(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 March 2012 (22.03.2012)

PCT

(10) International Publication Number
WO 2012/034273 A1

(51) International Patent Classification:
H04L 29/08 (2006.01)

(21) International Application Number:
PCT/CN2010/076936

(22) International Filing Date:
15 September 2010 (15.09.2010)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **EMPIRE TECHNOLOGY DEVELOPMENT LLC** [US/US]; 2711 Centerville Road, Suite 400, Wilmington, Delaware 19808 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **WANG, Zhanjie** [CN/CN]; No.190, Lingshui Road, Ganjingzi District, Dalian, Liaoning 116023 (CN).

(74) Agent: **UNITALEN ATTORNEYS AT LAW**; 7th Floor, Scitech Place, No.22, Jian Guo Men Wai Ave., Chao Yang District, Beijing 100004 (CN).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, 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) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (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, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

(54) **Title**: TASK ASSIGNMENT IN CLOUD COMPUTING ENVIRONMENT

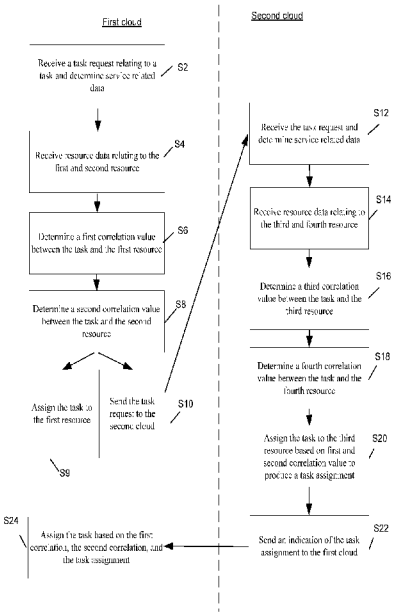


Fig. 2

(57) **Abstract**: Technologies are generally described for a system and method for assigning a task in a cloud. In some examples, the method may include receiving a task request relating to a task and determining service related data relating to the task based on the task request. In some examples, the method may include receiving resource data relating to a first and second resource in the cloud. In some examples, the method may include determining a first correlation value between the task and the first resource and a second correlation value between the task and the second resource based on the service related data and the resource data. In some examples, the method may include assigning the task to the first resource based on the first and second correlation value.

WO 2012/034273 A1

TASK ASSIGNMENT IN CLOUD COMPUTING ENVIRONMENT

BACKGROUND

[0001] Unless otherwise indicated herein, the materials described in this section are not prior art to the claims in this application and are not admitted to be prior art by inclusion in this section.

[0002] A network may include two or more computers in communication with each other through a communication link. In a cloud computing environment, resources in the form of processors, storage, and software may be shared among processors in communication with the cloud. A task request may be received by the cloud and one of the resources may be assigned to handle the task.

SUMMARY

[0003] In an example, a method for assigning a task in a cloud is described. In an example, the cloud includes at least a first resource and a second resource. In one example the method includes receiving, at a first processor, a task request relating to a task. In one example the method includes determining, at the first processor, service related data relating to the task based on the task request. In one example the method includes receiving, at a second processor, resource data relating to the first and second resource. In one example the method includes receiving, at a third processor, the service related data and the resource data. In one example the method includes determining, at the third processor, a first correlation value between the task and the first resource. In one example the first correlation value is based on the service related data and the resource data. In one example the method includes determining, at the third processor, a second correlation value between the task and the second resource. In one example the second correlation value is based on the service related data and the resource data. In one example the method includes assigning, by the third processor, the task to the first resource based on the first correlation value and the second correlation value.

[0004] In an example, a cloud effective to assign a task is described. In an example, the cloud includes a first resource, a second resource, a first processor, a

second processor in communication with the first resource and the second resource, and a third processor in communication with the first processor and the second processor. In an example, the first processor is effective to receive a task request relating to a task and determine service related data relating to the task based on the task request. In an example, the second processor is effective to receive resource data relating to the first and second resource. In an example, the third processor is effective to determine a first correlation value between the task and the first resource. In an example, the first correlation value is based on the service related data and the resource data. In an example, the third processor is effective to determine a second correlation value between the task and the second resource. In an example, the second correlation value is based on the service related data and the resource data. In an example, the third processor is effective to assign the task to the first resource based on the first and the second correlation value.

[0005] In an example, a system effective to assign a task is described. In an example, the system includes a first cloud including a first access point, a first resource and a second resource. In an example, the system includes a second cloud including a second access point, a third resource and a fourth resource. In an example, the first cloud is effective to receive a task request relating to a task. In an example, the first cloud is effective to determine service related data relating to the task based on the task request. In an example, the first cloud is effective to receive first resource data relating to the first and second resource. In an example, the first cloud is effective to determine a first correlation value between the task and the first resource. In an example, the first correlation value is based on the service related data and the first resource data. In an example, the first cloud is effective to determine a second correlation value between the task and the second resource. In an example, the second correlation value is based on the service related data and the first resource data. In an example, the first cloud is effective to send the task request to the second cloud based on the first and second correlation value. In an example, the second cloud is effective to receive the task request. In an example, the second cloud is effective to determine the service related data relating to the task based on the task request. In an example, the second cloud is effective to receive second resource data relating to the third and

fourth resource. In an example, the second cloud is effective to determine a third correlation value between the task and the third resource. In an example, the third correlation value is based on the service related data and the second resource data. In an example, the second cloud is effective to determine a fourth correlation value between the task and the fourth resource. In an example the fourth correlation value is based on the service related data and the second resource data. In an example, the second cloud is effective to assign the task to the third resource based on the first and the second correlation value to produce a task assignment.

[0006] The foregoing summary is illustrative only and is not intended to be in any way limiting. In addition to the illustrative aspects, embodiments, and features described above, further aspects, embodiments, and features will become apparent by reference to the drawings and the following detailed description.

BRIEF DESCRIPTION OF THE FIGURES

[0007] The foregoing and other features of this disclosure will become more fully apparent from the following description and appended claims, taken in conjunction with the accompanying drawings. Understanding that these drawings depict only several embodiments in accordance with the disclosure and are, therefore, not to be considered limiting of its scope, the disclosure will be described with additional specificity and detail through use of the accompanying drawings, in which:

Fig. 1 illustrates some example systems that can be utilized to implement task assignment in a cloud computing environment;

Fig. 2 depicts a flow diagram for example processes for implementing task assignment in a cloud computing environment;

Fig. 3 illustrates computer program products for implementing task assignment in a cloud computing environment; and

Fig. 4 is a block diagram illustrating an example computing device that is arranged to perform task assignment in a cloud computing environment in accordance with the present disclosure;

all arranged according to at least some embodiments described herein.

