



(12)发明专利

(10)授权公告号 CN 103838516 B

(45)授权公告日 2016.09.28

(21)申请号 201210484576.0

(22)申请日 2012.11.23

(65)同一申请的已公布的文献号
申请公布号 CN 103838516 A

(43)申请公布日 2014.06.04

(73)专利权人 中国科学院声学研究所
地址 100190 北京市海淀区北四环西路21号

专利权人 北京中科华影传媒技术有限公司

(72)发明人 王劲林 查奇文 郭秀岩 张武

(74)专利代理机构 北京法思腾知识产权代理有限公司 11318

代理人 杨小蓉 杨青

(51)Int.Cl.
G06F 3/06(2006.01)

(56)对比文件

- CN 101546601 A, 2009.09.30,
- CN 101566923 A, 2009.10.28,
- US 2010/0284411 A1, 2010.11.11,
- CN 101119374 A, 2008.02.06,

审查员 王丽娜

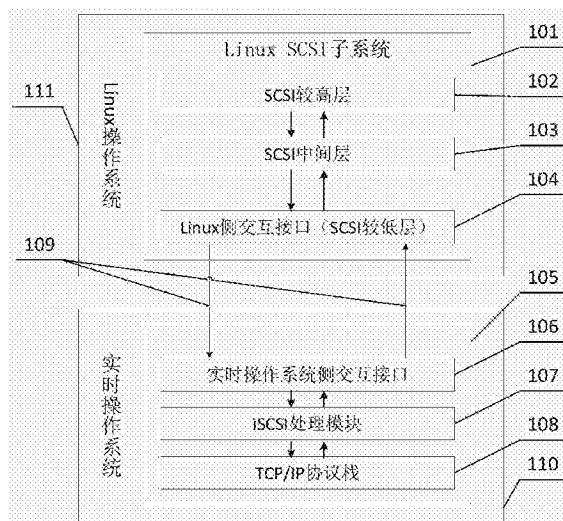
权利要求书3页 说明书7页 附图1页

(54)发明名称

一种用于多核处理器高效访问iSCSI磁盘阵列的方法及系统

(57)摘要

本发明公开了一种用于多核处理器高效访问iSCSI磁盘阵列的方法及系统,所述的方法包含:步骤101)在互不相交的处理器核心上分别运行Linux操作系统和实时操作系统步骤102)首先, Linux操作系统接收位于其上的应用层的请求信息,并将请求信息封装成SCSI命令字;然后, Linux操作系统指定读取或写入数据的设备LUN信息,设置用于建立Linux操作系统和实时操作系统之间的数据通道和操作通道的核间消息;最后, Linux操作系统将包含SCSI命令字的核间消息发送至实时操作系统;步骤103)实时操作系统接收核间消息,并将核间消息包含的SCSI命令采用iSCSI协议进行处理,再将处理后的得到的iSCSI数据包提交给实时操作系统上的TCP/IP协议栈,由TCP/IP协议栈完成iSCSI磁盘阵列数据的发送和接收。



1. 一种用于多核处理器高效访问iSCSI磁盘阵列的系统,所述系统包含:

运行于互不相交的处理器核心上的分别独立管理不同物理内存空间的Linux操作系统和实时操作系统;

所述Linux操作系统包含:上层应用层和Linux SCSI子系统;且所述Linux SCSI子系统从上向下依次包含:

SCSI较高层,该层用于接受来自上层应用层的请求并将其转换成SCSI请求,完成SCSI命令并将状态信息通知上层应用层;

SCSI中间层,该层用于提交SCSI命令;

Linux侧交互接口,用于实现Linux SCSI子系统较低层的功能,所述较低层功能包含:接收SCSI中间层的SCSI命令,将SCSI命令封装成多核核间消息,发送给实时操作系统;并负责与实时操作系统侧交互接口进行多核核间消息交互;

所述实时操作系统从上向下依次包含:

实时操作系统侧交互接口,用于与Linux侧交互接口进行多核核间消息交互;

iSCSI处理模块,用于负责iSCSI协议的处理;

TCP/IP协议栈,用于负责实现iSCSI磁盘阵列数据的收发;

其中,所述的多核核间消息用于建立Linux操作系统和实时操作系统之间的数据通道和操作通道。

2. 根据权利要求1所述的用于多核处理器高效访问iSCSI磁盘阵列的系统,其特征在于,所述实时操作系统侧交互接口进一步包含:

接收模块,用于接收来自Linux侧交互接口的多核核间消息;和

解析和调用模块,用于解析出SCSI中间层提交的SCSI命令,并调用iSCSI处理模块进行处理。

3. 根据权利要求1所述的用于多核处理器高效访问iSCSI磁盘阵列的系统,其特征在于,所述TCP/IP协议栈用于实现TCP/IP协议族中2-4层协议的处理;且该协议栈提供接口供上层应用使用,iSCSI处理模块通过这些接口实现数据的收发。

