



(12) 发明专利申请

(10) 申请公布号 CN 102281153 A

(43) 申请公布日 2011. 12. 14

(21) 申请号 201110156082. 5

(22) 申请日 2011. 06. 10

(30) 优先权数据

12/813, 085 2010. 06. 10 US

(71) 申请人 美国博通公司

地址 美国加州尔湾市奥尔顿公园路 16215 号, 92618-7013

(72) 发明人 韦尔·威廉·戴博

(74) 专利代理机构 深圳市顺天达专利商标代理有限公司 44217

代理人 蔡晓红 邹秋菊

(51) Int. Cl.

H04L 12/24 (2006. 01)

G06F 9/455 (2006. 01)

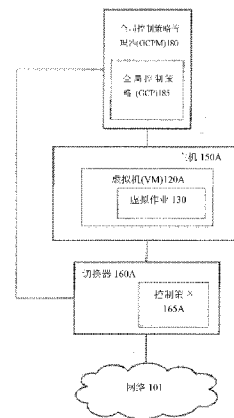
权利要求书 2 页 说明书 12 页 附图 9 页

(54) 发明名称

在网络中管理能效和控制机制的系统和方法

(57) 摘要

本发明涉及具有多个网络组件的网络中管理能效和控制机制的系统。所述系统包括全局控制策略管理器, 具有全局控制策略, 和所述多个网络组件中的至少一个相连, 其中, 所述全局控制策略管理器配置成用于: 接收来自所述多个网络组件中所述至少一个网络组件的电能信息; 分析所述电能信息; 基于所述全局控制策略和所述接收到的电能信息, 生成控制策略修改; 发送所述控制策略修改到所述多个网络组件中的所述至少一个网络组件。



1. 一种在具有多个网络组件的网络中管理能效和控制机制的系统,包括:  
全局控制策略管理器,具有全局控制策略,和所述多个网络组件中的至少一个相连,其中,所述全局控制策略管理器配置成用于:  
接收来自所述多个网络组件中所述至少一个网络组件的电能信息;  
分析所述电能信息;  
基于所述全局控制策略和所述接收到的电能信息,生成控制策略修改;  
发送所述控制策略修改到所述多个网络组件中的所述至少一个网络组件。
2. 根据权利要求1所述的系统,其特征在于,所述控制策略修改涉及执行能效以太网机制的设备的控制策略。
3. 根据权利要求1所述的系统,其特征在于,所述控制策略修改涉及与切换器上的端口相关的能效设置。
4. 根据权利要求1所述的系统,其特征在于,所述控制策略修改涉及增加通信量路径的能效。
5. 根据权利要求4所述的系统,其特征在于,所述通信量路径是指根据系统配置的改变来确定的新的通信量路径。
6. 根据权利要求1所述的系统,其特征在于,所述控制策略修改的生成是响应系统配置变化执行的。
7. 根据权利要求1所述的系统,其特征在于,所述系统配置变化是虚拟作业从网络中的一个虚拟机到网络中的另一个虚拟机的迁移。
8. 根据权利要求1所述的系统,其特征在于,所述系统还包括其上有虚拟作业运行的虚拟机,其中,所述全局控制策略基于所述虚拟作业的要求。
9. 一种在具有多个网络组件的网络中管理电能控制和能效机制的方法,其特征在于,所述方法包括:  
在全局控制策略管理器接收来自所述多个网络组件中至少一个网络组件的电能信息;  
分析所述电能信息;  
基于全局控制策略和所述接收到的电能信息生成控制策略修改;  
发送所述控制策略修改到所述多个网络组件中的至少一个网络组件。
10. 一种在网络中管理能效和控制机制的系统,其特征在于,包括:  
第一主机;  
所述第一主机控制的第一虚拟机;  
第二主机;  
所述第二主机控制的第二虚拟机;  
具有应用在其上的第一控制策略的第一组硬件组件,所述第一组硬件组件连接所述第一主机;  
具有应用在其上的第二控制策略的第二组硬件组件,所述第二组硬件组件连接所述第二主机;  
在所述第一虚拟机上运行的虚拟作业;  
全局控制策略管理器,其中,如果所述虚拟作业从所述第一虚拟机迁移到所述第二虚

拟机,所述全局控制策略管理器配置成重配置所述第二控制策略以匹配所述虚拟作业的要求。

## 在网络中管理能效和控制机制的系统和方法

### 技术领域

[0001] 本发明涉及使用虚拟机的网络的电能 (power) 消耗管理。

### 背景技术

[0002] 近些年,能耗持续增长的趋势已经加速。因此,各种各样的行业对能耗增加的影响变得越来越敏感。IT 基础设施也对此进行了行详细的研究。现在,许多多个公司正在考虑他们的 IT 系统的电能使用情况来判断能耗是否能降低。出于这个原因,已经出现了专注于节能网络的行业,这个行业着眼于解决整个 IT 设备(例如,个人计算机,显示器,打印机,服务器和网络组件等)的电能消耗增加的问题。

[0003] 现代的网络组件正在不断深入地贯彻能耗及能效 (energy consumption and efficiency, ECE) 的控制机制。一些能耗及能效 (ECE) 控制机制允许物理层组件进入或者退出低能模式。能耗及能效 (ECE) 控制策略在能耗及能效 (ECE) 控制使能物理层组件进入或者退出低能模式时起控制作用。设备控制策略在最大化节约的同时最小化对网络的影响中扮演重要的角色。

[0004] 相同的现代网络正在不断增加使用能耗及能效 (ECE) 使能的组件和虚拟组件。虚拟机管理器管理虚拟的环境,以不同的方式在不同的物理拓扑结构的顶端展开。

[0005] 在虚拟环境中展开应用的好处是虚拟应用与系统硬件间灵活的关系。现代虚拟机管理器 (VMMs) 能够基于不同的考虑,例如,安全性、高可用性、能效和运行情况快速将虚拟申请移动到不用的硬件系统中。

[0006] 通常,当虚拟应用迁移到与新的硬件设备一起工作时,虚拟应用不能充分地利用新硬件设备不同的特性和功能。进一步,新的硬件设备可能有与迁移来的虚拟作业冲突的可配置选项。当新的硬件设备加入到退出系统时,相同的问题将在小范围发生。

[0007] 即使当设备能够重新配置来适应迁移来的虚拟作业的新要求时,也没有一种统一的、全局 (global) 的方法来协调各种系统设备的控制策略。

[0008] 履行用于运行虚拟作业的系统的全局控制策略的挑战包括虚拟机配置的动态特性和其上展开该虚拟机的不同类型的拓扑结构,例如,跨越多个机器的虚拟机,复杂以及动态变化的网络拓扑结构和维护性能指标。

[0009] 因此,需要一种能克服上述缺点的方法。

### 发明内容

[0010] 基于本发明的一个方面,提供一种在具有多个网络组件的网络中管理能效和控制机制的系统,包括:

[0011] 全局控制策略管理器 (GCPM),具有全局控制策略 (GCP),和所述多个网络组件中的至少一个相连,其中,所述全局控制策略管理器 (GCPM) 配置成用于:

[0012] 接收来自所述多个网络组件中所述至少一个网络组件的电能信息;

[0013] 分析所述电能信息;

- [0014] 基于所述全局控制策略和所述接收到的电能信息,生成控制策略修改(modification);
- [0015] 发送所述控制策略修改到所述多个网络组件中的所述至少一个网络组件。
- [0016] 优选地,所述多个网络组件中的一个网络组件是切换器和路由器中的一个。
- [0017] 优选地,所述多个网络组件中的一个网络组件是主机。
- [0018] 优选地,所述控制策略修改涉及(direct to)执行能效以太网(Energy Efficient Ethernet EEE)机制的设备的控制策略。
- [0019] 优选地,所述控制策略修改涉及与切换器上的端口相关的能效设置。
- [0020] 优选地,所述控制策略修改涉及增加通信量路径的能效。
- [0021] 优选地,所述通信量路径是指基于系统配置的改变来确定的新的通信量路径。
- [0022] 优选地,所述控制策略修改的生成是响应系统配置变化执行的。
- [0023] 优选地,所述系统配置变化是虚拟作业从网络中的一个虚拟机到网络中的另一个虚拟机的迁移。
- [0024] 优选地,所述系统还包括其上有虚拟作业运行的虚拟机,其中,所述全局控制策略基于所述虚拟作业的要求。
- [0025] 优选地,所述系统还包括第二虚拟机,其中,所述控制策略修改的所述生成配置成在当虚拟作业从所述第一虚拟机迁移到所述第二虚拟机时发生。
- [0026] 优选地,虚拟作业从所述第一虚拟机迁移到所述第二虚拟机还包括将要在两个虚拟机上执行的虚拟作业的跨越(spanning)。
- [0027] 基于本发明的一个方面,提供一种在具有多个网络组件的网络中管理电能控制和能效机制的方法,所述方法包括:
- [0028] 在全局控制策略管理器(GCPM)接收来自所述多个网络组件中至少一个网络组件的电能信息;
- [0029] 分析所述电能信息;
- [0030] 基于全局控制策略(GCP)和所述接收到的电能信息生成控制策略修改;
- [0031] 发送所述控制策略修改到所述多个网络组件中的至少一个网络组件。
- [0032] 优选地,所述多个网络组件中的至少一个网络组件是虚拟机电能管理器(VMPM)。
- [0033] 优选地,所述的全局控制策略是基于虚拟作业的要求来创建的。
- [0034] 优选地,所述方法还包括在控制策略修改生成阶段前:
- [0035] 分析虚拟作业;
- [0036] 基于所述虚拟作业的所述分析生成全局控制策略。
- [0037] 优选地,所述控制策略修改涉及执行能效以太网机制的设备的控制策略。
- [0038] 基于本发明更进一步的方面,提供一种在网络中管理能效和控制机制的系统,包括:
- [0039] 第一主机;
- [0040] 所述第一主机控制的(host)第一虚拟机;
- [0041] 第二主机;
- [0042] 所述第二主机控制的第二虚拟机;
- [0043] 具有应用在其上的第一控制策略的第一组硬件组件,所述第一组硬件组件连接所

