



(12) 发明专利申请

(10) 申请公布号 CN 105260243 A

(43) 申请公布日 2016. 01. 20

(21) 申请号 201510727434. 6

(22) 申请日 2015. 10. 30

(71) 申请人 浪潮(北京)电子信息产业有限公司
地址 100085 北京市海淀区上地信息路2号
2-1号C栋1层

(72) 发明人 王荣廷

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227

代理人 罗满

(51) Int. Cl.
G06F 9/48(2006. 01)

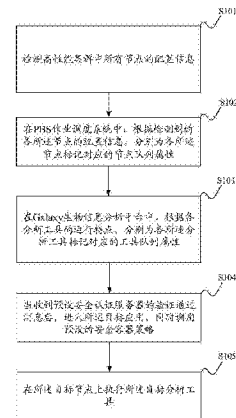
权利要求书1页 说明书4页 附图3页

(54) 发明名称

一种集群作业调度方法与系统

(57) 摘要

本发明公开了一种集群作业调度方法与系统,检测高性能集群中所有节点的配置信息,在PBS作业调度系统中,根据检测到的各所述节点的配置信息,分别为各所述节点标记对应的节点队列属性,在Gx galaxy生物信息分析平台中,根据各分析工具的运行特点,分别为各所述分析工具标记对应的工具队列属性,根据目标分析工具的工具队列属性匹配对应的目标节点,在所述目标节点上执行所述目标分析工具,将高性能集群, Galaxy生物信息分析平台, PBS作业调度系统相融合,将集群节点类型分类,针对Galaxy平台不同分析工具的运行特点,绑定对应的节点,使分析工具在适合的节点上工作,提高了Galaxy平台分析工具的运行效率。



1. 一种集群作业调度方法,其特征在于,包括:
检测高性能集群中所有节点的配置信息;
在 PBS 作业调度系统中,根据检测到的各所述节点的配置信息,分别为各所述节点标记对应的节点队列属性;
在 Gx1axy 生物信息分析平台中,根据各分析工具的运行特点,分别为各所述分析工具标记对应的工具队列属性;
根据目标分析工具的工具队列属性匹配对应的目标节点;
在所述目标节点上执行所述目标分析工具。
2. 根据权利要求 1 所述的集群作业调度方法,其特征在于,检测高性能集群中所有节点的配置信息前还包括:
在高性能集群中进行 PBS 作业调度系统和 Galaxy 生物信息分析平台的配置。
3. 根据权利要求 1 所述的集群作业调度方法,其特征在于,根据各分析工具的运行特点,分别为各所述分析工具标记对应的工具队列属性包括:
确定各分析工具适合运行的对应节点,分别将各所述分析工具对应节点的节点队列属性标记为该分析工具的工具队列属性。
4. 根据权利要求 1 所述的集群作业调度方法,其特征在于,根据目标分析工具的工具队列属性匹配对应的目标节点包括:
确定目标分析工具的工具队列属性,根据该工具队列属性匹配对应的节点队列属性标记的目标节点。
5. 根据权利要求 1~4 任一项所述的集群作业调度方法,其特征在于,所述节点的配置信息包括 CPU 数量,CPU 主频以及内存值。
6. 一种集群作业调度系统,其特征在于,包括:
配置信息获取模块,用于检测高性能集群中所有节点的配置信息;
节点队列属性标记模块,用于在 PBS 作业调度系统中,根据检测到的各所述节点的配置信息,分别为各所述节点标记对应的节点队列属性;
工具队列属性标记模块,用于在 Gx1axy 生物信息分析平台中,根据各分析工具的运行特点,分别为各所述分析工具标记对应的工具队列属性;
匹配模块,用于根据目标分析工具的工具队列属性匹配对应的目标节点;
执行模块,用于在所述目标节点上执行所述目标分析工具。
7. 根据权利要求 2 所述的集群作业调度系统,其特征在于,还包括:
配置模块,用于在高性能集群中进行 PBS 作业调度系统和 Galaxy 生物信息分析平台的配置。
8. 根据权利要求 6 和 7 任一项所述的集群作业调度系统,其特征在于,所述节点的配置信息包括 CPU 数量,CPU 主频以及内存值。

一种集群作业调度方法与系统

技术领域

[0001] 本发明涉及生物信息领域,特别是涉及一种基于集群作业调度方法与系统。

背景技术

[0002] 传统的 Galaxy 生物信息分析平台,一般是将 Galaxy 平台与高性能集群简单的进行整合,平台中的各个工具运行后,直接将作业提交到集群中运行。