4. 一种基于权利要求1-3任一项所述系统的用于多核处理器高效访问iSCSI磁盘阵列的方法,该方法基于位于多核处理器上的异构操作系统,所述的方法包含:

步骤101)首先,Linux操作系统接收位于其上的应用层的请求信息,并将请求信息封装成SCSI命令;然后,Linux操作系统指定读取或写入数据的设备LUN信息,设置用于建立Linux操作系统和实时操作系统之间的数据通道和操作通道的多核核间消息;最后,Linux操作系统将包含SCSI命令的多核核间消息发送至实时操作系统;

步骤102)实时操作系统接收多核核间消息,并将多核核间消息包含的SCSI命令采用iSCSI协议进行处理,再将处理后的得到的iSCSI数据包提交给实时操作系统上的TCP/IP协议栈,由TCP/IP协议栈完成iSCSI磁盘阵列数据的发送和接收。

5. 根据权利要求4所述的用于多核处理器高效访问iSCSI磁盘阵列的方法,其特征在于,所述多核核间消息包含七个字段,具体格式为:

第一个字段为发送消息的核标识,第二个字段为接收消息的核标识,第三个字段为SCSI命令,第四个字段为SCSI命令参数,第五个字段为操作执行返回结果,第六个字段为数据块指针,第七个字段为预留的状态位;

所述发送消息的核标识,用于指定由多核处理器的哪个核发送上层应用的读或写的请求消息;

所述接收消息的核标识,用于指定由多核处理器的哪个核接收多核核间消息;

所述SCSI命令参数,用于指定SCSI命令所需要的参数;

所述操作执行返回结果,用于表示实时操作系统解析出SCSI命令后调用iSCSI协议处理后的返回值。

6.根据权利要求5所述的用于多核处理器高效访问iSCSI磁盘阵列的方法,其特征在于,当上层应用发起读取iSCSI磁盘阵列数据操作时,所述步骤101)进一步包含:

步骤101-1)Linux操作系统上的应用程序提交磁盘读命令,经过Linux操作系统的SCSI子系统中SCSI较高层和SCSI中间层的处理后到达Linux侧交互接口;

步骤101-2)Linux侧交互接口将磁盘读命令封装成SCSI命令,指定读取数据的设备LUN和读取数据的长度;

步骤101-3)Linux侧交互接口填写多核核间消息,根据Linux侧交互接口所在核填写发送消息核标识,根据实时操作系统上的实时操作系统侧交互接口所在核填写接收消息核标识,根据操作类型填写SCSI命令和SCSI命令参数,清除操作执行返回结果,申请存放读取结果的内存块,并将内存块的物理地址填写到内存块指针中,完成上述操作后将多核核间消息发送给实时操作系统上的实时操作系统侧交互接口。

7.根据权利要求6所述的用于多核处理器高效访问iSCSI磁盘阵列的方法,其特征在于,所述步骤102)进一步包含:

步骤102-1)实时操作系统上的实时操作系统侧交互接口接收多核核间消息,并将SCSI命令和SCSI命令参数传递给实时操作系统上的iSCSI处理模块;

步骤102-2)实时操作系统上的iSCSI处理模块完成iSCSI协议处理后将iSCSI数据包提交给实时操作系统上的TCP/IP协议栈;

步骤102-4)实时操作系统上的TCP/IP协议栈完成iSCSI数据的发送,并等待iSCSI磁盘阵列的返回结果;

步骤102-5)实时操作系统上的TCP/IP协议栈接收iSCSI磁盘阵列的数据,并将数据返回给实时操作系统上的iSCSI处理模块,实时操作系统上的iSCSI处理模块根据多核核间消息中的数据块指针将数据写入数据块中,并将返回结果返回给实时操作系统上的实时操作系统侧交互接口;

步骤102-6)实时操作系统上的实时操作系统侧交互接口根据实时操作系统上的实时操作系统侧交互接口填写多核核间消息,根据Linux侧交互接口所在核填写接收消息核标识,根据实时操作系统上的实时操作系统侧交互接口所在核填写发送消息核标识,填写操作执行返回结果,完成上述操作后将多核核间消息发送给Linux侧交互接口;

步骤102-7)Linux侧交互接口解析核间消息,将多核核间消息中的返回结果和数据块8中的数据返回给Linux操作系统中的应用程序。

8.根据权利要求5所述的用于多核处理器高效访问iSCSI磁盘阵列的方法,其特征在于,当上层应用发起写入iSCSI磁盘阵列数据操作时,所述步骤101)进一步包含:

步骤101-a)Linux操作系统上应用程序提交磁盘写命令,经过Linux操作系统的SCSI子系统中SCSI较高层和SCSI中间层的处理后到达Linux侧交互接口;

步骤101-b)Linux侧交互接口将磁盘写命令封装成SCSI命令,指定写入数据的设备LUN;

