



(12) 发明专利申请

(10) 申请公布号 CN 102694868 A

(43) 申请公布日 2012.09.26

(21) 申请号 201210184950.5

(22) 申请日 2012.06.07

(71) 申请人 浪潮电子信息产业股份有限公司
地址 250014 山东省济南市高新区舜雅路
1036 号

(72) 发明人 王通

(51) Int. Cl.

H04L 29/08 (2006.01)

H04L 29/06 (2006.01)

H04L 12/26 (2006.01)

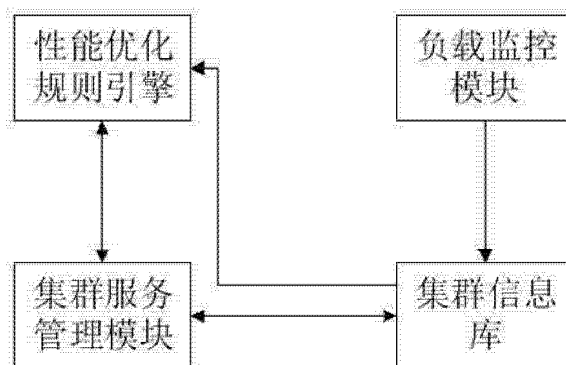
权利要求书 3 页 说明书 7 页 附图 4 页

(54) 发明名称

一种集群系统实现及任务动态分配方法

(57) 摘要

本发明提供一种集群系统实现及任务动态分配方法,利用模式方法来实现对硬件设备及服务的监控,用户需要监控某类硬件设备或服务时,仅需要在模式监控配置文件中填写该类硬件设备或服务的模式,模式分析模块分析该文件并即时检测系统中符合该模式的所有硬件设备或服务,生成系统中相应设备的设备文件名或服务名并写入具体设备及服务监控配置文件,当发生硬件热插拔或服务在线迁移操作包括数据卷在线迁移时,设备及服务检测模块被触发获取被热插拔的设备或自动迁移的服务,并自动修改配置文件,系统包括:模式分析模块、设备及服务检测模块。



1. 一种集群任务动态分配方法，包括集群服务动态配置装置和集群系统的实现方式，其特征在于将各类硬件设备及各类软件服务按预定义的模式分类，动态监控指定模式的具体硬件设备及服务信息，以此来实现服务器、存储设备的大规模动态实时监控，利用模式方法来实现对硬件设备及服务的监控，用户需要监控某类硬件设备或服务时，仅需要在模式监控配置文件中填写该类硬件设备或服务的模式，模式分析模块分析该文件并即时检测系统中符合该模式的所有硬件设备或服务，生成系统中相应设备的设备文件名或服务名并写入具体设备及服务监控配置文件，当发生硬件热插拔或服务在线迁移操作包括数据卷在线迁移时，设备及服务检测模块被触发获取被热插拔的设备或自动迁移的服务，并自动修改配置文件，系统包括：模式分析模块、设备及服务检测模块，其中：

模式分析模块是已有监控方式的改进，目前的监控方式需要用户写入具体的需要被监控的设备及服务在系统中的名字，改进后，用户只需要写入 希望被监控的某类或某几类设备及服务在系统中的定义，模式分析模块分析该模式监控配置文件，并依据此在系统中搜索具体的满足模式监控配置文件内的模式的硬件设备及服务，并写入具体设备及服务监控配置文件，以备监控模块使用；

设备及服务检测模块是对硬件设备热插拔及服务迁移的检测及报告模块，在如今服务器及存储设备中，部分硬件设备如硬盘都实现了热插拔，当热插拔操作发生时，该模块检测到热插拔操作，得到热插拔设备的具体信息，然后按照用户配置或者通知模式分析模块重新分析，生成具体设备及服务监控配置文件或对直接修改具体设备及服务监控配置文件，并通知监控模块，集群环境中发生故障时实现服务的在线迁移，当服务迁移时，设备及服务检测模块检测到被迁移走或迁入的服务，然后按照用户配置或者通知模式分析模块重新分析，生成具体设备及服务监控配置文件或对直接修改具体设备及服务监控配置文件，并通知监控模块；

模式分析模块分析用户填写的模式监控配置文件，并依据分析的结果即时检测系统中存在的属于用户指定模式的具体硬件设备及服务，并将具体硬件设备及服务信息填入具体设备及服务监控配置文件，实现了监控的动态性、实时性、可靠性、更好的扩展能力和可管理性；

设备及服务检测模块动态方式监控，不依赖已写好的具体设备监控配置文件，而是动态监测硬件设备热插拔及服务迁移，并将已经改变的硬件及服务环境信息反映到具体设备监控配置文件中，在用户改变系统硬件配置或服务迁移时，不需要自己修改配置文件，实现了监控的动态性、实时性、可靠性、更好的扩展能力和可管理性。

2. 根据权利要求 1 所述的方法，其特征在于集群服务动态配置装置，包括负载监控模块、性能优化规则引擎、集群服务管理模块以及集群信息库，其中：

负载监控模块，用以监控集群中各节点的负载情况，并将其存储至集群信息库，节点的负载情况包括节点资源利用率和节点的性能负载情况，其中，节点资源利用率，包括 CPU 的利用率、内存的利用率；节点的性能负载情况包括数据流量；

性能优化规则引擎，根据集群信息库中各节点的历史监控数据及当前监控数据，预测各节点未来的负载情况，以及根据预测结果和一预设的服务均衡策略制定集群中各节点上的服务分配方案；

