



(12)发明专利

(10)授权公告号 CN 108319359 B

(45)授权公告日 2020.01.03

(21)申请号 201810007242.1

(22)申请日 2018.01.04

(65)同一申请的已公布的文献号
申请公布号 CN 108319359 A

(43)申请公布日 2018.07.24

(66)本国优先权数据
201711497729.4 2017.12.29 CN

(73)专利权人 华中科技大学
地址 430074 湖北省武汉市洪山区珞喻路
1037号

(72)发明人 陆枫 石子倩 刘若雪 吴松
金海

(74)专利代理机构 北京海虹嘉诚知识产权代理
有限公司 11129
代理人 何志欣 侯越玲

(51)Int.Cl.

G06F 1/329(2019.01)

(56)对比文件

US 2012072481 A1,2012.03.22,
CN 104754063 A,2015.07.01,
CN 107370799 A,2017.11.21,
Yang Zhang.Offloading in Mobile
Cloudlet Systems with Intermittent
Connectivity.《IEEE》.2015,
何亚昕.移动云计算资源分配与功率控制研
究.《中国优秀硕士学位论文全文数据库》.2017,

审查员 廖瑜

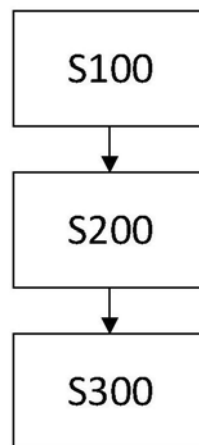
权利要求书3页 说明书14页 附图4页

(54)发明名称

一种基于微云的能耗优化方法

(57)摘要

本发明涉及一种基于微云的能耗优化方法,包括:构建一个微云系统,微云系统包括:至少两个微云和无线连接于微云的移动设备,微云能够为与之无线连接的移动设备提供云计算服务;获取微云系统中与移动设备相关的系统数据用于分析移动设备需要在本地执行的数据量并根据数据量设置移动设备的第一处理器的初始运行频率 $F_{Linitial}$;根据任务队列的可容忍延迟范围 T 使用马尔可夫决策过程动态决策任务队列中单个任务的调度策略以及设置与调度策略对应的移动设备的第一处理器的当前运行频率 F_L 以使移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低。



1. 一种基于微云的能耗优化方法,其特征在于,所述方法包括:

构建一个微云系统,所述微云系统包括:至少两个微云和无线连接于微云的移动设备,所述微云能够为与之无线连接的移动设备提供云计算服务;

获取所述微云系统中与所述移动设备相关的系统数据用于分析所述移动设备需要在本地执行的数据量并根据所述数据量设置所述移动设备的第一处理器的初始运行频率 $F_{Linitial}$;

根据任务队列的可容忍延迟范围 T 使用马尔可夫决策过程动态决策所述任务队列中单个任务的调度策略以及设置与所述调度策略对应的所述移动设备的第一处理器的当前运行频率 F_L 以使所述移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低。

2. 如权利要求1所述的方法,其特征在于,所述获取所述微云系统中与所述移动设备相关的系统数据用于分析出所述移动设备需要在本地执行的数据量并根据所述数据量设置所述移动设备的第一处理器的初始运行频率 $F_{Linitial}$ 包括:

根据所述至少两个微云的分布性质和持有所述移动设备的用户的移动特性计算出所述移动设备在随所述用户移动过程中可向所述微云系统迁移的任务数量期望 E_c ;

根据所述任务数量期望 E_c 获取所述移动设备需要在本地执行的数据量;

根据所述数据量设置所述移动设备的第一处理器的初始运行频率 $F_{Linitial}$,

其中,所述分布性质在二维空间上服从泊松分布,所述用户的移动特性服从随机移动模型。

3. 如权利要求1或2所述的方法,其特征在于,所述根据所述任务队列的可容忍延迟范围 T 使用马尔可夫决策过程动态决策所述任务队列中单个任务的调度策略以及设置与所述调度策略对应的所述移动设备的第一处理器的当前运行频率 F_L 以使所述移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低包括:

设置所述移动设备的状态空间和动作空间,其中,所述状态空间包括:执行的任务状态、可连接的微云的状态和所述第一处理器的运行频率的调节范围,所述动作空间包括第一动作和第二动作,所述第一动作表示任务在本地执行,所述第二动作表示将任务传输到所述微云执行;

根据所述状态空间和所述动作空间,通过值迭代得出最优的所述调度策略,所述移动设备根据调度策略设置所述第一处理器的当前运行频率 F_L 。

4. 如权利要求2所述的方法,其特征在于,所述获取所述微云系统中与所述移动设备相关的系统数据用于分析出所述移动设备需要在本地执行的数据量并根据所述数据量设置所述移动设备的第一处理器的初始运行频率 $F_{Linitial}$ 包括:

根据所述用户的移动区域的微云实际分布情况和分布性质,得出所述移动区域的微云分布密度 λ_c 和覆盖范围半径 R ,其中,所述分布性质在二维空间上服从泊松分布;

根据服从随机移动模型的所述移动特性,得出用户移动速度 V ;

获得所述移动设备上待执行的任务队列,所述任务队列共有 M 个任务,所述任务队列的可容忍延迟范围为 T ,每个任务的传输数据分别为 $\{DT_1, DT_2, \dots, DT_M\}$,每个任务的计算数据分别为 $\{DC_1, DC_2, \dots, DC_M\}$;

根据用户移动的速度和用户移动的起点与终点的距离和随机几何理论,计算出用户在移动过程中经过微云覆盖区域的覆盖时间 T_c ;

根据所述微云传输带宽 B_c 和所述微云的第二处理器的运行频率 F_c 计算出 M 个任务在所述微云执行的平均时间 \bar{T} ,根据所述覆盖时间 T_c 和所述平均时间 \bar{T} ,获取可迁移任务期望

$$E_c = \frac{T_c}{\bar{T}};$$

根据所述可迁移任务期望 E_c 计算出所述移动设备需要在本地执行的任务数量 $E_L = M - E_c$;

根据所述移动设备需要在本地执行的任务数量 E_L 、所述 M 个任务和所述每个任务的计算数据计算需要所述第一处理器在本地计算的数据量 D_L ,根据所述数据量 D_L 、所述可容忍延迟范围 T 和第一处理器的特性参数,设置所述初始运行频率 $F_{Linitial}$ 。

5.如权利要求4所述的方法,其特征在于,所述覆盖时间 T_c 的计算公式为:

$$T_c = \frac{L(1 - e^{-\pi R^2 \lambda_c})}{v}, \text{ 其中, } L \text{ 是用户移动的起点与终点间的距离};$$

$$\text{所述平均时间 } \bar{T} \text{ 的计算公式为: } \bar{T} = \frac{\sum_{i=1}^M DT_i + \sum_{i=1}^M DC_i}{B_c + F_c}。$$

6.如权利要求1或4所述的方法,其特征在于,所述根据所述任务队列的可容忍延迟范围 T 使用马尔可夫决策过程动态决策所述任务队列中单个任务的调度策略以及设置与所述调度策略对应的所述移动设备的第一处理器的当前运行频率 F_L 以使所述移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低包括:

设置用户状态空间

$$\Theta \in \{S = (i, N, F_L) \in S | i \in G, N \in \{0, 1, \dots, \bar{N}\}, F_L \in \{f_1, f_2, \dots, f_j\}\} \text{ 和}$$

动作空间 $A = \{0, 1\}$,

其中, i 表示当前执行的任务, N 表示当前可连接的微云个数, F_L 表示所述第一处理器的运行频率,所述动作空间 A 中的0表示任务在本地执行,所述动作空间 A 中的1表示将任务传输到所述微云执行;

根据微云的泊松分布特性和微云的分布密度 λ_c ,计算出用户当前可连接的微云个数概率 $P^N(N, N')$,得出微云的状态转移矩阵 \bar{N} ,

其中, N 为当前状态 S 可连接的微云个数, N' 为转移到下一个状态 S' 可能连接的微云个数;

根据当前任务的 DT_i 、 DC_i 以及所采取的动作 A_i ,计算出当前任务成功转移到下一个任务的概率 $P^G(G, G' | A_i)$,获得任务的状态转移矩阵 \bar{G} ;

