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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
1 October 2015 (01.10.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/147860 A1

(51) International Patent Classification:
H04L 12/24 (2006.01) *H04L 12/16* (2006.01)

(21) International Application Number:
PCT/US2014/032155

(22) International Filing Date:
28 March 2014 (28.03.2014)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **HEWLETT-PACKARD DEVELOPMENT COMPANY, L.P.** [US/US]; 11445 Compaq Center Drive W., Houston, Texas 77070 (US).

(72) Inventors: **WACKERLY, Shaun**; 8000 Foothills Boulevard, Roseville, California 95747 (US). **BRITT, Julie**; 8000 Foothills Boulevard, Roseville, California 95747 (US). **KRUEGER, Marjorie**; 8000 Foothills Boulevard, Roseville, California 95747 (US).

(74) Agents: **FERGUSON, Christopher Ward** et al.; Hewlett-Packard Company, Intellectual Property Administration, Mail Stop 35 3404 E. Harmony Road, Fort Collins, Colorado 80528 (US).

(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, BN, 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, IR, IS, JP, KE, KG, 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, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to the identity of the inventor (Rule 4.17(i))*

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

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

(54) Title: RESCHEDULING A SERVICE ON A NODE

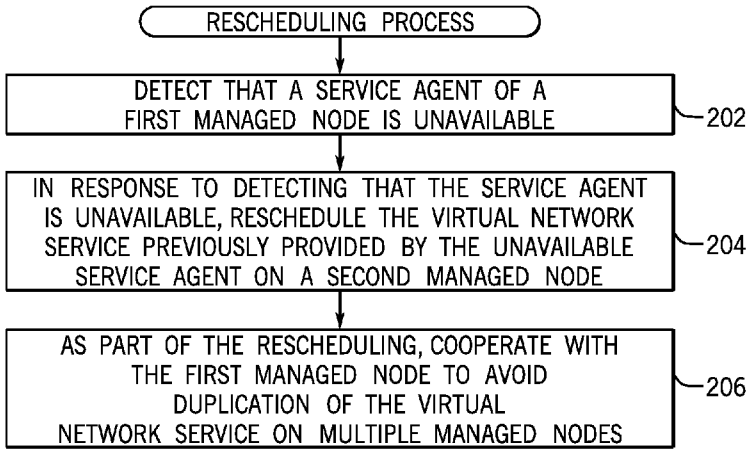

FIG. 2

(57) Abstract: A controller detects that an agent of a first node managed by the controller is unavailable, the agent providing a service accessible by a tenant of a cloud infrastructure that includes the controller and a plurality of nodes managed by the controller. In response to the detecting, the controller reschedules the service on a second node managed by the controller to continue to provide availability of the service to the tenant. As part of the rescheduling, cooperate, by the controller, with the first node to avoid duplication of the service on multiple nodes including the first and second nodes.

WO 2015/147860 A1

RESCHEDULING A SERVICE ON A NODE

Background

[0001] A network infrastructure composed of various network entities can be used by devices to communicate with each other. Examples of network entities include switches, routers, configuration servers (*e.g.* Dynamic Host Configuration Protocol or DHCP servers), and so forth.

[0002] Traditionally, the network infrastructure of a particular network is owned by a network operator. For example, an enterprise, such as a business concern, educational organization or government agency, can operate a network for use by users (*e.g.* employees, customers, etc.) of the enterprise. The network infrastructure of such network is owned by the enterprise.

[0003] In an alternative arrangement, instead of using a network operator's own network infrastructure to implement a network, the network operator can instead pay to use networking entities provided by a third party service provider. The service provider provides an infrastructure that includes various network entities accessible by customers (also referred to as "tenants") of the service provider. By using the infrastructure of the service provider, an enterprise would not have to invest in various components of a network infrastructure, and would not have to be concerned with maintenance of the network infrastructure. In this way, an enterprise's experience in setting up a network configuration is simplified. In addition, flexibility is enhanced since the network configuration can be more easily modified for new and evolving data flow patterns. Moreover, the network configuration is scalable to meet rising data bandwidth demands.

Brief Description Of The Drawings

[0004] Some implementations are described with respect to the following figures.

[0005] Fig. 1 is a block diagram of an example arrangement that includes a network cloud infrastructure and tenant systems, according to some implementations.

[0006] Fig. 2 is a flow diagram of a rescheduling process according to some implementations.

[0007] Fig. 3 is a flow diagram of a decommissioning process according to some implementations.

[0008] Fig. 4 is a flow diagram of a rejoin control process according to some implementations.

[0009] Fig. 5 is a block diagram of a controller according to some implementations.

Detailed Description

[0010] Fig. 1 is a block diagram of an example arrangement that includes a network cloud infrastructure 100, which may be operated and/or owned by a network service provider. The network cloud infrastructure 100 has customers (also referred to as "tenants") that operate respective tenant systems 102. Each tenant system 102 can include a network deployment that uses network entities of the network cloud infrastructure 100. The provision of network entities of a network cloud infrastructure by a network service provider to a tenant is part of a cloud-service model that is sometimes referred to as network as a service (NaaS) or infrastructure as a service (IaaS).

[0011] The network cloud infrastructure 100 includes both physical elements and virtual elements. The physical elements include managed nodes 106, which can include computers, physical switches, and so forth. The virtual elements in the network cloud infrastructure 100 are included in the managed nodes 106. More specifically, the managed nodes 106 include service agents 104 that provide virtual network services that are useable by the tenant systems 102 on demand.

[0012] A service agent 104 can be implemented as machine-readable instructions executable within a respective managed node 106. A service agent 104 hosts or provides a virtual network service that can be used in a specific network configuration of a tenant system 102. Each managed node 106 can include one or multiple service agents 104, and each service agent 104 can provide one or multiple virtual network services.