[0003] 但是 Galaxy 生物信息分析平台中各个工具都有对应的运行特点,适合在不同类型的节点上运行,但现在无法针对各个工具的运行特点,针对性的将作业提交到某一节点,导致 Galaxy 平台中分析工具的运行效率不高,同时造成了集群计算资源的浪费。

发明内容

[0004] 有鉴于此,本发明的主要目的在于提供一种集群作业调度方法与系统,可以提高 Galaxy 平台分析工具的运行效率。

[0005] 为实现上述目的,本发明提供了一种集群作业调度方法,包括:

[0006] 检测高性能集群中所有节点的配置信息;

[0007] 在 PBS 作业调度系统中,根据检测到的各所述节点的配置信息,分别为各所述节点标记对应的节点队列属性;

[0008] 在 Galaxy 生物信息分析平台中,根据各分析工具的运行特点,分别为各所述分析工具标记对应的工具队列属性;

[0009] 根据目标分析工具的工具队列属性匹配对应的目标节点;

[0010] 在所述目标节点上执行所述目标分析工具。

[0011] 优选地,检测高性能集群中所有节点的配置信息前还包括:

[0012] 在高性能集群中进行 PBS 作业调度系统和 Galaxy 生物信息分析平台的配置。

[0013] 优选地,根据各分析工具的运行特点,分别为各所述分析工具标记对应的工具队列属性包括:

[0014] 确定各分析工具适合运行的对应节点,分别将各所述分析工具对应节点的节点队列属性标记为该分析工具的工具队列属性。

[0015] 优选地,根据目标分析工具的工具队列属性匹配对应的目标节点包括:

[0016] 确定目标分析工具的工具队列属性,根据该工具队列属性匹配对应的节点队列属性标记的目标节点。

[0017] 优选地,所述节点的配置信息包括 CPU 数量,CPU 主频以及内存值。

[0018] 本发明还提供了一种集群作业调度系统,包括:

[0019] 配置信息获取模块,用于检测高性能集群中所有节点的配置信息;

[0020] 节点队列属性标记模块,用于在 PBS 作业调度系统中,根据检测到的各所述节点的配置信息,分别为各所述节点标记对应的节点队列属性;

[0021] 工具队列属性标记模块,用于在 Galaxy 生物信息分析平台中,根据各分析工具的

运行特点,分别为各所述分析工具标记对应的工具队列属性;

[0022] 匹配模块,用于根据目标分析工具的工具队列属性匹配对应的目标节点;

[0023] 执行模块,用于在所述目标节点上执行所述目标分析工具。

[0024] 优选地,所述集群作业调度系统还包括:

[0025] 配置模块,用于在高性能集群中进行 PBS 作业调度系统和 Galaxy 生物信息分析平台的配置。

[0026] 优选地,所述节点的配置信息包括 CPU 数量, CPU 主频以及内存值。

[0027] 应用本发明提供一种集群作业调度方法与系统,检测高性能集群中所有节点的配置信息,在 PBS 作业调度系统中,根据检测到的各所述节点的配置信息,分别为各所述节点标记对应的节点队列属性,在 Galaxy 生物信息分析平台中,根据各分析工具的运行特点,分别为各所述分析工具标记对应的工具队列属性,根据目标分析工具的工具队列属性匹配对应的目标节点,在所述目标节点上执行所述目标分析工具,将高性能集群, Galaxy 生物信息分析平台, PBS 作业调度系统相融合,将集群节点类型分类,针对 Galaxy 平台不同分析工具的运行特点,绑定对应的节点,使分析工具在适合的节点上工作,提高了 Galaxy 平台分析工具的运行效率。

附图说明

[0028] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据提供的附图获得其他的附图。

[0029] 图 1 为本发明一种集群作业调度方法实施例一的流程图;

[0030] 图 2 为本发明一种集群作业调度方法实施例一的原理示意图;

[0031] 图 3 为本发明一种集群作业调度方法实施例一的详细原理示意图;

[0032] 图 4 为本发明一种集群作业调度系统实施例二的结构示意图;

[0033] 图 5 为本发明一种集群作业调度系统实施例三的结构示意图。

具体实施方式

[0034] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0035] 实施例一:

[0036] 本发明提供了一种集群作业调度方法,图 1 示出了本发明集群作业调度方法实施例一的流程图,包括:

[0037] 步骤 S101:检测高性能集群中所有节点的配置信息;

[0038] 首先在高性能集群中,安装配置 PBS 作业调度系统,安装 Galaxy 生物信息分析平台,并安装分析工具如基因测序工具,检测 HPC 高性能集群中所有节点的配置信息,包括 CPU 数量 CPU 主频内存大小等配置信息,如对集群中节点 node1 的配置信息进行检测,检测

