



(12)发明专利

(10)授权公告号 CN 101694631 B

(45)授权公告日 2016.10.05

(21)申请号 200910235303.0

(22)申请日 2009.09.30

(65)同一申请的已公布的文献号

申请公布号 CN 101694631 A

(43)申请公布日 2010.04.14

(73)专利权人 曙光信息产业(北京)有限公司

地址 100193 北京市海淀区东北旺西路8号
院36号楼

(72)发明人 聂华 邵宗有 厉军 张海阔

李媛 张涛 李伟伟

(74)专利代理机构 北京德恒律师事务所 11306

代理人 梁永

(51)Int.Cl.

G06F 9/48(2006.01)

G06N 3/12(2006.01)

(56)对比文件

CN 1744046 A,2006.03.08,说明书第6页第1行至第10页第13行、图1,2.

CN 1877534 A,2006.12.13,说明书第3页倒数第7行至第4页第4行.

WO 2005069155 A2,2005.07.28,全文.

审查员 董洪梅

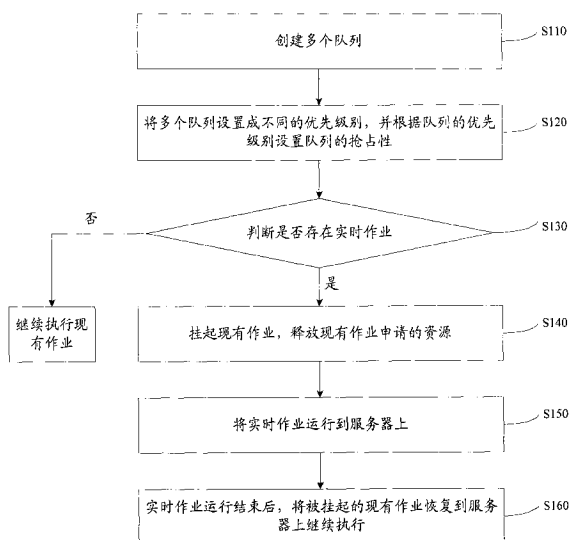
权利要求书2页 说明书4页 附图2页

(54)发明名称

实时作业调度系统及方法

(57)摘要

本发明提供了一种实时作业调度系统和方法。其中,该方法包括以下步骤:创建多个队列。将多个队列设置成不同的优先级别,并根据优先级别设置队列的抢占性。判断是否存在实时作业,如果不存在,则继续执行现有作业;如果存在,则将现有作业挂起,并释放掉现有作业申请的资源,然后将实时作业运行到服务器上。在实时作业运行结束后,再将被挂起的现有作业恢复到服务器上继续运行。本发明通过采用优先调度机制和抢占机制,实现了应用程序的实时性。



1. 一种作业调度的方法,其特征在于,所述方法包括以下步骤:

A. 创建多个队列;

B. 将所述多个队列设置成不同的优先级别,并根据所述队列的优先级别设置所述队列的抢占性;

C. 判断是否存在实时作业,如果存在,则执行步骤D,否则继续执行现有作业;

D. 将所述现有作业挂起,释放掉所述现有作业申请的资源;

E. 将所述实时作业运行到服务器上;以及

F. 所述实时作业运行结束后,将所述被挂起的现有作业恢复到所述服务器上继续运行,

其中,采用自适应算法设置所述队列的优先级别和抢占性,包括以下步骤:

收集集群信息,并查找所述集群信息中关于集群规模的信息;以及

根据所述集群规模设置所述队列的优先级别和抢占性,

其中,所述判断是否存在实时作业包括以下步骤:

A. 判断作业相关的用户是否具有提交实时作业的权限,如果具有所述权限,则执行步骤B1,否则判定所述作业为非实时作业;以及

B1. 判断所述作业的级别是否为高优先级,如果是高优先级,则所述作业为实时作业,否则为非实时作业,

其中,所述根据所述集群规模设置所述队列的优先级别和抢占性,包括:

如果所述集群规模小于集群规模的阈值,则将所述队列设置成两种优先级别,包括高优先级和低优先级,其中,所述高优先级队列的作业优先调度并抢占所述低优先级队列的作业的资源;以及

如果所述集群规模不小于所述集群规模的阈值,则将所述队列设置成三种优先级别,包括高优先级、中优先级和低优先级,其中,所述高优先级队列的作业优先调度并抢占所述中优先级队列的作业的资源 and 所述低优先级队列的作业的资源,所述中优先级队列的作业比所述低优先级队列的作业优先调度。