[0013] Virtual network services provided by service agents 104 can include any or some combination of the following: a switching service provided by a switch for switching data between devices at layer 2 of the Open Systems Interconnection (OSI) model; a routing service for routing data at layer 3 of the OSI model;a configuration service provided by a configuration server, such as a Dynamic Host Configuration Protocol (DHCP) server used for setting network configuration parameters such as Internet Protocol (IP) addresses for devices that communicate over a network; a security service provided by a security enforcement entity for enforcing a security policy; a domain name service provided by a domain name system (DNS) server that associates various information (including an IP address) with a domain name; and so forth.

[0014] Although examples of various network services are listed above, it is noted that service agents 104 can provide other types of virtual network services that are useable in a network deployment of a tenant system 102.

[0015] A virtual network service, or an agent that provides a virtual network service, constitutes a virtual network element in the network cloud infrastructure 100. The virtual network entities are “virtual” in the sense that the network entities are not physical entities within a network deployment of a respective tenant system 102, but rather entities (provided by a third party such as the network service provider of the network cloud infrastructure 100) that can be logically implemented in the network deployment.

[0016] More generally, a cloud infrastructure can include service agents 104 that provide virtual services useable in a tenant system 102. Such virtual services can

include services of processing resources, services of storage resources, services of software (in the form of machine-readable instructions), and so forth.

[0017] In the ensuing discussion, reference is made to provision of virtual network services. However, techniques or mechanisms according to some implementations can be applied to other types of virtual services provided by nodes of a cloud infrastructure.

[0018] When a fault occurs in the network cloud infrastructure 100 that causes a managed node 106 or a service agent 104 in a managed node 106 to go down (enter into a low power mode or off state, enter into a failed state, or otherwise enter into a state where the managed node 106 or service agent 104 becomes non-operational), a virtual network service may become temporarily unavailable. Examples of faults in the network cloud infrastructure 100 can cause a managed node 106 or an agent 104 to become unavailable can include any or some combination of the following: failure of a physical element such as a component in a managed node 106, an error during execution of machine-readable instructions, loss of communication over a physical network link, and so forth.

[0019] As another example, an administrator of the network cloud infrastructure 100 may issue an instruction to decommission a managed node 106, which will also cause a corresponding virtual network service to become unavailable. Decommissioning a managed node 106 refers to taking the managed node 106 out of service, which can be performed to repair, upgrade, or replace the decommissioned managed node 106, as examples. As discussed further below, decommissioning of a managed node 106 can be performed by a node decommissioner 116 executing in the controller 108 (or another system). The node decommissioner 116 can be implemented as machine-readable instructions.

[0020] In either scenario (a first scenario where a fault causes a managed node or service agent to go down, or a second scenario in which a managed node is decommissioned), a tenant system 102 that uses a virtual network service associated with the managed node 106 or service agent 104 that has gone down

may notice that the virtual network service has become unavailable (the virtual network service can no longer be used by the tenant system 102). The detection of the unavailability of the virtual network service by the tenant system 102 may cause disruption of operation of the tenant system 102.

[0021] If disruption is detected at the tenant system 102, an administrator of the tenant system 102 (or alternatively, an administrator of the network cloud infrastructure 100) may have to perform manual re-configuration of a network deployment at the tenant system 102 to address the disruption due to unavailability of the virtual network service. Such manual re-configuration may take a relatively long period of time, and also may be labor intensive.

[0022] In accordance with some implementations, a controller 108 in the network cloud infrastructure 100 is able to perform rescheduling of a virtual network service on a different managed node 106 in response to the controller 108 detecting that a service agent providing the virtual network service has become unavailable, in any of the scenarios discussed above. Rescheduling the virtual network service includes causing the virtual network service to be provided by a second service agent instead of by a first service agent (which has become unavailable). The first service agent is executed in a first managed node 106, while the second service agent is executed in a second managed node 106.

[0023] By performing the automatic rescheduling of the virtual network service on a different managed node 106, service disruption at a tenant system 102 can be avoided. From the perspective of the tenant system that uses the virtual network service provided by the service agent that has become unavailable, the virtual network service appears to be continually available during the rescheduling. As a result, seamless availability of the virtual network service is provided to the tenant system 102 in the presence of a fault or a decommissioning action that causes a service agent 104 to become unavailable.

[0024] The controller 108 can be a controller that manages the managed nodes 106 in the network cloud infrastructure 100. The controller 108 is able to direct

which virtual network services are provided by service agents on which managed nodes 106. Although just one controller 108 is shown in Fig. 1, it is noted that in other examples, the network cloud infrastructure 100 can include multiple controllers 108 for managing the managed nodes 106.

[0025] In some examples, the arrangement shown in Fig. 1 in which the controller 108 manages managed nodes 106 can be part of a software-defined networking (SDN) arrangement, in which machine-readable instructions executed by the controller 108 perform management of the managed nodes 106. In the SDN arrangement, the controller 108 is referred to as an SDN controller that is part of a control plane, while the managed nodes 106 are part of a data plane through which user or tenant traffic is communicated. User or tenant traffic does not have to be communicated through the control plane. The controller 108 is responsible for determining where (which of the managed nodes 106) a virtual network service is to be hosted, while a managed node is responsible for deploying a specific network service.

[0026] In some examples, communications between the controller 108 and the managed nodes 106 can be according to a Representational State Transfer (REST) protocol. In other examples, communications between the controller 108 and the managed nodes 106 can be according to other protocols.

[0027] The rescheduling of a virtual network service from a first managed node 106 to a second managed node 106 due to unavailability of a service agent can be performed by a scheduler 110 that executes in the controller 108. The scheduler 110 can be implemented as machine-readable instructions, in some examples.

