

服务器集群技术简介

背景

长期以来，科学计算、数据中心等领域一直是高端RISC服务器的天下，用户只能选择IBM、SGI、SUN、HP等公司的产品，不但价格昂贵，而且运行、维护成本高。

随着Internet服务和电子商务的迅速发展，计算机系统的重要性日益上升，对服务器可伸缩性和高可用性的要求也变得越来越来高。RISC系统高昂的代价和社会旺盛的需求形成强烈的反差。

集群技术的出现和IA架构服务器的快速发展为社会的需求提供了新的选择。它价格低廉，易于使用和维护，而且采用集群技术可以构造超级计算机，其超强的处理能力可以取代价格昂贵的中大型机，为行业的高端应用开辟了新的方向。

集群技术是一种相对较新的技术，通过集群技术，可以在付出较低成本的情况下获得在性能、可靠性、灵活性方面的相对较高的收益。

目前，在世界各地正在运行的超级计算机中，有许多都是采用集群技术来实现的。

集群是由一些互相连接在一起的计算机构成的一个并行或分布式系统。这些计算机一起工作并运行一系列共同的应用程序，同时，为用户和应用程序提供单一的系统映射。从外部来看，它们仅仅是一个系统，对外提供统一的服务。集群内的计算机物理上通过电缆连接，程序上则通过集群软件连接。这些连接允许计算机使用故障应急与负载平衡功能，而故障应急与负载平衡功能在单机上是不可能实现的。

服务器集群系统通俗地讲就是把多台服务器通过快速通信链路连接起来，从外部看来，这些服务器就像一台服务器在工作，而对内来说，外面来的负载通过一定的机制动态地分配到这些节点机中去，从而达到超级服务器才有的高性能、高可用。

集群的优点

- **高可伸缩性：**服务器集群具有很强的可伸缩性。随着需求和负荷的增长，可以向集群系统添加更多的服务器。在这样的配置中，可以有多个服务器执行相同的应用和数据库操作。
- **高可用性：**高可用性是指，在不需要操作者干预的情况下，防止系统发生故障或从故障中自动恢复的能力。通过把故障服务器上的应用程序转移到备份服务器上运行，集群系统能够把正常运行时间提高到大于99.9%，大大减少服务器和应用程序的停机时间。
- **高可管理性：**系统管理员可以从远程管理一个、甚至一组集群，就好象在单机系统中一样。

集群与管理

构建集群系统必须包含对系统及网络管理的两方面的考虑。服务器集群十分复杂，而复杂的技术又往往会引入许多人为的错误，因此系统应有网络资源管理、系统监测管理，并具有可以简化管理过程的工具。

如果仅仅把集群视为单一系统或把它视为分立的服务器，那么这种管理软件是不能胜任集群管理工作的。当我们观察集群上运行的一个应用程序时，需要站在单一系统角度；当我们试图区分、定位一个出错部件时，又需要站在分立服务器角度。如果管理系统不能提供必需的监测及管理能力，那么该集群是不能在重要的应用环境中投入使用的。

集群与成本

并不是所有的服务器都需要采用系统级的冗余。因此，我们需要对比一下是系统发生故障所造成的损失大，还是购买及管理一个集群系统的费用高，从中找出一个较好的选择。

为什么采用集群技术？

对于信息化时代的高等院校来讲，数据中心的建立是非常必要的（对于某些大学来讲，建立高性能计算中心也是非常重要的）。那么，我们到底需要什么样的服务器才能满足目前和未来发展的需要呢？

首先要有超强的计算能力，可以承受长时间、大量用户的并发访问。

其次，服务器系统的高可用性、易用性和易管理性对于用户也是极为重要的，如果系统出现故障，造成服务中断或者重要资料丢失，会给学校造成重大的损失。

第三，随着业务的增加、访问量的提高和数据的不断堆积，系统会变得越来越慢，所以硬件设备的更新换代是一个必不可少的工作。

从以上三点出发，可以考虑采用集群技术来实现。

采用集群技术的好处

- 1、**提高性能**：一些计算密集型应用，需要计算机要有很强的运算处理能力。这时，可以采用计算机集群技术来满足计算要求。
- 2、**降低成本**：在达到同等性能的条件下，采用计算机集群比采用同等能力的计算机所花的代价要小很多。
- 3、**提高扩展性**：采用传统服务器的用户如果需要大幅度扩展系统的能力，就必须购买昂贵的最新的服务器。如果该服务系统采用集群技术，则只需要将新的服务器加入集群中即可。
- 4、**增强可靠性**：集群技术可以使系统在故障发生时仍继续工作，将系统停运时间减到最小，大大提高了系统的可靠性。

集群的类型

集群系统 (Cluster)，有时也称为机群或群集系统。

集群技术本身有很多种分类，市场上的产品很多，也没有很标准的定义，较为常见的主要分为三种类型。

- 高可用性集群 (High Availability Cluster) / 容错集群 (Fail-over Cluster)
- 负载均衡集群 (Load balancing Cluster)
- 高性能计算集群 (High Performance Computing Cluster)

高可用性集群

当集群中的一个系统发生故障时，集群软件迅速做出反应，将该系统的任务切换到集群中其它正在工作的系统上执行。

考虑到计算机硬件和软件的易错性，高可用性集群的目的主要是为了使集群的整体服务尽可能可用。如果高可用性集群中的主节点发生了故障，那么这段时间内将由次节点代替它。次节点通常是主节点的镜像，所以当它代替主节点时，它可以完全接管其身份。

高可用性(HA)集群致力于使服务器系统的运行速度和响应速度尽可能快。它们通常利用在多台机器上运行的冗余节点和服务进行相互跟踪。如果某个节点失败，它的替补将在几秒钟或更短时间内接管它的职责。因此，对于用户而言，群集永远不会停机。

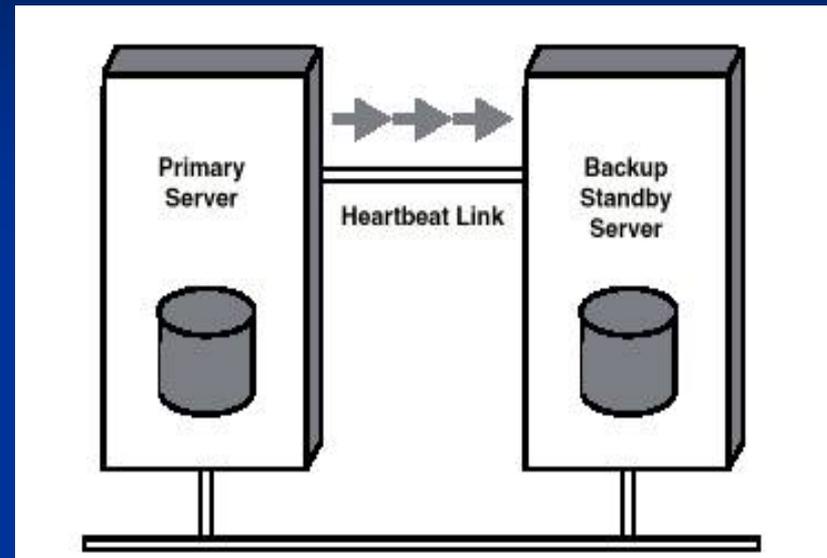
有些HA集群还可以实现节点间冗余应用程序。即使用户使用的节点出了故障，他所打开的应用程序仍将继续运行，该程序会在几秒之内迁移到另一个节点，而用户只会感觉到响应稍微慢了一点。但是，这种应用程序级冗余要求将软件设计成具有集群意识的，并且知道节点失败时应该做什么。