性能优化规则引擎，是通过将集群中各节点当前监控数据以及若干历史监控数据与一

预设的预测模型进行匹配,从而预测该各节点未来的负载情况,所述预测模型包括:自回归模型 AR(p)、滑动平均模型 MA(q)、自回归滑动平均模型 ARMA(p, q),在进行预测模型匹配之前,性能优化规则引擎还对所述负载监控模块采集到的负载样本进行预处理及平稳化处理;

考虑到预测的步数越多,预测的结果与实际值相差的越大,越久远的历史信息对预测结果的影响越小,因此,选择离当前观测值近的几个历史观测值和当前观测值,利用预测模型预测下一步的结果,主节点会根据各个节点的预测值进行判断,并进行服务的重新分配;

预设的服务均衡策略包括:配置集群中各节点的资源承载能力,分别判断预测出的各节点未来的负载是否超出其承载能力,如果超出,则在各节点间进行服务均衡,以确保各节点未来的负载不超出其承载能力;

集群服务管理模块,负责集群服务分配的相关操作,负责维护所述集群信息库,包括节点、服务、利用率、性能负载监控的信息,根据性能优化规则引擎制定的集群中各节点上的服务分配方案,调整集群中各节点的服务分配,它是唯一一个决定在整个集群进行服务更改的实体;

集群信息库,用以存储所述负载监控模块监控得到的集群中各节点的负载数据,所述节点的负载数据包括节点资源的利用率、节点的性能负载,所述节点资源的利用率包括 CPU 的利用率、内存的利用率;所述节点的性能负载包括数据流量;

进一步地,负载监控模块在监控到集群中的某节点宕机时,实时通知性能优化规则引擎,性能优化规则引擎在获知某节点宕机时,从所述集群信息库中获取该节点上的服务数据,并根据预测出的集群中其他节点未来的负载情况,将该节点上的服务数据分配给集群中其他节点,且确保集群中所述其他各节点未来的负载不超出其承载能力。

3. 根据权利要求 1 所述的方法,其特征在于集群系统的实现方式,包括一主节点和至少一从属节点,主节点和从属节点上均包括负载监控模块、性能优化规则引擎、集群服务管理模块、集群信息库、集群通信及成员关系模块、本地服务管理模块、集群信息库、服务代理模块以及隔离设备,其中:

主节点上和从属节点上的负载监控模块,用以监控其所在节点的负载情况,并将其存储至其所在节点的所述集群信息库;

主节点上和从属节点上的性能优化规则引擎,用以根据其所在节点上的集群信息库中存储的历史监控数据及当前监控数据,预测其所在节点未来的负载情况,并将其存储至其所在节点的集群信息库;主节点上的性能优化规则引擎,还用以根据其所在节点的集群信息库中存储的集群系统中各节点未来的负载情况的预测结果和一预设的服务均衡策略制定集群中各节点上的服务分配方案,并将其分别发送至所述从属节点的所述集群服务管理模块中;

主节点上和从属节点上的集群服务管理模块,用以将其所在节点上的集群信息库中的数据同步为集群系统中所有节点上的集群信息库中的数据的并集,以及根据主节点上的性能优化规则引擎制定的集群中各节点上的服务分配方案调整其所在节点上的服务分配情况;

集群通信及成员关系模块,用以构建高可用集群成员关系,以及负责集群中各节点之

间的消息通讯,消息通讯是指集群节点间的信息交互,包括节点、服务配置、监控信息等的交互及同步,构建高可用集群成员关系是,在高可用集群中的节点之间建立高可用关系,使得失效节点的服务能够被集群中其他节点接管;

本地服务管理模块,用以与服务代理模块交互,通过服务代理模块管理本节点服务;

服务代理模块,负责直接管理服务;

隔离设备,用以将失效节点隔离,防止其对集群造成不利影响;

进一步地,所述预设的服务均衡策略包括:配置集群中各节点的资源承载能力,分别判断预测出的各节点未来的负载是否超出其承载能力,如果超出,则在各节点间进行服务均衡,以确保各节点未来的负载不超出其承载能力;

进一步地,当从属节点宕机时,所述主节点的所述性能优化规则引擎根据所述主节点的所述集群信息库中存储的所述宕机的从属节点上的服务数据,以及预测出的集群中节点未来的负载情况,将所述宕机的从属节点上的服务数据分配给集群中节点,且确保集群中节点未来的负载不超出其承载能力;

进一步地,当主节点宕机时,其中一个从属节点替代该主节点成为集群系统中新的主节点;

新的主节点的所述性能优化规则引擎,根据新的主节点的集群信息库中存储的宕机的节点上的服务数据,以及预测出的新的主节点和集群中其他从属节点未来的负载情况,将宕机的节点上的服务数据分配给集群中节点,且确保集群中的节点未来的负载不超出其承载能力。

一种集群系统实现及任务动态分配方法

技术领域

[0001] 本发明涉及计算机集群技术领域,具体涉及一种集群系统实现及任务动态分配方法。

背景技术

[0002] 目前较为流行的集群服务器、集群存储监控软件如 ganglia 等大多需要在配置文件中指明需要监控的具体设备及具体服务,以硬盘为例,需要在配置文件中填写被监控硬盘在系统中的设备名。在大规模集群监控中,需要监控的硬件设备及服务数量极大,需要大规模填写监控配置文件,并且当今服务器、存储设备多实现了硬件设备热插拔方式,集群存储也多采用高可用方法实现在线迁移服务,因此硬件设备及服务常常动态变化,采用原有方式在热插拔硬件,人工修改监控配置文件,会对整个集群系统的实时监控不利,同时高可用软件自动迁移服务后,在人工修改监控配置文件之前,迁移的服务在目标节点上无法监控。

