

多核对分区操作系统影响研究

赵纯,龙翔,王雷

(北京航空航天大学 计算机学院,北京 100191)

摘要: 分区操作系统是综合化航空电子领域中的核心技术。随着单核性能极限的到来,处理器结构向着多核发展。将两者结合起来,在多核分区操作系统的基础上研究分析多核处理器结构为分区操作系统带来的影响。经分析实验数据得出多核处理器结构在多核处理器中共享 Cache 结构和内核中临界资源并发访问两方面对分区操作系统产生影响。

关键词: 多核;分区;操作系统

中图分类号: TP316

文献标识码: A

文章编号: 1674-7720(2012)02-0053-03

Design and implementation of scheduler for partition operating system on multi-core

Zhao Chun, Long Xiang, Wang Lei

(School of Computer Science and Engineering, University of Beihang, Beijing 100191, China)

Abstract: Partition operating system is the key technology of integrated modular avionics (IMA). With the coming of single CPU's bottleneck, CPU structure is developing toward multi-core. Based on the implementation of multi-core partition operating system, this paper revealed the influence to the partition operating system which is brought by multi-core CPU structure through some well-designed tests. Multi-core CPU structure affect operating system in two aspects, one is the sharing cache of different CPUs; and the other is the kernel lock which protected the concurrent access to critical resources from different partitions on different CPUs.

Key words: multi-core; partition; operating system

综合化航空电子 IMA (Integrated Modular Avionics) 系统是第四代航空电子的发展方向。综合化航空电子致力于最大化利用计算机硬件资源,减少航空电子设备对飞行器的负载,提高飞行器运载能力。

分区机制是综合化航空电子的核心技术。分区机制是指将系统资源(内存、I/O 等)进行配置,建立多个分区,分区中包含系统所分配的资源,分区与分区之间相互隔离。分区包括时间分区和空间分区。时间分区是指系统为分区之间提供时间维度的隔离,分区都有属于自己的处理器独立使用时间,并且时间不受其他分区影响;空间分区是指系统为分区之间提供空间维度的隔离,空间包括内存地址空间及 I/O 地址空间。通过空间分区隔离技术分区无法访问本分区外的空间资源,而外分区同样无法访问本分区内空间资源。不同应用运行于不同的分区内,应用之间由于分区间的隔离而无法感知到其他分区应用的存在而认为独享整个硬件资源。所以通过分区机制可以实现将运

行于不同硬件上的应用运行于同一硬件平台,并且相互之间互不影响,从而实现综合化航空电子系统。

单核处理器已经达到性能瓶颈,无法靠增加频率、带宽带来性能的提升。处理器结构正向多核发展。处理器这种结构上的变化在多方面影响操作系统的设计与实现,比如多核处理器系统中任务并发访问临界资源以及多核处理器结构中 Cache 一致性等。

多核处理器在硬件层面的变化会影响到操作系统及相关应用,本文通过实验分析了多核结构带来的影响。

1 多核分区操作系统

1.1 LESOS 操作系统

LESOS 是北航嵌入式实验室一款多核分区操作系统。该操作系统在硬件上支持现在常见的多核处理器 SMP (Symmetrical Multi-Processing) 结构,在软件上支持线程、进程、虚拟内存管理等功能。系统中相应的参数如表 1 所示。

技术与方法

Technique and Method

表 1 LESOS 系统相关服务性能

功能	时间/ μs	功能	时间/ μs
内存申请(1 MB)	4.85	销毁信号量	12.45
系统调用	0.67	消息队列创建	5.51
信号量创建	9.57	发送消息(1 B)	3.76
获得信号量	2.00	接收消息(1 B)	3.30
释放信号量	2.03	销毁消息队列	7.61

本文后续工作都基于此操作系统完成。

1.2 分区模型

分区操作系统中,分区是资源的载体,任务是操作的载体。系统为分区分配其所需的资源,如内存、I/O、处理器执行时间等;分区则通过具体的任务完成相应的功能。一个分区中包含有若干不同的任务,每个任务有不同的任务优先级;分区内部的所有任务共享此分区内部资源。系统中每个任务属于且只能属于一个分区,任务不能脱离分区单独存在。

