



(12) 发明专利申请

(10) 申请公布号 CN 106411552 A

(43) 申请公布日 2017.02.15

(21) 申请号 201510466806.4

(22) 申请日 2015.07.31

(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 顾忠禹

(74) 专利代理机构 北京康信知识产权代理有限公司
责任公司 11240

代理人 江舟 李灵洁

(51) Int. Cl.

H04L 12/24(2006.01)

H04L 12/46(2006.01)

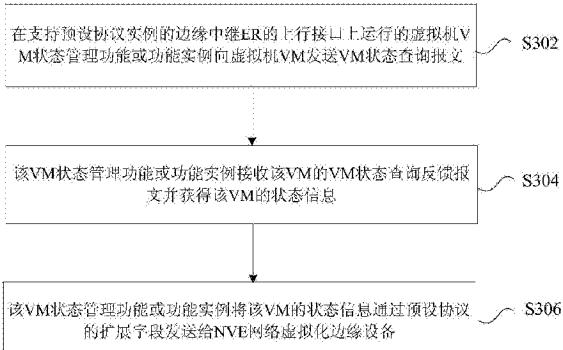
权利要求书2页 说明书10页 附图5页

(54) 发明名称

虚拟机状态管理方法及装置

(57) 摘要

本发明提供了一种虚拟机状态管理方法及装置，其中，该方法包括：在支持预设协议实例的边缘中继ER的上行接口上运行的虚拟机VM状态管理功能或功能实例向虚拟机VM发送VM状态查询报文，该VM状态管理功能或功能实例接收该VM的VM状态查询反馈报文并获得该VM的状态信息，该VM状态管理功能或功能实例将该VM的状态信息通过预设协议的扩展字段发送给NVE网络虚拟化边缘设备，解决了相关协议对虚拟机VM状态的管理还不完善的问题，实现了NV03环境下的虚拟机状态的实时管理。



1. 一种虚拟机状态管理方法,其特征在于,包括:

在支持预设协议实例的边缘中继 ER 的上行接口上运行的虚拟机 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文;

所述 VM 状态管理功能或功能实例接收所述 VM 的 VM 状态查询反馈报文并获得所述 VM 的状态信息;

所述 VM 状态管理功能或功能实例将所述 VM 的状态信息通过预设协议的扩展字段发送给 NVE 网络虚拟化边缘设备。

2. 根据权利要求 1 所述的方法,其特征在于,所述运行的虚拟机 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文,或者接收 VM 状态查询反馈报文包括:

所述 VM 状态管理功能或功能实例通过第一预定义的以太网帧向所述 ER 连接的 VM 发送所述状态查询报文,其中,在所述第一预定义的以太网帧为广播类型的情况下,所述 VM 状态管理功能或功能实例向所述 ER 连接的所有 VM 发送所述状态查询报文,在所述第一预定义的以太网帧为单播类型的情况下,所述 VM 状态管理功能或功能实例以指定的 VM 的 MAC 地址为目的地址向指定的 VM 发送所述 VM 状态查询报文;

所述 VM 以第二预定义的以太网帧发送 VM 状态查询反馈报文。

3. 根据权利要求 1 所述的方法,其特征在于,所述 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文之前,所述方法包括:

所述 VM 状态管理功能或功能实例接收 NVE 通过预设协议扩展的三元组 TLV 字段传送的状态查询命令信息,其中,所述状态查询命令信息用于指示所述 VM 状态管理功能或功能实例反馈 / 收集所述 VM 的状态信息。

4. 根据权利要求 1 所述的方法,其特征在于,所述 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文之后,包括:

接收状态查询反馈报文并获得所述 VM 的状态信息,并结合站点中保存的 VM 状态信息,形成 VM 的状态信息;

进一步,如果在预设时间段内没有收到所述 VM 回复的情况下,按照设定次数重复发送所述 VM 状态查询报文之后,在没有收到所述 VM 回复的情况下,判定所述 VM 不可达,形成不可达的 VM 状态信息。

5. 根据权利要求 1 所述的方法,其特征在于,所述 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文的发送时机包括以下至少之一:

在接收 NVE 发送的状态查询命令信息的情况下,所述 VM 状态管理功能或功能实例向 VM 发送 VM 状态查询报文;

定期发送 VM 状态查询报文。

6. 根据权利要求 1 至权利要求 5 任一项所述的方法,其特征在于,所述预设协议包括为:

虚拟机接口发现和配置协议 VDP,及所述 VDP 支持协议:分组数据汇聚协议 CDPC 和扩展能力端口模式协议 ECP。

7. 根据权利要求 1 或权利要求 2 所述的方法,其特征在于,所述虚拟机状态管理方法,包括:

所述 NVE 直接使用第三预定义的以太网帧,向虚拟机 VM 发送 VM 状态查询报文;所述 VM

使用第四预定义的以太网帧直接向 NVE 反馈 VM 状态信息。所述第三预定义的以太网帧和所述第四预定义的以太网帧不同于所述第一预定义的以太网帧和第二预定义的以太网帧的类型值。

8. 一种虚拟机状态管理装置，其特征在于，包括：

第一发送模块，用于在支持预设协议实例的边缘中继 ER 的上行接口上运行的虚拟机 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文；

第一接收模块，用于所述 VM 状态管理功能或功能实例接收所述 VM 的 VM 状态查询反馈报文并获得所述 VM 的状态信息；

第二发送模块，用于所述 VM 状态管理功能或功能实例将所述 VM 的状态信息通过预设协议的扩展字段发送给 NVE 网络虚拟化边缘设备。

9. 根据权利要求 8 所述的装置，其特征在于，所述第一发送模块包括：

预定义发送单元，用于所述 VM 状态管理功能或功能实例通过第一预定义的以太网帧向所述 ER 连接的 VM 发送所述状态查询报文，其中，在所述第一预定义的以太网帧为广播类型的情况下，所述 VM 状态管理功能或功能实例向所述 ER 连接的所有 VM 发送所述状态查询报文，在所述第一预定义的以太网帧为单播类型的情况下，所述 VM 状态管理功能或功能实例以指定的 VM 的 MAC 地址为目的地址向指定的 VM 发送所述 VM 状态查询报文；

所述第一接收模块还用于接收所述 VM 以第二预定义的以太网帧发送 VM 状态查询反馈报文。

10. 根据权利要求 8 所述的装置，其特征在于，所述装置包括：

第二接收模块，用于所述 VM 状态管理功能或功能实例接收 NVE 通过预设协议扩展的三元组 TLV 字段传送的状态查询命令信息，其中，所述状态查询命令信息用于指示所述 VM 状态管理功能或功能实例反馈 / 收集的所述 VM 的状态信息。

