

松散耦合分布式系统

鞠九滨 著



吉林大学出版社

松散耦合分布式系统

鞠九滨 著

吉林大学出版社

内 容 提 要

本书介绍用计算机网络特别是高速局域网络组成的分布式计算机系统的结构和实现,侧重于基本概念、基本原理和基本方法。全书共有七章。在介绍了典型的网络环境之后,详细介绍分布式操作系统,包括进程通信、命名与保护、控制、文件系统以及服务等。各主要章节中包括很多实例。最后一章集中介绍了国际上具有代表性的四个典型的分布式系统。

本书比较全面和系统地反映了国际上十多年来在这一领域取得的主要成果,特别是最近几年取得的最新成果。在有关章节中还介绍了作者及其合作者们的研究成果。

本书可作为研究生和大学高年级学生的教材,也可供有关科技人员参考。

松散耦合分布式系统

鞠 九 滨 著

责任编辑:张延

封面设计:甘利

吉林大学出版社出版
(长春市解放大路85号)

吉林省新华书店发行
吉林工学院印刷厂印刷

开本:787×1092毫米 1/16

1989年12月第1版

印张:16

1989年12月第1次印刷

字数:363千字

印数:1—600册

ISBN 7-5601-0419-3/TP·4

定价:3.40元

前 言

作者从1983年开始为研究生讲授分布式系统课程，每次的内容均有较大变化。本书是在这个基础上重新组织材料写成的。在一年半的重新编写过程中，由于国际上这一领域不断取得新的进展，一些著名分布式系统设计不断改进，因而就必须反复地修改原稿以跟上这一领域的前进步伐。这些说明分布式系统的研究正处于非常活跃的阶段。尽管其理论体系仍处于发展时期，但很多基础部分已趋于成熟。本书介绍的内容主要是一些基本概念、基本原理和基本方法，力求反映最新的研究结果，使用的参考资料也尽量选取有权威性的原著，以确保本书论述的内容的正确性、准确性和时代性。

按照国际学术界大多数人的观点，分布式计算系统可以分成紧密耦合的和松散耦合的两种，并且在现有的松散耦合式分布系统中，绝大多数是用计算机网络（主要是局部网络）支持的。所以，本书介绍这种系统的结构和实现，并且侧重于分布式操作系统，主要包括进程通信、命名和保护、控制机制、文件系统、提供的服务等问题，在说明基本原理和方法时所举的例子都是当前国际上已实现的有代表性的著名的系统。最后一章集中介绍了四个典型的分布式系统，以便使读者对设计分布式系统的几个方法和分布式系统的整体结构有个清楚的了解。由于篇幅所限，本书不讨论分布式数据库、分布式程序设计和分布式系统的应用等问题。

本书在很多章节中还介绍了JDCS分布式计算机系统的有关部分。这个系统是作者及其合作者们从1983年开始研制的，研究工作现仍在继续。作者的同事们和一大批研究生对这个项目作出贡献，主要人员有杨宏戟、李春阳、田颖声、金尚杰、李双丰、赵牧虹、张斯竞、周春光、张海帆、孙芙、王大明、陶杰、赵亮、房至一等。作者应该感谢国家科委和国家教委把该项研究列为“六五”科技攻关项目给予了资助，国家自然科学基金委员会在“七五”期间给予了资助。作者还应该感谢英国Durham大学计算机科学系主任K.H.Bennett教授1985年9月应邀来吉林大学指导这一研究，为研究生讲授了分布式系统课程，并与作者广泛讨论了分布计算领域的研究和教学的各个方面，对于本书的形成十分有益。此外，英国剑桥大学计算机实验室主任R.M.Needham教授、特肯大学计算实验室主任E.B.Spratt教授也都提供了有价值的资料，丰富了本书的内容。