2. 根据权利要求1所述的方法,其特征在于,所述集群规模的阈值为32。

3. 一种作业调度的系统,其特征在于,所述系统包括:

队列设置模块,用于将多个队列设置成不同的优先级别,并根据所述队列的优先级别设置所述队列的抢占性;

实时判断模块,用于判断是否存在实时作业;以及

挂起模块,用于当存在实时作业时,将现有作业挂起,并释放所述现有作业的资源,以及在所述实时作业运行结束后恢复所述现有作业的运行,

其中,所述队列设置模块包括:

信息获取模块,用于收集集群信息,并从所述集群信息中获取集群规模信息;

优先级别设置模块,用于根据所述集群规模信息设置所述队列的优先级别;以及

抢占性设置模块,用于根据所述队列的优先级别设置所述队列的抢占性,

其中,所述实时判断模块包括:

用户权限判断模块,用于判断作业相关的用户是否具有提交实时作业的权限;以及

优先级别判断模块,用于判断所述作业的优先级别是否是高优先级,如果是,则所述作

业为实时作业,否则为非实时作业,

其中,所述优先级别设置模块根据所述集群规模信息设置所述队列的优先级别,包括:

如果所述集群规模小于集群规模的阈值,则将所述队列设置成两种优先级别,包括高优先级和低优先级;

如果所述集群规模不小于集群规模的阈值,则将所述队列设置成三种优先级别,包括高优先级、中优先级和低优先级。

4.根据权利要求3所述的系统,其特征在于,所述抢占性设置模块根据所述队列的优先级别设置所述队列的抢占性,包括:

如果所述队列设置成两种优先级别,则将所述高优先级的队列设置成抢占者,所述低优先级的队列设置成被抢占者;

如果所述队列设置成三种优先级别,则将所述高优先级的队列设置成抢占者,所述中优先级的队列和所述低优先级的队列设置成被抢占者。

实时作业调度系统及方法

技术领域

[0001] 本发明涉及计算机技术高性能领域,特别涉及一种作业调度系统及方法。

背景技术

[0002] 目前,高性能服务器已经被广泛应用在科研、政府和国防部门,而且高性能服务器上的应用软件也越来越普及。多种应用软件的同时运行将会导致资源冲突,影响某些软件的性能,例如,实时软件对于相应时间有着非常强的要求,在某些情况下,超过截止时间的实时任务的结果将是无效结果,如气象软件,超过规定时间的气象数据将被视为无效数据。

[0003] 现有的作业调度系统采用公平的调度策略,没有引入优先级的概念,无法保证对反应速度具有较高要求的实时软件的快速运行。因此需要一种系统和方法解决上述问题。

发明内容

[0004] 针对相关技术中存在的一个或多个问题,本发明的目的在于提供一种实时作业调度系统及方法,以解决上述问题中的至少之一。

[0005] 为实现上述目的,根据本发明的一个方面,提供了一种实时作业调度方法,包括以下步骤:创建多个队列。将多个队列设置成不同的优先级别,并根据优先级别设置队列的抢占性。判断是否存在实时作业,如果不存在,则继续执行现有作业;如果存在,则将现有作业挂起,并释放掉现有作业申请的资源,然后将实时作业运行到服务器上。在实时作业运行结束后,再将挂起的现有作业恢复到服务器上继续运行。

[0006] 优选地,采用自适应算法设置队列的优先级别和抢占性,包括以下步骤:收集集群信息,并查找集群信息中关于集群规模的信息。根据集群规模设置队列的优先级别和抢占性。其中,如果集群规模小于集群规模的阈值,则将队列设置成两种优先级别,包括高优先级和低优先级,高优先级队列的作业优先调度并抢占低优先级队列的作业的资源;如果集群规模不小于集群规模的阈值,则将队列设置成三种优先级别,包括高优先级、中优先级和低优先级,高优先级队列的作业优先调度并抢占中优先级队列的作业的资源和低优先级队列的作业的资源,中优先级队列的作业比低优先级队列的作业优先调度。其中,集群规模的阈值为32。

[0007] 优选地,判断是否存在实时作业包括以下步骤:

[0008] A. 判断作业相关的用户是否具有提交实时作业的权限,如果具有权限,则执行步骤B,否则判定作业为非实时作业。

[0009] B. 判断作业的级别是否为高优先级,如果是高优先级,则作业为实时作业,否则为非实时作业。

[0010] 根据本发明的另一个方面,还提出一种实时作业调度系统,包括队列设置模块、实时判断模块和挂起模块。其中,队列设置模块用于将多个队列设置成不同的优先级别,并根据优先级别设置队列的抢占性。实时判断模块用于判断是否存在实时作业。挂起模块用于当存在实时作业时,将现有作业挂起,并释放现有作业的资源,以及在实时作业运行结束后

恢复现有作业的运行。

[0011] 优选地,队列设置模块包括信息获取模块、优先级别设置模块和抢占性设置模块。其中,信息获取模块用于收集集群信息,并从集群信息中获取集群规模信息。优先级别设置模块用于根据集群规模信息设置队列的优先级别。如果集群规模小于集群规模的阈值,则将队列设置成两种优先级别,包括高优先级和低优先级;如果集群规模不小于集群规模的阈值,则将队列设置成三种优先级别,包括高优先级、中优先级和低优先级。抢占性设置模块用于根据队列的优先级别设置队列的抢占性。如果队列被设置成两种优先级,则将高优先级的队列设置成抢占者,低优先级的队列设置成被抢占者;如果队列被设置成三种优先级,则将高优先级的队列设置成抢占者,将中优先级的队列和低优先级的队列设置成被抢占者。

[0012] 优选地,实时判断模块包括用户权限判断模块和优先级别判断模块。其中,用户权限判断模块用于判断作业相关的用户是否具有提交实时作业的权限。优先级别判断模块用于判断作业的优先级别是否是高优先级,如果是,则判断作业为实时作业,否则为非实时作业。

[0013] 借助本发明上述至少一个技术方案,通过采用优先调度机制和抢占机制,实现了应用软件的实时性。

附图说明

[0014] 本发明上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解,其中:

[0015] 图1为根据本发明的实时作业调度方法的流程图;以及

[0016] 图2为根据本发明的实时作业调度系统的结构图。

具体实施方式

[0017] 考虑到相关技术中存在的问题,本发明提供一种实时作业调度系统及方法。下面详细描述本发明的实施例,所述实施例的示例在附图中示出。应理解,下面通过参考附图描述的实施例是示例性的,仅用于解释本发明,而不能解释为对本发明的限制。

[0018] 图1为根据本发明的实时作业调度方法的流程图,如图1所示,该方法包括以下步骤:

[0019] 步骤S110,创建多个队列。

[0020] 步骤S120,将多个队列设置成不同的优先级别,并根据优先级别设置队列的抢占性。

[0021] 具体地,在本发明实施例中,采用自适应算法设置队列的优先级别和抢占性。当然,本领域技术人员还能够根据下述方案提出其他修改或变化,这些修改或变化均应包含在本发明的包含范围之内。

[0022] 首先,收集集群信息,并从集群信息中获取集群规模信息。然后,根据集群规模信息设置队列的优先级别和抢占性。如果集群规模小于集群规模的阈值,则将队列设置成两种优先级别,包括高优先级和低优先级,高优先级队列的作业优先调度并抢占低优先级队列的作业的资源;如果集群规模不小于集群规模的阈值,则将队列设置成三种优先级别,包

括高优先级、中优先级和低优先级，高优先级队列的作业优先调度并抢占中优先级队列的作业的资源 and 低优先级队列的作业的资源，中优先级队列的作业比低优先级队列的作业优先调度。其中，集群规模的阈值可随机设置，在本实施例中，设为32。

[0023] 步骤S130，判断是否存在实时作业，如果存在，则执行步骤S140，否则继续执行现有作业。

[0024] 具体地，对于如何判断是否存在实时作业，本发明提出的可能的判断方案如下，当然本领域技术人员还能够根据下述方案提出其他修改或变化，这些修改或变化均应包含在本发明的包含范围之内。

[0025] 首先，检查作业相关的用户的身份，判断用户是否具有提交实时作业的权限，如果没有，则判断该作业为非实时作业，否则，继续察看作业的级别，如果该作业的级别为高优先级，则判断该作业为实时作业，否则为非实时作业。

