



Sun StorageTek™ 5800 System

概述

Sun Microsystems, Inc.
www.sun.com

文件号码 820-5050-10
2008 年 5 月, 修订版 A

请将有关本文档的意见和建议提交至: <http://www.sun.com/hwdocs/feedback>

版权所有 © 2008 Sun Microsystems, Inc., 4150 Network Circle, Santa Clara, California 95054, U.S.A. 保留所有权利。

Sun、Sun Microsystems、Sun 徽标、Solaris 和 Sun StorageTek 是 Sun Microsystems, Inc. 在美国和其他国家/地区的商标或注册商标。

本服务手册所介绍的产品以及所包含的信息受美国出口控制法制约，并应遵守其他国家/地区的进出口法律。严禁将本产品直接或间接地用于核设施、导弹、生化武器或海上核设施，也不能直接或间接地出口给核设施、导弹、生化武器或海上核设施的最终用户。严禁出口或转口到美国禁运的国家/地区以及美国禁止出口清单中所包含的实体，包括但不限于被禁止的个人以及特别指定的国家/地区的公民。

对任何备用或替换 CPU 的使用仅限于对遵照美国出口法律出口的产品中的 CPU 进行修复或一对一的替换。除非经过美国政府授权，否则，严禁使用 CPU 进行产品升级。

本文档按“原样”提供，对于所有明示或默示的条件、陈述和担保，包括对适销性、适用性或非侵权性的默示保证，均不承担任何责任，除非此免责声明的适用范围在法律上无效。



请回收



Adobe PostScript

目录

前言 ix

产品概述 1

关于 5800 系统 1

5800 系统硬件 2

全单元和半单元配置 2

存储节点 7

服务节点 10

千兆位以太网交换机 12

网络配线架 13

5800 系统软件 14

捆绑软件 14

元数据 14

系统元数据 15

扩展元数据 15

布局算法 15

用户界面 15

应用程序接口 (Application Program Interface, API) 16

WebDAV 16

CLI 和 GUI 17

索引 19



图 1	5800 系统全单元前视图	3
图 2	双单元系统的后视图（显示了网络配线架连接）	4
图 3	5800 系统半单元前视图	6
图 4	存储节点前面板组件	8
图 5	存储节点后面板组件	10
图 6	服务节点前面板组件	11
图 7	服务节点后面板组件	12
图 8	千兆位以太网交换机	13
图 9	网络配线架	13

表

表 1	存储节点特性	7
表 2	存储节点前面板 LED 指示灯和开关描述	9
表 3	服务节点特性	10
表 4	千兆位以太网交换机与网络配线架的连接	13
表 5	5800 存储系统用户界面功能	16

前言

《Sun StorageTek 5800 System 概述》提供了有关 Sun StorageTek 5800™ System 的特性和功能的信息。其中介绍了可能的单元配置、构成 5800 系统的硬件组件，以及可用于访问和管理系统的多种用户界面。

印刷约定

字体	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 ls -a 列出所有文件。 % You have mail.
AaBbCc123	用户键入的内容，与计算机屏幕输出的显示不同	% su Password:
AaBbCc123	保留未译的新词或术语以及要强调的词。要使用实名或值替换的命令行变量。	这些称为 class 选项。 要删除文件，请键入 rm filename 。
新词术语强调	新词或术语以及要强调的词。	您 必须 成为超级用户才能执行此操作。
《书名》	书名	阅读 《用户指南》的第 6 章。

注 — 字符的显示会因浏览器设置的不同而异。如果字符无法正确显示，请将浏览器的字符编码更改为 Unicode UTF-8。

相关文档

下表列出了本产品的相关文档。您可以从以下位置获得所列出的联机文档：

<http://docs.sun.com/app/docs/prod/stortek.5800>

书名	文件号码	格式	所在位置
《Sun StorageTek 5800 System 1.1.1 发行说明》	820-5062-xx	PDF HTML	联机提供
《Sun StorageTek 5800 System Regulatory and Safety Compliance Manual》	819-3809-xx	PDF HTML	联机提供
《Sun StorageTek 5800 System 场地准备指南》	820-3723-xx	PDF HTML	联机提供
《Sun StorageTek 5800 System 管理指南》	820-5056-xx	PDF HTML	联机提供
《Sun StorageTek 5800 System API Reference Manual》	820-4796-xx	PDF HTML	联机提供
《Sun StorageTek 5800 System SDK Reference Manual》	820-4797-xx	PDF HTML	联机提供

文档、支持和培训

Sun 提供的服务	URL
文档	http://www.sun.com/documentation/
支持	http://www.sun.com/support/
培训	http://www.sun.com/training/

