

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 6 月 4 日 (04.06.2020)



(10) 国际公布号
WO 2020/108536 A1

(51) 国际专利分类号:
H04L 12/24 (2006.01)

(21) 国际申请号: PCT/CN2019/121297

(22) 国际申请日: 2019 年 11 月 27 日 (27.11.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201811430362.9 2018 年 11 月 28 日 (28.11.2018) CN

(71) 申请人: 深圳先进技术研究院 (SHENZHEN INSTITUTE OF ADVANCED TECHNOLOGY) [CN/CN]; 中国广东省深圳市南山区西丽大学城学苑大道 1068 号, Guangdong 518055 (CN)。

(72) 发明人: 叶志勇 (YE, Zhiyong); 中国广东省深圳市南山区西丽大学城学苑大道 1068 号, Guangdong 518055 (CN)。王洋 (WANG, Yang); 中国广东省深圳市南山区西丽大学城学苑大道 1068 号, Guangdong 518055 (CN)。须成忠 (XU, Chengzhong); 中国广东省深圳市南山区西丽大学城学苑大道 1068 号, Guangdong 518055 (CN)。

(74) 代理人: 深圳市科进知识产权代理事务所 (普通合伙) (SHENZHEN KEJIN INTELLECTUAL PROPERTY OFFICE); 中国广东省深圳市南山区粤兴三道二号虚拟大学园产业化基地 A701 室, Guangdong 518055 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB,

(54) Title: VIRTUAL NETWORK RESOURCE ALLOCATION METHOD AND SYSTEM AND ELECTRONIC DEVICE

(54) 发明名称: 一种虚拟网络资源分配方法、系统及电子设备

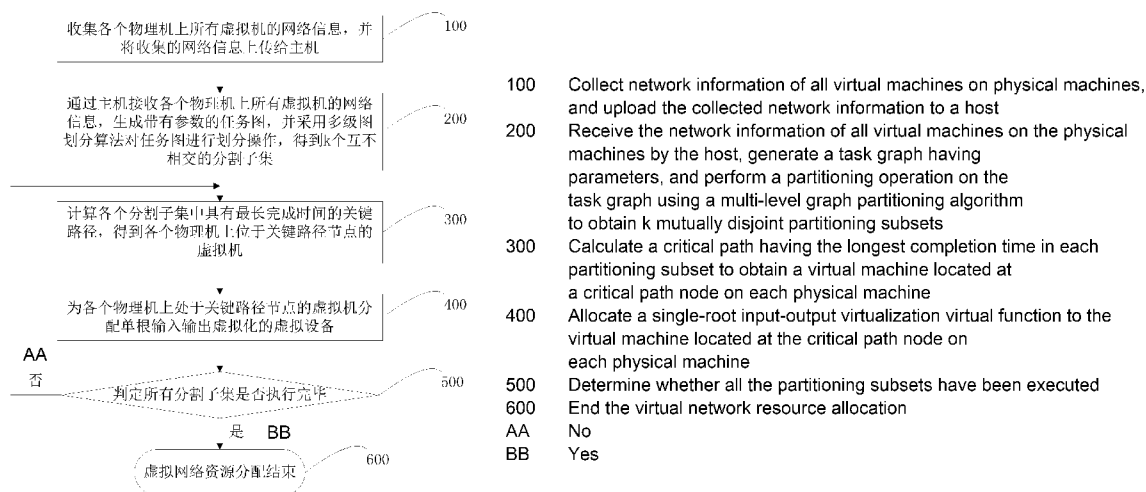


图 1

(57) Abstract: The present application relates to a virtual network resource allocation method and system and an electronic device. The method comprises: step a, generating a task graph according to network information of all virtual machines on physical machines, and partitioning the task graph using a multi-level graph partitioning algorithm to obtain k mutually disjoint partitioning subsets; step b, calculating a critical path having the longest completion time in each partitioning subset to obtain a virtual machine located at a critical path node; and step c, allocating a single-root input-output virtualization virtual function to the virtual machine located at the critical path node. In the present application, paravirtualization technology and single-root input-output virtualization technology are used in combination to optimize the network; by default, each virtual machine is assigned a paravirtualization virtual network adapter; single-root input/output virtualization virtual functions are only allocated to virtual machines located at critical path nodes

GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

after multi-level graph partitioning; thus, the purpose of improving the overall network bandwidth utilization and reducing the overall execution time of parallel applications is achieved.