[0026] 步骤S140，将现有作业挂起，释放掉现有作业申请的资源。

[0027] 步骤S150，将实时作业运行到服务器上。

[0028] 步骤S160，实时作业运行结束后，将被挂起的现有作业恢复到服务器上继续运行。

[0029] 下面结合附图2详细描述本发明的实时作业调度系统，该系统包括队列设置模块100、实时判断模块200和挂起模块300。其中，队列设置模块100用于将多个队列设置成不同的优先级别，并根据优先级别设置队列的抢占性。实时判断模块200用于判断是否存在实时作业。挂起模块300用于当存在实时作业时，将现有作业挂起，并释放现有作业的资源，以及在实时作业运行结束后恢复现有作业的运行。

[0030] 队列设置模块100包括信息获取模块110、优先级别设置模块120和抢占性设置模块130。其中，信息获取模块110用于收集集群信息，并从集群信息中获取集群规模信息。优先级别设置模块120用于根据集群规模信息设置队列的优先级别。如果集群规模小于集群规模的阈值，则将队列设置成两种优先级别，包括高优先级和低优先级；如果集群规模不小于集群规模的阈值，则将队列设置成三种优先级别，包括高优先级、中优先级和低优先级。抢占性设置模块130用于根据队列的优先级别设置队列的抢占性。如果队列被设置成两种优先级，则高优先级队列的作业优先调度并抢占低优先级队列的作业的资源；如果队列被设置成三种优先级，则高优先级队列的作业优先调度并抢占中优先级队列的作业的资源和低优先级队列的作业的资源，中优先级队列的作业比低优先级队列的作业优先调度。

[0031] 实时判断模块200包括用户权限判断模块210和优先级别判断模块220。其中，用户权限判断模块210用于判断作业相关的用户是否具有提交实时作业的权限。优先级别判断模块220用于判断作业的优先级别是否是高优先级，如果是，则判断作业为实时作业，否则为非实时作业。

[0032] 应理解，上述实施例仅是示意性的实施例，并不限制本发明仅能够通过上述实施例实现。本领域的普通技术人员还能够根据上述方案提出其他修改或变化，例如改变队列的优先级别的设置方法，这些修改或变化均应包含在本发明的包含范围之内。

[0033] 本发明通过采用优先调度机制和抢占机制，在作业层次上实现了应用程序的实时性，用户对作业的控制更方便。并且，在负载较大的高性能服务器上，本发明的性能可以不受影响。除此之外，将本发明结合作业调度系统的其他功能，可提高高性能服务器的吞吐量。

[0034] 尽管已经示出和描述了本发明的实施例,对于本领域的普通技术人员而言,可以理解在不脱离本发明的原理和精神的情况下可以对这些实施例进行多种变化、修改、替换和变型,本发明的范围由所附权利要求及其等同限定。

