8-29-2018

Task Allocation In a Computing Environment

Himadri Sekhar Paul

Arijit Mukherjee

Ansuman Banerjee

Swarnava Dey

Arpan Pal

See next page for additional authors

Creators
Himadri Sekhar Paul, Arijit Mukherjee, Ansuman Banerjee, Swarnava Dey, Arpan Pal, and Pubali Datta

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 2 977 898 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
29.08.2018 Bulletin 2018/35

(51) Int Cl.:
G06F 9/50 *(2006.01)*

(21) Application number: **15160431.1**

(22) Date of filing: **24.03.2015**

(54) **TASK ALLOCATION IN A COMPUTING ENVIRONMENT**

AUFTRAGSZUWEISUNG IN EINER COMPUTERUMGEBUNG

ATTRIBUTION DE TÂCHES DANS UN ENVIRONNEMENT INFORMATIQUE

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priority: **11.07.2014 IN 2281MU2014**

(43) Date of publication of application:
27.01.2016 Bulletin 2016/04

(73) Proprietor: **Tata Consultancy Services Limited**
Maharashtra (IN)

(72) Inventors:
- **Paul, Himadri Sekhar**
  700156 Rajarhat (IN)
- **Mukherjee, Arijit**
  700156 Rajarhat (IN)
- **Banerjee, Ansuman**
  700108 Kolkata (IN)
- **Dey, Swarnava**
  700156 Rajarhat (IN)
- **Pal, Arpan**
  700156 Rajarhat (IN)
- **Datta, Pubali**
  700156 Rajarhat (IN)

(74) Representative: **Zacco Sweden AB**
P.O. Box 5581
114 85 Stockholm (SE)

(56) References cited:
US-A- 6 098 091
US-A- 6 112 243
US-A1- 2007 220 520
US-A1- 2009 328 046
US-A1- 2012 284 383
US-B1- 6 393 433

- **MUSTAFA RAFIQUE M ET AL: "Symphony: A Scheduler for Client-Server Applications on Coprocessor-Based Heterogeneous Clusters", CLUSTER COMPUTING (CLUSTER), 2011 IEEE INTERNATIONAL CONFERENCE ON, IEEE, 26 September 2011 (2011-09-26), pages 353-362, XP032066083, DOI: 10.1109/CLUSTER.2011.46 ISBN: 978-1-4577-1355-2**

EP 2 977 898 B1

Description

TECHNICAL FIELD

[0001] The present subject matter relates, in general, to task allocation and, particularly but not exclusively, to task allocation in a computing environment.

BACKGROUND

[0002] Computing environments typically comprise a plurality of servers interconnected with each other. Such computing environments render support to large scale applications and facilitate high speed computation of such applications. In such computing environments, tasks or applications to be executed are allocated to the computing devices based on predetermined parameters. For example, the tasks may be allocated to the computing devices based on their corresponding processing power.

[0003] US 20120284383A1 discloses a computer-implemented method for workload management in a computer system by broadcasting a workload bid request to a plurality of compute nodes, wherein the workload bid request includes workload parameters characterizing the workload; the plurality of compute nodes each receiving the workload bid request from the first compute node, wherein each of the plurality of compute nodes uses the workload parameters included in the workload bid request to calculate a cost of running the workload; one or more individual compute nodes within the plurality of compute nodes each sending a workload bid to the first compute node, wherein each workload bid includes the cost of running the workload on the individual compute node sending the workload bid; and the first compute node receiving each workload bid and using each workload bid to select a target compute node to run the workload, wherein the target compute node is selected from the one or more individual compute nodes and the first compute node.

[0004] US 6112243 A discloses a method and apparatus for providing remote, distributed processing of a task by employing the Internet. A resource provider initiates the process by sending an application to a resource allocator requesting to be added to the resource allocator's list of providers. The resource allocator accepts or rejects a particular resource provider based on the application. If accepted, the resource provider waits for a task from the resource allocator. Upon receiving a task, the resource provider evaluates the currently available local resources. The resource provider determines whether or not it is currently able to handle the task in view of the available local resources. If the resource provider is able to handle the task, it accepts the task. The resource provider processes the task and returns the results to either the resource allocator or the original resource requester. If the resource provider is unable to handle the task in view of the current load, the resource provider notifies the resource allocator or the resource requester that it is unable to complete the task.

[0005] However, none of the above mentioned prior art documents teach or suggest a method for allocating a task in a computing environment with high accuracy in estimating the time taken for executing the task.

BRIEF DESCRIPTION OF DRAWINGS

[0006] The detailed description is described with reference to the accompanying figures. In the figures, the left-most digit(s) of a reference number identifies the figure in which the reference number first appears. The same numbers are used throughout the drawings to reference like features and components.

Figure 1 illustrates a network environment implementing a task allocation system for task allocation in a computing environment, in accordance with an implementation of the present subject matter.

Figure 2 illustrates a method for task allocation in a computing environment, in accordance with an implementation of the present subject matter.

DETAILED DESCRIPTION

[0007] The present subject matter relates to task allocation in a computing environment. The computing environment may be understood as an architecture comprising a plurality of computing devices, such as servers interconnected with each other. Examples of the computing environment may include, but are not limited to, a personal computing environment, a distributed computing environment, a grid computing environment, a utility computing environment, and a cloud computing environment.

[0008] Computing environments are typically deployed for processing large scale applications, for example, data mining applications involving computationally intensive computations. In computing environments, computations pertaining to such applications are typically performed by dividing the application into smaller tasks. The tasks are subsequently distributed for execution amongst the computing devices present in the computing environment. Typically, each task has an associated deadline, i.e., a time period by which the task is to be completely executed. Thus, while allocating the tasks in the computing environment, the corresponding deadlines are taken into account in order to ensure that the tasks are completed within their stipulated time.

[0009] In a known approach for allocating tasks in a computing environment, a worst-case execution time (WCET) for executing a given task is computed for each of the computing devices. The worst-case execution time may be understood as the maximum time taken by a computing device for completely executing a task by utilizing its processing resources to their full capacities. Thereafter, the task may be allocated to a computing

device based on the WCET corresponding to the computing device. However, in a case where the computing device is already performing other tasks, also referred to as active tasks, or is being used by a user for other applications, the processing resources to their fullest capacity may not be available for executing the task. As a result, the computing device may take more time to completely execute the application. Thus, subsequent actions based on the output of the complete execution of the task may be affected. Further, as the computing device has to simultaneously process the task in addition to the other tasks or applications, the computational time associated with the active tasks or applications may increase. As a result, the user utilizing the processing resources of the computing device may experience a degraded level of performance.

[0010] According to an embodiment of the present subject matter, systems and methods for task allocation in a computing environment are described. In accordance with the present subject matter, an estimated task execution time, i.e., time taken for completely executing a given task, is computed based on a WCET and a state transition model corresponding to a computing device. In an embodiment, the state transition model may indicate available processing resources of the computing device. The available processing resources, as used herein, indicate processing resources of the computing device which are not being utilized for any other processing by the computing device. Examples of the processing resources may include, but are not limited to, central processing unit (CPU) cycles, random access memory (RAM), cache, and Input Output (IO) speed, of the computing device. Thus, computation of the estimate task execution time based on the available processing resources may increase the accuracy in estimating the time taken for executing the task.

[0011] In an implementation, the computing device may receive a task execution estimation request message from a central sever. The task execution estimation request message may include the WCET corresponding to the computing device. Thereafter, the estimate task execution time may be computed based on the WCET and the state transition model. In an example, the state transition model includes a plurality of states corresponding to the computing device. In said example, each of the states may be annotated with information corresponding to the processing attributes. Further, the information may include a hold time corresponding to the state. The hold time may be understood as an average time period for which the computing device remains in the state.