(57) 摘要: 本申请涉及一种虚拟网络资源分配方法、系统及电子设备。所述方法包括: 步骤a: 根据物理机上所有虚拟机的网络信息生成任务图, 并采用多级图划分算法对所述任务图进行划分, 得到k个互不相交的分割子集; 步骤b: 计算各个分割子集中具有最长完成时间的关键路径, 得到位于关键路径节点的虚拟机; 步骤c: 为所述处于关键路径节点的虚拟机分配单根输入输出虚拟化的虚拟设备。本申请采用半虚拟化技术与单根输入输出虚拟化技术相结合的方式优化网络, 默认每个虚拟机都分配有半虚拟化的虚拟网卡, 而单根输入输出虚拟化的虚拟设备仅分配给多级图划分后的位于关键路径节点的虚拟机, 以达到提高总体网络带宽利用率和降低并行应用总体执行时间的目的。

一种虚拟网络资源分配方法、系统及电子设备

技术领域

本申请属于计算机虚拟化技术领域，特别涉及一种虚拟网络资源分配方法、系统及电子设备。

背景技术

由于近几年来计算机硬件成本的不断减低和单个计算机计算能力的不断加强，现在并行计算已越来越流行并被广泛接受。同时，为了数据中心资源的高效利用，往往采用虚拟化技术以提高数据中心的资源利用率。所谓虚拟化技术就是指能够在一台物理机上运行多台虚拟机技术，并且该虚拟机在用户、应用软件甚至操作系统看来，几乎与物理机没有区别。虽然这能大幅度提高资源利用率，但也由此引出了不同虚拟机间资源分配的问题，因此，如何有效的分配虚拟资源并最大程度的提高资源利用率一直受到广泛的关注和研究。

在当前的并行网络和分布式网络环境中，对网络资源的并行调度已经成为提升网络性能的一个极其重要的技术。但由于现阶段数据中心应用层与底层脱节、应用与网络脱节，这必将会引起网络资源调度的语义鸿沟，而使整个系统的运行缺乏协调性，从而导致网络资源的利用率不高。

发明内容

本申请提供了一种虚拟网络资源分配方法、系统及电子设备，旨在至少在一定程度上解决现有技术中的上述技术问题之一。

为了解决上述问题，本申请提供了如下技术方案：

一种虚拟网络资源分配方法，包括以下步骤：

步骤 a：根据物理机上所有虚拟机的网络信息生成任务图，并采用多级图划分算法对所述任务图进行划分，得到 k 个互不相交的分割子集；

步骤 b：计算各个分割子集中具有最长完成时间的关键路径，得到位于关

键路径节点的虚拟机；

步骤 c：为所述处于关键路径节点的虚拟机分配单根输入输出虚拟化的虚拟设备。

本申请实施例采取的技术方案还包括：所述步骤 a 前还包括：收集各个物理机上所有虚拟机的网络信息；所述网络信息的收集方式为：每个虚拟机通过读取 UNIX 操作系统为用户访问内核提供的伪文件系统获取网络信息，所述虚拟机获取的网络信息放入存储系统中，各个物理机存储系统读取所有虚拟机的网络信息，并将所述网络信息上传至主机。

本申请实施例采取的技术方案还包括：在所述步骤 a 中，所述根据物理机上所有虚拟机的网络信息生成任务图，并采用多级图划分算法对任务图进行划分具体为：所述主机接收所有物理机上传的网络信息，根据所述网络信息生成带有参数的任务图，并采用多级递归二分划分方法对所述任务图进行划分。

本申请实施例采取的技术方案还包括：所述采用多级递归二分划分方法对任务图进行划分具体为：所述多级递归二分划分方法包括粗化、初始划分和细化三个阶段，所述粗化阶段在输入任务图后产生一系列复杂度更小的聚合图，通过所述聚合图的最大邻接对的顶点降低任务图的复杂度，当所述任务图的复杂度降到预设顶点数量时结束粗化阶段；所述初始划分阶段使用启发式算法对任务图进行初始划分，所述细化阶段将初始划分后的任务图逐步恢复。

本申请实施例采取的技术方案还包括：在所述步骤 b 中，所述关键路径的计算公式为：

$$\begin{aligned} \forall path \quad P \in PathSet, \\ PathCompTime(CP) \geq PathCompTime(P) \end{aligned}$$

上述公式中，*path* 指的是任务图中的某一路径，*PathSet* 指的是任务图中

的全体路径集合，*PathCompTime* 是指路径中所有节点执行完成的时间，符号 *P* 是指任意路径，*CP* 是指关键路径。

本申请实施例采取的另一技术方案为：一种虚拟网络资源分配系统，包括物理机和主机，所述主机包括：

多级图划分模块：用于根据物理机上所有虚拟机的网络信息生成任务图，并采用多级图划分算法对所述任务图进行划分，得到 *k* 个互不相交的分割子集；

关键路径计算模块：用于计算各个分割子集中具有最长完成时间的关键路径，得到位于关键路径节点的虚拟机；

网络资源分配模块：用于为所述处于关键路径节点的虚拟机分配单根输入输出虚拟化的虚拟设备。

本申请实施例采取的技术方案还包括：所述物理机包括网络资源监控模块，所述网络资源监控模块用于收集物理机上所有虚拟机的网络信息，并将收集的网络信息上传给网络资源管理模块；所述网络信息的收集方式为：每个虚拟机通过读取 UNIX 操作系统为用户访问内核提供的伪文件系统获取网络信息，所述虚拟机获取的网络信息放入存储系统中，各个物理机存储系统读取所有虚拟机的网络信息，并将所述网络信息上传至网络资源管理模块。

