

集群系统基础



浪潮（北京）电子信息产业有限公司

inspur 浪潮

主要内容

集群系统简介

HPC集群系统简介

Linux 置配、装安统系群集

Intel 并行开发套件安装、配置

集群日常维护和故障判断

主要内容

集群系统简介

HPC集群系统简介

Linux 置配、装安统系群集

Intel 并行开发套件安装、配置

集群日常维护和故障判断

什么是集群？

集群是一组相互独立的、通过高速网络互联的计算机，它们构成了一个组，并以单一系统的模式加以管理。一个客户与集群相互作用时，集群像是一个独立的服务器。集群配置是用于提高可用性和可缩放性。

集群（cluster）技术是一种较新的技术,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，其任务调度则是集群系统中的核心技术。

构建集群的目的

- 1 提高性能
- 2 降低成本
- 3 提高可扩展性
- 4 增强可靠性

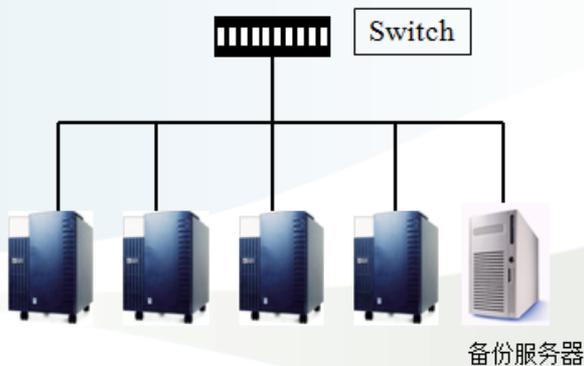
集群的分类

- 按照侧重点的不同，可以把集群分为三类
- 高可用集群
- 负载均衡集群
- 科学计算集群
(高性能计算High Performance Computing集群) 简称HPC集群

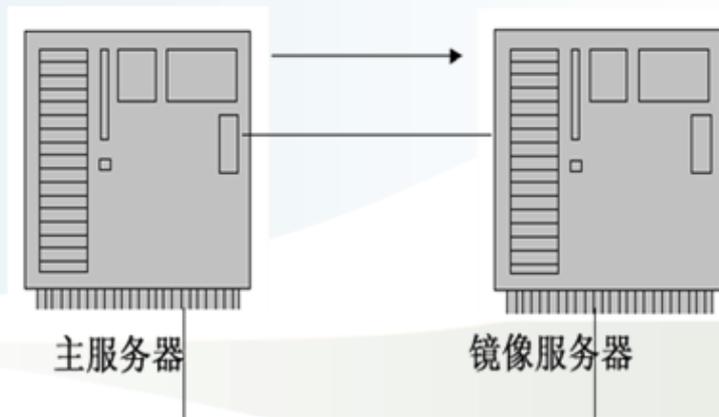
高可用集群

- 高可用集群
 - ▶ HA (High Available)
 - ▶ 主节点与从节点
 - ▶ 切换时间
- 实现方式
 - ▶ 镜像服务器双机
 - ▶ 双机与磁盘阵列柜
 - ▶ 多机HA集群

