



(21) 申请号 202310395751.7

(22) 申请日 2023.04.14

(71) 申请人 上海昊帆信息技术有限公司
地址 201800 上海市嘉定区科福路358_368
号4幢1层E区J3682室

(72) 发明人 张勇

(74) 专利代理机构 池州优佐知识产权代理事务
所(普通合伙) 34198
专利代理师 钱奥

(51) Int. Cl.

H04L 41/0677 (2022.01)

H04L 41/0663 (2022.01)

H04L 41/069 (2022.01)

H04L 43/10 (2022.01)

权利要求书2页 说明书5页

(54) 发明名称

一种计算机网络集群系统故障的诊断方法

(57) 摘要

本发明涉及一种计算机网络集群系统故障的诊断方法,包括以下步骤:A、检测节点故障:
a1、在集群节点之间的通信允许集群服务检测节点故障和状态改变并且将集群作为一个实体进行管理;a2、在节点到节点之间的通信中设立多重网络以便于消除单一点的故障;a3、在集群节点上运行的集群服务跟踪集群中节点的当前状态,并且确定组及其资源何时故障转移到备用节点上。该计算机网络集群系统故障的诊断方法,通过从3个方面对计算机网络集群系统故障进行诊断,诊断方向足够全面,能够快速、精准的定位到故障节点,并且给出了相应的解决方案,使得IT新手就能够上手解决这种计算机网络集群系统故障,大幅度降低了用人门槛和用人成本。

1. 一种计算机网络集群系统故障的诊断方法,其特征在于:包括以下步骤:

A、检测节点故障:

a1、在集群节点之间的通信允许集群服务检测节点故障和状态改变并且将集群作为一个实体进行管理;

a2、在节点到节点之间的通信中设立多重网络以便于消除单一点的故障;

a3、在集群节点上运行的集群服务跟踪集群中节点的当前状态,并且确定组及其资源何时故障转移到备用节点上;

其中,步骤a3中采用的通信以在每个节点的集群服务之间定期发送消息的形式进行,并将这些消息作为心跳信号;

a5、通过心跳信号,在每个节点上监察其他节点及它们的应用程序的可用性;

a6、如果服务器设有对心跳通信并做出响应,则正常工作的服务器会启动故障转移过稳(包括对故障服务器拥有的资源和应用程序的所有权进行仲裁);

a7、如果某个节点似乎发生了故障,则会在给定的时间内允许它以几种方式中的任何一种表明它仍处于正常运行当中,并且可以向其他正常的节点通信;

a8、如果它无法证明自己正在正常运行,则此时会将它移出集群,并定位处故障点出来;

B、检测资源故障,如果资源在当前主机节点上是不可操作的而且不能重新启动,则集群服务认为此资源发生了故障,应当让集群服务按以下方式检测资源故障:

b1、在周期性的时间间隔内,集群服务检查资源以了解资源是否可操作,并且集群服务通过定期调用资源监视器来完成此项任务;

b2、反之,资源监视器则依赖于每个资源的资源主动态链接库(DLL)执行一段过程以检测资源是否正常工作;

b3、资源DLL通过资源监视器将结果传给集群服务;

b4、如果资源DLL报告资源不可操作,那么集群服务会尝试重新启动资源;

b5、另外,可以指定集群服务在一给定时间间隔内尝试重新启动资源的次数,如果集群服务在指定的时间向陆内所做的尝试的数量超过了最大重新启动尝试次数,但是资源仍不可操作,那么集群服务会认为资源出了故障;

b6、同时,可以配置是否故障资源引起包含该资源的组故障转移到其他节点;

b61、如果故障资源配肯为引起包含该资源的组故障转移到其他节点,那么集群服务将尝试故障转移;

b62、如果故障转移尝试次数坦过组的阈值而资源仍然处在故障状态,则集群服务将尝试资源的重新启动;

b7、在资源的“故障重试周期”属性指定的周期结束后,将做重新启动尝试,集群服务开始对资源尝试重新启动然后故障转移;

其中,“lsAlive”轮询问隔通常比“LooksAlive”轮询问隔长;

C、看监控,看日志:

c1、通过clouderamanager监控和管理软件本身出问题,如没有任何数据,但集群还是好的,业务还在正常跑;

c2、监控软件是好的,从监控里发现了很多问题,如CPU飙高、内存飙高;Apache网站上

由类似于clouderamanager的Ambari (web-basetool) 集群内部应用故障;

使用依赖包工具Yarn, 进而产生一个ResourceManagerlog;

