

EterneCluster

技术白皮书

第一章 信息高可用性概述.....	3
一、 高可用性需求.....	3
二、 服务宕机的必然性.....	3
三、 高可用性解决方案.....	4
四、 高可用性工作原理.....	4
(一)、 高可用硬件配置.....	4
(二)、 高可用工作原理.....	5
第二章 EterneCluster 概述.....	7
一、 软件组成.....	7
(一)、 EterneCluster Console.....	7
(二)、 EterneCluster Core.....	7
(三)、 EterneCluster Monitor.....	7
(四)、 EterneCluster Resource DLL.....	8
二、 功能特点.....	8
(一)、 应用高可用性.....	8
(二)、 自我修复能力.....	8
(三)、 支持多条心跳路径.....	8
(四)、 共享磁盘的锁定.....	8
(五)、 本地失败恢复.....	9
(六)、 本地网卡失败转移.....	9
(七)、 支持多种群集工作模式.....	9
(八)、 支持自动迁回.....	9
(九)、 支持无为 ARP.....	9
(十)、 支持快速切换.....	9
(十一)、 支持双机双柜.....	10
(十二)、 直观的 GUI 用户界面.....	10
(十三)、 支持定制开发.....	10
第三章 EterneCluster 应用规格.....	11
一、 软硬件规格.....	11
(一)、 支持各种类型服务器.....	11
(二)、 支持各种类型的存储设备.....	11
(三)、 支持的操作系统平台.....	11
(四)、 提供丰富的应用软件支持.....	11
二、 群集规格配置.....	12
(一)、 双机集群 (1-1) 配置.....	12
(二)、 多机集群配置.....	12
第四章 术语附录.....	15

第一章 信息高可用性概述

一、高可用性需求

随着信息化建设的不断推进,企业的信息化已经非常普遍,各个企事业单位的活动越来越多的依赖于其关键的业务信息系统,这些业务信息系统对整个机构的运营和发展起着至关重要的作用。如果一些关键应用一旦发生宕机故障或应用停机,将给企业带来巨大的经济损失,并且由于业务的不可用而影响了企业的信誉,致使客户对企业失去信任,所造成的危害是致命的。

非正常的服务终止时间的长短会对用户造成不同的损失,或者说用户对服务停止所能承受的时间是不同的。通常取决于应用的类型,如果能够在几秒钟内修复错误,不会对一个在线联机事务(OLTP)处理系统构成影响,但是对火箭发射等实时应用,哪怕只停止一秒都是不可忍受的。

由于系统的任何一个部件都可能发生故障,因此系统设计师在构建系统时需要能够预见到系统可能会发生的各种故障,并进行软硬件的部署,在这些故障后发生能够迅速纠正。

可见,对那些需要保障信息安全和提供不间断的信息服务的企业而言,业务系统的容错性和不间断性显得尤为重要。如何保障各种关键应用持续运营,达到永续经营的良性循环,已成为当今企事业单位和 IT 领域急需解决的关键问题。

二、服务宕机的必然性

通常企业构造服务的硬件环境为单一服务器配置模式,一旦因为计划中宕机和计划外宕机,便会引起服务终止。

计划中的宕机通常指完成系统升级、应用迁移、部件更换等操作引起的宕机,这些计划中的宕机是企业维护服务正常工作不可避免的工作项目,虽然可以人为地采取一些措施减少损失,但依然会造成服务的终止。

计划外宕机通常是由于系统出错造成的。错误包括硬件、软件、系统和网络,或是系统运行外部环境原因等,这些情况包含了:硬件故障、操作系统崩溃死机、硬盘满、电压不稳、跳电、电源损坏、网络故障、软件漏洞(Bug)、应用出错。同样人为失误造成的故障也称为出错。当然,并非所有的出错会造成宕机,也不是所有的意外宕机都是由于部件出错造成的,灾难或其他意外情况同样会造成服务终止。计划外宕机通常会导致非常严重的后果,比如酒店的入店和结帐系统故障,所有的客户将无法 check-in 和 check-out;医院的电脑系统出错,将导致患者无法结账、医生无法得到患者信息,甚至无法进行手术;银行的记帐系统因电脑出错,客户将无法提现和存款。总之,机会外宕机所造成的损害将非常大。有些环境下,系统停止服务将导致事务处理无法进行,必将导致客户对企业信任度降低,甚至会给客户带来不可挽回的损失。

三、高可用性解决方案

通常，对于服务因错误造成的不可避免终止，企业都会安排相关的管理人员进行定时的监控，一旦发现服务不可用，便会立即进行处理排错。这是非常原始的高可用解决方案，存在诸多问题，譬如及时性问题，一旦错误在深夜发生，而人工不可能 24 小时监控；再譬如处理时间长短不可控，一旦错误造成的原因非常复杂，排错非常困难，需要非常长的时间，那么服务的终止将非常长。在排错过程中，人为操作失误将可能导致系统的永久瘫痪。另一方面，为了维持系统的高可用性而配备操作人员，劳务费用也非常昂贵。

