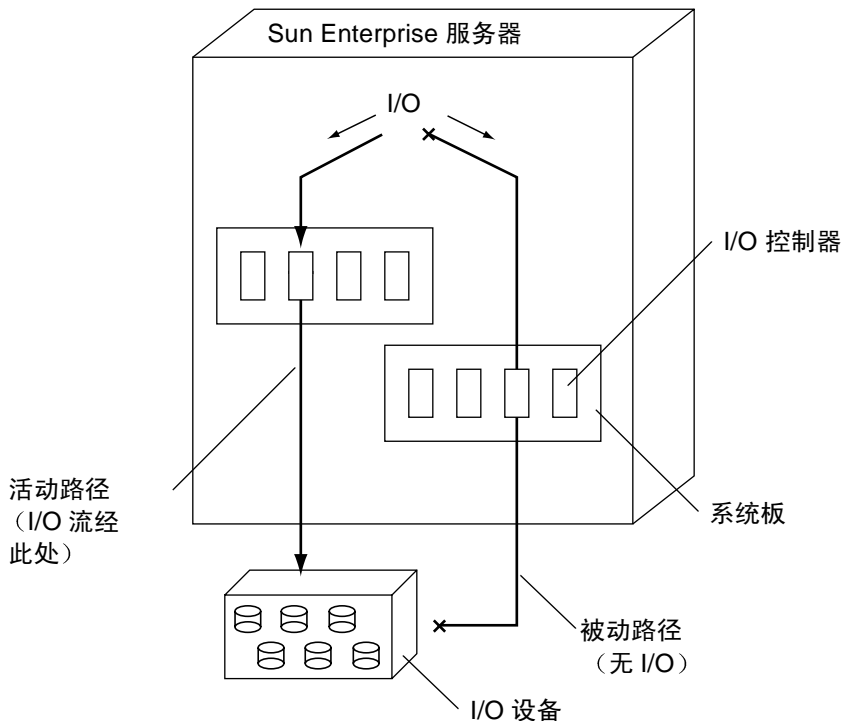


图 1-1 有替代路径的 I/O 设备

I/O 控制器是两条通往 I/O 设备的独立电子路径（互为替代路径）的一部分。AP 使得您可以设置和使用 Sun Enterprise 服务器上的替代路径。

对于磁盘控制器，如果在正常操作过程中检测到路径故障，会 *自动* 进行切换。对于网络控制器，必须手动切换路径（使用一条 AP 命令）。

AP 有两种用途。一种用途是防备 I/O 控制器出现故障。有了 AP 后，如果一个 I/O 控制器失败，可以切换到替代控制器。

对于 Sun StorEdge™ T3 磁盘上的磁盘元设备，提供有两条可行的物理 I/O 设备路径，磁盘路径组应用了路径优化。路径优化是指对特定设备的 I/O 流量的有效分配。如果其中一条物理 I/O 设备路径无法使用，无论是因设备故障还是因用户操作引起的，都将禁用路径优化。



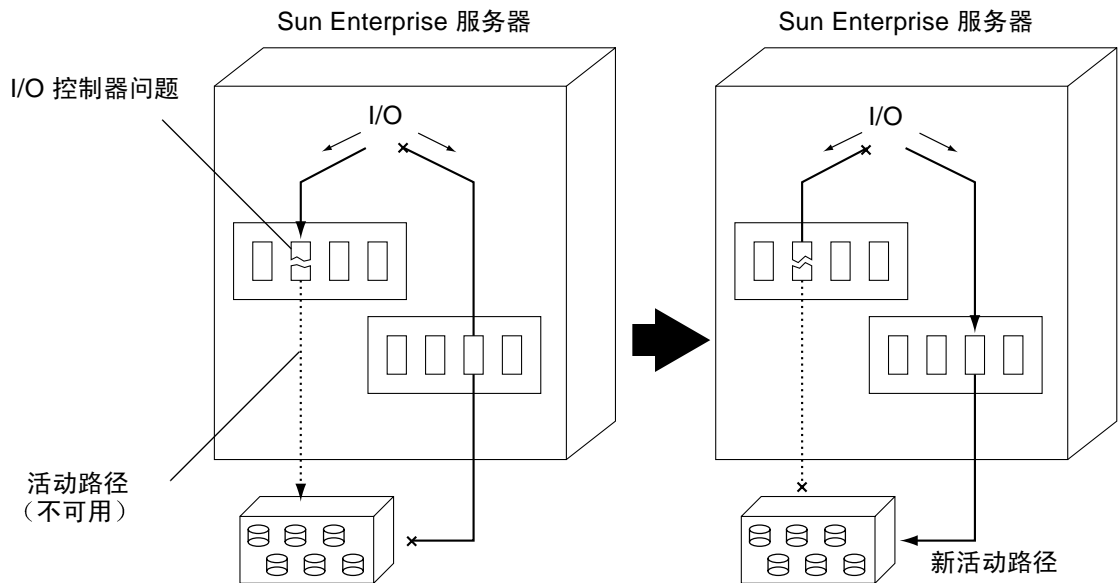


图 1-2 在 I/O 控制器出现故障后切换路径

---

注意 – 因为 T3 上仅有一条路径，因此自动切换功能禁用了路径优化算法。

---

AP 的第二个用途是支持 Dynamic Reconfiguration (DR)。DR 用于从操作系统上对系统板进行逻辑上的连接和分离，而无须中止和重新引导系统。

例如，可以利用 DR 从操作系统上分离系统板，物理拆除和维修该板，然后再把它重新插入并连接到操作系统上。执行这些操作时，无须中止该操作系统或终止任何用户应用程序。

如果想分离连接了 I/O 设备的板，并且该 I/O 设备是有替代路径的，那么必须首先使用 AP 将 I/O 流重定向到其它板上的控制器。

对于 T3 磁盘路径组，此操作将禁用路径优化，以允许非活动路径的 DR 分离。然后使用 DR 分离该系统板，从而无须中断 I/O 流。

在 Sun Enterprise 10000 上，对磁盘和网络设备进行 DR 操作过程中会自动进行切换，并假定另一块板上存在可用的替代控制器。但是，手动切换会在 DR 分离之前禁用路径优化。

在其它所有服务器上，必须手动进行切换。

下图显示了 AP 和 DR 之间的关系。

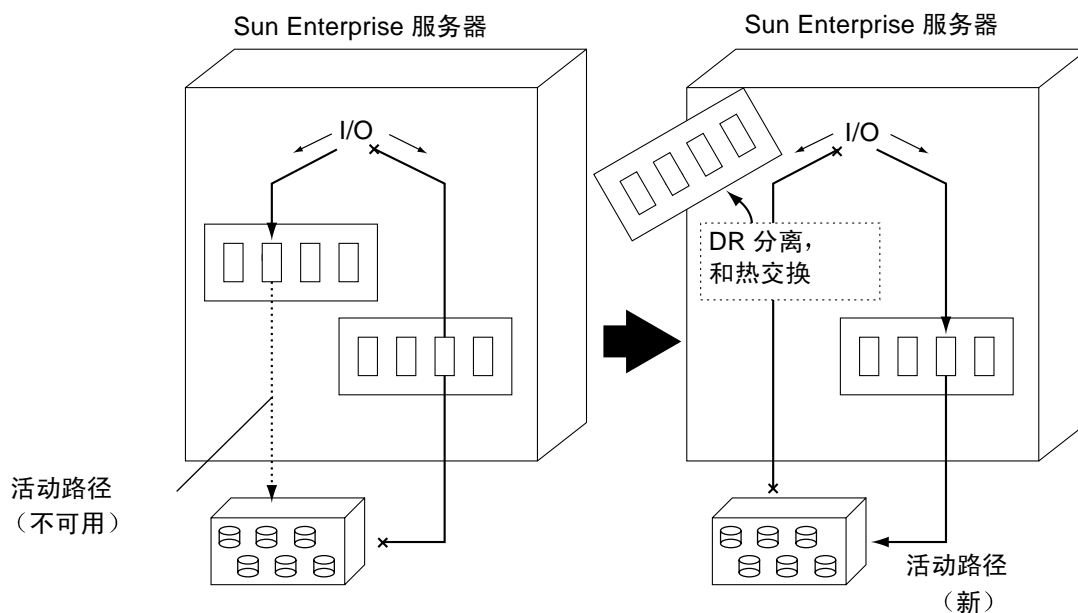


图 1-3 为 DR 分离操作切换路径

## Alternate Pathing 基本概念

本节讨论 AP 的基本概念并介绍本章要使用的术语。

### 物理路径

针对 AP 的用途而言，I/O 设备可以是磁盘也可以是网络。I/O 控制器是 I/O 设备的控制器卡。I/O 端口是控制器卡上的连接器。每块控制器卡上通常有两个或两个以上的端口。设备节点是 /devices 或 /dev 目录下的路径，用于指定物理设备，例如， /dev/dsk/c0t0d1s0。术语物理路径指的是从主机到磁盘或网络的电子路径。

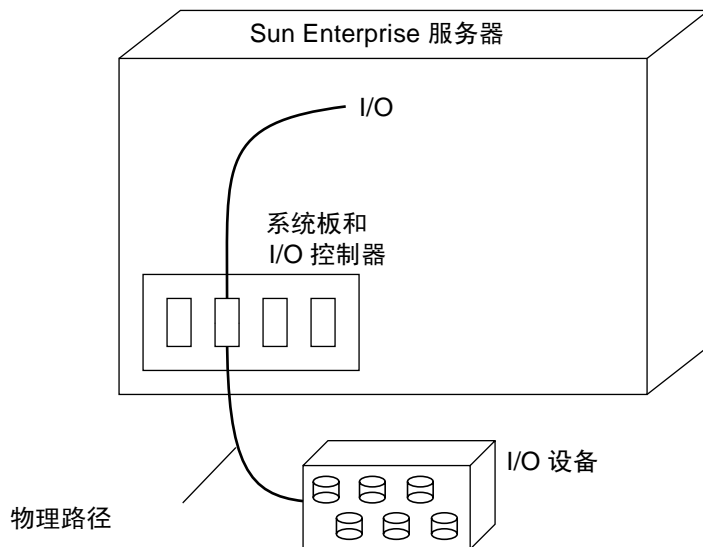


图 1-4 物理路径

可利用设备节点来引用物理路径，例如 `/dev/dsk/c0t1d1s0`。

## 元磁盘

*元磁盘*（如图 1-5 所示）是一种构造，使您能够利用两条物理路径中的任意一条来访问磁盘，而无须在脚本或程序内明确引用任何路径。

在脚本或程序中，可使用 AP 特定的设备节点（例如 `/dev/ap/dsk/mc0t1d1s0`）来引用元磁盘。有关更多信息，请参考第 21 页的“元磁盘的设备节点”。

在下图中，使用 AP 特定的设备节点执行磁盘 I/O，而不考虑当前哪个 `p1n`（`p1n:2` 或 `p1n:9`）端口正在处理 I/O。

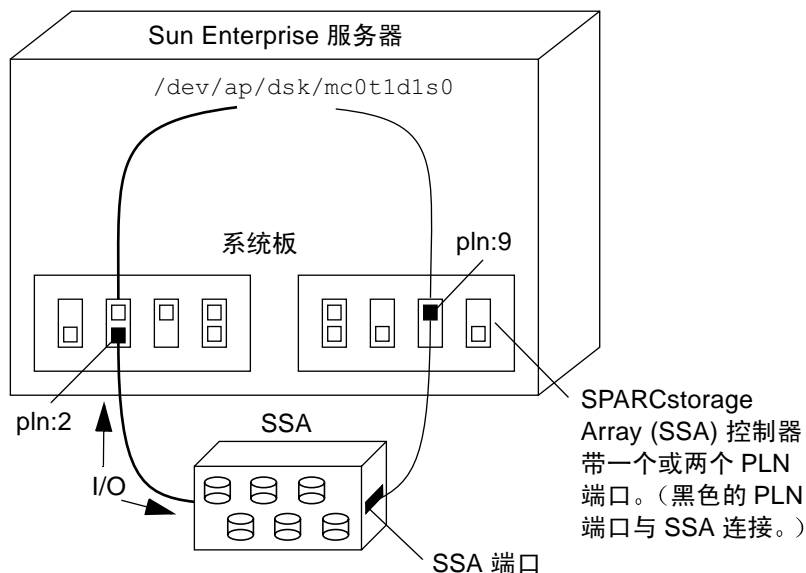


图 1-5 元磁盘示例

## Sun StorEdge T3 磁盘上的路径优化

对 T3 进行系统启动时，只要设备有两条可用的物理路径，就将对磁盘路径组执行路径优化算法。因设备故障或用户操作原因而禁用其中一条物理路径，将关闭受影响的磁盘路径组的路径优化。用户可使用 `apconfig(1M)` 命令或通过重新引导系统，来重新启动路径优化算法。除非两条物理 I/O 路径都可用，否则不能重新启用路径优化。有关详细信息，请参见第 29 页的“使用磁盘路径组和元磁盘”。

## 元网络

元网络（如图 1-6 所示）是一种构造，使您能够利用两条物理路径中的任意一条来访问网络，而无须在脚本或程序内明确引用任何路径。在脚本或程序中，可使用元网络接口名称（如 `mether1`）来引用元网络。有关更多信息，请参考第 45 页的“元网络接口”。