DETAILED DESCRIPTION

[0008] In the following detailed description, reference is made to the accompanying drawings, which form a part hereof. In the drawings, similar symbols typically identify similar components, unless context dictates otherwise. The illustrative embodiments described in the detailed description, drawings, and claims are not meant to be limiting. Other embodiments may be utilized, and other changes may be made, without departing from the spirit or scope of the subject matter presented herein. It will be readily understood that the aspects of the present disclosure, as generally described herein, and illustrated in the Figures, can be arranged, substituted, combined, separated, and designed in a wide variety of different configurations, all of which are explicitly contemplated herein.

[0009] This disclosure is generally drawn, *inter alia*, to methods, apparatus, systems, devices, and computer program products related to task assignment in a cloud computing environment.

[0010] Briefly stated, technologies are generally described for a system and method for assigning a task in a cloud. In some examples, the method may include receiving a task request relating to a task and determining service related data relating to the task based on the task request. In some examples, the method may include receiving resource data relating to a first and second resource in the cloud. In some examples, the method may include determining a first correlation value between the task and the first resource and a second correlation value between the task and the second resource based on the service related data and the resource data. In some examples, the method may include assigning the task to the first resource based on the first and second correlation value.

[0011] Fig. 1 illustrates some example systems that can be utilized to implement task assignment in a cloud computing environment. In some examples, as explained in more detail below, system 100 may include first cloud 120 and/or at least a second cloud 122 in communication with each other. In an example, each cloud 120, 122 may include one or more access points, such as task management module 118, through which processors outside the cloud can communicate with

resources in the cloud. For example, the access point may be configured to require a fee or subscription in order to allow communication with resources in the cloud.

[0012] In some examples, first cloud 120 may include two or more resources 106, 108, 110 and/or a scheduling system 114. Resources 106, 108, 110 may include, for example, processors, memories, routers, etc. In some examples, scheduling system 114 may be configured to receive a task request 102 from a requesting processor 126 and assign a resource to handle a task relating to task request 102 based on a scheduling algorithm as shown by task assignment 116. In some examples, scheduling system 114 may assign a resource in a transparent manner so that requesting processor 126 need not know which resource is being assigned to handle task request 102. As discussed in more detail below, in examples where scheduling system 114 determines that no resource in cloud 120 is related closely enough to the task, scheduling system 114 may be configured to send task request 102 to second cloud 122. In these examples, scheduling system 130 may be configured to assign a resource 132, 134, 136 in second cloud 122 to handle request 102.

[0013] In some examples, scheduling system 114 may include one or more software agents or modules that may be distributed among resources in cloud 120. In an example, scheduling system 114 may include at least one task management module 118, a resource monitoring module 112, a scheduling module 128, and/or a cloud interaction module 104, all in communication. In some examples, scheduling system 130 may include one or more software agents or modules that may be distributed among resources in cloud 122. In an example, scheduling system 130 may include at least one task management module 138, a resource monitoring module 146, a scheduling module 140, and/or a cloud interaction module 148, all in communication. In some examples, these modules may be implemented in one housing or distributed among two or more housings.

[0014] Discussion below will focus on cloud 120 though it should be clear that similar elements in cloud 122 or other clouds may have similar functions and operations. In an example, resource monitoring module 112 may be configured to receive resource data 113 relating to resources available in cloud 120. For example,

resource data 113 may include a capacity of resources 106, 108, 110, a number of processors, a size and type of memory, network traffic, usable disk space, etc.

[0015] In an example, task management module 118 may be configured to manage task request 102. For example, task management module 118 may be configured to extract and analyze task request 102 for service related data 119. In an example, service related data 119 may include a number of service requests, desired resources, desired response time, start time, communication volume, etc.

[0016] In an example, as discussed in more detail below, cloud interaction module 104 may be configured to receive task request 102 and send task request 102 to second cloud 122. In example, cloud interaction module 104 may be configured to send request 102 to second cloud 122 and receive task assignments 142 from second cloud 122. Based on the task assignments 142, cloud interaction module 104 may be configured to send information to scheduling module 128 indicating a particular resource in a particular cloud that may be selected to handle task request 102.

[0017] In an example, schedule module 128 may be configured to assign a particular resource to handle a task relating to task request 102 based on a determined correlation value between the task and the particular resource. In an example, the correlation value may indicate how much of a similarity exists between the task and the particular resource. For example, if a task should be assigned a particular type of resource, a higher correlation value may be determined between a resource of that type and the task. In an example, a correlation value between a task and a particular resource may be determined based on a correlation value between the task and tasks previously handled by the particular resource. For example, the task-resource correlation value between a task and a particular resource may increase with an increase of a sum of task-task correlation values between the task to be assigned and tasks previously assigned to the particular resource. In an example, the task-resource correlation value between a task and a particular resource may decrease with an increase of a sum of task-task correlation values between the task to be assigned and tasks assigned to any of the resources in the cloud or clouds.

[0018] In some examples, the following formula may be used by scheduling module 128:

$$C_{ip} = \frac{\sum_{k \in T_p} r_{ik}}{\sum_{j \in T} r_{ij}}, \sum_{j \in T} r_{ij} \neq 0$$

[0019] C_{ip} may be the correlation value between a task i and a resource p. A current task set may be T, a task set on a particular resource P may be T_p, r_{ik} may be the correlation value between task i and task k, r_{ij} may be the correlation value between task i and task j.

[0020] In some examples, scheduling module 128 may be configured to determine a correlation value between a task relating to task request 102 and two or more of resources 106, 108, 110 in one or more clouds 120, 122. In an example, scheduling module 128 may be configured to assign a task relating to task request 102 to the particular resource with a correlation value to the task that is a highest value, compared with other correlation values. In an example, scheduling module 128 may be configured to assign a task relating to task request 102 to the particular resource with a correlation value to the task that is above a threshold.

[0021] In an example, scheduling module 128 may assign a task relating to task request 102 based on evaluation rules that may be stored in, for example, one of resources 106, 108, 110. Evaluation rules may suggest an assignment of a resource based on factors relating to a performance of cloud 120 such as a communication load in cloud 120, load balancing among resources, scheduling overhead inside the cloud and to another cloud, response time from resources, service quality, reliability, fault-tolerance, and safety. In an example, data values used for the factors in the evaluation rules may be received by task management module 118 and resource monitoring module 112. In some examples, each data value for the relevant factors may be assigned a weight based on task request 102 and/or a state of cloud 120. In an example, if a task relating to task request 102 should receive a fast response, then the data value relating to the response time factor may receive a higher weight in the evaluation rule. In an example, if a task relating to task request 102 should receive a

high communication volume, then the value relating to the communication volume factor may receive a higher weight in the evaluation rule. A load status of cloud 120 may be used to affect a weight of the value of the load balancing factor.

