

# LoadDirector 白皮书



上海嵩卓信息技术有限公司

<http://songzhuo.com/>

2002 年 11 月

# 目 录

<b>前言</b>	<b>集群结束全球等待.....</b>	<b>3</b>
<b>第一章</b>	<b>LOADDIRECTOR——集群系统的核心.....</b>	<b>5</b>
<b>第二章</b>	<b>LOADDIRECTOR 的特点.....</b>	<b>7</b>
	2.1 LOADDIRECTOR 功能简介 .....	7
	2.2 LOADDIRECTOR 采用的三种路由算法 .....	9
	2.3 LOADDIRECTOR 采用的八种调度算法 .....	10
	2.4 LOADDIRECTOR 支持的 TCP/IP 协议 .....	12
	2.5 LOADDIRECTOR 为服务器提供安全保护 .....	13
<b>第三章</b>	<b>LOADDIRECTOR 负载均衡器的应用.....</b>	<b>14</b>
	3.1 LOADDIRECTOR 解决的问题 .....	14
	3.2 LOADDIRECTOR 应用范围 .....	16
	3.3 LoadDirector 应用实例.....	19
<b>第四章</b>	<b>技术规范 .....</b>	<b>20</b>
<b>附录</b>	<b>三种路由选择算法原理简介.....</b>	<b>20</b>

## 前言 集群结束全球等待

互联网络是 20 世纪最富有革命性意义的一种新技术,它不断地改变着人们的工作、学习、生活、娱乐和沟通方式。互联网络正在现实的基础上构筑并发展一个新的时空,它将从根本上改变现存的经济格局和商务模式。网络正成为社会环境中一个新的重要组成部分。

现在越来越多的个人在互联网上购物、娱乐、休闲、与人沟通、获取信息;越来越多的企业把他们与顾客和业务伙伴之间的联络搬到互联网上,通过网络来完成交易,建立与客户之间的联系。互联网的用户数和网络流量正以几何级数增长。由 Web、Mail、Media、DNS、Proxy、FTP 等应用服务器所提供的各种各样的服务是互联网的基石。但是,一旦系统超载或某一项服务意外中断,整个服务系统就会受到影响。特别对于电子商务系统而言,在当今充满竞争的市场里,服务响应缓慢乃至停顿不仅会直接造成经济上的损失,更会导致商业信誉的毁损。因此,无论是企业还是用户都渴望结束等待。

结束等待,你可以给服务器增加更多的内存,或是购买更大的服务器。但不管哪种情况,你总还是要用有停机和大量的支出代价来换取。以前,为了既满足迅速增长的需求,提供可靠的服务能力,又避免硬件在很短的周期内再次更新,唯一的解决办法就是购买比目前所需的有更大计算能力的服务器,但这会增加系统的初始投资和风险。一旦对系统服务能力需求增长到服务器本身的计算能力上限,那么,一个新的购买循环又将开始,旧的服务器必须被替换成计算能力更强大的新服务器。

集群技术,为结束全球等待的需求提供了一种走出这类不断升级淘汰循环的最佳途径。集群技术是将一群服务器组合在一起,共同完成任务,并且对外表现为单一资源的技术。在集群中,一组服务器分担原来由一台大型服务器承担的工作负载。用户访问集群仍像在访问一台高性能的服务器一样。由于集群系统在服务器、操作系统和应用软件上都有冗余,即使有一些服务器失效,其余的服务器也可以接管它们的工作,而对整个系统不会造成影响;如果工作负载超过当前集群系统的处理能力,无须替换原有设备,只需通过动态添加新的服务器到集群服务器组中就可以解决。

**集群为结束全球等待提供最佳方案！**

## 第一章 LoadDirector

### 集群系统的核心

LoadDirector 是集群系统的核心，它是一个独特的具有高可用性的智能化负载均衡设备。LoadDirector 负载均衡器（下面简称 LoadDirector）被放置在路由器和服务器阵列之间，不间断地对每一台后端服务器进行监控，以确保它们都处于正常的工作状态，并且 LoadDirector 会自动把外来的 Internet/Intranet 服务请求分派到最可用的服务器上。