高可用性集群拓扑结构

有两种典型的拓扑结构可以实现高可用性：主从服务器和活动第二服务器。

主从服务器

通常把一个服务器安排为“主”服务器，一个服务器为“第二”服务器；由主服务器为用户提供服务，第二服务器除了在主服务器出错时接管工作外，没有其它用处。



两台服务器通过一种被称为“心跳”（heartbeat）的机制进行连接，用于监控主服务器的状态，一旦发现主服务器宕机或出现不能正常工作的情况，心跳会通知第二服务器，接替出问题的主服务器。

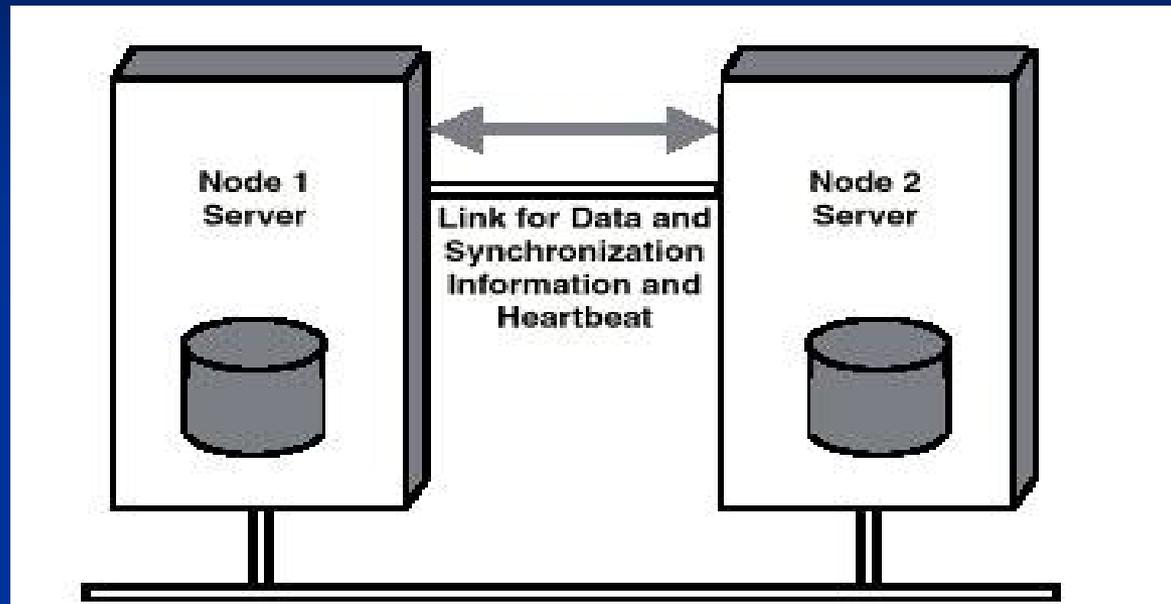
“心跳”可以通过专用线缆、网络链接等方式实现。

活动第二服务器

一个功能基本一致但成本低得多的方式是使第二服务器可以处理其它的应用程序，当主服务器发生故障时，第二服务器能够接管主服务器的工作。这种被称为“活动第二服务器”方法的主要优点是在保持使用第二服务器的同时，获得服务器冗余，而不是仅仅把第二服务器作为备份使用。这种方法可以降低集群系统的运行费用。

活动第二服务器有三种实现形式：“全部复制”、“0共享”和“全部共享”。

“全部复制”方式



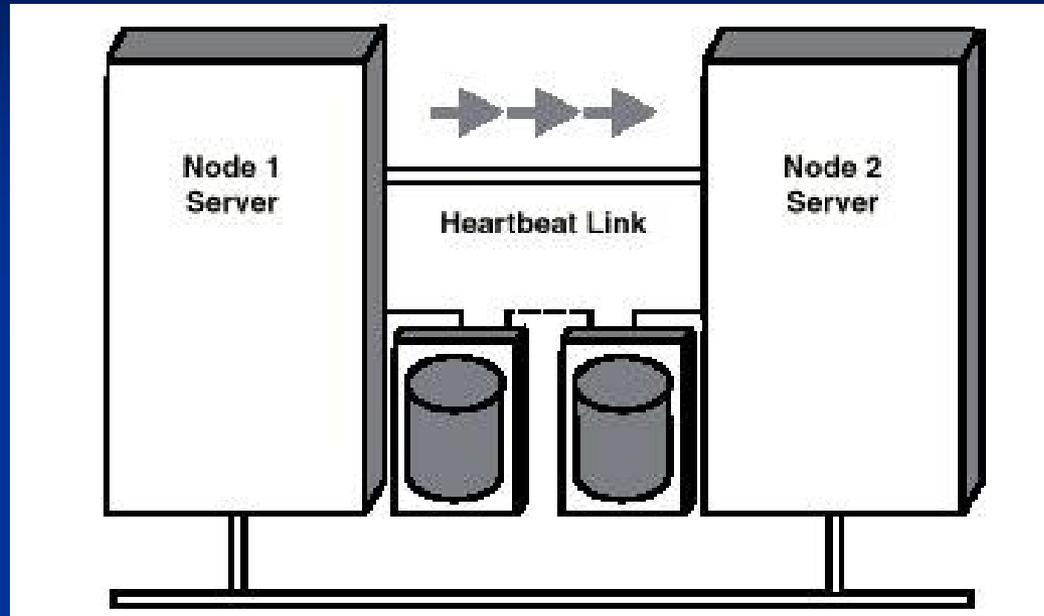
第一种方式称作“全部复制”，就是指彻底的服务器冗余。每个服务器都有自己的磁盘。数据不断地被拷贝到第二服务器的磁盘上，以保证故障发生时，第二服务器可以使用当前的数据。

“全部复制”方式的优缺点

这种方法增加了服务器及网络的负荷，可能会严重影响系统性能。另一个缺点是当一个服务器发生故障时，可能会有主、从服务器的不一致现象：某个磁盘上的事务处理可能并没有完全在另一个磁盘上得到继续。因为即便以最快的网络相连，两个服务器间信息传送仍会有一定的延迟。