Sun 欢迎您提出意见

Sun 致力于提高其文档的质量，并十分乐意收到您的意见和建议。您可以通过以下网址提交您的意见和建议：

<http://www.sun.com/hwdocs/feedback>

请在您的反馈信息中包含文档的书名和文件号码：

《Sun StorageTek 5800 System 概述》，文件号码 820-5050-10。

产品概述

本文档对 Sun StorageTek™ 5800 System 进行了概述。它包含以下各节：

- [第 1 页中的“关于 5800 系统”](#)
 - [第 2 页中的“5800 系统硬件”](#)
 - [第 14 页中的“5800 系统软件”](#)
-

关于 5800 系统

5800 系统是自包含存储单元，直接与网络连接。它不使用典型的数据和文件结构：没有要设置或管理的池、卷、逻辑单元号 (logical unit number, LUN) 或独立磁盘冗余阵列 (redundant arrays of independent disk, RAID)。

相反，5800 系统使用面向对象的方法存储固定内容的数据文件（从不进行修改的文件），如同它们是独立对象一样。存储系统根据这些**数据对象**的属性为每个数据对象指定一个称为对象 ID (Object ID, OID) 的唯一标识符。应用程序使用该 OID 查询和检索数据对象。

5800 系统使用**存储节点**的群集。每个存储节点都是单个服务器，其中具有 CPU 处理能力、RAM 和四个用于存储的串行高级技术附件 (Serial Advanced Technology Attachment, Serial ATA) 磁盘驱动器。每个存储节点都具有与其他存储节点相同的硬件和软件，但是各个存储节点均独立运行。该配置允许根据整个系统的可用处理能力来分布所有存储处理和数据路径操作，从而可增强系统的可靠性和性能。

每个 5800 系统都包括一个具有预配置软件和固件的**服务节点**。系统使用服务节点进行初始配置、故障排除和升级系统软件。

基本 5800 系统是一个**全单元配置**，包含 16 个存储节点、1 个服务节点、2 个千兆位以太网交换机、一个网络配线架 (patch panel)，以及预先安装的操作系统和软件。此外，还允许使用**半单元配置**，该配置只包含 8 个存储节点。可以将半单元配置扩展（缩放）为全单元配置。还可对全单元配置进行扩展，以创建**多单元配置**，也称为**蜂箱 (hive)**。在多单元配置中仅允许使用全单元。

5800 系统的特性包括：

- 使用命令行界面 (command-line interface, CLI) 和图形用户界面 (graphical user interface, GUI) 来监视系统性能和状态并执行管理任务。
- 可以对与各个数据对象相关联的元数据进行自定义，以：
 - 模仿数据库查询。利用应用程序编程接口 (application programming interface, API)，开发者可以编写用于存储、检索、查询和删除数据的应用程序。
 - 使用基于 Web 的分布式创作和版本控制 (Web-based Distributed Authoring and Versioning, WebDAV) 工具呈现模拟的分层目录结构。（多单元配置不支持 WebDAV。）
- 每个半单元或全单元具有两个虚拟 IP 地址：一个用于数据处理，另一个用于管理操作。集成的域名服务 (Domain Name Service, DNS) 可建立主机名。
- 可提高系统可靠性的自我修复操作。如果一个磁盘驱动器发生故障，系统会在系统中的其他磁盘上重新构建数据。
- 在系统运行时可轻松移除和更换的可热交换磁盘。
- 分布式数据存储模型，可利用 RAID 系统中常用的 Reed Solomon 算法来提高系统可靠性。
- 集成 Sun Solaris 10 操作系统 (Solaris 10 Operating System, Solaris OS)。