11. 根据权利要求 8 所述的装置，其特征在于，所述装置包括：

判定模块，用于在接收状态查询反馈的状态信息，并将结合站点中保存的 VM 状态信息，形成 VM 的状态信息；并且，如果在预设时间段内没有收到所述 VM 回复的情况下，按照设定次数重复发送所述 VM 状态查询报文之后，在没有收到所述 VM 回复的情况下，判定所述 VM 不可达，即形成 VM 不可达的状态信息；

所述第二发送模块，还用于向所述 NVE 报告所述 VM 不可达的状态信息。

12. 根据权利要求 8 所述的装置，其特征在于，包括：

所述第一发送模块，还用于在接收 NVE 发送的状态查询命令信息的情况下，所述 VM 状态管理功能或功能实例向 VM 发送 VM 状态查询报文；

所述第一发送模块，还用于定期发送 VM 状态查询报文。

13. 根据权利要求 8 至权利要求 12 任一项所述的装置，其特征在于，所述预设协议包括为：

虚拟机接口发现和配置协议 VDP，及所述 VDP 支持协议：分组数据汇聚协议 CDPC，扩展能力端口模式协议 ECP。

虚拟机状态管理方法及装置

技术领域

[0001] 本发明涉及通信领域,具体而言,涉及一种虚拟机状态管理方法及装置。

背景技术

[0002] 在数据中心中,如何通过虚拟化技术,例如虚拟机,提供虚拟网络 (Virtual Network, 简称为 VN) 是一个热门的话题,涉及到网络的虚拟化等众多方面,引起设备供应商和运营商的多方关注。例如,电气和电子工程师协会 (Institute for Electrical and Electronic Engineers, 简称为 IEEE) 推出 802.1Qbg(边缘虚拟桥),支持虚拟机 (IEEE 术语为虚拟站)。需要区分提供虚拟站的平台,称为 EVB 站。),如何连接到虚拟化平台外的交换机 (Edge Virtual Bridging, 简称为 EVB 桥)。图 1 是相关技术中的 EVB 的体系结构示意图,如图 1 所示,通过虚拟机接口发现和配置协议 (Virtual Station Interface discovery and configuration protocol, 简称为 VDP),分组数据汇聚协议 (Packet Data Convergence Protocol, 简称为 PDCP) 和扩展能力端口模式协议 (Extented Capability Port, 简称为 ECP) 等一系列协议,实现虚拟机接口 (Virtual Station Interface, 简称为 VSI) 到 EVB 桥的连接。另外,在国际互联网工程任务组 (The Internet Engineering Task Force 简称为 IETF) 中也专门成立了一个工作组基于层 3 的网络虚拟化 (Network Virtualization Overlay based-on L3Network, 简称为 NV03),对其进行研究。图 2 是相关技术中 NV03 虚拟网络提供的架构示意图,如图 2 所示,对具体如何实现 VN 的提供,有了一致的认识。

[0003] 如图 2 所示,有一个集中的网络虚拟化控制设备 (Network Virtualization Authority, 简称为 NVA),实现对业务提供的控制, VN 中的多个虚拟机 (Virtual Machine, 简称为 VM) 连接到网络虚拟化边缘设备 (Network Virtualization Edge, 简称为 NVE),并进一步与 NVE 中的该 VN 的虚拟路由转发功能 (Virtual Routing and Forwarding, 简称为 VRF) 进行关联,多个不同 NVE 种的同一 VN 的不同的 VRF,通过 NVE 直接的重叠网络隧道实现连接,从而实现 VN。当然,VN 可以通过 NVE(一般驻留在数据中心无关)和互联网实现连接。

[0004] 但对 NV03 如何实现 VN 的具体的解决方案,还莫衷一是。特别是对 VM 如何接入 NVE 这一段的控制协议,一种观点认为,可以通过 IEEE 的 VDP 来承担。直观的,VDP 是实现 VSI/VM 到 EVB 桥 (主要是和面向通道的桥接口 (Station-facing bridge port, 简称为 SBP)) 的关联工作,这和 VM 加入 NVE 的 VN 非常相似,是一种可能的解决方案。在 NV03 中有较多的讨论,例如,因为 VDP 是一种 L2 的协议,为了实现 NV03 的虚拟网络,需要进行 L3 地址信息的支持,以及需要支持 VN-ID 信息的传递等。而对 VM 状态的管理,目前还主要关注于 VM 的生成和删除等等的生命周期管理,具体的包括 VM 处于初始化状态、就绪状态和激活状态等。在激活状态下,就认为 VM 是正常工作的。则是与 VDP 所处的位置有关,或者进一步在相关技术体制下, Hypervisor 也就只能获得这些 VM 的信息。

[0005] 但是,在 VM 处于激活状态,或者运行状态下,VM 本身是需要执行特别的工作,或者

多项工作的,由于某种原因,这些工作或者功能有可能是不工作的,或者功能不能执行。这对 VN 中的其他 VM 而言,该 VM 实际上是不可用的,但是 VDP 技术体制不能提供对该问题的解决方案。

[0006] 针对相关技术中,相关协议对虚拟机 VM 状态的管理还不完善的问题,目前还没有有效的解决方案。

发明内容

[0007] 本发明提供了一种虚拟机状态管理方法及装置,以至少解决相关技术中相关协议对虚拟机 VM 状态的管理还不完善的问题。

[0008] 根据本发明的一个方面,提供了一种虚拟机状态管理方法,包括:

[0009] 在支持预设协议实例的边缘中继 ER 的上行接口上运行的虚拟机 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文;

[0010] 所述 VM 状态管理功能或功能实例接收所述 VM 的 VM 状态查询反馈报文并获得所述 VM 的状态信息;

[0011] 所述 VM 状态管理功能或功能实例将所述 VM 的状态信息通过预设协议的扩展字段发送给 NVE 网络虚拟化边缘设备。

[0012] 进一步地,所述运行的虚拟机 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文,或者接收 VM 状态查询反馈报文包括:

[0013] 所述 VM 状态管理功能或功能实例通过第一预定义的以太网帧向所述 ER 连接的 VM 发送所述状态查询报文,其中,在所述第一预定义的以太网帧为广播类型的情况下,所述 VM 状态管理功能或功能实例向所述 ER 连接的所有 VM 发送所述状态查询报文,在所述第一预定义的以太网帧为单播类型的情况下,所述 VM 状态管理功能或功能实例以指定的 VM 的 MAC 地址为目的地址向指定的 VM 发送所述 VM 状态查询报文;

[0014] 所述 VM 以第二预定义的以太网帧发送 VM 状态查询反馈报文。

[0015] 进一步地,所述 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文之前,所述方法包括:

[0016] 所述 VM 状态管理功能或功能实例接收 NVE 通过预设协议扩展的三元组 TLV 字段传送的状态查询命令信息,其中,所述状态查询命令信息用于指示所述 VM 状态管理功能或功能实例反馈 / 收集的所述 VM 的状态信息。

[0017] 进一步地,所述 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文之后,包括:

[0018] 接收状态查询反馈报文并获得所述 VM 的状态信息,并结合站点中保存的 VM 状态信息,形成 VM 的状态信息;

[0019] 进一步,如果在预设时间段内没有收到所述 VM 回复的情况下,按照设定次数重复发送所述 VM 状态查询报文之后,在没有收到所述 VM 回复的情况下,判定所述 VM 不可达,形成不可达的 VM 状态信息。

[0020] 进一步地,所述 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文的发送时机包括以下至少之一:

[0021] 在接收 NVE 发送的状态查询命令信息的情况下,所述 VM 状态管理功能或功能实例

向 VM 发送 VM 状态查询报文；

[0022] 定期发送 VM 状态查询报文。

[0023] 进一步地，所述预设协议包括为：

[0024] 虚拟机接口发现和配置协议 VDP，及所述 VDP 支持协议：分组数据汇聚协议 CDCP 和扩展能力端口模式协议 ECP。

[0025] 进一步地，所述虚拟机状态管理方法，包括：

[0026] 所述 NVE 直接使用第三预定义的以太网帧，向虚拟机 VM 发送 VM 状态查询报文；所述 VM 使用第四预定义的以太网帧直接向 NVE 反馈 VM 状态信息。所述第三预定义的以太网帧和所述第四预定义的以太网帧不同于所述第一预定义的以太网帧和第二预定义的以太网帧的类型值。

[0027] 根据本发明的另一个方面，还提供了一种虚拟机状态管理装置，包括：

[0028] 第一发送模块，用于在支持预设协议实例的边缘中继 ER 的上行接口上运行的虚拟机 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文；

[0029] 第一接收模块，用于所述 VM 状态管理功能或功能实例接收所述 VM 的 VM 状态查询反馈报文并获得所述 VM 的状态信息；

[0030] 第二发送模块，用于所述 VM 状态管理功能或功能实例将所述 VM 的状态信息通过预设协议的扩展字段发送给 NVE 网络虚拟化边缘设备。

[0031] 进一步地，所述第一发送模块包括：

[0032] 预定义发送单元，用于所述 VM 状态管理功能或功能实例通过第一预定义的以太网帧向所述 ER 连接的 VM 发送所述状态查询报文，其中，在所述第一预定义的以太网帧为广播类型的情况下，所述 VM 状态管理功能或功能实例向所述 ER 连接的所有 VM 发送所述状态查询报文，在所述第一预定义的以太网帧为单播类型的情况下，所述 VM 状态管理功能或功能实例以指定的 VM 的 MAC 地址为目的地址向指定的 VM 发送所述 VM 状态查询报文；

[0033] 所述第一接收模块还用于接收所述 VM 以第二预定义的以太网帧发送 VM 状态查询反馈报文。

[0034] 进一步地，所述装置包括：

[0035] 第二接收模块，用于所述 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文之前，所述 VM 状态管理功能或功能实例接收 NVE 通过预设协议扩展的三元组 TLV 字段传送的状态查询命令信息，其中，所述状态查询命令信息用于指示所述 VM 状态管理功能或功能实例反馈 / 收集的所述 VM 状态信息。

[0036] 进一步地，所述装置包括：

[0037] 判定模块，用于在接收状态查询反馈的状态信息，并将结合站点中保存的 VM 状态信息，形成 VM 的状态信息；并且，如果在预设时间段内没有收到所述 VM 回复的情况下，按照设定次数重复发送所述 VM 状态查询报文之后，在没有收到所述 VM 回复的情况下，判定所述 VM 不可达，即形成 VM 不可达的状态信息；

[0038] 所述第二发送模块，还用于向所述 NVE 报告所述 VM 不可达的状态信息。

[0039] 进一步地，所述第一发送模块，还用于在接收 NVE 发送的状态查询命令信息的情况下，所述 VM 状态管理功能或功能实例向 VM 发送 VM 状态查询报文；

[0040] 所述第一发送模块，还用于定期发送 VM 状态查询报文。

[0041] 进一步地,所述预设协议包括为:虚拟机接口发现和配置协议 VDP,及所述 VDP 支持协议:分组数据汇聚协议 CDCP,扩展能力端口模式协议 ECP。

[0042] 通过本发明,在支持预设协议实例的边缘中继 ER 的上行接口上运行的虚拟机 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文,该 VM 状态管理功能或功能实例接收该 VM 的 VM 状态查询反馈报文并获得该 VM 状态信息,该 VM 状态管理功能或功能实例将该 VM 的状态信息通过预设协议的扩展字段发送给 NVE 网络虚拟化边缘设备,解决了相关协议对虚拟机 VM 状态的管理还不完善的问题,实现了 NV03 环境下的虚拟机状态的实时管理。

附图说明

[0043] 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:

[0044] 图 1 是相关技术中的 EVB 的体系结构示意图;

[0045] 图 2 是相关技术中 NV03 虚拟网络提供的架构示意图;

[0046] 图 3 是根据本发明实施例的一种虚拟机状态管理方法的流程图;

[0047] 图 4 是根据本发明实施例的一种虚拟机状态管理装置的结构框图一;

[0048] 图 5 是根据本发明实施例的一种虚拟机状态管理装置的结构框图二;

[0049] 图 6 是根据本发明实施例的一种虚拟机状态管理装置的结构框图三;

[0050] 图 7 是根据本发明实施例的一种虚拟机状态管理装置的结构框图四;

[0051] 图 8 是根据本发明实施例的网络功能结构示意图;

[0052] 图 9 是根据本发明实施例的以太网帧结构示意图;

[0053] 图 10 是根据本发明实施例的 VM 状态收集命令以太网帧格式的详细定义示意图;

[0054] 图 11 是根据本发明实施例 VM 状态 VDP TLV 定义意图;

[0055] 图 12 是根据本发明实施例另一种 VM 状态 VDP TLV 定义意图;

[0056] 图 13 是根据本发明实施例基于 VDP 扩展的 VM 状态管理的流程示意图。

具体实施方式

[0057] 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。

[0058] 需要说明的是,本发明的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。

[0059] 在本实施例中提供了一种虚拟机状态管理方法,图 3 是根据本发明实施例的一种虚拟机状态管理方法的流程图,如图 3 所示,该流程包括如下步骤:

[0060] 步骤 S302,在支持预设协议实例的边缘中继 ER 的上行接口上运行的虚拟机 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文;

[0061] 步骤 S304,该 VM 状态管理功能或功能实例接收该 VM 的 VM 状态查询反馈报文并获得该 VM 的状态信息;