由于数据被完全复制，所以应用程序可以在任一服务器上运行，从而可以更好地平衡负载。另外，节点在空间上可以是分散的，节点间可通过广域网互连，物理上可以距离很远。

“0共享”方式

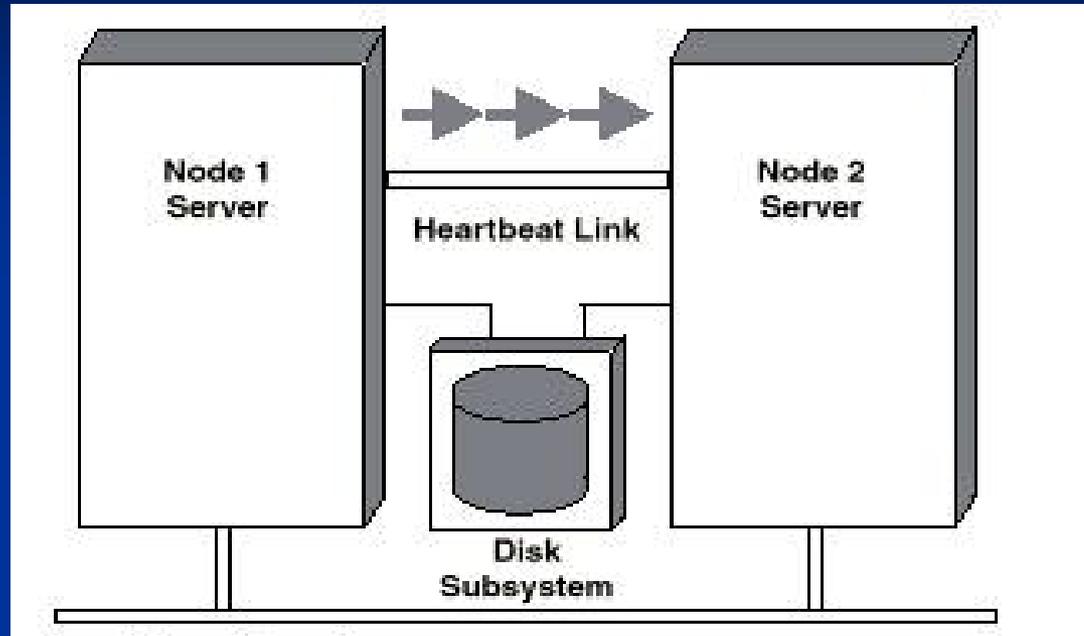


第二种方式是“0共享”，它是指两个服务器物理上连接到同一个磁盘组上，每个服务器都“拥有”自己的磁盘。在正常情况下，各服务器只能存取各自的数据。当一台服务器发生故障时，另一台服务器自动获得对方磁盘的读写权限，并对之进行操作。

“0共享”方式的优缺点

这种方式不必不停地在服务器间拷贝数据，从而大大降低了网络负荷。在这种方式下，磁盘是唯一可能产生导致长时间停机故障的地方。所以，在这种体系结构中，一般在磁盘子系统中采用RAID技术或者采用冗余镜象磁盘，以保证在磁盘出错时，应用程序和数据的可用性。

“全部共享”方式



第三种方式是“共享一切”，也就是说让多个服务器在同一时间共享同一磁盘。在这种方式中，所有与磁盘相连的服务器在正常运行时可在相同时刻共享磁盘存取通道。这种方式要求开发一个复杂的锁定管理软件，保证在一个时刻只有一个服务器在读写数据。

“全部共享”方式优缺点

“共享一切”方法也需要采用RAID技术或磁盘镜像来保障数据的安全。

由于每个服务器的数据被送到相同的磁盘上，所以不必在服务器间复制数据。这样在正常运作时，可以在不影响系统性能的前提下，提供高可用性。

高可用性集群对比一览表

集群方法	描述	优点	局限性
主从服务器	只是在主服务器发生故障时，第二服务器才能投入运行，接管一切。	易于实现。	成本高。因为第二服务器不能处理其它任务。
活动第二服务器	第二服务器也被用来运行任务处理。	成本低。因为第二服务器也能运行。	复杂性增加。
“全部复制”	每个服务器都有自己的磁盘。主、次服务器之间不停地进行数据拷贝。	高可用性和容错。适合于对可用性敏感的环境。	拷贝操作使网络及服务器负荷很大。可能会有发生不同步的风险。有故障发生时，可能会有丢失事件。应用程序需要全面的修改。
“0共享”	服务器连到相同的磁盘系上，但每个服务器都拥有属于自己的磁盘，如果某个服务器出错，它的磁盘将由另一服务器接管。	因为无需拷贝数据，所以降低了网络及服务器的一般运行开销。	通常需要磁盘镜像或RAID技术来补偿磁盘故障给系统带来的灾害。
“全部共享”	多服务器可同时共享磁盘存取。	低网络及服务器运行开销。由于磁盘故障而引发系统停机的风险被降低。	需要锁定管理软件；需要磁盘镜像或RAID技术。

负载均衡集群

对于高等院校来说，负载均衡集群是应用面最广、最有发展潜力的集群应用形式。

随着教育信息化的发展，用户需求成几何方式增长，从而导致网站的访问量不断增长。由于学校资源中很多是多媒体素材，包括音频、视频、动画等，会产生极大的网络流量负载和处理负载。在这种情况下，采用负载均衡集群是一个极好的选择。

负载均衡集群一般用于WEB服务器、代理服务器等。这种集群可以在接到请求时，检查接受请求较少，不繁忙的服务器，并把请求转到这些服务器上。

网络负载均衡功能增强了Web服务器、流媒体服务器和终端服务等Internet服务器程序的可用性和可伸缩性。

负载均衡集群的特点

所有节点对外提供相同的服务，这样可以实现对单个应用程序的负载均衡，而且同时提供了高可用性，性能价格比极高。

网络流量负载均衡是一个过程，它检查集群的入网流量，然后将流量分发到各个节点以进行适当处理。负载均衡网络应用服务要求群集软件检查每个节点的当前负载，并确定哪些节点可以接受新的作业。因此，集群中的节点(包括硬件和操作系统等)没有必要是一致的。

负载均衡集群提供了一个非常实用的解决方案。负载均衡集群使负载可以在计算机集群中尽可能平均地分摊处理。负载通常包括应用程序处理负载和网络流量负载。这样的系统非常适合向使用同一组应用程序的大量用户提供服务。每个节点都可以承担一定的处理负载，并且可以实现处理负载在节点之间的动态分配，以实现负载均衡。对于网络流量负载，当网络服务程序接受了太多入网流量，以致无法迅速处理，这时，网络流量就会发送给在其它节点上运行的网络服务程序。同时，还可以根据每个节点上不同的可用资源或网络的特殊环境来进行优化。

高性能计算集群

高性能计算集群具有响应海量计算的性能，主要应用于科学计算、大任务量的计算等。有并行编译、进程通讯、任务分发等多种实现方法。

因为高性能计算集群涉及为解决特定的问题而设计的应用程序，针对性较强，后面再做具体的介绍。

在集群的这三种基本类型之间，经常会发生混合。高可用性集群可以在其节点之间均衡用户负载。同样，也可以从要编写应用程序的集群中找到一个并行集群，使得它可以在节点之间执行负载均衡。从这个意义上讲，这种集群类别的划分只是一个相对的概念，而不是绝对的。

WINDOWS 集群解决方案

有网络负载均衡功能的 Windows 2003 和 Windows server 2008 为在分布和负载均衡的方式下建立关键且合乎要求的网站的工作提供了良好的基础结构。与组件服务的分布式应用程序特性和IIS的可伸缩性相结合，网络负载均衡有助于确保系统能够灵活地处理繁重的通信任务，同时，保持对服务器状态的监控，确保系统不停机。

WINDOWS 集群技术的特点

- 可伸缩性：可以加入更多的处理器或计算机以提高集群的计算能力。
- 高度的可用性：集群具有避免单点故障发生的能力。应用程序能够跨计算机进行分配，以实现并行运算与故障恢复，并提供更高的可用性。即便某一台服务器停止运行，一个由进程调用的故障应急程序会自动将该服务器的工作负荷转移至另一台服务器，以保证提供持续不断的服务。
- 易管理性：集群以单一系统映射的形式来面向最终用户、应用程序及网络，同时，也为管理员提供单一的控制点，而且这种单一控制点还可以是远程的。