[0028] The controller 108 can maintain node information 112 describing physical attributes of each managed node 106. The physical attributes of a managed node106 can include any or some combination of the following: number of processors, processor speed, type of operating system, storage capacity, and so forth. The controller 108 also maintains agent information 114, which relates to a service agent(s) of each managed node 106. The information pertaining to the

service agent(s) include information describing the capability of each service agent to host a respective virtual network service, information associating a service agent with a corresponding managed node 106, and other information relating to characteristics of each service agent. Service agents 104 can send their information to the controller 108, on a repeated basis, for inclusion in the agent information 114.

[0029] The node information 112 and agent information 114 can be stored in a storage medium within the controller 108, or in a storage medium outside the controller 108.

[0030] When a tenant system 106 wishes to employ a given virtual network service, the controller 108 can schedule the requested virtual network service on a selected service agent 104 residing on a corresponding managed node 106. More specifically, the tenant system 102 can submit a request for certain virtual network services. In response to the request, the controller 108 can determine which service agents 104 on which managed nodes 106 are to host the requested virtual network services.

[0031] Fig. 2 is a flow diagram of a process for rescheduling a virtual network service, in accordance with some implementations. The process can be performed by components (including the scheduler 110) in the controller 108. The controller 108 detects (at 202) that a first service agent 104 of a first managed node 106 is unavailable. As noted above, the unavailability of the first service agent 104 can be due to a fault in the network cloud infrastructure 100, or due to an explicit action to decommission the first managed node 106.

[0032] Detecting unavailability of a service agent can be based on checking for a heartbeat message from the service agent. If the controller 108 determines that the service agent 104 has not reported availability (by sending a heartbeat message), then the controller 108 makes a determination that the service agent is unavailable, and the status of the service agent 104 is marked accordingly. In some examples, the controller 108 can provide an alert service configured to send notification of an

unavailable service agent (along with other specified events) to a designated recipient, such as an administrator of the network cloud infrastructure 100.

[0033] In response to detecting that the first service agent 104 is unavailable, the scheduler 110 in the controller 108 reschedules (at 204) the virtual network service previously provided by the unavailable service agent 104 on a second managed node 106, to continue to provide availability of the virtual network service to a tenant system 102. As part of the rescheduling, the controller 108 cooperates (at 206) with the first managed node 106 to avoid duplication of the virtual network service on multiple nodes that include the first and second managed nodes.

[0034] In some implementations, the network cloud infrastructure 100 can be configured with a first physical network for communication of management traffic between the controller 108 and the managed nodes 106, and a second, different physical network for tenant data connections (for communicating data of network deployments of the tenant systems 102). A condition that results in the controller 108 losing contact with a service agent may not represent loss of the respective virtual network service to a tenant system 102 because of the separate management and tenant data networks. For example, if a managed node 106 loses its management network connectivity to the controller 108, it may appear to the controller 108 that the service agents 104 on that managed node 106 have become unavailable, even though the service agents are still running on the managed node 106, and thus providing tenant services to a tenant over the tenant data network. In this scenario, when the controller 108 reschedules the virtual network service of a first service agent to a second service agent, duplicate virtual network services (one provided by the first service agent and another provided by the second virtual service agent) may be provided for a network deployment of the tenant system 102.

[0035] The cooperation (206) between the controller 108 and the first managed node 106 to avoid duplication of a virtual network service can involve the following tasks, in some implementations. Both the first managed node 106 and the controller 108 are configured to detect loss of management connectivity. If the first managed node 106 detects the loss of management connectivity to the controller 108, then the

first managed node 106 can decommission all virtual network services on the first managed node 106, in anticipation of the controller 108 rescheduling such virtual network services on another managed node (or other managed nodes) 106. The process of decommissioning the virtual network services and rescheduling the virtual network services can be performed relatively quickly so that tenant systems 102 do not notice the temporary unavailability of the virtual network services. In addition, to prevent a "flapping rescheduling" condition (where the controller reschedules a virtual network service from a first managed node 106 to a second managed node 106, followed quickly by rescheduling the same virtual network service back from the second managed node 106 to the first managed node 106), the controller 108 can perform actions to prevent rejoinder of the first managed node 106 with which the controller 108 has recently lost communication, similar to actions performed according to Fig. 4 (discussed further below).

[0036] In addition to being able to reschedule virtual network services in response to detecting unavailability of service agents, the scheduler 110 of the controller 108 can also perform load balancing to balance workload across the managed nodes 106. Re-balancing workload across the managed nodes 106 can be accomplished by rescheduling, using the scheduler 110, virtual network services across different service agents 104 in the managed nodes 106. The network cloud infrastructure 100 may change over time, such as due to addition of new managed nodes 106 and/or new service agents 104. When the new managed nodes 106 and/or new service agents 104 register with the controller 108, the scheduler 110 can perform rescheduling of virtual network services to perform re-balancing of workload.

[0037] In some cases, new managed nodes and/or new service agents may possess greater performance characteristics or enhanced service features. By rescheduling virtual network services to such new managed nodes and/or new service agents, the controller 108 can better balance workload across the managed nodes 106, as well as to take advantage of enhanced performance characteristics or service features. Rebalancing virtual network services can also provide greater

reliability as more managed nodes 106 are deployed into the network cloud infrastructure 100. The ability of the network cloud infrastructure 100 to tolerate node failure without service interruption is a factor of the available unused service hosting capacity across the managed nodes 106. In network cloud infrastructure with N managed nodes capable of hosting virtual network services, if N-1 nodes fail, all services might end up hosted on the remaining node. As the failed nodes become available again, rebalancing allows the virtual network services to be redistributed across the available nodes, to achieve better usage of available resources for providing virtual network services.

[0038] Fig. 3 is a flow diagram of a node decommissioning process according to some implementations. The decommissioning process can be performed by the node decommissioner 116, in some examples, or by a different module, whether executing on the controller 108 or on another system. The node decommissioner 116 receives (at 302) a notification (such as from an administrator of the network cloud infrastructure 100 or another requester) that a given managed node 106 is to be taken offline.