本申请实施例采取的技术方案还包括：所述主机还包括网络资源管理模块，所述网络资源管理模块用于接收所有物理机上传的网络信息，并将网络信息输入多级图划分模块；所述多级图划分模块根据网络信息生成带有参数的任务图，并采用多级递归二分划分方法对所述任务图进行划分。

本申请实施例采取的技术方案还包括：所述采用多级递归二分划分方法对任务图进行划分具体为：所述多级递归二分划分方法包括粗化、初始划分和细化三个阶段，所述粗化阶段在输入任务图后产生一系列复杂度更小的聚合图，

通过所述聚合图的最大邻接对的顶点降低任务图的复杂度，当所述任务图的复杂度降到预设顶点数量时结束粗化阶段；所述初始划分阶段使用启发式算法对任务图进行初始划分，所述细化阶段将初始划分后的任务图逐步恢复。

本申请实施例采取的技术方案还包括：所述关键路径的计算公式为：

$$\begin{aligned} \forall path \quad P \in PathSet, \\ PathCompTime(CP) \geq PathCompTime(P) \end{aligned}$$

上述公式中，*path* 指的是任务图中的某一路径，*PathSet* 指的是任务图中的全体路径集合，*PathCompTime* 是指路径中所有节点执行完成的时间，符号 *P* 是指任意路径，*CP* 是指关键路径。

本申请实施例采取的又一技术方案为：一种电子设备，包括：

至少一个处理器；以及

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行上述的虚拟网络资源分配方法的以下操作：

步骤 a：根据物理机上所有虚拟机的网络信息生成任务图，并采用多级图划分算法对所述任务图进行划分，得到 *k* 个互不相交的分割子集；

步骤 b：计算各个分割子集中具有最长完成时间的关键路径，得到位于关键路径节点的虚拟机；

步骤 c：为所述处于关键路径节点的虚拟机分配单根输入输出虚拟化的虚拟设备。

相对于现有技术，本申请实施例产生的有益效果在于：本申请实施例的虚拟网络资源分配方法、系统及电子设备采用半虚拟化技术与单根输入输出虚拟

化技术相结合的方式优化网络，默认每个虚拟机都分配有半虚拟化的虚拟网卡，而单根输入输出虚拟化的虚拟设备仅分配给多级图划分后的位于关键路径节点的虚拟机，以达到提高总体网络带宽利用率和降低并行应用总体执行时间的目的。

附图说明

图 1 是本申请实施例的虚拟网络资源分配方法的流程图；

图 2 是本申请实施例的虚拟网络资源分配系统的结构示意图；

图 3 是本申请实施例的虚拟网络资源分配方法的硬件设备结构示意图。

具体实施方式

为了使本申请的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本申请进行进一步详细说明。应当理解，此处所描述的具体实施例仅用以解释本申请，并不用于限定本申请。

针对现有技术存在的问题，本申请实施例的虚拟网络资源分配方法采用半虚拟化技术与单根输入输出虚拟化技术(SR-IOV, Single-Root I/O Virtualization)相结合的方式优化网络，默认每个虚拟机都分配有半虚拟化的虚拟网卡，而单根输入输出虚拟化的虚拟设备(VF, Virtual Function)仅分配给多级图划分后的位于关键路径(Critical Path)节点的虚拟机，以达到提高总体网络带宽利用率和降低并行应用总体执行时间的目的。

具体地，请参阅图 1，是本申请实施例的虚拟网络资源分配方法的流程图。本申请实施例的虚拟网络资源分配方法包括以下步骤：

步骤100：收集各个物理机上所有虚拟机的网络信息，并将收集的网络信息上传给主机；

步骤 100 中，网络信息的收集方式为：每个虚拟机通过读取 UNIX 操作系统为用户访问内核提供的伪文件系统/proc/net/dev 来获取网络信息，虚拟机获取的网络信息放入存储系统（XenStore）中，各个物理机上分别设有网络资源监控模块，网络资源监控模块通过 XenStore 读取各个虚拟机的网络信息，并发送给网络资源管理模块。

步骤200：通过主机接收各个物理机上所有虚拟机的网络信息，生成带有参数的任务图，并采用多级图划分算法对任务图进行划分操作，得到k个互不相交的分割子集；

步骤200中，主机上包括网络资源管理模块和多级图划分模块，网络资源管理模块负责接收各个物理机上所有虚拟机的网络信息，并将网络信息输入多级图划分模块，生成带有参数的任务图，多级图划分模块采用多级递归二分划分方法对任务图进行划分，得到k个互不相交的分割子集，不同分割子集对应的任务分配给不同的物理机进行计算。本申请通过对任务图进行多级图划分，从而平衡各个物理机之间的计算量，并使各个物理机之间的总通信量最小。