Windows的两种集群

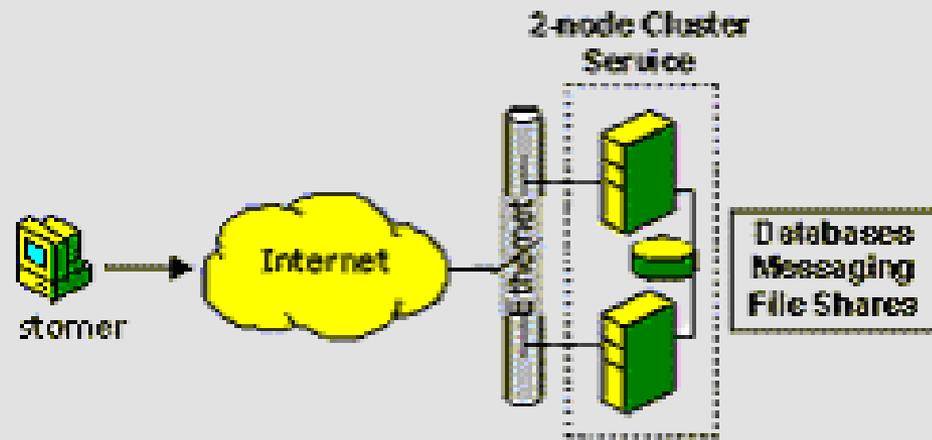
在Win2000操作系统中，微软引进了两种既能独立使用又能联合使用的集群技术，为用户提供了一套能够根据应用程序或服务需要而选用的高可用的、扩展能力强的集群解决方案。

- 高可用性集群：Windows 2000 高级服务器版支持双节点故障应急集群，数据中心服务器版支持四节点集群。
- 负载均衡集群：Windows 2000 高级服务器版可在多达32个节点的集群内均衡IP通信量。

负载均衡与高可用性联合使用

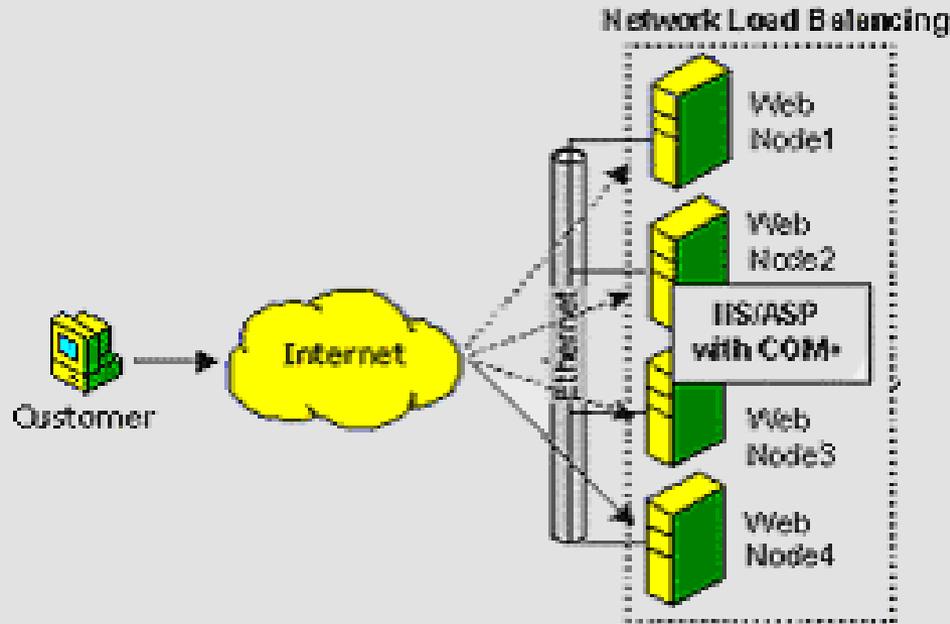
Windows 的这两种集群技术可联合使用以创造具有高度可伸缩性和可用性的多层应用服务系统。通过在前端的Web服务器区中配置网络负载均衡，并将后端的在线应用程序集群化，用户就能在不受服务器或应用程序单点故障困扰的情况下获得接近线性的可伸缩能力所带来的全部好处，并能够为满足需求而迅速做出调整。

高可用性集群系统实现



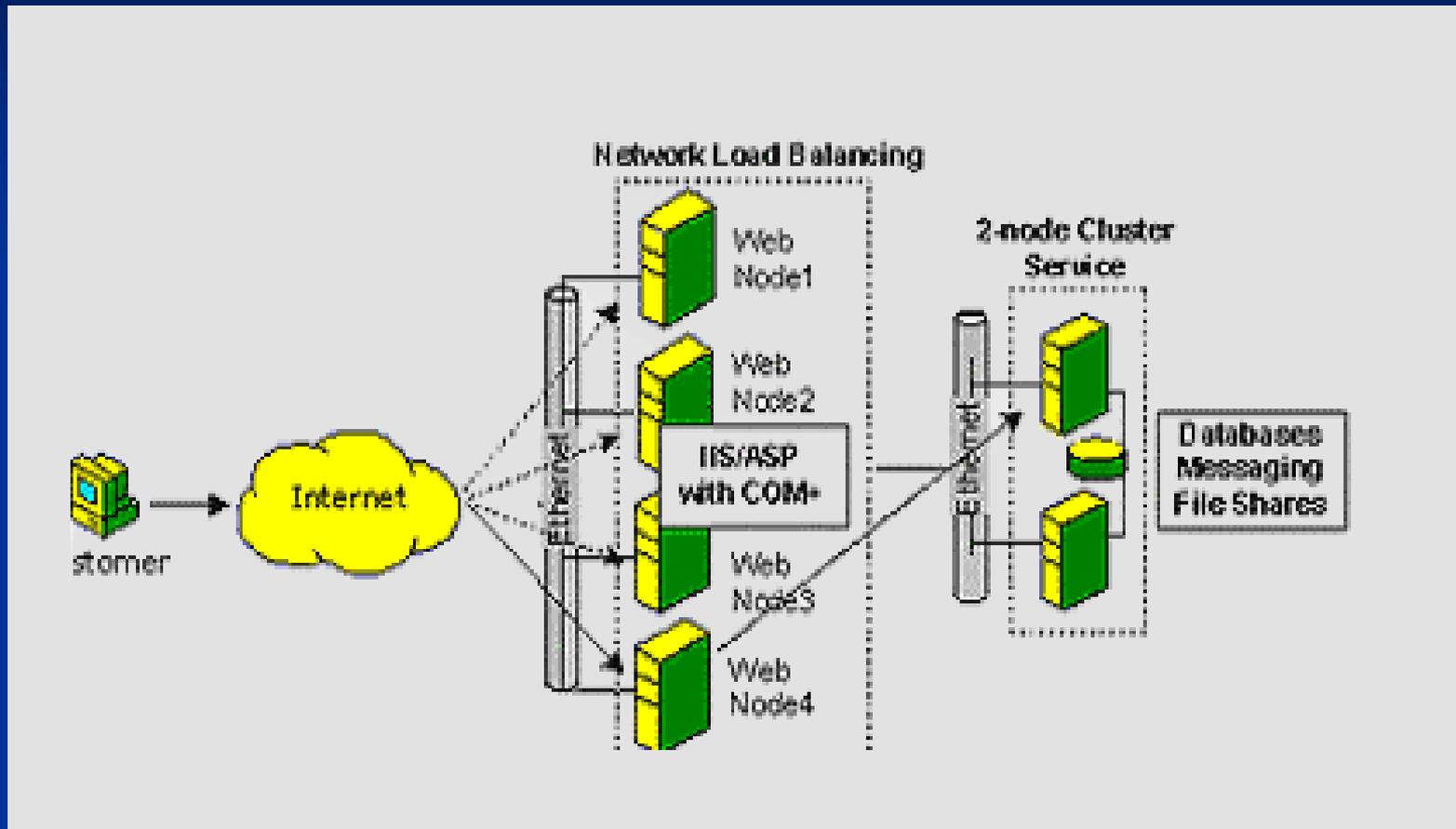
采用WIN2K ADVANCED SERVER操作系统，再配置集群服务，可实现应用系统的高可用性解决方案。