在下图中，使用 `mether1` 访问元网络，而不考虑当前哪个控制器（`hme1` 或 `qfe3`）正在处理元网络 I/O。

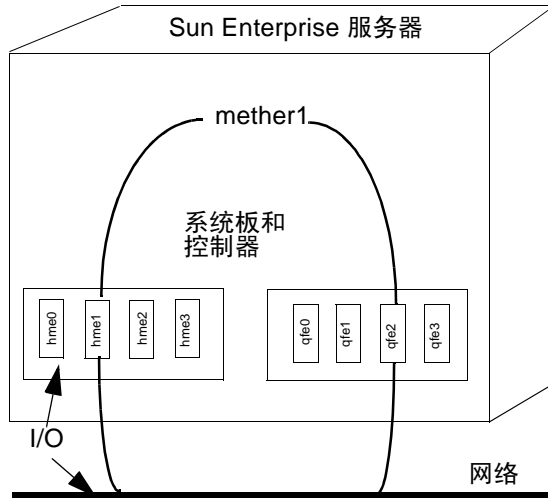


图 1-6 元网络

## 磁盘路径组

磁盘路径组 (如图 1-7 所示) 含有两条指向同一磁盘阵列的物理路径。当某条物理路径是路径组的一部分时, 就被称做 *替代路径*。指向磁盘的替代路径可以用其使用的 `pln` 端口或 `sf` 端口唯一地标识出来。当前正在处理 I/O 的替代路径称为 *活动替代路径*。

---

**注意** – 当 T3 磁盘路径组正在执行路径优化算法时, 两条物理路径都被标记为 *活动替代路径*。无论因何种原因丢失物理路径, 都将禁用路径优化。在此情况下, 只有一条路径被标记为 *活动替代路径*。

---

注意, 元磁盘 (例如, `/dev/ap/[r]dsk/mc?t?d?s?`) 为您提供了一种在脚本或程序中访问单个磁盘的方法, 而磁盘路径组则提供了一种在您运行 AP 命令时控制该磁盘路径的方法。例如, 若要执行切换操作, 即将活动替代路径从一条替代路径切换至另一条, 可以用 `apconfig(1M)` 命令引用磁盘路径组。

---

注意 – 启动切换操作将自动禁用 T3 上的路径优化。

---

将其中某一条替代路径指定为 *主路径*。尽管在执行切换操作时更改了活动替代路径，但主路径依然保持不变。可通过指定与主路径对应的 `p1n` 端口（例如 `p1n:2`）或 `sf` 端口（例如 `sf:2`）来引用磁盘路径组。有关确定 `p1n` 或 `sf` 端口名称的信息，请参考第 21 页的“元磁盘的设备节点”。

要切换磁盘路径组的活动替代路径，可使用如下命令：

```
# apconfig -P sf:2 -a sf:9
```

下图是一个示例，显示了使用 `apconfig(1M)` 命令切换磁盘路径组活动替代路径的结果。

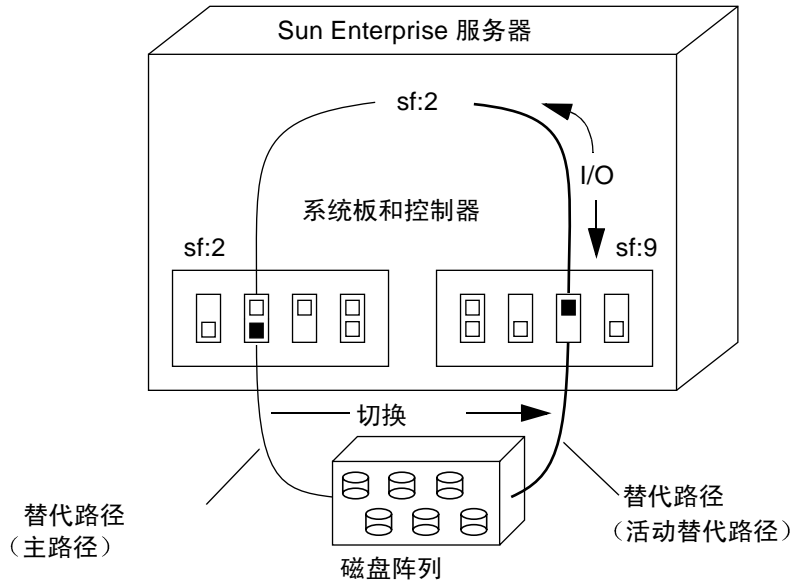


图 1-7 磁盘路径组切换

注意 – 此操作将禁用 T3 上的路径优化。要重新启用路径优化，使用：

```
# apconfig -P sf:2 -a sf:2 -a sf:9
```

## 网络路径组

网络路径组（如图 1-8 所示）含有两个连接同一物理网络的网络控制器。术语 *替代路径*、*活动替代路径* 和 *切换* 基本上与磁盘路径组中这些术语的含义相同。但是，网络路径组没有 *主路径*。

要指定网络路径组，可引用相应的元网络接口名称，例如 `mether1`。第 45 页的“元网络接口”中介绍了元网络接口名称。要切换网络路径组的活动替代路径，可使用如下命令：

```
#apconfig -a mether1 -a hme1
```

例如，图 1-8 显示了使用 `apconfig(1M)` 命令切换网络路径组活动替代路径的结果。

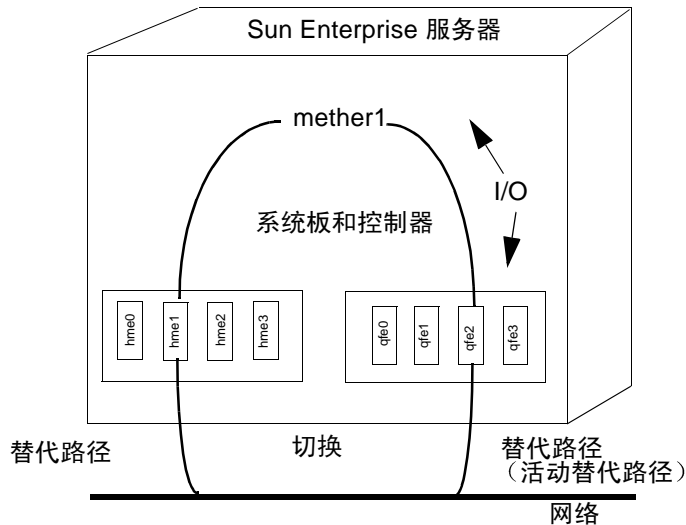


图 1-8 网络路径组

---

## 支持的软件版本

AP 2.3.1 支持 Solaris 2.6、Solaris 7 和 Solaris 8 操作环境。

AP 支持的磁盘网络设备和第三方软件产品列在 《Sun 硬件的 Solaris 8 10/00 发行说明补充资料》中。

如果已为磁盘创建替代路径，并且这些磁盘也使用了卷管理器，那么卷管理器只能使用它们的 AP 元磁盘名识别这些磁盘。这个要求使得 AP 能够在不干扰卷管理器的情况下切换主动路径。

您可以将引导磁盘和主网络接口设置为受控于 AP。这样，即使无法访问主网络或引导磁盘控制器，只要为这些设备定义了可用的替代路径，就可利用 AP 对系统进行无人引导。

---

## AP 配置示例

图 1-9 显示如何使用 AP 来支持以太网络和磁盘阵列。

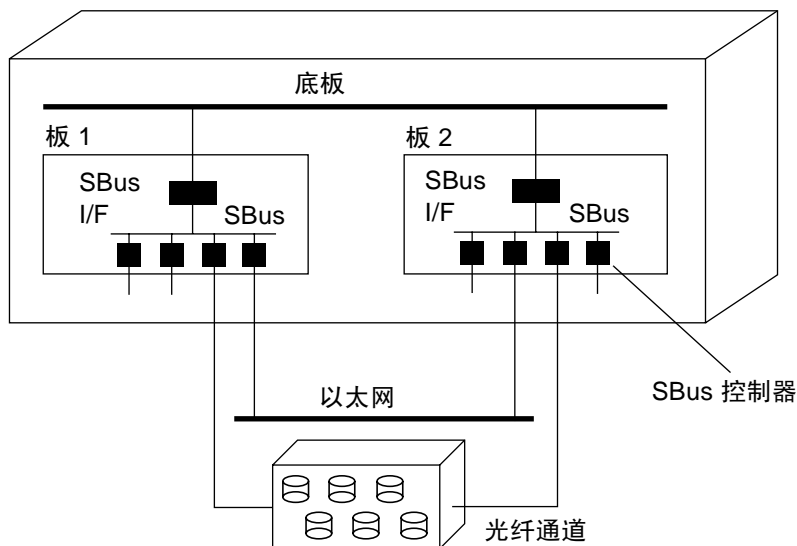


图 1-9 典型的 AP 配置



此例中，两个网络控制器（分别位于系统板 1 和系统板 2）连接在同一网络上。同样，两块板上的两个 SSA 控制器也与同一个 SSA 连接。在这种情况下，如果通过 DR 分离操作分离了板 1，AP 可以在不干扰任何可能正在进行的 I/O 操作的情况下切换到板 2。

AP 不等同于磁盘镜像。磁盘镜像主要是获得数据冗余，尽管有两条可用路径，但镜像的各方都只使用一条路径。

而 AP 则为镜像的各方建立两条可用路径，从而达到真正的路径冗余。要同时使用 AP 和磁盘镜像，必须配置卷管理器软件（例如 Sun Enterprise Volume Manager™ (SEVM)），以便它可以使用 AP 元磁盘路径。

下图显示了如何将 AP 与磁盘镜像结合使用的示例。

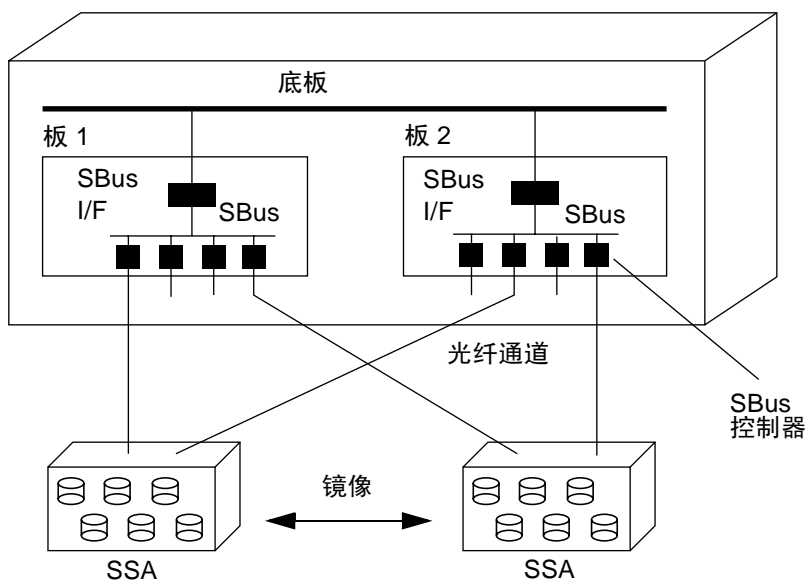


图 1-10 AP 和磁盘镜像

这种配置类型使得您可以将用于进行镜像的路径从一块板切换到另一块板，而不会中断磁盘镜像或任何活动的 I/O。

---

## AP 和域

所有的 Sun Enterprise 服务器都支持域。Sun Enterprise 10000 服务器支持动态系统域。但是，AP 不能跨域使用。

例如，假设一块包含控制器的板是某个路径组的一部分，而且您要将该板用 DR 移动到其它域内，那么只有在该板的替代路径当前处于非活动状态时，才能执行这个操作。这种情况下，将再也无法切换该板的替代路径。

# Alternate Pathing 数据库

---

本章介绍如何创建和管理用于维护 AP 配置状态的 AP 数据库。

---

## 管理数据库的副本

AP 维护着一个数据库，其中包含有关所有定义的元磁盘和元网络及其相应替代路径和属性的信息。要经常为该数据库设置多个副本。这样，如果给定的数据库副本无法访问或已损坏，AP 可以自动开始使用最近的，没有损坏的数据库副本。



---

**警告** – AP 引导磁盘至少需要一个当前的、没有损坏的 AP 数据库，否则系统无法引导。

---

必须为每个数据库副本分配至少 300 KB 的完整磁盘分区。分区过大也会浪费磁盘空间。在为 AP 数据库选择磁盘分区时，请牢记以下信息：

- 设置三到五个数据库副本。
- 依据磁盘的出厂配置，已将磁盘的第四分区分配给 AP 数据库，且不作其它任何用途。若您还没有将此分区移作它用，则是存储 AP 数据库副本的理想选择。
- 数据库副本彼此应该没有共同的 I/O 控制器。遵循此规则，在控制器出现故障的情况下，会具有最大的可用性。
- 如果已配置系统以使用 DR，则应将数据库副本驻留在连接到不同系统板的 I/O 控制器的磁盘上，以便在分离了其中一块系统板之后，还可以访问数据库副本。

- 如果要将在 AP 数据库副本放置在替代路径磁盘的分区上，则将使用 AP 元磁盘用以访问该分区的每条物理路径为该数据库各创建一个副本（共两次）。AP 工作时就像存在数据库的两个副本，而实际上只有一个副本，因为是通过两条路径对该磁盘进行访问。但是，这种工作方式并不会导致数据库不一致，因为 AP 始终更新并按顺序访问数据库副本。此工作方式也不会导致性能问题，因为并不会频繁访问 AP 数据库。

在 AP 版本低于 AP2.3 的 Sun Enterprise 10000 服务器上，SSP 上 AP 数据库中包含的信息子集会在引导时使用。此数据库包含有关引导磁盘的 AP 信息。如果要继续使用低于 AP2.3 的 AP 版本：