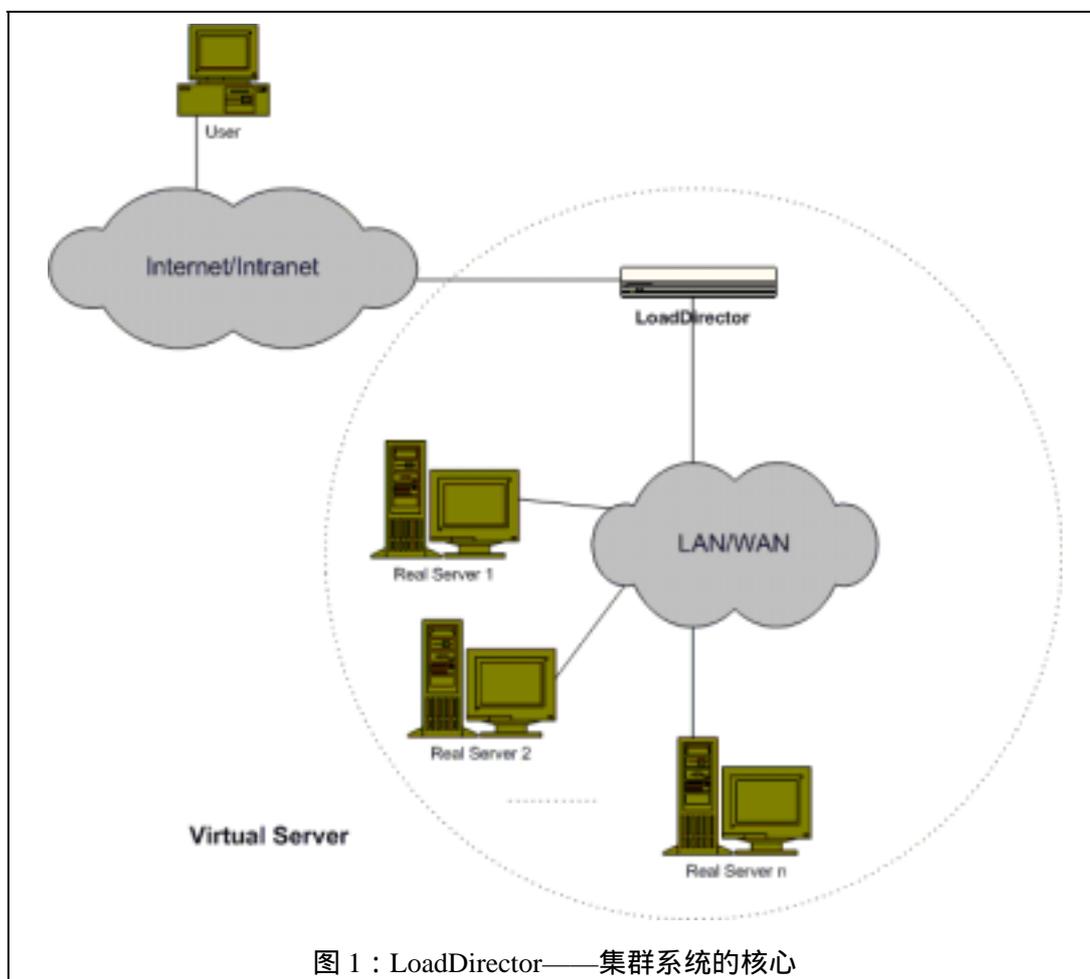


图 1：LoadDirector——集群系统的核心

LoadDirector 提供更经济的伸缩性和高度的容错，它使集群系统进行性能优化和为用户提供个性化的服务成为可能，它使用户享受到最好的服务。

## LoadDirector 给用户的承诺是高可伸缩、高可用和高性价比：

**高可伸缩：**如果集群中的服务器因访问请求过多而超载，就可以往集群中添加新的服务器，直到满足访问需求为止。采用 LoadDirector，网络管理员可以随时根据需要增加服务器，用以完全满足处理能力的要求。

**高可用：**服务器集群要比单个服务器系统更可靠。即使是在集群中的单个服务器出现故障或因为维护而停机时，服务器集群还是可以为不间断的服务。你不必为了某一台服务器的故障而关闭整个站点的服务，因为其他的服务器可以处理所有进来的访问请求。LoadDirector 负载均衡器能够自动地把新的访问请求重新分配到集群中其他的服务器上，这个过程对用户而言是透明的。负载均衡器是整个集群系统中的唯一可能引起系统崩溃的。为避免由于负载均衡器的死机引起系统的崩溃，LoadDirector 可以采用主从备份型的结构，一旦主负载均衡器出错，从负载均衡器就会自动接管其工作。

**高性价比：**用 LoadDirector 构建的服务器集群，可以采用普通的 PC 服务器，性能却可以和一台超级服务器相媲美。系统升级时，只需添加 PC 服务器，而不需要淘汰旧设备，还可以方便地把服务器从访问量少的集群迁移到访问量大的集群，使企业能够最有效地利用现有的硬件设备。

## 第二章 LoadDirector 的特点

### 2.1 LoadDirector功能简介

#### 1. 独创的负载均衡技术

在负载调度中，采用独创的通过 IP 隧道实现虚拟服务器（Virtual Server via IP Tunneling, VS/TUN）技术和通过直接路由实现虚拟服务器（Virtual Server via Direct Routing, VS/DR）技术，充分利用 Internet 访问请求与响应数据量不对称的特点，可调度极大的负载。

#### 2. 高可伸缩

VS/TUN 和 VS/DR 负载均衡技术最多可以调度数百台机架式服务器。当系统负载增加时，只需将普通的机架式服务器通过“热插拔”方式平滑地加入到集群服务器中，从而扩展整个系统的处理能力。

#### 3. 高可用、高可靠性

定时地检测网络服务器和其网络服务进程，将发生故障的服务器从集群中删除，新的服务请求不会被发送到出故障的服务器，而是由其他可用的服务器处理。这样，通过屏蔽有服务器的故障，提高整个网络服务系统的高可用性。LoadDirector 本身也可采用“Heartbeat”协议设计为主从备份的结构，一旦主调度器崩溃，备份调度器会主动接管集群虚拟 IP 地址，承担原来主调度器的职责。

#### 4. 易于维护管理

由 LoadDirector 负载调度产品构建和管理的集群系统，在进行软硬件升级和测试时，服务器能够被温和地切换到离线状态，而与此同时，其他服务器则仍然不间断地为用户和消费者提供服务。