目前，市场上提供多种高可用性解决的软件解决方案。主要分为双机和多机的解决方案，主要原理是在提供硬件冗余的系统基础之上，运行高可靠性软件而构成。高可用性软件自动检测服务的运行状态，一旦服务在主机上出现故障，高可用性软件将自动地把服务转到另一台服务器上，从而让服务持续可用。对系统应用状态的监控和服务故障的恢复，由高可用软件自动完成，减少了人员的负担。

四、高可用性工作原理

（一）、高可用硬件配置

高可用环境中，硬件是要求冗余的，包括两台服务器（我们称这种环境为双机容错），而部署多机高可用集群环境时，服务器则是多台，每台服务器拥有足够的 I/O 能力、内存容量、系统硬盘空间、网卡部件，使得配置能够在使用中符合系统的需求，并能够最大限度的降低系统停机时间。

存储系统通常也必须是冗余的，例如，将数据存放在 RAID 5 磁盘阵列中。当一个硬盘发生错误时，数据仍然可以正常访问，而且可以从其他硬盘中恢复出这个硬盘的内容（称为 RAID 重建）。

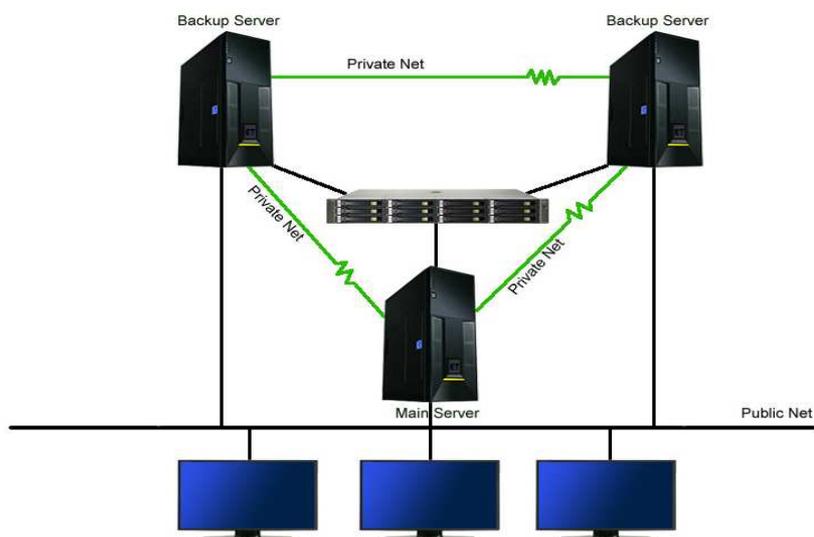


图 1 高可用系统的硬件拓扑结构

高可用系统的硬件拓扑结构如图 1 所示。

服务器：主服务器和备份服务器运行同一版本的操作系统，但其硬件配置可以不一致，主服务器为提供服务的服务器，备份服务器平时不提供服务，只有主服务器出现问题时，其能够接替主服务工作的服务器。

私网：用于在服务器之间传输 cluster 的信息网络结构，其主要是支持 TCP/IP 协议的以太网和光纤网。

公网：供客户端访问服务的网络，客户端可以通过公网访问到每台服务器。

共享存储：每个服务器都能访问的存储设备，同时也是服务数据存放的地方，一般采用磁盘阵列，为了对数据进行冗余保护，可根据客户需要做不同的 RAID 级别数据保护。磁盘阵列的接口形式可以是 iSCSI、FC、SAS、NAS 等。

（二）、高可用工作原理

每台服务器都接上磁盘阵列，客户的应用程序或数据都放入磁盘阵列中，客户通过浮动的 IP 地址访问服务。客户的服务平时在主节点上运行，当主节点出现故障（硬件故障如：网卡，卷，磁盘等出错；软件故障：应用不可用）时，第二台备用服务器自动接管其工作(即带入磁盘阵列的卷，接管浮动 IP 地址，最后启动应用)。如果第二台备用服务器不可用，其任务被第三台服务自动接管，如此类推。因此，一台服务器节点出现故障后，其他服务器可接管其服务，从而保证整个系统对外服务的正常，为企业 24 小时×365 天的关键业务应用提供了强大的保障。

（1）心跳信息原理

主备机通过心跳路径，周期性地发出检测数据包。如果主机出现故障，备机在规定的时间内没有收到来自主机的检测包后，会认为主机可能出现故障；这时备机自动检测设置中是否有第二条心跳途径，试图在第二条和主机建立心跳通信。如果第二条心跳途径同样失败后，备机则根据已设定的策略，启动备机的相关服务，完成服务的切换。

（2）浮动 IP 工作原理

浮动 IP 指客户端访问服务的 IP 地址，这个浮动 IP 地址挂载受到高可用软件的控制。正常情况下，由主服务器对外提供服务，高可用软件把此 IP 地址挂载到主服务器上。当主服务器出现故障后，服务发生切换，高可用软件会自动把该 IP 地址挂载到备份服务器上，用户访问的仍然是浮动 IP 地址。所以用户只会在切换的过程中发现有短暂的通信中断，经过一个短暂的时间，就可以恢复通信。