[0022] In an example, task management module 118 may be configured to receive task request 102. Task management module 118 may be configured to extract service related data 119 from task request 102 and send service related data 119 to scheduling module 128. In the example, scheduling module 128 may be configured to request resource data 113 from resource monitoring module 112. In an example, scheduling module 128 may be configured to determine a correlation value between a task relating to task request 102 and at least some devices in cloud 120 and/or cloud 122. Based on the determined correlation value, task management module 118 may be configured to assign task request 126 to be handled by a particular resource 106, 108, 110 with a highest relative correlation value. In an example where none of the resources in cloud 120 have a determined correlation value to task request 102 above a threshold, task scheduling module 128 may be configured to send task request 102 to cloud interaction module 104. In an example, scheduling module 128 may be configured to send task request 102 to cloud interaction module 104 when service related data 119 indicates that task request 102 should be handled by a particular resource not available in cloud 120.

[0023] In the example, cloud interaction module 104 may be configured to send task request 102 to cloud interaction module 148 in second cloud 122. Cloud interaction module 148 in scheduling system 130 of second cloud 122 may be configured to receive task request 102 and send task request 102 to scheduling module 140. In the example, scheduling module 140 may be configured to determine the service related data in the task relating to the task request. In the example, scheduling module 130 may be configured to receive resource data relating to resources 132, 134, 136. In the example, scheduling module 140 may be configured to determine a correlation value between task request 102 and resources 132, 134, 136 based on the service related data and the resource data. In examples where scheduling module 140 determines at least one correlation value between a particular resource and task request 102 above a threshold, scheduling module 140 may be

configured to generate task assignment 142. In these examples, cloud interaction module 148 may be configured to notify cloud interaction module 104 of task assignment 142.

[0024] In some examples, scheduling module 128 may be configured to receive the notification regarding task assignment 142 from second cloud 122 and other clouds. Scheduling module 128 may be configured to determine task assignment 116 based on task assignment 142 and an intra-cloud scheduling cost between cloud 120 and 122. For example, the intra-cloud scheduling cost may take into account delays associated in sending a request from one cloud to another cloud.

[0025] Among other possible benefits, as software agents or modules may be used, communication and transparent transfer of data among two or more clouds can be realized and support of cross-cloud management may be achieved. As the resource with a high correlation value may be chosen, a communication load either inter-cloud or intra-cloud may be effectively reduced. As information to compute the correlation values may be collected by distributed modules, communication cost complexity may be reduced.

[0026] Distributed scheduling modules may be used so that each part of a cloud may have the independent capability to assign a task. Decisions may also be distributed and performed by multiple clouds. Performance may not be affected in cases when multiple clouds merge into a bigger cloud or a cloud is divided into multiple smaller clouds. Scheduling overhead may be reduced. Evaluation rules allow for a tradeoff between users' needs and system performance.

[0027] Fig. 2 depicts a flow diagram for example processes for implementing task assignment in a cloud computing environment in accordance with at least some embodiments described herein. The process in Fig. 2 could be implemented using, for example, system 100 discussed above. An example process may include one or more operations, actions, or functions as illustrated by one or more of blocks S2, S4, S6, S8, S9, S10, S12, S14, S16, S18, S20 and/or S22. Although illustrated as discrete blocks, various blocks may be divided into additional blocks, combined into fewer blocks, or eliminated, depending on the desired implementation. Processing may begin at block S2.

[0028] At block S2, a first cloud may be configured to receive a task request relating to a task and determine service related data related to the task. Processing may continue from block S2 to block S4.

[0029] At block S4, the first cloud may be configured to receive resource data relating to the first and second resource. Processing may continue from block S4 to block S6.

[0030] At block S6, the first cloud may be configured to determine a first correlation value between the task and the first resource based on the service related data and the resource data. Processing may continue from block S6 to block S8.

[0031] At block S8, the first cloud may be configured to determine a second correlation value between the task and the second resource based on the service related data and the resource data. Processing may continue from block S8 to block S9 or block S10 depending upon the first and second correlation values.

[0032] At block S9, the first cloud may be configured to assign the task to the first resource.

[0033] At block S10, the first cloud may be configured to send the task request to the second cloud. Processing may continue from block S10 block S12.

[0034] At block S12, the second cloud may be configured to receive the task request and determine the service related data. Processing may continue from block S12 to block S14.

[0035] At block S14, the second cloud may be configured to receive resource data relating to the third and fourth resource. Processing may continue from block S14 to block S16.

[0036] At block S16, the second cloud may be configured to determine a third correlation value between the task and the third resource based on the service related data and the resource data in the second cloud. Processing may continue from block S16 to block S18.

[0037] At block S18, the second cloud may be configured to determine a fourth correlation value between the task and the fourth resource based on the service related data and the resource data in the second cloud. Processing may continue from block S18 to block S20.

[0038] At block S20, the second cloud may be configured to assign the task to the third resource based on first and second correlation value to produce a task assignment. Processing may continue from block S20 to block S22.

[0039] At block S22, the second cloud may be configured to send an indication of the task assignment to the first cloud. Processing may continue from block S22 to block S24.

[0040] At block S24, the first cloud may be configured to assign the task relating to the task request based on the first correlation value, the second correlation value, and the task assignment.

[0041] Fig. 3 illustrates computer program products 300 for implementing task assignment in a cloud computing environment arranged in accordance with at least some embodiments described herein. Program product 300 may include a signal bearing medium 302. Signal bearing medium 302 may include one or more instructions 304 that, when executed by, for example, a processor, may provide the functionality described above with respect to Figs. 1-2. Thus, for example, referring to system 100, one or more of the resources in cloud 120 may undertake one or more of the blocks shown in Fig. 3 in response to instructions 304 conveyed to the system 100 by medium 302.

[0042] In some implementations, signal bearing medium 302 may encompass a computer-readable medium 306, such as, but not limited to, a hard disk drive, a Compact Disc (CD), a Digital Video Disk (DVD), a digital tape, memory, etc. In some implementations, signal bearing medium 302 may encompass a recordable medium 308, such as, but not limited to, memory, read/write (R/W) CDs, R/W DVDs, etc. In some implementations, signal bearing medium 302 may encompass a communications medium 310, such as, but not limited to, a digital and/or an analog communication medium (e.g., a fiber optic cable, a waveguide, a wired communications link, a wireless communication link, etc.). Thus, for example, program product 300 may be conveyed to one or more modules of the system 100 by an RF signal bearing medium 302, where the signal bearing medium 302 is conveyed by a wireless communications medium 310 (e.g., a wireless communications medium conforming with the IEEE 802.11 standard).

[0043] Fig. 4 is a block diagram illustrating an example computing device 400 that is arranged to perform task assignment in a cloud computing environment in accordance with at least some embodiments described herein. In a very basic configuration 402, computing device 400 typically includes one or more processors 404 and a system memory 406. A memory bus 408 may be used for communicating between processor 404 and system memory 406.