步骤101-c)Linux侧交互接口填写多核核间消息,根据Linux侧交互接口所在核填写发送消息核标识,根据实时操作系统上的实时操作系统侧交互接口所在核填写接收消息核标识,根据操作类型填写SCSI命令和SCSI命令参数,清除操作执行返回结果,申请存放写入数据的内存块,并将内存块的物理地址填写到内存块指针中,完成上述操作后将多核核间消息发送给实时操作系统上的实时操作系统侧交互接口。

9.根据权利要求8所述的用于多核处理器高效访问iSCSI磁盘阵列的方法,其特征在于,所述步骤102)进一步包含:

步骤102-a)实时操作系统上的实时操作系统侧交互接口接收多核核间消息,并将SCSI命令和SCSI命令参数传递给实时操作系统上的iSCSI处理模块;

步骤102-b)实时操作系统上的iSCSI处理模块完成iSCSI协议处理后将iSCSI数据包提交给实时操作系统上的TCP/IP协议栈;

步骤102-c)实时操作系统上的TCP/IP协议栈完成iSCSI数据的发送,并等待iSCSI磁盘阵列的返回结果;

步骤102-d)实时操作系统上的TCP/IP协议栈接收iSCSI磁盘阵列的返回结果,并将返回结果返回给实时操作系统上的iSCSI处理模块,实时操作系统上的iSCSI处理模块处理返回结果,并将返回结果返回给实时操作系统上的实时操作系统侧交互接口;

步骤102-e)实时操作系统上的实时操作系统侧交互接口根据实时操作系统上的实时操作系统侧交互接口填写多核核间消息,根据Linux侧交互接口所在核填写接收消息核标识,根据实时操作系统上的实时操作系统侧交互接口所在核填写发送消息核标识,填写操作执行返回结果,完成上述操作后将多核核间消息发送给Linux侧交互接口;

步骤102-f)Linux侧交互接口解析多核核间消息,将多核核间消息中的返回结果返回给Linux操作系统中的应用程序。

一种用于多核处理器高效访问iSCSI磁盘阵列的方法及系统

技术领域

[0001] 本发明涉及一种基于多核处理器和异构操作系统的iSCSI发起端实现方法,具体涉及一种用于多核处理器高效访问iSCSI磁盘阵列的方法及系统。

背景技术

[0002] 随着计算机技术和网络技术的不断发展,传统单核处理器在功耗、并发、散热等方面暴露出的问题日益严重,已经不能满足各类网络应用对处理器性的迫切需求,因此具有较低功耗、更高并发性能的多核处理器被更多的用于计算机网络应用中。由于多核处理器由物理并行的多个处理核心构成,因此相对传统的单核处理器就有更强的并发性,同时低频率的处理核心可以更加简单的集成到同一颗处理器中,并能较好的控制功耗发热等问题,因此多核处理器相对传统通用处理器的优势正在逐渐扩大。所以,计算机网络技术中开始广泛的使用多核处理器来搭建各类高性能的网络应用。

[0003] 在各类网络视频应用中,磁盘阵列设备都是资源存储核心的重要组成部分,尤其是iSCSI磁盘阵列以其使用简单、性能突出等特性得到更加广泛的应用,所以在基于多核处理器为计算中心、基于iSCSI磁盘阵列为存储中心的新型视频类应用系统中,如何合理的结合多核处理器和iSCSI磁盘阵列,并有效的提高系统的整体性能,是增强网络视频应用系统服务质量的关键之一。现有技术对iSCSI磁盘阵列的访问需要根据SCSI子系统生产的SCSI命令字,并通过iSCSI协议进行封装,最终通过TCP/IP协议栈实现iSCSI Host端和iSCSI Target端的高效数据交互,而操作系统上的SCSI子系统由多个组成部分,因此SCSI命令执行流程冗余繁杂,会造成多核处理器处理资源的浪费,而多核处理器的实时执行环境执行效率较高,但是没有操作系统中各种功能的支持,无法直接实现SCSI命令、iSCSI协议的处理和执行,因此需要有效的整合多核处理器上Linux操作系统和实时执行环境的优势和特点,提高SCSI命令的处理效率。

发明内容

[0004] 本发明的目的在于,为克服上述问题,本发明提供一种用于多核处理器高效访问iSCSI磁盘阵列的方法及系统。

[0005] 为实现上述目的,本发明提供了一种用于多核处理器高效访问iSCSI磁盘阵列的系统,所述系统包含:

[0006] 运行于互不相交的处理器核心上的分别独立管理不同内存空间的Linux操作系统和实时操作系统;

[0007] 所述Linux操作系统包含:上层应用层和Linux SCSI子系统;且所述Linux SCSI子系统从上向下依次包含:

[0008] SCSI较高层,该层用于接受来自其上层的请求并将其转换成SCSI请求,完成SCSI命令并将状态信息通知上层;

[0009] SCSI中间层,该层用于提交SCSI命令;

[0010] Linux侧交互接口,用于实现Linux SCSI子系统较低层的功能,所述较低层功能包含:接收中间层的SCSI命令字,以及将iSCSI子系统的处理结果返回给中间层,并负责与实时操作系统侧交互接口进行核间信息交互;