[0003] 高性能计算技术的发展是伴随着计算机技术的发展而发展的,也就是说,从计算机技术诞生之日起,人们就在为追求更高计算能力的计算机系统而努力。在过去几十年间,可以说是高性能计算机体系结构和通信技术不断创新的年代,出现了包括 MPP (Massive Parallel Processing,海量并行处理结构)、SMP (Symmetric Multi-Processor,对称多处理器结构)、集群等各种各样的体系结构及网络互联技术。尤其是最近几年,集群技术发展迅速,已经成为构建超级计算机系统的主流架构之一。在最新发布的 TOP500 (TOP500 是世界最强超级计算机排名的简称,这个排行榜自 1993 年设立,每年 6 月、11 月两次统计发布世界上最强大的 500 个计算机系统。)中,超过 80% 的入选系统都采用了集群架构,并在数量上仍有不断增加的趋势,可见集群技术在高性能计算领域有着旺盛的生命力和广阔的发展前景。

[0004] 随着商业应用的不断发展,高性能集群技术在商业应用领域也得到了广泛的应用,由于商业用户希望自己的业务程序能够对外提供不间断的服务,把因软件、硬件,以及人为等原因造成的故障对业务的影响降低到最低程度,因此,要求如果某个节点失效,它的备用节点能够在几秒钟时间内接管它的职责。传统的集群服务高可用静态配置方法虽然能够实现服务的接管,但是却无法在剩余集群节点间实现接管服务的均衡分配,同时传统的方法不能根据负载监控状况进行服务的动态优化调整。

发明内容

[0005] 本发明的目的是提供一种集群系统实现及任务动态分配方法。

[0006] 本发明的目的是按以下方式实现的,将各类硬件设备及各类软件服务按预定义的模式分类,动态监控指定模式的具体硬件设备及服务信息,以此来实现服务器、存储设备的大规模动态实时监控,利用模式方法来实现对硬件设备及服务的监控,用户需要监控某类硬件设备或服务时,仅需要在模式监控配置文件中填写该类硬件设备或服务的模式,模式

分析模块分析该文件并即时检测系统中符合该模式的所有硬件设备或服务,生成系统中相应设备的设备文件名或服务名并写入具体设备及服务监控配置文件,当发生硬件热插拔或服务在线迁移操作包括数据卷在线迁移时,设备及服务检测模块被触发获取被热插拔的设备或自动迁移的服务,并自动修改配置文件,系统包括:模式分析模块、设备及服务检测模块,其中:

模式分析模块是已有监控方式的改进,目前的监控方式需要用户写入具体的需要被监控的设备及服务在系统中的名字,改进后,用户只需要写入 希望被监控的某类或某几类设备及服务在系统中的定义,模式分析模块分析该模式监控配置文件,并依据此在系统中搜索具体的满足模式监控配置文件内的模式的硬件设备及服务,并写入具体设备及服务监控配置文件,以备监控模块使用;

设备及服务检测模块是对硬件设备热插拔及服务迁移的检测及报告模块,在如今服务器及存储设备中,部分硬件设备如硬盘都实现了热插拔,当热插拔操作发生时,该模块检测到热插拔操作,得到热插拔设备的具体信息,然后按照用户配置或者通知模式分析模块重新分析,生成具体设备及服务监控配置文件或对直接修改具体设备及服务监控配置文件,并通知监控模块,集群环境中发生故障时实现服务的在线迁移,当服务迁移时,设备及服务检测模块检测到被迁移走或迁入的服务,然后按照用户配置或者通知模式分析模块重新分析,生成具体设备及服务监控配置文件或对直接修改具体设备及服务监控配置文件,并通知监控模块;

模式分析模块分析用户填写的模式监控配置文件,并依据分析的结果即时检测系统中存在的属于用户指定模式的具体硬件设备及服务,并将具体硬件设备及服务信息填入具体设备及服务监控配置文件, 实现了监控的动态性、实时性、可靠性、更好的扩展能力和可管理性;

设备及服务检测模块动态方式监控,不依赖已写好的具体设备监控配置文件,而是动态监测硬件设备热插拔及服务迁移,并将已经改变的硬件及服务环境信息反映到具体设备监控配置文件中,在用户改变系统硬件配置或服务迁移时,不需要自己修改配置文件,实现了监控的动态性、实时性、可靠性、更好的扩展能力和可管理性。

[0007] 一种集群服务动态配置装置,包括负载监控模块、性能优化规则引擎、集群服务管理模块以及集群信息库,其中:

负载监控模块,用以监控集群中各节点的负载情况,并将其存储至集群信息库,节点的负载情况包括节点资源利用率和节点的性能负载情况,其中,节点资源利用率,包括 CPU 的利用率、内存的利用率;节点的性能负载情况包括数据流量;

性能优化规则引擎,根据集群信息库中各节点的历史监控数据及当前监控数据,预测各节点未来的负载情况,以及根据预测结果和一预设的服务均衡策略制定集群中各节点上的服务分配方案;

性能优化规则引擎,是通过将集群中各节点当前监控数据以及若干历史监控数据与一预设的预测模型进行匹配,从而预测该各节点未来的负载情况,所述预测模型包括:自回归模型 AR(p)、滑动平均模型 MA(q)、自回归滑动平均模型 ARMA(p, q),在进行预测模型匹配之前,性能优化规则引擎还对所述负载监控模块采集到的负载样本进行预处理及平稳化处理。