[0039] In response to the notification, the node decommissioner 116 removes (at 304) the service agents of the given managed node 106 from a pool of available service agents. The pool of available service agents can be stored as part of the agent information 114 (Fig. 1). After removing the service agents of the given managed node 106 from the pool of available service agents, the controller 108 allows service agents 104 on the given managed node 106 to finish processing any remaining service requests.

[0040] The node decommissioner 116 can further notify (at 306) the scheduler 110 of the service agents that are removed from the pool of available service agents. This notification can cause the scheduler 110 to begin the computations relating to rescheduling of the virtual network services provided by the service agents that have been removed. Such computations can allow the rescheduling of the hosted virtual network services of the given managed node 106 to complete more quickly at a later time.

[0041] The node decommissioner 116 then notifies (at 308) the given managed node 106 to go offline so that the given managed node 106 can prepare to shut down or otherwise become inactive. This notification indicates to the given managed node 106 that the controller 108 is no longer controlling the given managed node 106.

[0042] Next, the node decommissioner 116 removes (at 310) information relating to the given managed node 106 and the corresponding service agents from the controller 108, such as by removing such information from the node information 112 and the agent information 114 (Fig. 1). Removing the information relating to the given managed node 106 and the corresponding service agents from the controller 108 triggers virtual network services provided by the service agents to be rescheduled by the scheduler 110 to another service agent (or other service agents).

[0043] The node decommissioner 116 further disconnects (at 312) the given managed node's control and data plane network interfaces so that the given managed node's service hosting capacity effectively ceases to exist from the network cloud infrastructure 100. The control plane network interface of the given managed node 106 is used to communicate with the controller 108, while the data plane interface of the given managed node 106 is used to communicate data with other network entities.

[0044] Fig. 4 is a flow diagram of a rejoin control process that can be performed by the node decommissioner 116, or by another module. The node decommissioner 116 tracks (at 402) recent removals of managed nodes (such as performed at 310 in Fig. 3). When information of a managed node 106 is removed from the controller 108, the node decommissioner 116 stores (at 404) information relating to the removed managed node 106 in a removal data structure (*e.g.* cache, log, etc.) that contains information of recently removed managed nodes. The data structure can store identifiers of the removed managed nodes, as well as time information indicating the latest time when each managed node 106 was removed from the view of the controller 108.

[0045] When a managed node 106 is notified that the controller 108 has removed the managed node from the controller's view, the managed node 106 may attempt to rejoin the controller 108. A managed node rejoining the controller 108 refers to the managed node 106 performing a registration procedure with the controller 108 to make the controller 108 aware of the presence and availability of the managed node 106. If the controller 108 allows the recently removed managed node 106 to fully rejoin the controller 108, then new virtual network services may be scheduled onto the rejoined managed node 106 even though the rejoined managed node 106 is being brought offline.

[0046] In accordance with some implementations, in response to receiving (at 406) a request from a given managed node 106 to rejoin the controller 108, the node decommissioner 116 checks (at 408) the removal data structure to determine if the removal data structure contains time information regarding when the given managed node 106 was removed. If the time information is in the removal data structure, then the node decommissioner 116 compares (at 410) the time information from the removal data structure with a current time to determine (at 412) if the elapsed time (time since removal of the given managed node) is greater than a specified threshold. If not, then the request to rejoin is denied (at 414) by the node decommissioner 116. If the elapsed time is greater than the specified threshold, then the node decommissioner 116 grants (at 416) the request to rejoin.

[0047] In some examples, the denial of the request to rejoin is a denial of the request to fully rejoin the recently removed managed node 106. The node decommissioner 116 can still allow rejoining of the recently removed managed node 106 in a partial capacity, where the recently removed managed node 106 is excluded from the pool of managed nodes on which virtual network services can be scheduled. However, the partially rejoined managed node 106 can remain operational to allow for interaction with an administrator through the controller 108, for example.

[0048] By using techniques or mechanisms according to some implementations, tenant cloud service availability is not interrupted by faults or node decommissioning

in the network cloud infrastructure 100. As a result, an administrator can fix infrastructure issues in the network cloud infrastructure 100 without interrupting service to tenants. By rescheduling services automatically, the execution of virtual network services can remain stable even if the underlying infrastructure is changing.

[0049] Fig. 5 is a block diagram of an arrangement of the controller 108 according to some implementations. The controller 108 can include one or multiple processors 502, which can be coupled to one or multiple network interfaces 504 (to allow the controller 108 to communicate over a network), and to a non-transitory machine-readable or computer-readable storage medium 506 (or multiple storage media). The storage medium or storage media 506 can store the scheduler 110 and the node decommissioner 116 in the form of machine-readable instructions, as well as the node information 112 and agent information 114. The scheduler 110 or node decommissioner 116 can be loaded from the storage medium or storage media 506 for execution on the processor(s) 502. A processor can include a microprocessor, microcontroller, processor module or subsystem, programmable integrated circuit, programmable gate array, or another control or computing device.

[0050] The storage medium (or storage media) 506 can include different forms of memory including semiconductor memory devices such as dynamic or static random access memories (DRAMs or SRAMs), erasable and programmable read-only memories (EPROMs), electrically erasable and programmable read-only memories (EEPROMs) and flash memories; magnetic disks such as fixed, floppy and removable disks; other magnetic media including tape; optical media such as compact disks (CDs) or digital video disks (DVDs); or other types of storage devices. Note that the instructions discussed above can be provided on one computer-readable or machine-readable storage medium, or alternatively, can be provided on multiple computer-readable or machine-readable storage media distributed in a large system having possibly plural nodes. Such computer-readable or machine-readable storage medium or media is (are) considered to be part of an article (or article of manufacture). An article or article of manufacture can refer to any manufactured single component or multiple components. The storage medium or media can be

located either in the machine running the machine-readable instructions, or located at a remote site from which machine-readable instructions can be downloaded over a network for execution.