[0011] 所述实时操作系统从上向下依次包含:

[0012] 实时操作系统侧交互接口,用于与Linux侧交互接口进行核间信息交互;

[0013] iSCSI处理模块,用于负责iSCSI协议的处理;

[0014] TCP/IP协议栈,用于负责实现iSCSI磁盘阵列数据的收发;

[0015] 其中,所述的核间信息用于建立Linux操作系统和实时操作系统之间的数据通道和操作通道。

[0016] 上述实时操作系统侧交互接口进一步包含:

[0017] 接收模块,用于接收来自Linux侧交互接口的多核核间消息;和

[0018] 解析和调用模块,用于解析出SCSI中间层提交的SCSI命令,并调用iSCSI处理模块进行处理。

[0019] 上述TCP/IP协议栈用于实现TCP/IP协议族中2-4层协议的处理;且该协议栈提供接口供上层应用使用,iSCSI子系统通过这些接口实现数据的收发。

[0020] 基于上述系统,本发明提供了一种用于多核处理器高效访问iSCSI磁盘阵列的方法,该方法基于位于多核处理器上的异构操作系统,所述的方法包含:

[0021] 步骤101)首先,Linux操作系统接收位于其上的应用层的请求信息,并将请求信息封装成SCSI命令字;然后,Linux操作系统指定读取或写入数据的设备LUN信息,设置用于建立Linux操作系统和实时操作系统之间的数据通道和操作通道的核间消息;最后,Linux操作系统将包含SCSI命令字的核间消息发送至实时操作系统;

[0022] 步骤102)实时操作系统接收核间消息,并将核间消息包含的SCSI命令采用iSCSI协议进行处理,再将处理后的得到的iSCSI数据包提交给实时操作系统上的TCP/IP协议栈,由TCP/IP协议栈完成iSCSI磁盘阵列数据的发送和接收。

[0023] 上述核间消息包含七个字段,具体格式为:

[0024] 第一个字段为发送消息的核标识,第二个字段为接收消息的核标识,第三个字段为SCSI命令,第四个字段为SCSI命令参数,第五个字段为操作执行返回结果,第六个字段为数据块指针,第七个字段为预留的状态位;

[0025] 所述发送消息的核标识,用于指定由多核处理器的哪个核发送上层应用的读或写的请求消息;

[0026] 所述接收消息的核标识,用于指定由多核处理器的哪个核接收核间消息;

[0027] 所述SCSI命令参数,用于指定SCSI命令所需要的参数;

[0028] 所述操作执行返回结果,用于表示实时操作系统解析出SCSI命令后调用iSCSI协议处理后的返回值。

[0029] 当上层应用发起读取iSCSI磁盘阵列数据操作时,所述步骤101)进一步包含:

[0030] 步骤101-1)Linux操作系统上的应用程序提交磁盘读命令,经过Linux操作系统的SCSI子系统中SCSI较高层和SCSI中间层的处理后到达Linux侧交互接口;

[0031] 步骤101-2)Linux侧交互接口将磁盘读命令封装成SCSI命令字,指定读取数据的设备LUN和读取数据的长度;

[0032] 步骤101-3)Linux侧交互接口填写核间消息,根据Linux侧交互接口所在核填写发送消息核标示,根据实时操作系统上的实时操作系统侧交互接口所在核填写接收消息核标示,根据操作类型填写SCSI命令和SCSI命令参数,清除操作执行返回结果,申请存放读取结果的内存块,并将内存块的物理地址填写到内存块指针中,完成上述操作后将核间消息发送给实时操作系统上的实时操作系统侧交互接口。

[0033] 上述步骤102)进一步包含:

[0034] 步骤102-1)实时操作系统上的实时操作系统侧交互接口接收核间消息,并将SCSI命令和SCSI命令参数传递给实时操作系统上的iSCSI处理模块;

[0035] 步骤102-2)实时操作系统上的iSCSI处理模块完成iSCSI协议处理后将iSCSI数据包提交给实时操作系统上的TCP/IP协议栈;

[0036] 步骤102-4)实时操作系统上的TCP/IP协议栈完成iSCSI数据的发送,并等待iSCSI磁盘阵列的返回结果;

[0037] 步骤102-5)实时操作系统上的TCP/IP协议栈接收iSCSI磁盘阵列的数据,并将数据返回给实时操作系统上的iSCSI处理模块,实时操作系统上的iSCSI处理模块根据核间消息中的数据块指针将数据写入数据块中,并将返回结果返回给实时操作系统上的实时操作系统侧交互接口;

[0038] 步骤102-6)实时操作系统上的实时操作系统侧交互接口根据实时操作系统上的实时操作系统侧交互接口填写核间消息,根据Linux侧交互接口所在核填写接收消息核标示,根据实时操作系统上的实时操作系统侧交互接口所在核填写发送消息核标识,填写操作执行返回结果,完成上述操作后将核间消息发送给Linux侧交互接口104;