根据采取的动作 A_i ,计算出所述第一处理器的频率转移概率 $P^F(F_L, F'_L | A_i)$,得出频率的状态转移矩阵 \bar{F} ,

其中, F_L 为在当前状态 S 时,所述第一处理器的运行频率, F'_L 为转移到下一个状态时移动设备可能采取的所述第一处理器的运行频率;

根据优化目标 $V_t(S) = \min_{\emptyset(S, A_i)} H_t(S, A_i)$,通过值迭代得出调度策略

$\phi_t^*(S) = \arg \min_{\phi(S, A_i)} H_t(S, A_i)$, 根据调度策略设置第一处理器的当前运行频率 F_L ,

其中, $H_t(S, A_i)$ 是用户在状态 S 时采取动作 A_i 时产生的能耗, $\phi(S, A_i)$ 是用户在当前状态 S 下采取动作 A_i 的概率。

7. 如权利要求6所述的方法, 其特征在于, 所述当前任务成功转移到下一个任务的概率的计算公式为:

$$P^G(G, G' | A_i) = \begin{cases} \left(\frac{DT_i}{B_c} + \frac{DC_i}{F_c} \right) \times \frac{3V}{R} \times \eta_a & \text{当 } A_i = 1 \text{ 时} \\ 1 & \text{当 } A_i = 0 \text{ 时} \end{cases},$$

其中, $\frac{3V}{R}$ 是用户与微云连接的平均时间, η_a 是当前连接的微云接收用户的任务迁移请求的概率。

8. 如权利要求7所述的方法, 其特征在于, 所述当前连接的微云接受用户的任务迁移请求的概率 η_a 是根据所述当前连接的微云的资源占用情况和/或任务排队情况得出的。

9. 如权利要求6所述的方法, 其特征在于, 所述优化目标 $V_t(S) = \min_{\phi(S, A_i)} H_t(S, A_i)$ 中的 $H_t(S, A_i) = C(S, A_i) + \gamma \sum P(S, S' | A_i) V_{t+1}(S')$,

其中, $C(S, A_i)$ 是在状态 S 时采取动作 A_i 时所述第一处理器产生的直接能耗, γ 是在所述下一个状态 S' 时所述第一处理器产生的能耗的折扣因子。

10. 一种微云系统, 其特征在于, 所述微云系统包括: 至少两个微云和无线连接于微云的移动设备, 所述微云能够为与之无线连接的移动设备提供云计算服务;

第一设备获取所述微云系统中与所述移动设备相关的系统数据用于分析所述移动设备需要在本地执行的数据量并根据所述数据量设置所述移动设备的第一处理器的初始运行频率 $F_{Linitial}$;

所述第一设备根据任务队列的可容忍延迟范围 T 使用马尔可夫决策过程动态决策所述任务队列中单个任务的调度策略以及设置与所述调度策略对应的所述移动设备的第一处理器的当前运行频率 F_L 以使所述移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低。

一种基于微云的能耗优化方法

技术领域

[0001] 本发明涉及云计算的任务调度策略技术领域,尤其涉及一种基于微云的能耗优化方法和系统。

背景技术

[0002] 随着移动设备端应用和数据的井喷式增长,移动设备的能耗问题日渐严峻。但是移动设备由容量十分有限的电池供电,电池能量密度发展大大滞后于移动设备的能耗增加速度。

[0003] 移动设备能耗的不断增加使得人们对移动设备节能问题日益关注。其中,最广泛使用的方法包括数据传输节能、任务迁移以及CPU频率调节。而本发明关注的任务迁移以及CPU频率调节领域同样也受到了影响。现有的任务迁移策略主要包括面向cloud端的任务迁移和面向cloudlet端的任务迁移。

[0004] 其中,面向cloud端的任务迁移主要是由远端的云服务器分担用户的任务。由于cloud的物理位置原因,存在以下缺点:

[0005] (1) 数据传输延迟较高,使得用户获取任务结果的时间过长;

[0006] (2) 数据传输能耗较高,采用远端cloud服务器有时反而会加剧移动端的能耗;

[0007] (3) 当使用CPU频率调节时,只能利用相邻任务在云端执行返回结果的时间差来调节频率,调频幅度较小,无法有效节能。

[0008] 面向cloudlet的任务迁移,主要是利用在二维空间上随机分布,离用户地理位置较近的cloudlet来分担用户的任务,达到减少网络延迟,降低用户能耗的目的。而现有的面向cloudlet的任务迁移存在以下缺点:

[0009] (1) 由于cloudlet分布的随机性与用户移动的随机性,现有研究还无法准确预判用户可迁移的任务量以及任务迁移成功与否,这使得cloudlet资源的利用率和任务迁移的成功率往往较低,不仅浪费了时间还加剧了移动设备端的数据传输能耗,不利于节能;

[0010] (2) 目前还没有将CPU频率调节应用到cloudlet环境下任务迁移中的研究,不能根据预判的可迁移任务量调节移动设备CPU频率,以致不能使得移动设备在保证任务完成时间的低频状态下运行,减少能耗的程度低。

发明内容

[0011] 针对现有技术之不足,本发明提供了一种基于微云的能耗优化方法,预先计算出了可迁移任务数量期望,提高了微云的资源利用率,并以此为依据设置移动设备的第一处理器的初始运行频率,使得移动设备在保证任务完成时间的低频状态下运行,较大程度地减少了移动设备的能耗;此外,用户移动过程中采用马尔可夫决策过程动态决策每一个任务的调度策略并调节移动设备的第一处理器的当前运行频率,不仅保证了任务的完成时间而且最小化了移动设备的能耗。

[0012] 根据一个优选实施方式,一种基于微云的能耗优化方法,所述方法包括:构建一个

微云系统,所述微云系统包括:至少两个微云和无线连接于微云的移动设备,所述微云能够为与之无线连接的移动设备提供云计算服务;获取所述微云系统中与所述移动设备相关的系统数据用于分析所述移动设备需要在本地执行的数据量并根据所述数据量设置所述移动设备的第一处理器的初始运行频率 $F_{Linitial}$;根据所述任务队列的可容忍延迟范围 T 使用马尔可夫决策过程动态决策所述任务队列中单个任务的调度策略以及设置与所述调度策略对应的所述移动设备的第一处理器的当前运行频率 F_L 以使所述移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低。

[0013] 根据一个优选实施方式,所述获取所述微云系统中与所述移动设备相关的系统数据用于分析出所述移动设备需要在本地执行的数据量并根据所述数据量设置所述移动设备的第一处理器的初始运行频率 $F_{Linitial}$ 包括:根据所述至少两个微云的分布性质和持有所述移动设备的用户的移动特性计算出所述移动设备在随所述用户移动过程中可向所述微云系统迁移的任务数量期望 E_c ;根据所述任务数量期望 E_c 获取所述移动设备需要在本地执行的数据量;根据所述数据量设置所述移动设备的第一处理器的初始运行频率 $F_{Linitial}$,其中,所述分布性质在二维空间上服从泊松分布,所述用户的移动特性服从随机移动模型。

[0014] 根据一个优选实施方式,所述根据所述任务队列的可容忍延迟范围 T 使用马尔可夫决策过程动态决策所述任务队列中单个任务的调度策略以及设置与所述调度策略对应的所述移动设备的第一处理器的当前运行频率 F_L 以使所述移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低包括:设置所述移动设备的状态空间和动作空间,其中,所述状态空间包括:执行的任務状态、可连接的微云的状态和所述第一处理器的运行频率的调节范围,所述动作空间包括第一动作和第二动作,所述第一动作表示任务在本地执行,所述第二动作表示将任务传输到所述微云执行;根据所述状态空间和所述动作空间,通过值迭代得出最优的所述调度策略,所述移动设备根据调度策略设置所述第一处理器的当前运行频率 F_L 。

