

NetWare SFT III V3.11

北京启扬软件有限公司技术部

一、技术概况

SFT III(系统三级容错)由主、辅两台镜像服务器构成。对网络上的用户而言仅主服务器是可见的,辅服务器保持在后备工作,但其具有与主服务器相同的基本内存映像和磁盘内容。若主服务器出现故障或停机,辅服务器会自动成为新的主服务器呈现在用户面前。对使用服务器的用户来说,这个过程是瞬间透明的。

NetWare SFT III 要求主、辅服务器的 386 或 486CPU 完全一致并且内存和磁盘容量以及监视器、监视器卡均相同。但并不要求这两台服务器处在同一网段上或具有相同的网卡,不过它们一定要在同一个网上,而且保证网上的用户从它们两者中都能获得信息包。

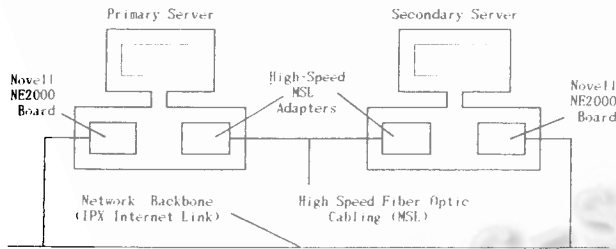


图 1 镜像服务器连接图

每台服务器有两类通信连接,其一为 IPX 网上连接,另一为镜像服务器连接 MSL。IPX 网上连接用于与用户通信和主、辅服务器间互发“服务器状态”信息包,以监测网络工作状态。

MSL 为主、辅两服务器间同步的双向点对点连接。诸如用户请求及应答等信息在 MSL 上传来传去。一旦某台服务器出现故障,MSL 便用来在其故障恢复并重新启动后同步其内存与磁盘存储内容。内存同步即活动的

主服务器将其 MEngine 存储映像的全部内容传至重新启动的辅服务器,紧接着传送整个内存内容,两台服务器间的正常活动便恢复。某服务器停工后,系统磁盘内容有了变化,这时的主服务器将传送所有磁盘更新内容至重新启动的辅服务器,该过程不会对正常网络工作造成冲击。当磁盘内容又完全重镜像后,系统便再次做到容错而受到保护。

主或辅服务器的指派可以变更,当两台服务器启动和同步后,首先激活的服务器便成为主服务器。服务器主辅确定后,如果主服务器发生故障,那么辅服务器将取代主服务器而向网络用户提供服务。

二、SFT III 的容错技术

NetWare SFT III V3.11 提供了比以往 NetWare 版本更强的容错特性。这里我们分如下题目讨论 SFT III 的功能。

1.IOEngine 和 MEngine

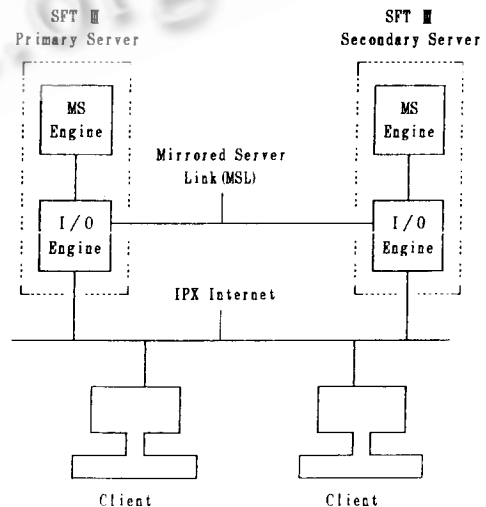


图 2 SFT III 映像图

NetWare SFT III 与标准 NetWare 的根本区别即 SFT III 使用一主一辅两台服务器镜像工作。每台服务器的软件操作系统分为两部分: IOEngine 处理服务器硬件, 未镜像; MEngine 基于 IOEngine 的输入。

MEngine 的内容是同一的(也就是说互为镜像), 两台服务器的 MEngine 具有相同的网间 IPX 号, MEngine 装载如 Oracle's Database Server for NetWare, Monitor 和 Btrieve 等与服务器硬件无关的应用、数据类 NetWare 可装载模块 NLM。