欢迎读者指正错误和提出宝贵的意见。

鞠九滨

于吉林大学计算机科学系

1989年8月

目 录

第一章 绪论	(1)
1.1 分布式系统的概念	(1)
1.1.1 什么是分布式系统	(1)
1.1.2 松散耦合分布式系统	(1)
1.1.3 异构型分布式系统	(3)
1.1.4 分布式系统的优点	(5)
1.1.5 分布式系统的新问题	(5)
1.2 分布式系统与计算机网络	(6)
1.2.1 文件系统	(7)
1.2.2 保护	(7)
1.2.3 执行地点	(8)
1.3 分布式系统的透明性	(9)
1.3.1 透明性概念	(9)
1.3.2 透明性与资源的最佳控制	(10)
1.3.3 透明性与异构性	(11)
1.3.4 透明性与局部自治性	(12)
1.3.5 透明性与网络互连	(12)
1.4 分布式系统结构	(13)
1.4.1 基本概念	(13)
1.4.2 分布式系统的组成	(15)
1.4.3 分布式操作系统	(17)
1.4.4 ISO参考模型	(18)
1.4.5 分布式系统的设计问题	(19)
第二章 进程通信	(21)
2.1 硬件互连和局部网络	(21)
2.1.1 传输介质	(21)
2.1.2 互连拓扑	(22)
2.1.3 介质访问控制	(23)
2.1.4 基本通信协议	(24)
2.1.5 计算机和网络的接口	(25)
2.2 以太网	(25)
2.2.1 结构	(25)
2.2.2 工作原理	(27)
2.2.3 物理层	(28)
2.2.4 数据链路层	(29)
2.3 剑桥环	(31)
2.3.1 原理与组成	(31)

2.3.2	物理层.....	(33)
2.3.3	数据链路层.....	(34)
2.4	网络层.....	(37)
2.4.1	数据报和虚拟电路.....	(38)
2.4.2	路由算法.....	(40)
2.4.3	拥挤控制.....	(42)
2.4.4	网络互连.....	(43)
2.4.5	网络前端机.....	(45)
2.5	传送协议.....	(46)
2.5.1	传送协议的功能.....	(46)
2.5.2	剑桥环的传送协议.....	(48)
2.6	XNS的网际网传送协议.....	(52)
2.6.1	网际网数据报协议.....	(53)
2.6.2	差错报告包.....	(56)
2.6.3	路由选择.....	(56)
2.6.4	回响包.....	(57)
2.6.5	虚拟电路服务.....	(57)
2.6.6	包交换协议.....	(58)
2.7	进程通信原语.....	(59)
2.7.1	信件传递.....	(59)
2.7.2	远程过程调用.....	(61)
2.7.3	IPC原语的实现问题.....	(63)
2.8	JDCS分布系统中的进程通信.....	(64)
2.8.1	小包驱动程序.....	(64)
2.8.2	基本块协议.....	(67)
2.8.3	远程过程调用.....	(68)
2.8.4	剑桥环与以太网的通信.....	(69)
2.8.5	虚拟电路协议.....	(70)
2.9	V核中的进程通信.....	(70)
2.9.1	IPC原语.....	(70)
2.9.2	组通信.....	(71)
2.9.3	核间通信协议.....	(72)
2.9.4	性能.....	(73)
第三章	命名与保护.....	(76)
3.1	分布系统中的命名.....	(76)
3.1.1	标识符.....	(76)
3.1.2	分布系统中的标识符.....	(77)
3.1.3	绝对名字和相对名字.....	(78)
3.1.4	面向机器的标识符.....	(79)
3.1.5	面向用户的名字.....	(81)
3.1.6	名字服务员.....	(82)

3.2	命名的例子	(82)
3.2.1	JDCS分布系统中的名字服务员	(83)
3.2.2	V系统中的命名	(84)
3.3	分布系统的保护与安全	(87)
3.3.1	计算机病毒	(87)
3.3.2	计算机犯罪	(88)
3.3.3	分布系统的保护与安全技术	(89)
3.4	数据加密	(90)
3.4.1	单密钥加密	(90)
3.4.2	网络各层的加密	(94)
3.4.3	密钥的分配	(95)
3.4.4	公开密钥	(95)
3.4.5	数字签名	(96)
3.5	访问控制	(98)
3.5.1	两种访问控制方法	(98)
3.5.2	分布系统中的权能	(99)
3.5.3	分布系统中访问控制的保护	(101)
3.6	保护的例子	(102)
3.6.1	剑桥分布计算系统中的保护与授权	(102)
3.6.2	AMOEBa中的保护	(105)
第四章	控制	(107)
4.1	分布式系统的资源管理	(107)
4.1.1	分散控制	(107)
4.1.2	控制空间	(108)
4.1.3	分散控制与通信	(110)
4.1.4	资源分配原则	(110)
4.2	进程管理	(111)
4.2.1	处理机的分配	(111)
4.2.2	调度	(112)
4.2.3	负载平衡	(113)
4.2.4	死锁	(114)
4.3	同步与互斥算法	(115)
4.3.1	同步机构的评价标准	(115)
4.3.2	分布式系统的同步机构	(116)
4.3.3	互斥算法	(118)
4.4	并发控制	(121)
4.4.1	处理事务处理的模型	(123)
4.4.2	可串行化调度	(124)
4.4.3	封锁法	(126)
4.4.4	时间戳的使用	(129)