述第一主机；

[0044] 具有应用在其上的第二控制策略的第二组硬件组件，所述第二组硬件组件连接所述第二主机；

[0045] 在所述第一虚拟机上运行的虚拟作业；

[0046] 全局控制策略管理器 (GCPM)，其中，如果所述虚拟作业从所述第一虚拟机迁移到所述第二虚拟机，所述全局控制策略管理器 (GCPM) 配置成重配置所述第二控制策略以匹配所述虚拟作业的要求。

[0047] 优选地：

[0048] 所述第一组硬件组件是第一组彼此连接、与所述第一主机和所述目的地节点相连的切换器，这样所述第一组硬件组件、第一主机和目的地节点构成了第一通信量路径；

[0049] 所述第二组硬件组件是第二组彼此连接、与所述第二主机和所述目的地节点相连的切换器，这样所述第二组硬件组件、第二主机和目的地节点构成了第二通信量路径；

[0050] 所述全局控制策略管理器配置成重配置所述第二控制策略，来推进关于第二通信量路径的目标。

[0051] 优选地，所述目标是提高的能效。

[0052] 优选地，所述目标是提高的性能。

#### 附图说明

[0053] 以下将参照附图结合实施例对本发明进行详细描述，更进一步地说明了本发明的原理，使与本技术领域有关的技术人员能够制造和使用本发明。

[0054] 图 1 是根据本发明的实施例，在具有多个网络组件的网络中管理电能控制和能效 (energy control and efficiency, ECE) 机制的系统的框图；

[0055] 图 2 是根据本发明的实施例，在具有多个网络组件的网络中管理电能控制和能效 (ECE) 机制的系统的框图，包括多个主机；

[0056] 图 3 是根据本发明的实施例，在具有多个网络组件的网络中管理电能控制和能效 (ECE) 机制的系统的框图，包括虚拟机电能管理器；

[0057] 图 4 是全局控制策略管理器的不同的实施例的框图；

[0058] 图 5 是根据本发明的实施例，在具有多个网络组件的网络中管理电能控制和能效 (ECE) 机制的系统的框图，包括跨越两台主机的虚拟机；

[0059] 图 6 示出了根据本发明的实施例的全局控制策略管理器的示例使用；

[0060] 图 7 示出了根据本发明的实施例、具有多个主机全局策略管理器的示例使用；

[0061] 图 8 示出了根据本发明的实施例、具有单个主机全局策略管理器的示例使用；

[0062] 图 9 是根据本发明的实施例，在具有多个网络组件的网络中管理电能控制和能效 (ECE) 机制的方法的流程图。

[0063] 本发明参照附图进行了描述。其中同一个附图标记在各幅附图中用于表示相同的或相似的部件。

#### 具体实施方式

[0064] 下文将要描述的各种实施例将要参考相应的附图，这些附图构成了实施例的一部

分,其中描述了实现本发明可能采用的各种实施例。应明白,还可使用其他的实施例,或者对本文列举的实施例进行结构和功能上的修改,而不会脱离本发明的范围和实质。

[0065] 后续部分将继续阐述本发明的其他特点与优点,通过这些阐述或者本发明的实施例,这些特点与优点在某种程度上是显而易见的。通过文字描述、附图及权利要求中描述的结构及指出的特征,可以了解到对本发明的优点。以上概括描述和后续的详细描述都是示例性的,目的是对本发明权利要求进行进一步解释。

[0066] 说明书中针对“一个实施例”、“实施例”、“一个示例实施例”等的引用,指的是描述的该实施例可包括特定的特征、结构或特性,但是不是每个实施例必须包含这些特定特征、结构或特性。此外,这样的表述并非指的是同一个实施例。进一步,在结合实施例描述特定的特定、结构或特性时,不管有没有明确的描述,已经表明将这样的特征、结构或特性结合到其它实施例中是属于本领域技术人员的知识范围内的。

[0067] 图 1A 是在具有虚拟机的网络中、管理节能和能效 (ECE) 以及控制机制的系统 100 的实施例的框图。示出了主机 150A 与网络切换器 (切换器) 160A 和全局控制策略管理器 (GCPM) 180 相连接。主机 150A 包括虚拟机 (VM) 120A。所示切换器 160A 与网络 101 相连,且具有控制策略 165A。所示的虚拟作业 130 在虚拟机 120A 中运行。

[0068] 在节能和能效 (ECE) 控制策略实现中、控制策略 165A 对于系统中的每一个硬件组件都是独立的。控制策略更进一步在下文中描述。

[0069] 在进一步的描述的实施例中,如果虚拟作业 130 置于新的硬件配置 (如果该新的硬件配置还存在系统配置变化),全局控制策略管理器 (GCPM) 180 将依照全局控制策略 (GCP) 185,为所述新的配置组件,重编独立的硬件控制策略 165。

[0070] 主机 150A 是具有各种各样的计算机系统配置的典型计算机,包括多核多处理机系统,微型计算机,大型计算机,连接或是集成分布式功能的计算机,以及可能嵌入任何虚拟设备的普及型或小型的计算机。在实施例中,更高容量 / 性能的主机 150A 可以称为计算机服务器,或者服务器。

[0071] 在此使用的网络 101,可以是任何一种或多种网络,包括但不限于局域网、媒体域网或者广域网如因特网。在实施例中,系统 100 是具有网络 101 连接的多台主机 150A 和切换器 160A 的较大的、互联数据中心子集。

[0072] 网络切换器 160A 是具有数据端口的典型的网桥设备,所述数据端口另具有路由 / 切换功能,例如,L3 切换器 / 路由。所述切换器可以有少到两个的数据端口,多到 400 个或更多的数据端口,而且可以从任何一个端口到另外一个端口以全双工管理通信量,有效地使任何端口作为输入,任何端口作为输出。据此,为了便于讨论,数据端口和与他们对应的链路可互换地称为数据信道、通信链路、数据链路等。因为图中的物理描述不应该解释为限制,切换器 160A 和主机 150A,在此使用到的可以包括主机 150A 和切换器 160A 可结合到物理设备中 (未示出)。切换器 160A 广义地包括在现代分层的切换器结构中使用的切换器逻辑。在一个实施例中,切换器 160A 功能能够与主机 150A 的功能在单个设备中物理结合,例如在服务器上整合 2- 端口 V- 切换器功能。在实施例中,切换器 160A 也可使用更高层的功能,如访问控制表 (ACLs) 和 L4 切换器 / 路由接口。

[0073] 虚拟机 (VM) 120A 是作为虚拟平台一部分的虚拟机的典型例子。本领域技术人员了解现代虚拟平台以及它的实现。具有虚拟机的系统一般提供了完整的系统平台,这个平

台支持完整的控制系统 (OS) 的执行。虚拟机操作系统可能在一个或者多个物理机上运行 (例如, 主机 150A), 它可以和其它虚拟机共享。

[0074] 如在此使用的能耗和能效 (ECE) 控制机制, 用来指控制设备的能耗和能效的各种现代技术。总的来说, 这些能耗和能效机制设计成用来降低能耗以及在维持可接受的运行时提高能效。

[0075] 一个能耗和能效控制机制的例子是配置成使用 IEEE802.3az 标准的机制, 也被称为能效以太网, 在此结合引用, 以作参考。EEE 是 IEEE 建议的标准, 设计用来在以太网网络中, 在选择的物理层设备 (PHY) 组上节能。应用到 EEE 使能的切换器的控制机制能监控切换器里的 TX 缓冲器, 和在特定的环境下指引切换器进入低能闲置 (low power idle, LPI) 状态。

[0076] 举例的物理层设备指的是在 EEE 标准下包括 10GBASE-TX 和 1000BASE-T 物理层设备, 以及新兴的 10GBASE-T 工艺和底板接口, 例如 10GBASE-KR。

[0077] 在实施例中, 不同的网络组件可能有不同的能耗和能效功能。例如, 有些组件可能没有能耗及能效功能, 而其他的可能只有基本的闲置功能, 例如, 闲置开或者闲置关闭。其它更复杂的组件可以由内部的“控制策略”来达到能耗和能效功能。