具体地，多级递归二分划分方法包括粗化、初始划分和细化三个阶段。粗化阶段在输入任务图后将产生一系列复杂度更小的聚合图，目的在于通过聚合图的最大邻接对的顶点以降低任务图的复杂度，当任务图的复杂度降到预设顶点数量（通常顶点数量为几百个，具体顶点数量可根据实际应用进行设定）时结束粗化阶段。在初始划分阶段，因为任务图的复杂度已经很小了，使用经典简单的启发式算法如K-L算法对任务图进行初始划分。在细化阶段，将初始划分后的任务图逐步恢复，将聚合在一起的顶点在不同的分割子集间移动，提高任务图的划分效果。

步骤300：计算各个分割子集中具有最长完成时间的关键路径，得到各个物理机上位于关键路径节点的虚拟机；

步骤300中，由于在同一物理机间的虚拟机通信开销可视为零，因此在任务图划分后的关键路径计算只计算不同分割子集间的通信开销。具体地，关键路径的计算公式为：

$$\forall path \quad P \in PathSet, \\ PathCompTime(CP) \geq PathCompTime(P) \quad (1)$$

公式（1）中，*path*指的是任务图中的某一路径，*PathSet*指的是任务图中的全体路径集合，*PathCompTime*是指路径中所有节点执行完成的时间，符号*P*是指任意路径，*CP*是指关键路径。

步骤400：为各个物理机上处于关键路径节点的虚拟机分配单根输入输出虚拟化的虚拟设备；

步骤 400 中，本申请实施例的网络资源分配方法是默认让每个虚拟机都分配有半虚拟化的虚拟网卡，而单根输入输出虚拟化的虚拟设备仅分配给多级图划分后的位于关键路径节点的虚拟机。其中半虚拟化的虚拟网卡与单根输入输出虚拟化的虚拟设备之间透明切换使用了多网卡绑定技术，单根输入输出虚拟化的虚拟设备被设置为主网卡，半虚拟化的虚拟网卡被设置为从网卡，当主网卡有效的时候虚拟机使用主网卡进行通信，只有在主网卡无效的时候才使用从网卡进行通信，从而减少网络开销。

步骤 500：判定所有分割子集是否执行完毕，如果所有分割子集执行完毕，执行步骤 600；否则，返回步骤 300 继续计算下一分割子集的关键路径计算，直到所有分割子集执行完毕；

步骤 600：虚拟网络资源分配结束。

请参阅图2，是本申请实施例的虚拟网络资源分配系统的结构图。本申请实施例的虚拟网络资源分配系统包括物理机和主机，物理机包括网络资源监控模块，主机包括网络资源管理模块、多级图划分模块、关键路径计算模块、网络资源分配模块和任务执行判断模块。

网络资源监控模块：用于收集各个物理机上所有虚拟机的网络信息，并将收集的网络信息上传给网络资源管理模块；其中，网络信息的收集方式为：每个虚拟机通过读取UNIX操作系统为用户访问内核提供的伪文件系统 `/proc/net/dev` 来获取网络信息，虚拟机获取的网络信息放入XenStore中，网络资源监控模块通过XenStore读取各个虚拟机的网络信息，并发送给网络资源管理模块。

网络资源管理模块：用于接收各个物理机上所有虚拟机的网络信息，并将网络信息输入多级图划分模块；

多级图划分模块：用于根据输入的网络信息生成带有参数的任务图，并采用多级图划分算法对任务图进行划分操作，得到 k 个互不相交的分割子集；其中，多级图划分模块采用多级递归二分划分方法对任务图进行划分，得到 k 个互不相交的分割子集，不同分割子集对应的任务分配给不同的物理机进行计算。本申请通过对任务图进行多级图划分，从而平衡各个物理机之间的计算量，并使各个物理机之间的总通信量最小。

具体地，多级递归二分划分方法包括粗化、初始划分和细化三个阶段。粗化阶段在输入任务图后将产生一系列复杂度更小的聚合图，目的在于通过聚合图的最大邻接对的顶点以降低任务图的复杂度，当任务图的复杂度降到只有几百个顶点时结束粗化阶段。在初始划分阶段，因为任务图的复杂度已经很小了，使用经典简单的启发式算法如K-L算法对任务图进行初始划分。在细化阶段，

将初始划分后的任务图逐步恢复，将聚合在一起的顶点在不同的分割子集间移动，提高任务图的划分效果。

关键路径计算模块：用于计算各个分割子集中具有最长完成时间的关键路径，得到各个物理机上位于关键路径节点的虚拟机；其中，由于在同一物理机间的虚拟机通信开销可视为零，因此在任务图划分后的关键路径计算只计算不同分割子集间的通信开销。具体地，关键路径的计算公式为：

$$\begin{aligned} \forall path \quad P \in PathSet, \\ PathCompTime(CP) \geq PathCompTime(P) \quad (1) \end{aligned}$$

公式（1）中，*path*指的是任务图中的某一路径，*PathSet*指的是任务图中的全体路径集合，*PathCompTime*是指路径中所有节点执行完成的时间，符号*P*是指任意路径，CP是指关键路径。