[0015] 根据一个优选实施方式,所述获取所述微云系统中与所述移动设备相关的系统数据用于分析出所述移动设备需要在本地执行的数据量并根据所述数据量设置所述移动设备的第一处理器的初始运行频率 $F_{Linitial}$ 包括:根据所述用户的移动区域的微云实际分布情况和分布性质,得出所述移动区域的微云分布密度 λ_c 和覆盖范围半径 R ,其中,所述分布性质在二维空间上服从泊松分布;根据服从随机移动模型的所述移动特性,得出用户移动速度 V ;获得所述移动设备上待执行的任務队列,所述任务队列共有 M 个任务,所述任务队列的可容忍延迟范围为 T ,每个任务的传输数据分别为 $\{DT_1, DT_2, \dots, DT_M\}$,每个任务的计算数据分别为 $\{DC_1, DC_2, \dots, DC_M\}$;根据用户移动的速度和用户移动的起点与终点的距离和随机几何理论,计算出用户在移动过程中经过微云覆盖区域的覆盖时间 T_c ;根据所述微云传输带宽 B_c 和所述微云的第二处理器的运行频率 F_c 计算出 M 个任务在所述微云执行的平均时间 \bar{T} ,

根据所述覆盖时间 T_c 和所述平均时间 \bar{T} ,获取可迁移任务期望 $E_c = \frac{T_c}{\bar{T}}$,根据所述可迁移任务期望 E_c 计算出所述移动设备需要在本地执行的任務数量 $E_L = M - E_c$;根据所述移动设备需要在本地执行的任務数量 E_L 、所述 M 个任务和所述每个任务的计算数据计算需要所述第一处理器在本地计算的数据量 D_L ,根据所述数据量 D_L 、所述可容忍延迟范围 T 和第一处理器的特性参数,设置所述初始运行频率 $F_{Linitial}$ 。

[0016] 根据一个优选实施方式,所述覆盖时间 T_c 的计算公式为: $T_c = \frac{L(1-e^{-\pi R^2 \lambda_c})}{V}$,其中, L 是用户移动的起点与终点间的距离;所述平均时间 \bar{T} 的计算公式为:

$$\bar{T} = \frac{\frac{\sum_{i=1}^M DT_i}{E_c} + \frac{\sum_{i=1}^M DC_i}{F_c}}{M}。$$

[0017] 根据一个优选实施方式,所述根据所述任务队列的可容忍延迟范围 T 使用马尔可夫决策过程动态决策所述任务队列中单个任务的调度策略以及设置与所述调度策略对应的所述移动设备的第一处理器的当前运行频率 F_L 以使所述移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低包括:设置用户状态空间

$\theta \in \{S = (i, N, F_L) \in S | i \in G, N \in \{0, 1, \dots, \bar{N}\}, F_L \in \{f_1, f_2, \dots, f_l\}\}$ 和动作空间 $A = \{0, 1\}$,其中, i 表示当前执行的任务, N 表示当前可连接的微云个数, F_L 表示所述第一处理器的运行频率,所述动作空间 A 中的 0 表示任务在本地执行,所述动作空间 A 中的 1 表示将任务传输到所述微云执行;根据微云的泊松分布特性和微云的分布密度 λ_c ,计算出用户当前可连接的微云个数概率 $P^N(N, N')$,得出微云的状态转移矩阵 \bar{N} ,其中, N 为当前状态 S 可连接的微云个数, N' 为转移到下一个状态 S' 可能连接的微云个数;根据当前任务的 DT_i 、 DC_i 以及所采取的动作 A_i ,计算出当前任务成功转移到下一个任务的概率 $P^G(G, G' | A_i)$,获得任务的状态转移矩阵 \bar{G} ;根据采取的动作 A_i ,计算出所述第一处理器的频率转移概率 $P^F(F_L, F'_L | A_i)$,得出频率的状态转移矩阵 \bar{F} ,

[0018] 其中, F_L 为在当前状态 S 时,所述第一处理器的运行频率, F'_L 为转移到下一个状态时移动设备可能采取的所述第一处理器的运行频率;根据优化目标

$$V_t(S) = \min_{\phi(S, A_i)} H_t(S, A_i) \quad , \quad \text{通过值迭代得出调度策略 } \phi_t^*(S) = \arg \min_{\phi(S, A_i)} H_t(S, A_i) \quad ,$$

根据调度策略设置第一处理器的当前运行频率 F_L ,其中, $H_t(S, A_i)$ 是用户在状态 S 时采取动作 A_i 时产生的能耗, $\phi(S, A_i)$ 是用户在当前状态 S 下采取动作 A_i 的概率。

[0019] 根据一个优选实施方式,所述当前任务成功转移到下一个任务的概率的计算公式

$$\text{为: } P^G(G, G' | A_i) = \begin{cases} \left(\frac{DT_i}{B_c} + \frac{DC_i}{F_c} \right) \times \frac{3V}{R} \times \eta_a & \text{当 } A_i = 1 \text{ 时} \\ 1 & \text{当 } A_i = 0 \text{ 时} \end{cases} \quad , \quad \text{其中, } \frac{3V}{R} \text{ 是用户与微云}$$

连接的平均时间, η_a 是当前连接的微云接收用户的任务迁移请求的概率。

[0020] 根据一个优选实施方式,所述当前连接的微云接受用户的任务迁移请求的概率 η_a 是根据所述当前连接的微云的资源占用情况和/或任务排队情况得出的。

[0021] 根据一个优选实施方式,所述优化目标 $V_t(S) = \min_{\phi(S, A_i)} H_t(S, A_i)$ 中的 $H_t(S, A_i) = C(S, A_i) + \gamma \sum P(S, S' | A_i) V_{t+1}(S')$,其中, $C(S, A_i)$ 是在状态 S 时采取动作 A_i 时所述第一处

理器产生的直接能耗, γ 是在所述下一个状态 S' 时所述第一处理器产生的能耗的折扣因子。

[0022] 根据一个优选实施方式, 所述微云系统包括: 至少两个微云和无线连接于微云的移动设备, 所述微云能够为与之无线连接的移动设备提供云计算服务; 第一设备获取所述微云系统中与所述移动设备相关的系统数据用于分析所述移动设备需要在本地执行的数据量并根据所述数据量设置所述移动设备的第一处理器的初始运行频率 F_{Linitial} ; 所述第一设备根据所述任务队列的可容忍延迟范围 T 使用马尔可夫决策过程动态决策所述任务队列中单个任务的调度策略以及设置与所述调度策略对应的所述移动设备的第一处理器的当前运行频率 F_L 以使所述移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低。

[0023] 本发明提供了一种基于微云的能耗优化方法至少具有以下有益效果:

[0024] (1) 本发明在微云分布服从泊松分布以及用户移动服从随机移动模型的情况下, 计算出了可迁移任务数量期望, 提高了微云的资源利用率, 并以此为依据设置移动设备的第一处理器的初始运行频率, 使得移动设备在保证任务完成时间的低频状态下运行, 较大程度地减少了移动设备的能耗;

[0025] (2) 本发明在用户移动过程中采用马尔可夫决策过程动态决策每一个任务的调度策略, 保证了任务迁移成功率, 避免了不必要的数据传输能耗;

[0026] (3) 本发明在动态调度任务过程中, 根据每一个任务的调度结果动态调节移动设备的第一处理器的当前运行频率, 保证移动设备端在任务可容忍延迟范围内的最低频状态运行, 不仅保证了任务的完成时间而且最小化了本地的计算能耗。

附图说明

[0027] 图1是本发明的方法的一个优选方式的流程图;

[0028] 图2是本发明的步骤200的子步骤的一个优选方式的流程图;

[0029] 图3是本发明的步骤300的子步骤的一个优选方式的流程图;

[0030] 图4是本发明的步骤350的子步骤的一个优选方式的流程图;

[0031] 图5是初始运行频率设置模块的模块示意图;

[0032] 图6是马尔可夫调度模块的模块示意图;

[0033] 图7是当任务在本地执行时的一个示例性状态转移示意图; 和

[0034] 图8是当任务在微云执行时的一个示例性状态转移示意图。

[0035] 附图标记列表

[0036]	200: 初始运行频率设置模块	210: 微云分布获取单元
[0037]	220: 用户移动速度获取单元	230: 任务情况获取单元
[0038]	240: 覆盖时间获取单元	250: 可迁移任务期望获取单元
[0039]	260: 本地执行任务数量获取单元	270: 初始运行频率设置单元
[0040]	300: 马尔可夫调度模块	310: 状态和动作空间设置单元
[0041]	320: 第一状态转移矩阵获取单元	330: 第二状态转移矩阵获取单元
[0042]	340: 第三状态转移矩阵获取单元	350: 当前运行频率调节单元