（3）切换条件

- 主机的硬件发生故障，譬如主机主板故障、主板掉电。
- 主机的操作系统发生故障。
- 运行于主机的服务出现故障。

- 主机的公共网络发生故障。

当发生以上故障时，高可用软件会及时检测到这些故障，主动把服务切换到相应的备份主机上。

(4) 切换过程

服务切换过程主要是把服务从主服务器上切换到备份服务上。

首先，在主服务器上主动释放相应的资源，主要步骤是：

1. 执行正常的服务关闭程序，这样可以保证服务的完整性及数据的安全，而且也不会再有其他服务器执行此项服务。

2. 释放共享磁盘，以便备份服务器能够访问共享磁盘。

3. 释放浮动 IP 地址，这样便可保证不会再有其他服务器使用此 IP 地址。

其次，由备份服务器接管相关的资源：

1. 获取存有服务的共享磁盘分区并系统检测与恢复，此过程可进一步保证在服务异常中断时共享磁盘的完整性。

2. 在备份服务器网络挂载浮动的 IP 地址，客户端会重新连接到原来的 IP 地址。这样，用户再去访问服务时，就没有必要再去重新配置或做修改。

3. 重新启动服务。

(5) 切换时间影响因素

切换时间受多种因素的影响，主要由 3 个方面决定：

1. 启动服务的实际时间，与环境配置的复杂性和数据库存储数据量相关，往往客户实际配置比较复杂，数据库比较庞大，因而启动服务就比较长；

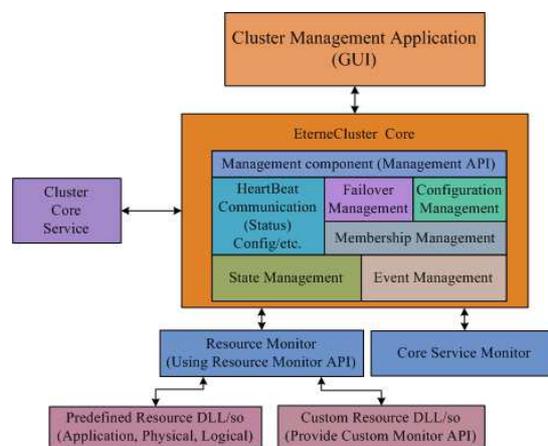
2. 共享磁盘阵列的挂载，当磁盘阵列的文件系统出现问题同时数据比较多时，挂载该分区时做文件系统检测就比较长；

3. 停止服务实际时间，其所受各种环境条件的影响和启动服务一样。

第二章 EterneCluster 概述

EterneCluster 产品是来自中国本土的，拥有自主知识产权的多机高可用产品，支持 Windows 操作系统平台，支持几乎所有的存储环境。具有稳定可靠、易于管理、备援方式灵活多样、节点和任务伸缩性强、保护用户投资、整合和优化用户系统环境、支持包括 iSCSI 存储环境等特点。

一、软件组成



软件主要分为 4 大部分构成：EterneCluster Console、EterneCluster Core、EterneCluster Monitor、EterneCluster Resource DLL。

(一)、EterneCluster Console

Console 是由是 JAVA 编写的用户界面，通过该界面可以管理整个 cluster 的所有节点和服务，同时直观地表示 cluster 的各个任务运行状态和细节。

(二)、EterneCluster Core

Core 是 Cluster 的核心服务程序。其负责执行来自 Console 的命令，协调各个节点的任务分配，传输和接收各个节点的心跳信息，切换和接管任务的管理中心，同时接收来自 Monitor 定时发送的资源状态信息，并进行解析和进行相关的操作。

(三)、EterneCluster Monitor

Monitor 负责监控资源的状态，定时把这些资源的信息传递给 Core，同时监控 Core 运行的状态，一旦发现 Core 出现问题，会立即重新启动 Core，从而保证 Core 不存在单点故障

的可能。

（四）、EterneCluster Resource DLL

Resource DLL 是 Eterne 提供主要的资源动态连接库，主要包含：IPAddress、Network Name、Physical Driver、Share、Application、Service，每个资源都有自己的属性、方法、事件。对于用户的特别服务，Eterne 可以为用户提供第三程序编写 DLL 服务，在用户提出需求后，由 Eterne 开发人员为用户开发第三方软件的资源编写。

二、功能特点

（一）、应用高可用性

提供高可用性解决方案是 EterneCluster 的核心功能，该解决方案提供双机热备系统，也提供多机的高可用解决方案。

（二）、自我修复能力

在集群系统的每一台服务器内，EterneCluster 具有两个核心进程，他们互相监控（如：系统宕机、EterneCluster 进程被杀掉、网络、SCSI 线缆是否断开），如果其中一个进程失败，另一个进程会立即进行恢复。