[0008] 考虑到预测的步数越多,预测的结果与实际值相差的越大,越久远的历史信息对预测结果的影响越小,因此,选择离当前观测值近的几个历史观测值和当前观测值,利用预测模型预测下一步的结果,主节点会根据各个节点的预测值进行判断,并进行服务的重新分配;

预设的服务均衡策略包括:配置集群中各节点的资源承载能力,分别判断预测出的各节点未来的负载是否超出其承载能力,如果超出,则在各节点间进行服务均衡,以确保各节点未来的负载不超出其承载能力;

集群服务管理模块,负责集群服务分配的相关操作,负责维护所述集群信息库,包括节点、服务、利用率、性能负载监控的信息,根据性能优化规则引擎制定的集群中各节点上的服务分配方案,调整集群中各节点的服务分配,它是唯一一个决定在整个集群进行服务更改的实体;

集群信息库,用以存储所述负载监控模块监控得到的集群中各节点的负载数据,所述节点的负载数据包括节点资源的利用率、节点的性能负载,所述节点资源的利用率包括 CPU 的利用率、内存的利用率;所述节点的性能负载包括数据流量;

进一步地,负载监控模块在监控到集群中的某节点宕机时,实时通知性能优化规则引擎,性能优化规则引擎在获知某节点宕机时,从所述集群信息库中获取该节点上的服务数据,并根据预测出的集群中其他节点未来的负载情况,将该节点上的服务数据分配给集群中其他节点,且确保集群中所述其他各节点未来的负载不超出其承载能力。

[0009] 一种集群系统实现方式,包括一主节点和至少一从属节点,主节点和从属节点上均包括负载监控模块、性能优化规则引擎、集群服务管理模块、集群信息库、集群通信及成员关系模块、本地服务管理模块、集群信息库、服务代理模块以及隔离设备,其中:

主节点上和从属节点上的负载监控模块,用以监控其所在节点的负载情况,并将其存储至其所在节点的所述集群信息库;

主节点上和从属节点上的性能优化规则引擎,用以根据其所在节点上的集群信息库中存储的历史监控数据及当前监控数据,预测其所在节点未来的负载情况,并将其存储至其所在节点的集群信息库;主节点上的性能优化规则引擎,还用根据其所在节点的集群信息库中存储的集群系统中各节点未来的负载情况的预测结果和一预设的服务均衡策略制定集群中各节点上的服务分配方案,并将其分别发送至所述从属节点的所述集群服务管理模块中;

主节点上和从属节点上的集群服务管理模块,用以将其所在节点上的集群信息库中的数据同步为集群系统中所有节点上的集群信息库中的数据的并集,以及根据主节点上的性能优化规则引擎制定的集群中各节点上的服务分配方案调整其所在节点上的服务分配情况;

集群通信及成员关系模块,用以构建高可用集群成员关系,以及负责集群中各节点之间的消息通讯,消息通讯是指集群节点间的信息交互,包括节点、服务配置、监控信息等的交互及同步,构建高可用集群成员关系是,在高可用集群中的节点之间建立高可用关系,使得失效节点的服务能够被集群中其他节点接管;

本地服务管理模块,用以与服务代理模块交互,通过服务代理模块管理本节点服务;

服务代理模块,负责直接管理服务;

隔离设备,用以将失效节点隔离,防止其对集群造成不利影响;

进一步地,所述预设的服务均衡策略包括:配置集群中各节点的资源承载能力,分别判断预测出的各节点未来的负载是否超出其承载能力,如果超出,则在各节点间进行服务均衡,以确保各节点未来的负载不超出其承载能力;

进一步地,当从属节点宕机时,所述主节点的所述性能优化规则引擎根据所述主节点的所述集群信息库中存储的所述宕机的从属节点上的服务数据,以及预测出的集群中节点未来的负载情况,将所述宕机的从属节点上的服务数据分配给集群中节点,且确保集群中节点未来的负载不超出其承载能力;

进一步地,当主节点宕机时,其中一个从属节点替代该主节点成为集群系统中新的主节点;

新的主节点的所述性能优化规则引擎,根据新的主节点的集群信息库中存储的宕机的节点上的服务数据,以及预测出的新的主节点和集群中其他从属节点未来的负载情况,将宕机的节点上的服务数据分配给集群中节点,且确保集群中的节点未来的负载不超出其承载能力。

[0010] 本发明的有益效果是:通过对节点负载进行监控,并据以对服务预先优化分配,可以平衡负载,提高并发服务,提高了整个集群系统的效能,降低因节点负载过重导致宕机的概率,可以防止单一节点接管宕机节点的所有服务可能造成的多米诺骨牌效应。

附图说明

[0011] 图1是集群服务动态分配装置组成示意图;

图2是集群服务动态分配方法流程示意图;

图3是模式分析及实时检测策略图;

图4是集群模式监控框图。

具体实施方式

[0012] 参照说明书附图对本发明的方法及装置作以下详细地说明。

[0013] 将各类硬件设备及各类软件服务按预定义的模式分类,动态监控指定模式的具体硬件设备及服务信息,以此来实现服务器、存储设备的大规模动态实时监控,利用模式方法来实现对硬件设备及服务的监控,用户需要监控某类硬件设备或服务时,仅需要在模式监控配置文件中填写该类硬件设备或服务的模式,模式分析模块分析该文件并即时检测系统中符合该模式的所有硬件设备或服务,生成系统中相应设备的设备文件名或服务名并写入具体设备及服务监控配置文件,当发生硬件热插拔或服务在线迁移操作包括数据卷在线迁移时,设备及服务检测模块被触发获取被热插拔的设备或自动迁移的服务,并自动修改配置文件,系统包括:模式分析模块、设备及服务检测模块,其中:

模式分析模块是已有监控方式的改进,目前的监控方式需要用户写入具体的需要被监控的设备及服务在系统中的名字,改进后,用户只需要写入希望被监控的某类或某几类设备及服务在系统中的定义,模式分析模块分析该模式监控配置文件,并依据此在系统中搜索具体的满足模式监控配置文件内的模式的硬件设备及服务,并写入具体设备及服务监控配置文件,以备监控模块使用;

设备及服务检测模块是对硬件设备热插拔及服务迁移的检测及报告模块,在如今服务器及存储设备中,部分硬件设备如硬盘都实现了热插拔,当热插拔操作发生时,该模块检测到热插拔操作,得到热插拔设备的具体信息,然后按照用户配置或者通知模式分析模块重新分析,生成具体设备及服务监控配置文件或对直接修改具体设备及服务监控配置文件,并通知监控模块,集群环境中发生故障时实现服务的在线迁移,当服务迁移时,设备及服务检测模块检测到被迁移走或迁入的服务,然后按照用户配置或者通知模式分析模块重新分析,生成具体设备及服务监控配置文件或对直接修改具体设备及服务监控配置文件,并通知监控模块;

模式分析模块分析用户填写的模式监控配置文件,并依据分析的结果即时检测系统中存在的属于用户指定模式的具体硬件设备及服务,并将具体硬件设备及服务信息填入具体设备及服务监控配置文件,实现了监控的动态性、实时性、可靠性、更好的扩展能力和可管理性;

设备及服务检测模块动态方式监控,不依赖已写好的具体设备监控配置文件,而是动态监测硬件设备热插拔及服务迁移,并将已经改变的硬件及服务环境信息反映到具体设备监控配置文件中,在用户改变系统硬件配置或服务迁移时,不需要自己修改配置文件,实现了监控的动态性、实时性、可靠性、更好的扩展能力和可管理性。

[0014] 通过深入研究注意到,节点的负载变化是一种时间序列,具有高度的自相似性,因此提出一种集群服务动态分配方法及装置,巧妙地利用时间序列对节点的负载进行预测,监控集群中各节点的负载情况,根据历史监测数据及当前的监控数据,预测各节点的未来负载情况,根据预测结果及时地协调服务分配,从而实现自适应的负载均衡。

[0015] 参见图 1,该图示出了集群服务动态配置装置,包括负载监控模块、性能优化规则引擎、集群服务管理模块以及集群信息库,其中:

负载监控模块,用以监控集群中各节点的负载情况,并将其存储至集群信息库。节点的负载情况可以包括节点资源的利用率、节点的性能负载情况等。节点资源的利用率可以是,例如 CPU 的利用率、内存的利用率等;节点的性能负载情况可以是,例如流量等。

[0016] 性能优化规则引擎,用以根据所述集群信息库中各节点的历史监控数据及当前监控数据,预测各节点未来的负载情况,以及根据预测结果和一预设的服务均衡策略制定集群中各节点上的服务分配方案。

[0017] 性能优化规则引擎,是通过将集群中各节点当前监控数据以及若干历史监控数据与一预设的预测模型进行匹配,从而预测该各节点未来的负载情况。所述预测模型可以是,例如自回归模型 AR(p)、滑动平均模型 MA(q)、自回归滑动平均模型 ARMA(p, q) 等等。在进行预测模型匹配之前,性能优化规则引擎还对所述负载监控模块采集到的负载样本进行预处理及平稳化处理。

[0018] 考虑到预测的步数越多,预测的结果与实际值相差的越大,越久远的历史信息对预测结果的影响越小,因此,应尽量选择离当前观测值近的几个历史观测值和当前观测值,利用预测模型预测下一步的结果。主节点会根据各个节点的预测值进行判断,并进行服务的重新分配。

[0019] 预设的服务均衡策略可以包括:配置集群中各节点的资源承载能力,分别判断预测出的各节点未来的负载是否超出其承载能力,如果超出,则在各节点间进行服务均衡,以

确保各节点未来的负载不超出其承载能力。

[0020] 集群服务管理模块,负责集群服务分配的相关操作,负责维护所述集群信息库,包括节点、服务、利用率、性能负载监控等信息,根据性能优化规则引擎制定的集群中各节点上的服务分配方案,调整集群中各节点的服务分配,它是唯一一个可以决定在整个集群进行服务更改的实体。

[0021] 集群信息库,用以存储所述负载监控模块监控得到的集群中各节点的负载数据。所述节点的负载数据可以包括节点资源的利用率、节点的性能负载等。所述节点资源的利用率可以是,例如 CPU 的利用率、内存的利用率等;所述节点的性能负载可以是,例如流量等。

[0022] 进一步地,负载监控模块在监控到集群中的某节点宕机时,通知性能优化规则引擎。性能优化规则引擎在获知某节点宕机时,从所述集群信息库中获取该节点上的服务数据,并根据预测出的集群中其他节点未来的负载情况,将该节点上的服务数据分配给集群中其他节点,且确保集群中所述其他各节点未来的负载不超出其承载能力。

[0023] 本发明还提供了一种集群系统实现方式,包括一主节点和至少一从属节点,主节点和从属节点上均包括负载监控模块、性能优化规则引擎、集群服务管理模块、集群信息库、集群通信及成员关系模块、本地服务管理模块、集群信息库、服务代理模块,以及隔离设备,如图 2 所示,其中:

主节点上和从属节点上的负载监控模块,用以监控其所在节点的负载情况,并将其存储至其所在节点的所述集群信息库。

[0024] 主节点上和从属节点上的性能优化规则引擎,用以根据其所在节点上的集群信息库中存储的历史监控数据及当前监控数据,预测其所在节点未来的负载情况,并将其存储至其所在节点的集群信息库;主节点上的性能优化规则引擎,还用以根据其所在节点的集群信息库中存储的集群系统中各节点未来的负载情况的预测结果和一预设的服务均衡策略制定集群中各节点上的服务分配方案,并将其分别发送至所述从属节点的所述集群服务管理模块中。

[0025] 主节点上和从属节点上的集群服务管理模块,用以将其所在节点上的集群信息库中的数据同步为集群系统中所有节点上的集群信息库中的数据的并集,以及根据主节点上的性能优化规则引擎制定的集群中各节点上的服务分配方案调整其所在节点上的服务分配情况。

[0026] 集群通信及成员关系模块,用以构建高可用集群成员关系,以及负责集群中各节点之间的消息通讯。消息通讯是指集群节点间的信息交互,包括节点、服务配置、监控信息等的交互及同步。构建高可用集群成员关系是,在高可用集群中的节点之间建立高可用关系,使得失效节点的服务能够被集群中其他节点接管。

[0027] 本地服务管理模块,用以与服务代理模块交互,通过服务代理模块管理本节点服务。

[0028] 服务代理模块,负责直接管理服务。

[0029] 隔离设备,用以将失效节点隔离,防止其对集群造成不利影响。

[0030] 进一步地,所述预设的服务均衡策略包括:配置集群中各节点的资源承载能力,分别判断预测出的各节点未来的负载是否超出其承载能力,如果超出,则在各节点间进行服

务均衡,以确保各节点未来的负载不超出其承载能力;

进一步地,当从属节点宕机时,所述主节点的所述性能优化规则引擎根据所述主节点的所述集群信息库中存储的所述宕机的从属节点上的服务数据,以及预测出的集群中节点未来的负载情况,将所述宕机的从属节点上的服务数据分配给集群中节点,且确保集群中节点未来的负载不超出其承载能力;

进一步地,当主节点宕机时,其中一个从属节点替代该主节点成为集群系统中新的主节点;

新的主节点的所述性能优化规则引擎根据新的主节点的集群信息库中存储的宕机的节点上的服务数据,以及预测出的新的主节点和集群中其他从属节点未来的负载情况,将宕机的节点上的服务数据分配给集群中节点,且确保集群中的节点未来的负载不超出其承载能力。