[0012] Subsequently, the estimate task execution time may be transmitted to the central server. As may be understood, the computing environment may include a plurality of computing devices. Thus, the central server may receive the estimate task execution time from each of the computing devices. The central server may then accept all the estimate task execution times which are less than the deadline of the task. Thereafter, the central server may compare the estimate task execution times for identifying a computing device having lowest estimated task execution time and the task may be allocated to the identified computing device.

[0013] Thus, as the time taken for executing the task is computed based on the available processing resources of the computing device, time taken for complete execution of the task may be determined with high accuracy. As a result, subsequent tasks dependent on the given task may be executed with more accuracy. Furthermore, as the task is executed by utilizing the available processing resources, the user experience may not be affected.

[0014] Figure 1 illustrates a task allocation system 100 for allocating tasks in a computing environment. The task allocation system 100 includes a central server 102 and a plurality of computing devices 104-1, 104-2, ..., and 104-N, hereinafter collectively referred to as the computing devices 104 and individually referred to as the computing device 104. The central server 102 and the computing device 104 may communicate with each other, through a network 106, according to an embodiment of the present subject matter.

[0015] In an implementation, the central server 102 may be implemented as one or more systems, such as a cloud server, a mainframe computer, a workstation, a multiprocessor system, a network computer, and a gateway server. In an example, the central server 102 may be provisioned to allocate tasks amongst the computing devices 104. Examples of the computing device 104 may include, but are not limited to, a mobile phone, a smart phone, a personal digital assistant (PDA), a tablet, a laptop, a workstation computer, a server, and a personal computer.

[0016] The network 106 may be a wireless network, a wired network, or a combination thereof. The network 106 can also be an individual network or a collection of many such individual networks, interconnected with each other and functioning as a single large network, e.g., the Internet or an intranet. The network 106 can be implemented as one of the different types of networks, such as intranet, local area network (LAN), wide area network (WAN), the internet, and such. Further, the network 106 may include network devices that may interact with the central server 102 and the computing devices 104 through communication links.

[0017] In one implementation, the central server 102 and the computing device 104 include processors 108-1 and 108-2, respectively. The processors 108-1 and 108-2, hereinafter collectively referred to as the processor 108, may be implemented as one or more microprocessors, microcomputers, microcontrollers, digital signal processors, central processing units, state machines, logic circuitries, and/or any devices that manipulate signals based on operational instructions. Among other capabilities, the processor(s) is configured to fetch and execute computer-readable instructions stored in the memory.

[0018] The functions of the various elements shown in

the figure, including any functional blocks labeled as "processor(s)", may be provided through the use of dedicated hardware as well as hardware capable of executing software in association with appropriate software. When provided by a processor, the functions may be provided by a single dedicated processor, by a single shared processor, or by a plurality of individual processors, some of which may be shared.

[0019] Also, the central server 102 and the computing device 104 include I/O interface(s) 110-1 and 110-2, respectively. The I/O interfaces 110-1 and 110-2, collectively referred to as I/O interfaces 110 may include a variety of software and hardware interfaces that allow the central server 102 and the computing device 104 to interact with the network 106, or with each other. Further, the I/O interfaces 110 may enable the central server 102 and the computing device 104 to communicate with other communication and computing devices, such as web servers and external repositories. The central server 102 and the computing device 104 may further include memory 112-1 and 112-2, respectively, collectively referred to as memory 112. The memory 112-1 and 112-2 may be coupled to the processor 108-1 and the processor 108-2, respectively. The memory 112 may include any computer-readable medium known in the art including, for example, volatile memory (e.g., RAM), and/or non-volatile memory (e.g., EPROM, flash memory, etc.).

[0020] The central server 102 and the computing device 104 include modules 114-1, 114-2 and data 116-1, 116-2, respectively, collectively referred to as modules 114 and data 116, respectively. The modules 114 include routines, programs, objects, components, data structures, and the like, which perform particular tasks or implement particular abstract data types. The modules 114 further include modules that supplement applications on the central server 102 and the computing device 104, for example, modules of an operating system.

[0021] Further, the modules 114 can be implemented in hardware, instructions executed by a processing unit, or by a combination thereof. The processing unit can comprise a computer, a processor, such as the processor 108, a state machine, a logic array, or any other suitable devices capable of processing instructions. The processing unit can be a general-purpose processor which executes instructions to cause the general-purpose processor to perform the required tasks, or the processing unit can be dedicated to perform the required functions.

[0022] In another aspect of the present subject matter, the modules 114 may be machine-readable instructions (software) which, when executed by a processor/processing unit, perform any of the described functionalities. The machine-readable instructions may be stored on an electronic memory device, hard disk, optical disk, or other machine-readable storage medium or non-transitory medium. In one implementation, the machine-readable instructions can be also be downloaded to the storage medium via a network connection. The data 116 serves, amongst other things, as a repository for storing data that may be fetched, processed, received, or generated by one or more of the modules 114.

[0023] In an implementation, the modules 114-1 of the central server 102 include a task allocation module 118 and other module(s) 120. In said implementation, the data 116-1 of the central server 102 includes computing device data 122 and other data 124. The other module(s) 120 may include programs or coded instructions that supplement applications and functions, for example, programs in the operating system of the central server 102. The other data 124 comprise data corresponding to one or more other module(s) 120.

[0024] Similarly, in an implementation, the modules 114-2 of the computing device 104 include a communication module 126, a task execution estimation module, and other module(s) 130. In said implementation, the data 116-2 of the computing device 104 includes task execution data 132 and other data 134. The other module(s) 130 may include programs or coded instructions that supplement applications and functions, for example, programs in the operating system of the computing device 104. The other data 134 comprise data corresponding to one or more other module(s) 130.

[0025] As mentioned previously, the central server 102 may be provisioned to allocate tasks amongst the computing devices 104. In an implementation, for allocating a task to a computing device, such as the computing device 104-1, from amongst the computing devices 104, the task allocation module 118 may compute a worst-case execution time (WCET) for each of the computing devices 104. The WCET may be understood as time taken by a computing device for completely executing the task when all the processing resources, for example, random access memory (RAM), central processing unit (CPU) cycles, and cache, of the computing device are fully available for executing the task. In an example, the task allocation module 118 may compute the WCET for each of the computing devices 104 based on corresponding computing device information. The computing device information may be understood as information pertaining to architecture or processing capabilities of the computing device and may include, but are not limited to, processor and RAM of the computing device. Upon computing the WCET, the task allocation module 118 may then transmit the corresponding WCETs to the computing devices 104. For instance, the task allocation module 118 may transmit a task execution estimation request message comprising the WCET to each of the computing devices 104.

[0026] In an implementation, the communication module 126 may receive the task execution estimation request message from the task allocation module 118. Based on the WCET, the task execution estimation module 128 may compute an estimate task execution time for executing the task. The estimate task execution time, as used herein, may be understood as time taken by the computing device for executing the task using the available processing resources. For instance, for a CPU which

is seventy percent loaded with other tasks, only thirty percent of the CPU would be available for executing the task. Thus, the task execution estimation module 128, in such a case, may only take into consideration thirty percent CPU power for computing the estimate task execution time.