具体实施方式

[0043] 下面结合附图进行详细说明。

[0044] 为了便于理解,在可能的情况下,使用相同附图标记来表示各附图中共同的相似元件。

[0045] 如在整篇本申请中所使用的那样,词语“可以”系容许含义(即,意味着有可能的)而不是强制性含义(即,意味着必须的)。类似地,词语“包括”意味着包括但不限于。

[0046] 短语“至少一个”、“一个或多个”以及“和/或”系开放式表达,它们涵盖操作中的关联与分离两者。例如,表述“A、B和C中的至少一个”、“A、B或C中的至少一个”、“A、B和C中的一个或多个”、“A、B或C”和“A、B和/或C”中的每个分别指单独A、单独B、单独C、A和B一起、A和C一起、B和C一起或A、B和C一起。

[0047] 术语“一种”或“一个”实体指的是该实体中的一个或多个。这样,术语“一”(或“一”)、“一个或多个”以及“至少一个”在本文中可以交换地使用。还应该注意,术语“包括”、“包含”和“具有”可以交换地使用。

[0048] 首先对本发明涉及的术语进行说明:

[0049] 微云,即Cloudlet,是位于互联网边缘的小型数据中心,可以被附近的设备使用。比如,为附近的设备提供计算和/或存储服务。Cloudlet相比于cloud在地理位置上离用户更近,能以更低的延迟为其附近的设备服务。Cloudlet还可能存在其他等效表述,例如,朵云、follow me cloud和mobile micro-cloud。虽然表述不同,但是在其与本发明的Cloudlet实质相同的情况下,也应该纳入本发明的保护范围。

[0050] 分布式微云环境:微云在二维空间上以一定的密度服从泊松分布,用户在行走过程中可以根据距离等因素选择离自己最近的微云连接并发起请求。

[0051] 状态空间:用户在移动以及任务迁移过程中所有状态组成的集合。

[0052] 动作空间:任务迁移采取的动作,有本地执行和微云执行之分。

[0053] 状态转移矩阵:用户在移动以及任务迁移过程中由当前状态转移到另一个状态的概率所组成的矩阵。

[0054] 马尔可夫过程:指决策者周期地或连续地观察具有马尔可夫性的随机动态系统,序贯地作出决策。即根据每个时刻观察到的状态,从可用的行动集中选用一个行动作出决策,系统下一个状态是随机的,并且其状态转移概率具有马尔可夫性。决策者根据新观察到的状态,再作新的决策,依此反复地进行。马尔可夫性是指一个随机过程未来发展的概率规律与观察之前的历史无关的性质。它又可简单叙述为状态转移概率的无后效性。状态转移概率具有马尔可夫性的随机过程即为马尔可夫过程。

[0055] 实施例1

[0056] 本实施例公开了一种基于微云的能耗优化方法,本发明的方法可以包括:构建一个微云系统,微云系统可以包括:至少两个微云。当移动设备进入一微云的覆盖范围内并无线连接该微云时,微云系统能够通过该微云为移动设备提供云计算服务。优选地,至少两个微云之间彼此是通信连接的。优选地,移动设备连接微云后,成为微云系统的一部分。优选地,所述微云系统还可以包括:管理装置。管理装置可以采集微云系统中各组成设备的信息,经过分析处理后形成系统数据。比如,管理装置可以获取至少两个微云的微云配置信息和微云状态信息。其中,微云配置信息可以包括:微云的连接带宽、微云的第二处理器的运

行频率和微云的有效通信半径中的至少一个。微云状态信息可以包括微云的资源占用情况和\或任务排队情况。又比如,管理装置还可以获取移动设备的移动设备配置信息和移动设备状态信息。移动设备配置信息可以包括第一处理器的特性参数。移动设备状态信息可以包括运动状态、任务队列的情况和任务队列的可容忍延迟范围 T 。尤其优选地,管理装置是至少两个微云中的至少一个。

[0057] 根据一个优选实施方式,本发明的方法还可以包括:获取微云系统的系统数据。系统数据可以用于分析相应的移动设备需要在本地执行的数据量。根据数据量可以设置移动设备的第一处理器的初始运行频率 $F_{Linitial}$ 。系统数据可以包括微云位置信息、微云分布密度、覆盖范围半径、用户移动速度、任务队列的任务个数和/或任务的数据大小。先设置第一处理器的初始运行频率,以改变系统对第一处理器的频率的默认设置。因为根据默认设置,任务队列中所有任务都将在本地执行,为了按时完成任务,这个频率是偏大的。而如果不设置初始运行频率,仅根据每个任务的调度来调节第一处理器的频率,这样可以调节的范围很小,因为每次只能在一个任务的范围内调节。所以设置初始运行频率可以根据任务队列的总体情况做一个对整体的预判,让第一处理的能耗节省达到最大。

[0058] 根据一个优选实施方式,本发明的方法还可以包括:根据任务队列的可容忍延迟范围 T ,使用马尔可夫决策过程动态决策任务队列中单个任务的调度策略。设置与调度策略对应的移动设备的第一处理器的当前运行频率 F_L ,以使移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低。可容忍延迟范围 T 是完成任务队列的截止时间。在设置第一处理器的设置初始运行频率之后,后续根据马尔可夫决策过程的动态决策可以根据实际情况不断地对当前运行频率进行迭代更新,使频率调节逐步达到最优。

[0059] 根据一个优选的实施方式,参见图1,本发明的方法可以包括:

[0060] 步骤S100:构建一个微云系统,微云系统包括:至少两个微云和移动设备。微云能够为与之无线连接的移动设备提供云计算服务;

[0061] 步骤S200:移动设备(或管理设备)获取微云系统中与移动设备相关的系统数据,用于分析移动设备需要在本地执行的数据量,并根据数据量设置移动设备的第一处理器的初始运行频率 $F_{Linitial}$;

[0062] 步骤S300:根据任务队列的可容忍延迟范围 T ,移动设备(或管理设备)使用马尔可夫决策过程动态决策任务队列中单个任务的调度策略以及设置与调度策略对应的移动设备的第一处理器的当前运行频率 F_L ,使得移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低。

[0063] 根据另一个可替代的优选实施方式,再次参见图1,本发明的方法可以包括:

[0064] 步骤S100:构建一个微云系统,微云系统可以包括:至少两个微云、管理设备和移动设备,其中,微云能够为与之无线连接的移动设备提供云计算服务;

[0065] 步骤S200:管理设备获取微云系统中与移动设备相关的系统数据,用于分析移动设备需要在本地执行的数据量,并根据数据量设置移动设备的第一处理器的初始运行频率 $F_{Linitial}$;

[0066] 步骤S300:根据任务队列的可容忍延迟范围 T ,管理设备使用马尔可夫决策过程动态决策任务队列中单个任务的调度策略以及设置与调度策略对应的移动设备的第一处理器的当前运行频率 F_L ,使得移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最

低。

[0067] 根据另一个可替代的优选实施方式,再次参见图1,本发明的方法可以包括:

[0068] 步骤S100:构建一个微云系统,微云系统可以包括:至少两个微云、管理设备和移动设备,其中,微云能够为与之无线连接的移动设备提供云计算服务,移动设备内集成有专用传感器;

[0069] 步骤S200:专用传感器获取微云系统中与移动设备相关的系统数据,用于分析移动设备需要在本地执行的数据量,并根据数据量设置移动设备的第一处理器的初始运行频率 $F_{Linitial}$;

[0070] 步骤S300:根据任务队列的可容忍延迟范围 T ,专用传感器使用马尔可夫决策过程动态决策任务队列中单个任务的调度策略以及设置与调度策略对应的移动设备的第一处理器的当前运行频率 F_L ,使得移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低。

[0071] 根据一个优选实施方式,步骤S200可以包括:

[0072] 根据至少两个微云的分布性质和持有移动设备的用户的移动特性计算出移动设备在随用户移动过程中可向微云系统迁移的任务数量期望 E_C ;

[0073] 根据任务数量期望 E_C 获取移动设备需要在本地执行的数据量;

[0074] 根据数据量设置移动设备的第一处理器的初始运行频率 $F_{Linitial}$;