到 node1 的配置是 :胖节点,高主频,大内存。

[0039] 步骤 S102 :在 PBS 作业调度系统中,根据检测到的各所述节点的配置信息,分别为各所述节点标记对应的节点队列属性 ;

[0040] 根据检测到的节点配置信息,在 PBS 作业调度系统中为节点标记不同的队列属性,如根据 node1 检测到的配置信息为其标记队列属性为 FAT-HF-HM(胖节点-高主频-大内存)。

[0041] 步骤 S103 :在 Gx1axy 生物信息分析平台中,根据各分析工具的运行特点,分别为各所述分析工具标记对应的工具队列属性 ;

[0042] 在 Gx1axy 生物信息分析平台中,确定各分析工具适合运行的对应节点,分别将各所述分析工具对应节点的节点队列属性标记为该分析工具的工具队列属性,如查找 Galaxy 平台中适合在 node1 上运行的分析工具,查找到分析工具 bwa 适合在 node1 上运行,为其标记队列属性 FAT-HF-HM。

[0043] 步骤 S104 :根据目标分析工具的工具队列属性匹配对应的目标节点 ;

[0044] 要使用目标分析工具时,确定目标分析工具的工具队列属性,根据该工具队列属性匹配对应的节点队列属性标记的目标节点,如使用 bwa 工具时,首先查找其对应队列属性的队列 FAT-HF-HM,根据队列 FAT-HF-HM,查找到队列属性为 FAT-HF-HM 的节点 node1。

[0045] 步骤 S105 :在所述目标节点上执行所述目标分析工具。

[0046] 在节点 node1 上开始执行分析工具 bwa。

[0047] 如图 2 所示,为本实施例的原理图,图 3 为本实施例具体实施过程原理图。

[0048] 应用本实施例提供的一种集群作业调度方法,检测高性能集群中所有节点的配置信息,在 PBS 作业调度系统中,根据检测到的各所述节点的配置信息,分别为各所述节点标记对应的节点队列属性,在 Gx1axy 生物信息分析平台中,根据各分析工具的运行特点,分别为各所述分析工具标记对应的工具队列属性,根据目标分析工具的工具队列属性匹配对应的目标节点,在所述目标节点上执行所述目标分析工具,将高性能集群, Galaxy 生物信息分析平台, PBS 作业调度系统相融合,将集群节点类型分类,针对 Galaxy 平台不同分析工具的运行特点,绑定对应的节点,使分析工具在适合的节点上工作,提高了 Galaxy 平台分析工具的运行效率。

[0049] 实施例二 :

[0050] 本发明还提供了一种集群作业调度系统,图 4 示出了本发明集群作业调度系统实施例二的结构示意图,包括 :

[0051] 配置信息获取模块 101,用于检测高性能集群中所有节点的配置信息 ;

[0052] 节点队列属性标记模块 102,用于在 PBS 作业调度系统中,根据检测到的各所述节点的配置信息,分别为各所述节点标记对应的节点队列属性 ;

[0053] 工具队列属性标记模块 103,用于在 Gx1axy 生物信息分析平台中,根据各分析工具的运行特点,分别为各所述分析工具标记对应的工具队列属性 ;

[0054] 匹配模块 104,用于根据目标分析工具的工具队列属性匹配对应的目标节点 ;

[0055] 执行模块 105,用于在所述目标节点上执行所述目标分析工具。

[0056] 本实施例中所述节点的配置信息包括 CPU 数量, CPU 主频以及内存值,应用本实施例提供的一种集群作业调度系统,检测高性能集群中所有节点的配置信息,在 PBS 作业

调度系统中,根据检测到的各所述节点的配置信息,分别为各所述节点标记对应的节点队列属性,在 GxIaxy 生物信息分析平台中,根据各分析工具的运行特点,分别为各所述分析工具标记对应的工具队列属性,根据目标分析工具的工具队列属性匹配对应的目标节点,在所述目标节点上执行所述目标分析工具,将高性能集群, Galaxy 生物信息分析平台, PBS 作业调度系统相融合,将集群节点类型分类,针对 Galaxy 平台不同分析工具的运行特点,绑定对应的节点,使分析工具在适合的节点上工作,提高了 Galaxy 平台分析工具的运行效率。

[0057] 实施例三:

[0058] 图 5 示出了本发明集群作业调度系统实施例三的结构示意图,对应于图 4,还包括:

[0059] 配置模块 100,用于在高性能集群中进行 PBS 作业调度系统和 Galaxy 生物信息分析平台的配置。

[0060] 本实施例中所述节点的配置信息包括 CPU 数量, CPU 主频以及内存值。

[0061] 需要说明的是,本说明书中的各个实施例均采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似的部分互相参见即可。对于系统类实施例而言,由于其与方法实施例基本相似,所以描述的比较简单,相关之处参见方法实施例的部分说明即可。

[0062] 最后,还需要说明的是,在本文中,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

[0063] 以上对本发明所提供的方法和系统进行了详细介绍,本文中应用了具体个例对本发明的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本发明的方法及其核心思想;同时,对于本领域的一般技术人员,依据本发明的思想,在具体实施方式及应用范围上均会有改变之处,综上所述,本说明书内容不应理解为对本发明的限制。

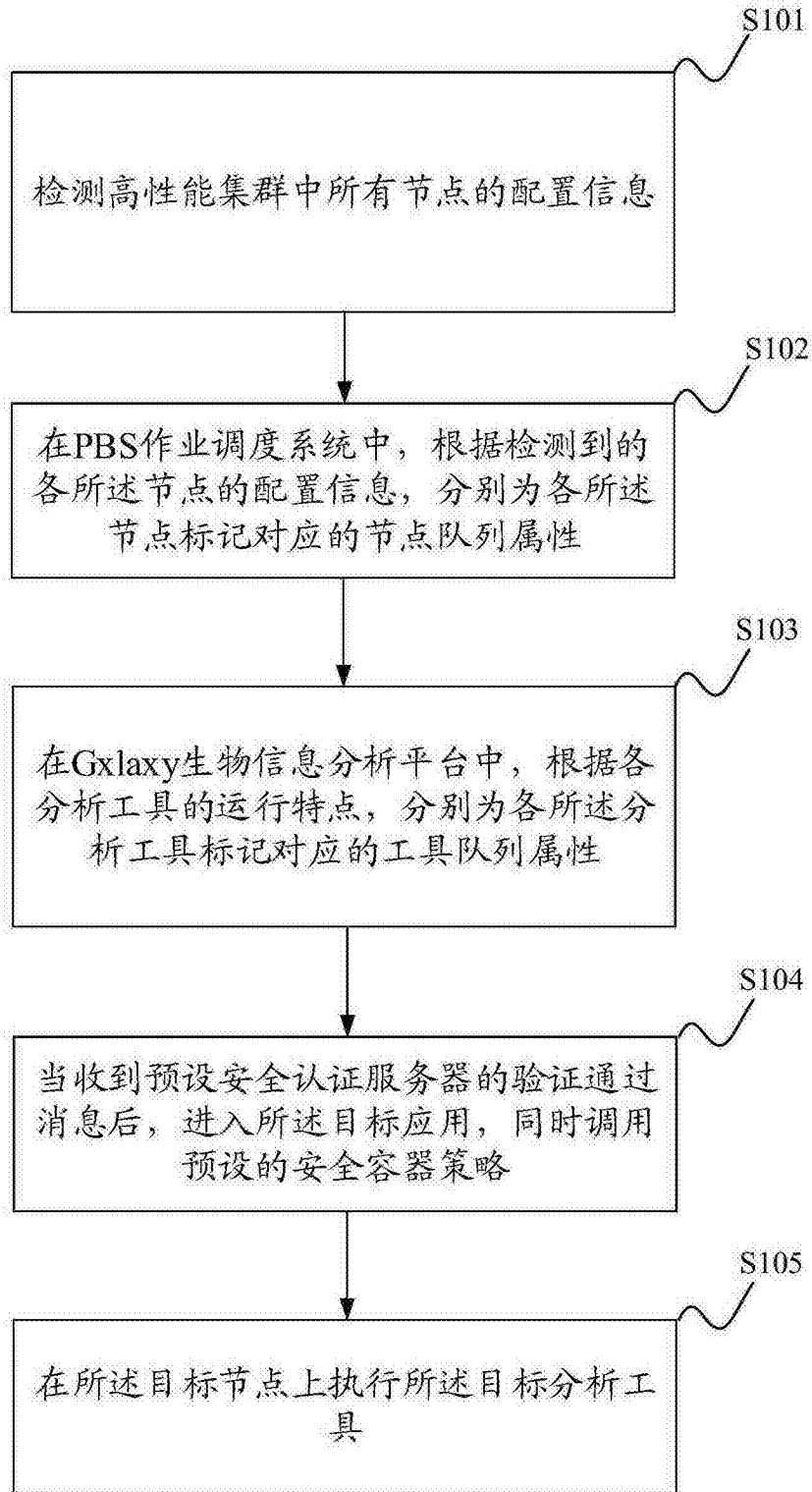


图 1

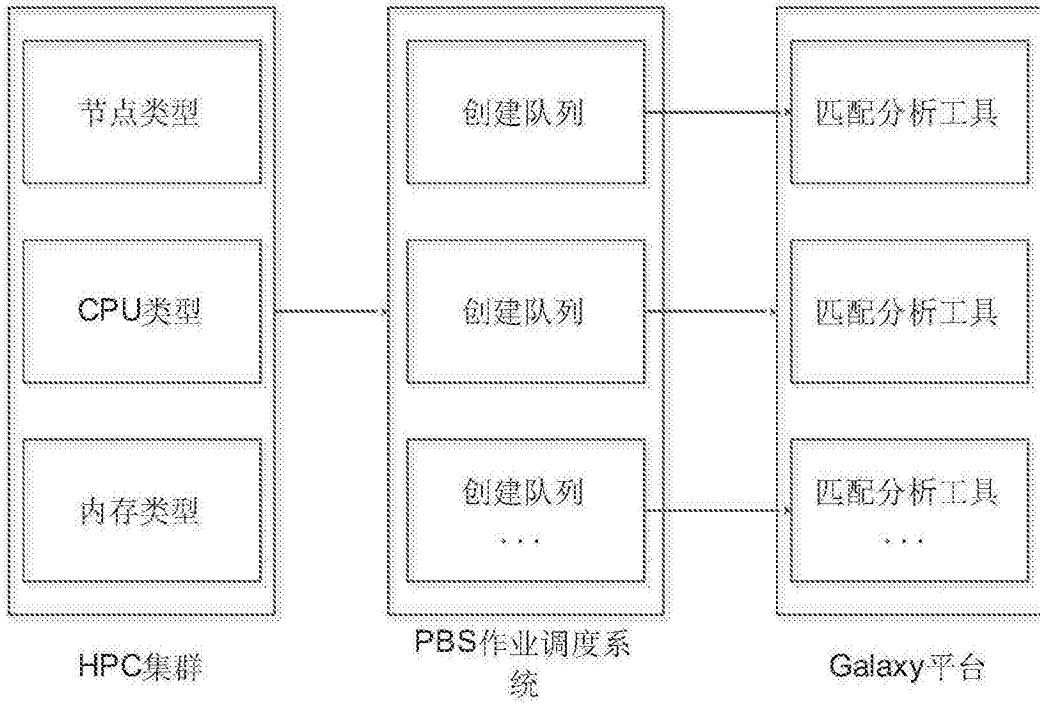


图 2

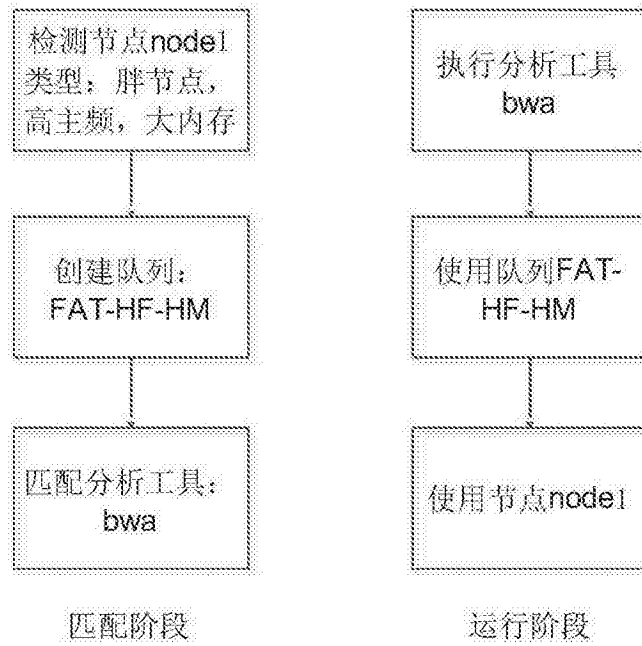


图 3



图 4

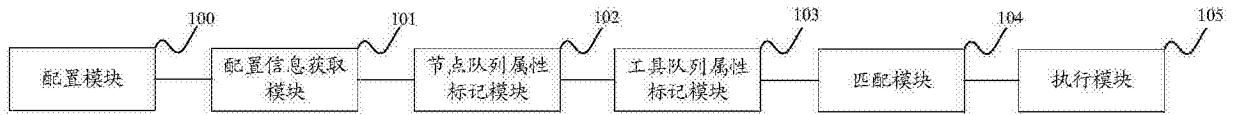


图 5