[0078] 如在此典型使用的, 控制策略广泛用于描述引导策略, 控制什么时候物理设备使用节能功能。例如, 切换器的控制策略可以描述什么时候和在什么情况下, 切换器进入和退出一个节能低能状态。控制策略可以用于控制系统中的一个或者多个物理或者虚拟的设备。控制策略 (也被称为物理控制策略或者设备控制策略), 例如, 为支持 EEE 的设备增加附加的控制层。

[0079] 全局控制策略管理器 (GCPM)

[0080] 作为此处使用的全局控制策略 (GCP) 185 的实施例是一配置策略, 该策略为单个设备控制策略发出 (set out) 高电平目标, 以此推进与虚拟作业 130 相关的不同结果。在另一个实施例中, 全局控制策略 185 具有用于推进与不同的组件相关联的结果的设置, 一般地包括主机 150A 和虚拟作业的组合。相关技术领域的人员, 在本发明的教导的基础上, 将会充分意识到 GCP185 中概述的全局系统控制策略可以具有不同的详细程度和应用到不同类型的相关情况中。

[0081] 基于实施例, 全局控制策略管理器 180 可以与系统 100 的硬件交互, 基于虚拟作业 130 的要求, 使用全局控制策略 185 来调整它们的 ECE 机制、控制策略和其它相关的控制。全局控制策略 185 的实施例还可以平衡虚拟作业 130 的要求与其他网络考虑, 例如, 运算和安全性等。在一实施例中, 可平衡改进 ECE 性能的全局控制策略管理器 180 的操作, 可将全局控制策略管理器 180 的操作与为整个系统设置的其他性能特性和目标、和其他专门的虚拟作业协调, 或使用为整个系统设置的其他性能特性和目标、和其他专门的虚拟作业更改所述全局控制策略管理器 180 的操作。

[0082] 在一个实施例中, 全局控制策略 185 可以由用户或者外部、自动化的程序手动生成。在另一个实施例中, 全局控制 185 是全局控制策略管理器 180 通过收集与虚拟作业 130 和网络 100 组件有关的数据, 以及自动地选择优选的全局策略来自动生成的。上述手动和自动的方法都可以结合起来使用。

[0083] 实现全局控制策略 185 时, 全局控制策略管理器 180 可以接收各种类型的关于网

络组件的 ECE/ 电能相关的信息（电能信息）。全局控制策略管理器 180 在实施例中，可以引导配置改变，来影响发送电能信息的组件。这些电能信息的例子包括物理层（PHY）信息、链路信息、ECE 控制策略信息和应用信息。相关技术领域的人员，在本发明的教导的基础上，将会充分意识到广范的信息、特性和策略等，将作为在此使用的电能信息。

[0084] 物理层（PHY）信息可以与网络组件本身的操作特性或者是性能有关，包括的特性如网络组件可用的支持的链路速率，网络组件可用的不同的操作模式（例如，子集模式）等。

[0085] 链路信息可以与网络组件间的连接的使用相关。链路信息的例子是通信量缓冲丰满度。在另外的例子中，所述链路信息可以包括突发（burstiness）参数（例如，链路上突发的规模，突发的时间间隔，链路闲置的时间等），这些链路信息实现对实际链路使用的确定。其它的实施例是随时间的链路容量的百分比，例如，如果在一段时间内，10G 链路的平均使用总是少于 1G，这是链路使用的有用测量。

[0086] ECE 策略参数可以与管理为网络组件设置的控制策略的分析与 / 或操作的一些参数有关。例如，当全局控制策略管理器 180 使用全局控制策略 185 为设备控制策略生成配置改变时，策略参数可以设置成管理设备的 ECE 操作，包括链路使用阈值，IT 策略，用户参数等。最终，应用信息可以与系统的应用特性相关，系统的应用特性可以管理网络组件的操作。有用的应用信息的例子包括在被分析的网络组件（例如，在没有虚拟化的 L2 切换器上）上运行的数据流，知晓在组件上运行的 AVB 数据流有助于确定是否是低能状态。

[0087] 应注意，本发明的原理能够广泛应用到各种各样的环境中，例如执行 ECE 的所有物理层（例如，背板、双绞线、光纤等）。此外，本发明的原理可以应用到标准或者非标准（例如，2.5G, 5G, 100M, 1G 和 10G 的光纤接口，PON 等）的链路速率，以及将来的链路速率（例如，40G, 100G, 400G, 兆兆位等）。还应注意，本发明的原理可以应用到特定的对称或不对称的链路。

[0088] 全局控制策略 185 基于系统操作，可以是固定的，或者是动态的。在全局控制策略 185 固定的系统中，即使进行调整，全局策略也不改变，全局策略（例如，在设备控制策略层）的物理实施可基于硬件的变化特性而变化。

[0089] 全局控制策略管理器（GCPM）180 也可以基于虚拟作业 130 的改变要求和公差，动态地修改全局控制策略 185。如上所述，全局控制策略管理器 180 配置用于收集系统运行和虚拟作业 130 信息。在一个实施例中，这些信息可以用来动态地、优选地修改全局控制策略 185。

[0090] 图 2 示出了具有连接到多个虚拟机 120A-B 以及切换器 160A-B 的 GCPM180 的系统 200。在一个实施例中，全局控制策略管理器 180 直接或间接（未示出）连接系统 200 的不同的物理组件、虚拟组件和软件组件。所述物理组件包括主机 150A-B 和切换器 160A-B，虚拟组件包括 VM120A-B 和虚拟作业 130，软件组件包括切换器 267A-B。

[0091] 应了解，接收的特定的电能信息组，在电能信息执行上的分析，以及基于所述电能信息执行 GCP 185 的过程，以及虚拟作业 130 的特点，将独立实施。不考虑数据的收集和析机制的使用，明显的是 GCPM 180 是使用 GCP 185 和来自网络组件的电能信息来引导设备控制策略 165A-B 的协调编程的。

[0092] 虚拟机电能管理器（VMPM）

[0093] 图 3 所示的系统 300 的全局控制策略管理器 180 与虚拟机电能管理器 (VMPM) 310A-B 相连。所示的虚拟机电能管理器 310A-B 与虚拟机 120A-B 分别连接,每个虚拟机 120A-B 再与切换器 160A-B 分别连接。

[0094] 虚拟机 120A-B 往往由虚拟机管理器 (也被称为管理程序) 管理,例如虚拟机管理器 (VMM) 320。本领域技术人员应了解在一些执行过程中,虚拟系统的虚拟控制平面 (例如,虚拟机管理器 320) 可以不知道系统的物理部分的不同方面。在虚拟管理器 320 中这种信息的缺少与 ECE 提高有非常重要的联系,例如,虚拟机管理器 320 没有关于不同系统组件的不同的 ECE 性能、控制策略和其他节能特征的信息。

[0095] 在一个实施例中,虚拟机电能管理器 310A-B 收集和转发与切换器 160A-B 相关的 ECE 信息到虚拟机管理器 320,来提高系统 300 的 ECE。在另一个实施例中,虚拟机电能管理器 310A-B 从网络组件中收集 ECE 信息,并使用收集到的信息来生成虚拟机管理器 320 的配置指令。虚拟机电能管理器 (VMPM) 310A-B 的例子可以在美国专利申请 No. TBD (代理备案号 #2875. 3930000),题为“虚拟机电能管理器”中找到,在此将其全文结合引用,以作参考 (“虚拟机电能控制器的应用”)。

[0096] 如图 3 所描述,在一个实施例中,全局控制策略管理器 180 可以从虚拟机电能管理器 310A-B 接收关于系统 300 组件 (切换器 160A-B) 的信息。在一个实施例中,全局控制策略管理器 180 也可以通过虚拟机电能管理器 310A-B 传递配置指令到网络组件。

[0097] GCPM 的放置

[0098] 如图 4 所示的系统 400,对于图 1 中 GCPM 180 不同的实施例,有二者选一的物理与逻辑的配置。每个对 GCPM 480A-E 的放置 (placement) 的描述是非限制性的,并提出能够独立作用或与其他 GCPM 480A-E 组件集合作用的放置。例如,系统 400 可具有单个 GCPM 480A,两个 GCPM 480A-B,或者是全部五个 GCPM 480A-E 组件。

[0099] 在一个实施例中,如图 4 所示,全局控制策略管理器 480A,作为主机 150A 的组件放置。取代如图 1A 所示的外部放置,全局控制策略管理器 480A 示出为主机 150A 的组件。如上所述,在一个实施例中,全局控制策略管理器 180,480A 可以与不容的网络设备连接,例如,切换器 160A 和系统 400 的其它组件。在主机 150A-B 中,全局控制策略管理器的实施例充当软件或者是硬件组件。