[0051] In the foregoing description, numerous details are set forth to provide an understanding of the subject disclosed herein. However, implementations may be practiced without some of these details. Other implementations may include modifications and variations from the details discussed above. It is intended that the appended claims cover such modifications and variations.

What is claimed is:

1. A method comprising:

detecting, by a controller including a processor, that an agent of a first node managed by the controller is unavailable, the agent providing a service accessible by a tenant of a cloud infrastructure that includes the controller and a plurality of nodes managed by the controller;

in response to the detecting, rescheduling, by the controller, the service on a second node managed by the controller to continue to provide availability of the service to the tenant; and

as part of the rescheduling, cooperating, by the controller, with the first node to avoid duplication of the service on multiple nodes including the first and second nodes.

2. The method of claim 1, wherein avoiding the duplication of the service comprises decommissioning the service on the first node.

3. The method of claim 1, wherein detecting that the agent of the first node is unavailable comprises determining that a message has not been received from the first node for greater than a specified time period.

4. The method of claim 1, wherein the agent of the first node is unavailable due to decommissioning of the first node.

5. The method of claim 4, further comprising:

in response to a notification of decommissioning of the first node,

removing agents on the first node from a pool of available agents; and

notifying the first node to go offline.

6. The method of claim 5, further comprising:

in response to the notification of decommissioning of the first node,

removing information pertaining to the first node from information maintained by the controller; and

triggering the rescheduling in response to removing the information pertaining to the first node.

7. The method of claim 1, wherein the rescheduling provides seamless availability of the service to the tenant such that the tenant is not aware of a temporary unavailability of the service due to the agent being unavailable.

8. The method of claim 1, wherein the service provided by the agent comprises a virtual network service for use in a network of a tenant system.

9. The method of claim 1, further comprising:

storing, by the controller, time information relating to when a given node was decommissioned; and

using, by the controller, the time information to prevent the given node from rejoining the controller in a capacity that allows services to be scheduled on the given node.

10. A system comprising:

a plurality of managed nodes; and

a controller comprising at least one processor to:

manage the plurality of managed nodes that include agents providing network services in a cloud infrastructure, the network services useable in networks of tenants of the cloud infrastructure;

detect that an agent of a first of the plurality of managed nodes is unavailable;

in response to the detecting, reschedule the service on a second of the plurality of managed nodes managed by the controller to continue to provide availability of the service to a tenant; and

wherein the first managed node is to decommission the service on the first managed node to avoid duplication of the service on multiple managed nodes.

11. The system of claim 10, wherein the controller is to further rebalance services across the plurality of managed nodes.

12. The system of claim 10, wherein the controller is to reschedule services onto particular managed nodes that have rejoined the controller after the particular managed nodes were previously removed.

13. The system of claim 10, wherein the controller is to further:

receive a notification that the first managed node is to go offline; and

in response to the notification, remove information of the first managed node and information of agents on the first managed node from the controller.

14. The system of claim 13, wherein the controller is to further:

store time information regarding when the first managed node was removed;

in response to receiving, from the first managed node, a request to rejoin the controller, use the time information to determine an elapsed time since the first managed node was removed; and

decide to grant or deny the request to rejoin based on the determined elapsed time.

15. An article comprising at least one non-transitory machine-readable storage medium storing instructions that upon execution cause a controller to:

detect that an agent of a first node of a plurality of nodes managed by the controller is unavailable, the agent providing a service accessible by a tenant of a cloud infrastructure that includes the controller and the plurality of nodes;

in response to the detecting, reschedule the service on a second of the plurality of nodes to provide seamless availability of the service to the tenant; and

as part of the rescheduling, cooperate with the first node to avoid duplication of the service on multiple nodes including the first and second nodes.