1.3 分区调度

双层调度模型选择一个要执行的任务,首先需要在分区一级进行调度,选择相应的分区作为候选分区,然后在候选分区的所有任务中选择相应的任务作为候选任务。由于在调度过程中需要在两个级别进行两次选择,所以称为双层调度。

首先在分区内部的调度中,调度器的调度单位是任务。此层调度使用的是基于优先级抢占式的调度算法,目的是使分区内部重要的任务能够及时得到处理。调度器在所需调度的分区内部所有的任务中选择优先级最高的任务作为此次调度的结果。

其次在分区之间的调度中,调度器的调度单位是分区。此层调度使用静态时间片轮转调度算法,目的是使各个分区之间都能够得到相应的运行时间,不会因为某些分区的异常而导致其他分区无法运行。调度器依据当前系统时间以及静态调度表选择适当分区作为此次调度的结果。

最后为了支持多核处理器,调度器为每个处理器核心维护一张静态调度表,这样每个处理器可以在自身的静态调度表中进行调度,从而避免相互影响。每个处理器需要获得一个要执行的任务时,首先通过处理器本身的静态表进行分区调度选择合适的分区,然后在选中的分区中进行分区内部的任务调度选择出合适的任务,此任务就是处理器下一个要执行的任务。

2 测试及数据

多核处理器与单核处理器结构上的不同会影响操作系统及上层应用程序的行为、性能。本文在前述工作的基础上通过以下实验研究分析多核处理器结构为分区操作系统所带来的影响。

2.1 测试环境

本文的实现工作是基于 Freescale 的 HPCNet-8641D

平台。平台中两颗 PowerPC E600 处理器运行于 SMP 模式。平台基本参数如下:处理器频率 1 GHz,系统总线频率 400 MHz,内存 512 MB。两颗处理器中分别包含指令、数据为 32 KB 的一级 Cache 以及共享 1 MB 的二级 Cache。此外,平台还有 Dual-UART、千兆以太网卡等外设接口。

2.2 多核分区对 Cache 的影响

由于多核处理器结构从物理层面上引入了共享 Cache,所以本文设计以下实验分析多核处理器结构在 Cache 方面对上层应用的影响。

首先,在单核分区操作系统上分别执行不同大小数据量的读、写、拷贝任务,测试访问不同大小数据量的相关操作执行时间。因为在多核处理器上,Cache 分为两级,一级为 32 KB 处理器本地 Cache,以及共享 1 MB 的二级 Cache,所以在设计实验时将数据量分为两类,一类是一级 Cache 内不同数据量的访问测试,另一类是大于一级 Cache 大小数据量访问的测试。其次,在多核分区操作系统上处于不同处理器上的分区同时运行上述任务,收集相关测试数据。最后,将两者同性质的数据进行比较,结果如图 1~图 6 所示。