[0100] 全局控制策略管理器 480C,如图 4 所示,作为虚拟机管理器 320 的组件放置,例如软件组件或者是“插件”。正如本领域技术人员所知,软件系统,例如,虚拟机管理器 320,可以允许安装追加 (add-on) 软件组件和允许该追加软件与主系统进行交互。在一个实施例中,全局控制策略管理器 480C 设计成作为插件安装到虚拟机管理器 320 软件系统中。正如本领域技术人员所知的,其它的软件间的交互是可能的。

[0101] 全局控制策略管理器 480D 如图 4 的一个实施例所示,作为切换器 160A 的组件放置,例如,分开的、与现有切换器软件 167A 相连的软件组件,或者作为切换器软件 167A 的“插件”组件 (未示出)。正如本领域的技术人员了解的那样,切换器 160A 可以已经安装了切换器软件 167A 和硬件逻辑,这些为切换器设备提供了不同的功能。在一个实施例中,全局控制策略管理器 480D 在切换器 160A 中作为硬件逻辑安装,它可以为系统 400 中的每个组件提供功能。在另一个实施例中,全局控制策略管理器 480D 在切换器 160A 中作为软件组件安装。在另一个实施例中,全局控制策略管理器 480E 被描述为与切换器 160A 连接的

外部组件,这个链路与图 1 所示的实施例中全局控制策略管理器 180 与主机 150A 直接相连成鲜明的对比。

[0102] 图 4 所示的放置不是限制性的。相关领域的技术人员将了解此处描述的全局控制策略管理器 180、480A-E 的功能可定位于在这里所描述的系统中的任何位置,作为软件或者硬件,或者是两者结合来实现。

[0103] 图 5 所示的系统 500 的实施例有虚拟机 (VM) 520,该虚拟机跨越多个物理主机 150A-B,虚拟机管理器 320 和全局控制策略管理器 180A。图 5 所示的实施例表示在虚拟机 520 形式的 ECE 管理的增加的复杂性。

[0104] 虚拟机 520 和其放置横越多个物理设备 (主机 150A-B) 并显示了可变的 ECE 环境。例如,虚拟端口 525A 和 525B 现在可以分别映射到物理端口 555A 和 555B,这些端口潜在地不同的 ECE 特性和控制策略。应注意的是,在图 5 中的实施例,每个主机 150A-B 被描述为分别与不同切换器 160A 和 160B 物理连接。如上所述,全局控制策略管理器 180A 在一个实施例中,可以协调 ECE 性能和在这些复杂连接间的控制策略。如上所述,全局控制策略管理器 180A 可以对系统 500 组件的控制策略重编程来跟随单个集成的全局控制策略。在另外一个实施例中,全局控制策略管理器 180A 生成的配置变化指向设备控制策略,且这些变化发送到设备上执行。

[0105] 在图 5 所示的一个实施例中,全局控制策略管理器 180A 生成的配置变化指向系统 500 的组件。在此的实施例描述的全局控制策略的方法是特别有用的,因为如上所述,虚拟机 520 跨越主机 150A-B。在一个实施例中,当虚拟机管理器 320 横跨虚拟机 520 分配虚拟作业 130 的运算时,全局控制策略管理器 180A 能够应用一致的全局控制策略到附加组件。在一个实施例中,如果虚拟机 520 从只在主机 150A 上运行扩展到在主机 150A-B 上运行,全局控制策略 (GCP) 185 能够实现下述 A1-A3 的功能:

[0106] A1. 重新编程应用于主机 150A 的组件的控制策略,例如虚拟端口 525A 和物理端口 555A。具有额外的主机应用于虚拟人物 130,可以影响应用与主机 150A 的策略。

[0107] A2. 重新编程与新主机 150B 相关的系统组件的控制策略,来跟随全局控制策略 185,例如虚拟端口 525B 和物理端口 555B。

[0108] A3. 全局控制策略管理器 188 执行逻辑可以实现影响在源主机 150A 和新主机 150B 的端口间内部通信量映射的控制策略变化。在完成横跨后,可经物理端口 555B,通过切换器 160B 来路由与虚拟作业 130 相关的通信量,可改变该控制策略以提高新的路径的性能。

[0109] 应注意,步骤 A1-A3 仅是用于说明而非限制性的。在不偏离本发明的精神的情况下,可以对上述步骤进行组合或是修改。

[0110] 例子

[0111] 在图 6 所示的例子中,主机 650 被描述成与切换器 660 和 670 相连接。主机 650 被描述为具有虚拟机 620A-B 和全局控制策略管理器 680,虚拟机 620B 有在其上运行的虚拟作业 630 和参考 (referencing) 全局控制策略 685 的全局控制策略管理器 680。切换器 660 有两个物理端口 665A-C,切换器 670 有两个物理端口 665D-E。切换器 660 和切换器 670 分别具有控制策略 667A-B。在这个实施例中,两个切换器都与网络 101 (未示出) 相连接。在这个例子中,下列特性适用:

[0112] C1. 系统 600 的 ECE 可以从 ECE 策略到虚拟作业 630 的应用中获得好处。例如, 虚拟作业 630 可以是电子商务的应用, 它主要服务单个地理区域, 这样的应用具有长期的闲置时间和特别的、半可预计 (semi-predictable) 的使用模式 (例如, 人口统计的时候所服务的网站是商店的次数)。

[0113] C2. 全局控制策略 685 是虚拟的控制策略, 它被设计来考虑虚拟任务 630 不同的特性。如此所说, 当生成全局控制策略 685 时, 可以考虑一个不同应用特性的宽广的范围。

[0114] C3. 控制策略 667A 是安装在切换器 660 上的、独立地应用的 ECE 策略。应了解, 本领域技术人员在本发明的教导下应了解安装在切换器 660 中的 ECE 控制策略可控制物理端口 665A-C 的 ECE 设置。

[0115] 在一个实施例中, 如上所述的一些系统 600 的组件涉及到了如下的时间表。时间表事件 T1-T7 是说明而非限制性的。本领域技术人员在本发明的教导下, 将了解到时间表事件 T1-T7 是可选择的, 或者可以不同的顺序发生。

[0116] T1. 系统 600, 在示例性实施例的开端中, 虚拟作业 630 在虚拟机 620B 中运行, 切换器 670 并未已经在系统 600 中执行。

[0117] T2. 全局控制策略 680 已依据虚拟作业 630 的要求在切换器 660 中设置了控制策略 667A。

[0118] T3. 在系统 600 的控制进程中, 确定需要新的虚拟机 (620A) 执行虚拟作业 630。如虚拟机电能管理器应用中所描述的那样, 所述确定可以由在主机 650 (未示出) 上操作的 VMPM, 或是通过 VMM 610 的正常操作执行。

[0119] T4. 基于增加的第二虚拟机 (620), 全局控制策略管理器 680 基于接收到的电能信息以及参考全局控制策略 685 确定, 切换器 660 的控制策略 (667A) 应该被修改。

[0120] T5. 使用全局控制策略管理器 680 与切换器 660 直接或间接连接, 全局控制策略管理器重新配置 (重变成) 控制策略 667A 来考虑增加的虚拟机 620A, 以及虚拟作业 630 的影响。这种方式, 全局控制策略 (GCP) 685 维持静态, 但是它在切换器 660 的物理控制策略 667A 上的执行是动态修改的。

[0121] T6. 作为系统 600 的硬件修改的附加例子, 如图 6 所示, 切换器 670 可在系统 600 中执行。这个附加的切换器增加了附加的数据通信量路径和三个以上的物理端口。

[0122] T7. 在实施例中, 在上述事件 T5 以后, 全局控制策略管理器 680 监控系统运行、电能信息, 并将虚拟作业 630 的要求与收集的 ECE 数据进行比较 -- 正如全局控制策略 (GCP) 685 所表达的那样。在实施例中, 如果数据被全局控制策略管理器 680 收集, 这表明全局控制策略 685 的优选的变化, 这个改变可以由全局控制策略管理器 680 来实现。这样, 不仅全局控制策略 685 的物理执行可以是动态的 (适应的), 如 T5 中记录的, 全局控制策略 685 本身也可以是动态修改来配合变化的环境。下面的因素 F1-F4 不是能修改 GCP 685 的环境的示例环境的限制性清单:

[0123] F1. 虚拟作业 630 需要增加的运行电平 (level)。

[0124] F2. 虚拟作业 630 需要增加的节能和能效电平。

[0125] F3. 新的网络组件的可用性允许虚拟作业 630 占用更多的系统资源。

[0126] 相关领域的技术人员在了解在此描述的基础上, 应该明白, 在一实施例中, 当确定是否或怎样改变全局控制策略 685 时, 可使用其他因素。

[0127] 图 7 示意了非限制性的系统 700 的例子,其中全局控制策略管理器 780 管理在具有网络拓扑结构(包括两台主机(750,755)、四个切换器(760A-B,770A-B)和客户端 790)的网络中的电能消耗。每台主机(750,755)分别控制虚拟机 720A-B,且虚拟作业 730 最初由虚拟机 720A 运行。