4.5	容错	(129)
4.5.1	冗余技术	(130)
4.5.2	原子事务处理	(131)
4.5.3	坚固存储器	(132)
4.5.4	两阶段提交协议	(133)
4.6	多副本更新	(137)
4.6.1	系统数据库	(137)
4.6.2	多副本更新机构的基本结构	(138)
4.6.3	同步表决	(139)
4.6.4	多数一致方法	(140)
4.6.5	循环令牌法	(142)
第五章	文件系统	(143)
5.1	文件服务员	(143)
5.1.1	服务员的构成	(143)
5.1.2	文件服务员的一些基本概念	(144)
5.1.3	文件服务员设计问题	(147)
5.2	文件的使用与管理	(148)
5.2.1	文件的操作	(148)
5.2.2	文件模型	(150)
5.2.3	原子事务处理的支持	(151)
5.2.4	文件管理	(152)
5.3	原子事务处理的实现	(153)
5.3.1	并发控制	(154)
5.3.2	可恢复文件的实现	(156)
5.3.3	事务处理的提交与完成	(157)
5.3.4	信息的可靠存贮	(159)
5.4	顾客/服务员通信	(160)
5.4.1	传送服务	(160)
5.4.2	远程调用	(160)
5.4.3	重复的请求	(161)
5.4.4	数据传送	(162)
5.4.5	服务员对无回答顾客的处理	(162)
5.5	文件系统结构	(163)
5.5.1	SUN微系统公司的网络文件系统	(163)
5.5.2	VICE文件系统	(165)
5.5.3	LOCUS分布式文件系统	(167)
5.6	JDCS分布系统的文件服务	(170)
5.6.1	第一文件服务站	(171)
5.6.2	第二文件服务站	(172)
5.6.3	文件服务站的使用	(173)
5.6.4	工作站间的文件传送	(174)

第六章 服务	(176)
6.1 处理机库	(176)
6.1.1 具有处理机库的分布式系统	(176)
6.1.2 处理机库的接口与控制	(177)
6.1.3 处理机库的管理	(179)
6.1.4 对话管理员	(181)
6.2 空闲工作站的共享	(182)
6.2.1 全局调度结构	(182)
6.2.2 工作环境	(183)
6.2.3 集中式调度	(184)
6.2.4 分散式调度	(186)
6.2.5 混合式调度	(187)
6.3 进程迁移与远程执行	(188)
6.3.1 进程迁移的一般要求	(188)
6.3.2 Sprite环境	(189)
6.3.3 进程迁移机构	(190)
6.3.4 透明的远程执行	(191)
6.3.5 性能和结论	(193)
6.4 长期调度	(194)
6.4.1 长期调度的目标	(194)
6.4.2 工作站的工作负载	(195)
6.4.3 上下算法	(196)
6.4.4 长期调度的实现	(198)
6.5 其他服务	(201)
6.5.1 主架机服务	(201)
6.5.2 简单服务	(203)
6.5.3 简单服务的实现	(205)
6.5.4 jDCS提供的服务	(206)
第七章 实例	(209)
7.1 LOCUS分布系统	(209)
7.1.1 程序的远程执行	(210)
7.1.2 动态重组	(213)
7.1.3 异构性	(215)
7.2 V分布系统	(216)
7.2.1 内核	(217)
7.2.2 输入/输出	(221)
7.2.3 命名	(222)
7.2.4 服务	(224)
7.2.5 应用	(224)
7.3 Clouds 分布系统	(226)
7.3.1 对象/进程模型	(226)

