

(19) 中华人民共和国国家知识产权局



(12) 发明专利申请

(10) 申请公布号 CN 105577414 A

(43) 申请公布日 2016. 05. 11

(21) 申请号 201410555238. 0

(22) 申请日 2014. 10. 17

(71) 申请人 中兴通讯股份有限公司

地址 518057 广东省深圳市南山区科技南路
55 号

(72) 发明人 陈露静 周智伟 刘磊

(74) 专利代理机构 北京康信知识产权代理有限
责任公司 11240

代理人 余刚 梁丽超

(51) Int. Cl.

H04L 12/24(2006. 01)

权利要求书2页 说明书8页 附图5页

(54) 发明名称

虚拟机 VM 资源弹性伸缩处理方法、装置及系
统

(57) 摘要

本发明提供了一种虚拟机 VM 资源弹性伸缩
处理方法、装置及系统，其中，该方法包括：由第
一节点或第一节点的相邻上游节点根据第一节点
中的 VM 的指标信息确定第一节点是否达到了进
行 VM 资源弹性伸缩处理的预定阈值；利用第一节
点的上游节点对第一节点进行 VM 资源弹性伸缩
处理。解决了相关技术中存在的对 NFVO 信息交互过
多，冲击大的问题，进而达到了减少与 NFVO 的信
息交互，降低对该 NFVO 的冲击的效果。

由第一节点或第一节点的相邻上游节点根据第一节点中的V
M的指标信息确定
第一节点达到了进行VM资源弹性伸缩处理的预定阈值

S202

利用第一节点的上游节点对第一节点进行VM资源弹性伸缩
处理

S204

1. 一种虚拟机 VM 资源弹性伸缩处理方法, 其特征在于, 包括 :

由第一节点或第一节点的相邻上游节点根据所述第一节点中的 VM 的指标信息确定所述第一节点达到了进行 VM 资源弹性伸缩处理的预定阈值 ;

利用所述第一节点的上游节点对所述第一节点进行所述 VM 资源弹性伸缩处理。

2. 根据权利要求 1 所述的方法, 其特征在于, 所述 VM 的指标信息包括以下信息至少之一 :

内存信息、中央处理器 CPU 信息、硬盘信息、宽带信息、所述 VM 所承载的业务指标信息、所述 VM 的处理能力。

3. 根据权利要求 1 所述的方法, 其特征在于, 当所述第一节点为 VM 时, 利用所述第一节点的上游节点对所述第一节点进行所述 VM 资源弹性伸缩处理包括 :

所述 VM 向虚拟化网络功能 VNF 发送用于请求进行所述 VM 资源弹性伸缩处理的第一请求消息 ;

所述 VNF 根据所述第一请求消息向虚拟化网络功能管理 VNFM 发送用于请求全局的资源访问接口进行预申请资源的第二请求消息 ;

所述 VNFM 根据所述第二请求消息向网络功能虚拟化编排 NFVO 发送 VM 资源变更请求 ;

所述 NFVO 根据所述 VM 资源变更请求对所述 VM 进行所述 VM 资源弹性伸缩处理。

4. 根据权利要求 1 所述的方法, 其特征在于, 当所述第一节点为虚拟化网络功能 VNF 时, 利用所述第一节点的上游节点对所述第一节点进行所述 VM 资源弹性伸缩处理包括 :

所述 VNF 向虚拟化网络功能管理 VNFM 发送用于请求全局的资源访问接口进行预申请资源的第二请求消息 ;

所述 VNFM 根据所述第二请求消息向网络功能虚拟化编排 NFVO 发送 VM 资源变更请求 ;

所述 NFVO 根据所述 VM 资源变更请求对所述 VNF 进行所述 VM 资源弹性伸缩处理。

5. 根据权利要求 1 所述的方法, 其特征在于, 当所述第一节点为虚拟化网络功能管理 VNFM 时, 利用所述第一节点的上游节点对所述第一节点进行所述 VM 资源弹性伸缩处理包括 :

所述 VNFM 向网络功能虚拟化编排 NFVO 发送用于请求全局的资源访问接口进行预申请资源的第三请求消息 ;

在进行所述预申请资源申请成功的情况下, 所述 VNFM 向所述 NFVO 发送 VM 资源变更请求 ;

所述 NFVO 根据所述 VM 资源变更请求对所述 VNFM 进行所述 VM 资源弹性伸缩处理。

6. 一种虚拟机 VM 资源弹性伸缩处理装置, 其特征在于, 包括 :

确定模块, 位于第一节点或第一节点的相邻上游节点中, 用于根据所述第一节点中的 VM 的指标信息确定所述第一节点达到了进行 VM 资源弹性伸缩处理的预定阈值 ;

处理模块, 位于所述第一节点的各个上游节点中, 用于对所述第一节点进行所述 VM 资源弹性伸缩处理。

7. 根据权利要求 6 所述的装置, 其特征在于, 所述处理模块包括 :

第一发送单元, 位于 VM 中, 用于当所述第一节点为所述 VM 时, 向虚拟化网络功能 VNF 发送用于请求进行所述 VM 资源弹性伸缩处理的第一请求消息 ;

第二发送单元, 位于所述 VNF 中, 用于根据所述第一请求消息向虚拟化网络功能管理

VNFM 发送用于请求全局的资源访问接口进行预申请资源的第二请求消息；

第三发送单元，位于所述 VNFM 中，用于根据所述第二请求消息向网络功能虚拟化编排 NFVO 发送 VM 资源变更请求；

第一处理单元，位于所述 NFVO 中，用于根据所述 VM 资源变更请求对所述 VM 进行所述 VM 资源弹性伸缩处理。

8. 根据权利要求 6 所述的装置，其特征在于，所述处理模块包括：

第四发送单元，位于虚拟化网络功能 VNF 中，用于当所述第一节点为所述 VNF 时，利用所述 VNF 向虚拟化网络功能管理 VNFM 发送用于请求全局的资源访问接口进行预申请资源的第二请求消息；

第五发送单元，位于所述 VNFM 中，用于根据所述第二请求消息向网络功能虚拟化编排 NFVO 发送 VM 资源变更请求；

第二处理单元，位于所述 NFVO 中，用于根据所述 VM 资源变更请求对所述 VNF 进行所述 VM 资源弹性伸缩处理。

9. 根据权利要求 6 所述的装置，其特征在于，所述处理模块包括：

第六发送单元，位于虚拟化网络功能管理 VNFM 中，用于当所述第一节点为所述 VNFM 时，利用所述 VNFM 向网络功能虚拟化编排 NFVO 发送用于请求全局的资源访问接口进行预申请资源的第三请求消息；