[0128] 列举的实施例如图 7 所示,切换器 760A-B 和 770A-B 分别具有控制策略 765A-B 和 775A-B。

[0129] 在虚拟机电能管理器应用中,示例性实施例描述了虚拟机管理器 710 和虚拟机 720A-B 从系统 700 的组件接收信息的系统和方法。这些收集到的信息由虚拟机电能管理器 715 用来为虚拟机管理器 710 生成配置指令。

[0130] 广泛地说,以下描述了参照图 7 的示例事件,其中在虚拟机 720A(使用第一通信量路径)上运行的虚拟作业 730 迁移到虚拟机 720B(使用了第二通信量路径)。这一迁移的传统方法一般将使得虚拟作业可适应环境。在下面所述的例子中,在第二通信量路径上,全局控制策略管理器 780 用来应用全局控制策略 785 到第二通信量路径上,全局控制策略 785 被设计来特别地提高虚拟作业 730。

[0131] 特别地,在上述示例事件在系统 700 中发生时,VMM 710,在管理虚拟机 720A-B 的过程中,将虚拟作业 730 从虚拟机 720A 迁移到虚拟机 720B。这个示例迁移事件有助于说明全局控制策略管理器 780 的实施例的功能。本领域的技术人员应该了解到,不同的原因可以影响从虚拟机 720A 迁移虚拟作业 730 到虚拟机 720B,例如,安全性和性能和高可用性。如在虚拟机电能管理器的相关应用中所知,虚拟机电能管理器 715 可以为虚拟机管理器 710 生成配置指令来指示这类型的迁移。

[0132] 基于上面所述的迁移事件,在本申请中,全局控制策略管理器 780 可以使用全局控制策略 785 来引导连接到新主机(755)、与虚拟作业 730 相关联的网络组件的重新配置。在一个实施例中,如果控制策略 775A-B 没有按虚拟作业 730 的要求最优化,全局控制策略管理器 780 可以重新配置控制策略 775A-B。这种对独立物理设备控制策略(775A-B)的重新配置以执行“全局的”全局控制策略 785 是所述实施例的特性。由于这种“全局的”达到,可认为全局控制策略管理器 780 执行控制策略,该控制策略跨越了切换器和系统 700 的其它组件,因此统一了 ECE 策略和改进了优选的 ECE 配置。

[0133] 一个实施例可以被认为是重新编程与迁移的虚拟作业 730 相关联的新物理环境。控制策略 775A-B 可以与全局控制策略管理器 780 的一个优选的形式相协调,而且能够实现更好的输出。在一个实施例中,简单地输入,而不像在传统的方法中要求虚拟作业 730 去适应新的硬件环境那样,全局控制策略管理器 780 调整硬件的环境以适应迁移的虚拟作业 730。

[0134] 可考虑上述示例的不同变化。在一个实施例中,全局控制策略 785 也可以设计成实现在系统 700 上运行的多个虚拟作业的要求间的结合。

[0135] 虚拟全局控制策略:通信量路由/切换器

[0136] 图 8 描述了一个非限制性的网络 800 的例子,其中全局控制策略管理器 880 使用全局控制策略 885 来管理在网络 800 中虚拟作业 830 的 ECE。网络 800 具有网络拓扑结构。该网络拓扑结构包括主机 850,四个切换器(860A-B,870A-B),虚拟机管理器(VMM)810 和客户端 890。由于此处的教导可以应用于通信量路径的选择,也可以一般地应用于处理这些功

能的全部组件,此处使用的术语路由、切换和路由 / 切换可以互换使用。

[0137] 用图 8 所示的一个实施例中,切换器 860A-B 控制其各自的切换器上的 ECE 性能的控制策略 865A-B。在这个例子中,切换器 860A-B 和 870A 具有控制策略 (865A-B, 875), 所述控制策略能够实现 EEE 的性能,以及切换器 870B 没有控制策略,也没有 EEE 性能。

[0138] 虚拟机管理器 810 在实施例中可以在各种网络通信量路径中引导通信量,例如,主机 850 到切换器 860A 到切换器 860B 到客户端 890,或者是主机 850 到切换器 870A 到切换器 860B 到客户端 890。本领域的技术人员知道这个路由 / 切换可以通过不同的方式来实现,包括数据包的标记和切换器基于标记路由 / 切换。由于此处的教导可以应用于通信量路径的选择,也可以一般地应用于处理这些功能的全部组件,此处使用的术语路由、切换和路由 / 切换可以互换使用。

[0139] 在这个实施例中,全局控制策略管理器 880 已经收集了控制策略和图 8 所示的所有硬件组件的特性,而且基于虚拟作业 830 的要求和全局控制策略 885 对他们进行了分析。本领域的技术人员应该了解的,基于在此给出的描述,全局控制策略管理器 880 可以与虚拟机管理器 880 合作,通过调整和编程控制策略 865A-B 和 875,使网络 800 的通信量路径成形。

[0140] 在这个例子中,由于切换器 860A-B 具有 ECE 性能和切换器 870B 上缺乏 ECE 性能,全局控制策略管理器 880 可以重新编程切换器 860A-B 控制策略 (865A-B),为虚拟作业 830 通信增加他们的 ECE 效率。在一个实施例中,全局控制策略管理器 880 可以通过改变控制策略 865A-B 的配置来“成形路径”和虚拟机管理器 810 可以管理沿着路径的通信路由。这个路由 / 切换可以由虚拟机管理器 810 使用上述数据包标记的方法来运行。本领域的技术人员应该了解的,其它通信量路径和方法可以被设想出来。

[0141] 图 8 还示意了在网络 800 中的第二非限制例子,在实施例中,在全局控制策略管理器 880 的 ECE 策略和性能协调中,其可以是“通信量知晓 (traffic aware)”。在这个例子中,虚拟作业 830 要求网络路径保持高优先权 (例如,视频流,VoIP) 通信量和低优先权 (例如,HTML) 通信量。本领域的技术人员将会理解,在不同的通信量类型上可以设计不同的 ECE 策略 / 方法。在第二例子中,全局控制策略管理器 880 可以设置控制策略 865A-B (控制切换器 860A-B),来缓冲低优先权标准通信量和快速地转发 VoIP 高优先权通信量。在这个实施例中,全局控制策略管理器 880 与图 8 中切换器 860A-B 和 870A-B 相连,以及基于 ECE 策略和其它信息可以接收特别的电能信息。在一个实施例中,全局控制策略管理器 880 的协调功能可以沿着一条选择的通信量路径来实现控制策略的统一。

[0142] 在另一个实施例中,全局控制策略管理器 880 也可以是有性能意识的 (performance-aware)——衡量在不同的路由 / 切换设备中重新配置控制策略的性能结论 (performance implication)。例如,如果性能考虑规定的话,如上所述的四个切换器 (860A-B, 870A-B) 可以具有全局控制策略管理器 880 重新配置控制策略,以不考虑 ECE 地增加所有通信量的速度。本领域的技术人员,了解在此所揭示的原理,将意识到上面所述的全局控制策略管理器 880 功能的应用和好处。

[0143] 这个基于虚拟作业 830 要求的全局控制策略配置,不是传统地 VMM 810 可用的。

[0144] 同样应注意的是,全局控制策略管理器 880 的一个实施例可以不依赖于 EEE 控制策略改变具有较低电能签名 (signature) 的链路的通信量,例如,将 1G 通信量从 10G 的链

路重路由到 1G 的链路并采用该方法提高的更长的反应时间。

#### [0145] 方法

[0146] 这个部分和图 9 总结了展示的一个举例方法 900 的框图描述的技术,方法 900 是在具有多个网络组件的网络中管理电能控制和能效 (ECE) 机制的方法。虽然方法 900 是参照本发明的一个实施例描述的,但其并非限制性的,其可在其他应用中使用。

[0147] 如图 9 所示,方法 900 的实施例从步骤 910 开始,在此全局控制策略管理器 (GCPM) 从多个网络组件中的一个网络组件接收电能信息。在一个实施例中,全局控制策略管理器 780 接收电能信息,例如从图 7 中的切换器 770A-B 中接收,在上面讨论过的 ECE 信息。一旦步骤 910 完成,方法 900 进行到步骤 920。

[0148] 在步骤 920 中,分析电能信息。在一个实施例中,电能信息例如来自切换器 770A-B 的 ECE 信息,由全局控制策略管理器 780 分析。一旦步骤 920 完成,方法 900 进行到步骤 930。

[0149] 在步骤 930 中,控制策略修改是基于全局控制策略和接收到的电能信息生成的。在一个实施例中,控制策略修改是基于 GCP 785 和虚拟作业 130 从主机 750 到主机 755 (电能信息) 的迁移的对控制策略修改。一旦步骤 930 完成,方法 900 进行到步骤 940。

[0150] 在步骤 940 中,将控制策略修改发送到多个网络组件中的一个网络组件。在一个实施例中,全局控制策略管理器 780 发送控制策略修改给切换器 770A-B,在此,修改控制策略 775A-B。一旦步骤 940 完成,方法 900 结束。