7.3.2 系统的实现.....	(229)
74. 异构型计算机系统 HCS.....	(231)
7.4.1 远程过程调用.....	(232)
7.4.2 命名.....	(234)
7.4.3 远程计算.....	(235)
7.4.4 文件系统.....	(237)
参考资料	(240)

第一章 绪 论

在这一章里我们将要回答下面的问题：什么是分布式计算机系统？它有哪些特点？它和计算机网络有什么区别？它由哪些部分组成？

1.1 分布式系统的概念

1.1.1 什么是分布式系统

分布式系统是分布式计算机系统、分布式计算系统、分布式（数据）处理系统的简称。关于它的定义，国内外很多学者做过研究，可概括如下：

分布式系统是由多个相互连接的计算机组成的一个整体，这些计算机在整个系统的控制下可合作执行一个共同任务，最少依赖于集中的过程、数据或硬件。

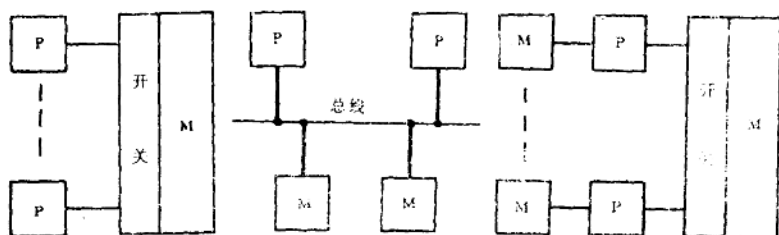
现在对这一定义进行说明：

- （1）系统是由多个处理器或计算机系统组成的；
- （2）这些计算机可以在物理上相邻的、使用机器内部总线或开关连接的处理器，通过共享主存进行通信；也可以是在地理上分开的，使用计算机网络（远程网或局部网）连接的计算机系统，使用信件进行通信；
- （3）这些计算机组成一个整体，对用户是透明的，即用户使用任何资源时不必知道这些资源在哪里；
- （4）一个程序可分散到各计算机上运行；
- （5）各计算机地位平等，除了受全系统的操作系统控制外，不存在主从控制和集中控制环节。

这种计算机系统属于多指令流多数据流（MIMD）结构。现存的分布式系统只具有一定程度的分布特征。

1.1.2 松散耦合分布式系统

从结构上分，有两大类分布式系统：紧密耦合系统和松散耦合系统。有共享主存贮器的分布式系统称作紧密耦合系统（如图1.1.1所示），通常也叫作多处理机系统。这种系统中的处理机也常常有自己的专用主存贮器。由于要求处理机和主存贮器之间有较高的通信速度，紧密耦合系统一般限于一个局部区域，各部分相距很近，可以说是在物理上分散的，但不是在地理上分散的。各处理机交换信息是通过共享主存进行的。本书不讨论这种系统。