[0027] In an implementation, the task execution estimation module 128 may compute the estimate task execution time based on the WCET and a state transition model corresponding to the computing device 104. In an example, the state transition model comprises a plurality of states of the computing device 104. In said example, each state is annotated with processing information corresponding to one or more processing resources of the computing device 104, when the computing device 104 is in that state. Examples of the processing resources may include, but are not limited to, RAM, CPU cycles, and cache. The processing information may indicate a percentage of availability of the processing resource. For instance, processing information annotated with a state S1 of the computing device 104 may correspond to available CPU cycles of the computing device 104. Further, each of the states may be annotated with a hold time which indicates an average time period for which the computing device 104 remains in that state. For developing the state transition model, in an example, usage of the processing resources of the computing device 104 may be logged in the form of a time series in a processing log. Thereafter, the state transition model may be developed based on the processing log using known clustering algorithms. In another example, the state transition model may be developed based on the processing log using known machine learning techniques. The state transition model may further include a transition probability corresponding to each interlinked states of the state transition model. In an example, the transition probabilities may be computed from the processing log based on a usage pattern of the computing device 104.

[0028] In an implementation, for computing the estimate task execution time, the task execution estimation module 128 may identify a current state of the computing device 104. The current state may be understood as the state in which the computing device 104 receives the task execution estimation request message. Upon identifying the current state, the task execution estimation module 128 may ascertain one or more paths arising out of the current state based on the state transition model. Each path from among the paths may comprise one or more states of the computing device 104. Thereafter, the task execution estimation module 128 may compute a path task execution time and a path traverse value for each of the paths.

[0029] The path task execution time may be understood as time taken by the computing device 104 for executing the task on the path. In an example, the path task execution time may be based on the processing information and the hold time corresponding to the states present in the path. In said example, the task execution estimation module 128 may compute a task completion time for each of the states, starting with the current state, in the path. The task completion time may be understood as percentage of the WCET completed in the state. Upon computing the task completion time, the task execution estimation module 128 may then compare the task completion time for the current state with the WCET. In a case where the task execution estimation module 128 ascertains the task completion time to be less than the WCET, the task execution estimation module 128 may further compute another task completion time corresponding to the state subsequent to the current state. Thereafter, the task execution estimation module 128 may sum up the task completion time for both the states and may then compare the sum with the WCET. In a case where the sum is less than the WCET, the task execution estimation module 128 may then reiterate the updation of the sum by computing and adding the task completion time for the next state and may subsequently compare the updated sum with the WCET. Thus, as may be understood, the task execution estimation module 128 may repeat the process of updation and comparison till the sum is equal to or greater than the WCET. In the reiteration process, for a state in which the updated sum becomes equal to or exceeds the WCET, hereinafter also referred as the last state, the task execution estimation module 128 may sum up the hold time corresponding to each state, starting from the current state, till the last state to obtain the path task execution time for the path. Further, in said implementation, the task execution estimation module 128 may compute the path traverse value corresponding to the path. The path traverse value may be understood as a value obtained by multiplying the transition probabilities of all the states starting from the current state to the last state.

[0030] Upon computing the path task execution time and the path traverse value for the paths, the task execution estimation module 128 may compare the path traverse values for determining a path having a highest path traverse value. The path task execution time corresponding to the path may then be selected as the estimate task execution time. The task execution estimation module 128 may then store the estimate task execution value in the task execution data. Subsequently, the communication module 126 may transmit the estimate task execution time to the central server 102.

[0031] In an example, the task allocation module 118 may receive the estimated task execution times from each of the computing devices 104. Thereafter, the task allocation module 118 ascertains a lowest estimated task execution time from amongst the task execution times. Subsequently, the task allocation module 118 may compare the lowest estimate task execution time with a target task execution time corresponding to the task. The target task execution time may be understood as a time within which the task is to be completed. In a case where the lowest estimate task execution time is less than the target task execution time, the task allocation module 118 may

allocate the task to the corresponding computing device 104.

[0032] Figure 2 illustrates a method 200 for allocating a task in a computing environment, according to an embodiment of the present subject matter. The method 200 may be implemented in a variety of computing systems in several different ways. For example, the method 200, described herein, may be implemented using the central server 102, as described above.

[0033] The method 200, completely or partially, may be described in the general context of computer executable instructions. Generally, computer executable instructions can include routines, programs, objects, components, data structures, procedures, modules, functions, etc., that perform particular functions or implement particular abstract data types. A person skilled in the art will readily recognize that steps of the method can be performed by programmed computers. Herein, some embodiments are also intended to cover program storage devices, e.g., digital data storage media, which are machine or computer readable and encode machine-executable or computer-executable programs of instructions, wherein said instructions perform some or all of the steps of the described method 200.

[0034] The order in which the method 200 is described is not intended to be construed as a limitation, and any number of the described method blocks can be combined in any order to implement the method, or an alternative method. Additionally, individual blocks may be deleted from the method without departing from the scope of the subject matter described herein. Furthermore, the methods can be implemented in any suitable hardware, software, firmware, or combination thereof. It will be understood that even though the method 200 is described with reference to the application server 102, the description may be extended to other systems as well.

[0035] With reference to the description of Figure 2, for the sake of brevity, the details of the components of the central server 102 are not discussed here. Such details can be understood as provided in the description provided with reference to Figure 1.

[0036] Referring to the Figure 2, at block 202, at each of a plurality of computing devices, a task execution estimation request message is received from a central server, where the task execution estimation request message comprises a worst-case execution time (WCET) corresponding to the computing device. The WCET may be understood as time taken by the computing device for executing the application when all the processing resources, for example, CPU cycles, RAM, and cache, are fully available. In an implementation, the WCET may be computed based on computing device information corresponding to the computing device. The computing device information may be understood as information pertaining to processing capabilities of the computing device and may include, but are not limited to, processor and RAM of the computing device.

[0037] At block 204, by each of the computing devices, an estimate task execution time for the task is computed based on the WCET and a state transition model corresponding to the computing device. The estimate task execution time may be understood as time taken by the computing device for executing the task using available processing resources. In an example, the estimate task execution time may be computed based on the state transition model. The state transition model may indicate the available processing resources of the computing device. In an example, the state transition model may comprise a plurality of state, such that each state may indicate the available processing resources of the computing device, when the computing device is in that state. In said example, each state may be annotated with processing information corresponding to one or more processing resources. For example, a state S 1 may be annotated with processing information corresponding to free CPU cycles and free cache of the computing device. Further, the state may also be annotated with a corresponding hold time. The hold time may be understood as a time period for which the computing device remains in that state in average. In an implementation, for computing the estimate task execution time, a current state of the computing device may be identified. Thereafter, one or more paths from the current state may be ascertained. Once the paths from the current state are ascertained, for each path, a path task execution time and a path traverse value may be computed. Thereafter, a path having the highest path traverse value may be determined. The path task execution time corresponding to the path having the highest path traverse value may then be ascertained to be the estimate task execution time.

[0038] At block 206, the estimate task execution time is transmitted to the central sever by each of the computing devices for allocation of the task based on the estimate task execution time. Upon receiving the estimate task execution time from the computing devices, a computing device having lowest estimate task execution time may be identified. Thereafter, the lowest estimate task execution time may be compared with a target task execution time. In a case where the lowest estimate task execution time is less than the target task execution time, the task may be allocated to the corresponding computing device.

[0039] Although implementations for methods and systems for task allocation in a computing environment are described, it is to be understood that the present subject matter is not necessarily limited to the specific features or methods described. Rather, the specific features and methods are disclosed as implementations for task allocation in a computing environment.