负载均衡集群系统实现



采用WIN2K ADVANCE SERVER操作系统，再加上WIN2K的负载均衡技术，不需要其它的硬件流量分配设备和流量管理软件，即可实现多个节点的流量管理、负载均衡。

负载均衡加高可用性解决方案



采用WIN2K ADVANCED SERVER 操作系统再加上负载均衡技术和高可用技术，可以在三层结构计算环境中实现对中间件的负载均衡和后台数据库的高可用性。

高性能计算集群相关介绍

- 目前，HPC的应用范围越来越广，除了一些特殊的应用需要高性能计算外，一些如：数据挖掘应用、图像处理业务、基因测序比对处理等领域也越来越需要高性能计算。
- 科学计算领域
 - 更快的速度
 - 更高的精度
 - 更大的范围
- 信息服务领域
 - 更好地整合、提高可管理性
 - 更高的吞吐率
 - 更高的可靠性

面临的问题

- 但是，想要得到一个高性能解决方案，往往意味着要投入大量的金钱。
 - 国内：
 - 曙光4000A（10万亿次运算能力），市场价格约为1亿人民币。
 - 曙光5000A（200万亿次运算能力），市场价格约为2亿人民币。
 - 国外能达到相同计算规模的超级计算机，价格大约是国内的5~6倍。
 - 一个小规模的高性能解决方案，也需要花上百万乃至上千万的成本投入。

低成本高性能计算的产生

- 一方面是有限的成本投入，一方面是对高性能计算的迫切需要。怎样在这两者之间取得一个平衡点？
- 因此，构建一个低成本、高性能的解决方案由此产生。

多种结构的比较

结构 \ 特性	SMP	DSM	MPP	Cluster
可扩展性	最差	差	强	很强
易管理性	很强	很强	一般	较差
网络能力	无需网络	无需网络	很强	强
并行编程方式	共享变量	共享变量、 消息传递	消息传递	消息传递
价格	贵	昂贵	昂贵	较低

环境搭建

- 硬件环境搭建

- 软件环境搭建

低成本考虑:

- 1) 从硬件上降低成本
- 2) 从软件上降低成本
- 3) 从日常维护费用上降低成本

高性能考虑:

- 1) 使用高频、多核CPU
- 2) 使用优化的数学库
- 3) 优化网络环境

硬件环境搭建

■ 硬件选购:

1) 计算节点:

对大量集群环境配置的经验性研究显示小型SMP平台最适合构建HPC集群，并且现在的多核技术已经达到平民化的程度，因此使用多路处理器的节点成为首选。

选用AMD处理器

- 在很多领域中，需要对大量的数据进行处理。这种模式对每个节点上的性能及I/O吞吐量提出了很高要求。AMD皓龙处理器的直连架构和超传输总线（HyperTransport）技术，保证了低内存延时和缓存的一致性，避免了系统扩展带来的性能瓶颈，可实现处理器核心之间的高速数据交换，提升节点性能，满足了应用的需求。以及其相对低廉的价格，是搭建一个低成本解决方案的重要因素。（更适合多核单节点的情况）

