

一个微机网络操作系统—— Uni NOS 的设计与实现

程 岩 朱逸芬

(合肥工业大学)

摘要 本文分别从计算机网络角度和分布式系统的角度, 考察了网络操作系统的定义和特点、实现方法, 并提出了基于 Unix 7⁺、从分布式角度出发的网络操作系统 UniNOS 的结构模式。UniNOS 引入了网络特别目录文件, 并扩充了文件描述字和进程标识号的意义, 使其具有全局性概念, 从而把本地操作和网络操作有机地统一起来, 做到了网络对用户的透明性。为支持分布式系统的开发, 系统增加了远程过程调用、消息传送机构, 加强了进程同步机构, 使得网络际的任意进程间的通讯和同步成为可能。随着 Unix 内核网络功能的增加, 附在其上的 Shell 命令也随之具有了远程访问的功能。

DESIGN AND IMPLEMENTATION OF UniNOS—— A MICROCOMPUTER NETWORK OPERATING SYSTEM

Abstract This paper discusses the definition, characteristic and implementing method of NOS in the scope of computer network system and distributed system, respectively. And it advances the network operating system UniNOS structure model in the scope of distributed system. UniNOS introduces a special network directory file and enlarges the area which the file descriptor and process identifier express. The file descriptor and process identifier have a globe meaning in UniNOS. Therefore, the local operation and network operation have the same interface to users, so the system has good network transparency to users. Remote process call, message-oriented interprocess communication and remote process fork are added into the system to support the development of distributed applications. UniNOS strengthens the synchronous process mechanism and makes it possible to communicate between any process and synchronize the operation in the network wide. With the network function being added into the Unix kernel, the Shell on the Unix kernel also has the remote access function.

一、引 言

微型机局部网已经普遍被认为是分布式计算机系统的物质和技术基础。

计算机网络是指多台计算机通过通讯设备和链路相互连接, 使任何一台计算机上的资源能够被近程或远程的各台计算机所共享; 分布式计算机系统是由多个分散计算机通过互连网

络构成的统一计算机系统，其中各个物理和逻辑资源元件既互相配合，又独立的在全系统范围内实现资源管理和在动态基础上实现任务或功能分配与并行地运行分布式程序。由此可见，局部网络在资源分布、互连拓扑、通信协议等方面与分布式要解决的问题是相同的，但它对于全局性管理、并行操作、自治控制等无硬性要求。所以说，局部网络可视为分布式计算机系统的初级阶段。

在微机局部网络上开发分布式系统，那么首先要求微机局网是开放式的，即留有进一步开发和扩充的接口；其次，要求尽量使原软件可以不加修改地运行于网络环境。因为，在微机局网上开发分布式系统，往往要运用原来软件作为开发工具和实施基础，所以有必要保证网络对原有软件的透明性。

目前，在我们所见到和用到的较成熟的局网产品中，都不能完全满足上述要求。例如，具有代表性的 3 Com Ethernet 局网类和 IBM PC 宽带网，它们的特点是：

3 Com Ethernet 将网络的实用命令直接提供到终端用户，在执行一系列命令后，原软件可运行于网络环境。但是它不是开放式的，未向用户提供任何开发接口，那么，在这样的局网上，进一步开发网络和分布式软件，则是一件困难的事了。

IBM PC 宽带网是一个开放式的局网系统，它为用户提供网络各层的网络调用，用户可用此调用来开发网络和分布式软件。但它的缺点是原有软件一个也不能运行于网络环境，所有的网络应用层都要开发。

为了支持分布式系统的开发，必须克服微机局网的现行产品的缺陷，从而设计开发出适合分布式系统开发的网络软件。

我们想到，所有的软件都是以系统调用方式调用操作系统来完成其功能的，如能使系统调用具有网络功能，即根据参数的不同，自动识别本地和网络操作，即从系统内部实现网络功能，那么就可满足以上的要求。