[0075] 其中,分布性质在二维空间上服从泊松分布,用户的移动特性服从随机移动模型。在微云分布服从泊松分布以及用户移动服从随机移动模型的情况下计算出可迁移任务数量期望,简化了计算过程,降低了对计算资源的占用率和系统能耗,并提升了整个系统的流畅性。

[0076] 优选地,参见图2,步骤S200可以包括以下子步骤:

[0077] 步骤S210:根据用户的移动区域的微云分布情况,得出移动区域的微云分布密度 λ_c 和覆盖范围半径 R ,其中,微云的分布性质在二维空间上服从泊松分布;

[0078] 步骤S220:根据服从随机移动模型的移动特性,得出用户移动速度 V ;

[0079] 步骤S230:获得移动设备上待执行的任务队列,任务队列共有 M 个任务,任务队列的可容忍延迟范围为 T ,每个任务的传输数据分别为 $\{DT_1, DT_2, \dots, DT_M\}$,每个任务的计算数据分别为 $\{DC_1, DC_2, \dots, DC_M\}$;

[0080] 步骤S240:根据用户移动的速度和用户移动的起点与终点的距离和随机几何理论,计算出用户在移动过程中经过微云覆盖区域的覆盖时间 T_c ;

[0081] 步骤S250:根据微云传输带宽 B_c 和微云的第二处理器的运行频率 F_c 计算出 M 个任务在微云执行的平均时间 \bar{T} ,根据覆盖时间 T_c 和平均时间 \bar{T} ,获取可迁移任务期望 $E_C = \frac{T_c}{\bar{T}}$;

[0082] 步骤S260:根据可迁移任务期望 E_C 计算出移动设备需要在本地执行的任务数量 $E_L = M - E_C$;

[0083] 步骤S270:根据移动设备需要在本地执行的任务数量 E_L 、 M 个任务和每个任务的计算数据计算需要第一处理器在本地计算的数据量 D_L ,根据数据量 D_L 、可容忍延迟范围 T 和第一处理器的特性参数,设置初始运行频率 $F_{Linitial}$ 。在随机模型下,根据随机理论计算出了用户可通过微云计算的任务数量期望,以此设置第一处理器的初始运行频率,减少了移动设

备能耗。在马尔可夫过程中,动态的根据任务调度情况调节第一处理器的当前运行频率,以确保任务的完成时间。相当于先做一个总体的预估,后续再动态的调整,逐步达到最优的状态。在使用马尔可夫决策过程的动态决策中考虑到了任务队列的可容忍延迟范围T,也就是任务队列的deadline,保证任务在deadline之前完成的情况下达到能耗最低。

[0084] 优选地,覆盖时间 T_C 的计算公式为: $T_C = \frac{L(1-e^{-\pi R^2 \lambda_c})}{v}$,其中,L是用户移动的起点与终点间的距离。

[0085] 优选地,平均时间 \bar{T} 的计算公式为: $\bar{T} = \frac{\sum_{i=1}^M DT_i + \sum_{i=1}^M DC_i}{E_C + F_C}$ 。

[0086] 优选地,数据量 D_L 的计算公式为: $D_L = \frac{E_L}{M} \sum_{i=1}^M DC_i$ 。

[0087] 优选地,初始运行频率 $F_{Linitial}$ 的计算公式为: $F_{Linitial} = D_L/T$ 。

[0088] 根据一个优选实施方式,使用马尔可夫决策过程动态决策总任务中的单个任务的调度策略以及设置与调度策略对应的移动设备的第一处理器的动态运行频率 F_L 可以包括:设置移动设备的状态空间和动作空间。状态空间可以包括:执行的任务状态、可连接的微云的状态和第一处理器的运行频率的调节范围。动作空间可以包括第一动作和第二动作,第一动作表示任务在本地执行,第二动作表示将任务传输到微云执行。移动设备可以使用马尔可夫决策过程并依据状态空间和动作空间,通过值迭代得出最优的调度策略。

[0089] 优选地,参见图3,步骤S300可以包括以下子步骤:

[0090] 步骤S310:设置用户状态空间

[0091] $\theta \in \{S = (i, N, F_L) \in S | i \in G, N \in \{0, 1, \dots, N\}, F_L \in \{f_1, f_2, \dots, f_J\}\}$ 和

[0092] 动作空间 $A = \{0, 1\}$,

[0093] 其中,i表示当前执行的任务,N表示当前可连接的微云个数, F_L 表示第一处理器的运行频率, f_J 的编号J为正整数,动作空间A中的0表示任务在本地执行,动作空间A中的1表示将任务传输到微云执行;

[0094] 步骤S320:根据微云的泊松分布特性和微云的分布密度 λ_c ,计算出用户当前可连接的微云个数概率 $P^N(N, N')$,得出微云的状态转移矩阵 \bar{N} ,

[0095] 其中,N为当前状态S可连接的微云个数, N' 为转移到下一个状态 S' 可能连接的微云个数;

[0096] 步骤S330:根据当前任务的 DT_i 、 DC_i 以及所采取的动作 A_i ,计算出当前任务成功转移到下一个任务的概率 $p^G(G, G' | A_i)$,获得任务的状态转移矩阵 \bar{G} ,

[0097] 步骤S340:根据采取的动作 A_i ,计算出第一处理器的频率转移概率 $p^F(F_L, F'_L | A_i)$,得出频率的状态转移矩阵 \bar{F} ,

[0098] 其中, F_L 为在当前状态S时,第一处理器的运行频率, F'_L 为转移到下一个状态时移动设备可能采取的第一处理器的运行频率;

[0099] 步骤S350:根据优化目标 $V_t(S) = \min_{\phi(S, A_i)} H_t(S, A_i)$,通过值迭代得出调度策略

$\phi_t^*(S) = \arg \min_{\phi(S, A_i)} H_t(S, A_i)$, 根据调度策略设置第一处理器的当前运行频率 F_L ;

[0100] 其中, $\phi(S, A_i)$ 是用户在当前状态 S 下采取动作 A_i 的概率。

[0101] 优选地, 当前任务成功转移到下一个任务的概率的计算公式为:

$$P^G(G, G' | A_i) = \begin{cases} \left(\frac{DT_i}{B_c} + \frac{DC_i}{F_c} \right) \times \frac{3V}{R} \times \eta_a & \text{当 } A_i = 1 \text{ 时} \\ 1 & \text{当 } A_i = 0 \text{ 时} \end{cases} \quad \text{。其中, } \frac{3V}{R} \text{ 是用户}$$

与微云连接的平均时间, η_a 是当前连接的微云接收用户的任务迁移请求的概率。

[0103] 优选地, 当前连接的微云接受用户的任务迁移请求的概率 η_a 是根据当前连接的微云的资源占用情况和/或任务排队情况得出的。

[0104] 根据一个优选实施方式, 在前任务成功转移到下一个任务的概率低于第一阈值且当前任务的计算数据量超过第二阈值之时, 成功转移到下一个

[0105] 优选地, 优化目标 $V_t(S) = \min_{\phi(S, A_i)} H_t(S, A_i)$ 中的 $H_t(S, A_i) = C(S, A_i) + \gamma \sum P(S, S' | A_i) V_{t+1}(S')$ 。其中, $H_t(S, A_i)$ 是用户在状态 S 下采取动作 A_i 时产生的能耗。 $C(S, A_i)$ 是在状态 S 下采取动作 A_i 时第一处理器产生的直接能耗。 γ 是在下一个状态 S' 下第一处理器产生的能耗的折扣因子。

[0106] 参见图7, 图7是当任务在本地执行时的一个示例性状态转移示意图。对于一个特定的应用, 它可以被区分为很多子任务, 如图7中的 G_1, G_2, G_3, G_4 。对于 G_1 , 当 $A=1$ 时表示它将被迁移到云端执行, 那么将可能有两种结果: 第一种: 以概率 $P^n P^G(1, 2)$ 迁移成功, 状态由 G_1 转移到 G_2 。第二种: 以概率 $1-P^n$ 迁移失败, 状态自循环到 G_1 。对于 G_2 , 当 $A=1$ 时表示它将被迁移到云端执行, 那么将可能有三种结果, 第一种: 以概率 $P^n P^G(2, 3)$ 迁移成功, 状态由 G_2 转移到 G_3 。第二种: 以概率 $P^n P^G(2, 4)$ 迁移成功, 状态由 G_2 转移到 G_4 。第三种: 以概率 $1-P^n$ 迁移失败, 状态自循环到 G_2 。对于 G_3, G_4 , 在 $A=1$ 的情况下, 若迁移成功则过程结束, 若失败则自循环到当前任务。