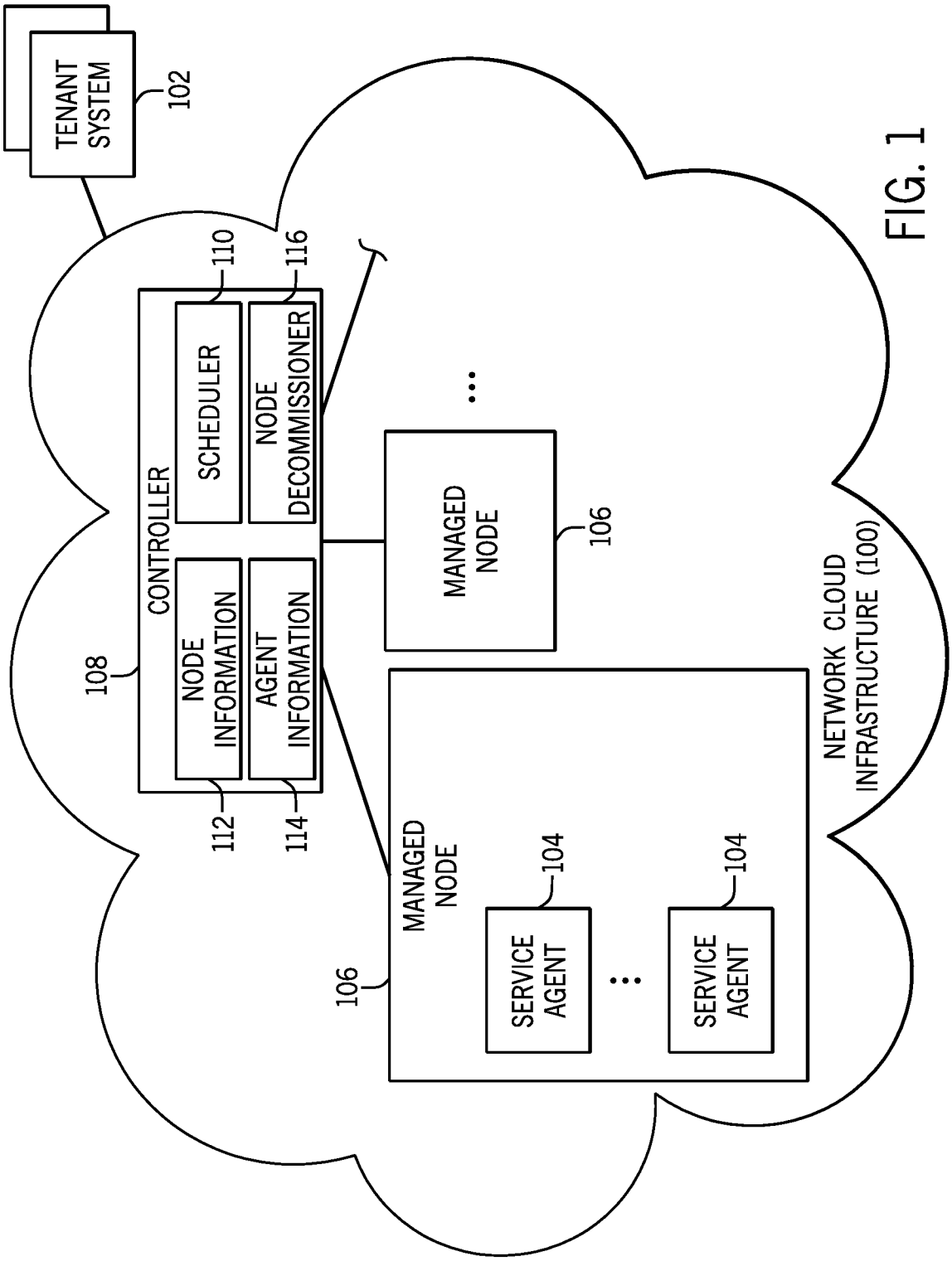
TENANT SYSTEM
102
NETWORK CLOUD INFRASTRUCTURE (100)
CONTROLLER
108
SCHEDULER
110
NODE DECOMMISSIONER
116
NODE INFORMATION
112
AGENT INFORMATION
114
MANAGED NODE
106
MANAGED NODE
106
SERVICE AGENT
104
SERVICE AGENT
104

FIG. 1

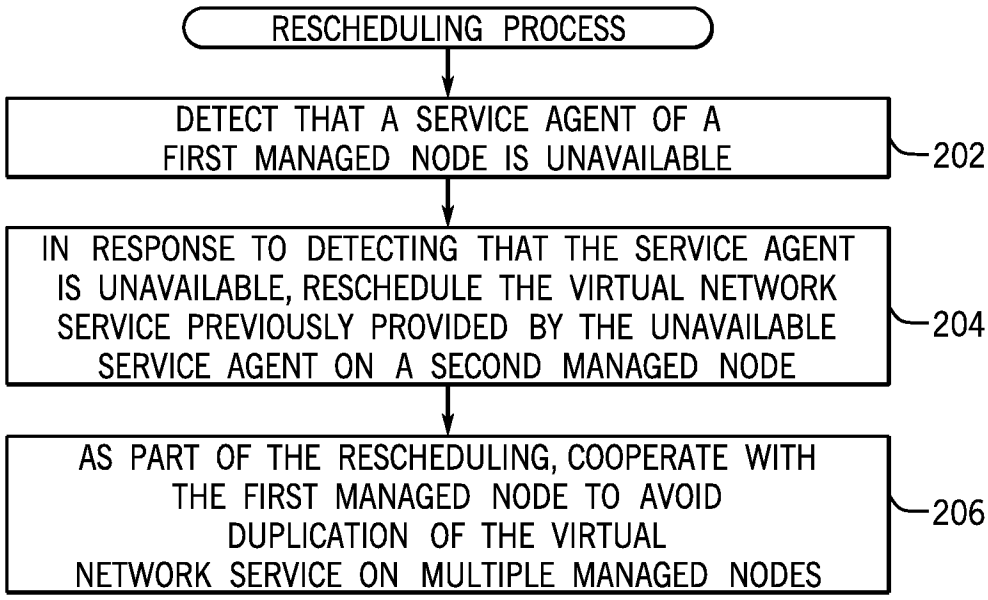
RESCHEDULING PROCESS
DETECT THAT A SERVICE AGENT OF A FIRST MANAGED NODE IS UNAVAILABLE
202
IN RESPONSE TO DETECTING THAT THE SERVICE AGENT IS UNAVAILABLE, RESCHEDULE THE VIRTUAL NETWORK SERVICE PREVIOUSLY PROVIDED BY THE UNAVAILABLE SERVICE AGENT ON A SECOND MANAGED NODE
204
AS PART OF THE RESCHEDULING, COOPERATE WITH THE FIRST MANAGED NODE TO AVOID DUPLICATION OF THE VIRTUAL NETWORK SERVICE ON MULTIPLE MANAGED NODES
206