第七发送单元，位于所述 VNFM 中，用于在进行所述预申请资源申请成功的情况下，向所述 NFVO 发送 VM 资源变更请求；

第三处理单元，位于所述 NFVO 中，用于根据所述 VM 资源变更请求对所述 VNFM 进行所述 VM 资源弹性伸缩处理。

10. 一种虚拟机 VM 资源弹性伸缩处理系统，其特征在于，包括权利要求 6 至 9 中任一项所述的装置。

虚拟机 VM 资源弹性伸缩处理方法、装置及系统

技术领域

[0001] 本发明涉及通信领域，具体而言，涉及一种虚拟机 VM 资源弹性伸缩处理方法、装置及系统。

背景技术

[0002] 在当前通讯领域，如何在电信设备中构建高效可靠的弹性伸缩服务是业内面临的重要问题和核心难题。

[0003] 目前大多数的解决方案在安全性和可靠性上均不是很完善，特别是虚拟机 (Virtual Machine，简称 VM) 的弹性伸缩指标监控部分，目前对虚拟机指标信息的监控多以集中式的单点监控方式为主，其中监控节点多放在上游，多以网络功能虚拟化编排 (Network Functions Virtualisation Orchestrator，简称 NFVO) 为主，这种情况对 NFVO 的冲击过大，要是上游节点出现问题，就很容易会引发单点故障。

[0004] 图 1 是相关技术中的资源控制结构图，并且，目前的监控方式一般都是虚拟机将本机的指标信息定时采集上报给上游节点，然后由上游节点判断是否达到了伸缩规则的阈值，此种方式的优点是对集中式的批量管理等比较方便，但是缺点也很明显，比如：

[0005] 1、消息在传送过程中出现丢包或者网络阻塞时将无法进行实时监控。

[0006] 2、将本机采集的指标信息交由上游节点进行伸缩判断的话，在精确性上也会有问题。

[0007] 针对相关技术中存在的对 NFVO 信息交互过多，冲击大的问题，目前尚未提出有效的解决方案。

发明内容

[0008] 本发明提供了一种虚拟机 VM 资源弹性伸缩处理方法、装置及系统，以至少解决相关技术中存在的对 NFVO 信息交互过多，冲击大的问题。

[0009] 根据本发明的一个方面，提供了一种虚拟机 VM 资源弹性伸缩处理方法，包括：由第一节点或第一节点的相邻上游节点根据所述第一节点中的 VM 的指标信息确定所述第一节点达到了进行 VM 资源弹性伸缩处理的预定阈值；利用所述第一节点的上游节点对所述第一节点进行所述 VM 资源弹性伸缩处理。

[0010] 优选地，所述 VM 的指标信息包括以下信息至少之一：内存信息、中央处理器 CPU 信息、硬盘信息、宽带信息、所述 VM 所承载的业务指标信息、所述 VM 的处理能力。

[0011] 优选地，当所述第一节点为 VM 时，利用所述第一节点的上游节点对所述第一节点进行所述 VM 资源弹性伸缩处理包括：所述 VM 向虚拟化网络功能 VNF 发送用于请求进行所述 VM 资源弹性伸缩处理的第一请求消息；所述 VNF 根据所述第一请求消息向虚拟化网络功能管理 VNFM 发送用于请求全局的资源访问接口进行预申请资源的第二请求消息；所述 VNFM 根据所述第二请求消息向网络功能虚拟化编排 NFVO 发送 VM 资源变更请求；所述 NFVO 根据所述 VM 资源变更请求对所述 VM 进行所述 VM 资源弹性伸缩处理。

[0012] 优选地,当所述第一节点为虚拟化网络功能 VNF 时,利用所述第一节点的上游节点对所述第一节点进行所述 VM 资源弹性伸缩处理包括:所述 VNF 向虚拟化网络功能管理 VNFM 发送用于请求全局的资源访问接口进行预申请资源的第二请求消息;所述 VNFM 根据所述第二请求消息向网络功能虚拟化编排 NFVO 发送 VM 资源变更请求;所述 NFVO 根据所述 VM 资源变更请求对所述 VNF 进行所述 VM 资源弹性伸缩处理。

[0013] 优选地,当所述第一节点为虚拟化网络功能管理 VNFM 时,利用所述第一节点的上游节点对所述第一节点进行所述 VM 资源弹性伸缩处理包括:所述 VNFM 向网络功能虚拟化编排 NFVO 发送用于请求全局的资源访问接口进行预申请资源的第三请求消息;在进行所述预申请资源申请成功的情况下,所述 VNFM 向所述 NFVO 发送 VM 资源变更请求;所述 NFVO 根据所述 VM 资源变更请求对所述 VNFM 进行所述 VM 资源弹性伸缩处理。

[0014] 根据本发明的另一方面,提供了一种虚拟机 VM 资源弹性伸缩处理装置,包括:确定模块,位于第一节点或第一节点的相邻上游节点中,用于根据所述第一节点中的 VM 的指标信息确定所述第一节点达到了进行 VM 资源弹性伸缩处理的预定阈值;处理模块,位于所述第一节点的各个上游节点中,用于对所述第一节点进行所述 VM 资源弹性伸缩处理。

[0015] 优选地,所述处理模块包括:第一发送单元,位于 VM 中,用于当所述第一节点为所述 VM 时,向虚拟化网络功能 VNF 发送用于请求进行所述 VM 资源弹性伸缩处理的第一请求消息;第二发送单元,位于所述 VNF 中,用于根据所述第一请求消息向虚拟化网络功能管理 VNFM 发送用于请求全局的资源访问接口进行预申请资源的第二请求消息;第三发送单元,位于所述 VNFM 中,用于根据所述第二请求消息向网络功能虚拟化编排 NFVO 发送 VM 资源变更请求;第一处理单元,位于所述 NFVO 中,用于根据所述 VM 资源变更请求对所述 VM 进行所述 VM 资源弹性伸缩处理。

[0016] 优选地,所述处理模块包括:第四发送单元,位于虚拟化网络功能 VNF 中,用于当所述第一节点为所述 VNF 时,利用所述 VNF 向虚拟化网络功能管理 VNFM 发送用于请求全局的资源访问接口进行预申请资源的第二请求消息;第五发送单元,位于所述 VNFM 中,用于根据所述第二请求消息向网络功能虚拟化编排 NFVO 发送 VM 资源变更请求;第二处理单元,位于所述 NFVO 中,用于根据所述 VM 资源变更请求对所述 VNF 进行所述 VM 资源弹性伸缩处理。