[0151] 在此,管理器的功能(例如,全局控制策略管理器 GCPM,虚拟机电能管理器 (VMPM),虚拟机管理器 (VMM) 等)可以在硬件、软件或者硬件软件的结合中实现。例如,基于在此给出的讨论,本领域技术人员可以理解,全局控制策略管理器功能可以使用计算机处理器、计算机逻辑和专用集成电路 (ASIC) 等中实现。相应地,任何执行在此描述的数据收集,策略管理,协调、分析功能的任何处理器都落入本发明的范围和精神下。例如,主机 150A-B 的实施例是使用处理器来实现主机功能的计算机服务器。

[0152] 进一步,在此描述的全局控制策略管理器功能可通过可由计算机处理器或上面列举的任何硬件设备执行的计算机程序指令来实施。计算机程序指令使得处理器执行在此描述的全局控制策略管理器的功能。计算机程序指令(例如,软件)可以存储在计算机或处理器可以访问的计算机可用的介质,计算机程序介质,或者是任何计算机可读的存储介质中。所述的介质包括存储器设备例如 RAM 或 ROM,再或者是类似的计算机存储介质,如计算机磁盘或者是 CD 盘或类似物。相应地,任何具有使得处理器执行数据收集,策略管理、协调、分析功能以及在此描述的相关功能的处理器的计算机代码的计算机存储介质都落入本发明的范围和精神。

#### [0153] 总结

[0154] 虽然以上描述了本发明的各种实施例,应当理解,其目的仅在于举例说明,而没有限制性。本领域的技术人员知悉,在不离开本发明的精神和范围情况下,在形式上和细节上还可做各种的改变。因此,本发明的保护范围不当仅局限于以上描述的任一实施例,而应该依照权利要求及其等同来限定。