网络资源分配模块：用于为各个物理机上处于关键路径节点的虚拟机分配单根输入输出虚拟化的虚拟设备；其中，本申请实施例的网络资源分配方法是默认让每个虚拟机都分配有半虚拟化的虚拟网卡，而单根输入输出虚拟化的虚拟设备仅分配给多级图划分后的位于关键路径节点的虚拟机。其中半虚拟化的虚拟网卡与单根输入输出虚拟化的虚拟设备之间透明切换使用了多网卡绑定技术，单根输入输出虚拟化的虚拟设备被设置为主网卡，半虚拟化的虚拟网卡被设置为从网卡，当主网卡有效的时候虚拟机使用主网卡进行通信，只有在主网卡无效的时候才使用从网卡进行通信，从而减少网络开销。

任务执行判断模块：用于判定所有分割子集是否执行完毕，如果所有分割子集执行完毕，则虚拟网络资源分配结束；否则，通过关键路径计算模块继续计算下一分割子集的关键路径计算，直到所有分割子集执行完毕。

图3是本申请实施例提供的虚拟网络资源分配方法的硬件设备结构示意图。

如图3所示，该设备包括一个或多个处理器以及存储器。以一个处理器为例，该设备还可以包括：输入系统和输出系统。

处理器、存储器、输入系统和输出系统可以通过总线或者其他方式连接，图3中以通过总线连接为例。

存储器作为一种非暂态计算机可读存储介质，可用于存储非暂态软件程序、非暂态计算机可执行程序以及模块。处理器通过运行存储在存储器中的非暂态软件程序、指令以及模块，从而执行电子设备的各种功能应用以及数据处理，即实现上述方法实施例的处理方法。

存储器可以包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需要的应用程序；存储数据区可存储数据等。此外，存储器可以包括高速随机存取存储器，还可以包括非暂态存储器，例如至少一个磁盘存储器件、闪存器件、或其他非暂态固态存储器件。在一些实施例中，存储器可选包括相对于处理器远程设置的存储器，这些远程存储器可以通过网络连接至处理系统。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

输入系统可接收输入的数字或字符信息，以及产生信号输入。输出系统可包括显示屏等显示设备。

所述一个或者多个模块存储在所述存储器中，当被所述一个或者多个处理器执行时，执行上述任一方法实施例的以下操作：

步骤 a：根据物理机上所有虚拟机的网络信息生成任务图，并采用多级图划分算法对所述任务图进行划分，得到 k 个互不相交的分割子集；

步骤 b：计算各个分割子集中具有最长完成时间的关键路径，得到位于关键路径节点的虚拟机；

步骤 c：为所述处于关键路径节点的虚拟机分配单根输入输出虚拟化的虚拟设备。

上述产品可执行本申请实施例所提供的方法，具备执行方法相应的功能模块和有益效果。未在本实施例中详尽描述的技术细节，可参见本申请实施例提供的方法。

本申请实施例提供了一种非暂态（非易失性）计算机存储介质，所述计算机存储介质存储有计算机可执行指令，该计算机可执行指令可执行以下操作：

步骤 a：根据物理机上所有虚拟机的网络信息生成任务图，并采用多级图划分算法对所述任务图进行划分，得到 k 个互不相交的分割子集；

步骤 b：计算各个分割子集中具有最长完成时间的关键路径，得到位于关键路径节点的虚拟机；

步骤 c：为所述处于关键路径节点的虚拟机分配单根输入输出虚拟化的虚拟设备。

本申请实施例提供了一种计算机程序产品，所述计算机程序产品包括存储在非暂态计算机可读存储介质上的计算机程序，所述计算机程序包括程序指令，当所述程序指令被计算机执行时，使所述计算机执行以下操作：

步骤 a：根据物理机上所有虚拟机的网络信息生成任务图，并采用多级图划分算法对所述任务图进行划分，得到 k 个互不相交的分割子集；

步骤 b：计算各个分割子集中具有最长完成时间的关键路径，得到位于关键路径节点的虚拟机；

步骤 c：为所述处于关键路径节点的虚拟机分配单根输入输出虚拟化的虚拟设备。

本申请实施例的虚拟网络资源分配方法、系统及电子设备采用半虚拟化技术与单根输入输出虚拟化技术相结合的方式优化网络，默认每个虚拟机都分配有半虚拟化的虚拟网卡，而单根输入输出虚拟化的虚拟设备仅分配给多级图

划分后的位于关键路径节点的虚拟机，以达到提高总体网络带宽利用率和降低并行应用总体执行时间的目的。

以上所述仅是本发明的优选实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本发明原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也应视为本发明的保护范围。

权 利 要 求

1、一种虚拟网络资源分配方法，其特征在于，包括以下步骤：

步骤 a：根据物理机上所有虚拟机的网络信息生成任务图，并采用多级图划分算法对所述任务图进行划分，得到 k 个互不相交的分割子集；

步骤 b：计算各个分割子集中具有最长完成时间的关键路径，得到位于关键路径节点的虚拟机；

步骤 c：为所述处于关键路径节点的虚拟机分配单根输入输出虚拟化的虚拟设备。

2、根据权利要求 1 所述的虚拟网络资源分配方法，其特征在于，所述步骤 a 前还包括：收集各个物理机上所有虚拟机的网络信息；所述网络信息的收集方式为：每个虚拟机通过读取 UNIX 操作系统为用户访问内核提供的伪文件系统获取网络信息，所述虚拟机获取的网络信息放入存储系统中，各个物理机存储系统读取所有虚拟机的网络信息，并将所述网络信息上传至主机。