1. 不要从 SSP 上删除 SUNWapssp 软件包。
2. 验证 SUNWapssp 的版本（对于所使用的、AP 2.3 之前的 AP 最新版本而言）是最新的。例如，如果您在一个域上运行 AP 2.0，而在另一个域上运行 AP 2.1，那么您的 SUNWapssp 软件包针对 AP 2.1 而言应是最新版本。如果该软件不是最新的版本，将导致无法在 UNIX 引导之前，引导 AP 控制的引导磁盘的替代路径。

---

## 定位数据库以高效使用 RAS

您应该考虑如何使用驻留 I/O 控制器的系统板，该控制器连接着存储 AP 数据库的磁盘。如果是由于要将系统板在域之间迁移，而经常拆卸它，那么就不应把 AP 数据库存储在与该板上驻留的控制器相连接的磁盘上。如果觉得有必要拆卸系统板，那么在 AP 试图写入不可用的数据库时，控制台将会接收到错误消息。这并非一个严重的问题。您可以随时再插上这块系统板，并立即重新同步旧的数据库。但是，如果在此期间您还将该板挂接到其它域上，这些域就可能将数据写入为 AP 数据库保留的磁盘分片上。

---

## 创建和删除数据库

下列 AP 示例假定命令搜索路径中已包括这些命令所在的目录。请参见第 43 页的“使用单用户模式”。

### ▼ 创建 AP 数据库的一个副本

- 使用 `apdb(1M)`:

```
# apdb -c /dev/rdisk/c0t1d0s4
```

其中:

`-c` 指定了要在其中创建数据库副本的原始磁盘分片（位于 `/dev/rdisk` 目录下）。必须为每个数据库副本指定整个磁盘分区。该磁盘分区必须至少有 300 KB。

## ▼ 删除 AP 数据库的副本

- 使用 `apdb(1M)`:

```
# apdb -d /dev/rdsk/c0t1d0s4 -f
# apconfig -D
#
```

其中:

`-d` 指定了要在其中删除数据库副本的原始磁盘分片（位于 `/dev/rdsk` 目录下）。

`-f`（强制）只在删除倒数第二个和最后一个数据库副本时才必须使用。

在本例中，使用 `apconfig -D` 验证是否已删除最后一个数据库副本。通常，使用 `apconfig -D` 查看有关现有 AP 数据库副本的信息。若无返回信息，`apdb(1M)` 命令一定已删除了最后一个数据库副本。

如果在删除最后一个数据库副本之后重新启动，那么所有的 AP 元设备都将不可用。因此在重新启动之前最好取消所有 AP 元设备的配置，否则在系统恢复时，将断开对元设备的引用（例如 `/etc/vfstab`）。有关更多信息，请参见第 38 页的“取消配置元磁盘”或第 51 页的“取消配置元网络”。



---

**警告** – 如果删除了最后一个数据库，而您的引导磁盘又具有替代路径，这意味着在系统崩溃或重新启动时将不可引导！因此，在删除了最后一个数据库之前，务必要在重新启动之前，利用 `apboot(1M)` 命令使引导磁盘脱离 AP 的控制。请参阅第 42 页的“使引导磁盘脱离 AP 控制”。

---

## 查看数据库信息

可以查看数据库中的信息，包括有关数据库副本、数据库内磁盘项以及数据库内网络项等的信息。

## ▼ 查看有关数据库副本的信息

- 使用 `apconfig -D`:

```
# apconfig -D

path: /dev/rdisk/c0t1d0s4
major: 32
minor: 12
timestamp: Thu Jul 27 16:24:27 1995
checksum: 687681819
corrupt: No
inaccessible: No
```

在本例中，只有一个 AP 数据库。该命令显示了此数据库的路径，及其主要编号、次要编号、时间戳和校验和。corrupt 字段表明该数据库是否损坏。（如果 corrupt 设为 Yes，数据不会按照校验和进行正确验证。）inaccessible 字段表明是否可以访问包含数据库的设备。

---

## 查看路径组信息

AP 数据库包含有关磁盘和网络路径组的信息。在最初定义路径组时（如第三章和第五章所述），其路径组定义是一个未提交的数据项。在提交路径组定义之前，与此项关联的元磁盘或元网络均无法使用。反之，在删除路径组定义时，必须在提交之后删除才能生效。这两种状态（未提交和提交）使您在操作继续执行之前，可以查阅该操作的效果。要提交尚未提交的数据项，可使用 `apdb -C` 命令。

---

**注意** – 未提交的项会一直保存在该数据库中，直到您提交或删除之。但升级例外。升级 AP 软件会删除未提交的项。

---

## ▼ 查看未提交的磁盘项

- 使用带有 `-S` 和 `-u` 选项的 `apconfig(1M)` 命令（如下所示），其中 `-S` 代表存储，`-u` 代表未提交：

```
# apconfig -S -u

c1      sf:0  P A
c2      sf:1
        metadiskname(s):
                mclt5d0  U
                mclt4d0  U
                mclt3d0  U
                mclt2d0  U
                mclt1d0  U
                mclt0d0  U
```

有关更多信息，请参见第三章。

## ▼ 查看已提交的磁盘项

- 使用带有 `-S` 选项的 `apconfig(1M)` 命令（如下所示），其中 `-S` 代表存储：

```
# apconfig -S

c1      pln:0  P A
c2      pln:1  A
        metadiskname(s):
                mclt5d0  R
                mclt4d0
                mclt3d0
                mclt2d0
                mclt1d0
                mclt0d0
```

有关更多信息，请参见第三章。



## ▼ 查看未提交的网络项

- 使用带有 `-N` 和 `-u` 选项的 `apconfig(1M)` 命令（如下所示），其中 `-N` 代表网络，`-u` 代表未提交：

```
# apconfig -N -u

metanetwork:    mether0  U
physical devices:
                 hme2    A
                 qfe0
```

有关更多信息，请参见第五章。

## ▼ 查看已提交的网络项

- 使用带有 `-N` 选项的 `apconfig(1M)` 命令，如下所示：

```
# apconfig -N

metanetwork:    mether3
physical devices:
                 hme4    A
                 qfe2
```

有关更多信息，请参见第五章。



# 使用元磁盘和磁盘路径组

---

只能为可通过两条路径访问的磁盘创建元磁盘和磁盘路径组。通常，要在不同的系统板上使用两个单独的控制器的。

---

**注意** – 当磁盘放置在 AP 控制下或者删除路径组时，AP 不会修改该磁盘上的数据（包含 AP 数据库副本的磁盘分片上的数据除外）。AP 不会对磁盘重新分区。如果删除了路径组，可以继续利用其物理设备名称访问数据。

---

---

## 元磁盘的设备节点

以下是磁盘设备的 *物理* 设备节点的两个示例：

- /dev/dsk/c0t0d0s0
- /dev/rdisk/c0t0d0s0

其中：

c 表示主机上（不是磁盘阵列上）的 I/O 端口

t 是磁盘阵列的总线

d 是该总线上包含的磁盘的目标 ID

s 是该磁盘的分片号

这些物理设备节点代表 SCSI 磁盘分区的特定物理路径。

其中：

c 是主机适配器号

t 是磁盘插槽的目标号

d 是磁盘号

s 是该磁盘的分片号

每个控制器端口都有一个端口号（例如 c0）和一个端口名（例如 p1n:2 或 sf:2）。端口名由端口类型和实例号组成，之间用冒号(:)分隔。这与以前版本的 AP 端口名不同。这种更改只适用于磁盘驱动程序命名约定，但不适合网络。请参考 path\_to\_inst(4) 以获得有关实例号的更多信息。

当磁盘插槽连接两个端口时，可以利用物理设备节点从任意一条路径访问它，例如 /dev/dsk/c0t0d0s0 或 /dev/dsk/c1t0d0s0。

元磁盘的设备节点源于路径组中主路径的物理设备节点。以下是元磁盘设备节点的两个示例：

- /dev/ap/dsk/mc0t0d0s0
- /dev/ap/rdisk/mc0t0d0s0

您可以看到，已添加了一个 ap 目录，m（其含义为“元”）置于设备规范之前。元磁盘的设备节点能够通过多条路径访问主要的物理磁盘驱动器。

---

## 自动切换元磁盘

在下列两种情况下，元磁盘可以自动从活动路径切换到替代路径：

- 活动路径失败。
- 含有连接活动路径的控制器的系统板已利用 DR 分离操作拆卸下来。只有在 Sun Enterprise 10000 服务器上执行 DR 分离操作时，可用自动切换。

对于 Sun StoreEdge T3 磁盘，在执行任何 DR 操作之前，首先应进行手动切换。这将禁用 T3 磁盘上的路径优化。此后，当两条物理路径再次可用时，可重新启用路径优化，方法是使用：

```
# apconfig -P sf:2 -a sf:2 -a sf:9
```

当活动路径失败时，只有在替代路径可用的情况下才尝试自动切换。然后将这条出故障的路径标记为不可用，或*经过试用*。您可以用 `apconfig-S` 命令识别经过试用的路径：

```
# apconfig -S

c1      sf:0  P A
c2      sf:1  T
        metadiskname(s) :
                mclt5d0
                mclt4d0
                mclt3d0
                mclt2d0
                mclt1d0
                mclt0d0
```

在本例中，当前不活动的路径 `sf:1` 被标记为 `T`，这表明该路径经过试用，但失败。

“经过试用”标记只对自动切换操作而言有意义，对于手动切换操作无效。通常，AP 2.3.1 不会尝试自动切换到经过试用的路径上。这样在两条路径均失败的情况下，可防止系统崩溃。

可以使用以下任意操作复位“经过试用”标记：

- 重新启动相应的域。
- 执行 DR 挂接操作将包含标记为经过试用的控制器的系统板挂接后，再执行 DR 分离操作。
- 手动复位特定控制器的“经过试用”标记。

可以按以下示例所示手动复位“经过试用”标记：

```
# apdisk -w sf:1
```

在本例中，`sf:1` 是一个“经过试用”标记为真的控制器。应谨慎使用 `apdisk-w` 功能。此命令只清除“经过试用”标记；不解决有关控制器或设备的任何潜在问题。此命令仅在已恢复失败的路径且不干涉 DR 操作或重新启动的情况下使用。注意您可以尝试执行手动切换到标记为经过试用的路径上。

---

## 磁盘可用性和性能折衷

在配置磁盘阵列和控制器之前，需要建立您的磁盘使用优先级。通过将性能折衷，或购买更多的硬件，可以更有效地使用磁盘资源。

设想有一个双端口的 SSA 磁盘阵列。此类设备可以连接一个或两个光纤通道磁盘控制器（SOC 控制器）。在 SSA 内，有多个目标。每个目标含有多个磁盘。每个磁盘被划分为多个磁盘分片。根据系统设置的不同，对于这些磁盘 I/O 资源的争用有若干种级别。

- 磁盘争用
- 目标争用（I/O 总线争用）
- 控制器争用

例如，假定您将磁盘划分为四个分片，跨这四个分片创建一个文件系统。尽管该文件系统跨越多个分片，这些分片均驻留在同一张磁盘上，您也可将该文件系统置于一个分片上。上述情况是一种不明智的配置，可能会导致在磁盘级上争用资源，因为每次读取或写入该文件系统均需要访问同一个磁盘。

可以将文件系统放在同一个目标的多张磁盘上。在这种情况下，会导致目标级的资源争用，因为每次读取或写入该文件系统均需要访问同一个目标。目标级的资源争用没有磁盘级的资源争用严重，但它依然是个较差的配置。

如果在同一个 SSA 上跨越多个目标创建文件系统，将会导致控制器级的资源争用。因为每次读取或写入该文件系统均需要访问同一个控制器。

通常，最好的方法是利用多个控制器，创建一个跨越多个 SSA 磁盘阵列的文件系统。但是，在磁盘访问速度和系统可用性之间要有一个折衷方案。文件系统使用的磁盘阵列越多，磁盘访问速度也越快。但是，如果这些磁盘阵列中的任意一个组件出现故障，文件系统将不可用。如果限制文件系统的磁盘阵列数量，将它限制在一个磁盘阵列上，会导致性能下降，但系统整体的可用性将增强。这是因为给定的文件系统中出现故障的组件减少。

假设您有六个磁盘控制器分别连接到三个双端口的 SSA 磁盘阵列。

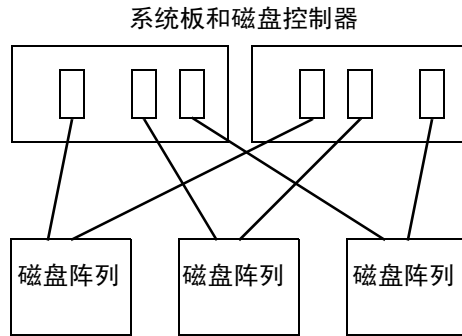


图 3-1 系统板和磁盘控制器

如果想实现最佳的可用性，可以利用 AP 为每个 SSA 建立替代路径。这样做的优点是可使用 DR 来挂接和分离系统板（以维修或升级这些系统板），而同时仍能够访问磁盘阵列上的文件系统。当然，若要这样做，必须将替代磁盘控制器（SOC 控制器）放置在不同的系统板上。一种安排是使用两块系统板，每块板上使用三个磁盘控制器。这是一种简洁的布置，并使您能够在需要分离其中一块板时，切换到另一块板上的控制器。同样还能使您通过分离和挂接单系统板，使在域内迁移磁盘资源变得相当简单。