[0107] 参见图8, 图8是当任务在微云执行时的一个示例性状态转移示意图。所用应用示例与图7一致。对于 G_1 , 当 $A=0$ 时表示它将在本地执行, 那么只会产生一种结果: 以概率 $P^G(1, 2)$ 由 G_1 转移到 G_2 , 状态不可能自循环。对于 G_2 , 当 $A=0$ 时表示它将在本地执行, 那么可能产生两种结果, 第一种: 以概率 $P^G(2, 3)$ 由 G_2 转移到 G_3 , 第二种: 以概率 $P^G(2, 4)$ 由 G_2 转移到 G_4 , 状态不可能自循环。

[0108] 根据一个优选的实施方式, 参见图4, 步骤S350可以包括以下子步骤:

[0109] 步骤S351: 判断当前任务在是否在云端执行, 若是, 则跳转至步骤S352, 若否, 则跳转至步骤S356;

[0110] 步骤S352: 移动设备选择距离最近的微云传输迁移的任务, 并转至步骤S353;

[0111] 步骤S353: 判断迁移的任务是否成功返回结果, 若是, 则跳转至步骤S354, 若否, 则跳转至步骤S357;

[0112] 步骤S354: 判断是否还有待处理的任務, 若是, 则跳转至步骤S355, 若否, 则步骤S350暂时结束, 等待后续任务;

[0113] 步骤S355:处理下一个任务,并转至步骤S351;

[0114] 步骤S356:移动设备将当前任务在本地执行,并转至步骤S354;

[0115] 步骤S357:增大移动设备的第一处理器的当前运行频率 F_L ,并将当前任务差值任务队列的尾部,并转至步骤S355。

[0116] 根据一个优选实施方式,本发明的方法还可以包括:在当前连接的微云不能将迁移的任务的计算结果按时反馈给移动终端之时,获取用户从当前连接的微云移动到下一个状态时可能连接的另一微云的信息并分析将该迁移的任务通过另一微云反馈给用户所需的最大第一时长,以及用户在本地重新计算迁移的任务所需的第二时长,当最大第一时长小于第二时长之时,移动设备被设置为在连接到另一微云时请求获取迁移的任务的计算结果。由于下一个状态可能连接的另一微云的数量可能是多个,多个另一微云的位置和传输速度的不同导致相应的第一时长的不同。因此,在其中的最大第一时长仍小于第二时长之时,用户有较高概率能够以比在本地重新计算迁移的任务更快的方式获取计算结果。特别是在迁移的任务的计算量大的情况下,能够防止计算资源的浪费和提高任务按时完成的几率。根据本发明容易想到的是,该实施方式可以是由微云、移动设备和管理装置中的至少一个执行。优选地,该过程由管理装置执行,以降低微云和/或移动设备的计算量。优选地,在采取本实施方式之前,还包括判断迁移的任务的计算量是否超过第三阈值的步骤,当迁移的任务的计算量超过第三阈值时,采取本实施方式;当迁移的任务的计算量未超过第三阈值时,移动设备在本地重新计算该迁移的任务。因为任务的计算量较小时,采用本实施方式可能反而会增大移动设备的计算量,比如本实施方式由移动设备执行的情况下,在本地重新计算迁移的任务可能需要0.01ms,但是执行本实施方式需要0.03ms,反而增大了移动设备的计算量。因此,先预判任务的计算量大小,当超过第三阈值时才执行本实施方式有助于提高系统的计算效率。尤其优选地,移动设备被设置为在连接到另一微云时请求获取迁移的任务的计算结果的条件是:最大第一时长小于第二时长且最大第一时长与第二时长之比小于第四阈值。考虑到连接另一微云时可能出现的变化情况导致实际的第一时长大于理论的第一时长,把移动设备被设置为在连接到另一微云时请求获取迁移的任务的计算结果的条件设置为当最大第一时长小于第二时长且最大第一时长与第二时长之比小于第四阈值时,可以进一步提高按时完成任务的几率。变化情况是用户的移动情况变化、设备间通信连接时间变化或者通信速度变化。比如,变化情况是用户速度变化、用户移动方向变化或者移动设备与另一微云建立通信时间延长。

[0117] 实施例2

[0118] 本实施例是对实施例1的进一步改进,重复的内容不再赘述。

[0119] 本发明还提供一种基于微云的能耗优化装置,所述装置包括:

[0120] 初始运行频率设置模块200,用于获取所述微云系统中与所述移动设备相关的系统数据用于分析所述移动设备需要在本地执行的数据量并根据所述数据量设置所述移动设备的第一处理器的初始运行频率 $F_{Linitial}$;和/或

[0121] 马尔可夫调度模块300,用于根据所述任务队列的可容忍延迟范围 T 使用马尔可夫决策过程动态决策所述任务队列中单个任务的调度策略以及设置与所述调度策略对应的所述移动设备的第一处理器的当前运行频率 F_L 以使所述移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低,

[0122] 其中,所述微云系统包括:至少两个微云和无线连接于微云的移动设备,所述微云能够为与之无线连接的移动设备提供云计算服务。

[0123] 优选地,初始运行频率设置模块200可以用于执行以下步骤:

[0124] 根据所述至少两个微云的分布性质和持有所述移动设备的用户的移动特性计算出所述移动设备在随所述用户移动过程中可向所述微云系统迁移的任务数量期望 E_C ;

[0125] 根据所述任务数量期望 E_C 获取所述移动设备需要在本地执行的数据量;

[0126] 根据所述数据量设置所述移动设备的第一处理器的初始运行频率 F_{Linitial} ;

[0127] 其中,所述分布性质在二维空间上服从泊松分布,所述用户的移动特性服从随机移动模型。

[0128] 优选地,马尔可夫调度模块300可以用于执行以下步骤:

[0129] 设置所述移动设备的状态空间和动作空间,其中,所述状态空间包括:执行的任务状态、可连接的微云的状态和所述第一处理器的运行频率的调节范围,所述动作空间包括第一动作和第二动作,所述第一动作表示任务在本地执行,所述第二动作表示将任务传输到所述微云执行;

[0130] 所述移动设备根据所述状态空间和所述动作空间,通过值迭代得出最优的所述调度策略。

[0131] 进一步地,参见图5,初始运行频率设置模块200可以包括:

[0132] 微云分布获取单元210,用于根据所述用户的移动区域的微云实际分布情况和分布性质,得出所述移动区域的微云分布密度 λ_c 和覆盖范围半径 R ,其中,所述分布性质在二维空间上服从泊松分布;

[0133] 用户移动速度获取单元220,用于根据服从随机移动模型的所述移动特性,得出用户移动速度 V ;

[0134] 任务情况获取单元230,用于获得所述移动设备上待执行的任务队列,所述任务队列共有 M 个任务,所述任务队列的可容忍延迟范围为 T ,每个任务的传输数据分别为 $\{DT_1, DT_2, \dots, DT_M\}$,每个任务的计算数据分别为 $\{DC_1, DC_2, \dots, DC_M\}$;

[0135] 覆盖时间获取单元240,用于根据用户移动的速度和用户移动的起点与终点的距离和随机几何理论,计算出用户在移动过程中经过微云覆盖区域的覆盖时间 T_C ;

[0136] 可迁移任务期望获取单元250,根据所述微云传输带宽 B_C 和所述微云的第二处理器的运行频率 F_C 计算出 M 个任务在所述微云执行的平均时间 \bar{T} ,根据所述覆盖时间 T_C 和所述

平均时间 \bar{T} ,获取可迁移任务期望 $E_C = \frac{T_C}{\bar{T}}$;

[0137] 本地执行任务数量获取单元260,根据所述可迁移任务期望 E_C 计算出所述移动设备需要在本地执行的任务数量 $E_L = M - E_C$;