Claims

1. A method for allocating a task in a computing environment, the method comprising:

receiving, at each of a plurality of computing devices (104), a task execution estimation request message from a central server (102), wherein the task execution estimation request message comprises a worst-case execution time (WCET) corresponding to a computing device (104), and wherein the WCET corresponds to time taken by the computing device (104) for completely executing a task when processing resources are fully available for executing the task and wherein the WCET for each of the plurality of computing devices (104) is computed by the central server (102) based on corresponding computing device information;

computing, by each of the plurality of computing devices (104), an estimate task execution time for the task based on the WCET and a state transition model corresponding to the computing device (104), wherein the state transition model indicates available processing resources corresponding to the computing device (104), wherein the computing comprises:

identifying a current state of the computing device (104), wherein the current state is a state in which the task execution estimation request message is received;

ascertaining one or more paths from the current state based on the state transition model, wherein the state transition model comprises a plurality of states of the computing device (104), and wherein each path comprises one or more states from amongst the plurality of states of the computing device (104);

computing, for each path, a path task execution time and a path traverse value, wherein the path task execution time indicates a time period taken by the computing device (104) for executing the task on the path, and wherein the path task execution time is based on processing information annotated with each of the one or more states corresponding to the path, and wherein the processing information corresponds to the available processing resources of the computing device (104) in the state, and wherein the path traverse value indicates a probability of the computing device (104) to traverse the path;

determining a path, from amongst the one or more paths, having a highest path traverse value;

selecting the path task execution time corresponding to the path as the estimate task execution time; and

transmitting, by each of the plurality of computing devices (104), the estimate task execution time to the central server (102);

allocating the task to the computing device (104) from amongst the plurality of computing devices (104) based on the estimate task execution time corresponding to the computing device (104); and

executing the task on the computing device (104) from amongst the plurality of computing devices (104).

2. The method as claimed in claim 1, wherein the processing information corresponds to at least one of a percentage of free computational cycles of a processor (108) of the computing device (104) and a hold time associated with each of the plurality of states, wherein the hold time indicates an average time period for which the computing device (104) is in a state.

3. A computing device (104) comprising:

a processor (108);

a task execution estimation module (128) coupled to the processor (108) to:

compute an estimate task execution time for a task based on a worst-case execution time (WCET) and a state transition model corresponding to the computing device (104), wherein the state transition model indicates available processing resources corresponding to the computing device (104);

a communication module (126) coupled to the processor (108) to:

transmit the estimate task execution time to a central server (102) for allocation of the task based on the estimate task execution time; and

wherein the communication module (126) further receives a task execution estimation request message from the central server (102), wherein the task execution estimation request message comprises the WCET corresponding to the computing device (104), and wherein the WCET corresponds to time taken by the computing device (104) for completely executing the task when processing resources are fully available for executing the task and wherein the WCET for each of the plurality of computing devices (104) is computed by the central server (102) based on corresponding computing device information;

wherein the task execution estimation module (128) is further configured to:

identify a current state of the computing

device (104) in which the task execution estimation request message is received;

ascertain one or more paths from the current state based on the state transition model, wherein the state transition model comprises a plurality of states of the computing device (104), and wherein each path comprises one or more states from amongst the plurality of states of the computing device (104);

compute, for each path, a path task execution time and a path traverse value, wherein the path task execution time indicates a time period taken by the computing device (104) for executing the task on the path, and wherein the path task execution time is based on processing information annotated with each of the one or more states, and wherein the processing information corresponds to the available processing resources of the computing device (104) in the state, and wherein the path traverse value indicates a probability of the computing device (104) to traverse the path;

determine a path, from amongst the one or more paths, having a highest path traverse value; and

select the path task execution time corresponding to the path as the estimate task execution time.

4. A task allocation system (100) for allocating tasks to a plurality of computing devices (104) comprising:

a central server (102) comprising:

a processor (108); and

a task allocation module (118) coupled to the processor (108) to:

compute a worst-case execution time (WCET) for executing a task for each of a plurality of computing devices (104) based on computing device information corresponding to each of the plurality of computing devices (104), wherein the WCET corresponds to time taken by computing device (104) for completely executing a task when processing resources are fully available for executing the task and wherein the WCET for each of the plurality of computing devices (104) is computed by the central server (102) based on corresponding computing device information;

transmit a task execution estimation request message to each of the plurality of computing devices (104), wherein the task execution estimation request message comprises corresponding WCET;

receive an estimate task execution time from each of the plurality of computing devices (104), wherein the estimated task execution time received from each of the plurality of computing devices (104) is computed by:

identify a current state of the computing device (104), wherein the current state is a state in which the task execution estimation request message is received;

ascertain one or more paths from the current state based on the state transition model, wherein the state transition model comprises a plurality of states of the computing device (104), and wherein each path comprises one or more states from amongst the plurality of states of the computing device (104);

compute, for each path, a path task execution time and a path traverse value, wherein the path task execution time indicates a time period taken by the computing device (104) for executing the task on the path, and wherein the path task execution time is based on processing information annotated with each of the one or more states corresponding to the path, and wherein the processing information corresponds to the available processing resources of the computing device (104) in the state, and wherein the path traverse value indicates a probability of the computing device (104) to traverse the path;

determine a path, from amongst the one or more paths, having a highest path traverse value; and

select the path task execution time corresponding to the path as the estimate task execution time;

transmit the executed task execution time to a central server;

allocate the task to a computing device from among the plurality of computing devices (104) based on the corresponding estimate task execution time of the computing device (104); and

execute the task on the computing device (104) from amongst the plurality of computing devices (104).

5. The task allocation system (100) as claimed in claim

4, wherein the task allocation module (118) is further configured to retrieve the computing device information corresponding to each of the plurality of computing devices (104) from a database.

6. The task allocation system (100) as claimed in claim 5, wherein the task allocation module (118) is further configured to:

select one or more computing devices from amongst the plurality of computing devices (104), wherein each of the computing device (104) from amongst the one or more computing devices has an estimate task execution less than a target task execution time corresponding to the task; and
ascertain the computing device (104) from amongst the one or more computing devices, wherein the computing device (104) has a lowest estimate task execution time for executing the task.

Patentansprüche

1. Verfahren für das Zuweisen eines Auftrags in einer Computerumgebung, welches Verfahren umfasst:

in jeder einer Vielzahl von Berechnungsvorrichtungen (104) Empfangen einer Auftragsausführungsschätzungs-Anforderungsnachricht von einem zentralen Server (102), wobei die Auftragsausführungsschätzungs-Anforderungsnachricht eine einer Berechnungsvorrichtung (104) entsprechende Worst-Case-Ausführungszeit (WCET) umfasst, und wobei die WCET einer Zeitdauer entspricht, die von der Berechnungsvorrichtung (104) zum vollständigen Ausführen eines Auftrags genommen wird, wenn Verarbeitungsressourcen für das Ausführen des Auftrags völlig verfügbar sind, und wobei die WCET für jede der Vielzahl von Berechnungsvorrichtungen (104) von dem zentralen Server (102) auf Grundlage entsprechender Berechnungsvorrichtungsinformationen berechnet wird;
Durch jede der Vielzahl von Berechnungsvorrichtungen (104) Berechnen einer Schätzungsauftragsausführungszeit für den Auftrag auf Grundlage der WCET und eines der Berechnungsvorrichtung (104) entsprechenden Zustandsübergangsmodells, wobei das Zustandsübergangsmodell der Berechnungsvorrichtung (104) entsprechende verfügbare Verarbeitungsressourcen angibt, wobei das Berechnen umfasst:

Identifizieren eines aktuellen Zustands der Berechnungsvorrichtung (104), wobei der aktuelle Zustand ein Zustand ist, in dem die Auftragsausführungs-Schätzungsanforderungsnachricht empfangen wird;
Ermitteln eines oder mehrerer Pfade aus dem aktuellen Zustand auf Grundlage des Zustandsübergangsmodells, wobei das Zustandsübergangmodell eine Vielzahl von Zuständen der Berechnungsvorrichtung (104) umfasst, und wobei jeder Pfad einen oder mehrere Zustände unter der Vielzahl von Zuständen der Berechnungsvorrichtung (104) umfasst;
Berechnen für jeden Pfad einer Pfad-Auftragsausführungszeit und eines Pfad-Traverse-Werts, wobei die Pfad-Auftragsausführungszeit eine Zeitdauer angibt, die von der Berechnungsvorrichtung (104) zum Ausführen des Auftrags im Pfad genommen wird, und wobei die Pfad-Auftragsausführungszeit auf Verarbeitungsinformationen, die miteinander annotiert sind, des einen oder mehrerer Zustände, dem Pfad entsprechend, beruht, und wobei die Verarbeitungsinformationen den verfügbaren Verarbeitungsressourcen der Berechnungsvorrichtung (104) in dem Zustand entsprechen, und wobei der Pfad-Traverse-Wert eine Wahrscheinlichkeit der Berechnungsvorrichtung (104) angibt, um den Pfad zu überqueren;
Bestimmen eines Pfads unter dem einen oder mehreren Pfaden mit einem höchsten Pfad-Traverse-Wert;
Auswählen der Pfad-Auftragsausführungszeit, die dem Pfad als der Schätzungsauftragsausführungszeit entspricht; und
Übertragen der Schätzungsauftragsausführungszeit an den zentralen Server (102) durch jede der Vielzahl von Berechnungsvorrichtungen (104);

Zuweisen des Auftrags zur Berechnungsvorrichtung (104) unter der Vielzahl von Berechnungsvorrichtungen (104) auf Grundlage der der Berechnungsvorrichtung (104) entsprechenden Schätzungsauftragsausführungszeit; und
Ausführen des Auftrags in der Berechnungsvorrichtung (104) unter der Vielzahl von Berechnungsvorrichtungen (104).

2. Verfahren nach Anspruch 1, wobei die Verarbeitungsinformationen zumindest einem eines Prozentsatzes von freien Berechnungszyklen eines Prozessors (108) der Berechnungsvorrichtung (104) und einer mit jedem der Vielzahl von Zuständen verknüpften Haltezeit entsprechen, wobei die Haltezeit

eine durchschnittliche Zeitdauer angibt, in der sich die Berechnungsvorrichtung (104) in einem Zustand befindet.

3. Berechnungsvorrichtung (104) umfassend:

einen Prozessor (108);
ein Auftragsausführungsschätzungsmodul (128), das mit dem Prozessor (108) gekoppelt ist, um:

eine Schätzungsauftragsausführungszeit für einen Auftrag auf Grundlage einer Worst-Case-Ausführungszeit (WCET) und eines Zustandsübergangsmodells entsprechend der Berechnungsvorrichtung (104) zu berechnen, wobei das Zustandsübergangsmodell verfügbare Verarbeitungsressourcen entsprechend der Berechnungsvorrichtung (104) angibt;
ein Kommunikationsmodul (126), das mit dem Prozessor (108) gekoppelt ist, um:

die Schätzungsauftragsausführungszeit an einen zentralen Server (102) für das Zuweisen des Auftrags auf Grundlage der Schätzungsauftragsausführungszeit zu übertragen; und
wobei das Kommunikationsmodul (126) ferner eine Auftragsausführungsschätzungs-Anforderungsnachricht vom zentralen Server (102) empfängt, wobei die Auftragsausführungsschätzungs-Anforderungsnachricht die der Berechnungsvorrichtung (104) entsprechende WCET umfasst, und wobei die WCET einer Zeitdauer entspricht, die von der Berechnungsvorrichtung (104) zum vollständigen Ausführen des Auftrags genommen wird, wenn Verarbeitungsressourcen für das Ausführen des Auftrags völlig verfügbar sind, und wobei die WCET für jede der Vielzahl von Berechnungsvorrichtungen (104) vom zentralen Server (102) auf Grundlage entsprechender Berechnungsvorrichtungsinformationen berechnet wird;

wobei das Auftragsausführungsschätzungsmodul (128) ferner dazu ausgelegt ist:

einen aktuellen Zustand der Berechnungsvorrichtung (104) zu identifizieren, in welchem die Auftragsausführungsschätzungs-Anforderungsnachricht empfangen wird;
einen oder mehrere Pfade vom aktuellen Zustand auf Grundlage des Zustandsübergangsmodells zu ermitteln, wobei das Zustandsübergangsmodell eine Vielzahl von Zuständen der Berechnungsvorrichtung (104) umfasst, und wobei jeder Pfad einen oder mehrere Zustände unter der Vielzahl von Zuständen der Berechnungsvorrichtung (104) umfasst;
für jeden Pfad eine Pfad-Auftragsausführungszeit und einen Pfad- Traverse-Wert zu berechnen, wobei die Pfad-Auftragsausführungszeit eine Zeitdauer angibt, die von der Berechnungsvorrichtung (104) zum Ausführen des Auftrags im Pfad genommen wird, und wobei die Pfad-Auftragsausführungszeit auf Verarbeitungsinformationen, die miteinander annotiert sind, des einen oder mehrerer Zustände beruht, und wobei die Verarbeitungsinformationen den verfügbaren Verarbeitungsressourcen der Berechnungsvorrichtung (104) in dem Zustand entsprechen, und wobei der Pfad-Traverse-Wert eine Wahrscheinlichkeit der Berechnungsvorrichtung (104) angibt, um den Pfad zu überqueren;
ein Pfad unter dem einen oder mehreren Pfaden mit einem höchsten Pfad-Traverse-Wert zu bestimmen; und
die Pfad-Auftragsausführungszeit, die dem Pfad als der Schätzungsauftragsausführungszeit entspricht, auszuwählen.

4. Auftragszuweisungssystem (100) für das Zuweisen von Aufträgen zu einer Vielzahl von Berechnungsvorrichtungen (104) umfassend:

einen zentralen Server (102) umfassend:

einen Prozessor (108); und
ein Auftragszuweisungsmodul (118), das mit dem Prozessor (108) gekoppelt ist, um:

eine Worst-Case-Ausführungszeit (WCET) zum Ausführen eines Auftrags für jede einer Vielzahl von Berechnungsvorrichtungen (104) auf Grundlage von jeder der Vielzahl von Berechnungsvorrichtungen (104) entsprechenden Berechnungsvorrichtungsinformationen zu berechnen, wobei die WCET einer Zeitdauer entspricht, die von Berechnungsvorrichtung (104) zum vollständigen Ausführen eines Auftrags genommen wird, wenn Verarbeitungsressourcen für das Ausführen des Auftrags völlig verfügbar sind, und wobei die WCET für jede der Vielzahl von Berechnungsvorrichtungen (104) auf Grundlage entsprechender