[0062] 步骤 S306,该 VM 状态管理功能或功能实例将该 VM 的状态信息通过预设协议的扩展字段发送给 NVE 网络虚拟化边缘设备。

[0063] 通过上述步骤,在支持预设协议实例的边缘中继 ER 的上行接口上运行的虚拟机 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文,该 VM 状态管理功能或功能实例接收该 VM 的 VM 状态查询反馈报文并获得该 VM 状态信息,该 VM 状态管理功能或功能实例将该 VM 的状态信息通过预设协议的扩展字段发送给 NVE 网络虚拟化边缘设备,解决了相关协议对虚拟机 VM 状态的管理还不完善的问题,实现了 NV03 环境下的虚拟机状态的实时管理。

[0064] 在本实施例中,该运行的虚拟机 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文,或者接收 VM 状态查询反馈报文包括:

[0065] 该 VM 状态管理功能或功能实例通过第一预定义的以太网帧向该 ER 连接的 VM 发送该状态查询报文,其中,在该第一预定义的以太网帧为广播类型的情况下,该 VM 状态管理功能或功能实例向该 ER 连接的所有 VM 发送该状态查询报文,在该第一预定义的以太网帧为单播类型的情况下,该 VM 状态管理功能或功能实例以指定的 VM 的 MAC 地址为目的地址向指定的 VM 发送该 VM 状态查询报文;

[0066] 该 VM 以第二预定义的以太网帧发送 VM 状态查询反馈报文。

[0067] 在本实施例中,该 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文之前,该 VM 状态管理功能或功能实例接收 NVE 通过预设协议扩展的三元组 TLV 字段传送的状态查询命令信息,其中,该状态查询命令信息用于指示该 VM 状态管理功能或功能实例反馈 / 收集的该 VM 的状态信息。

[0068] 在本实施例中,该 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文之后,接收状态查询反馈报文并获得该 VM 的状态信息,并结合站点中保存的 VM 状态信息,形成 VM 的状态信息;

[0069] 进一步,如果在预设时间段内没有收到该 VM 回复的情况下,按照设定次数重复发送该 VM 状态查询报文之后,在没有收到该 VM 回复的情况下,判定该 VM 不可达,形成不可达的 VM 状态信息;

[0070] 在本实施例中,该 VM 状态管理功能或功能实例向虚拟机 VM 发送 VM 状态查询报文的发送时机包括以下至少之一:

[0071] 在接收 NVE 发送的状态查询命令信息的情况下,该 VM 状态管理功能或功能实例向 VM 发送 VM 状态查询报文;

[0072] 定期发送 VM 状态查询报文。

[0073] 在本实施例中,该预设协议包括为:

[0074] 虚拟机接口发现和配置协议 VDP,及该 VDP 支持协议:分组数据汇聚协议 CDCP 和扩展能力端口模式协议 ECP。

[0075] 在本实施例中,该 NVE 直接使用第三预定义的以太网帧,向虚拟机 VM 发送 VM 状态查询报文;该 VM 使用第四预定义的以太网帧直接向 NVE 反馈 VM 状态信息。该第三预定义的以太网帧和第四预定义的以太网帧不同于该第一预定义的以太网帧和第二预定义的以太网帧的类型值。

[0076] 在本实施例中还提供了一种虚拟机状态管理装置,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。

[0077] 图4是根据本发明实施例的一种虚拟机状态管理装置的结构框图一,如图4所示,该装置包括:

[0078] 第一发送模块42,用于在支持预设协议实例的边缘中继ER的上行接口上运行的虚拟机VM状态管理功能或功能实例向虚拟机VM发送VM状态查询报文;

[0079] 第一接收模块44,用于该VM状态管理功能或功能实例接收该VM的VM状态查询反馈报文并获得该VM的状态信息;

[0080] 第二发送模块46,用于该VM状态管理功能或功能实例将该VM的状态信息通过预设协议的扩展字段发送给NVE网络虚拟化边缘设备。

[0081] 通过上述装置,在支持预设协议实例的边缘中继ER的上行接口上运行的虚拟机VM状态管理功能或功能实例向虚拟机VM发送VM状态查询报文,该VM状态管理功能或功能实例接收该VM的VM状态查询反馈报文并获得该VM状态信息,该VM状态管理功能或功能实例将该VM的状态信息通过预设协议的扩展字段发送给NVE网络虚拟化边缘设备,解决了相关协议对虚拟机VM状态的管理还不完善的问题,实现了NV03环境下的虚拟机状态的实时管理。

[0082] 图5是根据本发明实施例的一种虚拟机状态管理装置的结构框图二,如图5所示,该装置包括:该第一发送模块42包括:

[0083] 预定义发送单元52,用于该VM状态管理功能或功能实例通过第一预定义的以太网帧向该ER连接的VM发送该状态查询报文,其中,在该第一预定义的以太网帧为广播类型的情况下,该VM状态管理功能或功能实例向该ER连接的所有VM发送该状态查询报文,在该第一预定义的以太网帧为单播类型的情况下,该VM状态管理功能或功能实例以指定的VM的MAC地址为目的地址向指定的VM发送该VM状态查询报文;

[0084] 该第一接收模块44还用于接收该VM以第二预定义的以太网帧发送VM状态查询反馈报文。

[0085] 图6是根据本发明实施例的一种虚拟机状态管理装置的结构框图三,如图6所示,该装置包括:

[0086] 第二接收模块62,用于该VM状态管理功能或功能实例向虚拟机VM发送VM状态查询报文之前,该VM状态管理功能或功能实例接收NVE通过预设协议扩展的三元组TLV字段传送的状态查询命令信息,其中,该状态查询命令信息用于指示该VM状态管理功能或功能实例反馈/收集的该VM状态信息。

[0087] 图7是根据本发明实施例的一种虚拟机状态管理装置的结构框图四,如图7所示,该装置包括:

[0088] 判定模块72,用于在接收状态查询反馈的状态信息,并将结合站点中保存的VM状态信息,形成VM的状态信息;并且,如果在预设时间段内没有收到该VM回复的情况下,按照设定次数重复发送该VM状态查询报文之后,在没有收到该VM回复的情况下,判定该VM不可达,即形成VM不可达的状态信息;

[0089] 该第二发送模块46,还用于向该NVE报告该VM不可达的信息,其中,该不可达的信息指示该NVE刷新虚拟路由转发功能VRF的该VM不可达的状态信息。

[0090] 在本实施例中,该第一发送模块42,还用于在接收NVE发送的状态查询命令信息

的情况下,该 VM 状态管理功能或功能实例向 VM 发送 VM 状态查询报文;