P: 处理机, M: 主存储器

图1.1.1 loosely耦合分布式系统

在松散耦合系统中, 处理机没有共享主存储器, 其互连使用某种输入/输出机构, 如图1.1.2所示。这种系统中处理机虽然不共享主存储器, 但可共享外部设备。

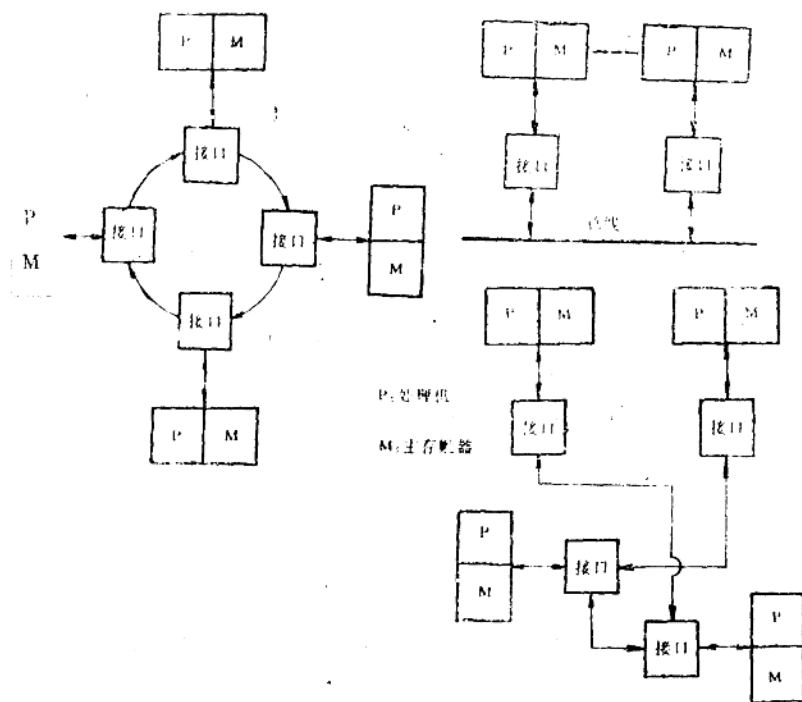


图1.1.2 松散耦合式分布系统

松散耦合系统由于处理机不共享主存储器, 只能用交换信件的方法进行通信。它可用远程网络(WAN)组成, 也可用局部网络(LAN)组成, 一般是在地理上分散的系统。目前绝大部分松散耦合分布系统都是由LAN组成的。本书以后各部分的分布系统除特别说明外均指松散耦合式。

根据组成松散耦合分布系统的计算机机种, 可以把松散耦合分布系统分为三种模式: 小型机模式、工作站模式和处理机库模式。

在小型机模式中, 系统由若干具有多用户的小型机(如VAX)组成, 每个用户在一个机器上登录, 可对其他机器进行远程访问。这种模式是集中式分时系统的简单扩充。

在工作站模式中，每个用户有自己的专用工作站，它通常配有功能很强的处理机、主存贮器和图形显示终端，有时还配有硬磁盘，几乎能完成所有的工作。它所以要加入分布式系统，是因为系统中有一个全局的文件系统可供各工作站访问。

处理机库模式是工作站模式的进一步发展。分时系统的中央处理机(CPU)数目对登录的用户数之比通常远小于1。这个比数，在工作站模式中近于1，而在处理机库模式中则远大于1。随着CPU越来越便宜，处理机库模式将越来越普遍。这种模式的主要思想是：当一个用户需要较多的计算时，将若干个CPU暂时分配给此用户，工作完成时再将这些CPU归还给处理机库，等待下一个请求。例如要重新编译10个子程序（分别在不同的文件上），这时可分配10个处理机并行完成这项工作。

还可将上述三种模式混合起来组成新的模式，例如一个系统中除了工作站之外还可有处理机库。

1.1.3 异构型分布式系统

分布式系统按照其组成的计算机种类可分成同构型分布式系统和异构型分布式系统。异构型分布式系统由若干不同的硬件或软件系统互连而成，而同构型分布式系统由相同或非常相似的硬件和软件组成。例如，由3个VAX、16个SUN和1个Symbolics LISP机组成的系统，由1个DEC-2060、1个IBM-4341和20个IBM PC-AT组成的系统，由12个Xerox D机（其中6个运行Interlisp而另外6个运行XDE软件组成的系统），均是异构型分布系统。只由SUN工作站（运行UNIX）组成的系统是一个同构型系统。

计算机的异构性在硬件和软件方面的主要表现如下：

1. 硬件异构性

首先是指令系统不同。这意味着一种机器上的程序模块不能在另一种不兼容的机器上执行。其次是数据表示方法不同。例如M68000和VAX类型的计算机都是按字节编址的，但是高位字节和低位字节的规定恰好相反。浮点表示方法也常不一样。第三是机器系统配置不一样，即使相同的CPU也可能有不同大小的主存和外部设备配置。

2. 软件异构性

即使硬件完全相同，如果运行的操作系统不同，则由这些机器组成的系统仍为异构型。软件不兼容首先表现在操作系统所提供的功能可能大不相同。其次是操作系统提供的系统调用在语法、语义和功能方面也不同。第三，文件系统不一样是经常遇到的现象。文件名的表示就是一个最简单的例子：Unix文件系统和DEC-VAX VMS文件系统都是分层结构，但Unix用符号/分隔路径名的各部分，而VMS文件系统中文件的表示是第一部分表示设备名，跟着一个：符号，最后的部分采取××××××.×××形式。这样，Unix程序可以产生和使用任何VMS文件名，反过来则不行。文件系统的差异还表现在其他方面。