根据以上思想，就产生了我们的课题，即基于 Unix 7+ 操作系统，从操作系统角度出发，设计网络操作系统 UniNOS，来满足在其上进行分布式系统开发的要求。

二、网络操作系统的概述

1. 网络操作系统的定义

网络操作系统从不同的角度来考虑，有着不同的理解和定义，也有着不同的实现方法。

(1) 从计算机网络角度来考虑

网络操作系统是为解决异种机连网中不同操作系统之间的协作问题而提出的。H.C.Fors-disk 等人认为^[1]：网络操作系统是针对如下问题提出的：

① 为了有效地使用网络资源，用户不仅要知道并掌握网络的存取机构，而且要会使用其主机上的操作系统；

② 各个主机上的资源一般不是互相兼容的；

③ 很难获得关于资源的可用性和怎样使用的信息；

④ 记帐和帐单问题，通常它们都由各机自行管理，所以往往用户要分别在不同机上建立多个户头。

所以说，网络操作系统的概念代表一个约定的方法，从而克服以上的困难，继而挖掘出网络的全部潜力。

网络操作系统的定义：

网络操作系统是由一组软件和相关协议的组合，它们将各个自治的主机由网络连接起来以达到方便有效的方式使用。

按照 A.S. 达南伯姆的观点^[2]：网络操作系统是由运行在各种不同主机上的一组用户程序来实现的，这些用户程序都是附于各自的操作系统之上的。

MIMAS 认为^[3]：网络操作系统是实现于通常本地操作系统之上的一层软件。

综上所述，网络操作系统是连接并协调网

络上各主机的操作系统的一组软件。实现方法是趋向于在各单机操作系统之上,实现一层用于连接各操作系统,提供网络功能的软件。

从以上对网络操作系统的定义和观点,可看到他们有个共同特点,那就是都只从计算机网络角度来分析和理解的,并未涉及到操作系统的本身,也未涉及到系统的应用环境。

(2) 从分布式系统角度来考虑

前面已谈到,计算机网络对于分布式系统所要求的全局性管理、并行操作;自治控制等无硬性要求。但是,发展到分布式资源共享计算机网,使它的资源能够在整体尺度上进行管理,并且文件和进程能够根据用户要求和系统负载情况在全网范围内迁移,而用户无需确切知道为其服务的各资源的物理位置,那么它就向符合前述定义的分布式计算机系统靠近了一步。即使这样,它们在并行性和自治性的要求上仍不是一样的。所以只有在计算机网络、尤其是局网中引入并行性和自治性,才有可能进一步发展分布式系统。

作为分布式操作系统开发的第一步:网络操作系统,应该在并行性和自治性上多加考虑,必须对单机操作系统作一些必要的修改和补充,所以从分布式角度来讲,网络操作系统是一个具有网络功能的操作系统,它应该在一定程度上提供自治性,并在全网范围内实现资源管理,至少要提供在动态基础上实现任务和功能分配的手段和开发基础,为并行地执行分布式程序提供开发手段和物质保证。所以只是在操作系统之上实现一层软件,来完成网络功能,那是很难满足进一步的分布式开发的要求,在效率上也是达不到要求的,而且单机操作系统中某些封闭性因素也是分布式开发的桎梏。

A.S. 达南伯姆在给出网络操作系统的定义后,也是这样定义分布式操作系统的^[2]:把实用网络操作系统的程序从操作系统之外并入操作系统之内,用具有网络功能的操作系统来代替各个主机上的操作系统,这样就产生了分布式操作系统。

2. 网络环境下的操作系统

计算机系统一旦加入计算机网后,其用户可用资源的范围和自己的用户范围都大大地扩充了。所以这就要求网络中的操作系统既要为本机用户提供服务,又要为网络用户提供使用本机资源的服务。要求网络操作系统具有普通操作系统所没有的通讯功能,并且要提供一定协议软件模块,支持可靠的、坚定的进程间、用户间的通讯;由于网际用户不仅要用本站资源,而且往往要用其他站资源,所以网络操作系统还应具备一个全局资源管理模块,完成本机到全网和全网到本机,也即局部→全局,全局→局部的操作映象,从而完成全局性资源管理。