多机HA集群



镜像服务器双机



双机与磁盘阵列柜



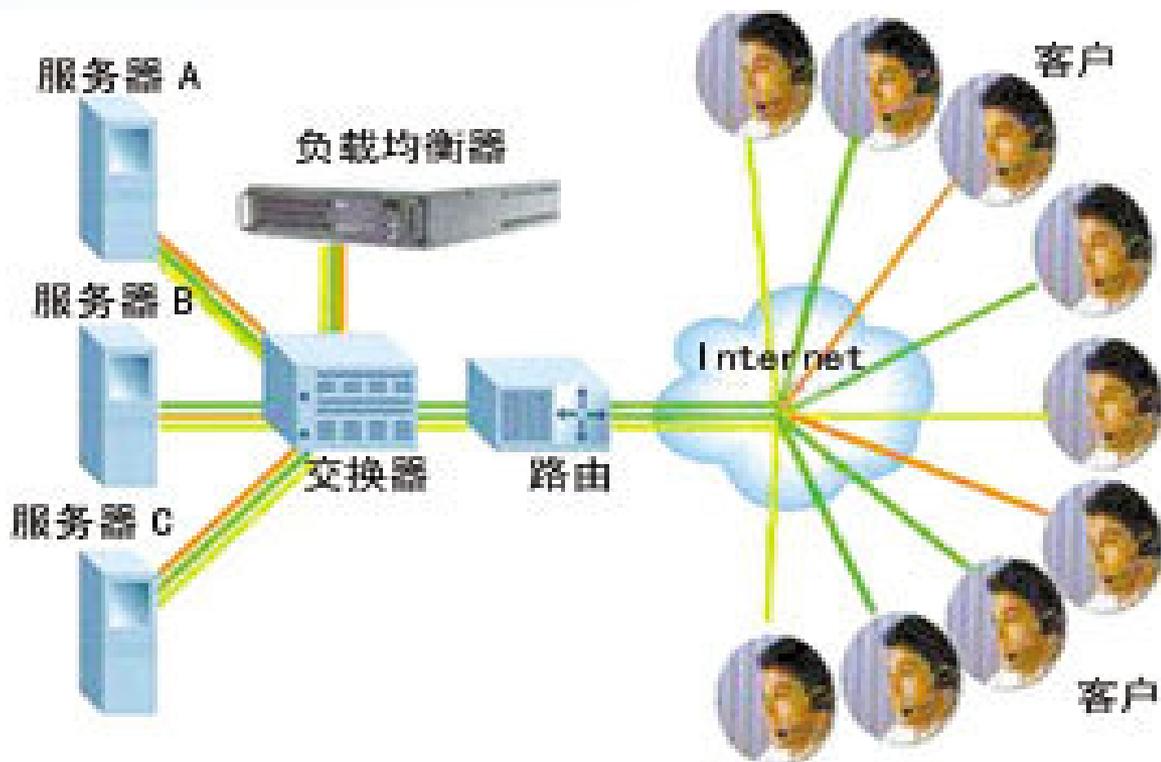
负载均衡集群

▶ 负载均衡的定义

- ▶ 使网络流量负载可以在计算机集群中尽可能平均地分摊处理

▶ 负载均衡的应用

- ▶ 海量WEB
- ▶ Email服务
- ▶ VOD视频点播
- ▶ 并行中间件等



科学计算集群（HPC集群）

- 科学计算集群的分类
 - ▶ 计算密集型
 - ▶ 通信密集型
 - ▶ 数据密集型
- 科学计算集群的应用领域
 - ▶ 有限元分析
 - ▶ 分子模拟
 - ▶ 图像处理
 - ▶ 流体仿真
 - ▶ 油藏模拟
 - ▶ 地震分析
 - ▶ 数值天气预报等

主要内容

集群系统简介

HPC集群系统简介

Linux 置配、装安统系群集

Intel 并行开发套件安装、配置

集群日常维护和故障判断

高性能计算

- **高性能计算 (High Performance Computing)**
 - 简单的说，高性能计算是计算机科学的一个分支，研究并行算法和开发相关软件，致力于开发高性能计算机（High Performance Computer）。高性能计算的衡量标准主要以计算速度（尤其是浮点运算速度）作为标准。
- **并行计算 (Parallel Computing)**
 - 并行计算（Parallel Computing）是指同时使用多种计算资源解决计算问题的过程，是提高计算机系统计算速度和处理能力的一种有效手段。

高性能计算的应用领域



OAK RIDGE NATIONAL LABORATORY

Government-Classified Work

Government - Research

(Severe) Weather Prediction & Climate Modeling



Instituto Nacional de Meteorología

España

Automotive Design & Safety



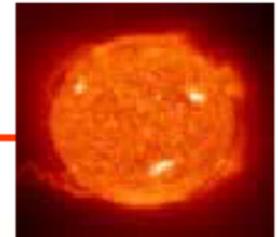
Drug Discovery & Genomic Research



Aircraft/Spacecraft Design & Fuel-Efficiency



Oil Exploration & Energy Research



Basic Scientific Research



高性能计算的分类

- 计算密集 (compute-intensive)

科学计算与数字模拟

- 数据密集 (data-intensive)

数字图书馆、数据仓库、图像处理

- 网络密集 (network-intensive)

网络应用

- 以上三种的混合

并行计算

- 并行计算，就是在并行机上所做的计算
 - 计算科学与传统的两种科学，即理论科学和实验科学，并立被认为是人类认识自然的三大支柱，他们彼此相辅相成地推动科学发展与社会进步。在许多情况下，或者是理论模型复杂甚至理论尚未建立，或者实验费用昂贵甚至无法进行时，计算就成了求解问题的唯一或主要的手段。
- 并行计算带来的优势
 - 降低单个问题求解的时间
 - 提高问题求解精度
 - 增加问题求解规模
 - 无法替代的模拟计算