## 5. 跨多种平台

可以将不同生产厂家、不同硬件配置、不同操作系统和不同服务软件的计算机通过 Internet/Intranet 联接在一起，作为一个单一的网站服务系统对外服务。

## 6. 成本低

集群中的实际服务器可以是入门级的台式服务器或机架式服务器，极大地降低系统的总体拥有成本。

## 7. 内置安全机制

LoadDirector 具备防火墙的功能。它使用包过滤、端口映射和网络地址转换等技术限制和拒绝对服务器的访问，它可以抵抗诸如拒绝访问服务等多种黑客攻击手段。此外，它还可以构建防火墙集群系统，提高防火墙的性能和伸缩性。

## 2.2 LoadDirector采用的三种路由算法

### 1. Virtual Server via Network Address Translation (VS/NAT)

NAT (网络地址映射) 是负载调度技术中比较常用的路由选择算法, 它把一段内部网络地址区间转换为一个外部网络地址及其端口。通过地址映射关系, LoadDirector 在重写外部请求报文的源和目的地址后, 根据预设的调度算法, 将请求分派给后端最可用的实际服务器; 实际服务器的应答最后再通过 LoadDirector 重写地址返回给客户, 完成整个负载调度过程。

### 2. Virtual Server via IP Tunneling (VS/TUN)

采用 NAT 技术时, 由于请求和应答报文都必须经过负载均衡器地址重写, 当客户请求越来越多时, 调度器的处理能力将成为系统新的瓶颈。为了解决这个问题, LoadDirector 采用独创的 VS/TUN 技术, 把请求通过 IP 隧道转发至实际服务器, 而实际服务器将应答直接返回给客户。由于一般网络服务应答比请求报文大许多, 采用 VS/TUN 技术后, 集群系统的吞吐量可以提高 10 倍以上。另外, VS/TUN 不要求实际服务器在一个网段中。

### 3. Virtual Server via Direct Routing (VS/DR)

VS/DR 通过改写报文的 MAC 将用户的请求原封不动的转发给后段最可用的服务器。同使用 VS/TUN 技术一样, VS/DR 技术使 LoadDirector 只须处理客户到服务器端的连接, 应答报文直接返回给客户。这种方法没有 IP 隧道的开销, 对集群中的实际服务器也没有必须支持 IP-Tunneling 协议的要求, 但是要求 LoadDirector 与实际服务器都有一块网卡连在同一物理网段上。

在实际应用过程中, 根据集群系统网络环境具体情况不同, LoadDirector 可在三种路由选择算法的组合模式中工作。有关上述算法的具体工作原理请参见附录。

## 2.3 LoadDirector采用的八种调度算法

针对各种性质的网络服务和实际服务器配置，LoadDirector 负载均衡器采用了如下八种负载调度算法：

### 1. 轮叫 (Round Robin)

LoadDirector 通过“轮叫”调度算法将外部请求按顺序轮流分配到集群中的实际服务器上，它均等地对待每一台服务器，而不管服务器上实际的连接数和系统负载。

### 2. 加权轮叫 (Weighted Round Robin)

LoadDirector 通过“加权轮叫”调度算法根据实际服务器的不同处理能力来调度访问请求。这样可以保证处理能力强的服务器处理更多的访问流量。LoadDirector 可以自动问询实际服务器的负载情况，并动态地调整其权值。

### 3. 最少链接 (Least Connections)

LoadDirector 通过“最少连接”调度算法动态地将网络请求调度到已建立的链接数最少的服务器上。如果集群系统的实际服务器具有相近的系统性能，采用“最小连接”调度算法可以较好地均衡负载。

### 4. 加权最少链接 (Weighted Least Connections)

在集群系统中的服务器性能差异较大的情况下，LoadDirector 采用“加权最少链接”调度算法优化负载均衡性能，具有较高权值的服务器将承受较大比例的活动连接负载。LoadDirector 可以自动问询实际服务器的负载情况，并动态地调整其权值。

### 5. 基于局部性的最少链接 (Locality-Based Least Connections)

“基于局部性的最少链接”调度算法是针对目标 IP 地址的负载均衡，目前主要用于 Cache 集群系统。该算法根据请求的目标 IP 地址找出该目标 IP 地址最近使用的服务器，若该服务器是可用的且没有超载，将请求发送到该服务器；若服务器不存在，或者该服务器超载且有服务器处于一半的工作负载，则用“最少

链接”的原则选出一个可用的服务器，将请求发送到该服务器。

#### 6. 带复制的基于局部性最少链接 ( Locality-Based Least Connections with Replication )

“带复制的基于局部性最少链接”调度算法也是针对目标 IP 地址的负载均衡，目前主要用于 Cache 集群系统。它与 LBLC 算法的不同之处是它要维护从一个目标 IP 地址到一组服务器的映射，而 LBLC 算法维护从一个目标 IP 地址到一台服务器的映射。该算法根据请求的目标 IP 地址找出该目标 IP 地址对应的服务器组，按“最小连接”原则从服务器组中选出一台服务器，若服务器没有超载，将请求发送到该服务器，若服务器超载；则按“最小连接”原则从这个集群中选出一台服务器，将该服务器加入到服务器组中，将请求发送到该服务器。同时，当该服务器组有一段时间没有被修改，将最忙的服务器从服务器组中删除，以降低复制的程度。