3、根据权利要求 2 所述的虚拟网络资源分配方法，其特征在于，在所述步骤 a 中，所述根据物理机上所有虚拟机的网络信息生成任务图，并采用多级图划分算法对任务图进行划分具体为：所述主机接收所有物理机上传的网络信息，根据所述网络信息生成带有参数的任务图，并采用多级递归二分划分方法对所述任务图进行划分。

4、根据权利要求 3 所述的虚拟网络资源分配方法，其特征在于，所述采用多级递归二分划分方法对任务图进行划分具体为：所述多级递归二分划分方法包括粗化、初始划分和细化三个阶段，所述粗化阶段在输入任务图后产生一系列复杂度更小的聚合图，通过所述聚合图的最大邻接对的顶点降低任务图的复

杂度，当所述任务图的复杂度降到预设顶点数量时结束粗化阶段；所述初始划分阶段使用启发式算法对任务图进行初始划分，所述细化阶段将初始划分后的任务图逐步恢复。

5、根据权利要求 1 至 4 任一项所述的虚拟网络资源分配方法，其特征在于，在所述步骤 b 中，所述关键路径的计算公式为：

$$\begin{aligned} \forall path \ P \in PathSet, \\ PathCompTime(CP) \geq PathCompTime(P) \end{aligned}$$

上述公式中，*path* 指的是任务图中的某一路径，*PathSet* 指的是任务图中的全体路径集合，*PathCompTime* 是指路径中所有节点执行完成的时间，符号 *P* 是指任意路径，*CP* 是指关键路径。

6、一种虚拟网络资源分配系统，包括物理机和主机，其特征在于，所述主机包括：

多级图划分模块：用于根据物理机上所有虚拟机的网络信息生成任务图，并采用多级图划分算法对所述任务图进行划分，得到 *k* 个互不相交的分割子集；

关键路径计算模块：用于计算各个分割子集中具有最长完成时间的关键路径，得到位于关键路径节点的虚拟机；

网络资源分配模块：用于为所述处于关键路径节点的虚拟机分配单根输入输出虚拟化的虚拟设备。

7、根据权利要求 6 所述的虚拟网络资源分配系统，其特征在于，所述物理机包括网络资源监控模块，所述网络资源监控模块用于收集物理机上所有虚拟机的网络信息，并将收集的网络信息上传给网络资源管理模块；所述网络信息的收集方式为：每个虚拟机通过读取 UNIX 操作系统为用户访问内核提供的伪文件系统获取网络信息，所述虚拟机获取的网络信息放入存储系统中，各个物理机存储系统读取所有虚拟机的网络信息，并将所述网络信息上传至网络资源管理模块。

8、根据权利要求 7 所述的虚拟网络资源分配系统，其特征在于，所述主机还包括网络资源管理模块，所述网络资源管理模块用于接收所有物理机上传的网络信息，并将网络信息输入多级图划分模块；所述多级图划分模块根据网络信息生成带有参数的任务图，并采用多级递归二分划分方法对所述任务图进行划分。

9、根据权利要求 8 所述的虚拟网络资源分配系统，其特征在于，所述采用多级递归二分划分方法对任务图进行划分具体为：所述多级递归二分划分方法包括粗化、初始划分和细化三个阶段，所述粗化阶段在输入任务图后产生一系列复杂度更小的聚合图，通过所述聚合图的最大邻接对的顶点降低任务图的复杂度，当所述任务图的复杂度降到预设顶点数量时结束粗化阶段；所述初始划分阶段使用启发式算法对任务图进行初始划分，所述细化阶段将初始划分后的任务图逐步恢复。

10、根据权利要求 6 至 9 任一项所述的虚拟网络资源分配系统，其特征在于，所述关键路径的计算公式为：

$$\begin{aligned} \forall path \quad P \in PathSet, \\ PathCompTime(CP) \geq PathCompTime(P) \end{aligned}$$

上述公式中，*path* 指的是任务图中的某一路径，*PathSet* 指的是任务图中的全体路径集合，*PathCompTime* 是指路径中所有节点执行完成的时间，符号 *P* 是指任意路径，*CP* 是指关键路径。