Berechnungsvorrichtungsinformationen vom zentralen Server (102) berechnet wird;
eine Auftragsausführungsschätzungs-Anforderungsnachricht an jede der Vielzahl von Berechnungsvorrichtungen (104) zu übertragen, wobei die Auftragsausführungsschätzungs-Anforderungsnachricht die entsprechende WCET umfasst;
eine Schätzungsauftragsausführungszeit von jeder der Vielzahl von Berechnungsvorrichtungen (104) zu empfangen, wobei die von jeder der Vielzahl von Berechnungsvorrichtungen (104) empfangene geschätzte Auftragsausführungszeit berechnet wird, indem:

ein aktueller Zustand der Berechnungsvorrichtung (104) identifiziert wird, wobei der aktuelle Zustand ein Zustand ist, in dem die Auftragsausführungs-Schätzungsanforderungsnachricht empfangen wird;
einen oder mehrere Pfade vom aktuellen Zustand auf Grundlage des Zustandsübergangsmodells zu ermitteln, wobei das Zustandsübergangsmodell eine Vielzahl von Zuständen der Berechnungsvorrichtung (104) umfasst, und wobei jeder Pfad einen oder mehrere Zustände unter der Vielzahl von Zuständen der Berechnungsvorrichtung (104) umfasst;
für jeden Pfad eine Pfad-Auftragsausführungszeit und ein Pfad-Traverse-Wert berechnet werden, wobei die Pfad-Auftragsausführungszeit eine Zeitdauer angibt, die von der Berechnungsvorrichtung (104) zum Ausführen des Auftrags im Pfad genommen wird, und wobei die Pfad-Auftragsausführungszeit auf Verarbeitungsinformationen, die miteinander annotiert sind, des einen oder mehrerer Zustände, dem Pfad entsprechend, beruht, und wobei die Verarbeitungsinformationen den verfügbaren Verarbeitungsressourcen der Berechnungsvorrichtung (104) in dem Zustand entsprechen, und wobei der Pfad-Traverse-Wert eine Wahrscheinlichkeit der Berechnungsvorrichtung (104) angibt, um den Pfad zu überqueren;
ein Pfad unter dem einen oder mehreren Pfaden mit einem höchsten Pfad-Traverse-Wert zu bestimmen; und
die Pfad-Auftragsausführungszeit, die dem Pfad als der Schätzungsauftragsausführungszeit entspricht, ausgewählt wird;

die ausgeführte Auftragsausführungszeit an einen zentralen Server übertragen wird;
der Auftrag einer Berechnungsvorrichtung unter der Vielzahl von Berechnungsvorrichtungen (104) auf Grundlage der entsprechenden Schätzungsauftragsausführungszeit der Berechnungsvorrichtung (104) zugewiesen wird; und
der Auftrag in der Berechnungsvorrichtung (104) unter der Vielzahl von Berechnungsvorrichtungen (104) ausgeführt wird.

5. Auftragszuweisungssystem (100) nach Anspruch 4, wobei das Auftragszuweisungsmodul (118) ferner für das Abrufen der jeder der Vielzahl von Berechnungsvorrichtungen (104) entsprechenden Berechnungsvorrichtungsinformationen von einer Datenbank ausgelegt ist.

6. Auftragszuweisungssystem (100) nach Anspruch 5, wobei das Auftragszuweisungsmodul (118) ferner dazu ausgelegt ist:

eine oder mehrere Berechnungsvorrichtungen unter der Vielzahl von Berechnungsvorrichtungen (104) auszuwählen, wobei jede der Berechnungsvorrichtung (104) unter der einen oder mehreren Berechnungsvorrichtungen eine Schätzungsauftragsausführung weniger als eine dem Auftrag entsprechende Auftragsausführungszeit aufweist; und
die Berechnungsvorrichtung (104) unter der einen oder mehreren Berechnungsvorrichtungen zu ermitteln, wobei die Berechnungsvorrichtung (104) eine niedrigste Schätzungsauftragsausführungszeit für das Ausführen des Auftrags aufweist.

Revendications

1. Procédé pour allouer une tâche dans un environnement informatique, le procédé comprenant :

la réception, à chacun d'une pluralité de dispositifs informatiques (104), d'un message de requête d'estimation d'exécution de tâche à partir d'un serveur central (102), le message de requête d'estimation d'exécution de tâche comprenant un temps d'exécution du pire cas

(WCET) correspondant au dispositif informatique (104), et le WCET correspondant au temps pris par le dispositif informatique (104) pour l'exécution complète d'une tâche lorsque des ressources de traitement sont entièrement disponibles pour l'exécution de la tâche, et le WCET pour chacun de la pluralité de dispositifs informatiques (104) étant calculé par le serveur central (102) sur la base d'informations de dispositif informatique correspondantes ;

le calcul, par chacun de la pluralité de dispositifs informatiques (104), d'un temps estimé d'exécution de tâche pour la tâche sur la base du WCET et un modèle de transition d'état correspondant au dispositif informatique (104), le modèle de transition d'état indiquant des ressources de traitement disponibles et correspondantes au dispositif informatique (104), le calcul comprenant :

l'identification d'un état actuel du dispositif informatique (104), l'état actuel étant un état dans lequel le message de requête d'estimation d'exécution de tâche est reçu;

la détermination d'un ou de plusieurs trajets à partir de l'état actuel sur la base du modèle de transition d'état, le modèle de transition d'état comprenant une pluralité d'états du dispositif informatique (104), et chaque trajet comprenant un ou plusieurs états parmi la pluralité d'états du dispositif informatique (104) ;

le calcul, pour chaque trajet, un temps d'exécution de tâche de trajet et une valeur de trajet parcouru, où le temps d'exécution de tâche de trajet indique une période de temps prise par le dispositif informatique (104) pour exécuter la tâche sur le trajet, et le temps d'exécution de tâche de trajet est basé sur des informations de traitement annotées avec chacun de l'un ou de plusieurs états correspondant au trajet, et les informations de traitement correspondant aux ressources de traitement disponibles du dispositif informatique (104) dans l'état, et la valeur de trajet parcouru indiquant une probabilité du dispositif informatique (104) pour parcourir le trajet ;

la détermination d'un trajet, parmi le ou plusieurs trajets, ayant une valeur de trajet parcouru la plus élevée ;

la sélection du temps d'exécution de tâche de trajet correspondant au trajet en tant que temps estimé d'exécution de tâche ; et

la transmission, par chacun de la pluralité de dispositifs informatiques (104), du temps estimé d'exécution de tâche au serveur central (102) ;

l'allocation de la tâche au dispositif informatique (104) parmi la pluralité de dispositifs informatiques (104) sur la base du temps estimé d'exécution de tâche correspondant au dispositif informatique (104) ; et

l'exécution de la tâche sur le dispositif informatique (104) parmi la pluralité de dispositifs informatiques (104).

2. Procédé selon la revendication 1, dans lequel les informations de traitement correspondent à au moins un pourcentage de cycles de calcul libres d'un processeur (108) du dispositif informatique (104) et un temps de maintien associé à chacun de la pluralité d'états, le temps de maintien indiquant une période de temps moyenne pendant laquelle le dispositif informatique (104) est dans un état.