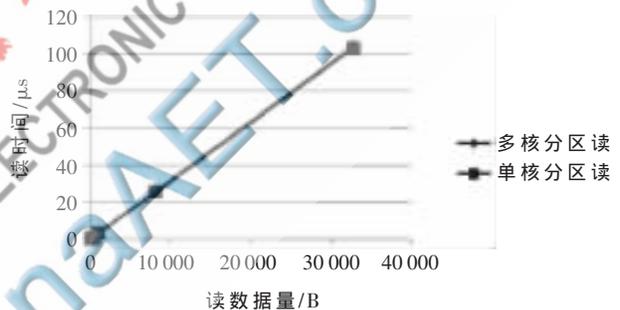


图 1 一级 Cache 内读性能对比

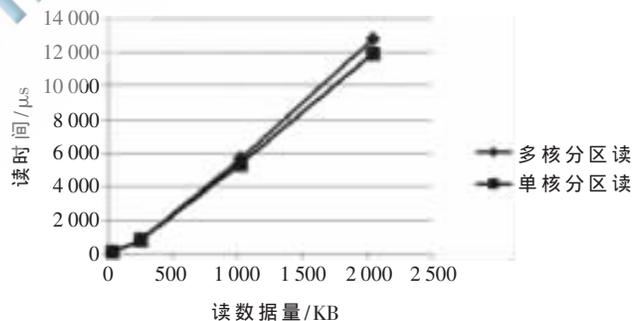


图 2 一级 Cache 外读性能对比

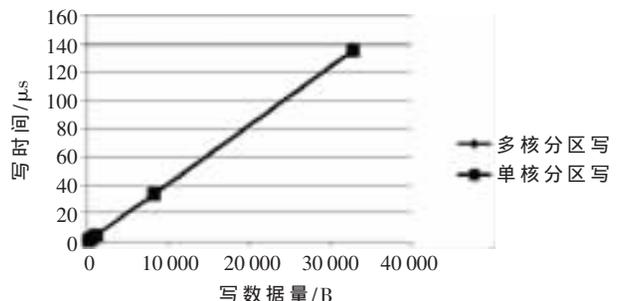


图 3 一级 Cache 内写性能对比

技术与方法

Technique and Method

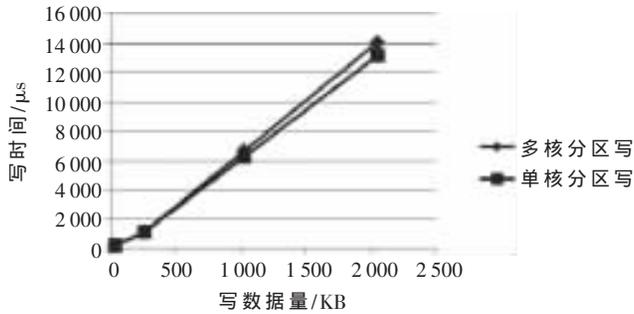


图4 一级 Cache 外写性能对比

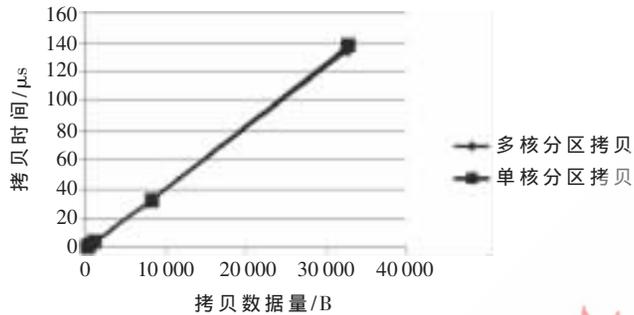


图5 一级 Cache 内拷贝性能对比

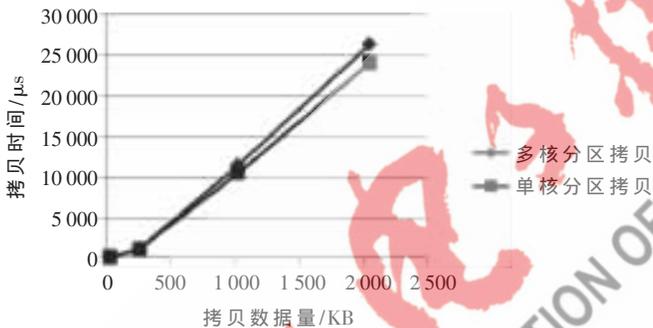


图6 一级 Cache 外拷贝性能对比

通过以上数据可以得出以下几点结论:(1)对于一级 Cache 容量(32 KB)大小以内的数据访问,单核分区系统与多核分区系统执行性能相差无几,这是因为此时对数据的访问大多集中于分区所在处理器本身的 Cache 中,所以多核分区性能并没有受到影响;(2)当任务访问数据量大于一级 Cache 容量(32 KB)后,多核分区中的数据访问任务性能要略低于单核分区中相关任务的性能。这是因为此时分区同时运行在不同的处理器上,分区中数据访问任务存在对共享的二级 Cache 的争用,导致分区数据访问过程中二级 Cache 命中率降低,从而引起多核分区中数据访问性能下降。

2.3 多核分区间对锁的竞争

在多核处理器结构系统中,运行于不同处理器上的分区会对某一临界资源并发访问。操作系统为保证此种并发访问结果的一致性,需要加锁以保护临界资源。在访问相应临界资源时,任务需要先获得相应互斥锁,然后再对资源进行访问、修改,最后释放互斥锁。以下实验研究分析多核处理器结构中增加锁开销带来的影响。