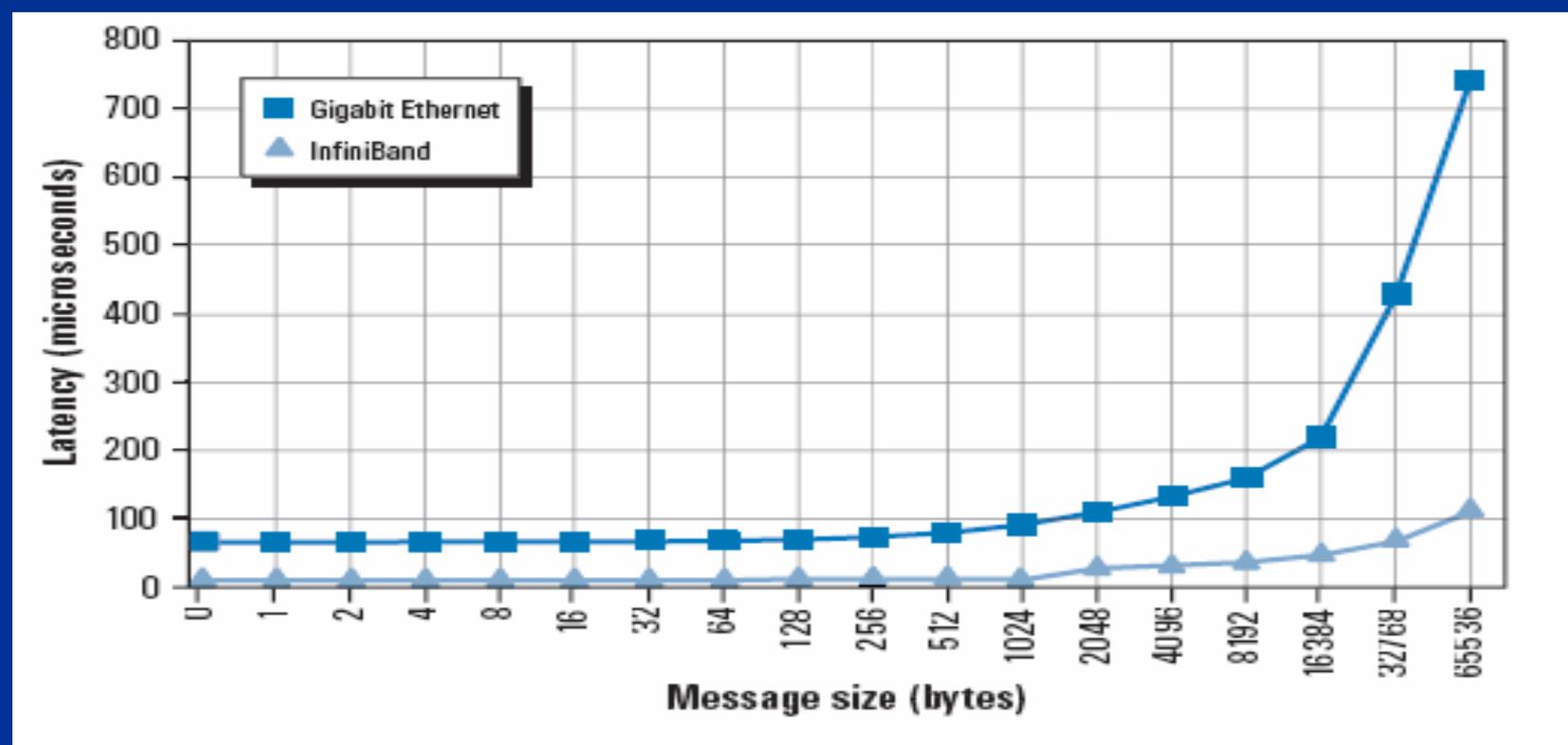
TOP 20

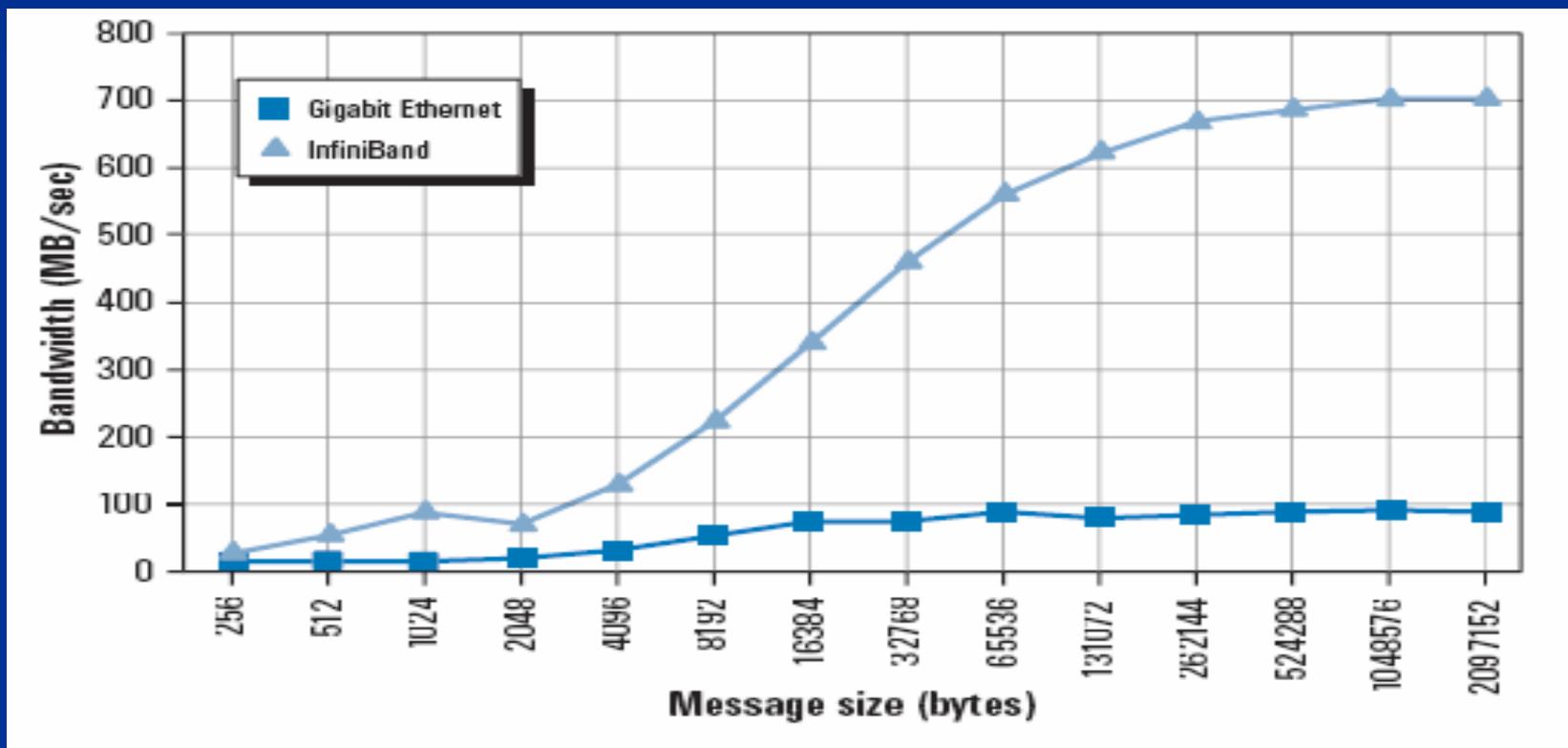
Rank	Site	Manufacturer	Computer	Year	Processors	RMax	RPeak	Power	Efficiency
1	DOE/NNSA/LANL	IBM	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz , Voltaire Infiniband	2008	122400	1026000	1375780	2345.5	74.58%
2	DOE/NNSA/LLNL	IBM	eServer Blue Gene Solution	2007	212992	478200	596378	2329.6	80.18%
3	Argonne National Laboratory	IBM	Blue Gene/P Solution	2007	163840	450300	557056	1260	80.84%
4	Texas Advanced Computing Center/Univ. of Texas	Sun Microsystems	SunBlade x6420, Opteron Quad 2Ghz, Infiniband	2008	62976	326000	503808	2000	64.71%
5	Oak Ridge National Laboratory	Cray Inc.	Cray XT4 QuadCore 2.1 GHz	2008	30976	205000	260200	1580.71	78.79%
6	Forschungszentrum Juelich (FZJ)	IBM	Blue Gene/P Solution	2007	65536	180000	222822	504	80.78%
7	New Mexico Computing Applications Center (NMCAC)	SGI	SGI Altix ICE 8200, Xeon quad core 3.0 GHz	2007	14336	133200	172032	861.63	77.48%
8	Computational Research Laboratories, TATA SONS	Hewlett-Packard	Cluster Platform 3000 BL460c, Xeon 53xx 3GHz, Infiniband	2008	14384	132800	172608	1602.24	76.94%
9	IDRIS	IBM	Blue Gene/P Solution	2008	40960	112500	139264	315	80.78%
10	Total Exploration Production	SGI	SGI Altix ICE 8200EX, Xeon quad core 3.0 GHz	2008	10240	106100	122880	442	86.34%
11	Government Agency	Hewlett-Packard	Cluster Platform 3000 BL460c, Xeon 53xx 2.66GHz, Infiniband	2007	13728	102800	146430	0	70.20%
12	NNSA/Sandia National Laboratories	Cray Inc.	Sandia/ Cray Red Storm, Opteron 2.4 GHz dual core	2007	26569	102200	127531	0	80.14%
13	EDF R&D	IBM	Blue Gene/P Solution	2008	32768	92960	111411	252	83.44%
14	IBM Thomas J. Watson Research Center	IBM	eServer Blue Gene Solution	2005	40960	91290	114688	448	79.60%
15	NERSC/LBNL	Cray Inc.	Cray XT4, 2.6 GHz	2007	19320	85368	100464	1150	84.97%
16	Information Technology Center, The University of Tokyo	Hitachi	Hitachi Cluster Opteron QC 2.3 GHz, Myrinet 10G	2008	12288	82984	113050	638.6	73.40%
17	Stony Brook/BNL, New York Center for Computational Sciences	IBM	eServer Blue Gene Solution	2007	36864	82161	103219	403.2	79.60%
18	ECMWF	IBM	Power 575, p6 4.7 GHz, Infiniband	2008	8320	80320	156416	1329.7	51.35%
19	RZG/Max-Planck-Gesellschaft MPI/IPP	IBM	Power 575, p6 4.7 GHz, Infiniband	2008	6720	80320	126336	1073.99	63.58%
20	Center for Computational Sciences, University of Tsukuba	Appro International	Appro Xtreme-X3 Server - Quad Opteron Quad Core 2.3 GHz, Infiniband	2008	10000	76460	92000	671.8	83.11%