[0138] 初始运行频率设置单元270,用于根据所述移动设备需要在本地执行的任务数量 E_L 、所述 M 个任务和所述每个任务的计算数据计算需要所述第一处理器在本地计算的数据量 D_L ,根据所述数据量 D_L 、所述可容忍延迟范围 T 和第一处理器的特性参数,设置所述初始运行频率 F_{Linitial} 。应当注意的是,本发明的初始运行频率设置模块200可以单独地工作、封装和出售。

[0139] 进一步地,参见图6,马尔可夫调度模块300可以包括:

[0140] 状态和动作空间设置单元310,用于设置用户状态空间

[0141] $\theta \in \{S = (i, N, F_L) \in S | i \in G, N \in \{0, 1, \dots, \bar{N}\}, F_L \in \{f_1, f_2, \dots, f_j\}\}$ 和

[0142] 动作空间 $A = \{0, 1\}$,

[0143] 其中, i 表示当前执行的任务, N 表示当前可连接的微云个数, F_L 表示所述第一处理器的运行频率, 所述动作空间 A 中的0表示任务在本地执行, 所述动作空间 A 中的1表示将任务传输到所述微云执行;

[0144] 第一状态转移矩阵获取单元320,用于根据微云的泊松分布特性和微云的分布密度 λ_c , 计算出用户当前可连接的微云个数概率 $P^N(N, N')$, 得出微云的状态转移矩阵 \bar{N} , 其中, N 为当前状态 S 可连接的微云个数, N' 为转移到下一个状态 S' 可能连接的微云个数;

[0145] 第二状态转移矩阵获取单元330,用于根据当前任务的 DT_i 、 DC_i 以及所采取的动作 A_i , 计算出当前任务成功转移到下一个任务的概率 $P^G(G, G' | A_i)$, 获得任务的状态转移矩阵 \bar{G} ;

[0146] 第三状态转移矩阵获取单元340,用于根据采取的动作 A_i , 计算出所述第一处理器的频率转移概率 $P^F(F_L, F'_L | A_i)$, 得出频率的状态转移矩阵 \bar{F} , 其中, F_L 为在当前状态 S 时, 所述第一处理器的运行频率, F'_L 为转移到下一个状态时移动设备可能采取的所述第一处理器的运行频率;

[0147] 当前运行频率调节单元350,用于根据优化目标 $V_t(S) = \min_{\phi(S, A_i)} H_t(S, A_i)$, 通过值迭代得出调度策略 $\phi_t^*(S) = \arg \min_{\phi(S, A_i)} H_t(S, A_i)$, 根据调度策略设置第一处理器的当前运行频率 F_L , 其中, $\phi(S, A_i)$ 是用户在当前状态 S 下采取动作 A_i 的概率。

[0148] 优选地, 本实施例的装置还适于执行前述实施例中的其他方法步骤, 以达到预期的技术效果。

[0149] 优选地, 初始运行频率设置模块200和/或马尔可夫调度模块300可以是专用集成电路、CPU、微处理器、存储器和传感器中的至少一种。初始运行频率设置模块200和马尔可夫调度模块300可以独立设置, 或者集成于一体。

[0150] 优选地, 微云分布获取单元210、用户移动速度获取单元220、任务情况获取单元230、覆盖时间获取单元240、可迁移任务期望获取单元250、本地执行任务数量获取单元260和初始运行频率设置单元270可以是专用集成电路、CPU、微处理器、存储器和传感器中的至少一种。

[0151] 优选地, 马尔可夫调度模块300、状态和动作空间设置单元310、第一状态转移矩阵获取单元320、第二状态转移矩阵获取单元330、第三状态转移矩阵获取单元340和当前运行频率调节单元350可以是专用集成电路、CPU、微处理器、存储器和传感器中的至少一种。

[0152] 实施例3

[0153] 本实施例是对实施例1、2及其结合的进一步改进, 重复的内容不再赘述。

[0154] 本发明还公开了一种处理器, 所述处理器被配置为执行以下步骤:

[0155] 步骤S200: 获取微云系统中与移动设备相关的系统数据, 用于分析移动设备需要

在本地执行的数据量,并根据数据量设置移动设备的第一处理器的初始运行频率 $F_{Linitial}$;

[0156] 步骤S300:根据任务队列的可容忍延迟范围 T ,使用马尔可夫决策过程动态决策任务队列中单个任务的调度策略以及设置与调度策略对应的移动设备的第一处理器的当前运行频率 F_L ,使得移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低。

[0157] 根据一个优选实施方式,所述处理器是移动设备的第一处理器。

[0158] 根据另一个可替代的优选实施方式,所述处理器是微云的第二处理器。

[0159] 优选地,该处理器还适于执行前述实施例中的其他方法步骤,以达到预期的技术效果。

[0160] 实施例4

[0161] 本实施例是对实施例1、2、3及其结合的进一步改进,重复的内容不再赘述。

[0162] 本发明还公开了一种传感器,所述传感器被配置为执行以下步骤:

[0163] 步骤S200:获取微云系统中与移动设备相关的系统数据,用于分析移动设备需要在本地执行的数据量,并根据数据量设置移动设备的第一处理器的初始运行频率 $F_{Linitial}$;

[0164] 步骤S300:根据任务队列的可容忍延迟范围 T ,使用马尔可夫决策过程动态决策任务队列中单个任务的调度策略以及设置与调度策略对应的移动设备的第一处理器的当前运行频率 F_L ,使得移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低。

[0165] 根据一个优选实施方式,所述传感器集成于移动设备之内。所述传感器连接于移动设备的第一处理器。

[0166] 根据另一个可替代的优选实施方式,所述传感器集成于微云之内。所述传感器连接于微云的第二处理器。

[0167] 根据另一个可替代的优选实施方式,所述传感器集成于管理装置之内。所述传感器连接于管理装置的第三处理器。

[0168] 优选地,该传感器还适于执行前述实施例中的其他方法步骤,以达到预期的技术效果。

[0169] 实施例5

[0170] 一种微云系统,所述微云系统包括:至少两个微云和无线连接于微云的移动设备,所述微云能够为与之无线连接的移动设备提供云计算服务;

[0171] 第一设备获取所述微云系统中与所述移动设备相关的系统数据用于分析所述移动设备需要在本地执行的数据量并根据所述数据量设置所述移动设备的第一处理器的初始运行频率 $F_{Linitial}$;

[0172] 所述第一设备根据所述任务队列的可容忍延迟范围 T 使用马尔可夫决策过程动态决策所述任务队列中单个任务的调度策略以及设置与所述调度策略对应的所述移动设备的第一处理器的当前运行频率 F_L 以使所述移动设备在可容忍延迟范围 T 内完成任务的情况下实现能耗最低。

[0173] 优选地,第一设备是第一处理器或者连接于第一处理器的传感器。

[0174] 如本文中所使用的那样,术语“自动的”及其变型是指当执行过程或操作时在没有实质性人工输入的情况下完成的任何过程或操作。然而,如果在执行该过程或操作之前接收到该输入,则该过程或操作可以是自动的,即使该过程或操作的执行使用了实质性或非实质性的人工输入。如果这样的输入影响该过程或操作的执行方式,则该人工输入被认为

是实质性的。准予执行该过程或操作的人工输入不视为“实质性的”。

[0175] 虽然已经详细描述了本发明,但是在本发明的精神和范围内的修改对于本领域技术人员将是显而易见的。这样的修改也被认为是本公开的一部分。鉴于前面的讨论、本领域的相关知识以及上面结合背景讨论的参考或信息(均通过引用并入本文),进一步的描述被认为是不必要的。此外,应该理解,本发明的各个方面和各个实施例的各部分均可以整体或部分地组合或互换。而且,本领域的普通技术人员将会理解,前面的描述仅仅是作为示例,并不意图限制本发明。

[0176] 已经出于示例和描述的目的给出了本公开的前述讨论。这并不意图将本公开限制于本文公开的形式。在前述的具体实施方式中,例如,为了简化本公开的目的,本公开的各种特征在一个或多个实施例、配置或方面中被组合在一起。实施例、配置或方面的特征可以以除上面讨论的那些之外的替代实施例、配置或方面组合。本公开的该方法不应被解释为反映本公开需要比每个权利要求中明确记载的更多特征的意图。相反,如以下权利要求所反映的,创造性方面在于少于单个前述公开的实施例、配置或方面的所有特征。因此,以下权利要求由此被并入本具体实施方式中,其中每个权利要求其自身作为本公开的单独实施例。