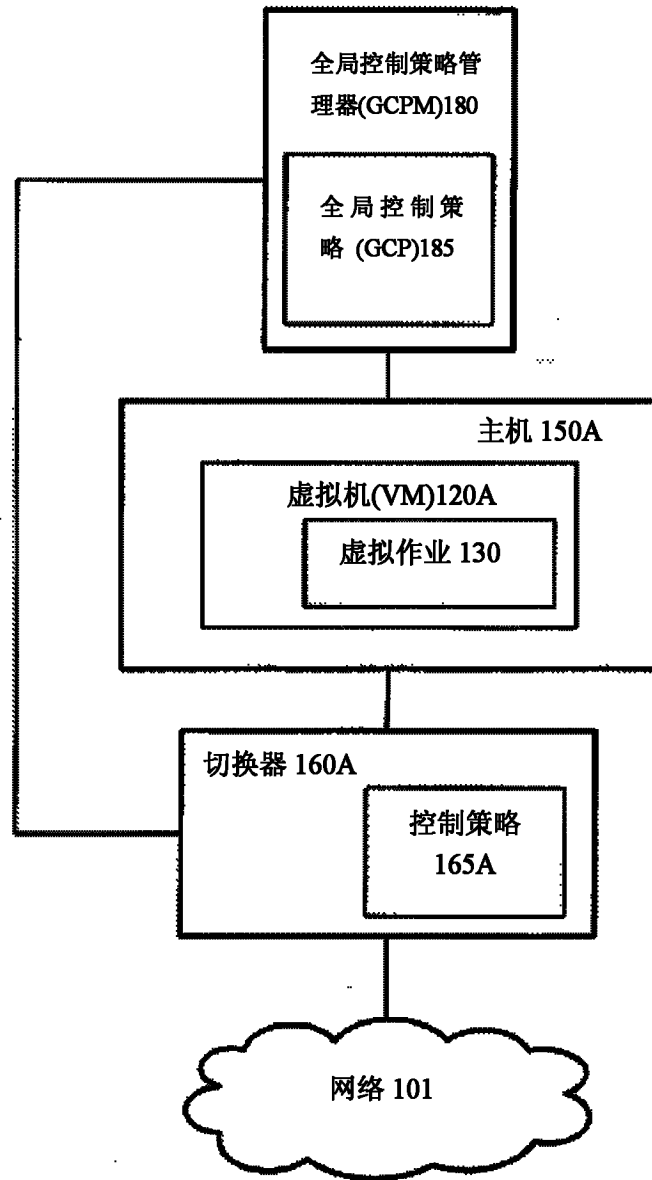


图 1

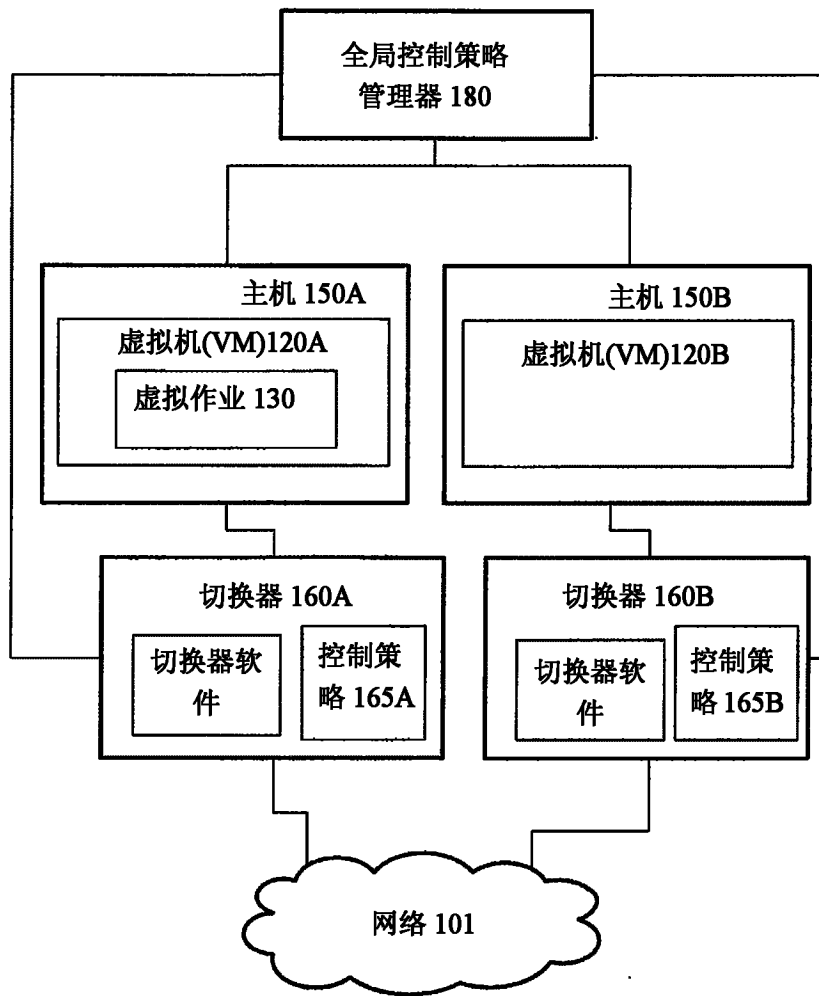


图 2

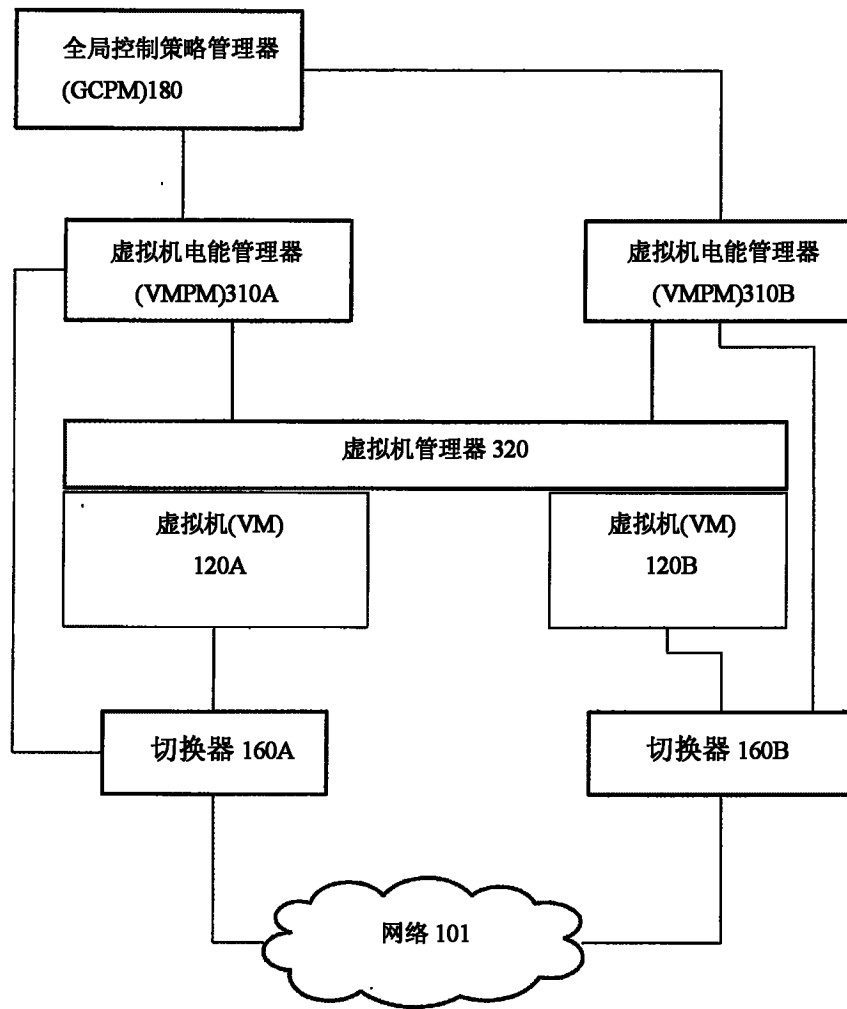


图 3

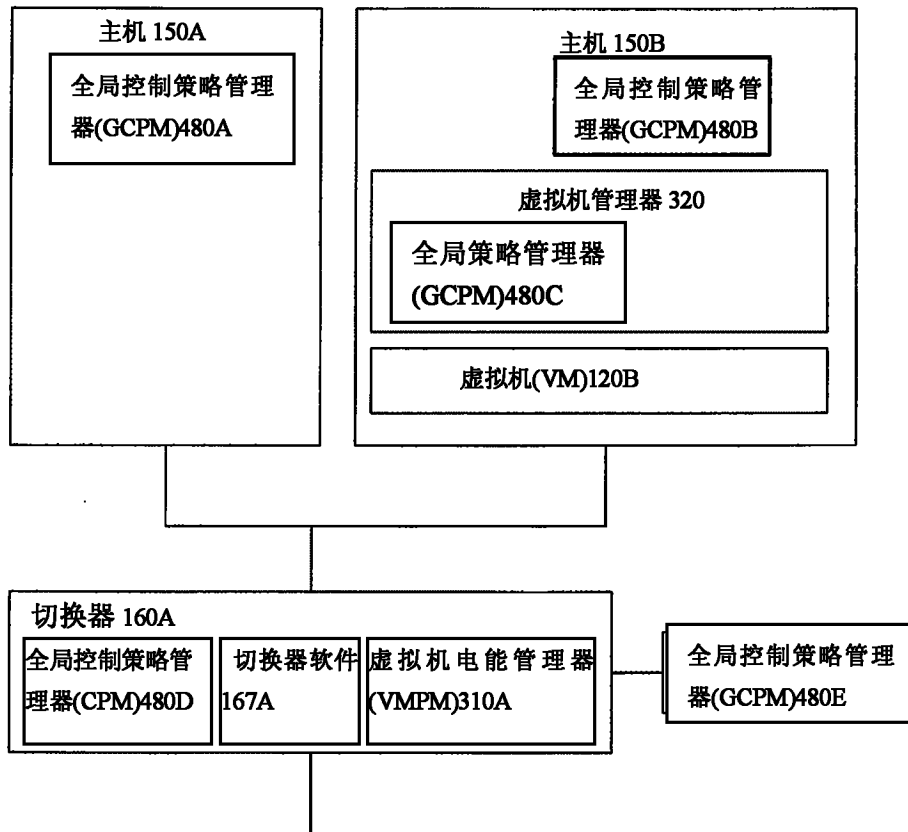


图 4

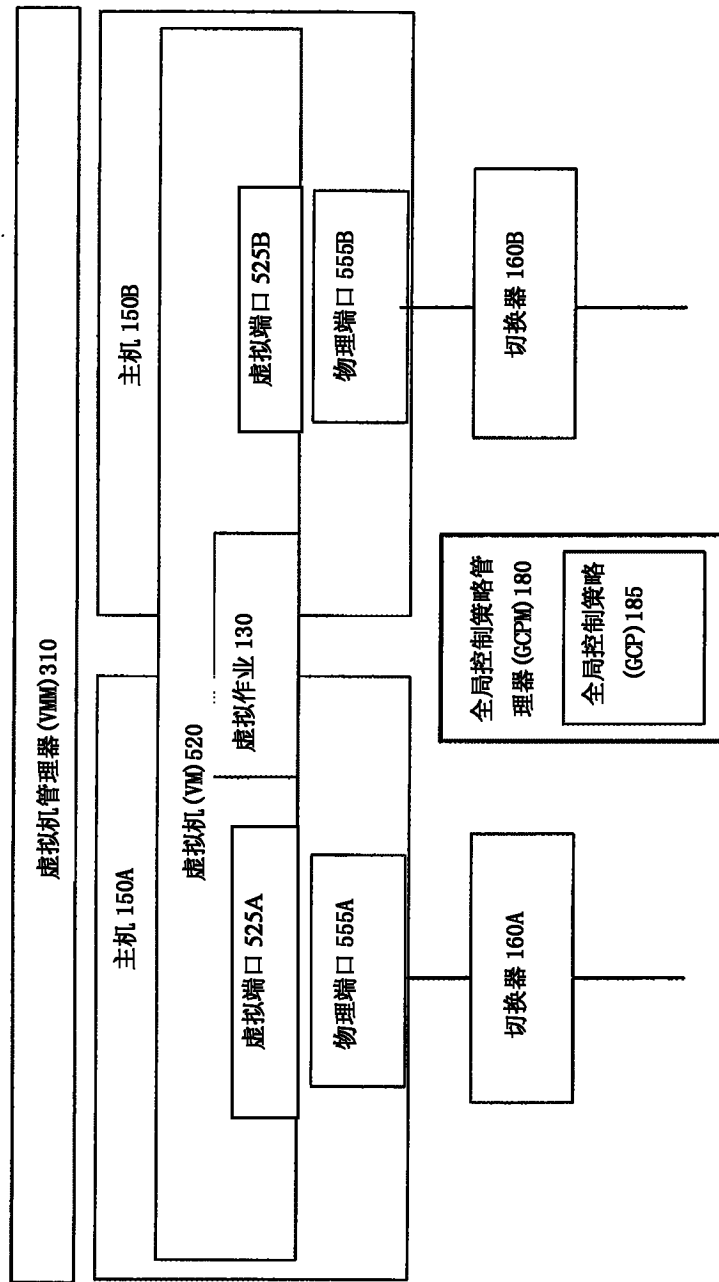