[0039] 步骤102-7)Linux侧交互接口解析核间消息,将核间消息中的返回结果和数据块8中的数据返回给Linux操作系统中的应用程序。

[0040] 当上层应用发起写入iSCSI磁盘阵列数据操作时,所述步骤101)进一步包含:

[0041] 步骤101-a)Linux操作系统上应用程序提交磁盘写命令,经过Linux操作系统的SCSI子系统中SCSI较高层和SCSI中间层的处理后到达Linux侧交互接口;

[0042] 步骤101-b)Linux侧交互接口将磁盘写命令封装成SCSI命令字,指定写入数据的设备LUN;

[0043] 步骤101-c)Linux侧交互接口填写核间消息,根据Linux侧交互接口所在核填写发送消息核标识,根据实时操作系统上的实时操作系统侧交互接口所在核填写接收消息核标示,根据操作类型填写SCSI命令和SCSI命令参数,清除操作执行返回结果,申请存放写入数据的内存块,并将内存块的物理地址填写到内存块指针中,完成上述操作后将核间消息发送给实时操作系统上的实时操作系统侧交互接口106;

[0044] 上述步骤102)进一步包含:

[0045] 步骤102-a)实时操作系统上的实时操作系统侧交互接口接收核间消息,并将SCSI命令和SCSI命令参数传递给实时操作系统上的iSCSI处理模块;

[0046] 步骤102-b)实时操作系统上的iSCSI处理模块完成iSCSI协议处理后将iSCSI数据包提交给实时操作系统上的TCP/IP协议栈;

[0047] 步骤102-c)实时操作系统上的TCP/IP协议栈完成iSCSI数据的发送,并等待iSCSI磁盘阵列的返回结果;

[0048] 步骤102-d)实时操作系统上的TCP/IP协议栈接收iSCSI磁盘阵列的返回结果,并将返回结果返回给实时操作系统上的iSCSI处理模块,实时操作系统上的iSCSI处理模块处理返回结果,并将返回结果返回给实时操作系统上的实时操作系统侧交互接口;

[0049] 步骤102-e)实时操作系统上的实时操作系统侧交互接口根据实时操作系统上的实时操作系统侧交互接口填写核间消息,根据Linux侧交互接口所在核填写接收消息核标识,根据实时操作系统上的实时操作系统侧交互接口所在核填写发送消息核标识,填写操作执行返回结果,完成上述操作后将核间消息发送给Linux侧交互接口;

[0050] 步骤102-f)Linux侧交互接口解析核间消息,将核间消息中的返回结果返回给Linux操作系统中的应用程序。

[0051] 与现有技术相比,本发明的技术优势在于:

[0052] 本发明提供了一种基于多核处理器和异构操作系统的iSCSI发起端实现方法,通过该方法实现iSCSI协议栈的卸载,并保持对上层应用兼容的SCSI调用接口,以实现高效快捷的iSCSI磁盘阵列操作和iSCSI磁盘阵列数据处理。

附图说明

[0053] 图1是本发明基于多核处理器和异构操作系统的iSCSI发起端实现方法结构图。

[0054] 图2是本发明基于多核处理器和异构操作系统的iSCSI发起端实现方法中各个模块交互时使用的消息体结构图。

具体实施方式

[0055] 下面结合附图和实施例对本发明进一步说明。

[0056] 本发明提供了一种基于多核处理器和异构操作系统的iSCSI发起端实现方法,由Linux侧交互接口、实时操作系统侧交互接口、iSCSI处理模块和TCP/IP协议栈完成iSCSI发起端的全部处理工作,操作系统之间和各个模块之间使用核间消息通信。其特征是:Linux侧交互接口运行于Linux操作系统上,实现Linux SCSI子系统较低层的功能,并负责与实时操作系统交互接口进行交互;实时操作系统交互接口运行于实时操作系统上,负责与Linux侧交互接口交互;iSCSI处理模块运行于实时操作系统上,负责iSCSI协议的处理;TCP/IP协议栈运行于实时操作系统上,负责实现iSCSI数据的收发。

[0057] Linux侧交互接口实现Linux SCSI子系统较低层的功能,接收Linux SCSI子系统中间层的SCSI命令,并将SCSI命令封装成多核核间消息,发送给实时操作系统侧交互接口。

[0058] 实时操作系统侧交互接口运行于实时操作系统上,负责接收来自Linux侧交互接口的多核核间消息,解析出Linux SCSI子系统中间层提交的SCSI命令,并调用iSCSI处理模块进行处理。

[0059] iSCSI处理模块运行于实时操作系统上,负责iSCSI协议的处理,并调用实时操作系统的TCP/IP协议栈与iSCSI target进行网络通信。

[0060] TCP/IP协议栈运行于实时操作系统上,实现了现有技术中Linux操作系统上TCP/IP协议栈的全部功能,并使用核间通讯机制和多核处理器提供的TCP/IP加速功能实现高速的iSCSI数据收发功能。