#### 7. 目标地址散列 ( Destination Hashing )

“目标地址散列”调度算法根据请求的目标 IP 地址，作为散列键 ( Hash Key ) 从静态分配的散列表找出对应的服务器，若该服务器是可用的且未超载，将请求发送到该服务器，否则返回空。

#### 8. 源地址散列 ( Source Hashing )

“源地址散列”调度算法根据请求的源 IP 地址，作为散列键 ( Hash Key ) 从静态分配的散列表找出对应的服务器，若该服务器是可用的且未超载，将请求发送到该服务器，否则返回空。

## 2.4 LoadDirector支持的TCP/IP协议

LoadDirector 负载均衡器能够支持绝大多数的 TCP 和 UDP 协议：

协议	内容
TCP	HTTP ,HTTPS ,FTP ,SMTP ,POP3 ,IMAP4 ,SSMTP ,POP3S , IMAPS , DNS , LDAP , SIP 等
UDP	DNS , NTP , ICP , SIP , RTP , 视频、音频流播放协议等

## 2.5 LoadDirector为服务器提供安全保护

在 LoadDirector 的设计中，本身就有许多固有的安全特性，这些设计的目的就在于使服务器能够抵御恶意的网络攻击，为用户的网络设备提供额外的安全保护，比如：

### 1. 安全网络地址转换（Secure Network Address Translation）

实际服务器可使用保留 IP 地址，对外部 Internet 不可见。对外可见的只有集群的虚拟 IP 地址和负载均衡的服务端口。这样，即使后端服务器中出现安全漏洞，因为后段服务器对外部不可见，所以也不会影响系统的安全性。

### 2. 使用 LoadDirector 可以严格控制虚拟 IP 地址上允许访问和禁止访问的端口。

### 3. 抗攻击保护

- 提供报过滤。
- 支持 OpenSSH 进行远程的基于命令行的管理。
- 支持基于 HTTPS 的基于浏览器式管理。
- 消除空闲连接(抵抗拒绝服务的攻击)。
- 进行源路由的跟踪（抵抗 IP 欺骗）。
- 保护自身和服务器免受 ICMP 的攻击。
- 自身能抵抗 SYN-Flooding 的攻击。
- 不运行不必要的守护进程（如 sendmail，ftpd，telnetd 等），和其它任何可能受攻击的服务进程。

## 第三章 LoadDirector 负载均衡器的应用

### 3.1 LoadDirector 解决的问题

总的来说，LoadDirector 可以解决的问题包含四个重要方面，但在实际应用中，LoadDirector 往往解决的是下面多个问题的综合。

#### 1. 服务器宕机问题

如果站点的服务器崩溃了，无论是 HTTP、DNS、Email 服务还是 FTP 服务，也无论是因为硬件还是因为软件的错误，系统都将无法提供正常的 Internet 服务。如果有两台提供相同服务的服务器，并通过 LoadDirector 在两者之间调度负载，即使一台服务器出错，LoadDirector 也会将 Internet 访问请求调度到正常工作的服务器上，站点的可用性是 100%。此外，LoadDirector 是自动诊断代理，它会周期性地测试服务器，以确认服务器是否正确地处理 Internet 请求。

#### 2. 系统伸缩性问题

从本地计算机商店购买的 PC 作为一般 Web 服务器的速度几乎和你所能买到任何计算机的速度相差无几。但是用来处理热门站点的请求还是远远不够的。如果通过 LoadDirector 将站点建立在服务器集群的架构之上，就可以建立你所要的任何强大的站点。当访问量增加时，只需增加更多的服务器。

#### 3. 计算机处理能力不相同问题

一个使用多个服务器来构建的大站点的服务器往往在不同的时期购买，这样会带来这样一个问题，即有些服务器的性能要比其它服务器要高得多。LoadDirector 可以根据这些服务器各自的处理能力来分配负载，在任何时刻给最大处理能力的系统更多的负载。同采用简单 Round-Robin 方法分配请求相比，这种智能化均衡在总吞吐量上能起到非常有价值的作用。

#### 4. 网站不同的内容或服务类型的访问量的差异很大的问题

一般来说，访问某些内容或服务的频率要比其它一些内容或服务要高得多。LoadDirector 可以将特定的访问请求发送给指定的服务器，以提高这些访问请求的处理效率。例如，热门的内容可以复制到多台服务器，而访问相当少的内容可以存在一台服务器上，这样就可以提高系统资源的利用率，缩短用户请求的响应时间。

## 3.2 LoadDirector 应用范围

得益于 LoadDirector 的 Internet 应用主要有 Web 站点（包括在线服务、Web hosting 服务、Application Hosting 服务等）、e-commerce（包括在线销售—B2C、在线交易—C2C、B2B 服务等）、大容量邮件、Cache、FTP、防火墙、DNS、ASP 等。

### 1. LoadDirector 结束 Web 等待

Web 可以说是 Internet 上应用最为广泛的服务。随着公司越来越多地将它们的业务通过基于 Web 的应用在网络进行，对高可伸缩、高可用的负载均衡技术的需求也快速增长。每天 24 小时每星期 7 天的正常运行不再只是目标，而是必须。Web 站点的集群解决方案几乎涉及 LoadDirector 所能解决的所有问题。