从分布式系统的角度看,网络操作系统应该提供故障恢复与检测。为拷贝一致性,全局性命名等分布式基本功能,从而为资源能够在整体上进行管理,而且文件和进程能够根据用户要求和系统负载情况在全网范围内迁移,而用户无需确切地知道为其服务的各资源的物理位置,为分布式操作系统的开发提供基础。

三、设计的主要思想

1. 从分布式角度来设计

据前所述,从不同的角度来考虑网络操作系统有着不同的理解,得出不同的结论,所以相应的设计和实现方法也有所不同。在设计我们的 UniNOS 时,结合本所的研究课题和方向,以及 UniNOS 将运行于局部网这一特征,在研制网络操作系统的同时,不仅考虑到网络应用的要求,要更多地考虑了操作系统本身和分布式系统的要求,本着提高全系统的效率,做到 UniNOS 对原来的 Unix 用户及其上的现存软件完全透明,为今后进一步的分布式开发奠定基础。根据 Unix 的特点,我们从分布式角度出发,采用了将网络功能结合到 Unix 内核中去,使之成为核心不可缺少的一部分来运行的方法而设计的。

众所周知，Unix 系统以它的短小精悍、层次分明的分级式树型文件系统、功能较强的 Shell 命令解释程序和与硬件无关性、高度的可开发性和可移植性而著称。基于 Unix 的特点，将网络功能结合到内核中去，使之成为内核的一部分；这有它的可行性和优越性。可行性在于：① Unix 系统短小精悍，而且大部分由高级语言编写，便于进行开发和移植；②我们已将 Dual 68000 系统上配制的 Unix 7+ 原代码还原成 C 语言，在此基础上做了较深入的解剖分析工作，这就为本方案的实施提供了物质和技术基础。其优越性在于：① 因为用户同 UniNOS 的接口通过标准 Unix 调用，所以可以把网络设计成完全透明的；② 由于保持了原系统与用户的界面，原 Unix 的 Shell 命令解释程序可运行于网络环境，所以无需再设计网络应用层；③ 因为网络程序是作为低级的操作系统子程序运行的，所以，UniNOS 的性能可由系统实现，设计者可作适当调节；④ 由于网络功能在系统内部实现，这就减少了用户空间和系统空间的数据流的交换所带来的系统开销，从而提高效率。

采用把网络功能结合到 Unix 内核中去的方法来设计网络操作系统，国外已有不少成功的经验。具有代表意义的有 Unix 4.2 BSD 系统和 Xenix 系统，它们已成功地用于多种 Unix 网络，但还有其不足之处。对于 Unix 4.2 BSD 系统来说，它所增加的套节字概念以及有关对套节字的系统调用就大大地影响了网络对用户的透明性。对于 Xenix 系统来说，其增加的远程文件操作服务模块实际上是位于 Unix 内核的表面，即位于用户空间。当一个远程文件操作请求到达本站，先由系统把请求送给远程文件服务模块，此模块根据该请求，调用内核系统为其服务，然后又将服务结果送交该模块。此模块又调用操作系统的通讯原语，把结果送回系统空间以发到网上。这样，控制权和数据流在系统与用户空间交替了两次，这势必影响了效率。另外，Xenix 支持的网络对用户界面

是一个超级根，其它各站主机的根在此超级根之下，这样的处理方法，或多或少地影响了原 Unix 上软件兼容性和网络的透明性。

所以，根据国外对 Unix 开发的经验，我们吸收了 Unix 4.2 BSD、Xenix 等的精华部分，继而提出 UniNOS 的结构模式。

2. UniNOS 的结构模式

从整体上来看，UniNOS 的结构模式如图 1 所示，分为三层：

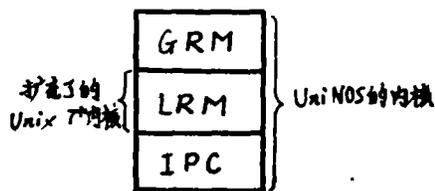


图 1 UniNOS 的结构模式

第一层：交互进程通讯层 IPC，该层是 UniNOS 与网的接口模块，也是通常的操作系统所缺少的机间通讯模块。此层向其他各层提供发送、接收信息功能，具体地完成网络通讯工作。