2) 网络

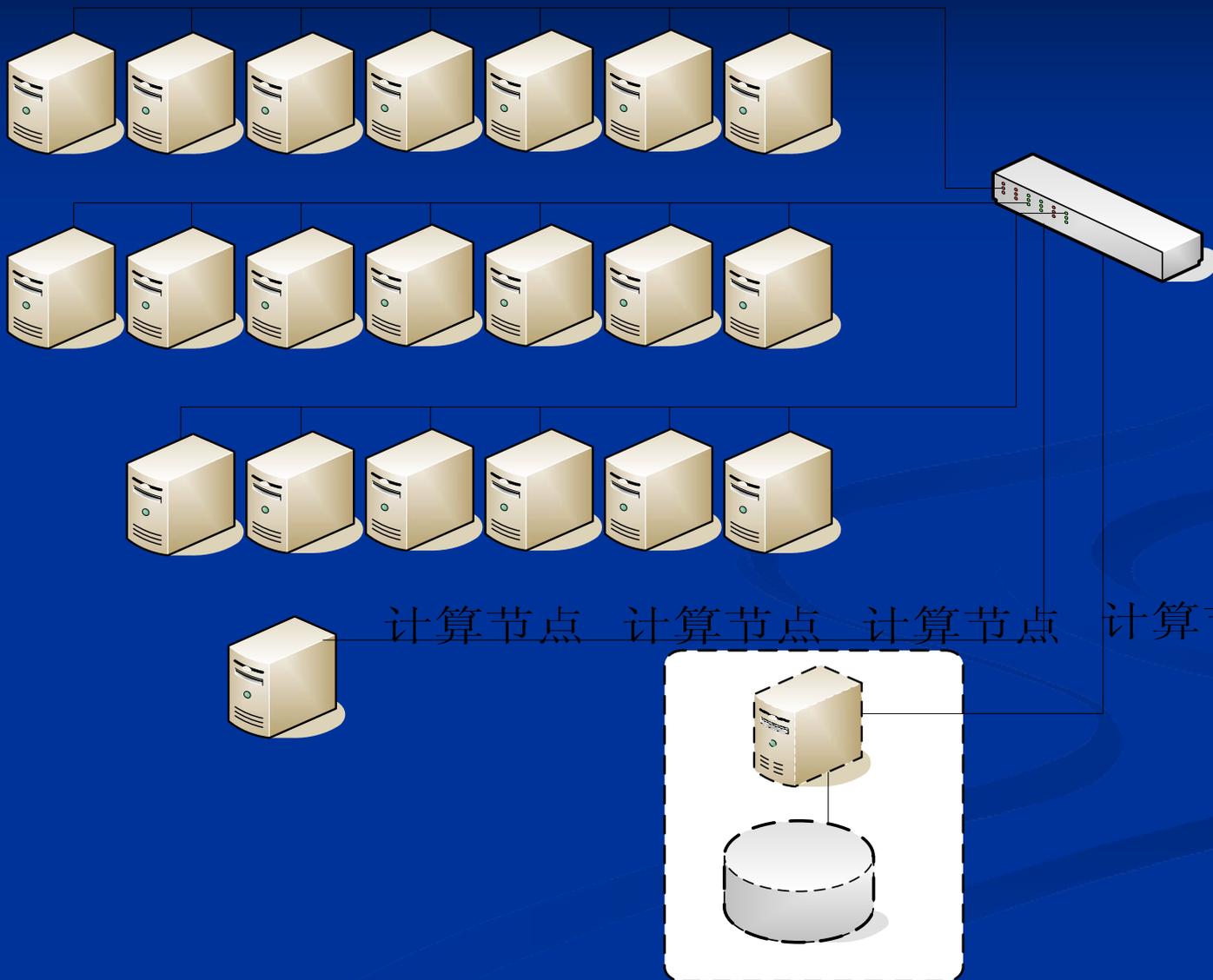
一般在大中型集群中，往往会使用 InfiniBand, Myrinet 等高速网络。但是他们的价格都不菲，针对于小型的集群搭建，使用千兆网络已经能够满足小型集群的网络通讯的需要。

网络比较





常见的HPC网络拓扑结构



计算节点 计算节点 计算节点 计算节点 计算节点

计算节点 计算节点 计算节点 计算节点 计算节点

软件环境基本搭建 (开源软件)

- 📎 操作系统: RHEL4 U5
- 📎 并行环境: MPI 1.2.7
- 📎 作业管理: OpenPBS 2.3.16
- 📎 监控软件: Ganglia 3.0.6
- 📎 数学库: GotoBLAS、Atlas、ACML

考虑因素

- 考虑到成本与实际环境的因素，可以使用开源软件对集群进行管理和相关的应用。
- 因为开源软件自身的局限性，不可能提供很好的集群管理机制。因此在较大规模集群方案中，如果还是使用开源的软件去对集群进行管理，日常的维护工作量就会明显增多。（建议使用专门的集群管理软件：CSM (Linux)、SUN Cluster 3.0、Dawning DCMM系统监控DCMS集群管理)

软件

- 系统应配置有标准的C、C++、Fortran和Java等主流编译器，将来还要在这基础上安装其它的并行编程环境，如PVM、HPF、OpenMP、串行程序自动并行化工具、并行调试器、并行程序监控工具和并行数学库函数。还可以安装一些针对与某些领域的应用软件，如Accelrys Material Studio材料计算软件等。

开发模式

- 支持串行程序、消息传递程序、SMP多线程程序、HPF数据并行程序、分布式共享虚存等多种编程模型。
- 根据所提供的编程环境，能够很快的进行并行程序的开发。

适用于VO的函数库以及软件

- 数学库: MKL、ACML、GotoBLAS、Atlas、BLAS、FFTW、IMSL、Lapack 等
- 编程环境: MPICH (MPI) /PVM, C、C++、Fortran77、Fortran90、Fortran99等程序的编译器
- 可用软件: MM5大规模天气数值预报模拟软件等
- 其它: Matlab在多核及集群计算机上的并行计算工具箱3.3: Parallel Computing Toolbox 3.3

性能测试

- 测试软件: Linpack (hpl 1.0a)
- 测试数学库: GotoBLAS 1.26、Atlas 3.6.0、MKL 10、ACML 4.0。

工作效率的提高

- 使用一台普通的计算机要进行40个小时的运算，理论上在20个节点集群环境内可以缩短至1个小时。

测试注意问题

- 如果是使用核心是Barcelona以上的AMD 四核处理器，在使用GotoBLAS数学库时，必须使用最新的1.26版本。
- 使用GotoBLAS数学库测试，在单节点上配置运行测试参数时，只需要配置如下：
mpirun -np 1 xhpl。则会同时调用8个核心进行计算。而使用其它数学库，要让8核同时进行计算，则需要明确配置8进程，如：
mpirun -np 8 xhpl 。

优点

1. 建设成本低廉，运算速度能达到中型计算的规模。
2. 建设周期短。
3. 结构易于变化，容易适应需求的改动。因此很容易通过扩展，升级，获得更好的性能。
4. 可通过Web进行监控，通过web进行远程管理等工作。
5. 维护工作量少，运行成本低。

建议

1. 经过以上的对比，怎样去选择一个合理的方案去搭建一个低价格高性能的集群需要多方面的考虑。
2. 如果确实需要达到比较大的运算速度，采用更好的网络环境，更多的运算节点和更快的cpu是最直接的方式。

3. 单台内集成更多的运算核心会得到更好的效果，但是瓶颈在于CPU访问内存的带宽并不能随着CPU个数的增加而有效增长，所以怎样在价格和性能上做一个很好的平衡，将是必须考虑的问题。
4. 针对于一些高吞吐量的运算来说，I/O瓶颈的也是一个必须面对的问题。

结束语

集群技术是一种通用的技术，其目的是为了了解决单机运算能力的不足、IO能力的不足、提高服务的可靠性、获得规模可扩展能力，降低系统整体的运行、升级、维护成本。只要在其他技术不能达到以上的目的的情况下，或者虽然能够达到以上的目的，但是成本过高的情况下，就可以考虑采用集群技术。