LoadDirector 不仅为 Web 站点提供了一个高可伸缩、高可用、高容错、非对称调度和个性化服务的虚拟服务器，而且还可为其它网络设备如防火墙等提供负载均衡。LoadDirector 内置的安全性也降低了系统的总体拥有成本。

### 2. LoadDirector 增强 e-commerce 的可用性、可靠性和安全性，提高 QoS

电子商务应用的持续增长给商务和 IT 管理都提出了严峻的挑战。从商务上讲，它已成为企业不得不从战略上考虑的重要运营手段；从技术上讲，服务器处理能力、伸缩性、固定连接和容错性等问题都是必须仔细考虑的。公司的品牌、服务质量都和站点的运行状况密切相关。

从这个意义上讲，电子商务对网络的基础架构提出了比一般 Web 站点更加独特的要求。从技术上讲，对服务器能力、伸缩性、容错可靠性有更高的要求。LoadDirector 为此提供了全面的解决方案。

除了具备单 LoadDirector 结构的强大性能、高可伸缩、安全等特点外，主从备份的双 LoadDirector 结构提供了更为强壮的可用性、容错性。

在电子商务应用中，还需要在服务器上始终保持用户访问站点的状态信息，如购物过程，这是非常重要的。在集群系统中，LoadDirector 负责维持用户和同一个服务器之间的这种固定连接。

### 3. LoadDirector 给大容量邮件系统带来性能、伸缩性和费效比等方面的优势

电子邮件是 ISP 们提供的主要服务，并且是关系到 ISP 提供服务质量的关键服务。不管 ISP 提供的电子邮件服务具备怎样的其它功能，如果不能保证邮件系统随时可用，则意味着客户的大量流失。邮件集群系统提出一种新型的 Internet 邮件系统构架，以负载均衡器为核心设备，具备集群产品高性能、高可用性和高可伸缩性的特点，可支持高达 1 千万的邮件用户。用多台机架式服务器构造邮件服务系统。我们的邮件集群系统支持 20 万用户只需 8 ~ 10 台机架式服务器即可。

### 4. LoadDirector 用于 Cache 集群系统方案

有效的网络 Cache 系统可以大大地减少网络流量、降低响应延时以及服务器的负载。但用大型服务器构造 Cache 服务器成本太高，而且一台大型服务器处理用户 Web 请求的能力和系统的伸缩能力有限。LoadDirector 可以将一组 Cache 服务器构造成对用户透明的一台高性能、高可靠的 Cache 服务器，可以极大地提高网络的带宽利用率和 Cache 系统的响应速度，适合于大中型 Internet 系统，如构建国家 Cache 网或 ISP 接入。

### 5. 其它应用

- **不间断服务和关键性应用需要可靠的容错机制**

许多客户的问题是由于意外情况而引起的主机崩溃会造成整个系统无法提供服务。发生崩溃后需要专人进行修复、重新配置，然后重新启动系统。LoadDirector 的目标就是可以提供不间断服务。

- **不断增长的业务发展需要灵活的可扩展性**

一些客户的问题是业务发展后计算机系统无法适应新的需求，简单地更换系统无法保护以前的投资，新的系统需要重新配置，应用程序需要移植/修改。这类客户包括小型的 ICP、电子商务和企业集团。

- **突发访问密集的应用**

视频点播系统，电子商务系统，多媒体查询系统。

- **信息家电系统/社区服务系统**

这可能是目前最有潜力的客户群。集群系统+ Linux 的嵌入式应用+ Home Network，可以有非常广大的市场。可以用某种机制，将各种信息家电结合，共享资源，统一管理方式等。

### 3.3 LoadDirector 应用实例

LoadDirector 的原型系统已被众多世界级的站点所采用，其中包括：

- 英国国家 JANET Cache 网
- real.com
- abc.com (美国广播电视公司)
- empas.com (韩国搜索引擎和门户网站)
- sourceforge.net
- www.linux.com
- www.netwalk.com
- ctrip.com (携程网)

特别值得一提的是：

1. 英国国家 Cache 网(www.cache.ja.net),它是为英国 150 多家大学和地区网络提供服务的国家级 Cache 网。其系统原为 56 台相互独立的 Cache 服务器，后采用 LoadDirector 的原型，将仅仅 30 台 Cache 服务器组成 Cache 集群系统，而其访问速度却比原来加快许多。
2. Real.com 是为其全球用户提供网上媒体服务。以往采用几台超级服务器为客户提供服务。采用 LoadDirector 原型后，使用 20 台双 CPU 的 PC 服务器的集群为其全球用户提供视频和音频服务。用 Real.com 技术主管的话来讲，LoadDirector 在性能、可靠性、可管理性上打败了其它任何商业化的负载调度产品。

## 第四章 技术规范

### 1. 性能：

配置 100M 网卡，采用 VS/TUN 或 VS/DR 路由算法，集群系统的总吞吐量可达 1Gbits/s；如配置千兆网卡，则系统的总吞吐量可达 10Gbits/s。

### 2. 后端服务器操作系统兼容性：

任何支持 TCP/IP 的操作系统，包括 Linux，各种 Unix（如 FreeBSD、Solaris、IBM AIX、HP Unix 等），Mac/OS 和 Windows NT/2000。