[0044] Depending on the desired configuration, processor 404 may be of any type including but not limited to a microprocessor (μP), a microcontroller (μC), a digital signal processor (DSP), or any combination thereof. Processor 404 may include one more levels of caching, such as a level one cache 410 and a level two cache 412, a processor core 414, and registers 416. An example processor core 414 may include an arithmetic logic unit (ALU), a floating point unit (FPU), a digital signal processing core (DSP Core), or any combination thereof. An example memory controller 418 may also be used with processor 404, or in some implementations memory controller 418 may be an internal part of processor 404.

[0045] Depending on the desired configuration, system memory 406 may be of any type including but not limited to volatile memory (such as RAM), non-volatile memory (such as ROM, flash memory, etc.) or any combination thereof. System memory 406 may include an operating system 420, one or more applications 422, and program data 424.

[0046] Application 422 may include a task assignment in cloud computing algorithm 426 that is arranged to perform the functions as described herein including those described previously with respect to Figs. 1 - 3. Program data 424 may include task assignment in cloud computing data 428 that may be useful for implementing a task assignment in cloud computing algorithm as is described herein. In some embodiments, application 422 may be arranged to operate with program data 424 on operating system 420 such that task assignment in a cloud computing environment may be provided. This described basic configuration 402 is illustrated in Fig. 5 by those components within the inner dashed line.

[0047] Computing device 400 may have additional features or functionality, and additional interfaces to facilitate communications between basic configuration

402 and any required devices and interfaces. For example, a bus/interface controller 430 may be used to facilitate communications between basic configuration 402 and one or more data storage devices 432 via a storage interface bus 434. Data storage devices 432 may be removable storage devices 436, non-removable storage devices 438, or a combination thereof. Examples of removable storage and non-removable storage devices include magnetic disk devices such as flexible disk drives and hard-disk drives (HDD), optical disk drives such as compact disk (CD) drives or digital versatile disk (DVD) drives, solid state drives (SSD), and tape drives to name a few. Example computer storage media may include volatile and nonvolatile, removable and non-removable media implemented in any method or technology for storage of information, such as computer readable instructions, data structures, program modules, or other data.

[0048] System memory 406, removable storage devices 436 and non-removable storage devices 438 are examples of computer storage media. Computer storage media includes, but is not limited to, RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium which may be used to store the desired information and which may be accessed by computing device 400. Any such computer storage media may be part of computing device 400.

[0049] Computing device 400 may also include an interface bus 440 for facilitating communication from various interface devices (e.g., output devices 442, peripheral interfaces 444, and communication devices 446) to basic configuration 402 via bus/interface controller 430. Example output devices 442 include a graphics processing unit 448 and an audio processing unit 450, which may be configured to communicate to various external devices such as a display or speakers via one or more A/V ports 452. Example peripheral interfaces 444 include a serial interface controller 454 or a parallel interface controller 456, which may be configured to communicate with external devices such as input devices (e.g., keyboard, mouse, pen, voice input device, touch input device, etc.) or other peripheral devices (e.g., printer, scanner, etc.) via one or more I/O ports 458. An example communication device 446

includes a network controller 460, which may be arranged to facilitate communications with one or more other computing devices 462 over a network communication link via one or more communication ports 464.

[0050] The network communication link may be one example of a communication media. Communication media may typically be embodied by computer readable instructions, data structures, program modules, or other data in a modulated data signal, such as a carrier wave or other transport mechanism, and may include any information delivery media. A "modulated data signal" may be a signal that has one or more of its characteristics set or changed in such a manner as to encode information in the signal. By way of example, and not limitation, communication media may include wired media such as a wired network or direct-wired connection, and wireless media such as acoustic, radio frequency (RF), microwave, infrared (IR) and other wireless media. The term computer readable media as used herein may include both storage media and communication media.

[0051] Computing device 400 may be implemented as a portion of a small-form factor portable (or mobile) electronic device such as a cell phone, a personal data assistant (PDA), a personal media player device, a wireless web-watch device, a personal headset device, an application specific device, or a hybrid device that include any of the above functions. Computing device 400 may also be implemented as a personal computer including both laptop computer and non-laptop computer configurations.

[0052] The present disclosure is not to be limited in terms of the particular embodiments described in this application, which are intended as illustrations of various aspects. Many modifications and variations can be made without departing from its spirit and scope, as will be apparent to those skilled in the art. Functionally equivalent methods and apparatuses within the scope of the disclosure, in addition to those enumerated herein, will be apparent to those skilled in the art from the foregoing descriptions. Such modifications and variations are intended to fall within the scope of the appended claims. The present disclosure is to be limited only by the terms of the appended claims, along with the full scope of equivalents to which such claims are entitled. It is to be understood that this disclosure is not limited to

particular methods, reagents, compounds compositions or biological systems, which can, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting.

[0053] With respect to the use of substantially any plural and/or singular terms herein, those having skill in the art can translate from the plural to the singular and/or from the singular to the plural as is appropriate to the context and/or application. The various singular/plural permutations may be expressly set forth herein for sake of clarity.

[0054] It will be understood by those within the art that, in general, terms used herein, and especially in the appended claims (*e.g.*, bodies of the appended claims) are generally intended as "open" terms (*e.g.*, the term "including" should be interpreted as "including but not limited to," the term "having" should be interpreted as "having at least," the term "includes" should be interpreted as "includes but is not limited to," etc.). It will be further understood by those within the art that if a specific number of an introduced claim recitation is intended, such an intent will be explicitly recited in the claim, and in the absence of such recitation no such intent is present. For example, as an aid to understanding, the following appended claims may contain usage of the introductory phrases "at least one" and "one or more" to introduce claim recitations. However, the use of such phrases should not be construed to imply that the introduction of a claim recitation by the indefinite articles "a" or "an" limits any particular claim containing such introduced claim recitation to embodiments containing only one such recitation, even when the same claim includes the introductory phrases "one or more" or "at least one" and indefinite articles such as "a" or "an" (*e.g.*, "a" and/or "an" should be interpreted to mean "at least one" or "one or more"); the same holds true for the use of definite articles used to introduce claim recitations. In addition, even if a specific number of an introduced claim recitation is explicitly recited, those skilled in the art will recognize that such recitation should be interpreted to mean at least the recited number (*e.g.*, the bare recitation of "two recitations," without other modifiers, means at least two recitations, or two or more recitations). Furthermore, in those instances where a convention analogous to "at

least one of A, B, and C, etc." is used, in general such a construction is intended in the sense one having skill in the art would understand the convention (*e.g.*, " a system having at least one of A, B, and C" would include but not be limited to systems that have A alone, B alone, C alone, A and B together, A and C together, B and C together, and/or A, B, and C together, etc.). It will be further understood by those within the art that virtually any disjunctive word and/or phrase presenting two or more alternative terms, whether in the description, claims, or drawings, should be understood to contemplate the possibilities of including one of the terms, either of the terms, or both terms. For example, the phrase "A or B" will be understood to include the possibilities of "A" or "B" or "A and B."