每个 SSA 必须购买两块 SOC 控制器。同时，在大型安装中，对于 SBus 插槽数目有一定的限制，这些插槽用于容纳连接大量 SSA 双重路径所需的 SOC 控制器。

## 磁盘镜像注意事项

如果使用第三方卷管理器，例如 SDS 或 VERITAS Volume Manager™(VxVM) 镜像磁盘，同时还想用 DR 分离系统板，那么必须配置您的卷和镜像，以使它们能与 AP 和 DR 协同工作。

例如，假设您有 12 块系统板，每块都有一个主机适配器（在下图中称为“控制器”）：

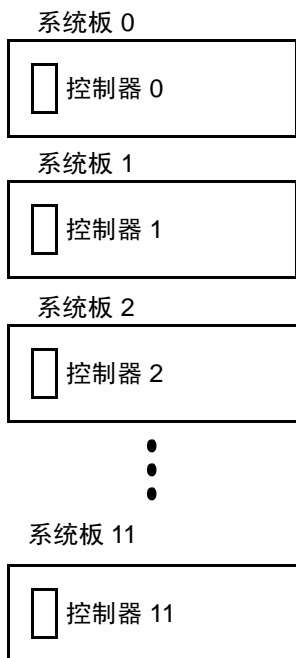


图 3-2 系统板和控制器的配置

您可能需要创建一个镜像卷。可考虑如下配置：

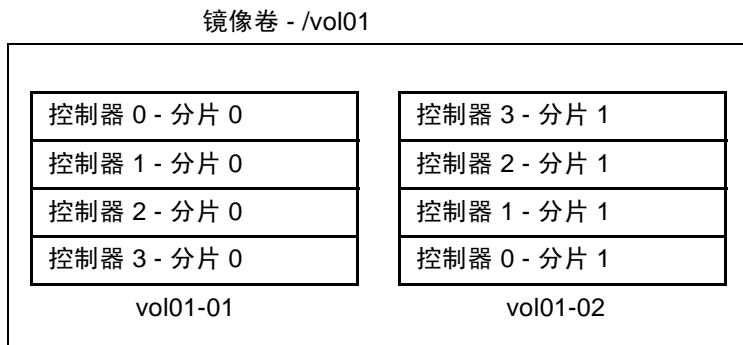


图 3-3 镜像卷示例 1



在示例 1 中，vol01-01 由一个四路分片构成，它可以通过驻留在四块独立系统板上的四个独立控制器进行访问。vol01-01 被镜像到也由一个四路分片构成的 vol01-02 上。例如，控制器 0 分片 0 被镜像到控制器 3 分片 1 上，等等。

考虑需要分离一块含有此四个控制器之一的系统板的情况。在分离该板之前，必须停用使用该板控制器的那一半镜像。则使用上述配置是不可能的。例如，如果您想分离板 0（含有控制器 0），将不得不同时停用镜像的双方，从而使文件系统无法访问。结果将导致无法按上述配置在任何一块系统板上使用 DR。

解决此问题的一种方法是镜像卷，以使同一块系统板的控制器不会同时出现在镜像的双方，例如：

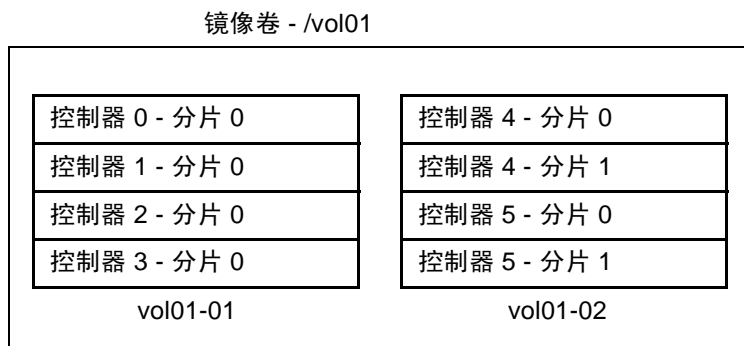


图 3-4 镜像卷示例 2

在示例 2 中，可以分离任何系统板（板 0 到板 5），但要首先停用使用该板控制器的那一半镜像。例如，要分离板 4（驻留有控制器 4），首先要停用 vol01-02。您依然可以访问文件系统，因为它可通过 vol01-01 访问。稍后在您重新挂接板 4 时，可以再次将 vol01-02 添加到镜像上。

但是，这种解决方案的问题是在镜像关闭时，系统由于单点故障极易损坏。如果一个磁盘出现故障，就没有镜像的备份磁盘。可以通过使用 AP 防止出现这种情况。可能要设置以下 AP 元设备：

- mc0 是控制器 0 和控制器 6 的元设备
- mc1 是控制器 1 和控制器 7 的元设备
- mc2 是控制器 2 和控制器 8 的元设备
- mc3 是控制器 3 和控制器 9 的元设备
- mc4 是控制器 4 和控制器 10 的元设备

- mc5 是控制器 5 和控制器 11 的元设备

为便于讨论，使用缩写。例如，元设备的完整名称可能是 mc0t0d0s0，并且它可能将物理设备 c0t0d0s0 和 c6t0d00s0 作为替代路径。

可考虑如下配置：

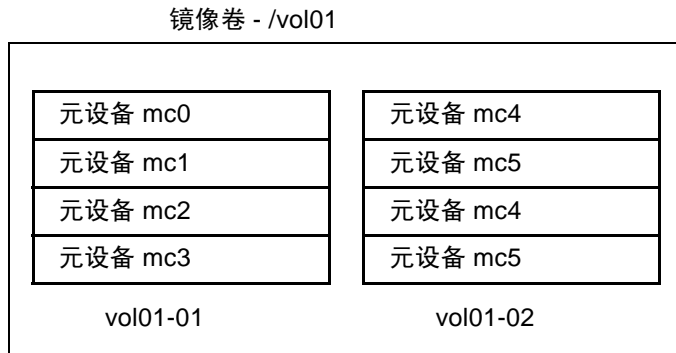


图 3-5 镜像卷示例 3

在示例 3 中，您可以分离任意一块板（板 0 到板 11），而不必取消镜像。这将降低单点故障所带来的危险。例如，要分离含有控制器 4 的系统板 4，首先要切换元设备 mc4 以使它能够使用板 10 的控制器 10。（只需用一个 AP 命令，`apconfig -P` 即可实现此操作。）

在本例中，随着您增大了 RAS 支持的级别（即磁盘 I/O 资源的可用性和防止受单点故障损坏的级别），将不得不在配置中增加系统板和控制器的数目。这等于增加系统的成本从而能够更好地支持 RAS 功能。

尽管这是一个假想的例子，主要目的还是让您了解在设置卷和镜像时必须考虑 AP 和 DR。否则，可能导致无法使用 AP 和 DR。如果使用第三方卷管理器，那么要跟踪构成卷的*物理*控制器和分片。卷管理器可用于自动选择物理组件，但这个选择过程并不考虑到 AP 和 DR 的情形。如果要确保物理组件与 AP 和 DR 兼容，必须明确地选择构成卷的物理组件。

---

## 使用磁盘路径组和元磁盘

---

注意 – 本节中的示例命令都使用的是 `pln` 端口（对于 SSA 磁盘阵列）。如果您有 Sun StorEdge™ A5000 或 T3 磁盘阵列 / 插槽，那么应将示例中的 `pln` 端口，指定为 `sf` 端口或 `fp` 端口（仅对于 Solaris 8 环境）。某些示例中为 T3 使用了 `sf` 端口。有关 Sun 支持设备的列表，请参考《*Sun Enterprise Server Alternate Pathing 2.3.1 发行说明*》。

---

### ▼ 创建磁盘路径组和元磁盘

1. 确定哪两个端口将成为该路径组的替代路径。
  - a. 可以使用 `apinst(1M)` 命令显示所有端口（例如 `pln:0` 和 `pln:1`）以及它们的磁盘设备节点（例如 `/dev/dsk/c1t0d0`）：

```
# apinst
pln:0
    /dev/dsk/c1t0d0
    /dev/dsk/c1t1d0
    /dev/dsk/c1t2d0
    /dev/dsk/c1t3d0
    /dev/dsk/c1t4d0
    /dev/dsk/c1t5d0

pln:1
    /dev/dsk/c2t0d0
    /dev/dsk/c2t1d0
    /dev/dsk/c2t2d0
    /dev/dsk/c2t3d0
    /dev/dsk/c2t4d0
    /dev/dsk/c2t5d0
```

- b. 必须了解您的系统硬件配置以识别两个端口何时连接到同一个磁盘阵列上。

在本例中，假设 SSA 包含六个磁盘和两个 SSA 端口。一个 SSA 端口连接 `pln` 端口 `c1`，另一个 SSA 端口连接 `pln` 端口 `c2`。

## 2. 使用带有 `-c`、`-p` 和 `-a` 选项的 `apdisk(1M)` 命令创建一个未提交的磁盘路径组:

```
# apdisk -c -p pln:0 -a pln:1
```

其中:

`-p` 指定主路径

`-a` 指定替代路径

`-c` 指定要创建的信息

`apdisk(1M)` 命令创建了一个元磁盘名, 以及 AP 数据库中所有必需的信息, 用于维护六个磁盘各自的两条替代路径。

## 3. 验证结果:

```
# apconfig -S -u

c1      pln:0  P A
c2      pln:1
        metadiskname(s):
            mc1t5d0  U
            mc1t4d0  U
            mc1t3d0  U
            mc1t2d0  U
            mc1t1d0  U
            mc1t0d0  U
```

`apconfig -S -u` 命令列出了未提交的元磁盘。

其中:

`-S` 只列出存储设备 (即只列出磁盘, 不列出网络)

`-u` 只列出未提交的设备

每个磁盘名后面的 `U` 表明该元磁盘项尚未提交。

`pln:0` 后面的 `P` 表明 `pln:0` 是主路径, `A` 表明 `pln:0` 是活动替代路径。

尽管可以更改活动替代路径, 主路径始终不变。主路径的意义在于, 它是最初的活动替代路径, 在命名元磁盘时使用它, 并通过它识别该元磁盘。在这种情况下, 主路径名 `c1t0d0` 会成为元磁盘名 `mc1t0d0` 的一部分。

4. 如果对上一步显示的结果满意，可使用带有 `-C` 选项的 `apdb(1M)` 命令提交这些未提交的数据库项。

```
# apdb -C
```

5. 继续之前，可通过带有 `-S` 选项的 `apconfig (1M)` 命令验证结果，以查看数据库中已提交的存储项：

```
# apconfig -S

c1      pln:0  P A
c2      pln:1
        metadiskname(s):
            mclt5d0
            mclt4d0
            mclt3d0
            mclt2d0
            mclt1d0
            mclt0d0
```

---

注意 – 在 T3 磁盘上，当有两条可用路径时，缺省情况下将创建对路径组的路径优化。执行上述命令将输出以下结果：

---

```
# apconfig -S

c1      sf:0  P A
c2      sf:1  A
        metadiskname(s):
            mclt5d0
            mclt4d0
            mclt3d0
            mclt2d0
            mclt1d0
            mclt0d0
```

如果在一个物理路径名下装入了一个分区，则应将该分区卸下并重新装入到元磁盘路径名下。

如果不卸下该分区，可能是因为它正被频繁使用，可以暂且不将该分区置于 AP 控制之下，直到准备好将系统关机，并进行维护然后重新启动，再执行此操作。在这种情况下，可以修改 `/etc/vfstab` 文件，以便系统重新启动后，该分区即位于某 AP 设备控制之下。如果将引导磁盘置于 AP 控制下，同时还需要使用第四章所述的 `apboot(1M)` 命令修改 `/etc/vfstab` 文件。

使用 `apconfig-S` 命令列出数据库中已提交的存储项。如本例所示，此列表与上一个列表基本相同，只是字符 `U` 不再出现在元磁盘名的后面，表明元磁盘不再处于未提交状态。

#### 6. 使用以下命令查看符号链接并验证结果：

```
# ls -l /dev/ap/dsk
total 8
lrwxrwxrwx 1 root          40 Jul 27 16:47 mclt0d0s0 ->
  ../../../../devices/pseudo/ap_dmd@0:128,blk
lrwxrwxrwx 1 root          40 Jul 27 16:47 mclt0d0s1 ->
  ../../../../devices/pseudo/ap_dmd@0:129,blk
lrwxrwxrwx 1 root          40 Jul 27 16:47 mclt0d0s2 ->
  ../../../../devices/pseudo/ap_dmd@0:130,blk
```

现在可以使用位于 `/dev/ap/dsk` 和 `/dev/ap/rdisk` 目录下的所需设备节点了。

#### 7. 修改使用物理设备节点（即以 `/dev/dsk` 或 `/dev/rdisk` 开始的路径）的所有引用以使用对应的元磁盘设备节点（即以 `/dev/ap/dsk` 或 `/dev/ap/rdisk` 开始的路径）。

## ▼ 从主路径切换到替代路径

在任何时候都可以执行切换，即使设备正在使用 I/O 也可以切换。可以尝试执行此切换过程以验证您理解该过程并且系统设置正确，而不是出现紧急情况时才进行切换。



---

**警告** – 在切换路径时，AP 不会检查能否将该路径上传输的数据切换到新路径上，尽管能确定该路径是处于分离或脱机状态。可以通过执行 `prtvtoc(1M)` 命令之类的 I/O 操作在切换路径之前验证该路径的状态。如果切换到一条不能正常工作的路径上，AP 不会生成任何错误或警告消息。如果将引导磁盘的路径切换到一条工作异常的路径，且如果该路径无法立即切换回来，那么系统可能崩溃。