### 3. 网络协议支持：

支持各种 TCP/IP 服务。

### 4. 管理环境支持：

SMTP，OpenSSH，动态/静态网络监控，经过调度的批处理，系统状态报告和事件通知警报。

### 5. 网络管理和监控：

基于浏览器的 SSL 管理界面，远程加密登录，实时网络流量统计和报表生成，OpenSSH 文件传输以及其他管理监控工具。

### 6. 动态内容支持：

ASP (Active Server Pages), VB (Visual Basic Script), ActiveX, JSP (Java Server Page), PHP, VRML, CGI, Cool Talk, Net Meeting, Real Audio, Real Video, Netshow, Quick Time, PointCast, 以及任何 HTTP 封装数据。

### 7. Web 服务器兼容性：

任何 Web 服务器，包括 Apache，Microsoft Internet Information Server，Netscape Enterprise Web Server 等。

## 8. 硬件形态：

体 积	428.6 × 222.5 × 44.4 mm <sup>3</sup> ( 1U 标准机架式机箱 )
重 量	4.8 公斤
网 卡	100 BASE-TX × 4 , 千兆 Ethernet 网卡 ( 可选 )
闪 存	8MB
内 存	128MB ( 可扩展至 1GB )
温 度	0 — 45
湿 度	10% — 96% ( 非冷凝 )
电 源	110VAC、220VAC ( 可切换 ) , 50/60Hz

## 附录 三种路由选择算法原理简介

### 1. Virtual Server via Network Address Translation ( VS/NAT )

由于 IPv4 中 IP 地址空间的日益紧张和安全方面的原因，很多网络使用保留 IP 地址 (10.0.0.0/255.0.0.0，172.16.0.0—172.31.255.255 和 192.168.0.0/255.255.0.0)。它们在 Internet 上未被分配，是为组建内部网络预留的。当内部网络中的主机要访问 Internet 或被 Internet 访问时，就需要采用网络地址映射技术 ( Network Address Translation, 以下简称 NAT )，将内部地址转化为对访问者可见的外部地址。