（三）、支持多条心跳路径

支持网卡(私网/公网)、磁盘心跳等多种心跳方式,多重保障心跳检测途径,配置多条心跳路径可以避免系统的单点故障,进一步提高了系统的可靠性。

（四）、共享磁盘的锁定

软件管理共享磁盘上的数据，以防止多个服务器在同一时间访问数据。EterneCluster 在逻辑设备级（卷）上控制对数据的访问，管理员可以不必担心群中的其它服务器访问数据时，可能会带来的数据访问冲突。EterneCluster 自动在被应用程序定义为共享资源的磁盘卷上设置 SCSI-2/SCSI-3 持久保留。当被保护的的应用程序由一个服务器被移动/转换到另一个服务器时，EterneCluster 控制这些锁定，以保证激活服务器对共享卷的访问。

在主系统发生故障的情况下，次节点系统将能够在磁盘上建立 SCSI 锁定，并在备份的系统上将资源投入使用。

（五）、本地失败恢复

在应用程序的检测时间间隔执行检测，以察看资源本身是否失效。如果快速检查和深入检查均局部告失败，系统将根据用户设定的策略尝试局部恢复资源。如果尝试成功，资源将不会向下一优先级的节点进行失效切换。如果局部恢复尝试失败，系统将向下一优先级的节点进行失效切换。

（六）、本地网卡失败转移

在主服务器上，可以加上另外一个 NIC 作为主服务器的备份 NIC。当主 NIC 出现故障时，EterneCluster 将可切换的 IP 地址让由本地的备份 NIC 接管，本地资源不需要切换，有效的减少的切换时间。

（七）、支持多种群集工作模式

不仅支持传统的单活(Active/Passive)、双活(Active/Active)，还支持 N 和 N+1 多节点集群备份。目前 EterneCluster 最多支持 32 个节点的群集工作模式。

（八）、支持自动迁回

当首选节点发生故障后又恢复正常，如果允许自动迁回，备份节点在指定的恢复时间窗口内将资源组切换回首选节点运行；禁止自动迁回时，资源组将保留在备份结点上，管理员可在合适的时间手工切换。

（九）、支持无为 ARP

当浮动 IP 切换到其它网卡上挂载时，Cluster 会主动进行 ARP 广播通知客户端该 IP 的 MAC 地址发生改变，让客户端主动更新映射关系，这样客户端便会连接到相关的服务上，避免因浮动 IP 切换造成客户端无法访问的情况。

（十）、支持快速切换

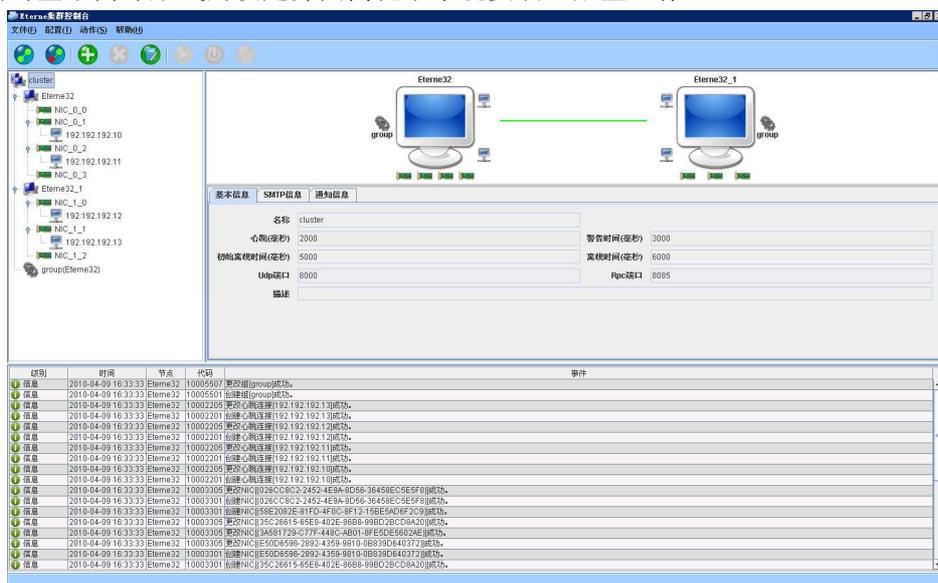
针对不需要切换的资源，如中间层程序，IIS 服务等应用，Cluster 提供本地监控及本机恢复功能（如重启资源和重启服务器），因为这些资源本身无需切换，只要主机备机同时启动即可，这样可以节省切换时间，快速切换实现了应用信息服务系统的永不停顿。

（十一）、支持双机双柜

配合 Veritas Storage Foundation Basic 的免费版本的卷管理器的动态磁盘组功能，EterneCluster 可很好的支持双机双柜的解决方案。

（十二）、直观的 GUI 用户界面

提供直观的 GUI 用户界面，用户可以交互式地对不同的集群系统进行配置、监控和远程管理，实时地显示出主机系统及服务的状态，大大减少了系统因用户人为的错误而引起的损失，同时基于向导形式的设定界面简化了系统安装，配置工作。