通过观察, c21、看是否在所有节点都失败了, 如, 全局变量、执行的脚本、参数有没有写对, yarn组件对应mapreduce的全局配置是不是出错, 就是ResourceManager的配置;

c22、看是否某些节点失败, 如, 看在失败的节点上NodeManager的日志, 是不是作业总在同一台节点上失败, 是资源枯竭还是每个配置有问题, 若不是, 在看是不是mapreduce作业没有配置好, 如: AppMaster, Mapper, Reducer, 即可诊断出故障节点。

2. 根据权利要求1所述的一种计算机网络集群系统故障的诊断方法, 其特征在于: 所述步骤a6中的仲裁是使用质询和辩护协议来执行的。

3. 根据权利要求1所述的一种计算机网络集群系统故障的诊断方法, 其特征在于: 所述步骤b3中提到的资源监视器的具体操作为: 设置“LooksAlive”和“lsAlive, 通过轮询间隔来指定集群服务检查资源故障的频率, 利用集群服务在每个“lsAlive”时间间隔内请求对资源的状态进行比“LooksAlive”间隔内所做的更彻底的检查。

4. 根据权利要求1所述的一种计算机网络集群系统故障的诊断方法, 其特征在于: 所述步骤b7中提到的“故障重试周期”属性的单位是ms, 但按分钟顺序选择值;

另外, 还可以选择比资源重新启动周期属性值大的或者相等的值并强制执行这个规则。

一种计算机网络集群系统故障的诊断方法

技术领域

[0001] 本发明涉及计算机网络集群技术领域,具体为一种计算机网络集群系统故障的诊断方法。

背景技术

[0002] 计算机集群是一组计算机,如此一起工作使得它们可以作为一个单一的系统查看。与网格计算机不同,计算机集群将每个节点设置为执行相同的任务,由软件控制和调度。集群的组件通常通过快速局域网相互连接,每个节点(用作服务器的计算机)运行自己的操作系统实例。在大多数情况下,所有节点都使用相同的硬件和相同的操作系统,尽管在某些设置中(例如使用开源集群应用程序资源)。

[0003] 不同的操作系统可以用于每台计算机,或不同的硬件,通常部署集群是为了提高单台计算机的性能和可用性,同时通常比具有可比速度或可用性的单台计算机更具成本效益。计算机集群的出现是许多计算趋势融合的结果,包括低成本微处理器、高速网络和用于高性能分布式计算的软件的可用性,它们具有广泛的适用性和部署范围,从具有少量节点的小型企业集群到世界上一些最快的超级计算机,例如IBM的红杉。在集群出现之前,具有模块化冗余的单机容错大型机受雇;但集群的前期成本较低,网络结构的速度提高,有利于集群的采用。

[0004] 然而,与高可靠性大型机相比,集群的扩展成本更低,但错误处理的复杂性也增加了,因为集群中的错误模式对运行的程序并不透明,一旦出现故障,就需要计算机高手对故障进行排查、定位,不仅繁琐,而且用人门槛和用人成本大幅提升,因此,本申请提出了一种计算机网络集群系统故障的诊断方法。

发明内容

[0005] 针对现有技术的不足,本发明提供了一种计算机网络集群系统故障的诊断方法,具备全面、快速定位、降低用人门槛,降低用人成本等优点。

[0006] 为实现上述目的,本发明提供如下技术方案:一种计算机网络集群系统故障的诊断方法,包括以下步骤:

[0007] A、检测节点故障:

[0008] a1、在集群节点之间的通信允许集群服务检测节点故障和状态改变并且将集群作为一个实体进行管理;

[0009] a2、在节点到节点之间的通信中设立多重网络以便于消除单一点的故障;

[0010] a3、在集群节点上运行的集群服务跟踪集群中节点的当前状态,并且确定组及其资源何时故障转移到备用节点上;

[0011] 其中,步骤a3中采用的通信以在每个节点的集群服务之间定期发送消息的形式进行,并将这些消息作为心跳信号;

[0012] a5、通过心跳信号,在每个节点上监察其他节点及它们的应用程序的可用性;

[0013] a6、如果服务器设有对心跳通信并做出响应,则正常工作的服务器会启动故障转移过稳(包括对故障服务器拥有的资源和应用程序的所有权进行仲裁);

[0014] a7、如果某个节点似乎发生了故障,则会在给定的时间内允许它以几种方式中的任何一种表明它仍处于正常运行当中,并且可以向其他正常的节点通信;

[0015] a8、如果它无法证明自己正在正常运行,则此时会将它移出集群,并定位处故障点出来;

[0016] B、检测资源故障,如果资源在当前主机节点上是不可操作的而且不能重新启动,则集群服务认为此资源发生了故障,应当让集群服务按以下方式检测资源故障:

[0017] b1、在周期性的时间间隔内,集群服务检查资源以了解资源是否可操作,并且集群服务通过定期调用资源监视器来完成此项任务;