在负载调度技术中，NAT 技术也为许多的厂商所采用，如 Cisco, F5,等。

NAT 技术的原理如下图：

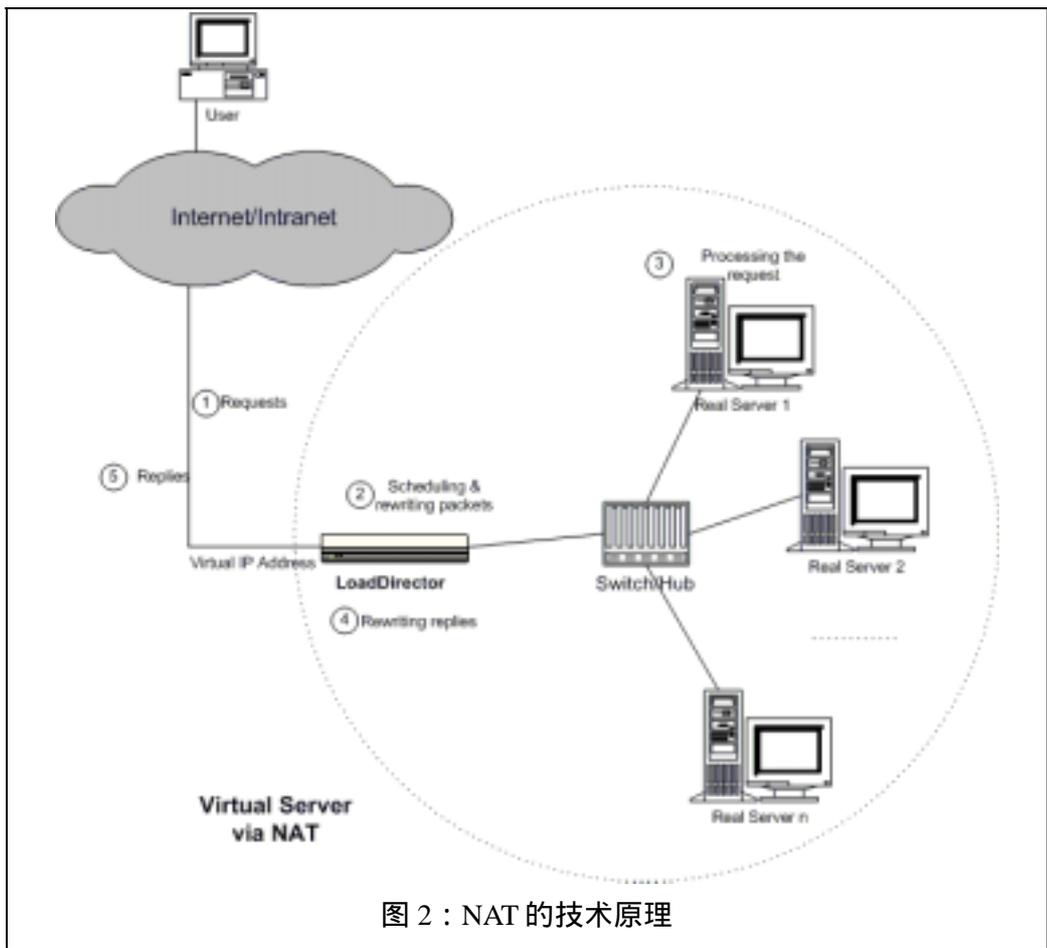


图 2：NAT 的技术原理

## 用户请求

用户发出访问由服务器集群提供的服务的请求，以虚拟 IP 地址（负载均衡器对外的 IP 地址）为目标地址的请求包到达负载均衡器。

## 调度和重写请求包

负载均衡器接受到请求包后，对照虚拟服务器规则表，检查包中的目的地址和端口，如果它们与虚拟服务相符，就按照某种调度算法从服务器集群中选择一个实际服务器，并且在 Hash 表中增加一项以记录建立的连接。然后包的源地址改写为调度器的地址，目标地址和端口被重新写为所选择的服务器的地址和端口，并且被向前转送给实际服务器。当接受到的请求包属于已有连接时，所选的服务器会在 Hash 表中找到，包会被重写并转送给所选的服务器。

## 处理请求

实际服务器接收到请求，处理完毕后，将应答发送给负载均衡器

## 重写应答

当应答返回时，负载均衡器用虚拟服务的地址和端口重写应答的源地址和端口，目的地址改和端口写为访问用户的地址和端口。当连接终止或超时，连接记录会从 Hash 表中删除。

## 应答返回

应答返回时，它的源地址和端口就变为虚拟服务器对应的源地址和端口。

基于 NAT 的的集群系统可以适合许多服务器的性能要求。如果负载均衡器成为系统新的瓶颈，可以有两种方法解决这个问题：混合方法和 IP Tunneling 或 Direct Routing。在 DNS 混合集群系统中，有若干个负载均衡器，每个负载均衡器带自己的服务器集群，同时这些负载均衡器又通过 RR-DNS 组成简单的域名。但更好的解决方法是 IP Tunneling 或 Direct Routing。

## 2. Virtual Server via IP Tunneling ( VS/TUN )

在基于 NAT 的集群系统中，请求和应答的数据包都需要通过负载均衡器，当实际服务器的数量超过 20 时，负载均衡器将成为集群系统新的瓶颈。许多 Internet 请求都有这样的特点：即请求包较短而应答往往包含大量的数据。如果能将请求和应答分开处理，即在负载均衡器中只通过一次，将极大地提高系统的吞吐量。

在基于 IP Tunneling 技术的集群系统中，负载均衡器只将请求调度到不同的实际服务器，实际服务器将应答的数据直接返回给用户。这样，负载均衡器就可以处理巨量的请求，它可以调度百台以上的服务器，而且它也不再是系统的瓶颈。其原理如下图所示。