3. Dispositif informatique (104) comprenant :

un processeur (108) ;

un module d'estimation d'exécution de tâche (128) couplé au processeur (108) pour :

calculer un temps estimé d'exécution de tâche pour une tâche sur la base d'un temps d'exécution du pire cas (WCET) et un modèle de transition d'état correspondant au dispositif informatique (104), le modèle de transition d'état indiquant des ressources de traitement disponibles et correspondantes au dispositif informatique (104) :

un module de communication (126) couplé au processeur (108) pour :

transmettre le temps estimé d'exécution de tâche à un serveur central (102) pour l'allocation de la tâche sur la base du temps estimé d'exécution de tâche ; et

dans lequel le module de communication (126) reçoit en outre un message de requête d'estimation d'exécution de tâche à partir du serveur central (102), le message de requête d'estimation d'exécution de tâche comprenant le WCET correspondant au dispositif informatique (104), et le WCET correspondant au temps pris par le dispositif informatique (104) pour l'exécution complète de la tâche lorsque des ressources de traitement sont entièrement disponibles pour l'exécution de la tâche, et le WCET pour chacun de la pluralité de dispositifs informatiques (104) étant calculé par le serveur central (102) sur la base d'informations de dispositif de calcul correspondantes ;

dans lequel le module d'estimation d'exécution de tâche (128) est en outre configuré à :

l'identification d'un état actuel du dispositif informatique (104), dans lequel le message de requête d'estimation d'exécution de tâche est reçu ;
la détermination d'un ou de plusieurs trajets à partir de l'état actuel sur la base du modèle de transition d'état, le modèle de transition d'état comprenant une pluralité d'états du dispositif informatique (104), et chaque trajet comprenant un ou plusieurs états parmi la pluralité d'états du dispositif informatique (104) ;
le calcul, pour chaque trajet, d'un temps d'exécution de tâche de trajet et une valeur de trajet parcouru, où le temps d'exécution de tâche de trajet indique une période de temps prise par le dispositif informatique (104) pour exécuter la tâche sur le trajet, et le temps d'exécution de tâche de trajet est basé sur des informations de traitement annotées avec chacun de l'un ou de plusieurs états, et les informations de traitement correspondant aux ressources de traitement disponibles du dispositif informatique (104) dans l'état, et la valeur de trajet parcouru indiquant une probabilité du dispositif informatique (104) pour parcourir le trajet ;
la détermination d'un trajet, parmi l'un ou plusieurs trajets, ayant une valeur de trajet parcouru la plus élevée ; et
la sélection du temps d'exécution de tâche de trajet correspondant au trajet en tant que temps estimé d'exécution de tâche.

4. Système d'allocation de tâches (100) pour allouer des tâches à une pluralité de dispositifs informatiques (104) comprenant :

un serveur central (102) comprenant :

un processeur (108) ; et
un module d'allocation de tâche (118) couplé au processeur (108) pour :

le calcul d'un temps d'exécution de pire cas (WCET) pour l'exécution d'une tâche pour chacun d'une pluralité de dispositifs informatiques (104) sur la base d'informations de dispositif informatique correspondant à chacun de la pluralité de dispositifs informatiques (104), le WCET correspondant au temps pris par le dispositif informatique (104) pour l'exécution complète d'une tâche lorsque des ressources de traitement sont entièrement disponibles pour l'exécution de la tâche, et le WCET pour chacun de la pluralité de dispositifs informatiques (104) étant calculé par le serveur central (102) sur la base d'informations de dispositif de calcul correspondantes ;
la transmission d'un message de requête d'estimation d'exécution de tâche à chacun de la pluralité de dispositifs informatiques (104), le message de requête d'estimation d'exécution de tâche comprenant un WCET correspondant ;
la réception d'un temps estimé d'exécution de tâche à partir de chacun de la pluralité de dispositifs informatiques (104), le temps estimé d'exécution de tâche reçu de chacun de la pluralité de dispositifs informatiques (104) étant calculé par:

l'identification d'un état actuel du dispositif informatique (104), l'état actuel étant un état dans lequel le message de requête d'estimation d'exécution de tâche est reçu;
la détermination d'un ou de plusieurs trajets à partir de l'état actuel sur la base du modèle de transition d'état, le modèle de transition d'état comprenant une pluralité d'états du dispositif informatique (104), et chaque trajet comprenant un ou plusieurs états parmi la pluralité d'états du dispositif informatique (104) ;
le calcul, pour chaque trajet, un temps d'exécution de tâche de trajet et une valeur de trajet parcouru, où le temps d'exécution de tâche de trajet indique une période de temps prise par le dispositif informatique (104) pour exécuter la tâche sur le trajet, et le temps d'exécution de tâche de trajet est basé sur des informations de traitement annotées avec chacun de l'un ou plusieurs états, et les informations de traitement correspondant aux ressources de traitement disponibles du dispositif informatique (104) dans l'état, et la valeur de trajet parcouru indiquant une probabilité du dispositif informatique (104) pour parcourir le trajet ;
la détermination d'un trajet, parmi

l'un ou plusieurs trajets, ayant une valeur de trajet parcouru la plus élevée ; et
la sélection du temps d'exécution de tâche de trajet correspondant au trajet en tant que temps estimé d'exécution de tâche ;

la transmission du temps d'exécution de tâche exécutée à un serveur central ;
l'allocation de la tâche à un dispositif informatique parmi la pluralité de dispositifs informatiques (104) sur la base du temps estimé d'exécution de tâche correspondant au dispositif informatique (104) ; et
l'exécution de la tâche sur le dispositif informatique (104) parmi la pluralité de dispositifs informatiques (104).

5. Système d'allocation de tâches (100) selon la revendication 4, dans lequel le module d'allocation de tâches (118) est en outre configuré à récupérer les informations de dispositif informatique correspondant à chacun de la pluralité de dispositifs informatiques (104) à partir d'une base de données.

6. Système d'allocation de tâches (100) selon la revendication 5, dans lequel le module d'allocation de tâches (118) est en outre configuré pour :

sélectionner un ou plusieurs dispositifs informatiques parmi la pluralité de dispositifs informatiques (104), chacun des dispositifs informatiques (104) parmi l'un ou plusieurs dispositifs informatiques ayant une exécution de tâche d'estimation inférieure à un temps d'exécution de tâche cible correspondant à la tâche; et
déterminer le dispositif informatique (104) parmi l'un ou plusieurs dispositifs informatiques, le dispositif informatique (104) ayant un temps estimé d'exécution de tâche le plus bas pour exécuter la tâche.