首先,构造分区任务,分区任务在执行过程中循环调用 malloc 函数,申请 1 MB 的内存,然后调用 free 函数释放刚申请的内存,在任务执行过程中分别统计调用 malloc 及 free 函数所使用的时间。其次,按照不同的分区之间相关系数将构造的分区任务分别运行于不同处理器的分区之中。最后,测试不同用例中分区任务的运行情况,统计、分析实验数据。

实验按照以下方式构造不同相关系数的分区任务,相关系数如图 7 所示。不同处理器上的两个完全相同的分区同时运行,此时两个分区由于行为完全一致,所以两个分区之间的相关系数为 1。然后将一个处理器上的分区与之前一个分区错开 20% 的执行时间,两个分区之间只有 80% 的重叠部分,这样两个分区之间的相关系数为 0.8。依此类推,当两个分区执行时间完全不重叠时,给定时间内只有一个分区运行不会受其他分区影响,所以此时分区之间的相关系数为 0。

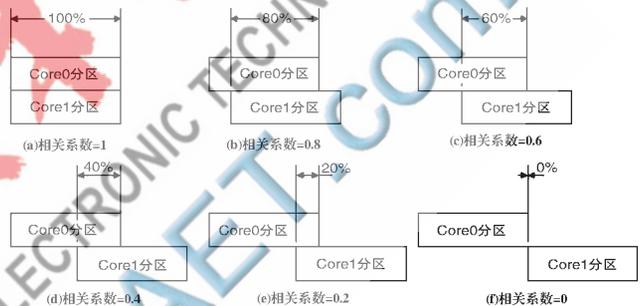


图7 分区相关系数

图 8 所示为数据整理,从数据中可以看出在多核分区系统中,不同处理器上的分区运行任务的相关系数越高,其相互之间的影响越大,性能损失也越大。造成这种情况的原因是由于同时运行于不同处理器上的分区,随着相关系数的增加对相关资源共享程度就越高,因而就会有一部分性能损失在获得锁和释放锁中。在最坏情况下,如果不同分区中所有资源都需要互斥访问,那么分区并发执行将由于锁的原因退化成顺序执行。

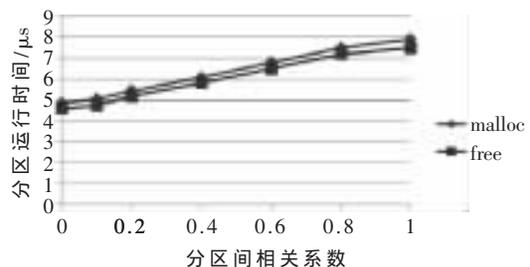


图8 分区相关系数影响

经过上述两个对比实验可以得到以下两点结论,首先,多核处理器由于在硬件上处理器之间需要共享 Cache,并发运行的分区会对共享 Cache 争用,从而影响其他分区的运行结果。此种情况由硬件结构决定,软件可做工作很少。其次,多核分区系统在内核中保护临界

技术与方法 Technique and Method

资源的锁开销会影响分区运行的性能。为了避免此种性能损失,系统在安排分区时可以进行优化,将相关性不高的分区安排在同时运行,避免高相关性分区同时运行,从而降低对临界资源的争用,提升系统性能。

参考文献

- [1] Airlines electronic engineering committee(AEEC).Avionics application software standard interface(ARINC specification 653-1)[S].ARINC, Inc., 2003.
- [2] RUSHBY J.Partitioning in avionics architectures:requirements, mechanisms and assurance[R]. Technical Report NASA CR-1999-209347, SRI International, California, USA, 1999.
- [3] POWELL M L, KLEIMAN S R, BARTON S, et al. SunOS multi-thread architecture[A]. USENIX Winter Conference[C],

1991.

- [4] KEPECS J. Lightweight processes for UNIX implementation and applications[A]. USENIX Summer Conference[C]. 1985:299-308.
- [5] BLACK R, FLETCHER M. Next generation space avionics: a highly reliable layered system implementation[A]. IEEE, 2004.

(收稿日期:2011-09-20)

作者简介:

赵纯,男,1988年生,硕士在读,主要研究方向:操作系统研究。

龙翔,男,1963年生,教授,主要研究方向:操作系统研究。

王雷,男,1969年生,副教授,主要研究方向:操作系统研究。

电子技术应用
APPLICATION OF ELECTRONIC TECHNIQUE
www.ChinaAET.com