5800 系统硬件

5800 系统出厂时即完整安装在机架中，并已在默认配置中装入软件。

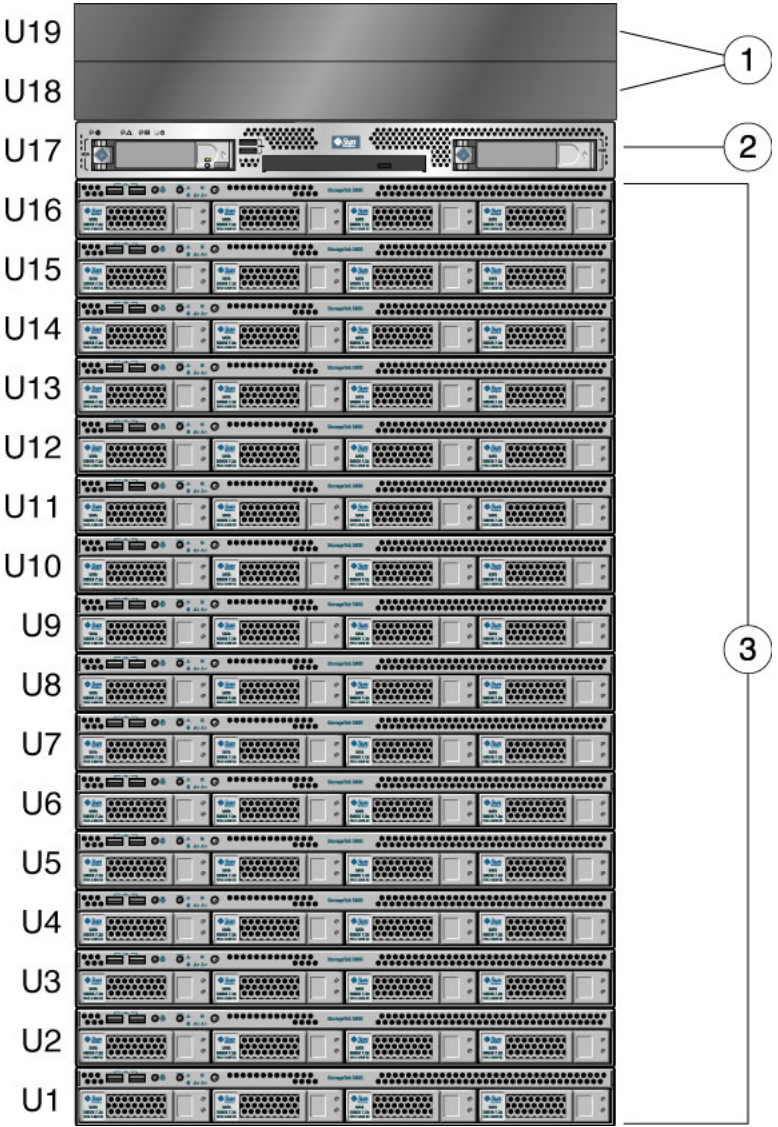
本节包含以下主题：

- [第 2 页中的“全单元和半单元配置”](#)
- [第 7 页中的“存储节点”](#)
- [第 10 页中的“服务节点”](#)
- [第 12 页中的“千兆位以太网交换机”](#)
- [第 13 页中的“网络配线架”](#)

全单元和半单元配置

全单元是 5800 系统的基本构建块。全单元包含一个服务节点、16 个存储节点、2 个千兆位以太网交换机和一个网络配线架。[图 1](#) 显示了一个全单元的前视图。多单元系统（称为**蜂箱**）中的其他全单元都是相同的。