[0055] In addition, where features or aspects of the disclosure are described in terms of Markush groups, those skilled in the art will recognize that the disclosure is also thereby described in terms of any individual member or subgroup of members of the Markush group.

[0056] As will be understood by one skilled in the art, for any and all purposes, such as in terms of providing a written description, all ranges disclosed herein also encompass any and all possible subranges and combinations of subranges thereof. Any listed range can be easily recognized as sufficiently describing and enabling the same range being broken down into at least equal halves, thirds, quarters, fifths, tenths, etc. As a non-limiting example, each range discussed herein can be readily broken down into a lower third, middle third and upper third, etc. As will also be understood by one skilled in the art all language such as "up to," "at least," "greater than," "less than," and the like include the number recited and refer to ranges which can be subsequently broken down into subranges as discussed above. Finally, as will be understood by one skilled in the art, a range includes each individual member. Thus, for example, a group having 1-3 cells refers to groups having 1, 2, or 3 cells. Similarly, a group having 1-5 cells refers to groups having 1, 2, 3, 4, or 5 cells, and so forth.

[0057] While various aspects and embodiments have been disclosed herein, other aspects and embodiments will be apparent to those skilled in the art. The various aspects and embodiments disclosed herein are for purposes of illustration and

are not intended to be limiting, with the true scope and spirit being indicated by the following claims.

CLAIMS

What is claimed is:

1. A method for assigning a task in a cloud, wherein the cloud comprises at least a first resource and a second resource, the method comprising:

receiving, at a first processor, a task request relating to a task;

determining, at the first processor, service related data relating to the task based on the task request;

receiving, at a second processor, resource data relating to the first and second resource;

receiving, at a third processor, the service related data and the resource data;

determining, at the third processor, a first correlation value between the task and the first resource, wherein the first correlation value is based on the service related data and the resource data;

determining at the third processor, a second correlation value between the task and the second resource, wherein the second correlation value is based on the service related data and the resource data;

assigning, by the third processor, the task to the first resource based on the first correlation value and the second correlation value.

2. The method as recited in claim 1, wherein the first, second, and third processors are all in the same housing.

3. The method as recited in claim 1, further comprising assigning, by the third processor, the task to the first resource based on an evaluation rule.

4. The method as recited in claim 1, wherein the evaluation rule is based on a performance of the cloud.

5. The method as recited in claim 1, wherein the first correlation value

relates to a task previously handled by the first resource and relates to a task previously handled by the second resource.

6. The method as recited in claim 1, wherein:

the service related data includes at least one of a number of task requests received by the first processor, a resource requested in the task request, a response time requested in the task request, a start time requested in the task request, or a communication volume requested in the task request; and

the resource data includes at least one of a capacity of the first and second resource, a number of processors in the first and second resource, a size of memory in the first and second resources, network traffic to the first and second resources, or disc space in the first and second resource.

7. A cloud effective to assign a task, the cloud comprising:

a first resource;

a second resource;

a first processor;

a second processor in communication with the first resource and the second resource; and

a third processor in communication with the first processor and the second processor;

wherein

the first processor is effective to receive a task request relating to a task and determine service related data relating to the task based on the task request;

the second processor is effective to receive resource data relating to the first and second resource;

the third processor is effective to determine a first correlation value between the task and the first resource, wherein the first correlation value is based on the service related data and the resource data;

the third processor is effective to determine a second correlation value between the task and the second resource, wherein the second correlation value is based on the service related data and the resource data; and

the third processor is effective to assign the task to the first resource based on the first and the second correlation value.

8. The cloud as recited in claim 7, wherein the first, second, and third processors are all disposed in the same housing.

9. The cloud as recited in claim 7, wherein the third processor is effective to assign the task to the first resource based on an evaluation rule.

10. The cloud as recited in claim 7, wherein the evaluation rule is based on a performance of the cloud.

11. The cloud as recited in claim 7, wherein the first correlation value relates to a task previously handled by the first resource and relates to a task previously handled by the second resource.

12. A system effective to assign a task, the system comprising:

a first cloud including a first access point, a first resource and a second resource;

a second cloud including a second access point, a third resource and a fourth resource;

wherein the first cloud is effective to

receive a task request relating to a task;

determine service related data relating to the task based on the task request;

receive first resource data relating to the first and second resource;

determine a first correlation value between the task and the first resource, wherein the first correlation value is based on the service related data and the first resource data;

determine a second correlation value between the task and the second resource, wherein the second correlation value is based on the service related data and the first resource data;

send the task request to the second cloud based on the first and second correlation value;

the second cloud is effective to

receive the task request;

determine the service related data relating to the task based on the task request;

receive second resource data relating to the third and fourth resource;

determine a third correlation value between the task and the third resource, wherein the third correlation value is based on the service related data and the second resource data;

determine a fourth correlation value between the task and the fourth resource, wherein the fourth correlation value is based on the service related data and the second resource data; and

assign the task to the third resource based on the first and the second correlation value to produce a task assignment.

13. The system as recited in claim 12, wherein in the first cloud:

a first processor is effective to receive the task request relating to the task and determine the service related data;

a second processor is effective to receive the first resource data relating to the first and second resource;

a third processor is effective to determine the first and second correlation value; and

a fourth processor is effective to send the task request to the second cloud.

14. The system as recited in claim 13, wherein the first, second, third and fourth processors are all disposed in the same housing.

15. The system as recited in claim 13,

wherein in the second cloud:

a fifth processor is effective to receive the task request relating to the task and determine the service related data;

a sixth processor is effective to receive the second resource data relating to the third and fourth resource;

a seventh processor is effective to determine the third and fourth correlation value; and

an eight processor is effective to receive another task request to the second cloud; wherein

the fifth and the eight processors are disposed in different housings.

16. The system as recited in claim 12, wherein:

the second cloud is effective to send an indication of the task assignment to the first cloud; and

the first cloud is effective to assign the task based on the first correlation value, the second correlation value, and the task assignment.

17. The system as recited in claim 16, wherein the first cloud is effective to assign the task based on a scheduling cost between the first cloud and the second cloud.

18. The system as recited in claim 12, wherein the first cloud is effective to assign the task to the first resource based on an evaluation rule.

19. The system as recited in claim 18, wherein the evaluation rule is based on a performance of the first cloud.

20. The system as recited in claim 12, wherein the first correlation value relates to a task previously handled by the first resource.