任何直接与硬件通信或明显涉及硬件的模块必须装载在 IOEngine, 如网卡的 LAN 驱动模块, 硬盘驱动模块, MSL 驱动模块, 打印机服务模块以及磁带备份应用程序等。IOEngine 和 MEngine 具有如下特征:

- IOEngine 和 MEngine 处在同一存贮空间的不同存贮段, 禁止内存共享。
- SFT III 服务器同步后, 装载在 MEngine 的 NLM 会自动镜像。
- 装载在 IOEngine 的 NLM 不镜像。

主服务器的 IOEngine 控制整个逻辑服务器, 它转换所有网络请求信息包、硬盘 I/O 中断等为 SFT III 事件, 然后提交到两台服务器的 MEngine。

若某用户需要访问服务器上的资源, 一个请求包将通过网络发送到主服务器 IOEngine, 主服务 IOEngine 收到该请求包后, 通过 MSL 发一个复制包到辅服务 IOEngine。主、辅服务器的 IOEngine 又将该请求送至各自 MEngine。每个 MEngine 接收到此事件, 对其相同地作相应处理, 并有相同的结果应答各自的 IOEngine。虽然主、辅服务器 MEngine 对各自 IOEngine 作了相同应答, 但仅主服务器 IOEngine 应答到用户。

2. 故障处理

故障处理可防止服务器网卡、缆线等系统硬件故障导致系统瘫痪。SFT III 从主服务器切换到辅服务器时, 用户可能感到短促的停顿, 但 SFT III 的故障处理能力保证服务器连续不断地工作。

如果硬件故障使得主服务器停机, 辅服务器便变为主服务器并立即通知网络上的用户访问 MEngine 的最佳路径所在。这时的主服务器将发送一信息包到每个用户工作站, 告知新的路径, 工作站外壳程序读到该通知

后便立即改变访问路径。

主、辅服务器切换过程中可能会丢失信息包, 不过 LAN 通信协议早已考虑了对丢失包的处理, 用户在未收到应答信息包时, 便会重发请求包。这个过程非常之快, 与用户的联接不会中断。

下面分情况详细地描述 SFT III 怎样处理出错的工作过程。

(1) 主服务器硬件故障

- 辅服务器发现镜像服务器连接 MSL 超时, 而且网络上无应答, 因此断定主服务器出现故障停止工作。
- 辅服务器控制台显示信息通告系统管理员主服务器停工, 辅服务器切换成主服务器。
- 辅服务器并发送消息给每个用户告知它现在取代主服务器。
- 用户包重取路径送往新主服务器(先前的辅服务器)。
- SFT III 跟踪服务器出错后的任何磁盘变化。
- 操作员解决故障并重新启动出错过服务器, 两服务器便同步其存贮映像和磁盘。

• (新)主服务器通过镜像服务器连接 MSL 传送磁盘变化内容更新修复的服务器, 并镜像磁盘内容。

(2) 辅服务器硬件故障

- 主服务器发现镜像服务器连接 MSL 超时, 而且网络上无应答, 因此断定辅服务器出现故障停止工作。
- 主服务器控制台显示信息通告系统管理员辅服务器停工, 主服务器未变更。
- SFT III 跟踪辅服务器出错后的任何磁盘变化。
- 系统管理员修复并重新启动辅服务器后, SFT III 识别两服务器不同步了。
- 主服务器通过镜像服务器连接 MSL 传送存贮映像和磁盘变化内容更新辅服务器并重新使两服务器同步。

(3) 镜像服务器连接 MSL 故障

- 辅服务器侦测到 MSL 上没有信息传来, 但 IPX 网络连接上仍有“服务器状态”包传到, 这说明主服务器仍在工作。
- 辅服务器便停工关机, 因为主服务器仍在为用户服务, 但同步却不能进行了。
- SFT III 保持跟踪辅服务器停机后的磁盘变化。

