

CpuMemSets在Linux操作系统中的实现（1）PDF转换可能丢失图片或格式，建议阅读原文

[https://www.100test.com/kao\\_ti2020/144/2021\\_2022\\_CpuMemSets\\_c103\\_144497.htm](https://www.100test.com/kao_ti2020/144/2021_2022_CpuMemSets_c103_144497.htm) 一、前言 非一致性内存访问

（Non-Uniform Memory Access）结构是分布式共享内存（Distributed Shared Memory）体系结构的主要分支，它通过结合分布式内存技术和单一系统映像（SSI）技术，实现了SMP系统的易编程性和MPP系统的易扩展性的折中，已成为当今高性能服务器的主流体系结构之一。目前国外著名的服务器厂商都先后推出了基于NUMA架构的高性能服务器，如HP的Superdome、SGI的Altix 3000、Origin 3000、IBM的x440、NEC的TX7、AMD的Opteron等。随着NUMA架构的高性能服务器被逐渐推广，系统软件针对这种分布式共享内存架构的特点，在调度器、存储管理和用户级接口等方面进行了大量的优化工作。例如，SGI的Origin 3000 ccNUMA系统在许多领域得到了广泛应用，是个非常成功的系统，为了优化Origin 3000的性能，SGI的IRIX操作系统在其上实现了CpuMemSets，通过将应用与处理器和内存的绑定，充分发挥NUMA系统本地访存的优势。Linux社区在自己的NUMA项目中也实现了CpuMemSets，并且在SGI的Altix 3000的服务器中得到实际应用。 本文将以SGI的ProPack v2.2为研究对象，分析CpuMemSets在Linux-2.4.20中的具体实现

。CpuMemSets是SGI进行的一个开放源码项目，由针对Linux2.4内核的补丁、用户库、python模块和runon等命令共四部分组成，以实现处理器和内存块的分区为目标，控制系统资源（处理器、内存块）面向内核、任务以及虚拟存储区

的分配，为 dplace、RunOn 等 NUMA 工具提供支持，最终优化 Linux 系统的 NUMA 性能。二、相关工作 分区技术

( Partition ) 最初出现在大型机 ( MainFrame ) 上，如今被广泛应用到服务器领域，支持在单个服务器上运行一个操作系统的多个实例或者多个操作系统的多个实例，主要特点是机器独立、屏障可靠、单点管理。在分区技术支持下，当前多台服务器运行的多个操作系统就可以在同一地点的一台服务器上同时运行，优于在一个组织中四处分散用多个服务器来支持不同的操作系统，从而有效地实现了服务器整合。支持分区技术的服务器可以当作应用服务器，运行Windows平台供市场部门使用；同时还可以运行Linux系统供工程部门使用。还可以在大多数用户运行Windows 2000 Advanced Server系统的同时，在另一个分区中为发展组测试其它操作系统；或者所有节点都应用在一个操作系统环境下。各种分区实现技术的主要差别体现在分区故障隔离手段（硬件或软件）、分区资源粒度、分区资源灵活性以、虚拟分区资源以及对动态分区重构的支持等方面。典型的有IBM的LPAR和DLAPAR（AIX 5L 5.1）、HP的nPartitions和vPartitions（HP-UX 11i）、SUN的Dynamic Domains（Solaris 8）、以及Compaq的Alpha Servers（Tru64 Unix 5.1）。但是，针对NUMA系统采用的分区技术与NUMA系统本身具有的单系统映像优势是矛盾的。

100Test 下载频道开通，各类考试题目直接下载。详细请访问 [www.100test.com](http://www.100test.com)