应该指出，使用不同种类的机器通常是不可避免的。现代的科学技术使得具有不同性能的计算机型号层出不穷。一个单位对计算机性能的要求不断变化和提高，不可能总是满足于已获得的单一型号的计算机。这就使得异构型分布式系统不可避免。

为了使读者对异构型分布系统有个具体了解, 现以作者及其同事们在吉林 大学 从 1983年开始研制的JDCS松散耦合异构型分布式系统(见图1.1.3) 为例加以说明。该

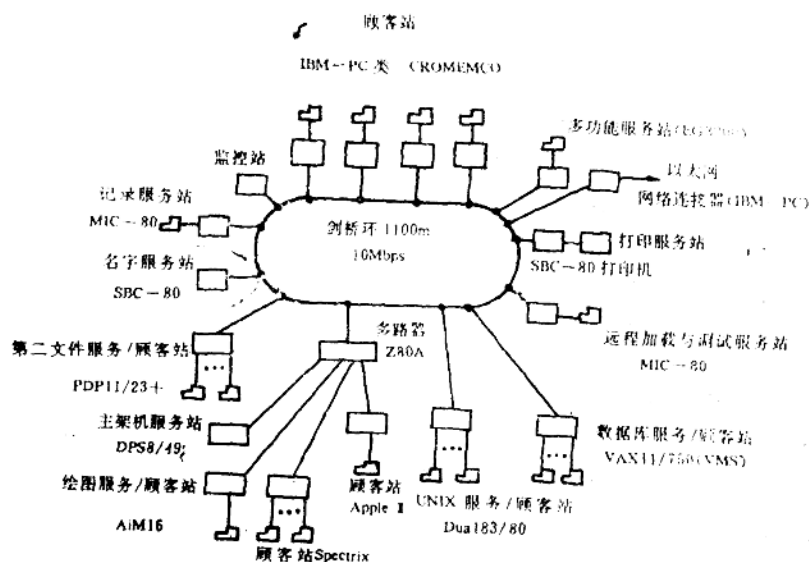


图1.1.3 JDCS分布式系统组成

系统用剑桥环局部网络将各种不同型号的机器连接起来, 随着机器的种类及数目的增加, 系统规模不断扩大。到写本书时, 该系统已连接了12种计算机, 各系统配置情况见表1.1.1, 这些计算机都不是兼容的。从处理机来看, 有 Z80(Z80A)、MC68000、8088、8086、PDP11、VAX11、6502以及DPS8等, 这些处理机具有不同的总线结构和

表1.1.1 JDCS分布式系统计算机配置

型号	处理机	主存容量	软盘容量	硬盘容量	操作系统
Dual 83/80	MC68000	1MB	1MB×1	80MB	UNIX(V7)
Spectrix	68000	1MB	—	170MB	UNIX(V7)
VAX11/750	VAX-11/750	4MB	500KB×2	465MB×2	VMS
PDP11/23+	LSI-11/23	256KB	1MB×1	30MB	RT-11, TSX-Plus
DPS8	DPS8	4MB	—	1200MB	GCOS8
IBM-PC	8088	640KB	360KB×2	10/15MB	PCDOS, CCDOS
Ai-M16	8086	512KB	1MB×1	10MB	Kudos
CROMEMCO	Z80A	64KB	170KB×2	—	CDOS
APPLEII	6502	64KB	160KB×2	—	—
EG3200	Z80A	64KB	340KB×2	13MB	CP/M2.2b
MIC-80	Z80	48KB	—	—	LEVEL 2BASIC
SBC-80	Z80A	6KB	—	—	—

输入输出信号。它们的键盘终端也不完全一样，在显示控制和字符等方面都有所不同。各计算机除了指令系统不一样外，操作系统也不完全相同，有UNIX、CP/M、Kudos、PCDOS、VMS、RT-11、GCOS8、CDOS等。使用的语言有C、Z80汇编、8088汇编、PDP11汇编等。C语言也有差别。文件系统也有很大差别，文件表示、结构、操作都不尽相同。机器的配置规模差别很大，最简单的SBC-80单板机、MIC-80机是无磁盘机微机，最大的是Honeywell DPS8中型主架机（配有1200MB硬盘），还有小型机VAX11/750、PDP11/23*（配有磁带驱动器）和超级微型机Dual 83/80及Spectrix。