图 5

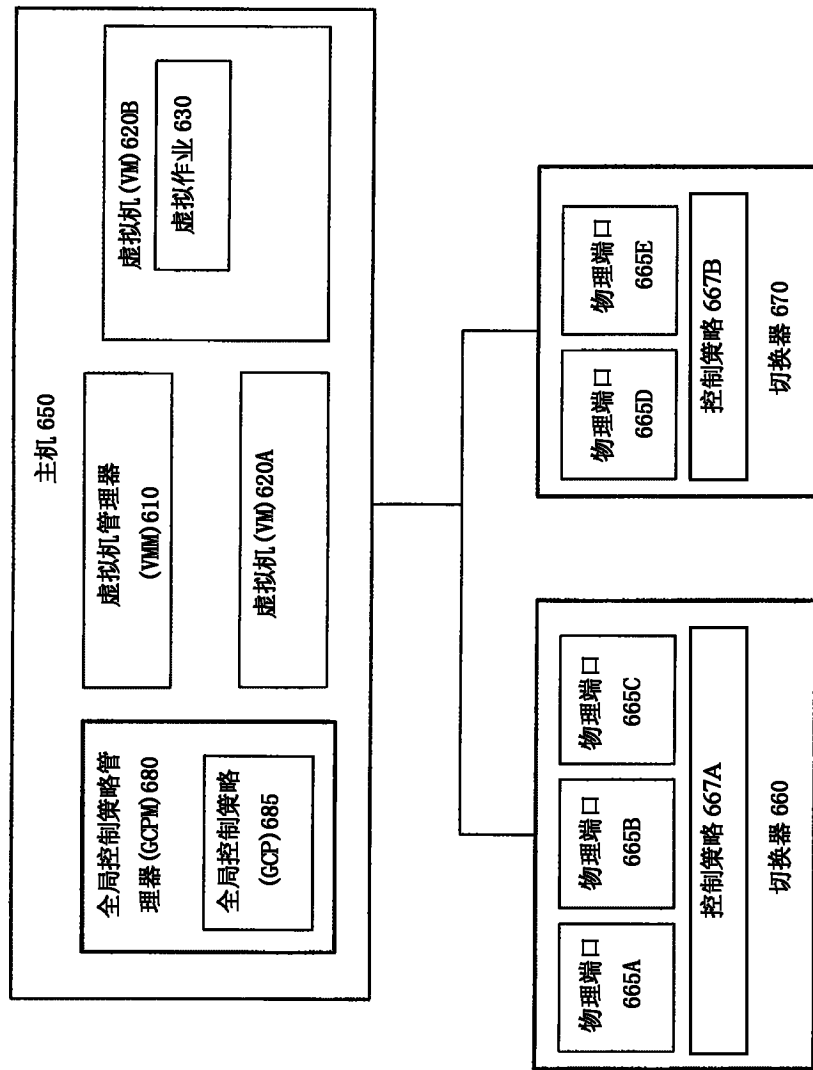


图 6

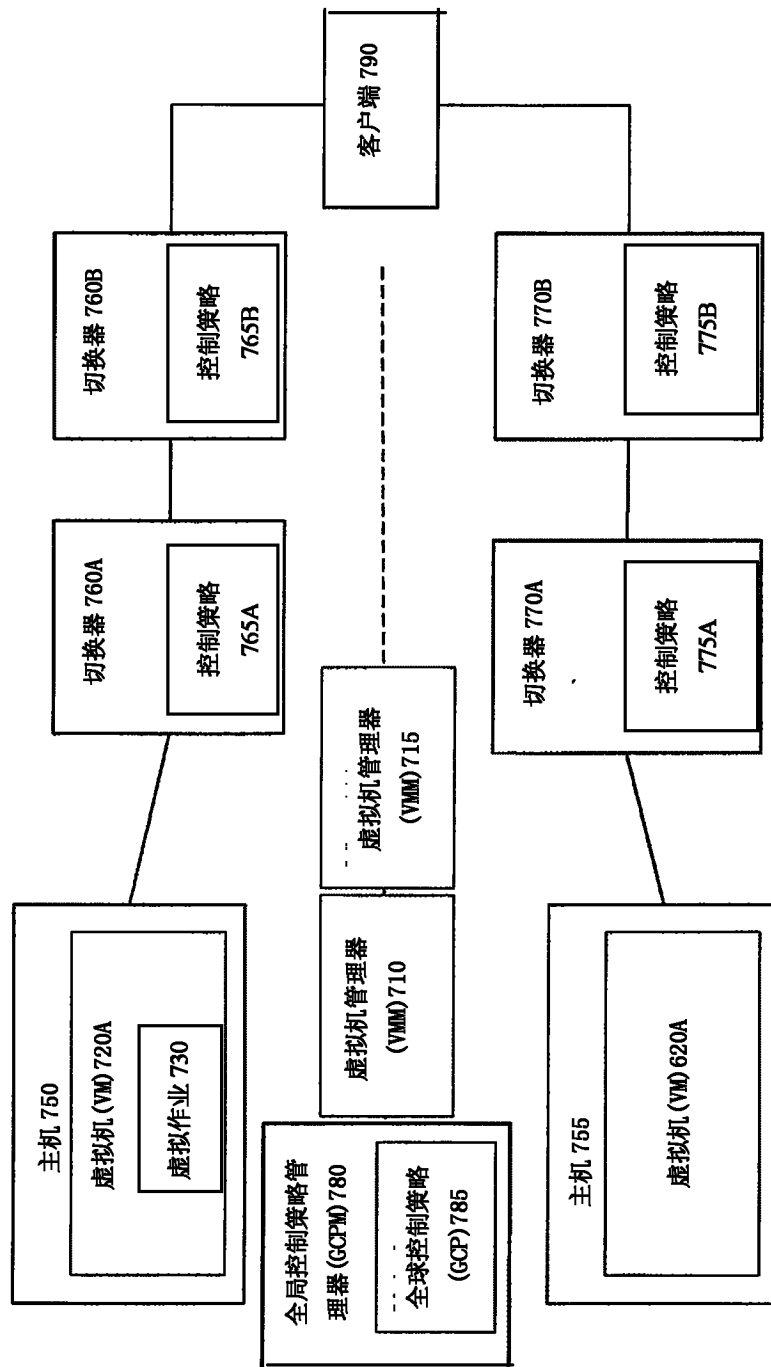


图 7

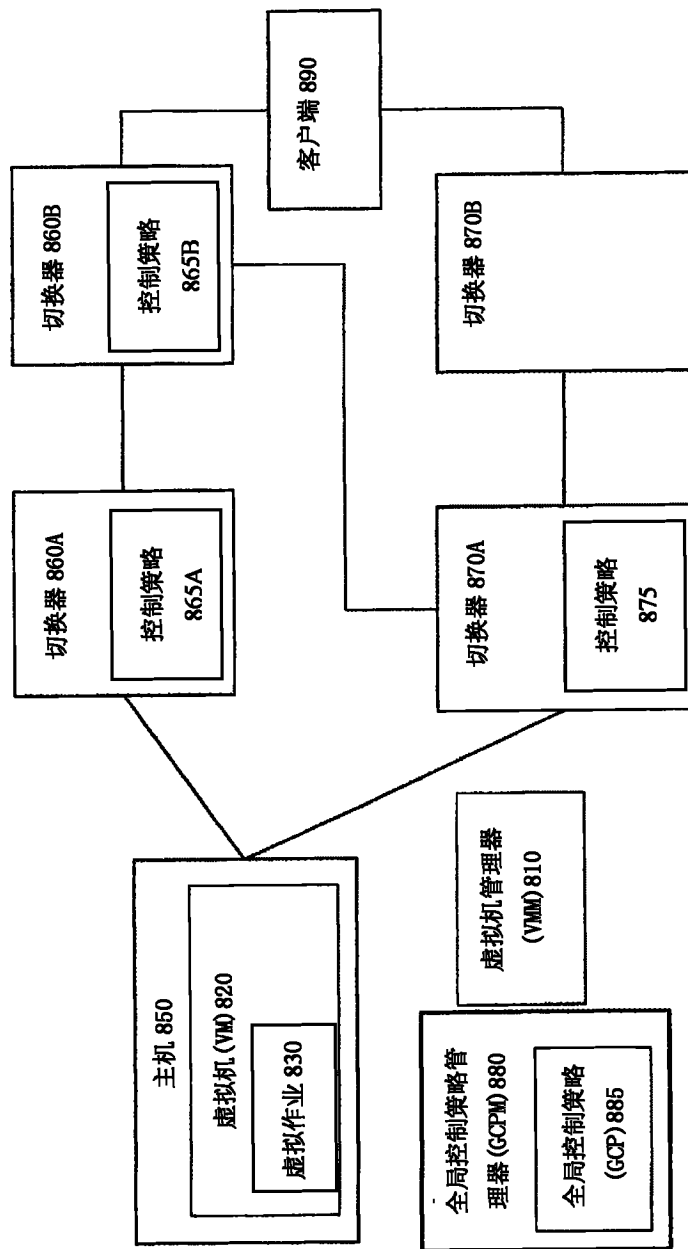


图 8

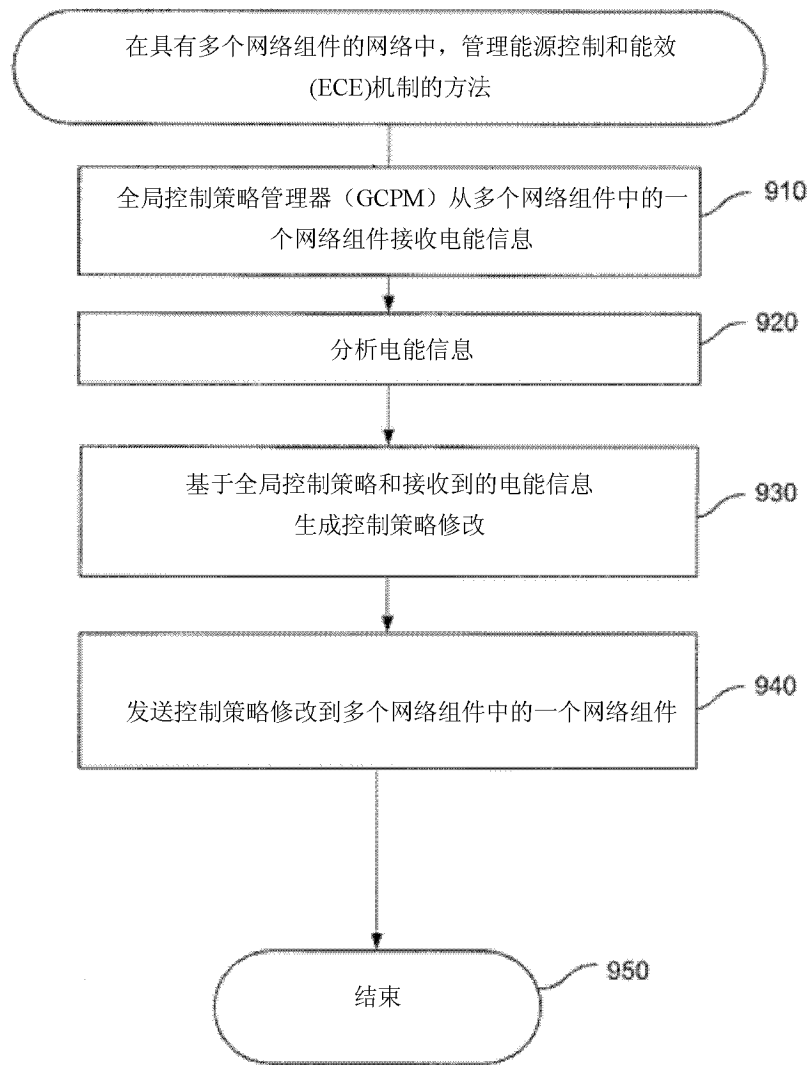


图 9