[0177] 而且,虽然本公开的描述已经包括对一个或多个实施例、配置或方面以及某些变型和修改的描述,但是其他变型、组合和修改也在本公开的范围内,例如在本领域技术人员的技能和知识范围内,在理解了本公开之后。旨在获得在允许的程度上包括替代实施例、配置或方面的权利,所述权利包括那些要求保护的替代的、可互换的和/或等效的结构、功能、范围或步骤的权利,无论这种替代的、可互换的和/或等效的结构、功能、范围或步骤是否在本文中公开,并且无意公开奉献任何可专利的主题。

[0178] 需要注意的是,上述具体实施例是示例性的,本领域技术人员可以在本发明公开内容的启发下想出各种解决方案,而这些解决方案也都属于本发明的公开范围并落入本发明的保护范围之内。本领域技术人员应该明白,本发明说明书及其附图均为说明性而非构成对权利要求的限制。本发明的保护范围由权利要求及其等同物限定。

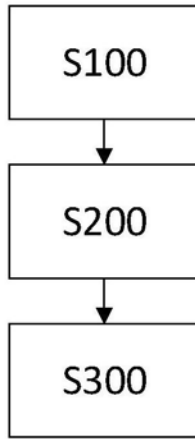


图1

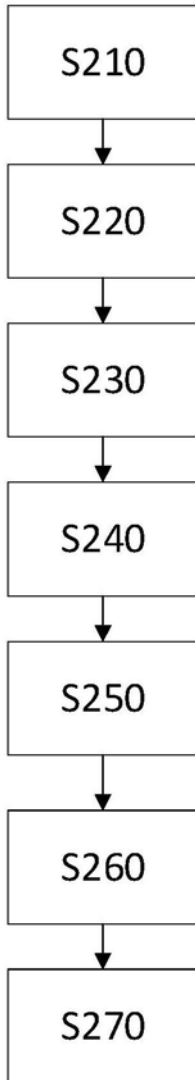


图2

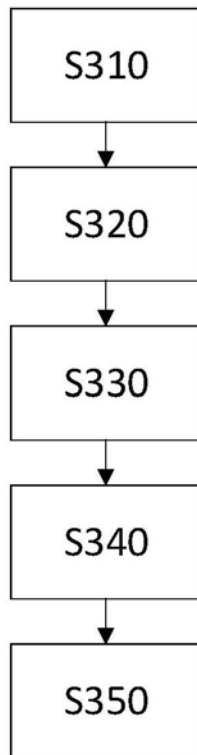


图3

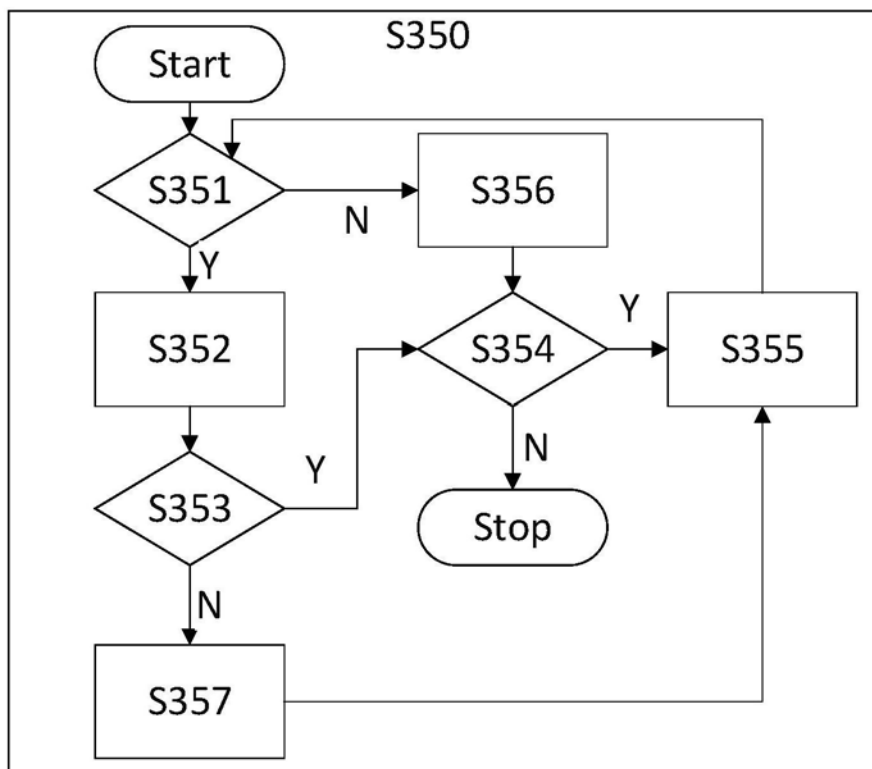


图4

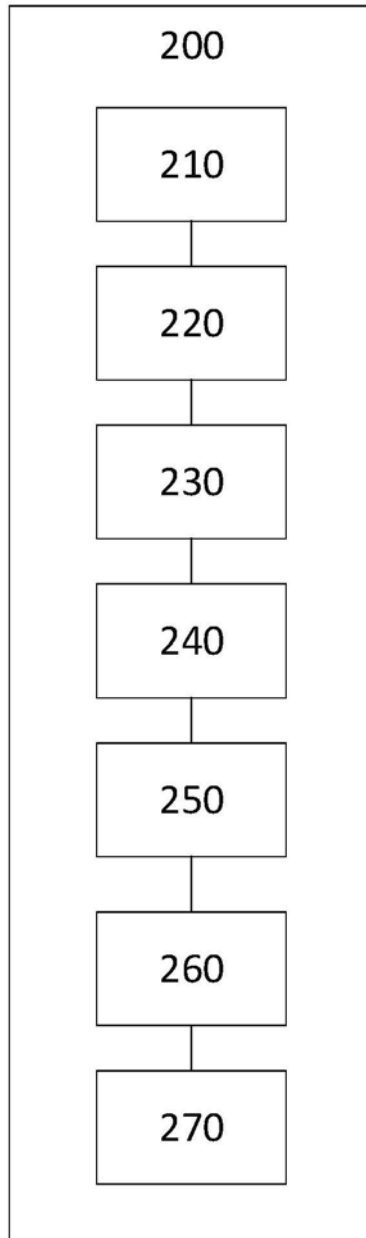


图5

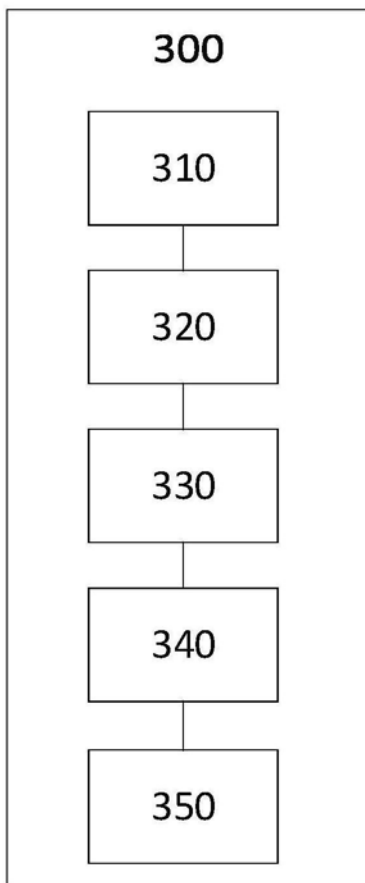


图6

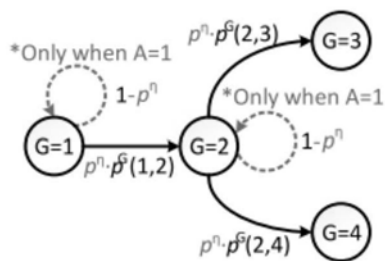


图7

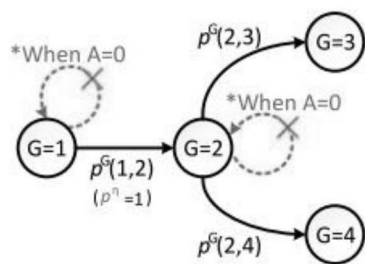


图8