FIG. 2

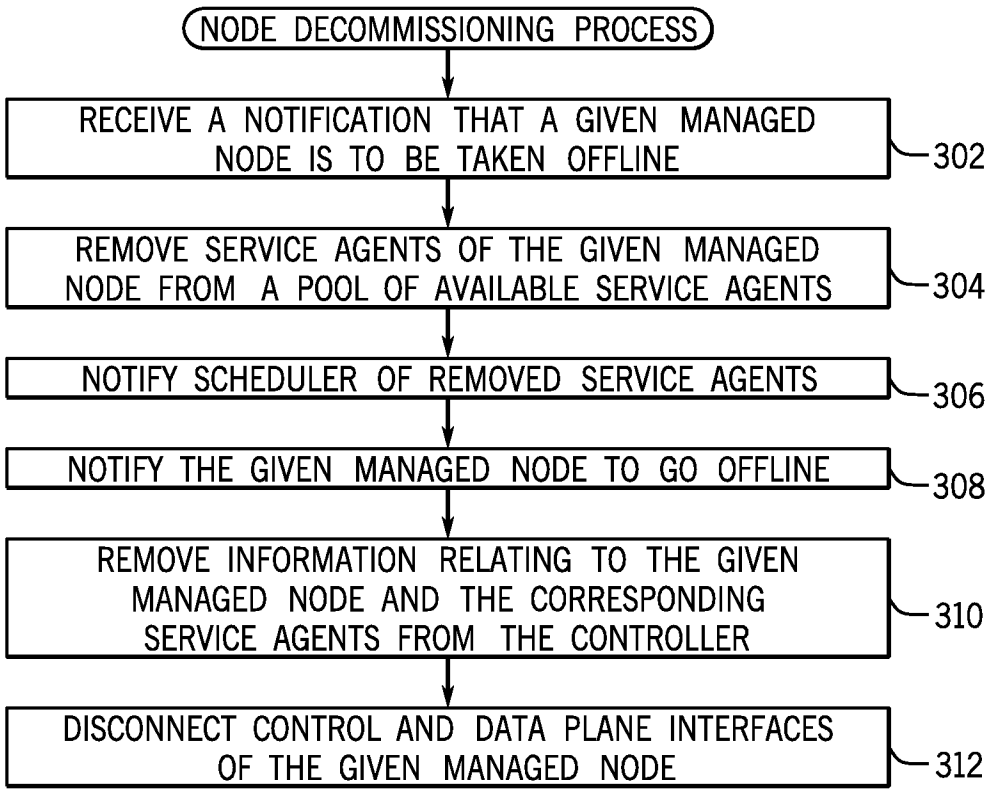
NODE DECOMMISSIONING PROCESS
RECEIVE A NOTIFICATION THAT A GIVEN MANAGED NODE IS TO BE TAKEN OFFLINE
302
REMOVE SERVICE AGENTS OF THE GIVEN MANAGED NODE FROM A POOL OF AVAILABLE SERVICE AGENTS
304
NOTIFY SCHEDULER OF REMOVED SERVICE AGENTS
306
NOTIFY THE GIVEN MANAGED NODE TO GO OFFLINE
308
REMOVE INFORMATION RELATING TO THE GIVEN MANAGED NODE AND THE CORRESPONDING SERVICE AGENTS FROM THE CONTROLLER
310
DISCONNECT CONTROL AND DATA PLANE INTERFACES OF THE GIVEN MANAGED NODE
312

FIG. 3

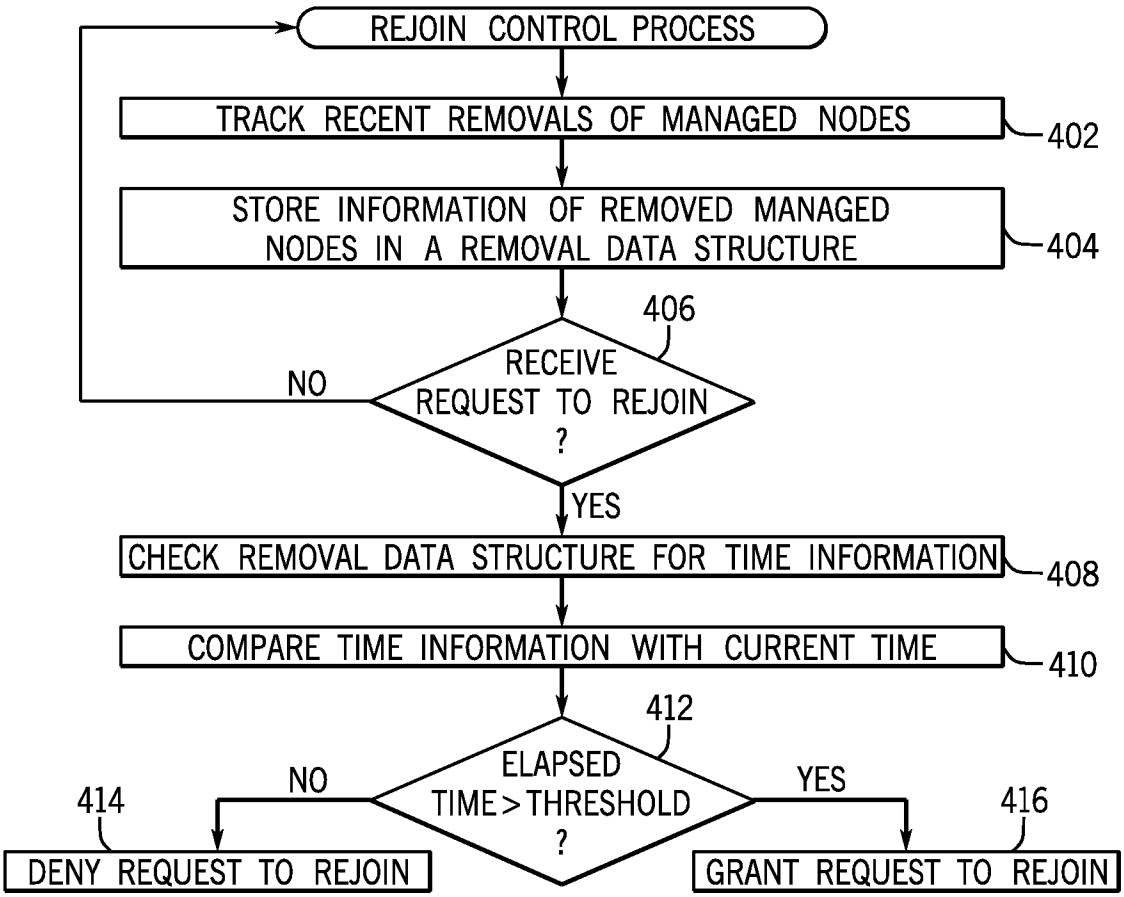
REJOIN CONTROL PROCESS
TRACK RECENT REMOVALS OF MANAGED NODES
402
STORE INFORMATION OF REMOVED MANAGED NODES IN A REMOVAL DATA STRUCTURE
404
406
RECEIVE REQUEST TO REJOIN ?
NO
YES
CHECK REMOVAL DATA STRUCTURE FOR TIME INFORMATION
408
COMPARE TIME INFORMATION WITH CURRENT TIME
410
412
ELAPSED TIME > THRESHOLD ?
NO
YES
414
DENY REQUEST TO REJOIN
416
GRANT REQUEST TO REJOIN