[0017] 优选地,所述处理模块包括:第六发送单元,位于虚拟化网络功能管理 VNFM 中,用于当所述第一节点为所述 VNFM 时,利用所述 VNFM 向网络功能虚拟化编排 NFVO 发送用于请求全局的资源访问接口进行预申请资源的第三请求消息;第七发送单元,位于所述 VNFM 中,用于在进行所述预申请资源申请成功的情况下,向所述 NFVO 发送 VM 资源变更请求;第三处理单元,位于所述 NFVO 中,用于根据所述 VM 资源变更请求对所述 VNFM 进行所述 VM 资源弹性伸缩处理。

[0018] 根据本发明的再一方面,提供了一种系统,包括上述任一项所述的虚拟机 VM 资源弹性伸缩处理装置。

[0019] 通过本发明,采用由第一节点或第一节点的相邻上游节点根据所述第一节点中的 VM 的指标信息确定所述第一节点达到了进行 VM 资源弹性伸缩处理的预定阈值;利用所述第一节点的上游节点对所述第一节点进行所述 VM 资源弹性伸缩处理,解决了相关技术中存在的对 NFVO 信息交互过多,冲击大的问题,进而达到了减少与 NFVO 的信息交互,降低对

该 NFVO 的冲击的效果。

附图说明

- [0020] 此处所说明的附图用来提供对本发明的进一步理解,构成本申请的一部分,本发明的示意性实施例及其说明用于解释本发明,并不构成对本发明的不当限定。在附图中:
- [0021] 图 1 是相关技术中的资源控制结构图;
- [0022] 图 2 是根据本发明实施例的虚拟机 VM 资源弹性伸缩处理方法的流程图;
- [0023] 图 3 是根据本发明实施例的虚拟机 VM 资源弹性伸缩处理装置的结构框图;
- [0024] 图 4 是根据本发明实施例的虚拟机 VM 资源弹性伸缩处理装置中处理模块 34 的结构框图一;
- [0025] 图 5 是根据本发明实施例的虚拟机 VM 资源弹性伸缩处理装置中处理模块 34 的结构框图二;
- [0026] 图 6 是根据本发明实施例的虚拟机 VM 资源弹性伸缩处理装置中处理模块 34 的结构框图三;
- [0027] 图 7 是根据本发明实施例的系统的结构框图;
- [0028] 图 8 是根据本发明实施例的 VM 弹性伸缩过程图;
- [0029] 图 9 是根据本发明实施例的 VNF 弹性收缩过程图;
- [0030] 图 10 是根据本发明实施例的 VNFM 弹性伸缩过程图。

具体实施方式

- [0031] 下文中将参考附图并结合实施例来详细说明本发明。需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。
- [0032] 在本实施例中提供了一种虚拟机 VM 资源弹性伸缩处理方法,图 2 是根据本发明实施例的虚拟机 VM 资源弹性伸缩处理方法的流程图,如图 2 所示,该流程包括如下步骤:
- [0033] 步骤 S202,由第一节点或第一节点的相邻上游节点根据第一节点中的 VM 的指标信息确定第一节点达到了进行 VM 资源弹性伸缩处理的预定阈值;
- [0034] 步骤 S204,利用第一节点的上游节点对第一节点进行 VM 资源弹性伸缩处理。
- [0035] 通过上述步骤,由第一节点或第一节点的相邻上游节点根据第一节点中的 VM 的指标信息确定第一节点达到了进行 VM 资源弹性伸缩处理的预定阈值;利用第一节点的上游节点对第一节点进行 VM 资源弹性伸缩处理,实现了由节点本身或节点的相邻上游节点对节点进行状态监控,减少了对 NFVO 节点的冲击,解决了相关技术中存在的对 NFVO 信息交互过多,冲击大的问题,进而达到了减少与 NFVO 的信息交互,降低对该 NFVO 的冲击的效果。
- [0036] 其中,上述 VM 的指标信息包括以下信息至少之一:内存信息、中央处理器 CPU 信息、硬盘信息、宽带信息、VM 所承载的业务指标信息、VM 的处理能力。
- [0037] 在一个优选的实施例中,当第一节点为 VM 时,利用第一节点的上游节点对第一节点进行 VM 资源弹性伸缩处理包括:VM 向虚拟化网络功能 VNF 发送用于请求进行 VM 资源弹性伸缩处理的第一请求消息;VNF 根据第一请求消息向虚拟化网络功能管理 VNFM 发送用于请求全局的资源访问接口进行预申请资源的第二请求消息;VNFM 根据第二请求消息向网络功能虚拟化编排 NFVO 发送 VM 资源变更请求;NFVO 根据 VM 资源变更请求对 VM 进行 VM 资

源弹性伸缩处理。从而实现了在对 VM 进行 VM 资源弹性伸缩处理时,极大的减少了和 NFVO 之间的信息交互,降低了对 NFVO 的冲击。

[0038] 在一个优选的实施例中,当第一节点为虚拟化网络功能 VNF 时,利用第一节点的上游节点对第一节点进行 VM 资源弹性伸缩处理包括:VNF 向虚拟化网络功能管理 VNFM 发送用于请求全局的资源访问接口进行预申请资源的第二请求消息;VNFM 根据第二请求消息向网络功能虚拟化编排 NFVO 发送 VM 资源变更请求;NFVO 根据 VM 资源变更请求对 VNF 进行 VM 资源弹性伸缩处理。从而实现了在对 VNF 进行 VM 资源弹性伸缩处理时,极大的减少了和 NFVO 之间的信息交互,降低了对 NFVO 的冲击。

[0039] 在一个优选的实施例中,当第一节点为虚拟化网络功能管理 VNFM 时,利用第一节点的上游节点对第一节点进行 VM 资源弹性伸缩处理包括:VNFM 向网络功能虚拟化编排 NFVO 发送用于请求全局的资源访问接口进行预申请资源的第三请求消息;在进行预申请资源申请成功的情况下, VNFM 向 NFVO 发送 VM 资源变更请求;NFVO 根据 VM 资源变更请求对 VNFM 进行 VM 资源弹性伸缩处理。从而实现了在对 VNFM 进行 VM 资源弹性伸缩处理时,极大的减少了和 NFVO 之间的信息交互,降低了对 NFVO 的冲击。

[0040] 在本实施例中还提供了一种虚拟机 VM 资源弹性伸缩处理装置,该装置用于实现上述实施例及优选实施方式,已经进行过说明的不再赘述。如以下所使用的,术语“模块”可以实现预定功能的软件和 / 或硬件的组合。尽管以下实施例所描述的装置较佳地以软件来实现,但是硬件,或者软件和硬件的组合的实现也是可能并被构想的。