[0091] 该第一发送模块 42,还用于定期发送 VM 状态查询报文。

[0092] 在上述实施例中,该预设协议包括为:虚拟机接口发现和配置协议 VDP,及该 VDP 支持协议:分组数据汇聚协议 CDCP,扩展能力端口模式协议 ECP。

[0093] 在另外一个实施例中,还提供了一种数据中心中 NV03 环境下、基于 VDP 的虚拟机状态管理的实现方法和系统。具体地,包括:

[0094] ER 的上行接口支持 VDP 实例,同时支持 VDP 对 VM 的状态的收集的功能实体 / 实例。

[0095] 进一步,具体的,状态收集功能实体,通过特定的以太网广播帧发送状态征询报文。ER 所连接的所有 VM,在收到该状态查询报文后,以该 VM 的 MAC 地址,以状态收集功能实体的 MAC 地址为目的地址发送状态信息。

[0096] 所述特定的以太网帧,通过特定的以太网帧类型进行定义。

[0097] 优选的,VM 状态收集功能实例,也可以通过单播的以太网帧,来进行状态收集,具体的,使用 VM 的 MAC 地址直接发送单播的状态查询报文。

[0098] VM 状态收集功能实例,在收集到 VM 的状态信息后,通过 VDP 协议的扩展,将信息发送给 NVE。

[0099] 进一步,VM 状态管理功能实体 / 实例可以接收 NVE 通过 VDP 扩展 TLV 传送的状态查询命令信息。

[0100] 进一步,VM 状态收集功能实例在发送状态查询报文后,在规定的时间内没有收到回复,则可以再进行设定次数的报文发送,如果没有收到回复,则判定该 VM 不可达。需要向 NVE 报告 VM 不可达信息,并刷新 NVE 中的 VRF 的 VM 状态信息。

[0101] 进一步,VM 状态收集功能实例状态查询报文的发送时机,可以是 NVE 发送来的状态查询,或者 VM 状态收集功能实例本身定义的保活 (keepalive) 时钟设定触发的。

[0102] 经过这样的扩展,可以基于 VDP 协议的扩展实现在 NV03 环境下的虚拟机状态的动态刷新和管理。

[0103] 图 8 是根据本发明实施例的网络功能结构示意图,如图 8 所示,其中,为了收集 VM 的运行状态,最基本的方法是与之进行通信,显式地询问其状态。在 EVB 的情况下,由于可以使用 VDP 来进行相关的信息传递,因此,一个可行的解决方案,是由驻留 VDP 功能实体的 ER 的上行接口 (Uplink Relay Port,简称为 URP) 上,再运行一个 VM 状态收集功能实例,来实现 VM 的状态的收集。

[0104] 需要说明的是,在 VM 和 NVE 之间,也可以运行其他的协议,来实现状态的收集。

[0105] 具体的,由于 VSI 和 SBP 之间,存在 L2 报文传输通道,这样本发明,就是使用 ER 和 VM 之间的 L2 隧道来进行状态收集。

[0106] 具体的,是定义一种新的 / 指定类型以太网帧格式,专门用于 URP 和 VM 之间进行信息交互。新的以太网帧,是指不同于现有的以太网类型,是专门用于 VM 状态信息交互的一种以太网类型。一般而言在大规模部署时,这个类型需要 IEEE 的批准。具体的数值,是在 IEEE 可以提供的类型中选择一个,并需要 IEEE 批准。一个具体的参考例子是,在以太网上承载 PPP 协议 (Point to Point Protocol over Ethernet,简称为 PPPoE) 发现阶段报文的以太网类型值为 0x8863(即 16 进制的 8863)。

[0107] 需要说明的是,这里需要 2 个指定类型的以太网帧,一个用于查询命令;另外一个用于查询命令的状态反馈。当然,都需要 IEEE 的批准。

[0108] 图 9 是根据本发明实施例的以太网帧结构示意图,如图 9 所示,而具体的状态的信息获得,还需要做进一步的命令定义。具体的,是通过该以太网帧的载荷部分来实现。在载荷部分,定义一个新的报头。图 10 是根据本发明实施例的 VM 状态收集命令以太网帧格式的详细定义示意图,如图 10 所示。

[0109] 其中编码字段定义具体实现的功能。包括:状态询问,可以询问单个 VM,或者所有 VM。这可以通过以太网帧的目的地址来实现,或者区分。具体的,如果是广播地址,则询问域中的所有 VM 的状态。如果是单播地址,则指定的 VM 返回其状态信息。

[0110] Keepalive 命令,定期发送,询问 VM 的状态,如果在几个 keepalive 命令后,仍然收不到 VM 的反馈,则认为该 VM 是不可达的。

[0111] VN-ID 字段,用于区分不同的 VN 信息,这个是可选的。

[0112] Session-ID,用于区分不同的通信,具体的 VM 返回状态信息时也需要返回该 session-ID 信息,以区分是哪一个状态询问。

[0113] 另外一个选项是,VM 地址 / 名称,或者一组 VM 的地址 / 名称,用于指示具体是针对操作的 VM。

[0114] 版本信息,可以定义为 001,用于区分后续的版本。

[0115] 长度字段,用于指示状态查询报头长度。不包括以太网帧的头部。

[0116] 为实现的简洁性,可以选择 2 个以太网帧类型,其中一个用于命令,另外一个用于状态信息的传递。

[0117] 或者也可以通过编码类型,来进行命令和状态报文的区分。例如,0x00,0x01,用于表示前述的 2 个命令;用 0x04,0x05,表示状态的反馈信息。对这样的报文,可以进一步定义具体的状态编码。例如,0x00 表示 VM 处于正常工作状态;0xFF 表示,VM 不可用,以及其他编码表示其他不同的状态。

[0118] 这里,仅仅做一个示范,具体的定义值,可以根据实际情况进行选择,而不影响本发明的本质内容。

[0119] 由于 VDP 用于关联 VSI 和 SBP,并没有相关的内容来传递 VM 的具体的工作状态,因此,需要对 VDP 进行扩展,实现 VM 的状态的传递。

[0120] 具体的,可以定义一种新的 VDP TLV 类型,VM 状态 TLV。可以用保留的 TLV 值 0x06-0x7E 中选择一个。

[0121] 图 11 是根据本发明实施例 VM 状态 VDP TLV 定义意图,如图 11 所示,而具体的 TLV (Type Length Value) 定义,如图 11 所示。具体的包括 TLV 类型,以及 TLV 信息长度和状态字段。

[0122] 进一步,该 TLV 包括命令字段,用于指示 NVE 的 VM 状态查询命令。或者,优选的,图 12 是根据本发明实施例另一种 VM 状态 VDP TLV 定义意图,如图 12 所示,可以通过状态字段的前 4 位用于状态查询命令的定义,而后 4 位用于 VM 状态的定义。