---

1. 使用带有 `-S` 选项的 `apconfig(1M)` 命令查看当前配置:

```
# apconfig -S

c1      pln:0  P A
c2      pln:1
        metadiskname(s):
            mclt5d0
            mclt4d0
            mclt3d0
            mclt2d0
            mclt1d0
```

在本例中, `pln:0` 是活动替代路径, 因为它后面有一个字符 `A`。同时它还是主路径, 因为它后面还有一个字符 `P`。

对于 `T3`, 应该有类似下列的输出:

```
# apconfig -S

c1      sf:0  P A
c2      sf:1
        metadiskname(s):
            mclt5d0
            mclt4d0
            mclt3d0
            mclt2d0
            mclt1d0
```

在本例中, `sf:0` 是活动替代路径, 因为它后面有一个字符 `A`。 `sf:0` 也是主要路径, 因为它后面还有一个字符 `P`。

---

注意 - 当有两条可用路径时, 缺省情况下将在 `T3` 磁盘上启动路径优化。也可能只有一条路径, 在此情况下, 上例中, `sf:1` 将不会作为活动替代路径列出。

---

2. 要执行切换, 可使用带有 `-P` 和 `-a` 选项的 `apconfig(1M)` 命令:

```
# apconfig -P pln:0 -a pln:1
```

---

**注意** – 如果 T3 上原来启用了路径优化，此操作将禁用它。

---

-P 指定了主路径，从而可 *识别路径组*，此路径组是要更改其活动替代路径的组。因此，上例中的 -P pln:0 标识出 pln:0 是活动路径的路径组。

-a 指定了要切换为活动替代路径的那条替代路径。

3. 可以使用带 -S 选项的 apconfig(1M) 命令验证结果，以查看数据库中已提交的元磁盘：

```
# apconfig -S

c1      pln:0  P
c2      pln:1  A
        metadiskname(s):
                mclt5d0
                mclt4d0
                mclt3d0
                mclt2d0
                mclt1d0
```

---

**注意** – 执行 AP 切换后，仅有一条活动路径。将禁用对 T3 的路径优化。

---

活动替代路径已被切换为 pln:1。

注意，不必提交切换操作。



## ▼ 切换回主路径

1. 可以使用以下命令切换回主路径:

```
# apconfig -P pln:0 -a pln:0
# apconfig -S

c1      pln:0  P A
c2      pln:1
        metadiskname(s):
                mclt5d0
                mclt4d0
                mclt3d0
                mclt2d0
                mclt1d0
```

---

注意 – 仍将禁用对 T3 的路径优化。上一条命令的输出应该类似于:

---

```
# apconfig -P pln:0 -a pln:0
# apconfig -S

c1      sf:0   P A
c2      sf:1
        metadiskname(s):
                mclt5d0
                mclt4d0
                mclt3d0
                mclt2d0
                mclt1d0
```

第一个 `apconfig(1M)` 命令切换了主控制器为 `pln:0` 的路径组的活动替代路径。切换后的活动替代路径变为 `pln:0`。

2. 要重新启用 T3 上的路径优化, 可使用:

```
# apconfig -P sf:0 -a sf:0 -a sf:1
```

## ▼ 删除磁盘路径组和元磁盘

1. 如果您的引导磁盘位于 AP 控制之下，那么可以使用 `apboot(1M)` 命令使它脱离 AP 控制，具体操作如第 42 页的“使引导磁盘脱离 AP 控制”中所述。

无须从引导磁盘卸下任何已装入的文件系统，因为 `apboot(1M)` 将这些文件系统放置在物理设备的顶部。

2. 卸下位于 AP 元磁盘顶部的文件系统（不是装入引导磁盘的文件系统）。

您的脚本和程序可能引用以下格式的元磁盘：

`/dev/ap/dsk/mc?t?d?s?` 和 `/dev/ap/rdisk/mc?t?d?s?`

这些引用必须分别转换为以下格式的物理设备引用：

`/dev/dsk/c?t?d?s?` 和 `/dev/rdisk/c?t?d?s?`

通常，对元磁盘的引用会出现在以下位置：

`/etc/vfstab`

`/etc/system`

`/etc/dumpadm.conf`

引用磁盘的任何应用程序或脚本

3. 使用带 `-d` 选项的 `apdisk(1M)` 命令指定要删除的路径组的主路径：

```
# apdisk -d pln:0
```

4. 要验证结果，可使用带 `-S` 选项的 `apconfig(1M)` 命令查看数据库中已提交的元磁盘项：

```
# apconfig -S

c1      pln:0  P A
c2      pln:1
        metadiskname(s):
            mclt5d0  D
            mclt4d0  D
            mclt3d0  D
            mclt2d0  D
            mclt1d0  D
            mclt0d0  D
```

如果该路径组不是以前提交的，那么使用 `apdisk-d` 命令会将它从数据库中删除掉。但是，如果该路径是以前提交的，`apdisk-d` 命令只是简单地将它标记为已删除，直至您下次向数据库提交项之后，删除操作才能完成。在上例中，`pln:0` 路径组是以前提交的，所以字母 `D` 指明它被标记为删除。

5. 使用 `apdb(1M)` 命令提交数据项，从而完成删除操作：

```
# apdb -C
```

6. 可以用带 `-S` 选项的 `apconfig(1M)` 命令验证删除：

```
# apconfig -S
```

---

注意 – 如果删除未提交，可以取消删除。要取消删除，可使用 `apdisk -z` 指定您以前所指定的同一端口。

---

## ▼ 取消配置元磁盘

- 将脚本引用从以下格式:

`/dev/ap/dsk/mc?t?d?s?` 和 `/dev/ap/rdisk/mc?t?d?s?`

分别转换为:

`/dev/dsk/c?t?d?s?` 和 `/dev/rdsk/c?t?d?s?`

通常, 对元磁盘的引用会出现在以下位置:

`/etc/vfstab`

`/etc/system`

`/etc/dumpadm.conf`

任何引用已配置的元磁盘的应用程序或脚本。

## ▼ 重新配置元磁盘

此步骤假设您以前创建了磁盘路径组和元磁盘, 并在随后取消了对元磁盘引用的配置。如果只是想重新配置元磁盘接口, 可使用此步骤。

- 将物理设备引用转换回元磁盘引用; 从以下形式:

`/dev/dsk/c?t?d?s?` 和 `/dev/rdsk/c?t?d?s?`

分别转换为:

`/dev/ap/[r]dsk/mc?t?d?s?`

通常, 对磁盘设备的引用会出现在以下位置:

`/etc/vfstab`

`/etc/system`

`/etc/dumpadm.conf`

任何引用磁盘或其它应用程序的应用程序或脚本, 或引用现在位于新配置的元设备控制下的物理磁盘设备的脚本

# 使用 AP 引导设备

本章介绍如何使引导磁盘具有替代路径。

## 将引导磁盘置于 AP 控制之下

现在通过将引导磁盘置于 AP 控制下，可以在所有的 Sun Enterprise 服务器上执行无人值守（自动）的系统引导（即使在引导磁盘的控制器出现故障时）。

可以使用 Dynamic Reconfiguration (DR) 在所有的 Sun Enterprise 服务器上分离系统板，即使系统板上有引导磁盘的控制器也可以将它分离。为此，必须使用不同系统板的两个控制器为引导磁盘建立替代路径。但是请注意，如果主网络的控制器与引导磁盘的一个控制器位于同一块系统板上，则必须使主网络也具有替代路径。否则，将无法使用 DR 分离该板。

### ▼ 将引导磁盘置于 AP 控制之下

1. 为引导磁盘创建一个 AP 路径组。

此过程在第三章中已做介绍。

2. 使用 `apboot(1M)` 定义新的 AP 引导设备。

`apboot (1M)`:

- 修改了 `/etc/vfstab`、`/etc/system` 和 `/etc/dumpadm.conf`。

例如：

```
# apboot mc2t0d0
```

其中：

mc2t0d0 是引导磁盘的元磁盘名。

- 检查 `/etc/vfstab` 并用元磁盘名 `/dev/ap/dsk/mc2t0d0*` 替换物理磁盘的物理设备名（例如 `/dev/ap/dsk/c2t0d0*` 或 `/dev/dsk/c1t0d0*`）。

不要用引导磁盘的元磁盘手动替换 `/etc/vfstab` 下的物理设备。而应使用 `apboot(1M)` 命令以确保完成全部所需的更改。

- 检查 `/etc/vfstab`，以确定是否要将交换设备更改为元设备。如果确定，会将交换设备转换为元设备。
- 编辑 `/etc/system`，以便在适当的时间加载 AP 引导磁盘所需的内核驱动程序。
- 检查转储设备的配置，并在必要时调用 `dumpadm(1M)` 命令，将转储设备配置为元设备。
- 更新 OpenBoot™ PROM 引导设备属性，从而列出每条替代路径的物理路径。

---

**注意** – 如果选择取消此功能（使用 `apboot -o` 命令），在 UNIX 引导之前，将禁用自动重新选择受 UNIX 控制的引导磁盘的替代路径。

---

### 3. 将已装入引导磁盘内的文件系统置于 AP 控制之下。

通常，在引导过程中，装入的文件系统被分开装在两张单独的磁盘上（由于磁盘空间所限）。如果将引导磁盘置于 AP 控制之下（使用 `apboot(1M)` 命令），必须手动编辑 `/etc/vfstab` 文件以便将引导过程中装入的其它文件系统也置于 AP 控制之下。

在 `/etc/vfstab` 文件中，必须为要置于 AP 控制下的所有其它装入点更改 `device to mount` 和 `device to fsck` 的路径。

例如：

```
# device          device          mount    FS    fsck
mount            mount
# to mount          to fsck        point    type  pass
at boot          options
#...
/dev/ap/dsk/mc1t34d0s1 -                -        swap    -
no              -
/dev/ap/dsk/mc1t34d0s0 /dev/ap/rdisk/mc1t34d0s0 /    ufs    1
no              -
/dev/ap/dsk/mc1t34d0s6 /dev/ap/rdisk/mc1t34d0s6 /usr  ufs    1
no              -
/dev/ap/dsk/mc1t34d0s7 /dev/ap/rdisk/mc1t34d0s7 /export/home ufs
2    yes    -
swap          -                -        /tmp    tmpfs    -
yes          -
#...
```

### 4. 此时，重新启动系统，开始使用 AP 引导设备。

## ▼ 使镜像引导磁盘具有替代路径

镜像引导磁盘是磁盘管理软件的一个主要功能。此步骤的目的是通知 AP，存在一个镜像的引导磁盘。当使用也具有替代路径的镜像引导磁盘时，该引导磁盘就有四条潜在的物理路径（镜像双方各有两条）。如果想最大限度地防止控制器出现故障，这是一种很好的配置方案。执行下面的步骤有两个优点。

- AP 确保始终将适当的替代路径指定为活动路径，即使您使用不同的引导设备路径进行引导。为此，首先必须将该引导磁盘置于 AP 控制之下。然后为引导磁盘的镜像创建一个路径组。

- AP 还必须确保引导过程中需要自动切换路径时，所有这四条路径均可作为替代路径。镜像系统中的缺省路径为：`primary1`、`mirror1`、`primary2`、`mirror2`。这与 AP2.2 或更早的版本不同，其目的在于改进冗余度和适用性。在具有替代路径、非镜像的系统中，缺省顺序是：`primary root`、`alternate root`。

1. 使 AP 控制引导磁盘的具体步骤如“将引导磁盘置于 AP 控制之下”所述。
2. 为引导磁盘的镜像创建一个 AP 路径组。  
此过程在第三章中已做介绍。
3. 通知 AP 关于引导磁盘镜像的信息。

```
# apboot -m mc3t0d0
```

在本例中，`mc3t0d0` 是引导磁盘镜像的元磁盘。

4. 用您的磁盘管理软件创建引导磁盘的镜像（使用两张元磁盘）。

## ▼ 使镜像引导磁盘脱离 AP 控制

- 使用 `apboot(1M)` 取消对 AP 镜像引导设备的定义。

```
# apboot -u mc3t0d0
```

## ▼ 使引导磁盘脱离 AP 控制

- 使用 `apboot(1M)` 命令指定适当的物理设备节点。

```
# apboot c2t0d0
```

在此命令中，`c2t0d0` 是引导磁盘一条替代路径的物理设备节点（即在 `/etc/vfstab` 中指定的节点）。

- `apboot(1M)` 编辑 `/etc/system` 文件以删除强制加载的 AP 内核驱动程序模块，因为当引导磁盘不是 AP 设备时不会再需要这些模块。
- `apboot(1M)` 重新配置交换、转储设备和 OpenBoot PROM `boot-device` 属性以便在必要时使用适当的物理设备路径。





---

**警告** – 如果将引导磁盘置于 AP 控制下，而后又决定要删除 AP 软件包（使用 `pkgrm(1M)` 命令），那么必须首先使用 `apboot(1M)` 命令使该引导磁盘脱离 AP 的控制。如果没有在执行删除 AP 的 `pkgrm` 之前将该引导磁盘脱离 AP 的控制，则磁盘会变得无法引导。

---

---

## AP 引导顺序

本节简要介绍了在具有替代路径的引导磁盘上引导 Sun 服务器时，各种事件的流向。这些事件的顺序表明，在引导过程中，如果有必要执行切换，引导磁盘控制器是如何实现自动切换的。引导顺序如下所述：