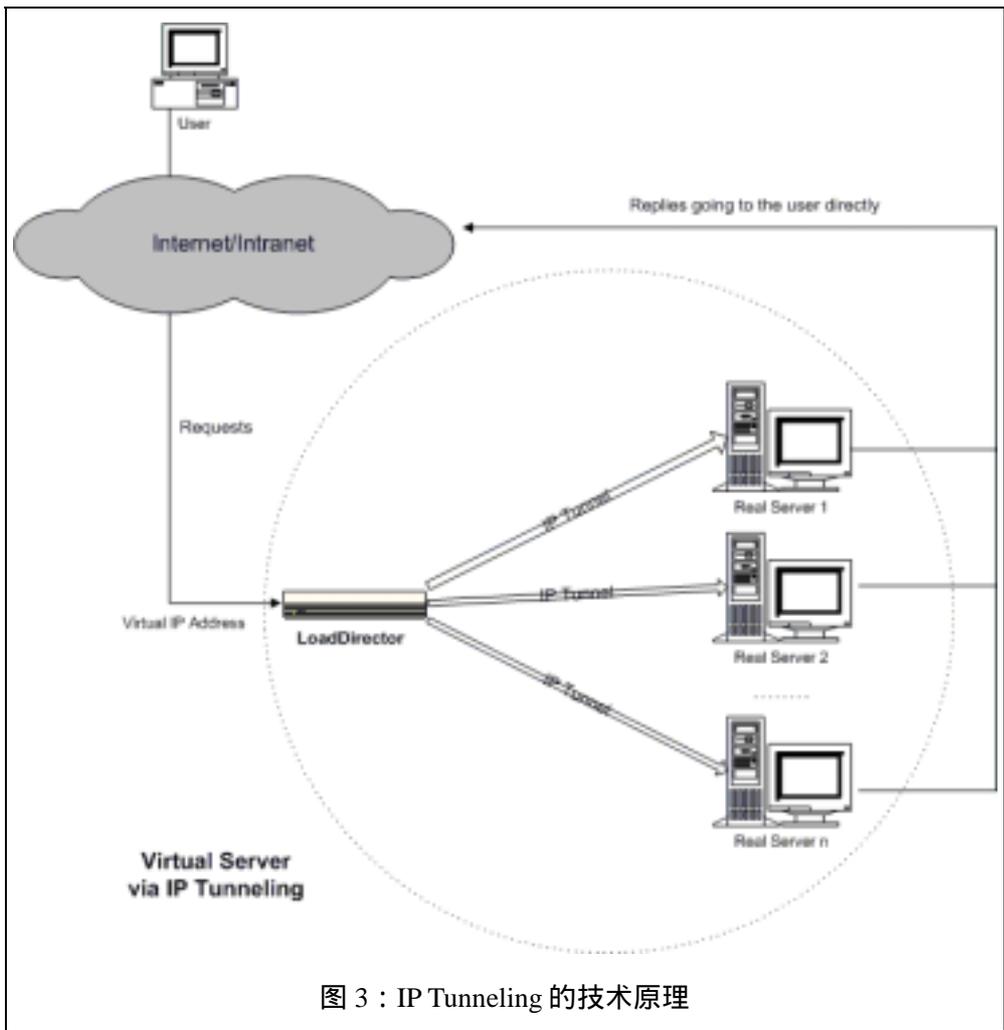


图 3：IP Tunneling 的技术原理

这样，使用 IP Tunneling 将极大地增加负载均衡器调度的服务器的最大数量。即使负载均衡器只有 100Mbps 的全双工网卡，虚拟服务器的最大吞吐量可以达到几千兆 bps。IP tunneling 技术可以用来构建高性能超级服务器。

IP Tunneling 技术对服务器有限制，即所有的服务器必须支持“IP Tunneling”协议。

### 3. Virtual Server via Direct Routing (VS/DR)

同 IP Tunneling 技术一样，Direct Routing 只处理客户到服务器端的连接，应答的数据可以直接从独立的网络路由返回给客户。这可以极大地提高虚拟服务器的伸缩性。

同 IP Tunneling 技术相比，这种方法没有 IP 隧道的开销，但是要求负载均衡器与实际服务器都有一块网卡连在同一物理网段上。其结构如下图。

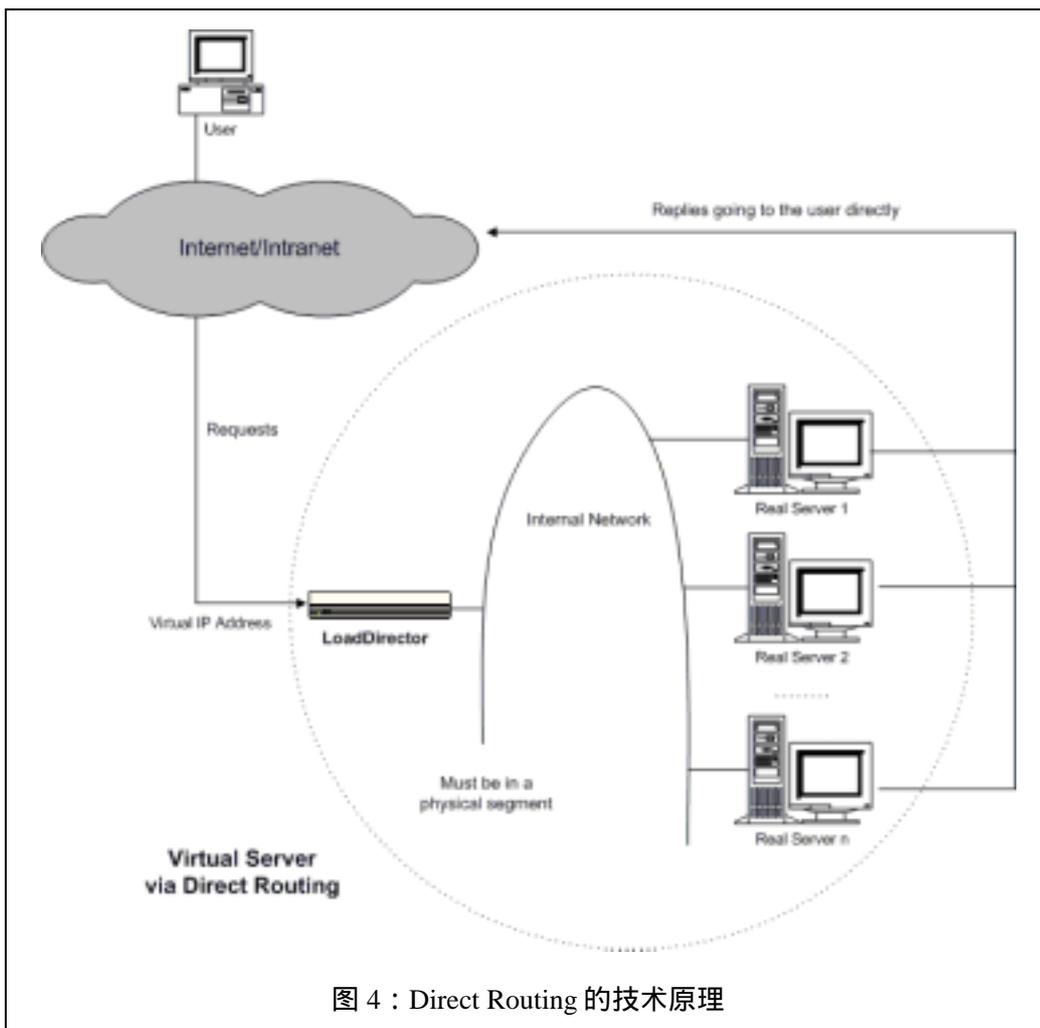


图 4：Direct Routing 的技术原理