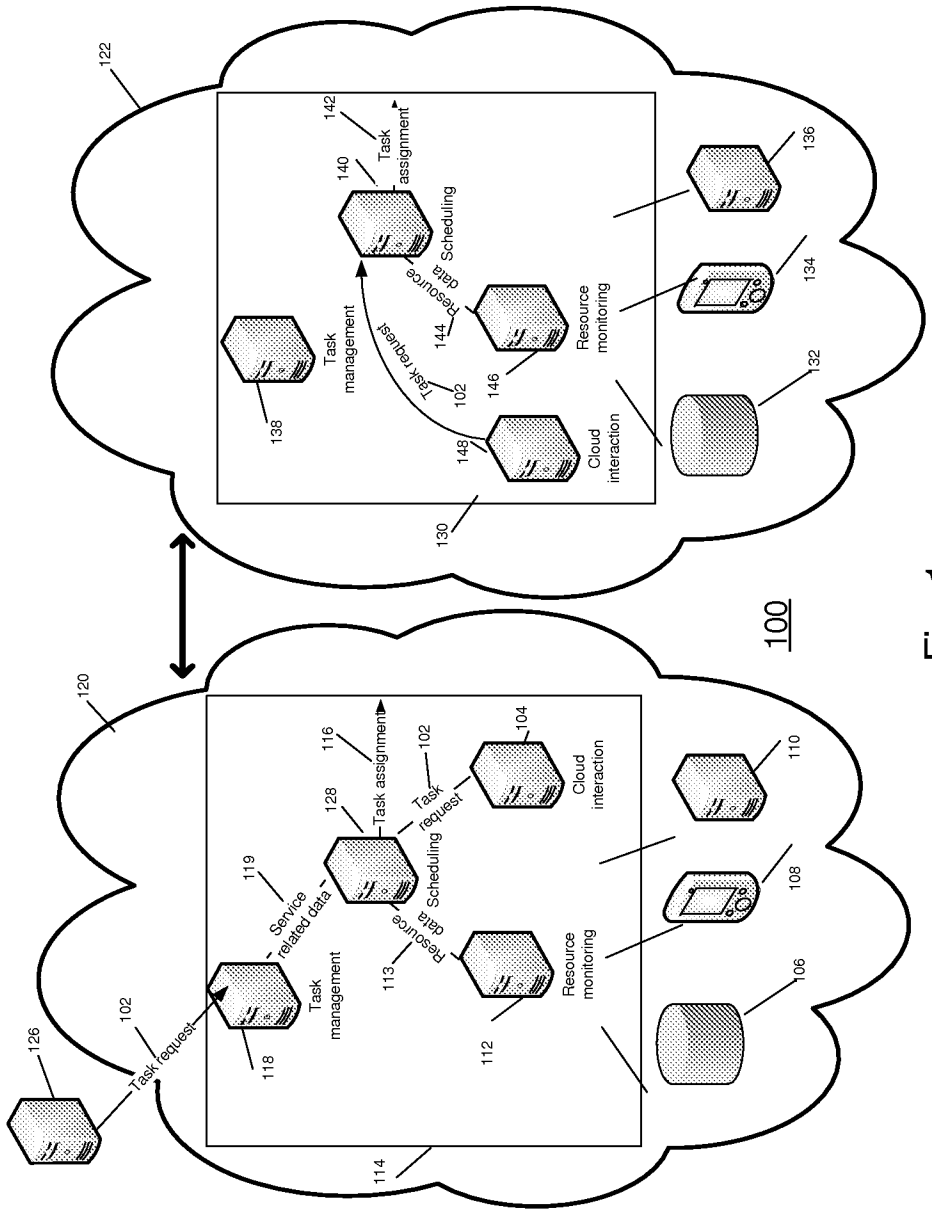

122
142
Task assignment
140
Scheduling
Resource data
144
Task request
102
146
Resource monitoring
Task management
138
148
Cloud interaction
130
136
134
132
100
120
116
Task assignment
102
Task request
104
Cloud interaction
128
Scheduling
119
Service related data
Resource data
113
Task management
118
Resource monitoring
112
110
108
106
126
102
Task request
114