1. 缺省情况下，要通过在 **OpenBoot™ (OBP)** `boot-device` 属性中指定的第一台设备引导系统。注意，此设备的路径可能与引导磁盘的最后一个活动替代路径不同。
2. 在几秒钟到几分钟（不超过三分钟，时间的长短取决于您的固件）之后，会检测到从第一台设备引导失败。然后 OBP 会转向属性中列出的下一台引导设备。此过程会继续进行直到设备引导成功或 OBP 运行完这些设备。
3. 在重新引导成功后，AP 为成功的设备建立活动替代路径。

---

## 使用单用户模式

通常，在完全引导 Sun Enterprise 服务器时，可以使用位于 `/usr/sbin` 目录下的 AP 命令版本。但是，如果您的服务器由于引导过程没有完全完成，进入单用户模式，可以使用位于 `/sbin` 目录下的 AP 命令。位于 `/sbin` 下的 AP 命令版本与 AP 守护程序服务（此服务在单用户模式下不可用）无关。如果由于出现与 AP 相关的问题导致系统进入单用户模式，那么可以利用 `/sbin` 命令执行所需的 AP 操作解决这个问题。

两种与 AP 相关的问题可以导致系统进入单用户模式。

- 如果假定两条路径指向同一张磁盘（按照 AP 数据库），但这些路径实际却指向不同的磁盘，且该磁盘需要在引导过程中装入，那么服务器将无法找到它并进入单用户模式。只有在更改了路径组的物理配置，又没有运行 AP 命令来更新该数据库时，才会出现这种情况。

- 如果某张磁盘的活动替代路径变得无法访问，而在引导过程中又需要插入该磁盘，那么服务器将进入单用户模式。如果一张磁盘包含在引导过程中装入的文件系统，那么在该引导过程中就需要插入这张磁盘；即，该磁盘有一些项位于 `/etc/vfstab` 文件中。

这些情况仅在使用磁盘时出现，而对于网络则不会出现这些情况。在这两种情况下，均可以使用 `/sbin` 下的 `AP` 命令解决该问题。

# 使用元网络路径组

---

要使用 AP 元网络，网络路径组内的两种物理网络必须是同一种媒体类型，还必须位于同一子网中。例如，网络路径组可以由两个以太网或两个 FDDI 网络组成，而不能由一个以太网和一个 FDDI 网组成。以太网网络可由不同类型的以太网组成。例如，hme 和 qfe 可以属于同一个路径组。

网络路径组中的替代路径必须物理连接到同一网络上。例如，以太网控制器必须连接到同一子网。

当存在多个物理网络连接时，一次只能有一个活动控制器。控制器必须位于不同的系统板上，以便在执行 DR 操作（例如 DR 分离）时不影响所有潜在的活动替代路径。

本节中的 AP 切换步骤显示了如何切换活动替代控制器。

---

## 元网络接口

元网络接口名来源于替代路径所属的网络类型。以太网元网络接口名的格式为 methex，其中  $x$  是实例号，例如 mether0。FDDI 元网络接口名的格式为 mfdrix，其中  $x$  是实例号，例如 mfdri0。

在创建元网络接口时必须使用属于同一媒体类型的两个网络接口；例如，可以使用 hme0 和 qfe2，或 nf0 和 nf1。但是，不能同时使用 hme0 和 nf1。某些示例会遵循。

- 假定网络控制器 hme0 和 qfe1 连接到同一个以太网子网。元网络 methernet0 可以包括这两个控制器。对于所有类型的以太网控制器（hme、qfe、le 等等），只要它们位于同一子网，均可以混合使用。
- FDDI 网络可以是 SAS 也可以是 DAS。在创建元网络接口时可以将 SAS 和 DAS 混合配置。

---

## 使用文件路径组

### ▼ 创建网络路径组和元网络

---

注意 – 此步骤不适宜主网络。要使主网络具有替代路径，请参见第 53 页的“使主网络接口具有替代路径”。

---

#### 1. 使用带 -c 选项的 apnet(1M) 命令：

```
# apnet -c -a hme0 -a qfe2
# apconfig -N -u

metanetwork:    methernet0  U
physical devices:
                 hme0 A
                 qfe2
```

apnet(1M) 命令为两个物理设备 hme0 和 qfe2 创建了网络路径组和元网络接口名 methernet0。

apconfig(1M) 列出了数据库中未提交的网络项。

-N 指定了应列出的网络数据库项。

-u 指定了应列出的未提交项。

2. 如果您对创建的网络路径组满意，则提交该项：

```
# apdb -C
# apconfig -N

metanetwork:    mether0
physical devices:
                 hme0 A
                 qfe2
```

apdb -C 提交了数据库项。

apconfig -N 列出了数据库中已提交的网络项。

此表与步骤 1 中的表大致相同，不同的是 mether0 后不再出现字符 U。

3. 删除所有直接使用路径组两个成员的操作（参见 ifconfig(1M) 命令）。

a. 如果物理接口当前处于“plumbed”状态，那么除了下面两种情况外，都要对该物理接口执行 unplumb 操作：

- 它不是主网络接口。
- 它不是将要用于配置元网络的接口。

如果将要执行 unplumb 的接口是主网络接口，或者是要用于配置元网络的接口，那么请遵循第 53 页的“使主网络接口具有替代路径”中的步骤。

可以按以下示例对物理接口执行 unplumb 操作：

```
# ifconfig hme0 down unplumb
```

通常，在系统引导过程中利用文件 /etc/hostname.xxxx 配置网络接口，其中 xxxx 是接口名称（例如 hme0）。此文件含有 IP 地址或与该接口关联的主机名。应该为所有建立了 AP 替代路径的接口删除或重命名对应的 /etc/hostname.xxxx 文件，因为不允许直接使用替代路径。

---

**注意 - IPv6:** 对于 AP 而言，无论在什么情况下使用 hostname.xxxx，均可以使用 hostname6.xxxx 文件。如果您的系统上有 IPv4 和 IPv6，则需要确保每个文件中的项彼此是一致的。有关 IPv6 的更多信息，请参见《System Administration Guide》的“Volume 3”。

---

- b. 为要在系统重新启动时配置的所有元网络创建一个 `/etc/hostname.metherx` 文件（如 `/etc/hostname.mether0`）。

此文件应该含有元网络 IP 地址或与该接口关联的或主机名。只需重命名文件：

```
# mv /etc/hostname.hme0 /etc/hostname.mether0
```

网络接口的正常操作状态是：使用时为“plumbed”，而不使用时为“unplumbed”。当通过 `/etc/hostname.*` 自动配置网络接口时，接口就处于上述任一状态。在手动配置网络接口时，可以将网络状态置于“暂时的plumbed”状态。由于这不是正常操作模式，网络接口不可能处于这种状态。

在 AP 网络配置中，不要将元网络置于这种状态。

只有某个网络元设备和所有此类的其它网络元设备都处于“unplumbed”状态或“plumbed”状态时，才可以删除它。否则，AP 会忽略删除请求，并根据您的配置情况，显示以下格式的警报消息：

```
WARNING:mether_setphyspath: APUNSET busy
WARNING:ap_db_commit: mfddi3 not deleted, metadvice returned
error 16
```

- c. 如果使用的是 FDDI，则必须为元网络指定一个唯一的 MACID。

MACID 可通过 `ifconfig(1M)` 命令的 `ether` 参数进行设置。首先检查每条替代路径的 MACID。可以通过查看每条替代路径并检查 `ether` 字段实现上述操作。然后，创建一个不与任何替代路径相同的 MACID。

---

注意 – 分配媒体访问控制标识符 (MACID) 在《IEEE Std. 802-1990》和《RFC1340, “Assigned Numbers”, July 1992》中有介绍。在生成 AP 网络接口的 MACID 时, 新的 48 位硬件地址应从 IEEE Standards Office (345 East 47th Street, New York, N.Y. 10017) 处获得。但是, 也可以通过将某个元接口替代元件的现有 MACID 的第一个字节加 2 来“创建”一个新标识号 (例如 8:0:20:xx:xx:xx 变为 A:0:20:xx:xx:xx)。创建一个标识号以后, 最重要的是验证同一子网内没有其它硬件是所创建地址的合法用户。

---

如下例所示:

```
#!/sbin/sh
/sbin/ifconfig mfddi0 ether A:0:20:68:6d:62
```

一旦创建了 S19macid, 可使用 chmod 命令将它的属性设置为 744(rwxr--r--)。元网络 MACID 用于配置该元网络的活动物理接口。在执行接口的 AP 切换和 DR 板插入时, 必须要使用此 MACID 以防止在该网络上出现相同的 MACID。缺省情况下, 元网络是活动替代路径引导时的 MACID。要确保在引导时将 MACID 设置正确, 以超级用户的身份, 创建下列文件:

```
/etc/rcs.d/S19macid.
```

```
#!/sbin/sh
/sbin/ifconfig mfddix ether mfddix_macid
```

用正确的元网络设备号替换 mfddix (使用 apconfig-N 命令获取该设备号)。用实际的以太网数目替换 mfddix\_macid。

4. 以常规方式启动元网络，但使用元网络名而不是物理网络名。可通过重新启动计算机或手动配置网络，执行上述操作，如下例所示：

```
# ifconfig mether0 plumb
# ifconfig mether0 inet 136.162.65.30 up netmask + broadcast +
Setting netmask of mether0 to 255.255.255.0
# ifconfig -a
lo0: flags=849<UP,LOOPBACK,RUNNING,MULTICAST> mtu 8232
    inet 127.0.0.1 netmask ff000000
mether0: flags=843<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 136.162.65.30 netmask ffffffff broadcast 136.162.65.255
    ether 0:0:be:0:8:c5
```

此时，设备节点（如 /dev/mether）可通过 Solaris 命令（如 snoop(1M)）访问网络。

## ▼ 切换网络路径组

---

注意 – 即使在网络流量很大时也可以切换网络路径组。

---

- 使用 apconfig(1M) 命令：

```
# apconfig -P mether0 -a hme2
# apconfig -N

metanetwork:    mether0
physical devices:
                 hme0
                 hme2  A
```

其中：

-P 指定该路径组

-a 指定要切换为活动的替代路径。

上表显示主动替代路径已切换为 hme2，因为 hme2 后面的字符 A 表明了这一点。不必提交切换操作。



## ▼ 删除网络路径组和元网络

1. 删除所有对相应元网络的使用操作，并使用带 `-d` 选项的 `apnet(1M)` 命令：

```
# ifconfig mether0 down unplumb
# apnet -d mether0
# apconfig -N

metanetwork:    mether0  D
physical devices:
                 hme0
                 hme2  A
```

在 `apconfig -N` 命令生成的列表中，`mether0` 后面的字母 `D` 表明将该路径组标记为删除。

2. 使用带 `-C` 选项的 `apdb(1M)` 命令提交数据库中的项：

```
# apdb -C
# apconfig -N
#
```

`apconfig-N` 命令不生成列表，表明网络路径组（本例中以前存在的唯一路径组）已被删除。

如果删除未提交，可以取消删除。要撤消删除，可使用 `apnet -z`，指定以前要删除的同一元网络接口。

在使用 `apnet-m -r` 或 `apnet -m -a` 命令时，AP 会将当前的路径组配置标记为已删除，并创建一个新的未提交的路径组定义。

一旦使用 `apdb -C` 命令提交了对数据库的更改，新的定义会替代旧的定义。

3. 删除 `/etc/hostname.metherx` 文件，其步骤如“取消配置元网络”中所述。

## ▼ 取消配置元网络



---

**警告** – 本步骤需要您重新引导计算机。如果还没有准备好重新引导计算机，则不要执行此步骤。

---

---

注意 - IPv6: 在所有的示例中, 均用 `hostname6.xxxx` 替换 `hostname.xxxx` 文件。

---

1. 验证主网络接口是否为 `mether0` (在本例中):

```
# cat /etc/nodename
eng2
# cat /etc/hostname.mether0
eng2
#
```

2. 重命名 `hostname.xxxx` 文件以便在引导时自动配置该网络:

```
# mv /etc/hostname.mether0 /etc/hostname.qfe0
```

3. 重新引导。

## ▼ 重新配置元网络



---

警告 - 本步骤需要您重新引导计算机。如果还未准备好重新引导计算机, 则不要执行此步骤。

---

---

注意 - IPv6: 在所有的示例中, 均用 `hostname6.xxxx` 替换 `hostname.xxxx` 文件。

---

1. 验证主网络接口是否为 `qfe0` (在本例中):

```
# cat /etc/nodename
eng2
# cat /etc/hostname.qfe0
eng2
#
```

2. 重命名 `hostname.xxxx` 文件以便在引导时自动配置该网络:

```
# mv /etc/hostname.qfe0 /etc/hostname.mether0
```

### 3. 重新引导。

---

## 使主网络接口具有替代路径

位于 Sun Enterprise 服务器和网络上其它计算机之间的主网络接口是用于传递与该服务器主机名相关联的地址的接口。标识主网络的一种方法是查找 `/etc/hostname.metherx` 文件直至找到所需的网络，该网络必须包含与 `/etc/nodename` 文件中所找到的主机名匹配的主机。所找到的对应 `metherx` 网络（例如 `mether0`）即为主网络。