[0123] 需要说明的是,也可以利用 VDP 的现有关联 TLV 定义的过滤信息字段来进行状态的字段定义。

[0124] 通过上述的定义和描述,实现了 VM 状态的信息收集和如何通过 VDP 的扩展,来实

现将 VM 状态信息,传递给 NVE。

[0125] 而具体的 VM 状态的传递时机,可以是 VM 在状态发生变化时,主动发送给 NVE,或者 VM 状态信息收集功能实例在获得 VM 状态发生变化时,进行传递,也可能是 NVE 主动发起,或者定期对 VM 进行状态询问。

[0126] 图 13 是根据本发明实施例基于 VDP 扩展的 VM 状态管理的流程示意图,如图 13 所示,包括如下步骤。

[0127] 步骤 S1001、可选的, NVE 向 VM 状态管理功能实体 / 实例发送 VM 状态查询信息。

[0128] 步骤 S1002、VM 状态管理功能实体 / 实例,在 VM 的 keepalive 计时器时间到,或者在接收到 NVE 发送的 VM 状态查询命令或者 keepalive 命令时,也或者 VM 状态管理功能实体 / 实例设定的定时状态查询时间到,启动 VM 状态查询。

[0129] 步骤 S1003、发送 VM 状态查询报文。

[0130] 步骤 S1004、VM 返回其状态信息。

[0131] 步骤 S1005. VM 状态管理功能实体 / 实例将收到的 VM 状态信息,发送给 NVE,或者做记录、以备后用。或者如果在设定的时间内,再次发送该命令,重复设定次数的命令再次发送,如果仍然不能收到 VM 的回复,则判定该 VM 为不可达。

[0132] 步骤 S1006、VM 状态管理功能实体 / 实例将所述 VM 状态信息,包括 VM 反馈的信息,或者判定的 VM 不可达信息,发送给 NVE。

[0133] 其他实施例

[0134] 上述实施例中,是充分保证利用现有协议的情况下解决方案,即利用现有的 VDP 协议的扩展,来完成 VM 的实时状态查询,而不存在不一致。

[0135] 另外一个解决方案,是在 VDP 协议运行的状态报告之外,NVE 直接和 VM 进行通信,获得 VM 的实时状态信息。

[0136] 具体的方案是,定义另外 2 个不同的以太网帧类型,用于 NVE 和 VM 之间的状态查询和反馈。

[0137] 具体的帧的定义和前述的定义类似。不同的是,由于 NVE 支持多个虚拟网络 VN,所以报文中的 VN-ID 字段是必选项,保证 VM 的状态信息送到正确的 VN。

[0138] 而具体的实现方面是类似的。例如可以通过广播向 NVE 所连接的 VN 中的所有 VM 发送查询命令,或者通过单播地址向某个特定的 VM 发送查询命令。而 VM 一般都是用单播地址反馈状态信息。

[0139] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到根据上述实施例的方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质(如 ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,或者网络设备等)执行本发明各个实施例所述的方法。

[0140] 本发明的实施例还提供了一种存储介质。可选地,在本实施例中,上述存储介质可以被设置为存储用于执行上述实施例的方法步骤的程序代码:

[0141] 可选地,在本实施例中,上述存储介质可以包括但不限于 :U 盘、只读存储器 (ROM, Read-Only Memory)、随机存取存储器 (RAM, Random Access Memory)、移动硬盘、磁碟或者