Fig. 1

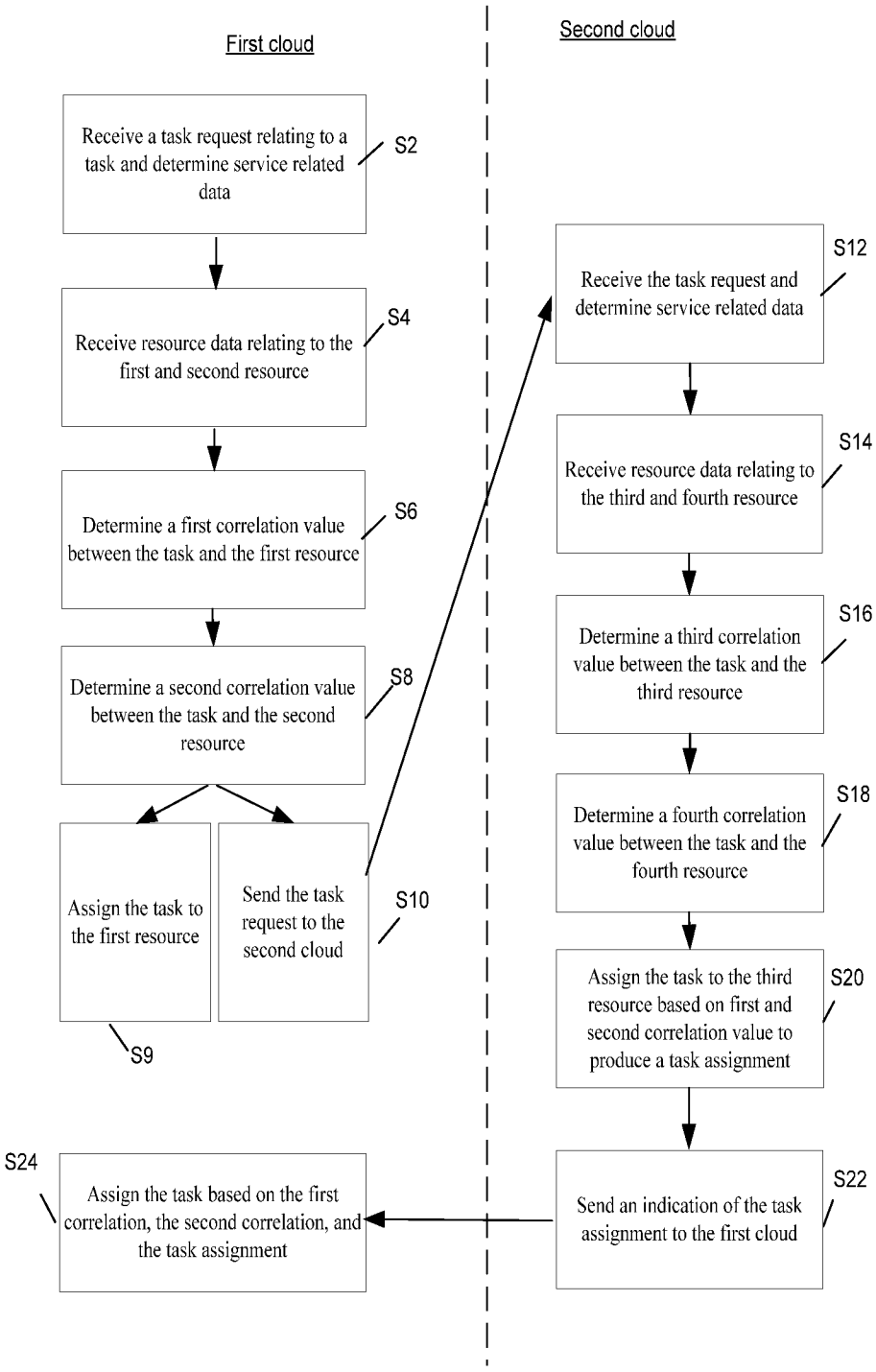

First cloud
Second cloud
Receive a task request relating to a task and determine service related data
S2
Receive resource data relating to the first and second resource
S4
Determine a first correlation value between the task and the first resource
S6
Determine a second correlation value between the task and the second resource
S8
Assign the task to the first resource
Send the task request to the second cloud
S10
S9
Receive the task request and determine service related data
S12
Receive resource data relating to the third and fourth resource
S14
Determine a third correlation value between the task and the third resource
S16
Determine a fourth correlation value between the task and the fourth resource
S18
Assign the task to the third resource based on first and second correlation value to produce a task assignment
S20
Send an indication of the task assignment to the first cloud
S22
S24
Assign the task based on the first correlation, the second correlation, and the task assignment

Fig. 2

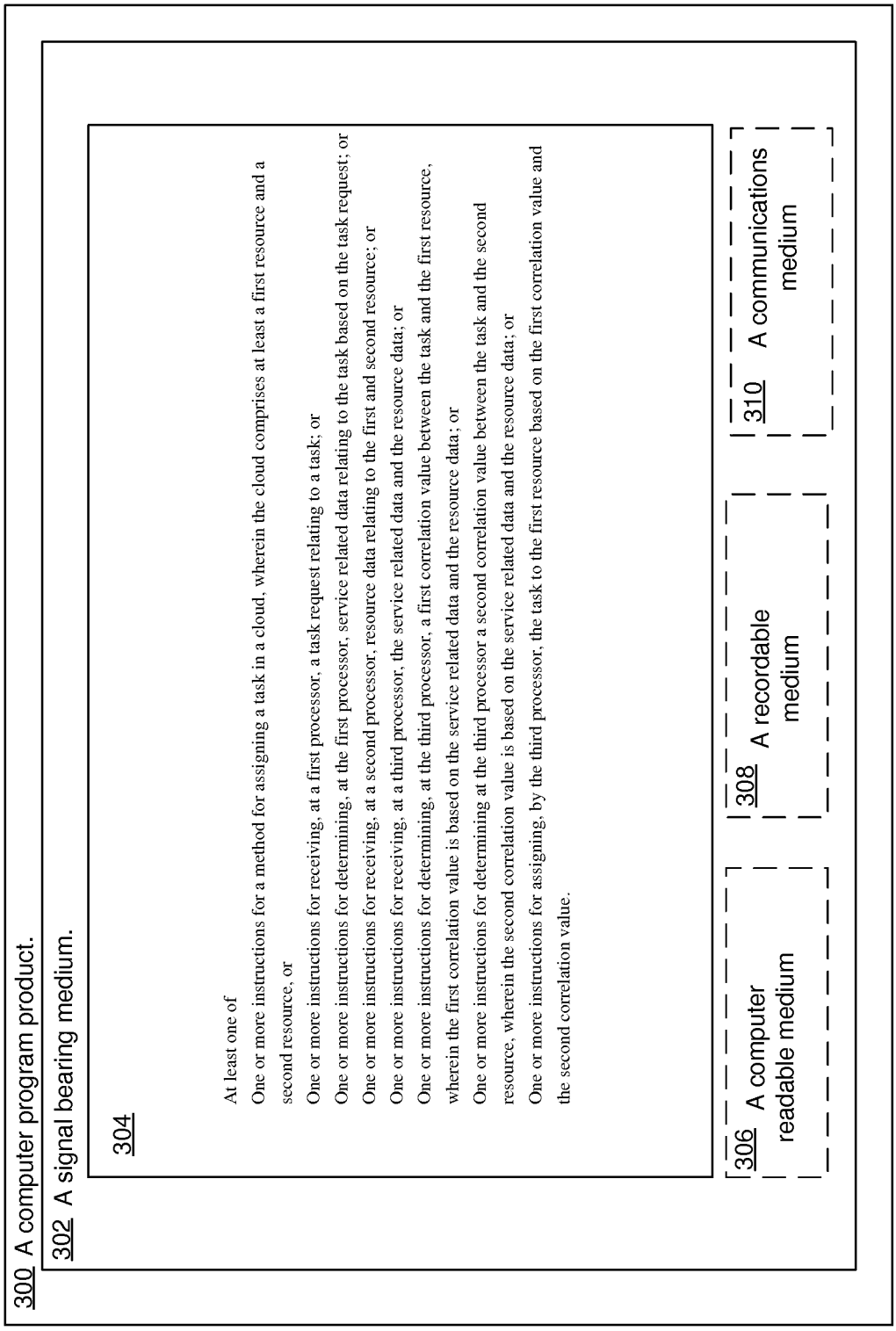

300 A computer program product.
302 A signal bearing medium.
304
At least one of
One or more instructions for a method for assigning a task in a cloud, wherein the cloud comprises at least a first resource and a second resource, or
One or more instructions for receiving, at a first processor, a task request relating to a task; or
One or more instructions for determining, at the first processor, service related data relating to the task based on the task request; or
One or more instructions for receiving, at a second processor, resource data relating to the first and second resource; or
One or more instructions for receiving, at a third processor, the service related data and the resource data; or
One or more instructions for determining, at the third processor, a first correlation value between the task and the first resource, wherein the first correlation value is based on the service related data and the resource data; or
One or more instructions for determining at the third processor a second correlation value between the task and the second resource, wherein the second correlation value is based on the service related data and the resource data; or
One or more instructions for assigning, by the third processor, the task to the first resource based on the first correlation value and the second correlation value.
306 A computer readable medium
308 A recordable medium
310 A communications medium