可以使主网络拥有替代路径。主网络是唯一可以在引导时自动切换的网络接口。引导过程中，如果主网络的活动替代路径出现故障，系统会尝试为该网络寻找工作正常的替代路径。

配置具有替代路径的网络时，千万不要在主驱动程序处于活动时配置元网络。

在为当前正在使用的网络配置 AP 时，在取消对网络接口的配置并重新配置 AP 接口的过渡期内，会无法使用 Sun Enterprise 服务器的网络服务。

执行此转换时，应使用以下任一步骤，这些步骤按优先顺序显示：

- 创建相应的 AP 数据库项，再创建一个新的 `/etc/hostname.xxx` 文件，并删除（或重命名）对应的 `/etc/hostname.xxx` 文件，然后重新引导 Sun Enterprise 服务器。此步骤的详细示例，请参见下一节：“为当前网络配置 AP”。
- 设置一个脚本文件执行 Sun Enterprise 服务器上的转换。
- 用另一个网络接口登录到 Sun Enterprise 服务器以便无法在将置于 AP 控制下的网络接口上使用网络服务时，能够启用要执行的命令。

---

## 为当前网络配置 AP

以下示例显示了为当前使用的主网络配置 AP 的首选步骤。本示例假定已有一个名为 `eng5` 的 Sun Enterprise 服务器，它的主网络接口位于 `mether0`，并假定您想有一个由 `qfe0` 和 `hme2` 组成的元网络接口。如果无法确信元网络应由哪些网络接口组合组成，可以使用 `snoop -d` 以确定同一个子网上配置了哪些网络。

## ▼ 为主网络创建网络路径组和元网络



---

**警告** – 本步骤需要您重新引导计算机。如果还没有准备好重新引导计算机，则不要执行此步骤。

---

---

**注意** – IPv6: 在所有的示例中，均用 `hostname6.xxxx` 替换 `hostname.xxxx` 文件。

---

1. 验证主网络接口是否为 `qfe0`:

```
# cat /etc/nodename
eng5
# cat /etc/hostname.qfe0
# eng5
```

2. 创建新的网络路径组并提交所做的更改:

```
# apnet -c -a qfe0 -a hme2
# apdb -C
```

3. 查看 AP 数据库中已提交的网络项，验证新路径组:

```
# apconfig -N
metanetwork:    mether0
physical devices:
                qfe0  A
                hme2
```

4. 重命名 `hostname.xxxx` 文件以便在引导时自动配置该网络:

```
# mv /etc/hostname.qfe0 /etc/hostname.mether0
```

5. 通过重新启动计算机来取消对元网络接口的配置并配置好物理网络接口。

## ▼ 删除主网络的网络路径组和元网络



---

**警告** – 本步骤需要您重新引导计算机。如果还没有准备好重新引导计算机，则不要执行此步骤。

---

---

**注意** – IPv6: 在所有的示例中，均用 `hostname6.xxxx` 替换 `hostname.xxxx` 文件。

---

1. 验证主网络接口是否为 `mether0`（在本例中）:

```
# cat /etc/nodename
eng5
# cat /etc/hostname.mether0
eng5
```

2. 重命名元网络接口的配置文件:

```
# mv /etc/hostname.mether0 /etc/hostname.qfe0
```

3. 重新引导。
4. 删除 AP 数据库中的项:

```
# apnet -d mether0
# apdb -C
# apconfig -N
#
```

## ▼ 取消配置主网络的元网络



---

**警告** – 本步骤需要您重新引导计算机。如果还没有准备好重新引导计算机，则不要执行此步骤。

---

---

**注意** – IPv6: 在所有的示例中，均用 `hostname6.xxxx` 替换 `hostname.xxxx` 文件。

---

1. 验证主网络接口是否为 `me0`（在本例中）：

```
# cat /etc/nodename
eng5
# cat /etc/hostname.mether0
eng5
```

2. 重命名 `hostname.xxxx` 文件以便在引导时自动配置该网络：

```
# mv /etc/hostname.mether0 /etc/hostname.qfe0
```

3. 重新引导。

## ▼ 重新配置主网络的元网络



---

**警告** – 本步骤需要您重新引导计算机。如果还没有准备好重新引导计算机，则不要执行此步骤。

---

---

**注意** – IPv6: 在所有的示例中，均用 `hostname6.xxxx` 替换 `hostname.xxxx` 文件。

---

1. 验证主网络接口是否为 qfe0（在本例中）：

```
# cat /etc/nodename
eng5
# cat /etc/hostname.qfe0
eng5
```

2. 重命名 `hostname.xxx` 文件以便在引导时自动配置该网络：

```
# mv /etc/hostname.qfe0 /etc/hostname.mether0
```

3. 重新引导。





# AP 和 DR 之间的交互作用

---

本章介绍了 Alternate Pathing (AP) 和 Dynamic Reconfiguration (DR) 之间的关系。

---

## 一起使用 DR 和 AP

Dynamic Reconfiguration 和 Alternate Pathing 可以紧密地合作。DR 可以在不中止操作系统的情况下挂接和分离系统板，具体信息如 《*Sun Enterprise 10000 Dynamic Reconfiguration User Guide*》中所述。AP 使您可以 *切断* 对正在分离的系统板上控制器的使用，或 *切换* 到已挂接的系统板上的控制器。

在 Sun Enterprise 10000 服务器上，AP *自动* 切换在要分离的系统板上拥有活动控制器的每个磁盘和网络元设备（假定在另一块系统板上存在替代路径）。同时，在 Sun Enterprise 10000 服务器上，AP 会阻止您手动切换到在 DR 分离操作中已处于隔离状态的系统板上的控制器。

---

**注意** – 在 T3 磁盘上进行任何 DR 操作之前，应先使用如下的 AP 切换命令 *手动* 禁用路径优化：apconfig(1M)。

---

在 Sun Enterprise 服务器而不是 Sun Enterprise 10000 服务器上，在分离系统板之前必须 *手动* 切换磁盘和网络元设备（如有必要）。

下列 AP 命令是显示要分离的系统板上的 sf:1 控制器（如 DE 标志所示），因此，不能切换到该控制器：

```
# apconfig -S

c1      sf:0  P A
c2      sf:1  DE
        metadiskname(s) :
                mclt5d0
                mclt4d0
                mclt3d0
                mclt2d0
                mclt1d0
```

同样，下列 AP 命令显示了处于隔离状态的系统板的 sf:1 控制器（如 DR 标志所示），因此，不能切换到该控制器：

```
# apconfig -S

c1      sf:0  P A
c2      sf:1  DR
        metadiskname(s) :
                mclt5d0
                mclt4d0
                mclt3d0
                mclt2d0
                mclt1d0
```

只在 Sun Enterprise 10000 服务器上，会通知 AP，系统板处于 DR 隔离状态。

要分离驻留有路径组的活动控制器所在的系统板时，可以在 DR 分离操作前，甚至在 DR 分离操作中，手动切换到另一块系统板的控制器上。

但是，对于除 Sun Enterprise 10000 服务器以外的计算机，必须在分离操作将要完成或分离操作要失败时执行上述切换操作。在这种情况下，可以先执行切换，再重试分离操作。

---

**注意** – 即使 AP 不能立即访问板，DR 挂接操作也可以完成。在使用 apconfig(1M) 命令切换到新板之前，必须验证此物理设备是否存在。

---

有关 DR 的更多信息，请参考《*Sun Enterprise 10000 Dynamic Reconfiguration User Guide*》或《*Sun Enterprise 6x00, 5x00, 4x00, and 3x00 Systems Dynamic Reconfiguration User Guide*》。

---

## 维护正确的 AP 状态

对于除 Sun Enterprise 10000 服务器以外的计算机，如果挂接或分离一块驻留磁盘或网络路径组的 I/O 控制器所在的系统板，必须运行 `apconfig -F` 命令。此命令设置或清除该板的分离标志 (DE)，以便该标志可以正确指明是要挂接还是要分离该系统板。

在 Sun Enterprise 10000 服务器上，在挂接或分离操作后无须使用 `apconfig -F` 命令，因为在 DR 操作完成后会自动设置或清除 DE 标志。

如果要分离包含有网络控制器的系统板，且自从上一次引导后还未使用该网络设备，那么必须运行 `apconfig -F` 命令以通知系统该网络设备不再可用。

如果还没有加载相应的 AP 元驱动程序，`apconfig -N` 命令可能会错误地指明某个网络控制器驻留在已分离的系统板上（或错误地指明它驻留在挂接的系统板上）。使用 `apconfig -F` 命令可以确保 `apconfig -N` 显示的信息是正确的。



# AP 组件

---

AP 包含以下组件：

- AP 命令 - 控制各种进程和 AP 选项的程序指令。
- AP 管理程序 - ap(7D) 管理 AP 数据库并在必要时与元驱动程序交互作用。它通过 `ioctl`s 接收请求并通过更新数据库或调用元驱动程序中的入口点处理这些请求。
- AP 元驱动程序 - 将在元驱动程序中重新路由 I/O 访问替代路径的低层功能。

所有使用正确元磁盘的应用程序 I/O 请求，都会通过元驱动程序，将这些请求传送到物理设备驱动程序上。因此，元驱动程序可以确定要使用哪条物理路径，给定的路径是否不再有效等等。元驱动程序依据 AP 管理程序和 AP 数据库的信息做出相应的决定。



# AP 手册页

---

AP 手册页位于 Sun Enterprise 服务器文档资料集的 《*Alternate Pathing 2.3.1 Reference Manual*》 部分，以及联机文档资料中（在安装 AP 软件包之后）。以下是 AP 手册页列表：

- `ap(1M)` – Alternate Pathing 概述
- `apboot(1M)` – 定义一个 AP 引导设备
- `apcheck(1M)` – 确定 AP SCSI 元设备的可用性
- `apconfig(1M)` – 显示并管理 AP 配置
- `apdb(1M)` – 管理 AP 数据库的副本
- `apdisk(1M)` – 管理 SCSI 磁盘的 AP
- `apinst(1M)` – 标识 SCSI 总线控制器
- `apnet(1M)` – 管理网络的 AP
- `ap(7D)` – AP 驱动程序
- `ap_dmd(7D)` – AP 磁盘元驱动程序
- `methernet(7D)` – AP 网络元驱动程序
- `mfddi(7D)` – AP 网络元驱动程序





## 驱动程序层

下图（和括号中的示例）显示了在 AP 控制磁盘设备时使用的驱动器层。

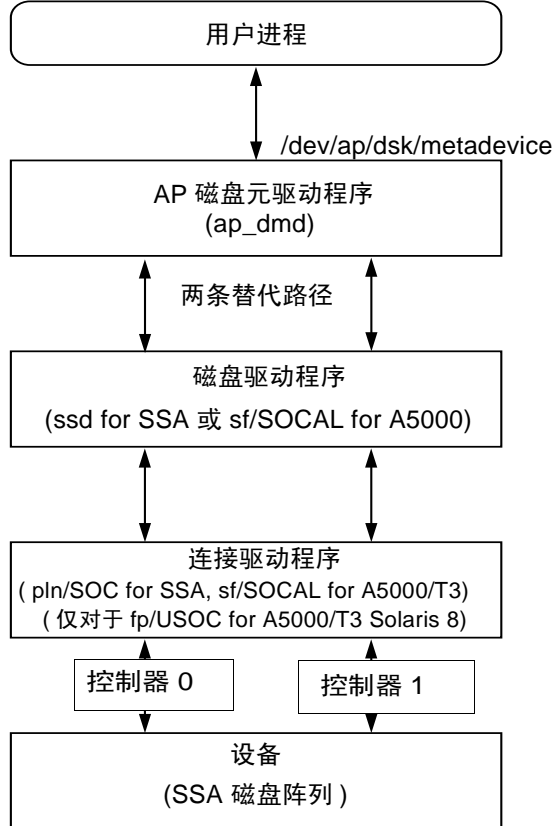


图 C-1 AP 磁盘驱动程序层

用户进程引用元磁盘，从而能够访问 AP 磁盘元驱动程序。AP 磁盘元驱动程序控制了物理磁盘驱动程序的两个实例，而该驱动程序又控制连接驱动程序（或控制器驱动程序）的两个实例。连接驱动程序控制物理设备。

图 C-2 显示了在 AP 控制网络时使用的驱动程序层。括号中是驱动程序层的示例。

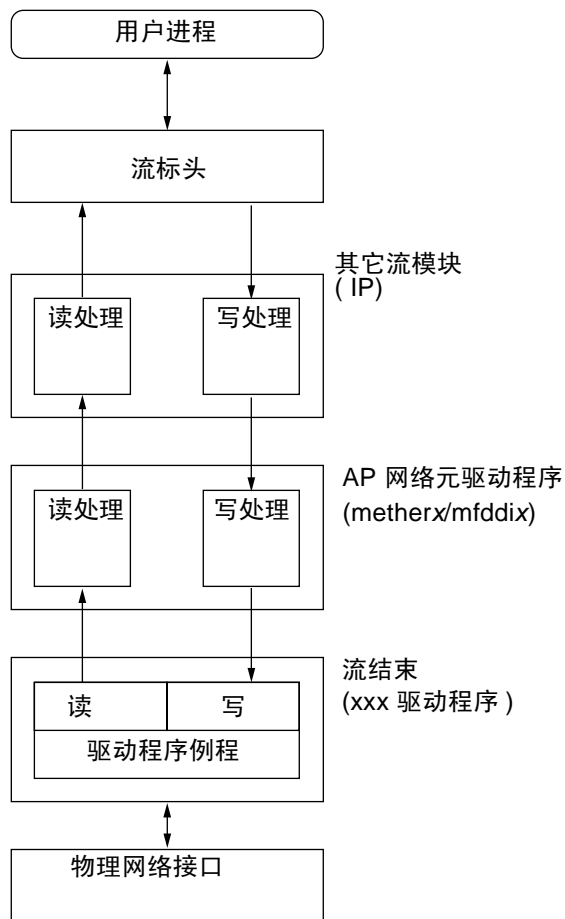


图 C-2 AP 网络驱动程序层

用户进程引用元网络，从而能够访问流标头。AP 网络元驱动程序被插入到高级读 / 写处理组件和物理驱动程序例程之间的流中。

# 词汇表

---

<b>活动替代路径</b>	当前为路径组处理 I/O 的替代路径。
<b>AP 数据库 (或简称数据库)</b>	一种由 AP 子系统维护的数据库。AP 数据库包含维护配置替代路径所需的全部信息。
<b>替代路径</b>	路径组内的一条物理路径。参见主路径。
<b>已提交的数据库项</b>	当前正在被 AP 使用的 AP 数据库项，用以管理对磁盘或网络的访问。(与 <i>未提交的数据库项</i> 相对而言。)
<b>磁盘阵列</b>	硬件外设内的磁盘系列。磁盘阵列能够通过一个或两个光纤通道模块访问它包含的磁盘。
<b>磁盘阵列控制器</b>	一个驻留在主机系统的控制器，它有一个或两个光纤通道模块。
<b>磁盘阵列端口</b>	可连接到磁盘阵列控制器的光纤通道模块，该磁盘阵列控制器由一个驱动程序对维护；例如 SSA 的 soc/pln。
<b>光纤通道模块</b>	磁盘阵列控制器上的光纤链路连接 (OLC) 模块，可以连接磁盘阵列端口。
<b>元磁盘</b>	可以访问磁盘的两条基本物理路径组的抽象磁盘。
<b>元网络</b>	可以访问网络的两条基本物理路径组抽象网络。
<b>最佳路径</b>	参见 <i>路径优化</i> 。
<b>路径组</b>	两条替代路径的组合，能够访问同一台设备或设备集。