1.1.4 分布式系统的优点

- (1) 增加性能增加处理部件的数目提高并行程度，可以得到要求的性能；
- (2) 可扩充性：随用户需求的增长，包括功能方面的要求和性能方面的要求，增加新节点数，不必替换整个系统；
- (3) 可靠性：由于控制、数据、软件和硬件的分散性（不存在集中环节），资源冗余以及结构上可动态重组，提高了可靠性；
- (4) 资源共享：系统中的软件、硬件资源如外部设备、文件系统和数据库等可为多个用户所共享；
- (5) 经济性：由于可扩充性，可以避免较大的初始投资，用多个微小型机代替一个大型主架机可以获得很好的性能价格比；
- (6) 适应性：对很多当用场合，如银行、铁路、商业等本来就分散而又必须相互协调的部门很适应。

由于分布式系统有这么多优点，所以发展很快。用户的需求极大地刺激了分布系统的发展，而现代半导体技术及计算机网络通信技术的巨大进展为分布式系统的发展提供了技术保证。

1.1.5 分布式系统的新问题

由于分布式系统包含多个（可能是不同种类的）分散的、自治的处理资源，要想把它们组织成一个整体最有效地完成一个共同的任务，比起通常集中式的单一计算机系统要困难得多，需要解决很多新的问题。

首先，由于处理资源的多重性，系统可能产生的差错类型和次数都比集中式系统多。最明显的是所谓**部分失效**问题：系统中某一个处理资源出故障而其他计算机尚不知道；单一计算机任何一部分出故障时将停止整个计算。另一个例子是多副本信息一致性问题。可见多重性资源使得差错处理和恢复问题变得很复杂。多重性还给系统资源管理带来新困难。

其次，系统资源在地理上的分散性带来很多新问题。在松散耦合系统中，使用信件在进程之间通信，而通信将产生不可预测的、有时是巨大的延迟，特别是在远程网络所组成的分布式系统中更是这样。分布式的状态信息和不可预知的信件延迟，使得系统的控制和同步问题变得复杂；要想及时地、完整地搜集到系统各方面的信息是很困难的，从而使处理机进行最佳调度很困难。在分布式系统中，用户或程序常常要访问远程资源如数据、程序或进程等，跟访问本地资源不同，也变得复杂了。例如可用open、close、

read或write调用访问本地文件,而访问远程文件时则需执行一个文件传送协议,并且作以该文件改名的操作。

在异构型分布系统中,由于各种不同资源的数据表示和编码、控制方式等均不相同,就产生了翻译、命名、保护和共享等新问题。

由于上述原因,分布式系统的研制,特别是软件的验证、调试、测量和维护问题变得很复杂。这正是分布式系统研制者要解决的主要问题。

1.2 分布式系统与计算机网络

分布式系统不一定由计算机网络组成,紧密耦合分布式系统就是这样。但松散耦合分布式系统是用计算机网络(远程网或局部网)组织起来的。松散耦合分布式系统与计算机网络系统都是多机系统,从硬件组成来看没有什么区别,很多作者对此不加以区别。例如世界上著名的计算机网络ARPANET,很多学者认为它也是分布系统。另一种观点认为,充其量它只能叫作计算机网络,不能叫分布式系统。本书采取后一种观点。这是因为,根据分布式系统的定义,尽管分布式系统由多个计算机组成,但从用户或编程人员角度看到的应该是一个单一的计算机系统整体;而计算机网络就不是这样,用户必须用命令明确指定一个计算机使用,指定一个文件从什么地方传送到另一个地方。粗略地说,如果用户能说明他在用哪一个计算机,则他是在使用一个计算机网络而不是分布式系统。一个真正的分布式系统的用户不必知道他的程序正在哪个机器上运行,他的文件在哪里存放,等等。使分布式系统具有这样性质的是它的软件:分布式操作系统。