光盘等各种可以存储程序代码的介质。

[0142] 可选地，在本实施例中，处理器根据存储介质中已存储的程序代码执行上述实施例的方法步骤。

[0143] 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。

[0144] 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。

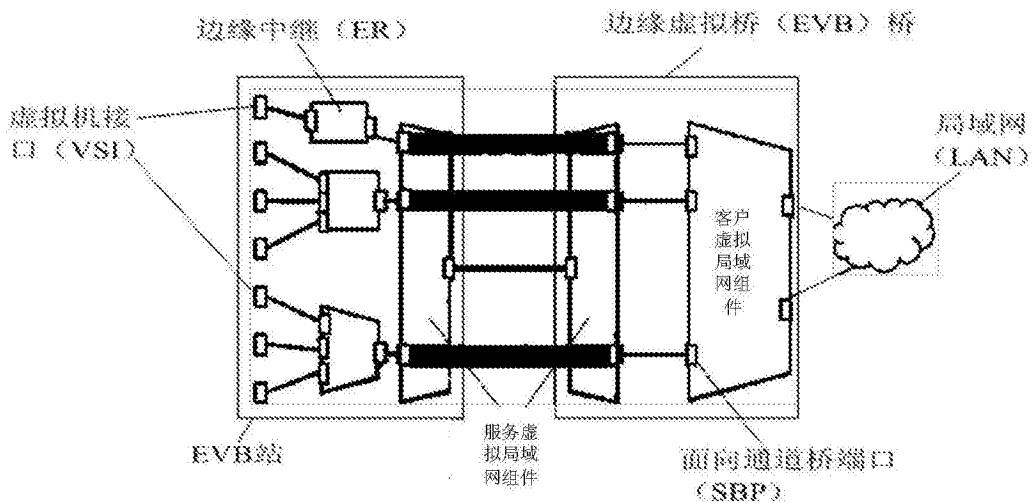


图 1

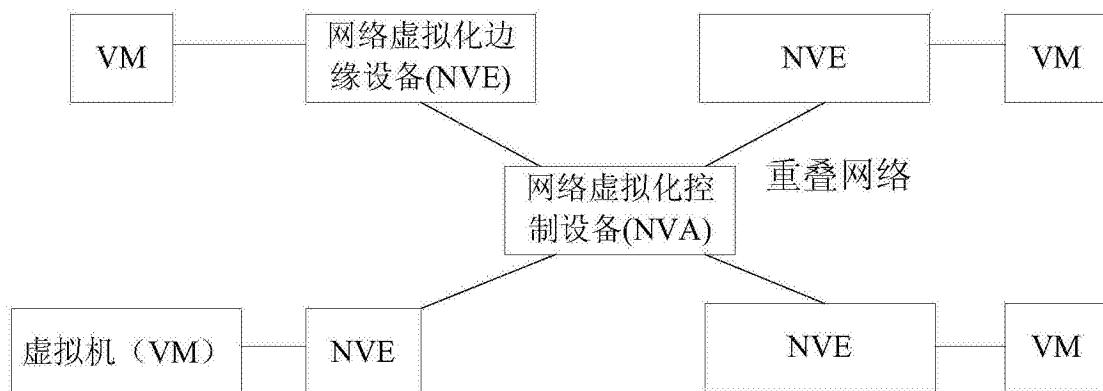


图 2

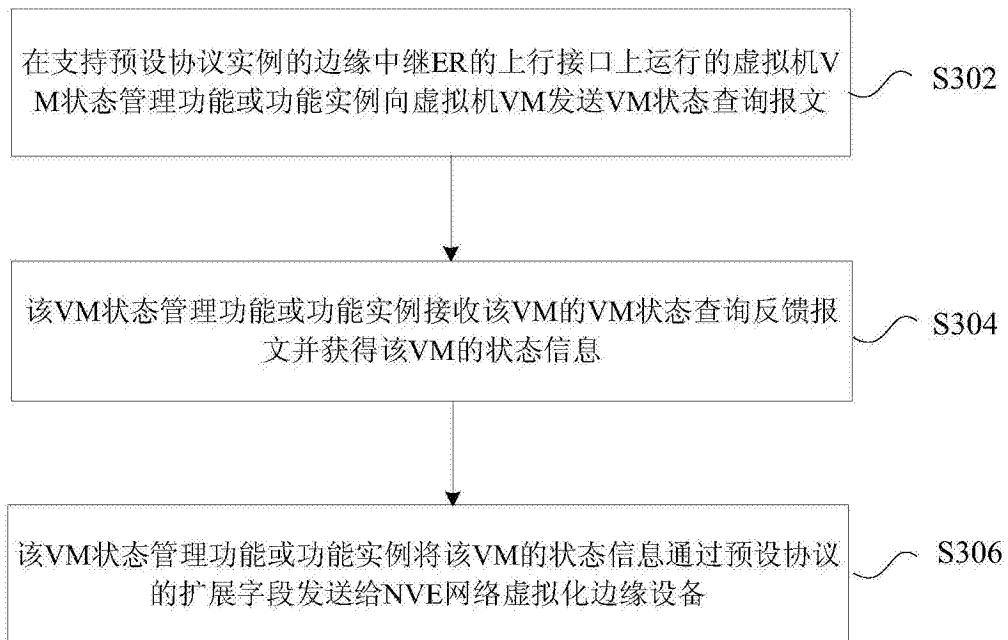


图 3

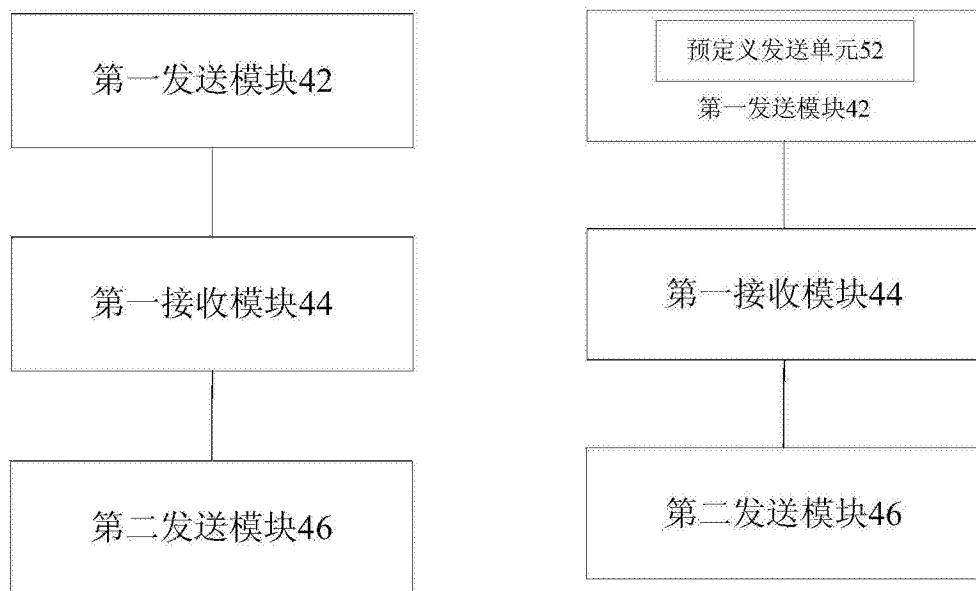