（十三）、支持定制开发

EterneCluster 可以结合用户需求，开发特定应用或硬件资源的监控程序，与用户应用环境更紧密地结合，更好的监控服务运行状态。

第三章 EterneCluster 应用规格

一、软硬件规格

（一）、支持各种类型服务器

支持 DELL, HP, PowerLeader, IBM, 浪潮, 联想等 X86 架构的所有服务器, 各个节点的服务器类型和硬件配置可以不一致。

（二）、支持各种类型的存储设备

可利用 SAN, SCSI, NAS 既存的资源进行系统配置, 支持各种类型的磁盘阵列, 包括 Infortrend, HP, IBM, SUN, HDS 等。

（三）、支持的操作平台

目前 EterneCluster 支持 Window 系列和 Linux 系列的操作平台, 下面是具体支持的操作平台:

- Windows 系列操作系统
 - Windows 2000 系列
 - Windows 2003 系列
 - Windows 2008 系列
- Linux 系列操作系统
 - RedHat Enterprise Linux 4.0/5.0
 - Suse Linux Enterprise 10.0/11.0
 - Ubuntu Linux 8.0

（四）、提供丰富的应用软件支持

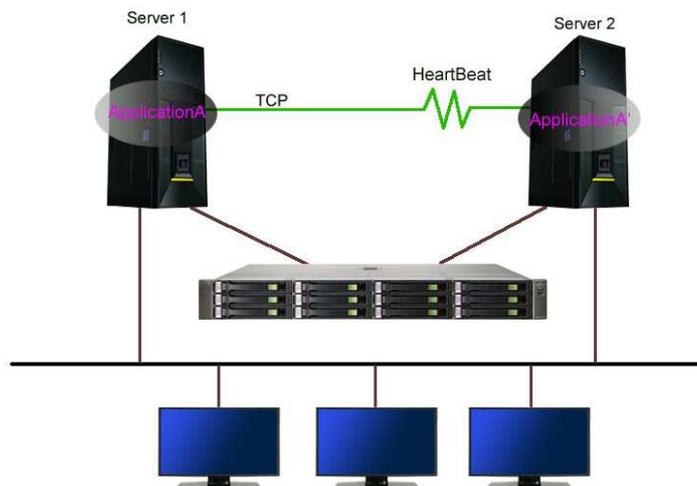
支持多种应用的高可用, 并能与重要的应用如数据库: Microsoft SQL Server、Exchange 2000/2003、Oracle、文件服务器等紧密配合。EterneCluster 支持的应用包括:

- ✓ 数据库: MS SQL Server、Oracle, Sybase 等。
- ✓ 文件服务器:NFS, Samba 等。
- ✓ WEB 应用服务:IIS, Apache 等。
- ✓ 邮件服务器: Exchange 2003/2007 等。
- ✓ 用户自定义。

二、群集规格配置

EterneCluster 支持多种类型的配置，主要分为 2 大类：双机集群和多机集群。

（一）、双机集群（1-1）配置

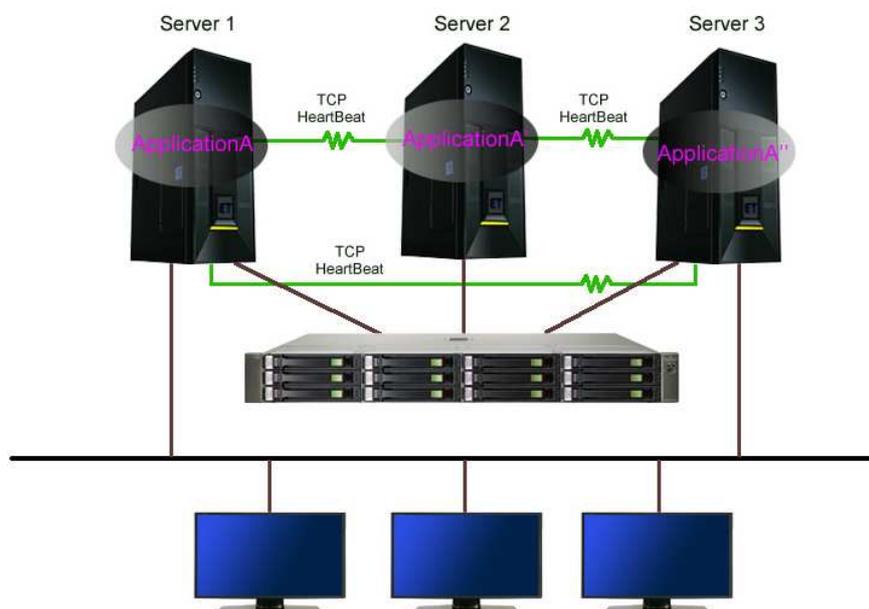


简单双机集群是目前大多用户采用的高可用环境，简单的说就是两台服务器加一台磁盘阵列，通过 EterneCluster 软件实现主从工作方式的双机环境。主从就是 Active-Standby 工作模式，容错软件作为不可缺少部分起到监控系统状态并在系统故障时，自动做出相应的反应，保证整个系统提供服务的不间断。同样双机集群可以实现 Active-Active 工作模式，即两台服务器分别运行不同的服务，并且互相作为备份，一旦某一台的系统出现故障，另外一台便会自动进行服务的接管。