[0041] 图 3 是根据本发明实施例的虚拟机 VM 资源弹性伸缩处理装置的结构框图,如图 3 所示,该装置包括确定模块 32 和处理模块 34。下面对该装置进行说明。

[0042] 确定模块 32,位于第一节点或第一节点的相邻上游节点中,用于根据第一节点中的 VM 的指标信息确定第一节点达到了进行 VM 资源弹性伸缩处理的预定阈值;处理模块 34,位于第一节点的各个上游节点中,连接至上述确定模块 32,用于对第一节点进行 VM 资源弹性伸缩处理。其中,上述 VM 的指标信息可以包括以下信息至少之一:内存信息、中央处理器 CPU 信息、硬盘信息、宽带信息、VM 所承载的业务指标信息、VM 的处理能力。

[0043] 图 4 是根据本发明实施例的虚拟机 VM 资源弹性伸缩处理装置中处理模块 34 的结构框图一,如图 4 所示,该处理模块 34 包括第一发送单元 42、第二发送单元 44、第三发送单元 46 和第一处理单元 48。下面对该处理模块 34 进行说明。

[0044] 第一发送单元 42,位于 VM 中,用于当第一节点为 VM 时,向虚拟化网络功能 VNF 发送用于请求进行 VM 资源弹性伸缩处理的第一请求消息;第二发送单元 44,位于 VNF 中,连接至上述第一发送单元 42,用于根据第一请求消息向虚拟化网络功能管理 VNFM 发送用于请求全局的资源访问接口进行预申请资源的第二请求消息;第三发送单元 46,位于 VNFM 中,连接至上述第二发送单元 44,用于根据第二请求消息向网络功能虚拟化编排 NFVO 发送 VM 资源变更请求;第一处理单元 48,位于 NFVO 中,连接至上述第三发送单元 46,用于根据 VM 资源变更请求对 VM 进行 VM 资源弹性伸缩处理。

[0045] 图 5 是根据本发明实施例的虚拟机 VM 资源弹性伸缩处理装置中处理模块 34 的结构框图二,如图 5 所示,该处理模块 34 包括第四发送单元 52、第五发送单元 54、和第二处理单元 56。下面对该处理模块 34 进行说明。

[0046] 第四发送单元 52,位于虚拟化网络功能 VNF 中,用于当第一节点为 VNF 时,利用

VNF 向虚拟化网络功能管理 VNFM 发送用于请求全局的资源访问接口进行预申请资源的第二请求消息；第五发送单元 54，位于 VNFM 中，连接至上述第四发送单元 52，用于根据第二请求消息向网络功能虚拟化编排 NFVO 发送 VM 资源变更请求；第二处理单元 56，位于 NFVO 中，连接至上述第五发送单元 54，用于根据 VM 资源变更请求对 VM 进行 VM 资源弹性伸缩处理。

[0047] 图 6 是根据本发明实施例的虚拟机 VM 资源弹性伸缩处理装置中处理模块 34 的结构框图三，如图 6 所示，该处理模块 34 包括第六发送单元 62、第七发送单元 64、和第三处理单元 66。下面对该处理模块 34 进行说明。

[0048] 第六发送单元 62，位于虚拟化网络功能管理 VNFM 中，用于当第一节点为 VNFM 时，利用 VNFM 向网络功能虚拟化编排 NFVO 发送用于请求全局的资源访问接口进行预申请资源的第三请求消息；第七发送单元 64，位于 VNFM 中，连接至上述第六发送单元 62，用于在进行预申请资源申请成功的情况下，向 NFVO 发送 VM 资源变更请求；第三处理单元 66，位于 NFVO 中，连接至上述第七发送单元 64，用于根据 VM 资源变更请求对 VM 进行 VM 资源弹性伸缩处理。

[0049] 图 7 是根据本发明实施例的虚拟机 VM 资源弹性伸缩处理系统的结构框图，如图 7 所示，该虚拟机 VM 资源弹性伸缩处理系统 72 包括上述任一项的虚拟机 VM 资源弹性伸缩处理装置 74。

[0050] 为了解决相关技术中存在的对 NFVO 信息交互过多，冲击大的问题，在本发明实施例中还提供了一种资源按需供给和动态管理的方法，通过本发明实施例中的方法可以减轻对上游节点 NFVO 的冲击，有效消除单点故障，并减少消息交互，同时让采集的指标信息更加精确，从而实现了构建高效、安全可靠的弹性伸缩服务的目的。

[0051] 在该实施例中通过使原来统一管理虚拟资源的 NFVO 节点下沉，通过权限下放分层管理的方式，来减小对 NFVO 的冲击，并提供资源管理接口给 VNFM，使其也能访问预占资源、VM 状态、VM 已使用资源等全局虚拟资源信息，从而不用每次都通过 NFVO 访问。

[0052] 优选实施例一当上述第一节点为 VM 时的 VM 资源弹性伸缩处理：

[0053] 最下游的 VM 资源创建好弹性伸缩规则后，各个 VM 是通过服务器负载均衡 (Server Load Balancing，简称为 SLB) 的方式绑定在一起组成的虚拟化网络功能 (Virtualised Network Function，简称为 VNF)，多指网元侧，相对于监控 VNF 层级的指标信息而言，监控 VNF 下每个 VM 的指标信息显得更为精确和及时，这里的 VM 监控指标主要由内存、CPU、硬盘、带宽等系统信息组成，通过每个 VM 自身进行实时状态监控，并由自身判断是否达到弹性伸缩条件，若是 VM 自身无法进行判断是否达到了弹性伸缩规则的条件，则将监控的信息数据交由上游的 VNF 节点进行判断。若是达到伸缩条件，则发送第一请求消息给 VNF 进行弹性伸缩处理，并发送请求给虚拟化网络功能管理 (Virtualised Network Function Manager，简称为 VNFM)，当系统整体负荷较低的情况下，如果达到了缩容条件，则由 VNF 发送第二请求消息给 VNFM，并触发弹性伸缩请求，VNFM 发送 VM 资源请求变更消息给 NFVO，通过对 VM 进行一系列状态迁移操作，释放 VM 资源，然后对 VNF 进行 SLB 处理，降低系统总体拥有成本 (Total Cost of Ownership，简称为 TCO)；当系统整体负荷较高的情况下，如果达到了扩容条件，则由 VNF 发送第二请求消息给 VNFM，VNFM 发送 VM 资源请求变更消息给 NFVO，并根据 VNF 绑定的 VM 配置，通过对 VM 进行一系列状态迁移操作，启动相应的 VM，然后对 VNF 进行 SLB 处理，以此扩展、扩大系统处理能力。