图 4

图 5

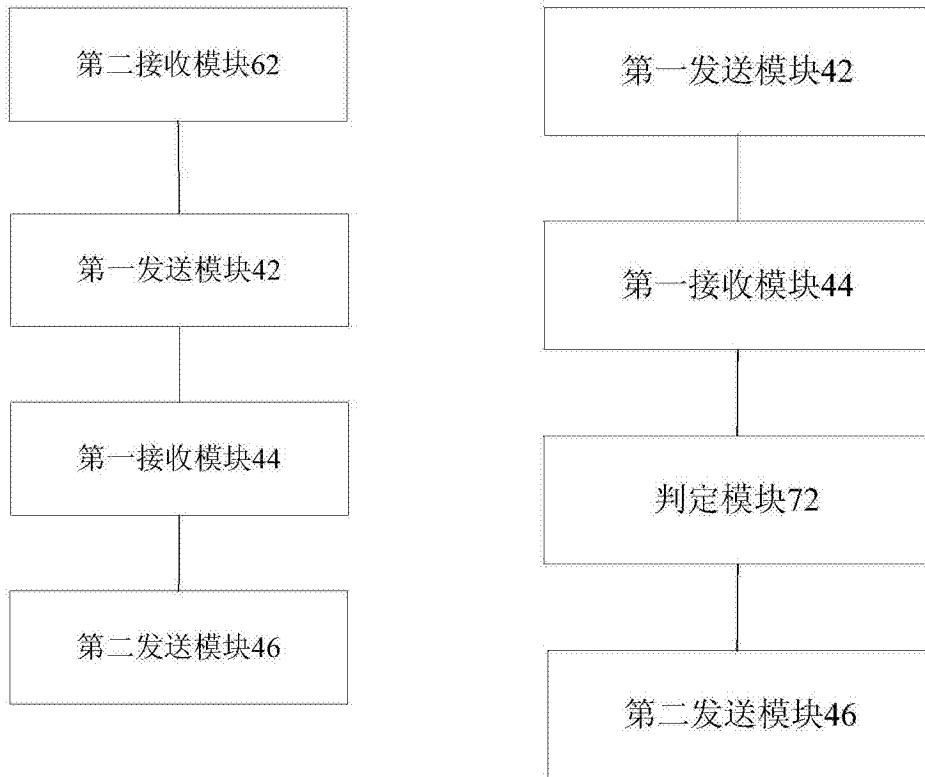


图 6

图 7

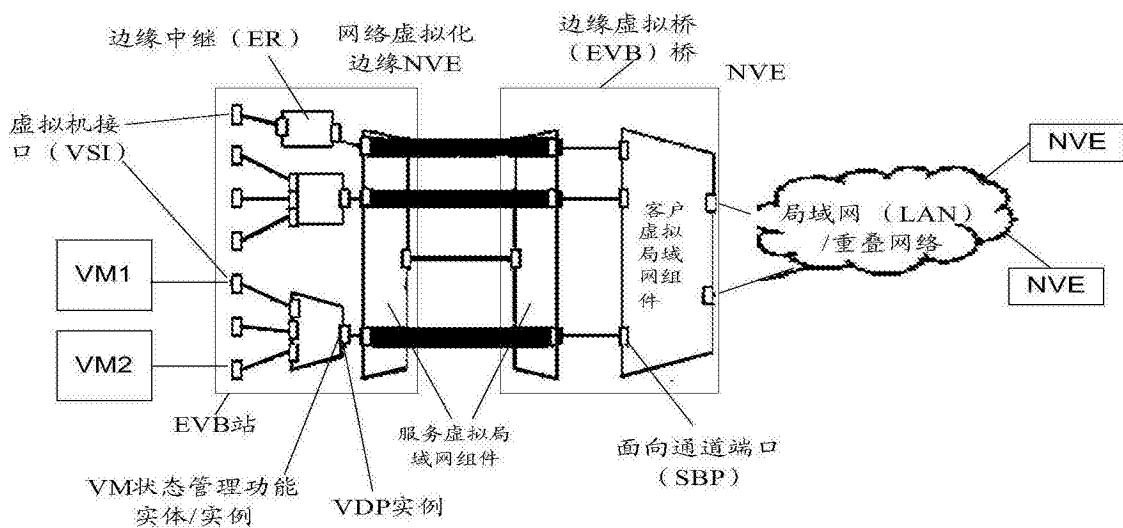


图 8



图 9

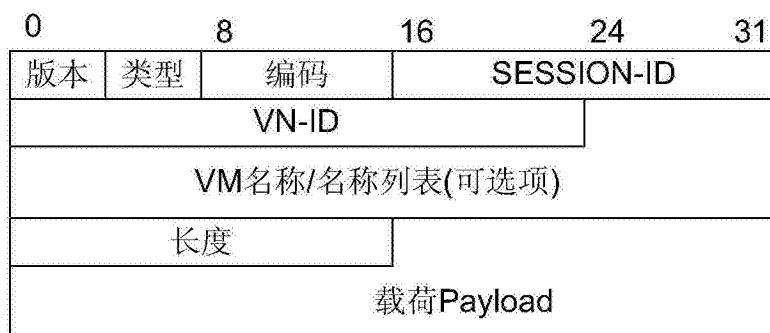


图 10

TLV类型 (7位)	TLV信息长度 (9位)	状态 (8位)
---------------	-----------------	------------

图 11

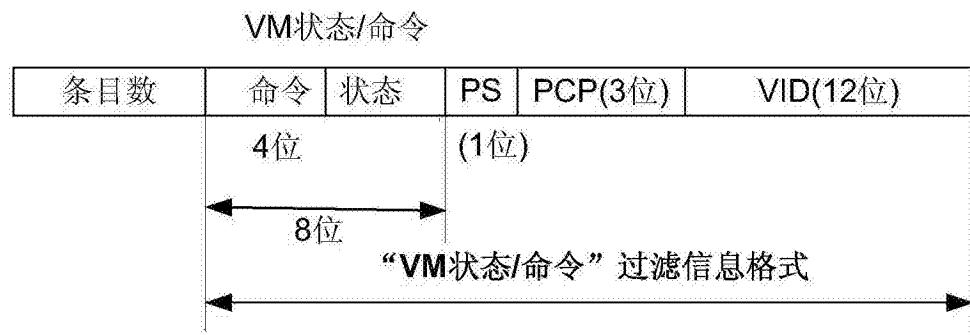


图 12

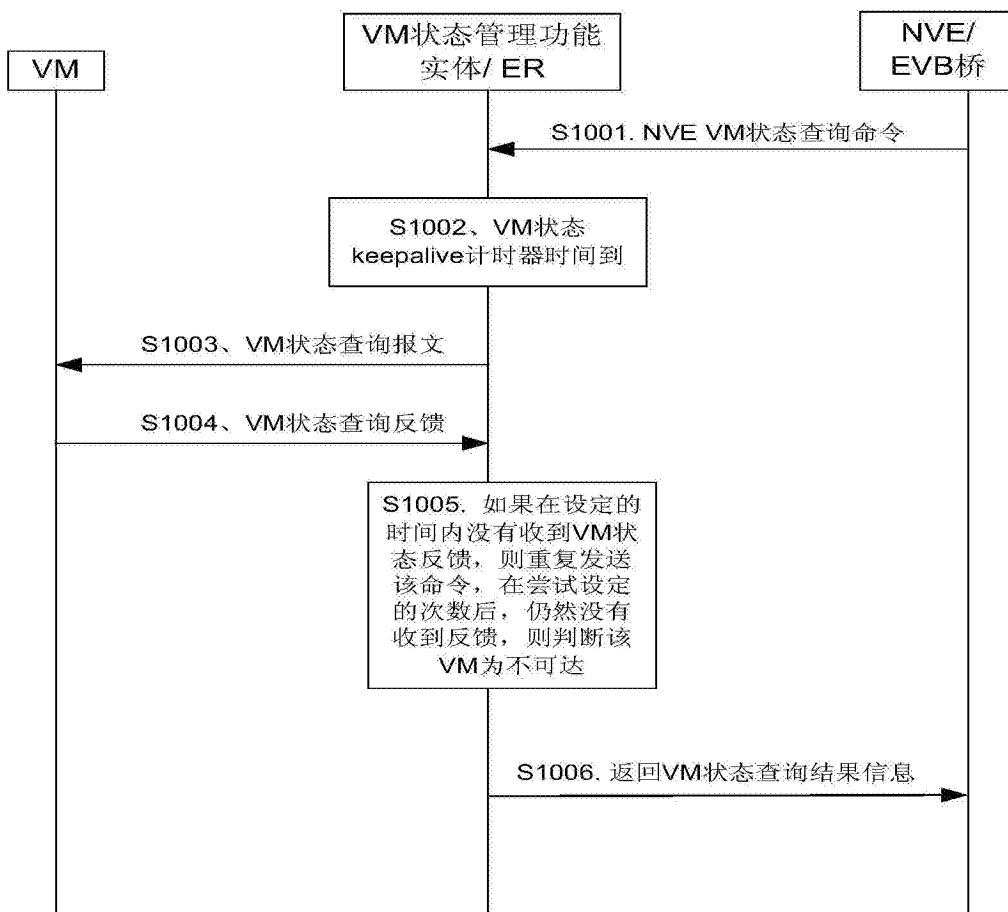


图 13