Fig. 3

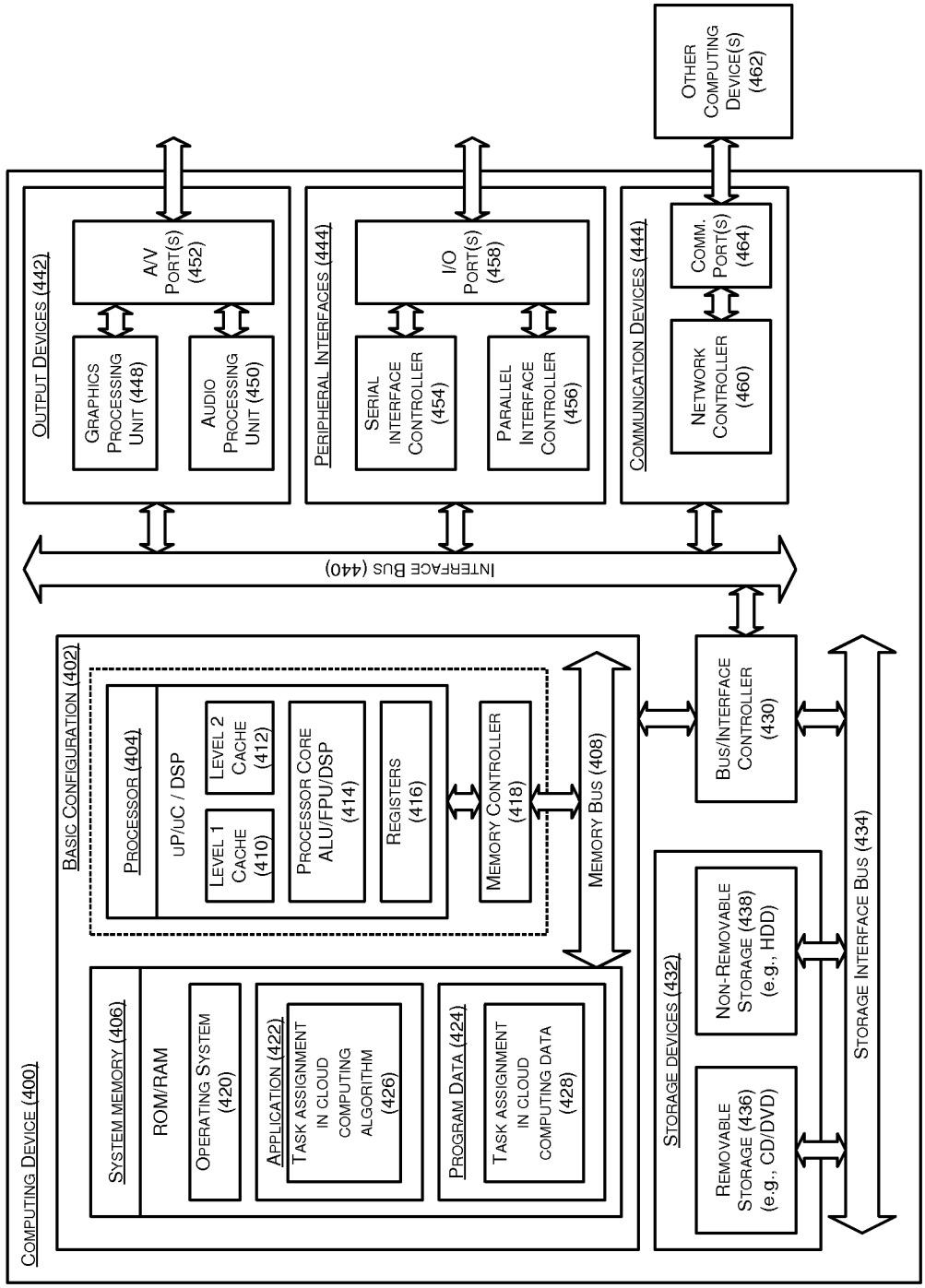

COMPUTING DEVICE (400)
BASIC CONFIGURATION (402)
PROCESSOR (404)
UP/UC / DSP
LEVEL 1 CACHE (410)
LEVEL 2 CACHE (412)
PROCESSOR CORE ALU/FPU/DSP (414)
REGISTERS (416)
MEMORY CONTROLLER (418)
SYSTEM MEMORY (406)
ROM/RAM
OPERATING SYSTEM (420)
APPLICATION (422)
TASK ASSIGNMENT IN CLOUD COMPUTING ALGORITHM (426)
PROGRAM DATA (424)
TASK ASSIGNMENT IN CLOUD COMPUTING DATA (428)
MEMORY BUS (408)
BUS/INTERFACE CONTROLLER (430)
INTERFACE BUS (440)
OUTPUT DEVICES (442)
GRAPHICS PROCESSING UNIT (448)
AUDIO PROCESSING UNIT (450)
A/V PORT(S) (452)
PERIPHERAL INTERFACES (444)
SERIAL INTERFACE CONTROLLER (454)
PARALLEL INTERFACE CONTROLLER (456)
I/O PORT(S) (458)
COMMUNICATION DEVICES (444)
NETWORK CONTROLLER (460)
COMM. PORT(S) (464)
OTHER COMPUTING DEVICE(S) (462)
STORAGE DEVICES (432)
REMOVABLE STORAGE (436) (e.g., CD/DVD)
NON-REMOVABLE STORAGE (438) (e.g., HDD)
STORAGE INTERFACE BUS (434)

FIG. 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2010/076936

A. CLASSIFICATION OF SUBJECT MATTER

H04L 29/08 (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)

IPC: H04L, H04W, H04Q, 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)

CNPAT,WPI,EPODOC,CNKI,IEEE: cloud, distribut+,task, assign+, schedul+, request+, resource, service,data, document?, file?, correllat+, relat+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN101753608A (CHINA MOBILE COMMUNICATION CORP.) 23 Jun. 2010（23.06.2010）the whole document	1-20
A	CN101662483A (SHENGDA COMPUTER SHANGHAI CO., LTD.) 03 Mar.2010（03.03.2010）the whole document	1-20
A	CN101719842A (INST. SOFTWARE CHINESE ACAD. SCI.) 02 Jun. 2010（02.06.2010）the whole document	1-20
A	CN101557551A (CHENGDU HUAWEI SAIMEN TIEKE TECHNOLOGY LTD.) 14 Oct.2009（14.10.2009）the whole document	1-20
A	US2010/0030866A1 (AMEYA COMPUTING, INC.) 04 Feb.2010（04.02.2010）the whole document	1-20

☐ 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	Date of mailing of the international search report
31 May 2010（31.05.2010）	**23 Jun. 2011 (23.06.2011)**
Name and mailing address of the ISA/CN The State Intellectual Property Office, the P.R.China 6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China 100088 Facsimile No. 86-10-62019451	Authorized officer XIA,Kan Telephone No. (86-10)62413461

Form PCT/ISA /210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2010/076936

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN101753608A	23.06.2010	none	
CN101662483A	03.03.2010	none	
CN101719842A	02.06.2010	none	
CN101557551A	14.10.2009	none	
US2010/0030866A1	04.02.2010	none	