• 当系统管理员修复 MSL 并启动辅服务器后，SFT III 识别出两服务器不同步。

• 主服务器便通过 MSL 传送存储映像和磁盘变化内容更新辅服务器以重新同步两服务器。

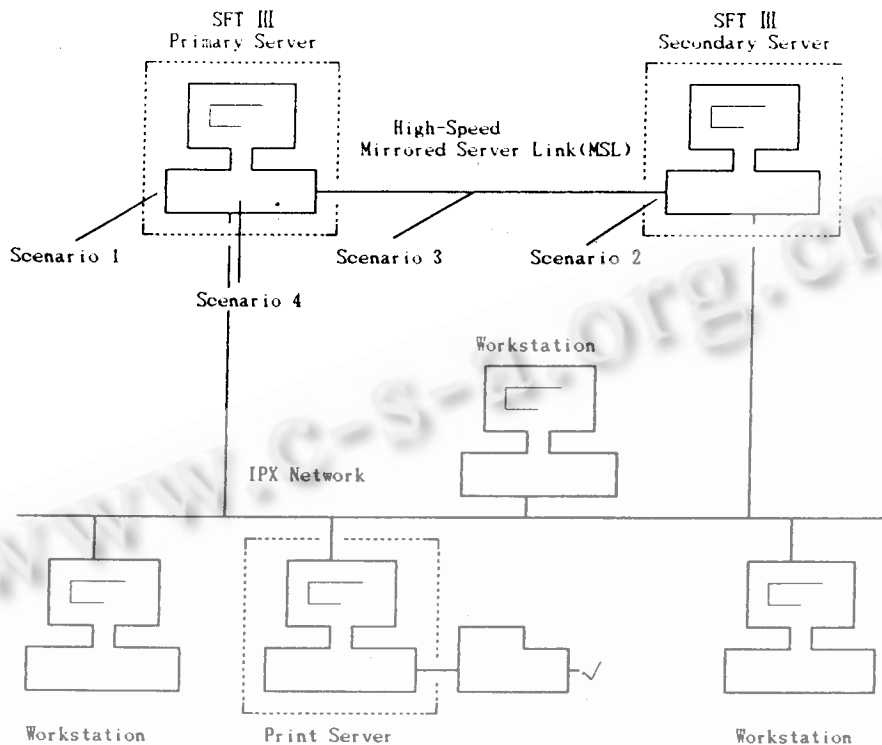


图 3 SFT III 故障处理示意图

(4) 主服务器网卡故障

- 主服务器操作系统检测到网卡出错。
- 操作系统检测辅服务器网卡状态，如果辅服务器的网卡工作正常而服务器仍然镜像着，辅服务器便接管对网络用户的服务。
- 服务器控制台将通知系统管理员该出错信息，并登记入记录文件。
- SFT III 保持跟踪出错服务器停工后的任何磁盘变化。
- 故障服务器重新启动并成为辅服务器。
- 主服务器通过 MSL 传送存储映像和磁盘变化内容更新故障服务器以重新同步两服务器。

3. 双 CPU(可选)

NetWare SFT III 为双 CPU 提供可选支持。由于 SFT III 划分为二个 Engine(IOEngine 和 MEngine)，所以每个 Engine 可在分别的 CPU 上运行，来建立一个双

CPU 系统。

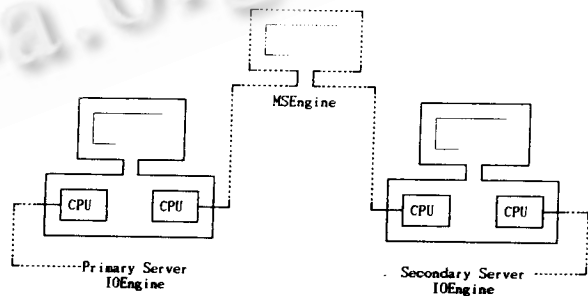


图 4 双 CPU 映像示意图

NetWare SFT III 运行在双 CPU 环境时，一个 CPU 负责所有与硬件有关的 I/O，另一个 CPU 执行文件服务、NetWare 核心协议(NCP)、及其它镜像的应用程序。