11、一种电子设备，包括：

至少一个处理器；以及

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行上述 1 至 5 任一项所述的

虚拟网络资源分配方法的以下操作：

步骤 a：根据物理机上所有虚拟机的网络信息生成任务图，并采用多级图划分算法对所述任务图进行划分，得到 k 个互不相交的分割子集；

步骤 b：计算各个分割子集中具有最长完成时间的关键路径，得到位于关键路径节点的虚拟机；

步骤 c：为所述处于关键路径节点的虚拟机分配单根输入输出虚拟化的虚拟设备。

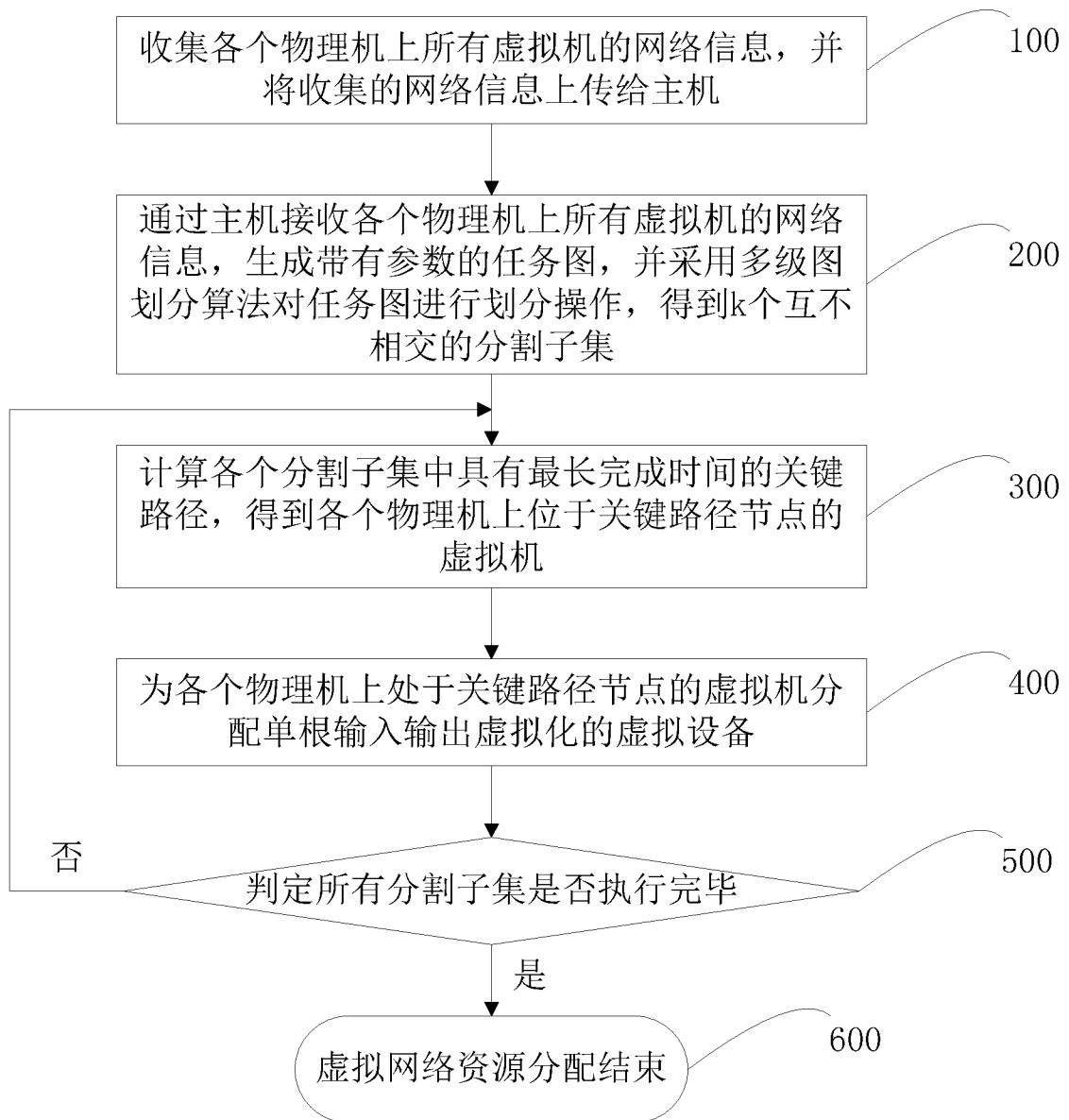


图 1

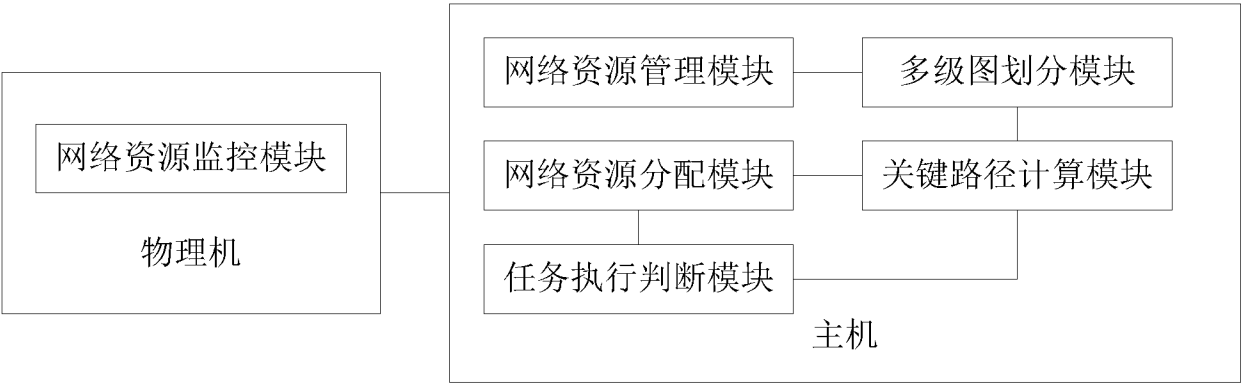


图 2

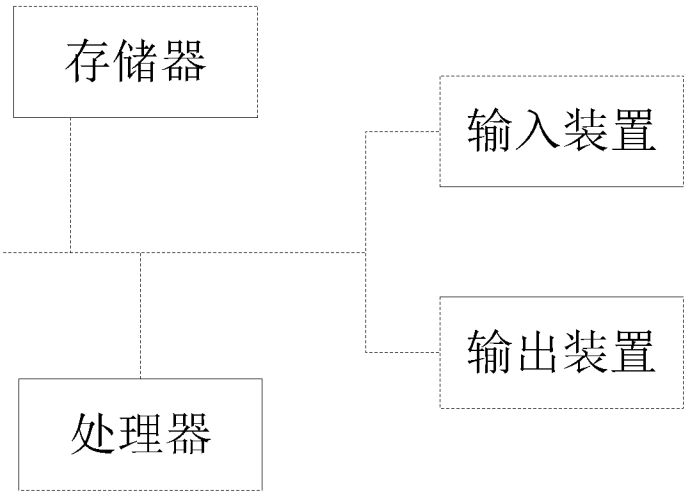


图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/121297

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/24(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04L, G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, CNKI, VEN, WOTXT, USTXT, EPTXT: 虚拟机, 图, 子集, 最长, 路径, 链路, 关键, 重要, 节点, 单根输入输出, SR-IOV, virtual machine, graph, subset, path, link, key, important, node

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109412865 A (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY) 01 March 2019 (2019-03-01) claims 1-11	1-11
A	CN 105357322 A (INSTITUTE OF INFORMATION ENGINEERING, CHINESE ACADEMY OF SCIENCES) 24 February 2016 (2016-02-24) entire document	1-11
A	CN 108205461 A (ARRAY NETWORKS CO., LTD.) 26 June 2018 (2018-06-26) entire document	1-11



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

17 February 2020

Date of mailing of the international search report

27 February 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/121297

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)		Publication date (day/month/year)
CN	109412865	A	01 March 2019	None		
CN	105357322	A	24 February 2016	CN	105357322	B 26 April 2019
CN	108205461	A	26 June 2018	None		

国际检索报告

国际申请号

PCT/CN2019/121297

A. 主题的分类 H04L 12/24 (2006.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类														