计算机网络在产生初期(60年代末)并没有操作系统,只有若干面向功能的协议,如虚拟终端协议和文件传送协议等。后来(70年代初期和中期)为了更好地发挥资源共享的潜力,研制了网络操作系统。这种网络操作系统一般不具备分布式操作系统的全部功能。典型的例子是用局部网络把若干个人计算机和共用的打字机服务器、文件服务器连接起来。这种系统一般具有以下特点:

(1) 每个计算机都运行自己的操作系统,而不是运行共同的、系统范围的操作系统的一部分;

(2) 每个用户通常工作在自己的计算机上,以“远程登录”的方式使用不同的机器,而不是动态地给CPU分配进程;

(3) 用户知道他们的文件都放在哪里,在机器之间移动文件时必须明确地使用“文件传送”命令;

(4) 系统根本没有或具有很少的容错能力。

所以说是软件而不是硬件决定了一个多机系统是分布式系统还是计算机网络。但是,无论如何,网络操作系统是分布式操作系统的先驱,可以发展成分布式操作系统。最近几年,一些研究人员将一批运行UNIX操作系统的小型机用局部网络连接起来运行网络操作系统,其中有的已具备分布式系统的特征。

区别计算机网络和分布式系统最重要的一个方法是如何知道用户事实上是否在使用多个计算机。下面从三个方面来说明这个问题。

1.2.1 文件系统

在把两个或更多的不同计算机系统连到一起时，首先一个问题是如何把它们的文件系统合并到一起。可以有三种方案：

第一种是完全不合并，计算机A上的程序不能用系统调用访问计算机B上的文件，而是运行一个文件传送程序，将B上的文件复制到A上来，然后再在A上访问此文件。USENET网络就是一个著名的例子，它使用的文件传送程序是UNIX的uucp。

第二种方法是网络操作系统的方法，把不同的文件系统连接起来，一个机器上的程序可以使用路径名打开另一个机器上的文件。例如

```
open( "/machine1/pathname", READ);
```

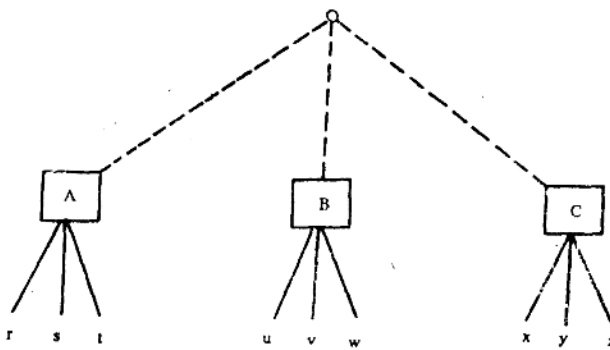
或

```
open( "machine1!pathname", READ);
```

或

```
open( "/../machine1/pathname", READ);
```

其中最后一种是用在Newcastle连接以及 Netix 网络操作系统中，将若干 UNIX 系统连到一起，在所有的UNIX 系统的根目录上创建一个虚拟的“超级目录”，“/··”表示从本地根目录开始上升一级到



“超级目录”上，然后再下降到“machine”的根目录上。图1.2.1示出计算机A、B、C的三个根目录和它们的“超级目录”。为了访问计算机C上的文件X，可以使用这样的系统调用：

```
open( "/../C/x",  
      READ-ONLY)
```

图1.2.1 用虚拟的超级目录访问远程文件

在Newcastle系统中，这种“命名树”实际上是很通用的，因为“machine1”事实上可以是任何目录，所以可把某机器的目录树作为“叶子”挂到“虚拟树”上的任何地方。

在分布式系统中，所有各子系统的文件系统组成一个整体系统，只存在着一个二进制程序的“bin”目录、一个password文件、一个……。任何机器上的程序想要读password文件时都使用系统调用

```
open( "/etc/password", READ-ONLY)
```

而不必指出该文件在哪里。分布式操作系统LOCUS使用这种方法。可见，这种采用全局命名空间的办法是很方便的。此外，在这种系统中，操作系统可以在各机器之间自由移动文件，保持各磁盘空间的有效利用。也可以在必要时设置任意数目的副本文件。

1.2.2. 保护

UNIX和其他许多操作系统给每个用户赋予一个唯一的内部标识符。文件系统中的