（二）、多机集群配置

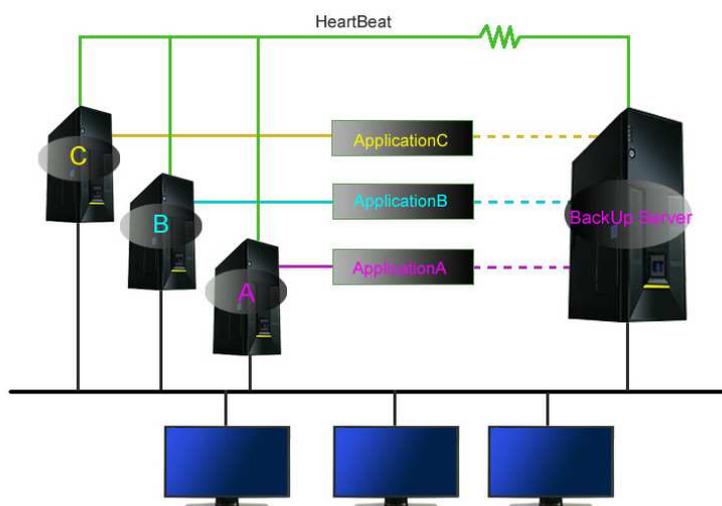
多节点集群在未来会有大量需求，大家对集群的认识目前还在双机环境下。而实际上，在有多多个关键应用都具有高可用需求，在一个机房内建立多个双机系统显然是很浪费的，管理的复杂度又高，多节点集群可以有效的解决这个问题。对于多节点高可用，可以理解为多机互备，多个服务器连接在一个共享存储设备上，同时运行多个不同应用，在其中任意服务器出现故障时，其它服务器根据备援策略进行接管服务，保证整个集群中的服务都能高可用。比较常用的多机集群配置主要有 3 个类型：1-N、N-1、M-N。

(1) 1-N 配置



这个配置方式是一个标准的多机配置方案。在这个 cluster 中，每台服务器都能运行客户的应用服务，平时仅由主服务器提供应用服务，其余服务器都是备用服务器。一旦主服务器发生故障，cluster 依照切换策略，会把任务切换到顺位的备份服务器中，如果顺位服务器也发生故障，应用服务会继续按照结点优先级切换到下一台，直到服务能够接管。这种配置方式一般用于那些要保证应用服务 100% 能够 24 小时在线的客户环境。

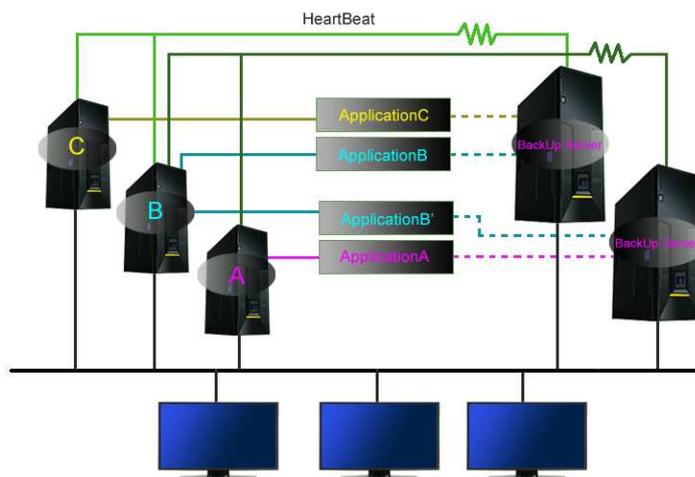
(2) N-1 配置



这个配置方式是所有的服务器都在一个 cluster 中，其中一个服务器作为 cluster 中其它

服务的备份服务器，一旦某个主服务器系统出现故障，任务会按照策略把该服务器的任务切换到该备份服务器。某种意义上，其为双机系统的扩展版本，但这种体系结构配置对客户有多个服务的环境是非常优化的系统配置，同时能够让整个集群集中管理并且能够达到高可用。

(3) M-N 配置



这个方式中多个生产服务器工作在一个集群中，通过合理定义备援策略，使环境中服务器相互备援，任何一个任务停止工作都将由其它服务器接管任务，是一种设备利用率最佳的部署方式。该部署方式同样有效的优化了系统结构，需要注意的是在定义备援策略时，仔细分析各服务器处理能力、任务的兼容性问题。

在 M-N 方式中当主节点发生故障时，允许多个备份节点被按照一定的优先级次序恢复一个资源组。对于在一对节点上的资源恢复，如果节点 A 发生故障了，资源将会切换到节点 B 上；如果节点 B 再发生故障而节点 A 仍然不可用，资源将会失效切换到节点 C 上。多个备份节点被指定一个优先级。在上面的例子中，节点 A 有最高的优先级，节点 B 有第二优先级，而节点 C 有最低优先级。EterneCluster 按优先级次序测验节点来决定在故障切换时哪一个服务器将进行工作。