[0018] b2、反之,资源监视器则依赖于每个资源的资源主动态链接库(DLL)执行一段过程以检测资源是否正常工作;

[0019] b3、资源DLL通过资源监视器将结果传给集群服务;

[0020] b4、如果资源DLL报告资源不可操作,那么集群服务会尝试重新启动资源;

[0021] b5、另外,可以指定集群服务在一给定时间间隔内尝试重新启动资源的次数,如果集群服务在指定的时间向陆内所做的尝试的数量超过了最大重新启动尝试次数,但是资源仍不可操作,那么集群服务会认为资源出了故障;

[0022] b6、同时,可以配置是否故障资源引起包含该资源的组故障转移到其他节点;

[0023] b61、如果故障资源配肯为引起包含该资源的组故障转移到其他节点,那么集群服务将尝试故障转移;

[0024] b62、如果故障转移尝试次数坦过组的阈值而资源仍然处在故障状态,则集群服务将尝试资源的重新启动;

[0025] b7、在资源的“故障重试周期”属性指定的周期结束后,将做重新启动尝试,集群服务开始对资源尝试重新启动然后故障转移;

[0026] 其中,“ls Alive”轮询问隔通常比“Looks Alive”轮询问隔长;

[0027] C、看监控,看日志:

[0028] c1、通过cloudera manager监控和管理软件本身出问题,如没有任何数据,但集群还是好的,业务还在正常跑;

[0029] c2、监控软件是好的,从监控里发现了很多问题,如CPU飙高、内存飙高;Apache网站上由类似于cloudera manager的Ambari(web-basetool)集群内部应用故障;

[0030] 使用依赖包工具Yarn,进而产生一个Resource Manager log;

[0031] 通过观察,c21、看是否在所有节点都失败了,如,全局变量、执行的脚本、参数有没有写对,yarn组件对应map reduce的全局配置是不是出错,就是Resource Manager的配置;

[0032] c22、看是否某些节点失败,如,看在失败的节点上Node Manager的日志,是不是作业总在同一台节点上失败,是资源枯竭还是每个配置有问题,若不是,在看是不是map reduce作业没有配置好,如:App Master,Mapper,Reducer,即可诊断处故障节点。

[0033] 进一步,所述步骤a6中的仲裁是使用质询和辩护协议来执行的。

[0034] 进一步,所述步骤b3中提到的资源监视器的具体操作为:设置“Looks Alive”和“ls Alive,通过轮询问隔来指定集群服务检资资源故障的频率,利用集群服务在每个“ls

Alive”时间间隔内请求对资源的状态进行比“Looks Alive”间隔内所做的更彻底的检查。

[0035] 进一步,所述步骤b7中提到的“故障重试周期”属性的单位是ms,但按分钟顺序选择值;

[0036] 另外,还可以选择比资源重新启动周期属性值大的或者相等的值并强制执行这个规则。

[0037] 与现有技术相比,本发明提供了一种计算机网络集群系统故障的诊断方法,具备以下有益效果:

[0038] 该计算机网络集群系统故障的诊断方法,通过从3个方面对计算机网络集群系统故障进行诊断,诊断方向足够全面,能够快速、精准的定位到故障节点,并且给出了相应的解决方案,使得IT新手就能够上手解决这种计算机网络集群系统故障,大幅度降低了用人门槛和用人成本。

具体实施方式

[0039] 下面将结合本发明实施例,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。

[0040] 在本实施例中,一种计算机网络集群系统故障的诊断方法,包括以下3个大步骤:

[0041] 第一个步骤A:检测节点故障。

[0042] a1、在集群节点之间的通信允许集群服务检测节点故障和状态改变并且将集群作为一个实体进行管理;

[0043] a2、在节点到节点之间的通信中设立多重网络以便于消除单一点的故障;

[0044] a3、在集群节点上运行的集群服务跟踪集群中节点的当前状态,并且确定组及其资源何时故障转移到备用节点上;

[0045] 其中,步骤a3中采用的通信以在每个节点的集群服务之间定期发送消息的形式进行,并将这些消息作为心跳信号。

[0046] a5、通过心跳信号,在每个节点上监察其他节点及它们的应用程序的可用性;

[0047] a6、如果服务器设有对心跳通信并做出响应,则正常工作的服务器会启动故障转移过稳(包括对故障服务器拥有的资源和应用程序的所有权进行仲裁);

[0048] 其中,仲裁是使用质询和辩护协议来执行的。

[0049] a7、如果某个节点似乎发生了故障,则会在给定的时间内允许它以几种方式中的任何一种表明它仍处于正常运行当中,并且可以向其他正常的节点通信;