[0061] Linux侧交互接口和实时操作系统侧交互接口通过多核核间消息建立数据和操作

通道;消息中包含七个字段,第一个字段为发送消息的核标识,第二个字段为接收消息的核标识,第三个字段为SCSI命令,第四个字段为SCSI命令参数,第五个字段为操作执行返回结果,第六个字段为数据块指针,第七个字段为预留的状态位。发送消息的核标识用于指定由多核处理器的哪个核发送该消息。接收消息的核标识用于指定由多核处理器的哪个核接收该消息。

[0062] SCSI命令是Linux SCSI子系统中间层向下提交给Linux侧交互接口的,主要包括以下命令:Test unit ready、Inquiry、Request sense、Read capacity、Read和Write。

[0063] 各个命令所需要的参数具体如下:

[0064] Test unit ready命令包括需要检查状态设备的LUN标志;

[0065] Inquiry命令包括需要查询摘要信息设备的LUN标志;

[0066] Request sense命令包括需要查询sense数据设备的LUN标志;

[0067] Read capacity命令包括需要查询最大容量设备的LUN标志;

[0068] Read命令包括读的数据长度,读数据存储的内存的物理地址以及所读数据在磁盘上的位置信息;

[0069] Write命令包括写的数据长度,写数据存储的内存的物理地址以及所写数据在磁盘上的位置信息;

[0070] 各个命令的返回结果具体如下:

[0071] Test unit ready命令的执行结果,包括状态正常和错误报告;

[0072] Inquiry命令的执行结果,包括SCSI标志查询数据结构体;

[0073] Request sense命令的执行结果,包括SCSI sense数据结构体;

[0074] Read capacity命令的执行结果,包括查询设备最大容量信息的结构体;

[0075] Read命令的执行结果,包括执行结果和数据地址、长度的结构体;

[0076] Write命令的执行结果,包括执行的结果和写入长度的结构体。

[0077] 实施例

[0078] 实施例选用Cavium Networks公司的OCTEON系列处理器平台。使用Linux操作系统111运行Linux侧交互接口104,使用Cavium Networks公司提供的简单执行环境SE(Simple Executive)作为本发明中的实时操作系统110运行实时操作系统侧交互接口106、iSCSI处理模块107和TCP/IP协议栈108。

[0079] Linux操作系统111和实时操作系统110所运行的处理核心互不相交,且这2个操作系统分别管理独立、不重叠的内存空间。

[0080] 运行于Linux操作系统111上的Linux侧交互接口104完成SCSI子系统中较低层的功能,并将SCSI命令字传递给实时操作系统110;实时操作系统110上的实时操作系统侧交互接口106完成核间消息的交互功能;实时操作系统110上的iSCSI处理模块107实现iSCSI协议栈的卸载和实际处理工作;实时操作系统110上的TCP/IP协议栈108实现iSCSI数据的高速收发。

[0081] Linux操作系统111上应用程序发起iSCSI磁盘阵列读取操作的步骤是:

[0082] Linux操作系统111上应用程序提交磁盘Read命令,经过Linux操作系统111的SCSI子系统101中SCSI较高层102和SCSI中间层103的处理后到达Linux侧交互接口104。

[0083] Linux侧交互接口104将磁盘Read命令封装成SCSI命令字,指定读取数据的设备

LUN和读取数据的长度等信息。

[0084] Linux侧交互接口104填写核间消息,核间消息的构成如图2所示,根据Linux侧交互接口104所在核填写发送消息核标示201,根据实时操作系统110上的实时操作系统侧交互接口106所在核填写接收消息核标示202,根据操作类型填写SCSI命令203和SCSI命令参数204,清除操作执行返回结果205,申请存放读取结果的内存块208,并将内存块208的物理地址填写到内存块指针206中,完成上述操作后将核间消息发送给实时操作系统110上的实时操作系统侧交互接口106。

[0085] 实时操作系统110上的实时操作系统侧交互接口106接收核间消息,并将SCSI命令203和SCSI命令参数204传递给实时操作系统110上的iSCSI处理模块107。

[0086] 实时操作系统110上的iSCSI处理模块107完成iSCSI协议处理后将iSCSI数据包提交给实时操作系统110上的TCP/IP协议栈108。

[0087] 实时操作系统110上的TCP/IP协议栈108完成iSCSI数据的发送,并等待iSCSI磁盘阵列的返回结果。

[0088] 实时操作系统110上的TCP/IP协议栈108接收iSCSI磁盘阵列的数据,并将数据返回给实时操作系统110上的iSCSI处理模块107,实时操作系统110上的iSCSI处理模块107根据核间消息中的数据块指针206将数据写入数据块208中,并将返回结果返回给实时操作系统110上的实时操作系统侧交互接口106。