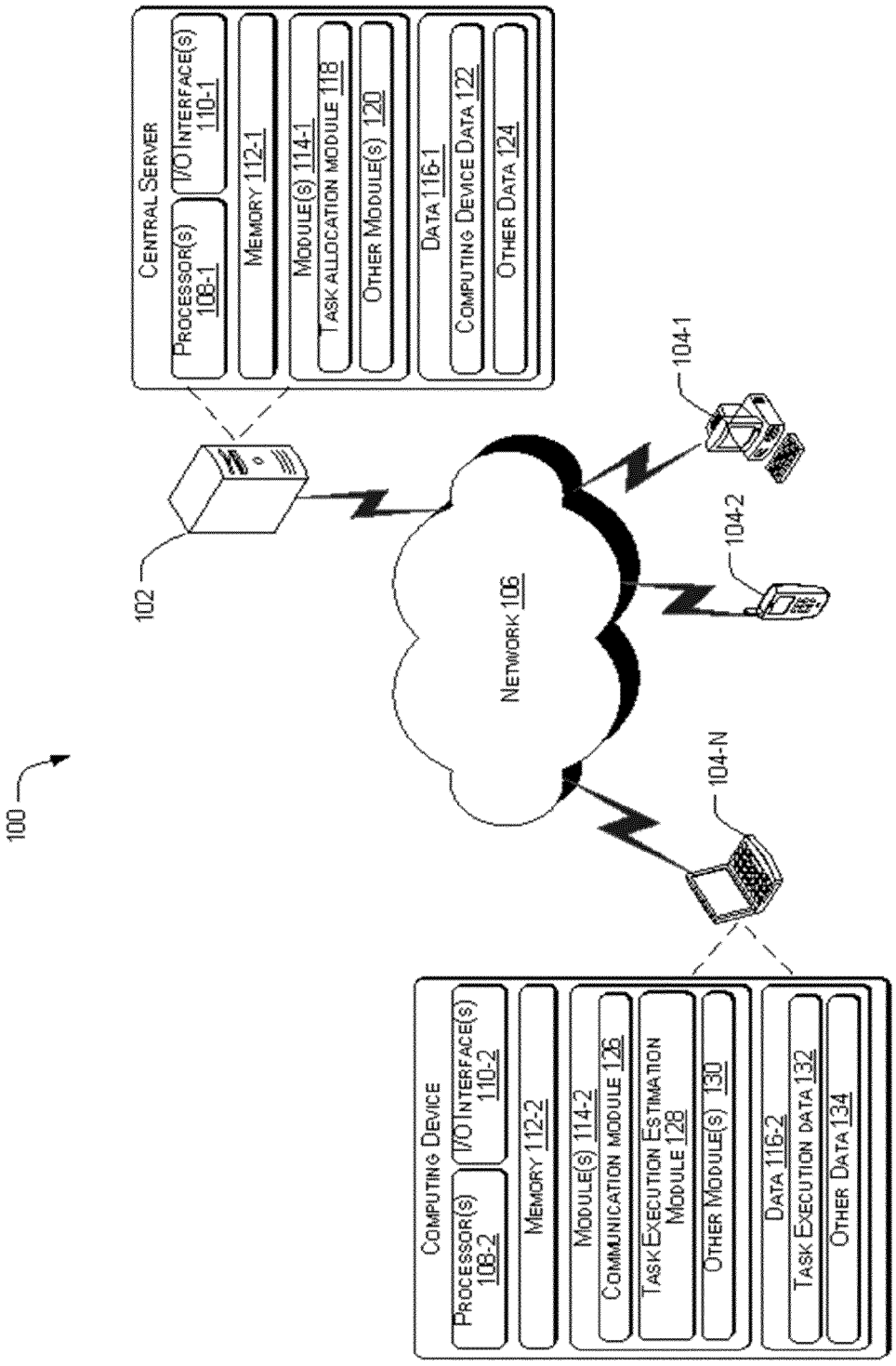
100
CENTRAL SERVER
PROCESSOR(S) 108-1
I/O INTERFACE(S) 110-1
MEMORY 112-1
MODULE(S) 114-1
TASK ALLOCATION MODULE 118
OTHER MODULE(S) 120
DATA 116-1
COMPUTING DEVICE DATA 122
OTHER DATA 124
102
NETWORK 106
104-1
104-2
104-N
COMPUTING DEVICE
PROCESSOR(S) 108-2
I/O INTERFACE(S) 110-2
MEMORY 112-2
MODULE(S) 114-2
COMMUNICATION MODULE 126
TASK EXECUTION ESTIMATION MODULE 128
OTHER MODULE(S) 130
DATA 116-2
TASK EXECUTION DATA 132
OTHER DATA 134

Figure 1

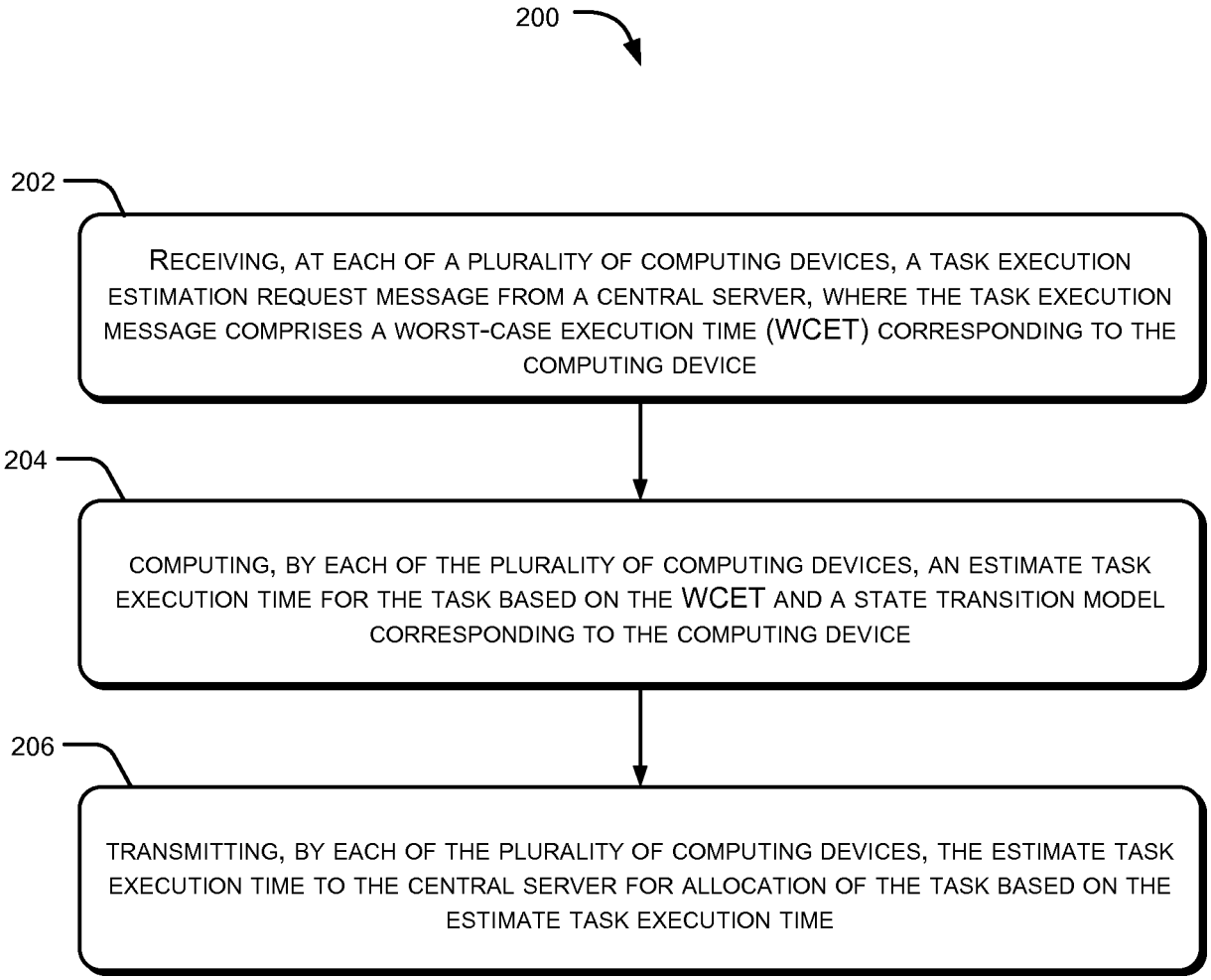
200
202
RECEIVING, AT EACH OF A PLURALITY OF COMPUTING DEVICES, A TASK EXECUTION ESTIMATION REQUEST MESSAGE FROM A CENTRAL SERVER, WHERE THE TASK EXECUTION MESSAGE COMPRISES A WORST-CASE EXECUTION TIME (WCET) CORRESPONDING TO THE COMPUTING DEVICE
204
COMPUTING, BY EACH OF THE PLURALITY OF COMPUTING DEVICES, AN ESTIMATE TASK EXECUTION TIME FOR THE TASK BASED ON THE WCET AND A STATE TRANSITION MODEL CORRESPONDING TO THE COMPUTING DEVICE
206
TRANSMITTING, BY EACH OF THE PLURALITY OF COMPUTING DEVICES, THE ESTIMATE TASK EXECUTION TIME TO THE CENTRAL SERVER FOR ALLOCATION OF THE TASK BASED ON THE ESTIMATE TASK EXECUTION TIME

Figure 2

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 20120284383 A1 **[0003]**
- US 6112243 A **[0004]**