[0050] a8、如果它无法证明自己正在正常运行,则此时会将它移出集群,并定位处故障点出来。

[0051] 需要注意的是,在集群中,有多种事件都可能导致节点无法响应心跳消息。如计算机故障、网络接口故障、网络故障,甚至可能是由于少有的高峰活动期。

[0052] 因此,当所有节点进行通信时,配置数据库管理器会向每个节点发送全局性的配置数据库更新。而当心跳通信失败时,日志管理锦还会将配置数据库的变更保存到仲裁资源中。这保证了幸存的节点可以在恢复过程中访问最新的集群配置和本地节点的注册表数

据。

[0053] 另外要注意的是,采用故障检测算法相当保守。

[0054] 因此,它会尽量多地给那些明显发生故障的节点以质询的机会,然后才会进入故障转移过程。如果导致心跳响应失败的原因是暂时的,避免故障转移所可能造成的潜在影响当然是再好不过了。

[0055] 但是,由于无法知道这样的节点还将沉默多长时间,故该节点可能遭受长时期的故障影响。因此,在经过一个合理的时间段后就应该启动故障转移过程。

[0056] 第二个步骤B:检测资源故障,如果资源在当前主机节点上是不可操作的而且不能重新启动,则集群服务认为此资源发生了故障,应当让集群服务按以下方式检测资源故障。

[0057] b1、在周期性的时间间隔内,集群服务检查资源以了解资源是否可操作,并且集群服务通过定期调用资源监视器来完成此项任务;

[0058] b2、反之,资源监视器则依赖于每个资源的资源主动态链接库(DLL)执行一段过程以检测资源是否正常工作;

[0059] b3、资源DLL通过资源监视器将结果传给集群服务;

[0060] b4、如果资源DLL报告资源不可操作,那么集群服务会尝试重新启动资源;

[0061] b5、另外,可以指定集群服务在一给定时间间隔内尝试重新启动资源的次数,如果集群服务在指定的时间向陆内所做的尝试的数量超过了最大重新启动尝试次数,但是资源仍不可操作,那么集群服务会认为资源出了故障;

[0062] b6、同时,可以配置是否故障资源引起包含该资源的组故障转移到其他节点;

[0063] b61、如果故障资源配肯为引起包含该资源的组故障转移到其他节点,那么集群服务将尝试故障转移;

[0064] b62、如果故障转移尝试次数坦过组的阈值而资源仍然处在故障状态,则集群服务将尝试资源的重新启动。

[0065] b7、在资源的“故障重试周期”属性指定的周期结束后,将做重新启动尝试,集群服务开始对资源尝试重新启动然后故障转移。

[0066] 其中,资源监视器的具体操作为:设置“Looks Alive”和“ls Alive,通过轮询间隔来指定集群服务检资资源故障的频率,利用集群服务在每个“ls Alive”时间间隔内请求对资源的状态进行比“Looks Alive”间隔内所做的更彻底的检查。并且,需要额外注意的是,“ls Alive”轮询问隔通常比“Looks Alive”轮询问隔长。

[0067] 而且“故障重试周期”属性的单位是ms,但是却是按分钟顺序选择值;另外,还可以选择比资源重新启动周期属性值大的或者相等的值并强制执行这个规则。

[0068] 第三个步骤C:看监控,看日志

[0069] c1、通过cloudera manager监控和管理软件本身出问题,如没有任何数据,但集群还是好的,业务还在正常跑;

[0070] c2、监控软件是好的,从监控里发现了很多问题,如CPU飙高、内存飙高;Apache网站上由类似于cloudera manager的Ambari(web-basetool)集群内部应用故障;

[0071] 使用依赖包工具Yarn,进而产生一个Resource Manager log;

[0072] 通过观察,c21、看是否在所有节点都失败了,如,全局变量、执行的脚本、参数有没有写对,yarn组件对应map reduce的全局配置是不是出错,就是Resource Manager的配置;

[0073] c22、看是否某些节点失败,如,看在失败的节点上Node Manager的日志,是不是作业总在同一台节点上失败,是资源枯竭还是每个配置有问题,若不是,在看是不是map reduce作业没有配置好,如:App Master,Mapper,Reducer,即可诊断故障节点。

[0074] 这个步骤的中心思想,就是看监控和日志。

[0075] 本实施例在使用时,通过从3个方面对计算机网络集群系统故障进行诊断,诊断方向足够全面,能够快速、精准的定位到故障节点,并且给出了相应的解决方案,使得IT新手就能够上手解决这种计算机网络集群系统故障,大幅度降低了用人门槛和用人成本。

[0076] 文中出现的电器元件均与主控器及电源电连接,主控器可为计算机等起到控制的常规已知设备,且现有公开的电力连接技术,不在文中赘述。

[0077] 需要说明的是,在本文中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,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。

[0078] 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物限定。