图 1 5800 系统全单元前视图



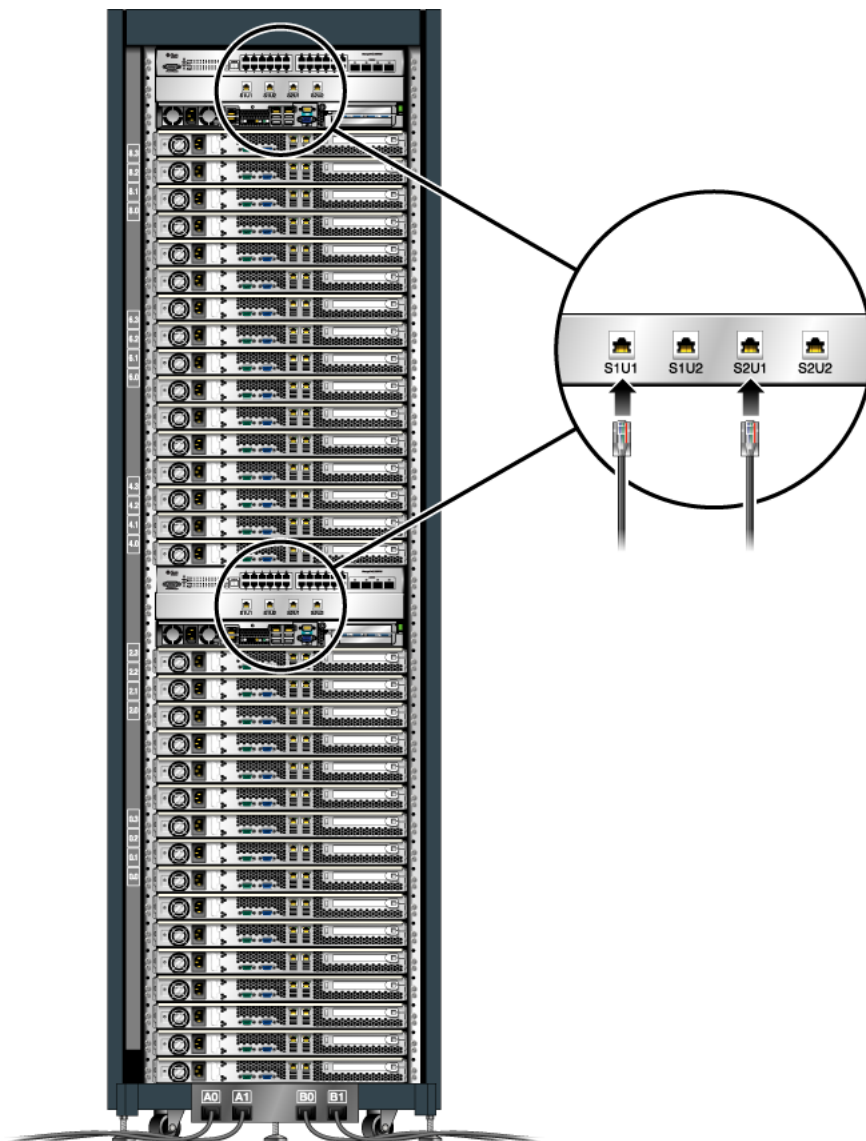
图例

- 1 填充面板位于两个朝后的千兆位以太网交换机的前面，在靠下的交换机后面有一个朝后的网络配线架
- 2 服务节点
- 3 16 个存储节点

安装了千兆位以太网交换机，其端口朝向机柜的后部。填充面板安装在机柜的前部，与两个交换机的高度相同。网络配线架安装在机柜后部靠下的交换机后面（请参见图 2）。这两个千兆位以太网交换机都与服务节点、所有的存储节点和网络配线架相连。

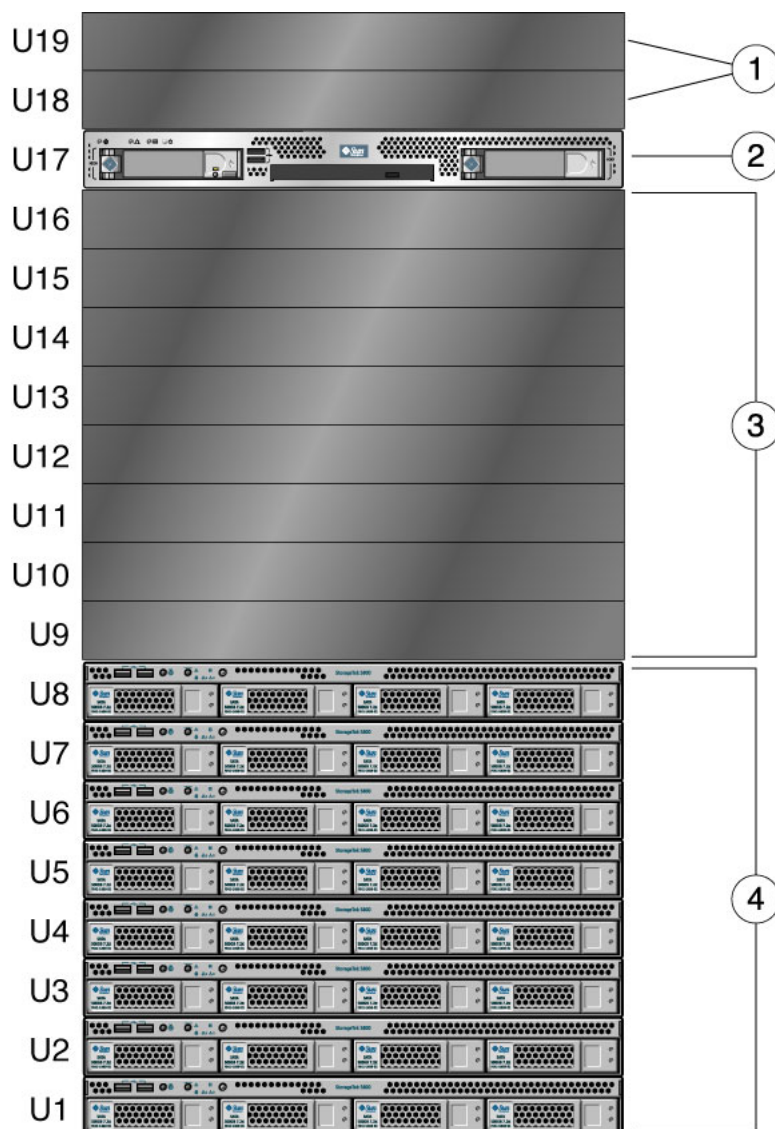
编号为 101 的存储节点位于单元的底部，存储节点的编号随着单元的升高而递增，节点 102 位于节点 101 的上方，再靠上则是节点 103，依此类推。

图 2 双单元系统的后视图（显示了网络配线架连接）



如果您选择仅使用半单元（八个存储节点）来填充机柜，则在插槽 U9 至 U16 的前面使用填充面板，如图 3 中所示。半单元配置为独立配置，不能与另一个全单元组合使用。要向蜂箱添加另一个全单元，必须首先添加其他存储节点，以将半单元扩展为全单元配置。因为半单元的存储节点数量较少，因此它的固有可靠性不如具有 16 个存储节点的全单元。

图 3 5800 系统半单元前视图



图例

-
- 1 填充面板位于两个朝后的千兆位以太网交换机的前面，在靠下的交换机后面有一个朝后的网络配线架
 - 2 服务节点
 - 3 8 个填充面板
 - 4 8 个存储节点
-