[0054] 综上可知, NFVO 配合 VNFM、VNF 完成 VM 的弹性伸缩。该过程由 VM 发起, VNFM 收到 VNF 的创建请求后, 计算所需修改的资源列表, 向 NFVO 申请分配资源, 完成资源的预占和分配, VNFM 启动新增 VM 并向 VNF 通知完成。图 8 是根据本发明实施例的 VM 弹性伸缩过程图, 如图 8 所示, 该过程包括如下步骤:

[0055] 步骤 S802: 云数据中心平台部署完成, VM 资源创建好弹性伸缩规则, 虚拟网络部署完成, VM 信息采集完成;

[0056] 步骤 S804: VM 自身进行状态监控以及弹性判断, 若 VM 自身无法进行判断是否达到了弹性伸缩规则的条件, 则将监控的信息数据交由上游的 VNF 节点进行判断;

[0057] 步骤 S806: 若达到伸缩条件, 则由 VM 发送第一请求消息给 VNF 进行弹性伸缩处理;

[0058] 步骤 S808: VNF 发送请求对 VM 资源进行弹性伸缩处理的弹性伸缩请求给 VNFM;

[0059] 步骤 S810: VNFM 向 NFVO 发送 VM 资源变更请求, 其中, VNFM 解析 VNF 发起的创建 VM 请求, 请求中携带 VM 名称和操作、以及创建 VM 的个数;

[0060] 如果 VM 已经存在, 当 VM 处于异常状态时, 直接上报错误;

[0061] 如果 VM 已经存在, 当 VM 处于运行状态时, 直接上报成功;

[0062] 如果 VM 已经存在, VM 处于其他正常状态, 则 VNFM 启动 VM 到正常运行状态, 然后上报成功;

[0063] 如果 VM 不存在, 则进入下一步;

[0064] VNFM 向 NFVO 请求全局的资源访问接口进行预申请资源, 资源申请成功上报消息并进入下一步。若 NFVO 无应答超时, 则流程终止并上报错误; 如果预申请资源失败, 则上报失败, 流程终止; VNFM 发送资源分配请求给 NFVO, 资源分配成功上报消息并进入下一步; 若 NFVO 无应答超时, 则流程终止并上报错误; 如果资源分配失败, 则上报失败, 流程终止;

[0065] 步骤 S812: NFVO 成功创建 VM, 最终使 VM 达到初始状态, VNFM 生命周期业务向 VNF 发送启动创建的 VM; 若启动 VM 失败, 则上报错误; VNFM 上报消息返回 VNF 创建结果, 携带操作动作和 VM 名称, 以及每个 VM 的操作结果; VNF 对新增 VM 进行 SLB 处理。

[0066] 优选实施例二当上述第一节点为 VNF 时的 VM 资源弹性伸缩处理

[0067] VNF 节点除了有可能需要判断下游的 VM 指标信息外, 还需要同时判断并监控自身的指标信息, 这里的监控指标主要涉及内存、CPU、硬盘、带宽等系统信息以及网元侧所承载的业务指标信息, 比如每秒事务处理次数 (Transactions Per Second, 简称为 TPS), 某种业务的负荷情况等, 若是 VNF 自身无法进行判断是否达到了弹性伸缩规则的条件, 则将监控的所有信息数据交由上游的 VNFM 节点进行判断, 若是达到伸缩条件, 则由 VNF 进行弹性伸缩处理, 并发送相关请求给 VNFM, 当系统整体负荷较低的情况下, 如果达到了缩容条件, 则发送第二请求消息给 VNFM, 并触发弹性伸缩请求, VNFM 发送 VM 资源变更请求给 NFVO, 通过对 VM 进行一系列状态迁移操作, 释放 VM 资源, 然后对 VNF 进行 SLB 处理, 降低系统 TCO; 当系统整体负荷较高的情况下, 如果达到了扩容条件, 则发送第二请求消息给 VNFM, VNFM 然后发送 VM 资源变更请求给 NFVO, 并根据 VNF 绑定的 VM 配置, 通过对 VM 进行一系列状态迁移操作, 启动相应的 VM, 然后对 VNF 进行 SLB 处理, 以此扩展、扩大系统处理能力。

[0068] 由上述可知, NFVO 配合 VNFM 完成 VM 的弹性伸缩。该过程由 VM 发起, VNFM 收到 VNF 的创建请求后, 计算所需修改的资源列表, 向 NFVO 申请分配资源, 完成资源的预占和分

配, VNFM 启动新增 VM 并向 VNF 通知完成。图 9 是根据本发明实施例的 VNF 弹性收缩过程图, 如图 9 所示, 该过程包括如下步骤:

[0069] 步骤 S902: 云数据中心平台部署完成, VNF 资源创建好弹性伸缩规则, 虚拟网络部署完成, VNFM 信息采集完成;

[0070] 步骤 S904: VNF 自身弹性判断, 若 VNF 自身无法进行判断是否达到了弹性伸缩规则的条件, 则将监控的信息数据交由上游的 VNFM 节点进行判断;

[0071] 步骤 S906: 若达到伸缩条件, 则由 VNF 发送请求进行 VM 资源弹性伸缩处理的弹性伸缩处理消息给 VNFM;

[0072] 步骤 S908: VNFM 解析 VNF 发起的创建 VM 请求, 请求中携带 VM 名称和操作、以及创建 VM 的个数;

[0073] 如果 VM 已经存在, 当 VM 处于异常状态时, 直接上报错误;

[0074] 如果 VM 已经存在, 当 VM 处于运行状态时, 直接上报成功;

[0075] 如果 VM 已经存在, VM 处于其他正常状态, 则 VNFM 启动 VM 到正常运行状态, 然后上报成功;

[0076] 如果 VM 不存在, 则进入下一步;

[0077] VNFM 向 NFVO 发送第二请求消息, 请求全局的资源访问接口进行预申请资源, 资源申请成功上报消息并进入下一步; 若 NFVO 无应答超时, 则流程终止并上报错误; 如果预申请资源失败, 则上报失败, 流程终止;

[0078] VNFM 发送 VM 资源变更请求给 NFVO, 资源分配成功上报消息并进入下一步。若 NFVO 无应答超时, 则流程终止并上报错误; 如果资源分配失败, 则上报失败, 流程终止;