[0089] 实时操作系统110上的实时操作系统侧交互接口106根据实时操作系统110上的实时操作系统侧交互接口106填写核间消息,根据Linux侧交互接口104所在核填写接收消息核标示201,根据实时操作系统110上的实时操作系统侧交互接口106所在核填写发送消息核标示202,填写操作执行返回结果205,完成上述操作后将核间消息发送给Linux侧交互接口104。

[0090] Linux侧交互接口104解析核间消息,将核间消息中的返回结果205和数据块208中的数据返回给Linux操作系统111中的应用程序。

[0091] Linux操作系统111上应用程序发起iSCSI磁盘阵列写入操作的步骤是:

[0092] Linux操作系统111上应用程序提交磁盘Write命令,经过Linux操作系统111的SCSI子系统101中SCSI较高层102和SCSI中间层103的处理后到达Linux侧交互接口104。

[0093] Linux侧交互接口104将磁盘Write命令封装成SCSI命令字,指定写入数据的设备LUN等信息。

[0094] Linux侧交互接口104填写核间消息,根据Linux侧交互接口104所在核填写发送消息核标示201,根据实时操作系统110上的实时操作系统侧交互接口106所在核填写接收消息核标示202,根据操作类型填写SCSI命令203和SCSI命令参数204,清除操作执行返回结果205,申请存放写入数据的内存块208,并将内存块208的物理地址填写到内存块指针206中,完成上述操作后将核间消息发送给实时操作系统110上的实时操作系统侧交互接口106。

[0095] 实时操作系统110上的实时操作系统侧交互接口106接收核间消息,并将SCSI命令203和SCSI命令参数204传递给实时操作系统110上的iSCSI处理模块107。

[0096] 实时操作系统110上的iSCSI处理模块107完成iSCSI协议处理后将iSCSI数据包提交给实时操作系统110上的TCP/IP协议栈108。

[0097] 实时操作系统110上的TCP/IP协议栈108完成iSCSI数据的发送,并等待iSCSI磁盘

阵列的返回结果。

[0098] 实时操作系统110上的TCP/IP协议栈108接收iSCSI磁盘阵列的返回结果,并将返回结果返回给实时操作系统110上的iSCSI处理模块107,实时操作系统110上的iSCSI处理模块107处理返回结果,并将返回结果返回给实时操作系统110上的实时操作系统侧交互接口106。

[0099] 实时操作系统110上的实时操作系统侧交互接口106根据实时操作系统110上的实时操作系统侧交互接口106填写核间消息,根据Linux侧交互接口104所在核填写接收消息核标示201,根据实时操作系统110上的实时操作系统侧交互接口106所在核填写发送消息核标示202,填写操作执行返回结果205,完成上述操作后将核间消息发送给Linux侧交互接口104。

[0100] Linux侧交互接口104解析核间消息,将核间消息中的返回结果205返回给Linux操作系统111中的应用程序。

[0101] 总之,本发明提供的一种基于多核处理器和异构操作系统的iSCSI initiator实现方法,具体是指在一颗多核处理器上同时运行Linux操作系统和实时操作系统,在Linux操作系统上保持iSCSI系统对Linux SCSI子系统的接口兼容,在实时操作系统上实现具体的iSCSI协议处理。方法包括:实现方法由Linux侧交互接口、实时操作系统侧交互接口、iSCSI处理模块和TCP/IP协议栈组成;Linux侧交互接口运行于Linux操作系统上,完成Linux SCSI子系统较低层的功能,并提供Linux操作系统和实时操作系统的交互功能;实时操作系统交互接口和iSCSI处理模块运行于实时操作系统上,提供实时操作系统和Linux操作系统的交互功能。Linux侧交互接口和实时操作系统侧交互接口通过多核核间消息建立数据和信令通道。该iSCSI initiator的实现方法兼容了Linux SCSI子系统接口,同时提供了高性能的iSCSI协议处理能力。在本发明提供的方法中由Linux侧交互接口、实时操作系统侧交互接口、iSCSI处理模块和TCP/IP协议栈完成iSCSI发起端的全部处理工作,操作系统之间和各个模块之间使用核间消息通信。Linux侧交互接口运行于Linux操作系统上,实现Linux SCSI子系统较低层的功能,并负责与实时操作系统交互接口进行交互;实时操作系统交互接口运行于实时操作系统上,负责与Linux侧交互接口交互;iSCSI处理模块运行于实时操作系统上,负责iSCSI协议的处理;TCP/IP协议栈运行于实时操作系统上,负责实现iSCSI数据的收发。

[0102] 最后所应说明的是,以上实施例仅用以说明本发明的技术方案而非限制。尽管参照实施例对本发明进行了详细说明,本领域的普通技术人员应当理解,对本发明的技术方案进行修改或者等同替换,都不脱离本发明技术方案的精神和范围,其均应涵盖在本发明的权利要求范围当中。