在一个机柜中可以装配 5800 系统的三种配置：

- 具有 16 个存储节点的全单元
- 具有 8 个存储节点的半单元
- 共有 32 个存储节点的两个全单元

具有两个以上全单元的系统必须使用更多机柜容纳。

存储节点

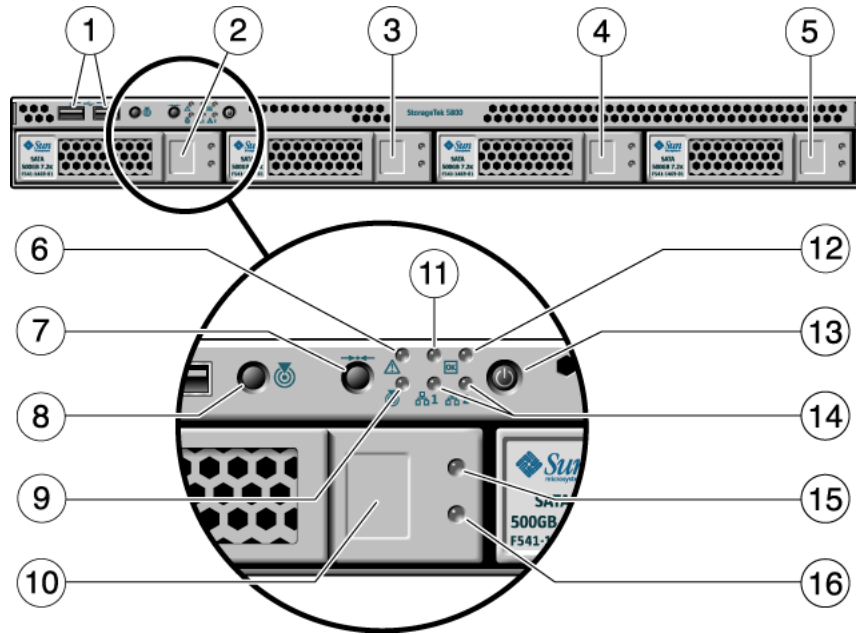
表 1 列出了 5800 系统存储节点的主要特性。

表 1 存储节点特性

组件	描述
CPU	• 一个单核 AMD Opteron 处理器 • 处理器频率：2.2 GHz • 1 MB 2 级高速缓存
内存	3 GB，采用两个 1 GB ECC DIMM 和两个 512 MB ECC DIMM
硬盘驱动器	四个 500 GB 串行 ATA 磁盘驱动器
电源	350W
网络 I/O	两个 10/100/1000BASE-T 千兆位以太网端口
系统管理	智能平台管理接口 (Intelligent Platform Management Interface, IPMI) 1.5 兼容的服务处理器模块

图 4 显示了 5800 系统存储节点的前面板组件。按下存储节点前部的定位器开关会使节点前后部的定位器 LED 指示灯闪烁，所以您可以从机柜的前后部轻松地识别某个特定节点。表 2 解释了存储节点开关和 LED 指示灯的功能及特性。

图 4 存储节点前面板组件



图例

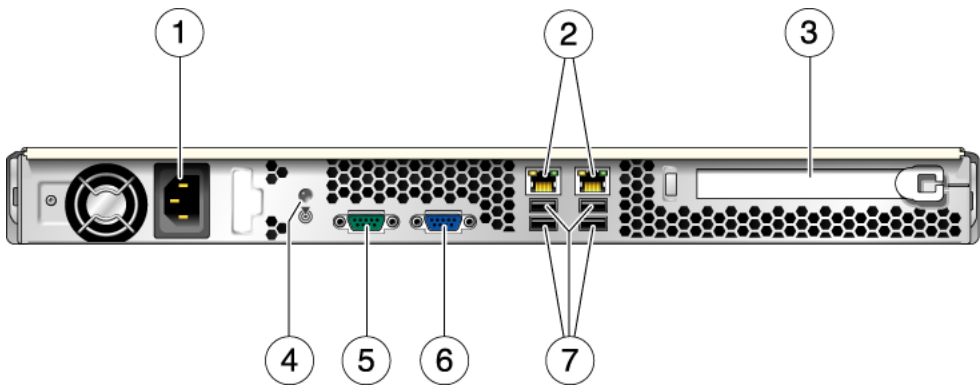
1	USB 端口（未使用）	9	定位器 LED 指示灯
2	驱动器 0	10	释放杆按钮
3	驱动器 1	11	未使用
4	驱动器 2	12	电源 LED 指示灯
5	驱动器 3	13	电源开关
6	存储节点故障 LED 指示灯	14	未使用
7	复位开关	15	驱动器故障 LED 指示灯
8	定位器开关	16	驱动器活动 LED 指示灯