[0079] 步骤 S910: NFVO 成功创建 VM, 最终使 VM 达到初始状态, VNFM 生命周期业务向 VNF 发送启动创建的 VM; 若启动 VM 失败, 则上报错误; VNFM 上报消息返回 VNF 创建结果, 携带操作动作和 VM 名称, 以及每个 VM 的操作结果; VNF 对新增 VM 进行 SLB 处理。

[0080] 优选实施例三当上述第一节点为 VNFM 时的 VM 资源弹性伸缩处理:

[0081] VNFM 节点除了有可能需要判断下游的 VNF 指标信息外, 还需要同时判断并监控自身的指标信息, 这里的监控指标主要涉及内存、CPU、硬盘、带宽等系统信息以及网元处理能力, 若是达到伸缩条件, 则进行弹性伸缩处理, 当系统整体负荷较低的情况下, 如果达到了缩容条件, 则触发弹性伸缩请求, 发送 VM 资源请求变更消息给 NFVO, 通过对 VM 进行一系列状态迁移操作, 释放 VM 资源, 然后对 VNFM 进行 SLB 处理, 降低系统 TCO; 当系统整体负荷较高的情况下, 如果达到了扩容条件, 则触发弹性伸缩请求, 发送资源请求变更消息给 NFVO, 并根据 VNFM 绑定的 VM 配置, 通过对 VM 进行一系列状态迁移操作, 启动相应的 VM, 然后对 VNFM 进行 SLB 处理, 以此扩展、扩大系统处理能力。

[0082] 综上可知, NFVO 配合完成 VNFM 的弹性伸缩。该过程由 VM 发起, VNFM 收到 VNF 的创建请求后, 计算所需修改的资源列表, 向 NFVO 申请分配资源, 完成资源的预占和分配, VNFM 启动新增 VM 并向 VNF 通知完成。图 10 是根据本发明实施例的 VNFM 弹性伸缩过程图, 如图 10 所示, 该过程包括如下步骤:

[0083] 步骤 S1002: 云数据中心平台部署完成, VNFM 资源创建好弹性伸缩规则, 虚拟网络部署完成, VNFM 信息采集完成;

[0084] 步骤 S1004: VNFM 自身弹性判断, 若 VNFM 自身无法进行判断是否达到了弹性伸缩

规则的条件，则将监控的信息数据交由上游的 NFVO 节点进行判断；若达到伸缩条件，则由 VNFM 发送用于请求进行 VM 资源弹性伸缩处理的弹性伸缩处理消息给 NFVO；

[0085] 步骤 S1006：VNFM 解析自身发起的创建 VM 请求，请求中携带 VM 名称和操作、以及创建 VM 的个数：

[0086] 如果 VM 已经存在，当 VM 处于异常状态时，直接上报错误；

[0087] 如果 VM 已经存在，当 VM 处于运行状态时，直接上报成功；

[0088] 如果 VM 已经存在，VM 处于其他正常状态，则 VNFM 启动 VM 到正常运行状态，然后上报成功；

[0089] 如果 VM 不存在，则进入下一步；

[0090] VNFM 向 NFVO 发送第三请求消息，请求全局的资源访问接口进行预申请资源，资源申请成功上报消息并进入下一步；若 NFVO 无应答超时，则流程终止并上报错误；如果预申请资源失败，则上报失败，流程终止；VNFM 发送 VM 资源变更请求给 NFVO，资源分配成功上报消息并进入下一步；若 NFVO 无应答超时，则流程终止并上报错误；如果资源分配失败，则上报失败，流程终止；

[0091] 步骤 S1008：NFVO 成功创建 VM，上报消息返回 VNFM 创建结果，携带操作动作和 VM 名称，以及每个 VM 的操作结果；VNFM 对新增 VM 进行 SLB 处理。

[0092] 在上述各实施例中，各个层级的 VM 通过自身及上游节点的实时监控及弹性伸缩判断，尽量减少各层级的频繁实时信息交互，更加精确的实时检查并监控 VM 状态指标，检查如果达到弹性伸缩的条件，触发弹性伸缩动作，对 VM 自动进行状态迁移，调整 VM 的状态，通过云数据中心操作内存、CPU、硬盘、带宽等 VM 资源，达到扩展网元能力或者降低 TCO。

[0093] 相对于相关技术，本发明实施例中提出的电信设备虚拟化技术下的弹性伸缩状态迁移方法，不仅能减轻对上游的冲击，还能有效消除单点故障，减少消息交互，同时让采集的指标信息更加精确，在提升弹性伸缩服务安全性的同时，而且能降低设备资源的使用成本，而且可以得到各种定制化的服务，包括虚拟服务动态迁移、VM 按需分配。进而实现了更加安全、可靠、灵活、高效的信息运作的目的。

[0094] 显然，本领域的技术人员应该明白，上述的本发明的各模块或各步骤可以用通用的计算装置来实现，它们可以集中在单个的计算装置上，或者分布在多个计算装置所组成的网络上，可选地，它们可以用计算装置可执行的程序代码来实现，从而，可以将它们存储在存储装置中由计算装置来执行，并且在某些情况下，可以以不同于此处的顺序执行所示出或描述的步骤，或者将它们分别制作成各个集成电路模块，或者将它们中的多个模块或步骤制作成单个集成电路模块来实现。这样，本发明不限制于任何特定的硬件和软件结合。

[0095] 以上所述仅为本发明的优选实施例而已，并不用于限制本发明，对于本领域的技术人员来说，本发明可以有各种更改和变化。凡在本发明的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