第四章 术语附录

- **节点(Node):** 运行相同版本操作系统和 Eterne Cluster service 的计算机系统,并对外部 client 提供特定的服务,如 SQL 服务和 IIS 服务等,服务通常可以分为下面两种.
 - **首选节点(Preferred Node):** 指集群中指定的运行资源组的首选的节点,往往配置比较好,该节点一般需要配置:2 块网卡,接上盘阵的 SCSI 或光纤
 - **备份节点(Backup Node):** 备份节点硬件配置可以不同于首选(当然也可以相同),但是要求运行相同的操作系统,同时必须安装和主节点上相同的应用,当首选节点出现故障时能够接管其资源组的工作。
- 首选节点是相对于一个特定的资源组,如果是多个节点的高可用性方案,就会有多个首选的节点。那么就可能出现这种情况,对一个 Resource Group1 来说 A 服务器是首选节点,但是对 Resource Group2 来说 A 服务器可能就是备份节点。
- **浮动 IP 地址(Floating IP address):** 一个服务将其服务通过特定的 IP 地址来输出,这个 IP 地址就是浮动 IP 地址,这个地址将会绑定在提供主服务的服务器的一个正常的公网上,当服务进行切换时,浮动 IP 地址也将进行切换,这样对外界来说,该 IP 地址仍然是可以访问的,由于这个 IP 地址是在多台服务器之间来回变动(如果发生服务不可用等故障时),所以叫做浮动 IP 地址。
- **主机网络名(Network Name):** 和上面的浮动 IP 地址类似,客户端大多是通过主机名字来访问服务的.这样我们就必须提供一个浮动的主机名字,我们将这个名字称为主机别名.当浮动 IP 地址发生切换时候,主机别名也会相应的进行浮动,但是相对于浮动 IP 地址来说,这个别名是被捆绑在其上的。
- **心跳(Heart Beat):** 是两个组件之间用来判断对方状态的机制,当然也能够通过心跳来交换一些简单的信息.心跳在 EterneCluster 中主要分成两种
 - **监控心跳: (Monitor Heart Beat),** 这种心跳是存在于监控(Monitor)和 EterneCluster 之间的,用于交换 Agent 信息的
 - **EterneCluster 心跳(EterneCluster Heart Beat):** 这种心跳是存在于多个 EterneCluster 之间的,用来交换 EterneCluster 间信息和判断对方是否有故障的
- **私网(Private Net):** 用以交换服务器间心跳的网络。它可通过一条独立的网线来传送心跳。为了避免单点故障,一般是配置 2 条或者 3 条私网,太多的私网将不会导致系统消耗更多的资源用于私网的通讯,从这点上说,2 条私网是比较合适的
- **公网(Public Net):** 供客户端访问服务的网络。连接公有网最好的办法是用独立的网线连接,同样为了提高系统的可用性,我们可以配置 2 条到多条公网,这样当其中一条公网发生故障的时候,我们可以不用进行服务的服务器之间的切换,而只需要进行浮动 IP 地址的切换(就在原来的这台服务器上的多个公网间切换浮动 IP 地址。
- **客户端(Client):** 是指通过公网访问服务的计算机系统。
- **共享磁盘(Shared Disk):** 共享磁盘一般为磁盘阵列,可以被所有的服务器访问,上面存放着服务的数据或主程序。当主节点访问时该磁盘时,备份节点将不能