表 2 存储节点前面板 LED 指示灯和开关描述

LED 指示灯/开关	描述
存储节点故障 LED 指示灯	系统中存在故障时呈黄色/琥珀色。
复位开关	重新引导存储节点。
电源 LED 指示灯	电源接通时呈绿色。
定位器 LED 指示灯	按下定位器开关时呈白色并闪烁。当服务节点收到使 LED 指示灯亮起的机箱指示器命令时，该 LED 指示灯也会闪烁。 注： 系统后部有一个相应的定位器 LED 指示灯。
驱动器故障 LED 指示灯	存在驱动器故障时呈琥珀色。 注： 驱动器故障 LED 指示灯由 5800 系统软件激活，以指示必须对单元中的组件进行维修。
驱动器活动 LED 指示灯	没有发生活动时呈稳定绿色。对磁盘进行访问时，该 LED 指示灯会即刻关闭。访问没有停止时，该 LED 指示灯会持续闪烁。当托架上不存在任何驱动器时，该 LED 指示灯熄灭。
电源开关	为系统提供电源。
定位器开关	使前部和后部的定位器 LED 指示灯闪烁，从而使您可以在完全装配的机柜的后部定位存储节点。

图 5 显示了 5800 系统存储节点的后面板组件。

图 5 存储节点后面板组件



图例

1	电源连接器	5	串行端口
2	千兆位以太网端口	6	VGA 端口
3	盲板	7	USB 端口
4	定位器 LED 指示灯		

服务节点

服务节点是具有一个 250 GB 串行 ATA 磁盘驱动器的 Sun Microsystems Sun Fire™ X2100 M2 服务器。5800 系统使用服务节点进行初始配置和故障排除，以及升级系统软件。系统不使用服务节点访问数据对象。表 3 列出了服务节点的主要组件。

表 3 服务节点特性

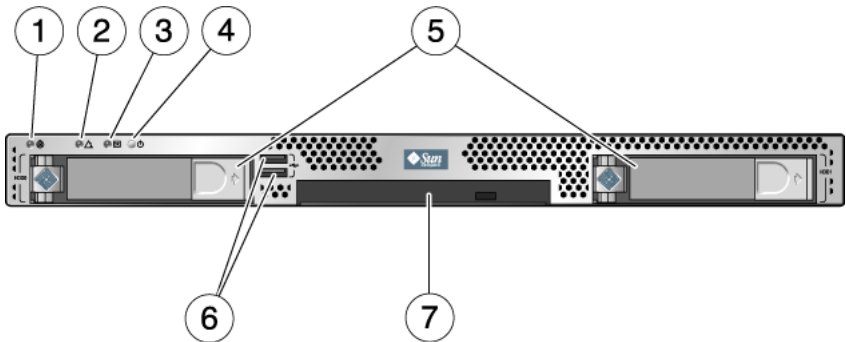
组件	描述
CPU	• 一个双核 AMD Opteron 处理器 • 处理器频率：1.8 GHz • 两个 1 MB 2 级高速缓存
内存	2 GB，采用四个 512 MB ECC DIMM
介质存储	DVD-ROM 驱动器
硬盘驱动器	一个 250 GB 串行 ATA

表 3 服务节点特性（续）

组件	描述
电源	345W PSU
网络 I/O	四个 10/100/1000BASE-T GB 以太网端口，两个 Broadcom 和两个 NVidia。（5800 系统使用两个 Broadcom 端口。）
系统管理	IPMI 2.0 兼容的服务处理器模块

图 6 显示了服务节点的前面板。

图 6 服务节点前面板组件

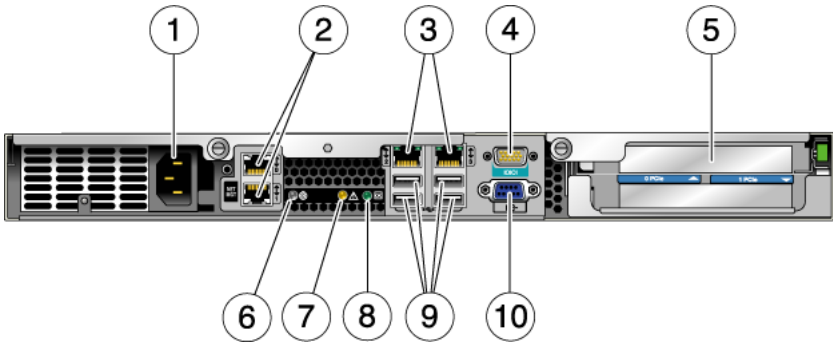


图例

1 定位器按钮 /LED 指示灯	5 硬盘驱动器托架。单个磁盘驱动器位于左侧驱动器托架门锁的后面。
2 服务指示器 LED 指示灯	6 USB 端口
3 电源 LED 指示灯	7 DVD 驱动器
4 电源开关	

图 7 显示了服务节点的后面板。

图 7 服务节点后面板组件



图例

1	电源连接器	6	定位器 LED 指示灯
2	Broadcom 以太网端口	7	服务指示器 LED 指示灯
3	NVIDIA 以太网端口	8	电源 LED 指示灯
4	串行端口	9	USB 端口
5	未使用	10	VGA 端口