第二层：局部资源管理层 LRM，由该层管理本地资源，无论是本站或是它站，对本地资源的操作请求都是由该层完成。

第三层：全局性资源管理 GRM 层，由该层来分析用户命令和程序，进行必要的转换和相应的处理。它通过调用其他各层，实现全局性资源管理，完成全局↔局部操作的相互映象，为系统维护着一张记录网上各站状态的表，并负责完成网络站的登录和上下网的操作。

四、UniNOS 的实现

1. IPC 层

IPC 层是由发送和接收这两个进程具体实现的。这两个进程随着上网操作的执行而产生，随着下网操作的执行而消亡，此层的操作对象是三个队：发送队 sq。所有的要发出去的报文都经过此队，由发送进程发生；接收队 rq，

所有从网上收到的报文全部排入此队，以待上层处理；故障恢复队 **crq**，所有试图发向某一故障点的报文，如有必要，都将排入此队，以待故障恢复后，放入 **sq** 队。IPC 层为上层提供了发送和接收原语。

IPC 层通讯采用了两相确认式算法，既发送进程在发出一报文后，必须等应答，如应答正确，则发下一报文，如应答不正确，则重发此包。在应答超时或重发次数超量的情况下，则报告出错，作故障处理。

2. 局部资源管理 LRM 层

LRM 层是扩充了的 Unix 资源管理模块。

(1) 文件管理

UniNOS 的文件系统就全网来说，是一个分布式文件系统。单机上的树型文件结构覆盖了网上所有站的文件系统。在 UniNOS 中，远程文件的存取对用户来说是透明的。

在 UniNOS 中，为了达到网络对用户完全透明，引入了网络特别目录文件的概念。

所谓网络特别目录文件，就是在根目录下建立一个网络特别目录 **net**，**net** 下的各目录就是网络特别目录，它实际上是网上各站的固有名。**net** 下所有站名目录，都与各站根目录相连，用户所看到的网络各站，就是他的 **net** 下的各个目录。用户是通过对 **net** 下各站名目录的访问，来实现对他站文件系统的访问，实现网络操作。上网之前，**net** 是一个空目录，一旦上网后，从他站获得的回答信息，来建立各个网络特别目录文件。

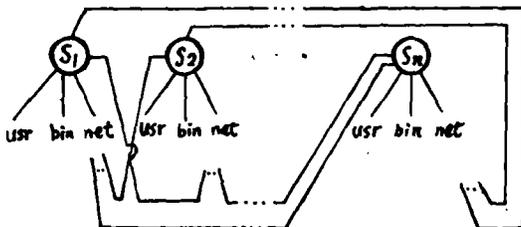


图 2 UniNOS 的树型文件结构

在 UniNOS 中，将原 Unix 的文件描述字的意义做了扩充，使其不仅描述了文件所在的

打开表的位置，也隐含了文件所在站的站号，这样使得对文件系统的系统调用，系统都有判别是否远程操作的依据。

(2) 进程管理

1) 公共服务进程

所谓公共服务进程就是远程服务进程只有一个，它为所有的远程用户、远程进程所共享，由它来统一地完成各个站、各个用户发来的服务请求，并送回结果。

采用公共服务进程有以下优点：

- ① 减少了系统为各个用户又分代办进程的开销，简化了管理；
- ② 使系统实现起来简单、统一；
- ③ 节省了 **proc** 表资源；
- ④ 减少了远程进程间的通讯频率。

2) 远程文件的执行

在网络资源共享中，有一个很重要的共享概念，那就是软件共享。用户远程调用的过程，也就是一个远程文件执行的过程。

为使远程调用与本地调用统一起来，使用户感到调用方式相同，UniNOS 引入一个远程进程又分的概念。所谓远程进程又分就是父进程在指定的远程机上又分出一个子进程，从而完成远程过程调用。实际上，在又分一个远程进程时，是公共服务进程代替父进程又分子进程的，而并不是真将父进程的图象送过去。远程调用的模式如图 3 所示：