前一个 CPU 为 IOEngine 服务,后一个 CPU 为 MEngine 服务。

双 CPU 系统可以改善高负荷系统的性能。通过两个 CPU 提高系统的性能,系统负荷大小是非常关键的,而且还要保证两个 CPU 同时工作。

4.SFT III 服务器角色指派

单个 SFT III 服务器在某时刻可以是:主镜像的主服务器、镜像的主服务器或镜像的辅服务器。两服务器镜像后,它们保持相互通信并告知各自的主、辅角色。当辅服务器发现主服务器不再响应它发出的询问时,便必须确定主服务器是否已瘫痪。若主服务器停止工作,辅服务器将变为主服务器在没有镜像伙伴情况下工作。

5.网间 IPX 号

与标准的 NetWare 一样,SFT III NetWare 服务器也有一个网间 IPX 号。一个服务器的网间 IPX 号既是该服务器的地址,也是用户请求网络服务发送信息包的远程目的地址。NetWare 中之所以需要网间 IPX 号,原因之一就是需要一个网络服务器必须有一个固定不变的地址。这样,即使有其它局网联入或退出,IPX 用户可通过转换路径继续与服务器通信,因为服务器的远程目的地址未变。

SFT III 扩展这种动态路径变换能力来提供高级容错,其关键是网络配置镜像服务器。两个镜像服务器共享一个 MEngine 网间 IPX 号。

6.服务器故障切换

客户向镜像服务器申请网络服务时,总是将信息包发往主服务器 IOEngine,因为主服务器 IOEngine 是通向 MEngine 的最佳路径。主服务器 IOEngine 接收所有的网络服务请求,控制送到两服务器 MEngine 中的所有事件序列,并保持两个 MEngine 同步。主服务器 IOEngine 还将发给用户应答信息包。

两个服务器 MEngine 接收同样的事件请求,并产生同样的结果给 IOEngine,但辅服务 IOEngine 丢弃所有的应答信息包。因此,客户仅从主服务器 IOEngine 处接收到应答信息,用户原始的请求信息包也是发至该 IOEngine 的。对于用户来说,镜像服务器同其它 NetWare 服务器一样。用户向某个地址发送一个请求信息并从该地址处获得一个应答信息。向辅服务器 IOEngine 复制请求信息包和同步两个服务 MEngine

所有事件对网络用户来说是透明的。

主服务器发生故障后,辅服务器立即切换为主服务器。从网络用户角度来看,只是发生了简单的路径变化。新主服务器(先前的辅服务器)的 IOEngine 要广告通往主服务器 MEngine 的路径变化。它给网络用户发送一个特殊的信息包,将新的路径告诉用户。新的主服务器现在将向用户发送应答信息包,而不象做为辅服务器时那样,丢弃所有的应答信息包。

通常用户向标准的网络 NetWare 服务器请求服务时,如果信息包丢失,会重发该请求包。IPX 网络用户能够响应网络路径变化,并在丢失请求或应答信息包时,做补救工作。

当发生主服务器切换时,NetWare SFT III 只是简单利用网络通讯协议本身处理动态路径变化的能力。

说明:

NetWare SFT III 是一个操作系统,在 NetWare V3.11 的基础上,增加了一些可靠性措施,防止因服务器硬件故障造成数据丢失。NetWare SFT III V3.11 专控版为 250 用户,除打印实用例程外,其它例程与 NetWare V3.11 相同。NetWare SFT III 实用例程向下兼容其它 NetWare 系统。

硬件要求

- 两台 Intel 386 / 486 CPU 服务器(建议使用 Novell 认证的服务器)
- 每个服务器主频 25M 以上,最少 12MB RAM。
- 两个服务器的 RAM 和磁盘空间完全一致。
- 两套镜像服务器连接(MSL)通讯卡及电缆。
- 每个服务器配有网络卡。
- 两服务器的监视卡及监视器相同。(均为 EGA、CGA、Super VGA 卡,彩显或单显)

软件要求:

- 两台服务器必须采用一致的 DOS 3.1 或以上版本系统(用于硬盘引导)
- Netware SFT III V3.11
- MSL 驱动程序
- 双 CPU 驱动程序(可选)
- 每个服务器配有网络驱动程序
- 每个网卡配有驱动程序