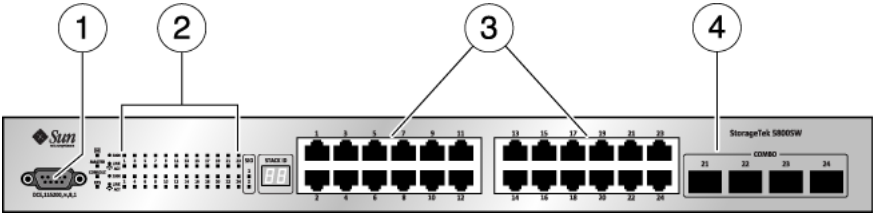
千兆位以太网交换机

半单元或全单元 5800 系统包含两个千兆位以太网交换机。交换机允许系统从单个物理以太网连接（带有冗余备份）以两个虚拟 IP (virtual IP, VIP) 地址的方式进行寻址：一个虚拟 IP 地址用于数据处理，另一个用于管理。此外，通过利用芯片组（这些芯片组支持基于散列表的路由信息的基本数据包标头分析），交换机还可实现负载平衡功能，以用于存储和检索流入和流出存储节点的数据。

图 8 显示了交换机的组件。其中一个交换机被指定为主交换机，另一个被指定为备用交换机。默认情况下，底部的交换机为处于活动状态的主交换机，顶部的交换机为处于备用模式的辅助交换机。如果主交换机发生故障，则辅助交换机会自动接管控制权，并成为主交换机。如果主交换机恢复为联机，则会恢复其控制权。

存储节点 1 至 16 连接到各交换机的以太网端口 1 至 16。服务节点连接到各交换机的端口 17。交换机则通过各交换机的端口 23 和 24 互相连接，以进行心跳通信。有关如何将主交换机和辅助交换机连接到网络配线架的说明，请参见表 4。

图 8 千兆位以太网交换机



图例

1	串行端口
2	端口连接状态 LED 指示灯
3	以太网端口
4	未使用

表 4 千兆位以太网交换机与网络配线架的连接

交换机和以太网端口	网络配线架端口
主交换机 - 端口 21	S1U2 - 仅供 Sun 服务人员使用
主交换机 - 端口 22	S1U1
辅助交换机 - 端口 21	S2U2 - 仅供 Sun 服务人员使用
辅助交换机 - 端口 22	S2U1

网络配线架

5800 系统后部的单个网络配线架可为网络提供所有挂接点。图 9 显示了网络配线架的端口配置。

图 9 网络配线架



以太网端口指定为 SxUy，其中：

S = 交换机
x = 交换机号
U = 上行链路
y = 端口号

S1 为主交换机，S2 为辅助交换机。当主交换机发生故障时，辅助交换机便开始起作用。S1U2 和 S2U2 连接仅供 Sun 服务人员使用。

5800 系统软件

本节介绍 5800 系统软件。本节包含以下主题：

- [第 14 页中的“捆绑软件”](#)
- [第 14 页中的“元数据”](#)
- [第 15 页中的“布局算法”](#)
- [第 15 页中的“用户界面”](#)

捆绑软件

需要将 5800 系统软件作为组合的包（而不是以单个组件的形式）来进行管理和升级。该包中的软件组件包括：

- 5800 系统软件
- Sun Solaris 10 操作系统（x86 版本）
- 所有的 Sun Solaris 10 修补程序
- 基本输入输出系统 (Basic Input Output System, BIOS)
- 服务器管理子卡 (Server Management Daughter Card, SMDC) 固件

即使有许多组件，每个组件都有其自己独特的软件、BIOS 和固件，但它们仍作为整个单元进行升级。即使仅有一个组件需要升级，该包的版本号也会变更，并且必须升级整个包。

元数据

元数据是描述数据对象的信息。5800 系统在分布式数据库中存储有关所有数据对象的元数据。用户可以根据指定给他们的元数据发出查询来搜索数据库并找到对象。5800 系统允许使用两种类型的元数据：系统元数据和扩展元数据。

系统元数据

将数据对象存储在 5800 系统中时，5800 系统会自动将系统元数据指定给每个数据对象。系统元数据包括每个对象的唯一标识符，称为对象 ID 或 OID。5800 系统中包含的应用程序编程接口 (application programming interface, API) 可以使用该 OID 检索对象。此外，系统元数据还包括创建时间、数据长度和数据散列。

扩展元数据

与系统元数据相比，扩展元数据可更进一步描述各个数据对象。例如，如果 5800 系统中存储的数据包括医疗记录，则扩展元数据属性可能会包括患者姓名、就诊日期、医生姓名、医疗记录编号和保险公司。用户可以使用这些属性发出查询来检索数据对象。例如，通过查询可以检索给定医生和特定保险公司的所有记录（数据对象）。

布局算法

5800 系统使用 5 + 2 编码在多个存储节点和磁盘中存储数据对象。系统最多可容忍两个丢失的数据或奇偶校验片段。当磁盘或存储节点发生故障后，系统会将数据和 / 或奇偶校验分布到其他存储节点和磁盘。经过一个重新构建周期后，系统可以容忍另外两个丢失的数据或奇偶校验片段。