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) H04L, G06F 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) CNABS, CNTXT, CNKI, VEN, WOTXT, USTXT, EPTXT: 虚拟机, 图, 子集, 最长, 路径, 链路, 关键, 重要, 节点, 单根输入输出, SR-IOV, virtual machine, graph, subset, path, link, key, important, node														
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 109412865 A (深圳先进技术研究院) 2019年 3月 1日 (2019 - 03 - 01) 权利要求1-11</td> <td>1-11</td> </tr> <tr> <td>A</td> <td>CN 105357322 A (中国科学院信息工程研究所) 2016年 2月 24日 (2016 - 02 - 24) 全文</td> <td>1-11</td> </tr> <tr> <td>A</td> <td>CN 108205461 A (华耀中国科技有限公司) 2018年 6月 26日 (2018 - 06 - 26) 全文</td> <td>1-11</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 109412865 A (深圳先进技术研究院) 2019年 3月 1日 (2019 - 03 - 01) 权利要求1-11	1-11	A	CN 105357322 A (中国科学院信息工程研究所) 2016年 2月 24日 (2016 - 02 - 24) 全文	1-11	A	CN 108205461 A (华耀中国科技有限公司) 2018年 6月 26日 (2018 - 06 - 26) 全文	1-11
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求												
PX	CN 109412865 A (深圳先进技术研究院) 2019年 3月 1日 (2019 - 03 - 01) 权利要求1-11	1-11												
A	CN 105357322 A (中国科学院信息工程研究所) 2016年 2月 24日 (2016 - 02 - 24) 全文	1-11												
A	CN 108205461 A (华耀中国科技有限公司) 2018年 6月 26日 (2018 - 06 - 26) 全文	1-11												
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。														
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件														
国际检索实际完成的日期 2020年 2月 17日		国际检索报告邮寄日期 2020年 2月 27日												
ISA/CN的名称和邮寄地址 中国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 王晓玲 电话号码 86-(010)-62089464												

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/121297

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	109412865	A	2019年 3月 1日	无	
CN	105357322	A	2016年 2月 24日	CN 105357322	B 2019年 4月 26日
CN	108205461	A	2018年 6月 26日	无	