<b>路径优化</b>	对特定设备的 I/O 流量的有效分配。
<b>物理路径</b>	从主机到磁盘或网络的电子路径。
<b>主路径</b>	磁盘路径组中最初为活动替代路径的替代路径。在创建元磁盘的名称时要使用主路径的名称。出现切换时，不更改主路径。参见替代路径。
<b>SSA</b>	一种 SPARCstorage 阵列，是硬件外设内的磁盘系列。SSA 能够使用两个端口访问它所包含的每一张磁盘。
<b>切换</b>	为指定的路径组创建新活动路径的操作。切换不会更改主路径。
<b>T3</b>	挂接到主机的 Sun StorEdge 插槽通过 FC-AL 连接到 PCI HBA，或使用 GBIC 适配器连接到 SBus HBA。AP 2.3.1 对 T3 的 I/O 路径分配进行了优化。
<b>未提交的数据库项</b>	AP 数据库中还未提交的项，因此当前也未生效。如果已创建了路径组但还未提交数据库项，则 AP 当前无法使用该路径组管理对磁盘或网络的访问。如果已删除了以前提交的路径组，但还未提交数据库项，AP 仍可使用该路径组管理对磁盘或网络的访问。

# 索引

---

## 符号

A (活动替代指示符), 30

Alternate Pathing (AP)

和 Dynamic Reconfiguration (DR), 3

和 Dynamic Reconfiguration (DR), 59

Alternate Pathing 命令, 65

Alternate Pathing 数据库, 13

AP

路径优化, 2, 3, 6, 7, 22, 31, 33, 35, 59

手动切换, 22

自动切换, 23

AP 管理程序, 63

AP 元驱动程序, 63

单用户模式, 43

DR 交互作用, 13, 59

分离系统板, 3

挂接系统板, 3

管理状态, 61

命令列表, 65

配置, 典型, 10

手动切换, 3

引导顺序, 43

域, 12

与磁盘镜像, 11

支持的 Solaris 环境, 10

apboot 示例, 39, 42

apboot -m, 42

apboot -u, 42

apconfig -P -a -a

重新启用路径优化, 22

apconfig 示例

apconfig -D, 16, 17

apconfig -N, 19, 47, 50, 51, 54

确保它显示正确的信息, 61

apconfig -N -u, 19, 46

apconfig -P -a, 35, 50

禁用路径优化, 33

apconfig -S, 18, 23, 31, 33, 34, 35

apconfig -S -u, 18, 30

apdb 示例, 47

apdb -C, 31, 37, 51, 54

apdb -c -f, 15

apdb -d -f, 16

AP 的用途, 1

apdisk 示例

apdisk -c -p -a, 30

apdisk -d, 36

apdisk -w, 23

apdisk -z, 37

AP 和 DR, 3

AP 和 DR 之间的交互作用, 59

apinst 示例, 29

AP 简介, 1

AP 命令

/usr/sbin 与 /sbin, 43

命令列表, 65

apnet 和撤消删除, 51

## apnet 示例

apnet -c -a -a, 46, 54

apnet -d, 51

AP 与磁盘镜像, 11

AP 支持的设备, 10

AP 状态, 管理, 61

bin, /usr/sbin 与 /sbin, 43

DE (已分离) 标志, 60

清除, 61

## DR

手动切换, 22

AP 交互作用, 13, 59

隔离状态, 59

和自动切换, 59

切换路径, 59

DR (隔离状态) 标志, 60

drvconfig 示例, 32

Dynamic Reconfiguration (DR)

和 Alternate Pathing (AP), 3, 59

## FDDI

和 MACid, 48

切换路径组, 50

元网络名称, 46

FDDI 的 MACid, 48

forcing (-f) 数据库

删除, 16

I/O 控制器 (已定义), 4

I/O 设备 (已定义), 4

LE 元网络名称, 46

P (主路径指示符), 30

pkgrm 和 AP, 43

plumb 网络, 50

T (经过试验的) 标志, 23

unplumb 网络, 51

vfstab, 修改, 39

## B

标识

主网络, 53

标志, 经过试验的, 23

不可访问的数据库, 确定是否, 17

不可用 (经过试验的) 路径, 23

## C

查看

路径组信息, 17

提交的数据库项

对于网络, 19

磁盘, 18

未提交的数据库项

磁盘, 18

对于磁盘, 30

对于网络, 19

已提交的数据库项

对于磁盘, 31

撤消删除, 37, 51

重新分区, 不由 AP 执行, 21

创建

数据库, 14, 15

网络路径组, 46

元设备, 29

磁盘

AP 下的引导磁盘, 39

磁盘路径组, 7

镜像引导磁盘

脱离 AP 控制, 42

引导磁盘

脱离 AP 控制, 42

引导磁盘, AP 和镜像, 41

元磁盘, 5

元磁盘路径, 5

自动故障修复, 3

自动切换, 22

磁盘镜像

与 AP, 11

磁盘路径组, 7

与元磁盘比较, 7

## D

单用户模式

调用它的原因, 43

和 AP, 43  
典型的 AP 配置, 10

## F

分离系统板和 AP, 3  
复位经过试验的标志, 23

## G

隔离状态 (DR) 标志, 60  
挂接系统板和 AP, 3  
管理 AP 状态, 61  
管理程序, AP 管理程序, 63  
故障修复, 自动, 3

## H

活动替代  
    指示符 (A), 30  
活动替代路径, 7

## J

节点  
    定义, 4  
    示例, 5  
接口  
    元网络接口, 45  
经过试验的标志, 23  
    复位经过试验的标志, 23  
镜像引导磁盘  
    脱离 AP 控制, 42  
镜像引导磁盘, AP 和, 41

## K

控制器 (已定义), 4

## L

路径  
    不可用 (经过试验的), 23  
    对于元设备, 22  
    切换之前验证, 32  
    确定端口, 29

元磁盘的, 5  
    在 DR 过程中切换, 3  
路径优化, 2, 3, 31, 59  
    重新启用, 6, 9, 22, 35  
    禁用, 3, 6, 22  
    缺省情况, 33  
    手动切换, 3  
    自动切换, 2, 23

## 路径组

    标识要切换的路径组, 34  
    查看路径组信息, 17  
    创建网络路径组, 46  
    磁盘路径组, 7  
    磁盘路径组, 使用, 29  
    删除网络路径组, 51  
    网络路径组, 9, 45

## M

命令  
    /usr/sbin 与 /sbin, 43  
    列表, 65

## P

配置, 典型, 10  
配置具有替代路径的网络, 53

## Q

启动网络, 50  
切换  
    从主要路径切换到替代路径, 32  
    DR 过程中处于隔离状态, 59  
    切换操作 (已定义), 7  
    切换网络路径组 (以太网或 FDDI), 50  
    切换元磁盘, 自动, 22  
    示例 (对磁盘而言), 33  
    手动切换和 DR, 22  
    验证路径, 32  
    引导时自动切换, 43  
    自动和 DR, 59  
    自动切换和 DR, 59  
切换之前验证路径, 32

清除 DE（已分离）标志，61

驱动程序

AP 元驱动程序，63

## R

软件包，删除 AP 软件包，43

## S

删除

磁盘路径组，36

数据库，16

网络路径组，51

已提交的 / 未提交的数据库项，37

删除，撤消，37, 51

设备（已定义），4

设备节点

定义，4

示例，5

物理设备节点，21

元设备的设备节点，22

识别

替代路径，7

使用 ifconfig 取消 unplumb 和 AP，47

手动切换

DR，3

路径优化，3, 22

数据库

不可访问，确定是否，17

查看

时间戳，17

数据库信息，17

提交的项，18

提交的项（对于网络而言），19

未提交的项，18

未提交的项（对于网络而言），19

未提交的项（对于元磁盘），30

创建数据库，14, 15

forcing (-f) 删除数据库，16

删除数据库，16

时间戳，17

数据库大小，推荐，13

数据库分区推荐方法，13

数据库副本，数目，13

损坏，确定是否，17

提交的项，17

未提交的项，17

用于创建数据库的原始磁盘分片，15

用于删除数据库的原始磁盘分片，16

数据库上的时间戳，查看

时间戳，17

损坏数据库，确定是否，17

## T

替代路径，2, 7

识别，7

网络配置，53

指定，30

主网络，54

替代路径的端口，确定，29

提交的数据库项，17

查看，18

网络项，查看，19

图例

AP 和磁盘镜像，11

磁盘路径组，8

典型的 AP 配置，10

具有替代路径的 I/O 设备，2

元磁盘，6

脱离

镜像引导磁盘脱离 AP 控制，42

引导磁盘脱离 AP 控制，42

## W

网络

Alternately Pathing 主网络，54

多个网络和 AP，45

配置具有替代路径的网络，53

启动网络，50

切换路径组（以太网或 FDDI），50

确保 apconfig -N 显示的信息正确，61

删除

直接使用物理路径，47

删除物理网络接口的配置文件，54



- 删除主网络的 Alternate Pathing, 51, 52, 55, 56
- unplumb, 51
- 网络路径组, 9
  - 创建网络路径组, 46
  - 删除网络路径组, 51
- 元网络, 6
- 元网络接口, 6, 45
- 主网络
  - 标识, 53
  - 注意事项, 53
- 网络路径组, 9, 45
  - 创建, 46
  - 删除, 51
- 未经 AP 修改的数据, 21
- 未提交的数据库项, 17
  - 查看, 18
    - 对于网络, 19
    - 对于元磁盘, 30
  - 删除, 37
- 文件
  - /etc/system, 39
  - /etc/vfstab, 39
  - /etc/hostname.xxx, 47, 53
  - /etc/hostname.xxxx, 53
  - /etc/nodename, 53
  - /etc/system, 42
  - /etc/vfstab, 42
  - hostname.xxxx, 52, 54, 56, 57
- 物理路径, 4
  - 删除直接使用 (网络), 47
- 物理设备节点
  - 概述, 21
  - 引用, 修改 AP, 32
- 物理网络接口
  - 删除配置文件, 54
- 无人参与的引导, 概述, 10

## X

- 系统
  - 通知系统网络设备不可用, 61
- 系统 (/etc/system), 修改, 39
- 修改 vfstab, 39

## Y

- 已分离的 (DE) 标志, 60
  - 清除, 61
- 引导, 无人参与的, 10
- 引导磁盘
  - AP 和引导磁盘, 39
  - 镜像和 AP, 41
  - 脱离 AP 控制, 42
- 引导时, 自动切换, 43
- 引导时自动切换, 概述, 43
- 引导顺序, 43
- 引用设备节点
  - 修改 AP, 32
- 引用设备节点, 修改 AP, 32
- 以太网
  - 切换路径组, 50
  - 元网络名称, 46
- 已提交的数据库项, 17
  - 查看, 18
  - 磁盘项, 查看, 31
  - 删除, 37
  - 网络项, 查看, 19
- 元磁盘, 5
  - 查看未提交的数据库项 (对于元磁盘), 30
  - 使用元磁盘, 29
  - 修改物理设备节点引用, 32
  - 元磁盘与磁盘路径组, 7
- 元设备, 21
  - 创建, 29
  - 设备节点, 概述, 22
- 原始磁盘分片
  - 用于创建数据库, 15
  - 用于删除数据库, 16
- 元网络, 6, 45
  - 接口, 6, 45
  - 元网络接口, 6
  - 元网络名称, 46
- 域和 AP, 12

## Z

在 DR 过程中切换路径, 3

支持的设备, 10

支持的 Solaris 版本, 10

主路径

标识路径组, 34

指定, 30

指示符 (P), 30

主路径定义, 8

主网络

Alternately Pathing, 54

标识, 53

和 AP, 53

删除 Alternate Pathing, 51, 52, 55, 56

自动故障修复, 2

自动切换, 3, 23

元磁盘, 22

在 DR 过程中, 59