并行计算机

- 并行计算机，就是由多个处理单元组成的计算机系统，这些处理单元相互通信和协作以快速、高效求解大型复杂问题。

处理单元有多少

处理单元的功能有多强

处理单元之间怎样连接

处理单元的数据如何传递

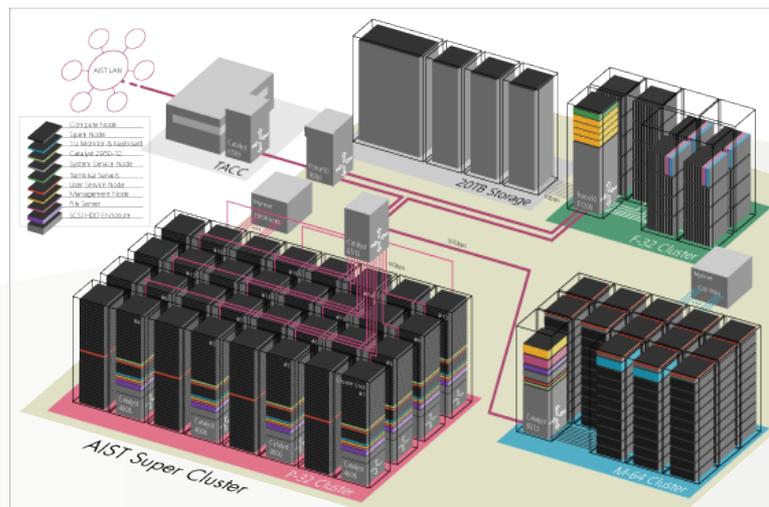
各处理单元如何相互协作

并行程序如何编写



并行计算机的发展

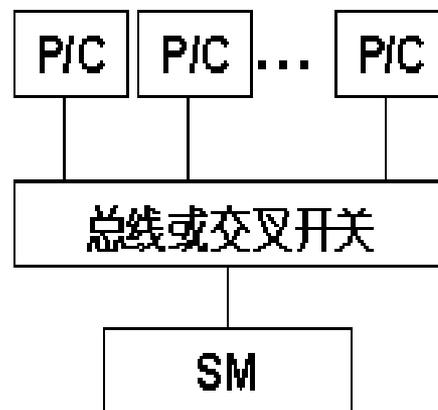
- (1)并行机的萌芽阶段（1964-1975）
- (2)向量机/ SMP的发展和鼎盛阶段(1976-1990)
- (3)MPP出现和蓬勃发展阶段（1990-1995）
- (4)Cluster出现，并成为并行计算机的主流（1995年后）



对称多处理机系统(SMP)

● SMP

- ▶ 对称式共享存储:任意处理器可直接访问任意内存地址,且访问延迟、带宽、机率都是等价的;系统是对称的;
- ▶ 微处理器:一般少于64个;
- ▶ 处理器不能太多,总线和交叉开关的一旦作成难于扩展;
- ▶ 例子: IBM R50, SGI Power Challenge, SUN Enterprise,

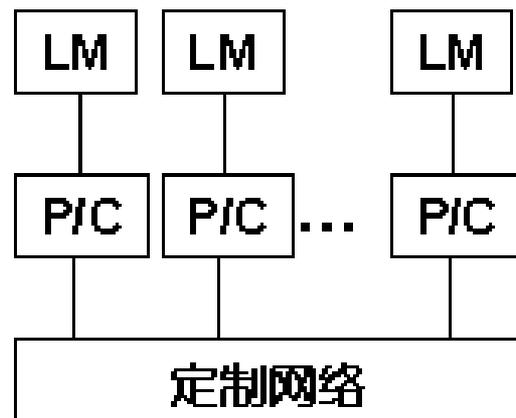


(b) SMP, 物理上单一地址空间

大规模并行计算机系统(MPP)

● MPP

- ▶ 物理和逻辑上均是分布内存
- ▶ 能扩展至成百上千个处理器(微处理器或向量处理器)
- ▶ 采用高通信带宽和低延迟的互连网络(专门设计和定制的)
- ▶ 一种异步的MIMD机器; 程序系由多个进程组成, 每个都有其私有地址空间, 进程间采用传递消息相互作用;
- ▶ 代表: CRAY T3E(2048), ASCI Red(3072), IBM SP2



(c) MPP, 物理/逻辑上多地址空间

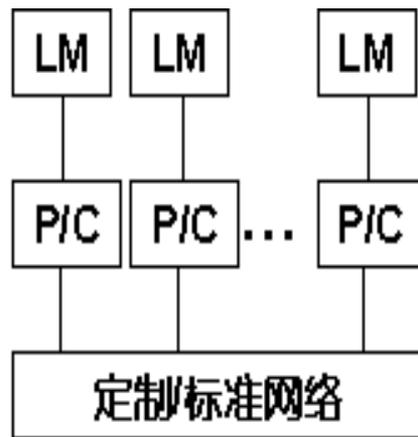
集群系统(Cluster)

● Cluster

- ▶ 每个节点都是一个完整的计算机
- ▶ 各个节点通过高性能网络相互连接
- ▶ 网络接口和I/O总线松耦合连接
- ▶ 每个节点有完整的操作系统

Cluster由于无可比拟的性价比优势占据主流位置。

更高的计算能力
系统的高可用性
良好的可扩展性
更高的性价比



(e) Cluster/COW, 物理/逻辑上多地址空间

衡量系统性能的主要指标

MFLOPS（百万次浮点运算每秒）

- (Million Floating-point Operations per Second, 每秒百万个浮点操作)衡量计算机系统的主要技术指标之一。
- 理论浮点峰值是该计算机理论上能达到的每秒钟能完成浮点计算最大次数，它主要是由CPU的主频决定的，
- 理论浮点峰值=CPU主频×CPU每个时钟周期执行浮点运算的次数×系统中CPU核心数目
- 比如一台配置2颗E5-2670的服务器的理论浮点运算值=2.6（GHz）*8*16=332.8GFLOPS 即3328亿次浮点运算每秒
- 1P=1000GFLOPS 即万亿次
- 1G=1000MFLOPS 即10亿次
-

Linux HPC集群系统体系结构