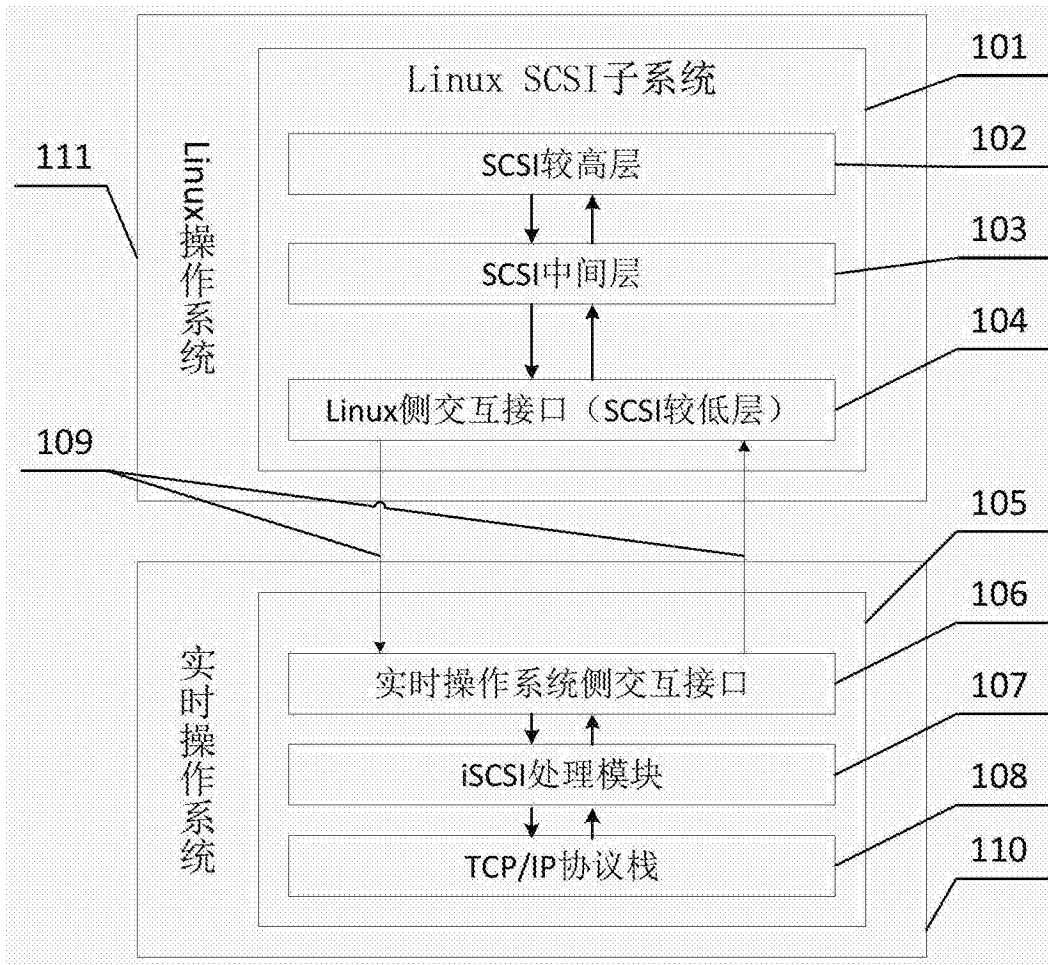


图1

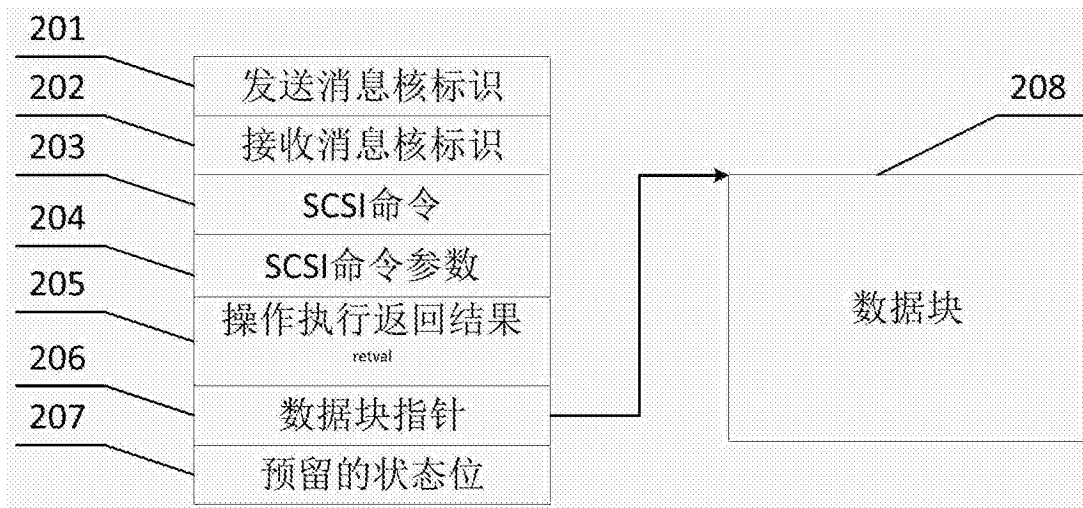


图2