[0031] 除说明书所述的技术特征外,均为本专业技术人员的已知技术。

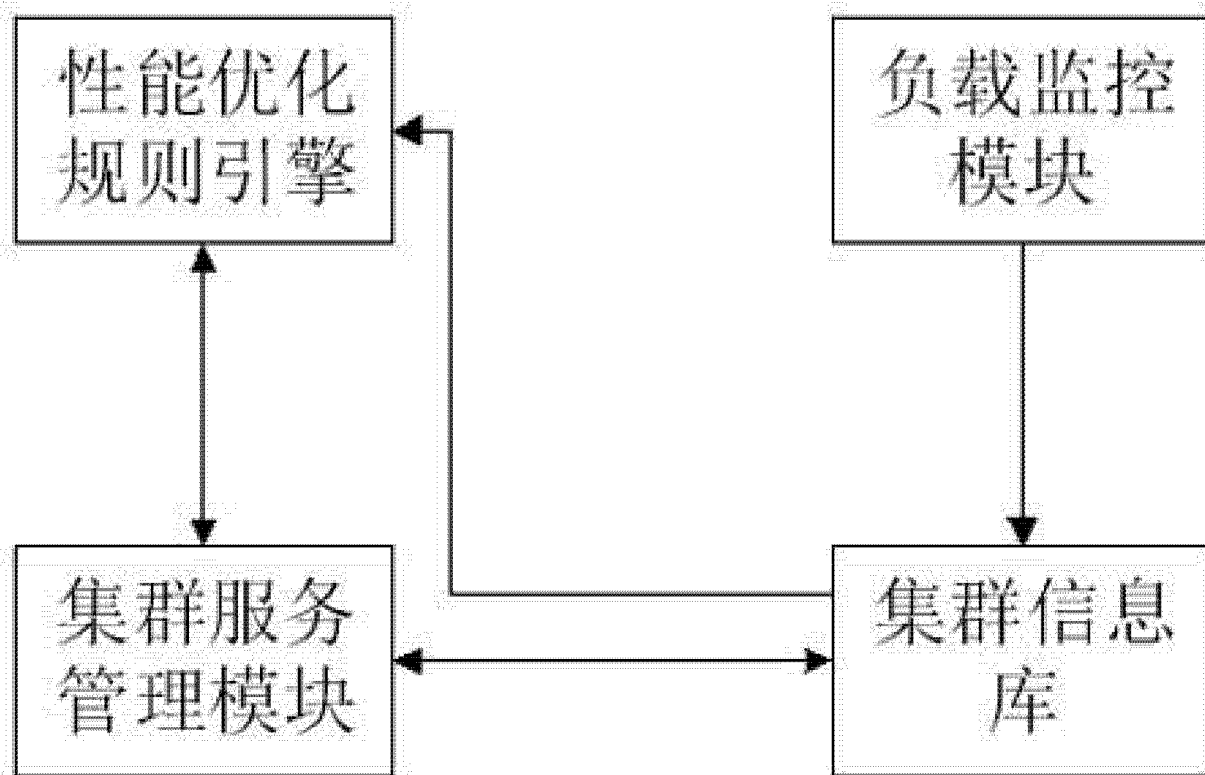


图 1

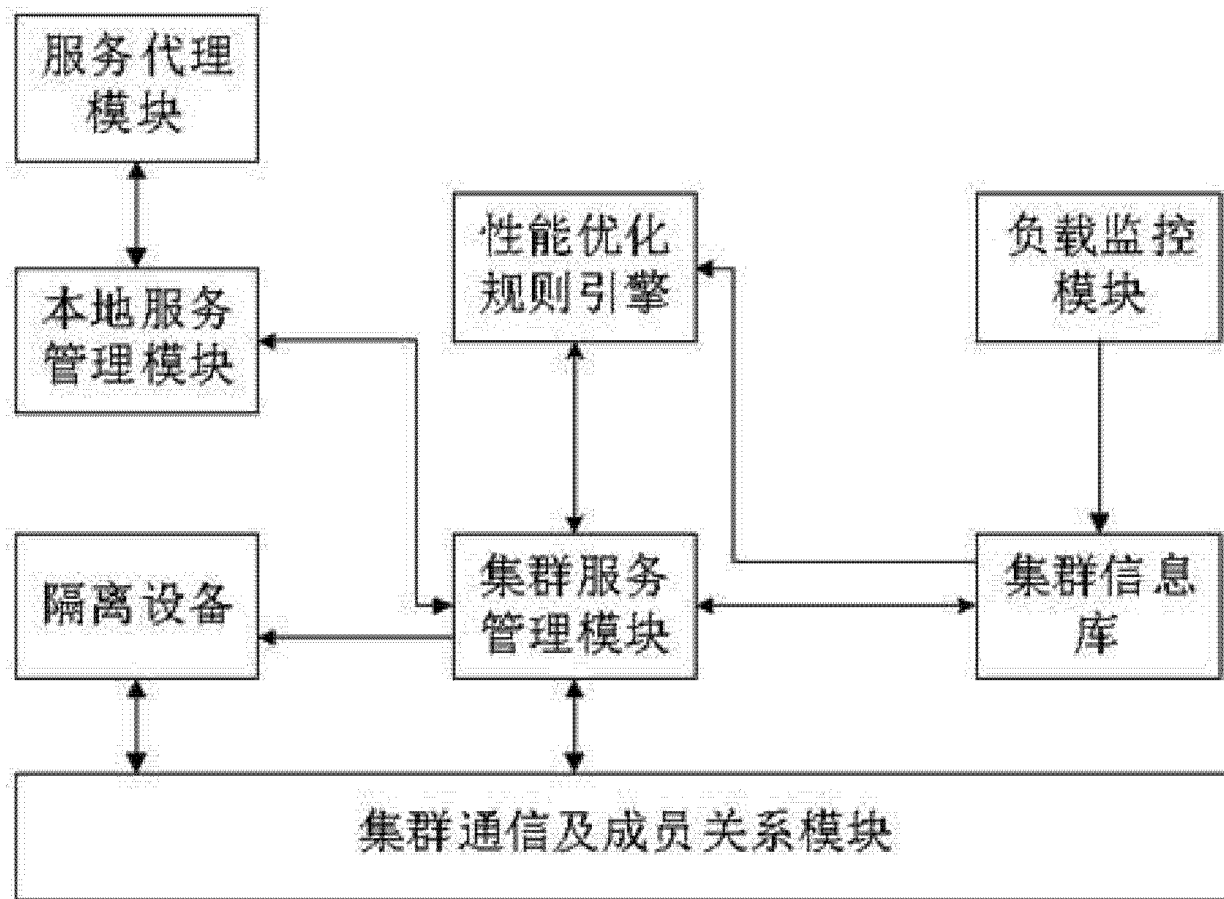