访问共享磁盘，即为独占访问模式。备份节点只有在将 Resource Group 切换过来后才能访问该磁盘。

- **应用层服务(Application service):** 简称为“应用服务”,这些服务(譬如 oracle, Exchange, SQL 等)是在指定的主节点上运行,客户端可通过指定的浮动 IP 地址来访问,主服务由 EterneCluster 软件来管理,当主节点出现故障时,在很短的宕机时间内,备份节点会将主服务接管过去。多个主服务也可以分布在多个不同主节点上,也就是说两台服务器都是主节点,但是是不同服务的主节点,当然他们也同时是备份节点(也是针对不同的服务),这样做可以提高系统的可靠性,同时能最大限度的利用硬件资源。
- **应用服务监控(Monitor for application service):** 简称为“监控”,用来监测应用层服务的可用性,并把这些信息及时汇报给 EterneCluster 核心。监控默认 3 秒(可由策略来设置)汇报给 EterneCluster 核心,如果联系 5 次(可由策略来设置),即 15 秒 EterneCluster 核心没有收到代理的信息或者代理返回应用已经不可用,那么 EterneCluster 核心就认为应用已经不可用, EterneCluster 核心便会进行相应的处理动作。至于汇报间隔和次数可以通过管理界面进行配置。
- **事件(Event):** 指系统中出现了有价值的状态变化,如发现硬件和软件方面的改变等等。
- **通知(Notification):** 是指将符合特定的条件的事件通过不同的方式通知用户,通知讲采用模块化的设计利用回调函数的方式来实现,这样有利于以后的功能扩充,现在通知采用如下几种方式:
 - **日志(Log):** 事件将记录在一个特定的日志文件中(各平台都实现)。
 - **系统日志(System Log):** 对不同的系统将特定级别的事件加入其系统日志中。
- **EterneCluster 核心(EterneCluster Core):** 简称为“核心”,它是软件的核心部分。它完成了绝大部分的工作。
- **监控服务(Monitor Service):** 简称为“监控”,它的目的是监控 EterneCluster 核心,当 EterneCluster 核心出现异常或故障时,监控服务将重新启动 EterneCluster 核心。该部件是为实现 EterneCluster 软件本身的高可用性而存在,其设计并不复杂。同时其也被 EterneCluster 核心监控,一旦发现监控服务异常,其立即重启监控服务。事实上监控服务和核心是相互监控的。
- **资源监控服务(Resource Monitor Service):** 简称为“资源监控”,它将监控服务器上所有相关的资源,采用外部 dll/so 方式实现,并把这些信息汇报给 EterneCluster 核心。
- **管理界面(GUI):** 提供一个用户友善的界面,这个界面将用 JAVA 来实现,其通过和 EterneCluster Core 通讯来取得配置信息和服务器和 Resource Group 的当前状态并且反映在界面上。
- **切换(Failover):** 是指把发生故障的节点上的 Resource Group 转移到另一节点上。故障切换会在三种不同的情况下发生,人工,自动,或者在特定的时间(由管理软件设定)。自动故障切换包含三个阶段: 1、故障发现。2、资源重新定位。3、重新启动应用程序(一般是故障切换过程中最耗费时间的)。切换又分成下面两种:
 - **远程切换: (Remote Failover):** 是由指定 Resource Group 的主节点发起的一个 Resource Group 的切换过程。通常来说,有两种可能会触发此项操作,一、因为某种错误导致 Resource Group 的中断或主节点重起失败。

二、主节点需要离线转去执行系统维护工作。

- **本地切换: (Local Failover):** 在主节点上, 如果有多个公网网卡 (NIC), 那么我们可以设置其他公网网卡作为主节点的备份 NIC。当主 NIC 出现故障时, EterneCluster 将会执行本地切换, 本地切换会将浮动 IP 地址从故障的公网 NIC 上切换到能正常工作的备份 NIC 上, 这样能减少由于应用在多机上切换带来的巨大的时间消耗。
- **故障恢复: (Failback):** 是故障切换的一种特例, 指由主节点启动的一个 Resource Group 的切换过程。当主节点解除故障, 正常运行后, 如果用户指定自动迁回, 它便会要求备份节点在指定的恢复时间窗口内将资源组切换回来运行, 否则资源组将保留在当前节点上, 等待由管理员在合适的时间进行切换。此操作的主要用途在于: 对客户端来说, 可以继续使用原来由主节点提供的服务响应级别, 这可能是由于原来的主节点的硬件资源好于备份节点, 这样能给客户提供更好的服务响应。
- **资源(Resource):** 我们将在节点中提供一定功能, 使得应用能够正常运行需要的部件都称之为资源。包括物理组件和逻辑组件, 例如磁盘、网络名、IP 地址、数据库、Web 站点、应用程序以及任何其他可以联机 and 脱机的实体。EterneCluster 支持如下几种资源:
 - **IP 地址(IP Address):** 就是前面说的浮动 IP 地址。
 - **主机网络名(network name):** 就是前面说的主机别名
 - **物理驱动盘(Physical driver):** 是应用要存放数据的物理磁盘或者分区, 就是前面说的共享磁盘。
 - **应用(Application):** 就是前面说的应用, 应用是一个十分广泛的概念, 可以是 Windows 下的某些服务或者是一些第三方的应用软件等等从资源所处的位置来说, 我们可以将资源分成下面两类:
 - **内部资源(Internal resource):** 就是这些资源存在于本地资源组上。如前面所说的 IP 等等
- **资源依赖性(Resource dependencies):** 是指资源的可用性通常依赖于其它资源的可用性, 资源间的依赖关系通过资源依赖树(dependency tree)来描述。资源依赖树描述资源启动的先后序列和哪些资源需要共同切换。资源依赖性仅限内同一个资源组内, 不同资源组间不允许有资源依赖关系。
- **资源组(Resource Group):** 是应用和它所依赖的所有资源的组成的一个单元, 是进行故障切换(FailOver)的最小单元, 一个资源组存在于一个节点上。在 EterneCluster 中, 切换是以 Resource Group 为单位的, 切换策略设置是基于 Group 为基础的, 包括首选节点(Preferred Node), Failback 时间等。
- **策略(Policy):** 我们将可以通过配置文件可以进行更改的部分都称之为策略。如心跳超时判断, 服务重启次数定义, 服务器之间的优先级定义等