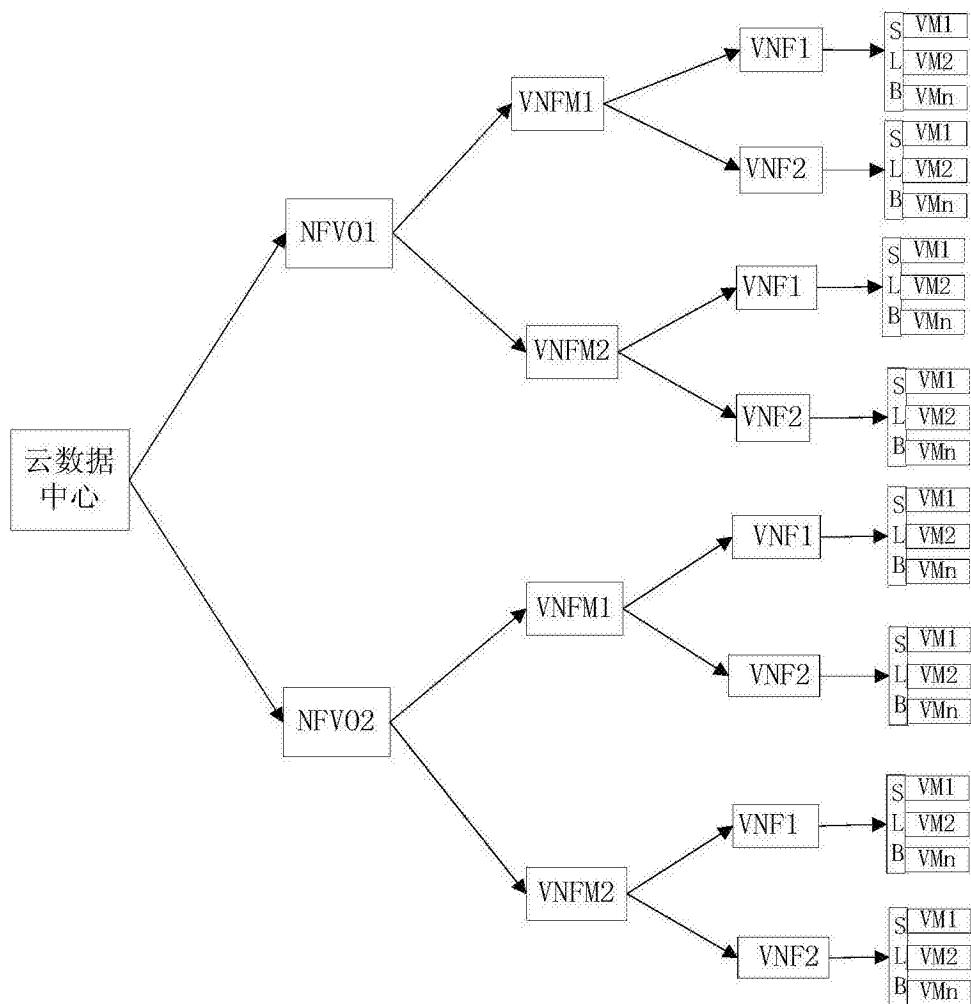


图 1

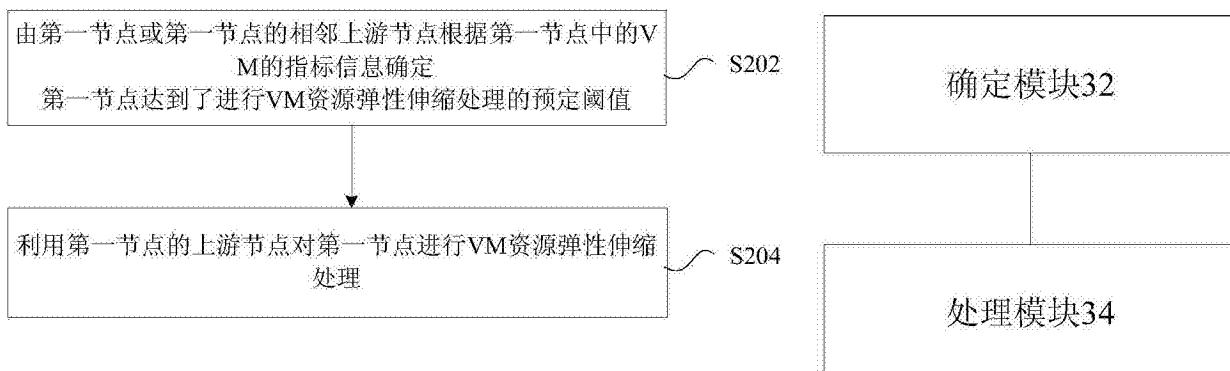


图 2

图 3



图 4



图 5



图 6

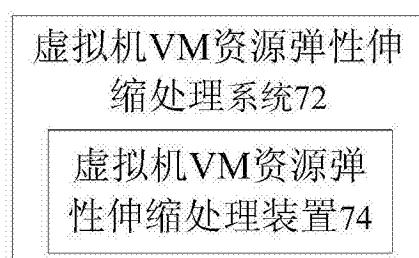