图 2

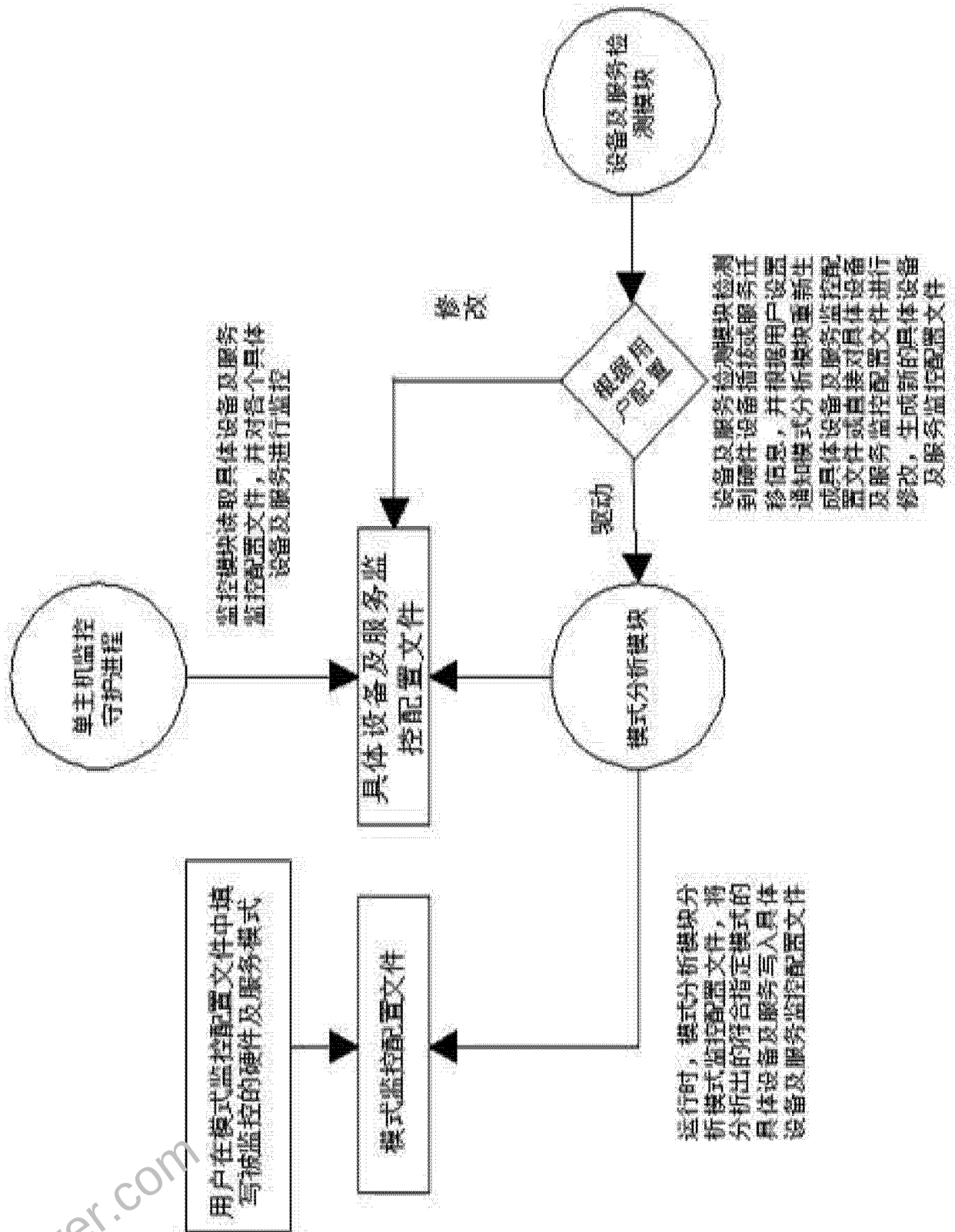


图 3

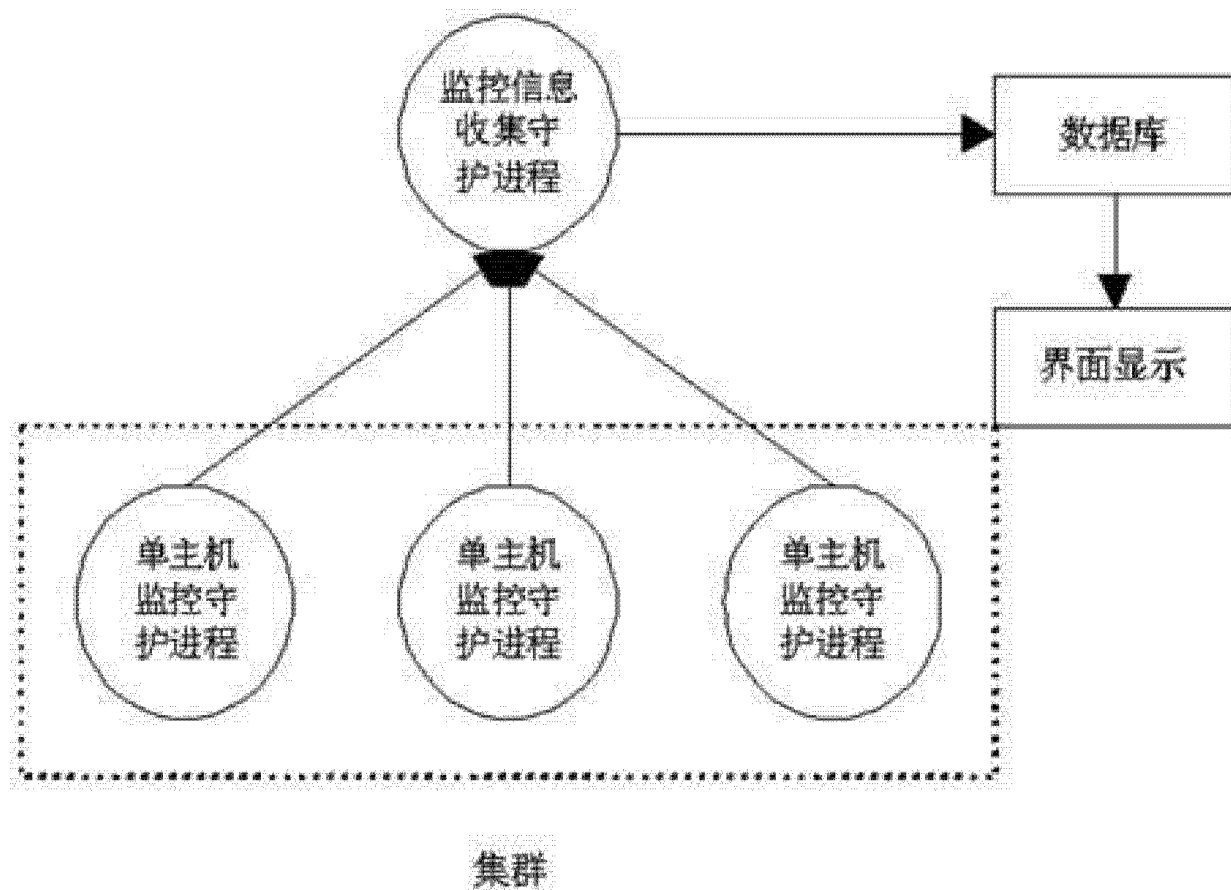


图 4