当数据对象进入系统时，千兆位以太网交换机会将存储请求定向至存储节点，该节点会将对象分段，然后将这些片段分布到系统中的不同磁盘。5800 系统将数据分成数据和奇偶校验块。布局算法随后从数千种不同的可能布局中确定放置这些块的位置。

用户界面

5800 系统导出两个虚拟 IP (virtual IP, VIP) 地址：一个用于数据处理，另一个用于管理功能。与系统进行交互无需具备底层硬件的知识。相反，您可以将其作为单个大型系统进行访问。

5800 系统提供了许多不同的用户界面，用于进行数据处理和管理功能，如表 5 中所示。

表 5 5800 存储系统用户界面功能

界面	数据处理功能	管理功能
API	可以执行所有存储、检索、查询和删除功能	无
WebDAV	• 功能与 API 相同，但不具有查询功能 • 可以根据定义的元数据属性模式呈现数据对象的虚拟目录结构视图 • 不受多单元配置支持	无
CLI	无（删除全单元上所有数据和元数据的功能除外）	可以执行大多数系统管理任务
GUI	无（删除全单元上所有数据和元数据的功能除外）	可以执行大多数系统管理任务

应用程序接口 (Application Program Interface, API)

基于 Java™ 和 C 语言的应用程序接口 (application program interface, API) 提供了具有其他查询语义的基本存储和检索命令。

您可以使用 API 或 WebDAV（二者均可通过数据 VIP 地址获得）在 5800 系统上执行数据处理任务。

利用 Java 和 C 语言 API，您可以通过 Java 和 C 客户机库存储、检索、查询和删除数据及元数据。客户机软件开发工具包 (software developer's kit, SDK) 中提供了样例应用程序和命令行例程，它们可演示 5800 系统的功能，并提供不错的编程示例。

此外，SDK 还提供了可以在 Sun Solaris、Red Hat Enterprise Linux 和 Microsoft Windows 操作系统上运行的仿真程序。仿真程序模仿 5800 系统的行为，允许您对软件或应用程序进行测试。有关 SDK 的更多信息，请参见《Sun StorageTek 5800 System SDK Reference Manual》。有关 Java 和 C 客户机 API 的更多信息，请参见《Sun StorageTek 5800 System API Reference Manual》。

WebDAV

基于 Web 的分布式创作和版本控制 (Web-based Distributed Authoring and Versioning, WebDAV) 协议是对 HTTP/1.1 协议的一组扩展，允许您在远程 Web 服务器上读取、添加和删除文件。可以在 5800 系统中设置虚拟文件系统视图，从而可以使用 WebDAV 浏览系统上的数据文件，如同这些文件存储在分层路径结构中一样。

例如，您可以设置虚拟文件系统视图，该视图中包含一个医疗记录文件夹，并包含许多不同医生或医院的子文件夹。然后，您就可以使用 Web 浏览器查看这些文件，并在该文件夹和子文件夹中添加和删除新文件。

多单元配置不支持 WebDAV。

CLI 和 GUI

可以使用命令行界面 (command-line interface, CLI) 或图形用户界面 (graphical-user interface, GUI) 执行管理任务。有关使用上述任一界面完成管理任务的信息，请参阅《Sun StorageTek 5800 System 管理指南》。

索引

A

API, 16

B

半单元配置, 2, 5

布局算法, 15

C

存储对象算法, 15

存储访问

管理 IP 地址, 15

数据 IP 地址, 15

存储节点, 7

D

电源 LED 指示灯, 9

电源开关, 9

定位器 LED 指示灯, 9

定位器开关, 9

对象 ID (object ID, OID), 1

F

复位开关, 9

G

故障 LED 指示灯, 9

管理 IP 地址

关于, 15

I

IP 地址

管理, 15

数据, 15

J

机架

填充, 6

填充面板位置, 4

交换机

安装, 4

服务节点连接到, 12

机架位置, 4

节点连接, 4

连接, 12

特性, 12

组件, 12

节点

存储, 7

服务, 10

节点, 服务, 10

K

开关

电源, 9

定位器, 9

复位, 9

扩展元数据, 15

L

LED 指示灯

- 电源, 9
- 定位器, 9
- 故障, 9
- 驱动器故障, 9
- 驱动器活动, 9

P

配置

- 半单元, 2, 5
- 全单元, 2, 4

Q

- 驱动器故障 LED 指示灯, 9
- 驱动器活动 LED 指示灯, 9
- 全单元配置, 2, 4

S

- 数据 IP 地址
 - 关于, 15

W

WebDAV

- 访问数据通过, 16
- 关于, 16
- 浏览文件, 16
- 网络配线架, 13

X

系统

- 关于, 1
- 特性, 2
- 元数据, 15
- 系统节点, 7
- 系统特性, 2
- 系统硬件, 2

Y

- 硬件概述, 2
- 元数据
 - 关于, 14
 - 系统, 15