图 7

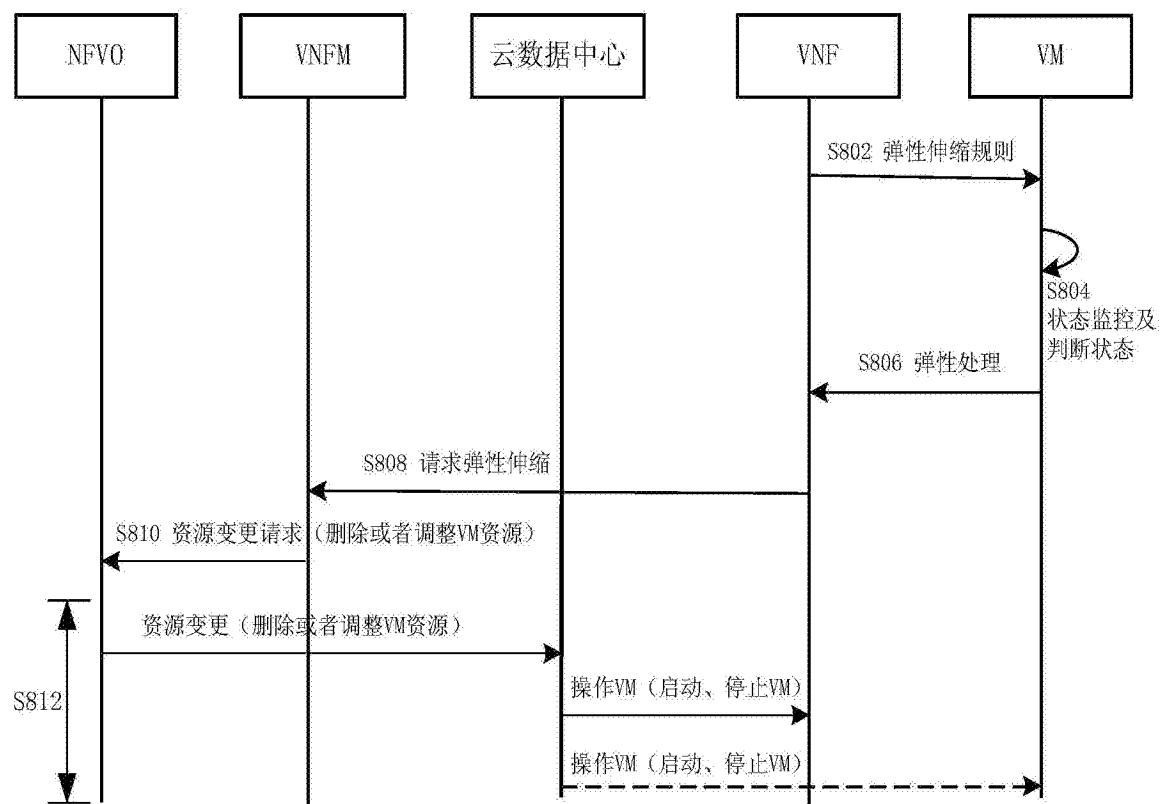


图 8

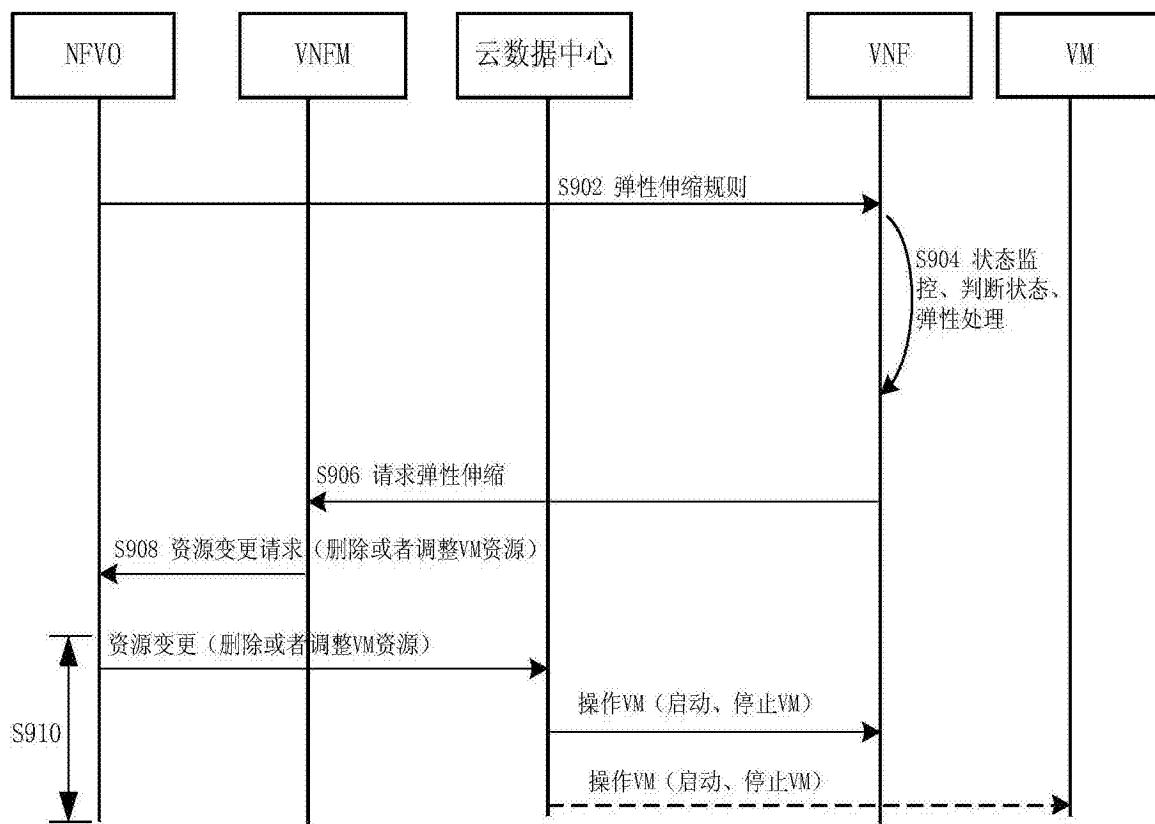


图 9

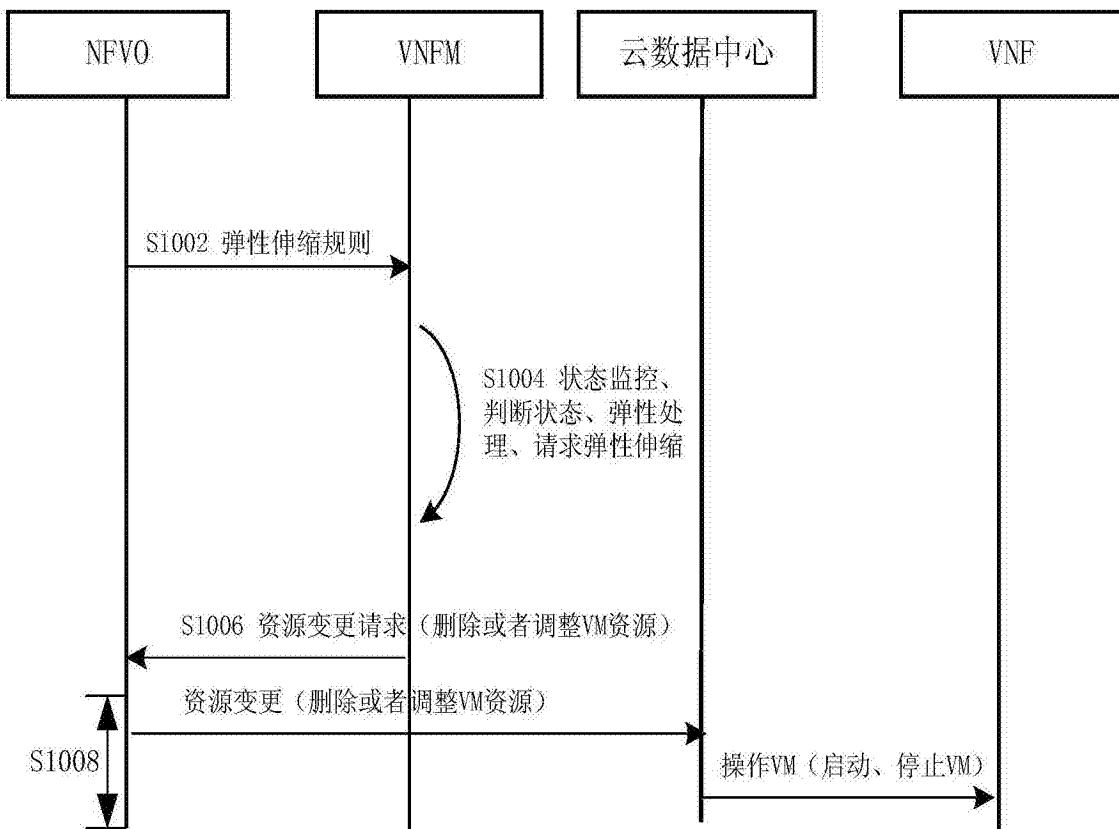


图 10