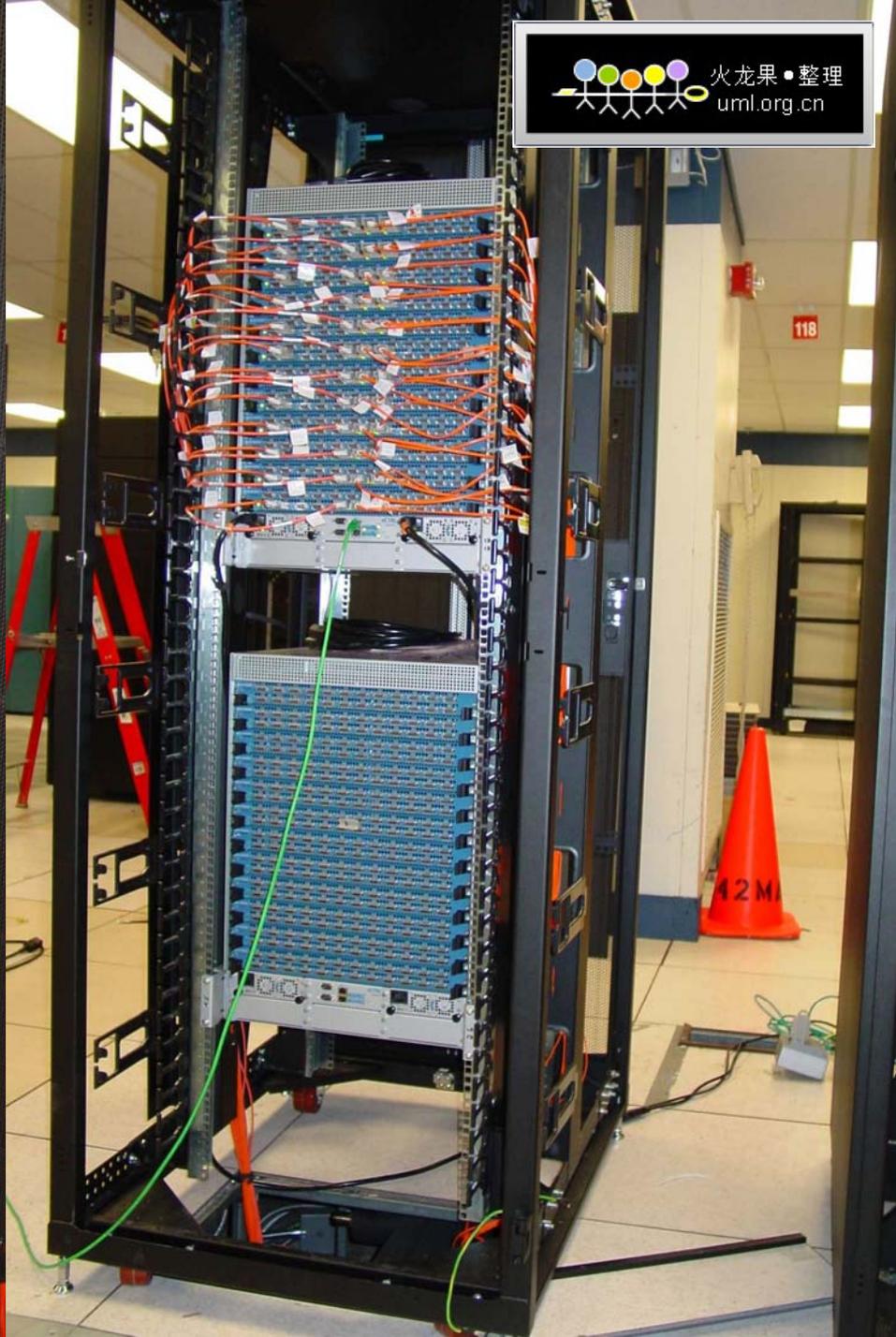
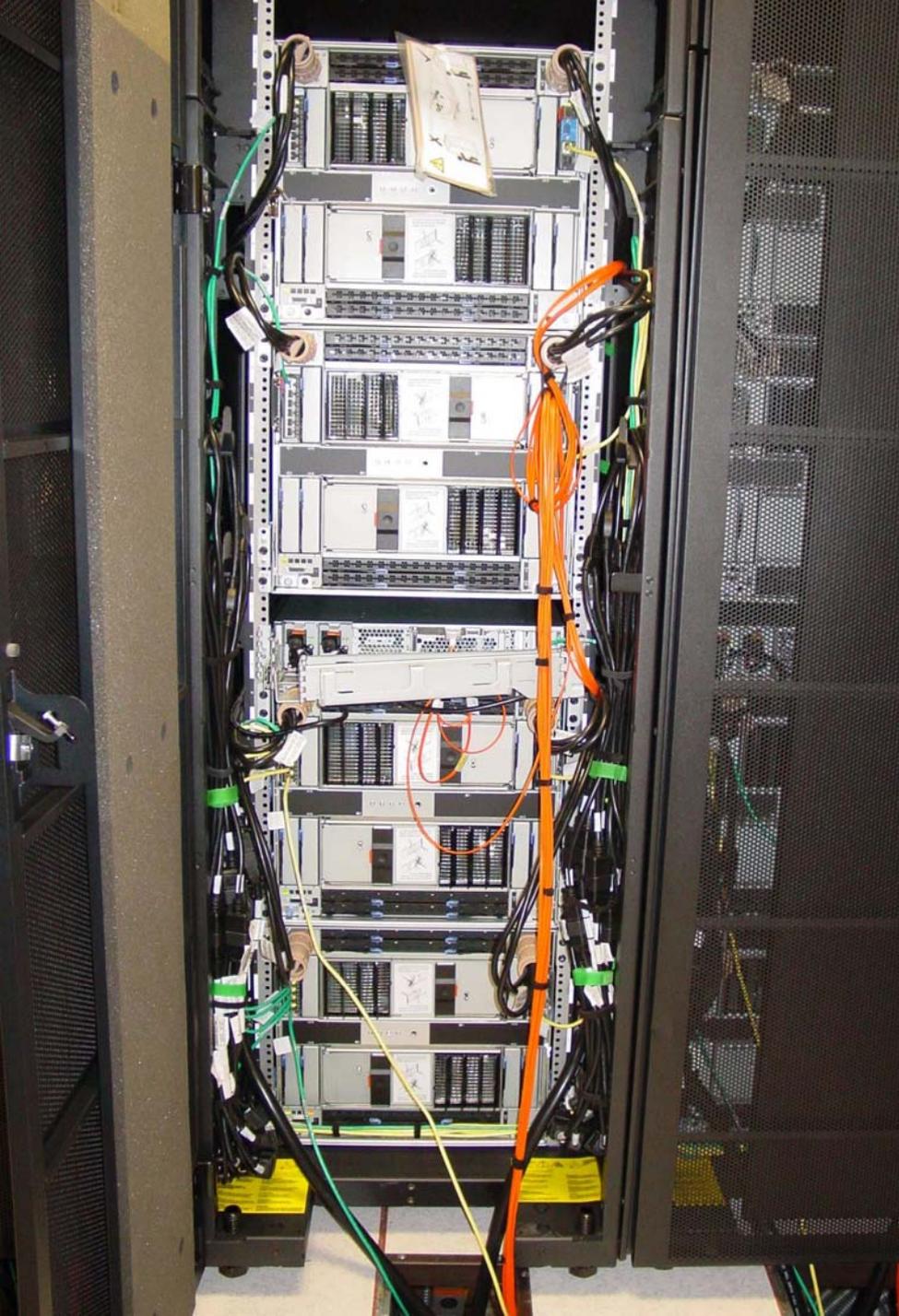
备注



Roadrunner 10^{15}

First to sustained Petaflop/sec





118

42M

Roadrunner Installation 09/06 to 12/06