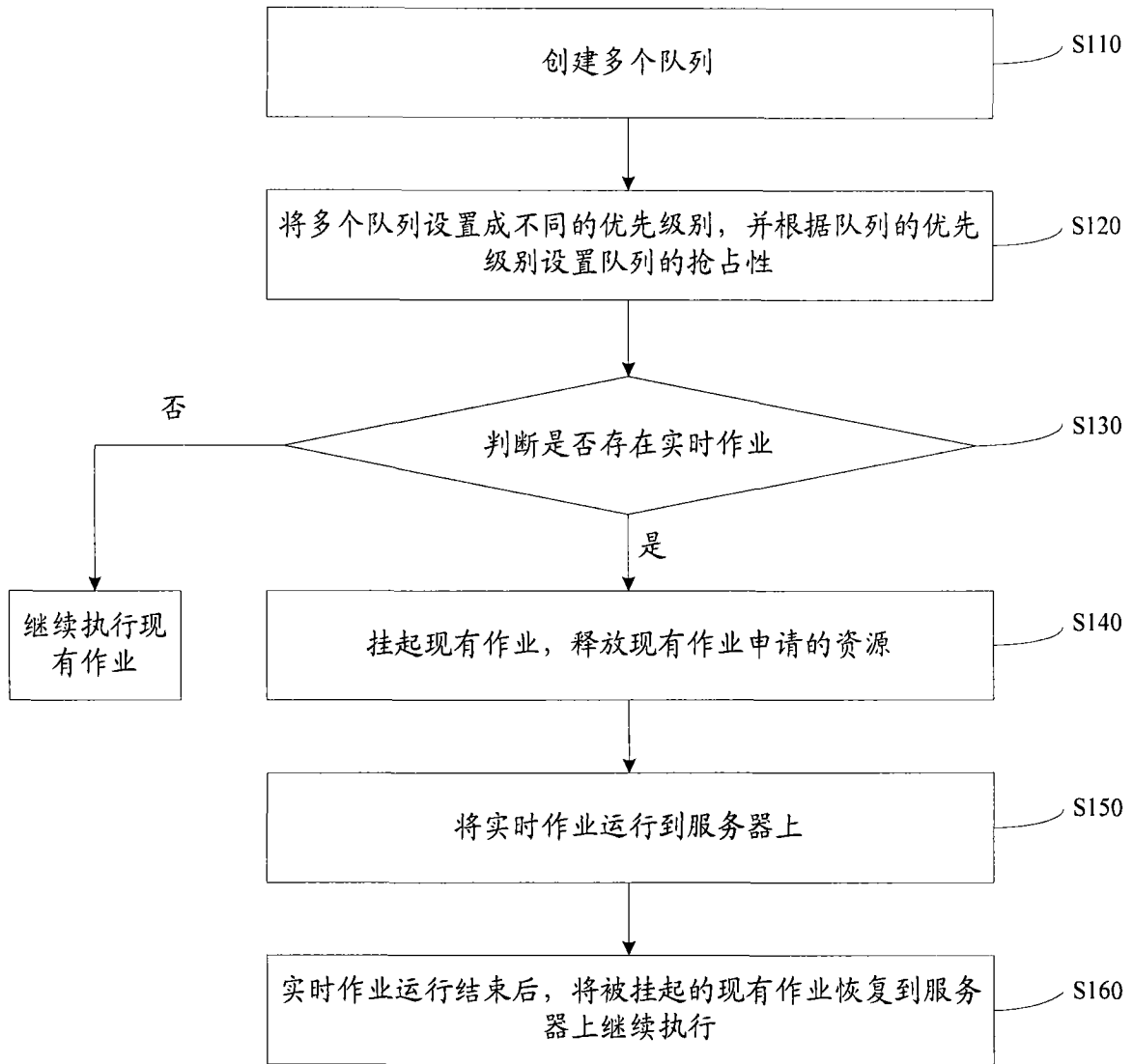


图1

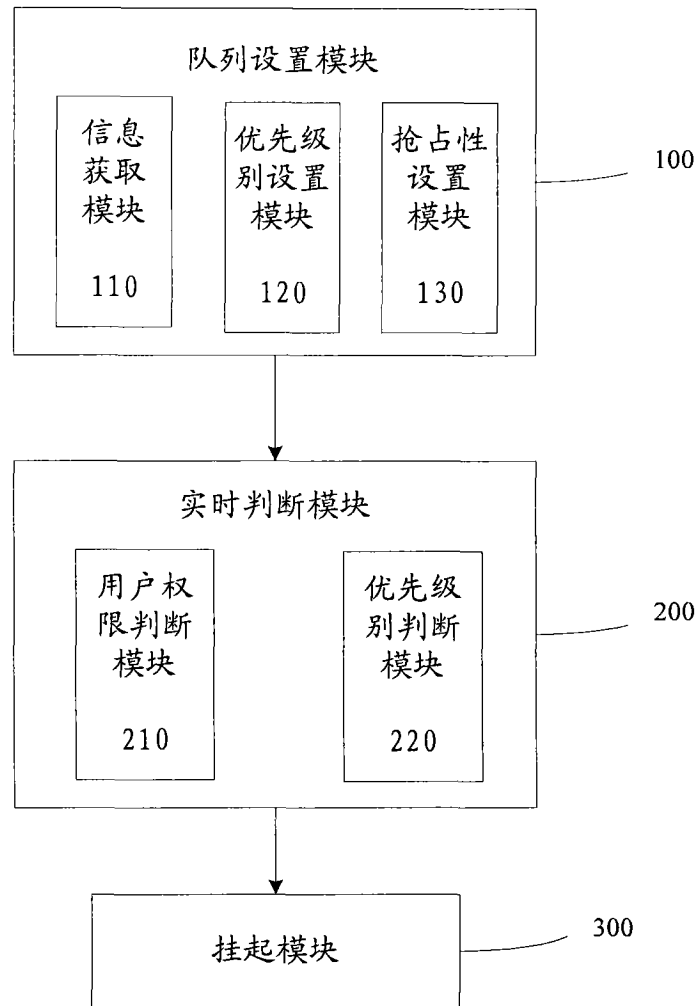


图2