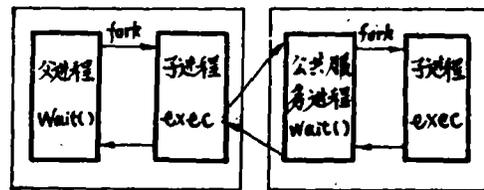


图 3 远程调用模式

3) 进程通讯机构

Unix 虽然提供了事件机构和 **pipe** 机构的进程通讯机构，但它们都存在着应用面窄、通讯量小的缺陷。在分布式环境下，任意进程的通讯发生得很频繁，而且通讯量有时很大。显

然 Unix 提供的进程通讯机构在功能和效率方面都达不到要求。UniNOS 在保存原有的进程通讯机构的前提下,增加了消息传送机构,从而较好地满足了分布式环境要求,加强了 UniNOS 的进程通讯机构的功能。

消息的发送就是将指定的消息挂到接受进程的消息链的队尾,接收消息就是从本进程的消息链队中按规定取一个消息处理之。消息传送机构又新加了三个系统调用向用户开放;

`sendmsg()`: 向指定进程发送消息。

`getmsg()`: 接收指定进程最早发来的消息。

`reply()`: 回答

UniNOS 所提供的消息传送机构具有很大程度的灵活性和可靠性,效率也是较高的。

4) 进程同步机构

从根本上来说,Unix 缺少一种完善的进程同步机构。尤其当系统进入分布式网络系统环境,这种缺陷更为突出。在一个分布式网络系统中,在执行某一分布式算法时,必须与网际的许多进程协商,从而得出最优决策。这就需要有一个功能很强的同步机构,UniNOS 又新加了用户级睡眠唤醒机构,借此满足进程同步的要求。

用户级睡眠唤醒机构大大地加强了系统的进程的同步功能。

3. 全局性资源管理 GRM

GRM 层为系统维持着一张记录当前全网各站的状态表:网络状态表 NST。此表各站有之,而且各站的 NST 都必须是一致的。随着各站的上下网,故障的发生和恢复,此表将统一地及时地更新着。

GRM 层还为系统管理提供了上网 `onnet` 和下线 `offnet` 两条命令,以完成上下网的操作,用户不得使用它们。

GRM 层的主要任务还完成全局↔局部的

操作映象。

当操作进入系统后,由 GRM 首先判别是否远程操作。如果是,则装好请求服务包,发向网上,如果不是,则装入正常处理。这就是局部→全局的映象。

全局→局部映象是由公用服务进程完成的。它不断地检查接收队 `rq`,实现完成具体请求,然后再发回服务结果。

结 束 语

UniNOS 已成功地运行于本所的 Dual 68000 微机(三台)网上,由于它与物理层有着统一的接口,可以支持任何物理层,只要主机配的是 Unix 7+ 操作系统,它与物理层的通讯方式无关。由于 Unix 的 Shell 命令有了网络功能,所以用户可以不用管网的存在,而用网上各地站的资源,透明性很好,原 Unix 上的软件可不加修改地运行于配制 UniNOS 的网络环境。总之,UniNOS 的研制为下一步分布式操作系统的研究与开发提供了良好的环境和基础,为本所的分布式数据库的研究提供了物质和技术基础。

本系统尚有一些工作有待于进一步的完善和开发:① 文件的多拷贝的一致性、全局性动态负载平衡;② 将同种机网扩展到异种机网。

主要参考文献

- [1] Hang C. Fordick etc al: "Operating Systems for Computer Networks," <Computer> Vol. 11, №1, Jan 1978
- [2] A. S. Tanenbaum: "Computer Networks" 1981
- [3] G. S. Blair, et al: "MIMAS—A network operating system for strathnet" The 3rd Inter. Conf. on Distributed Computing Systems 1982.