FIG. 4

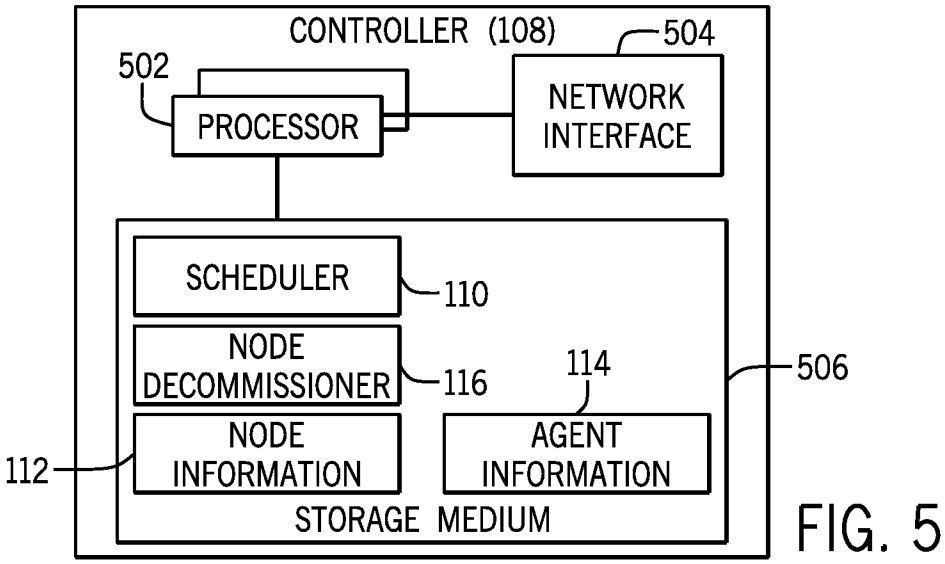
CONTROLLER (108)
504
502
PROCESSOR
NETWORK INTERFACE
SCHEDULER
110
NODE DECOMMISSIONER
116
114
506
NODE INFORMATION
112
AGENT INFORMATION
STORAGE MEDIUM

FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2014/032155

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/24(2006.01)i, H04L 12/16(2006.01)i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04L 12/24; G06F 15/16; H04J 1/16; H04W 40/12; H04J 3/22; H04Q 7/00; H04W 40/10; G06F 15/177; H04L 12/16

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Korean utility models and applications for utility models
Japanese utility models and applications for utility models

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
eKOMPASS(KIPO internal) & Keywords: network, controller, agent, cloud, rescheduling, duplication, time

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 02-23362 A1 (NETMOTION WIRELESS, INC.) 21 March 2002 See page 4, lines 11-22; page 35, lines 1-14; and page 42, line 23 - page 43, line 11; and figures 1-2.	1-15
A	KR 10-2013-0142426 A (KT CORPORATION et al.) 30 December 2013 See paragraphs [0008], [0010], [0103]-[0105]; and figure 8.	1-15
A	US 2009-0319647 A1 (CHRISTOPHER DYSON WHITE et al.) 24 December 2009 See paragraphs [0009]-[0010], [0675]-[0676]; and figure 35.	1-15
A	US 2006-0045005 A1 (ROBERT S. BLACKMORE et al.) 02 March 2006 See paragraphs [0263]-[0269]; and figures 1-2.	1-15
A	US 2002-0159410 A1 (JOSEPH P. ODENWALDER et al.) 31 October 2002 See paragraphs [0060]-[0070]; and figures 5-6.	1-15

☐ 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
10 December 2014 (10.12.2014)	**11 December 2014 (11.12.2014)**

Name and mailing address of the ISA/KR
International Application Division
Korean Intellectual Property Office
189 Cheongsa-ro, Seo-gu, Daejeon Metropolitan City, 302-701,
Republic of Korea
Facsimile No. +82-42-472-7140

Authorized officer
KIM, Seong Woo
Telephone No. +82-42-481-3348

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/US2014/032155

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 02-23362 A1	21/03/2002	AU 2001-89010 A1	26/03/2002
		AU 2003-205094 A1	30/07/2003
		AU 8901001 A	26/03/2002
		CA 2421609 A1	21/03/2002
		CA 2474089 A1	24/07/2003
		CA 2570093 A1	29/12/2005
		EP 1364296 A1	26/11/2003
		EP 1364296 A4	15/09/2004
		EP 1466434 A1	13/10/2004
		EP 1767024 A2	28/03/2007
		JP 2004-509539 A	25/03/2004
		JP 2005-515700 A	26/05/2005
		JP 2008-508837 A	21/03/2008
		US 2002-0098840 A1	25/07/2002
		US 2003-0120811 A1	26/06/2003
		US 2003-0182431 A1	25/09/2003
		US 2005-0223114 A1	06/10/2005
		US 2005-0223115 A1	06/10/2005
		US 2006-0009213 A1	12/01/2006
		US 2006-0123079 A1	08/06/2006
		US 2007-0038759 A1	15/02/2007
		US 2007-0265000 A1	15/11/2007
		US 6546425 B1	08/04/2003
		US 6981047 B2	27/12/2005
		US 7136645 B2	14/11/2006
		US 7293107 B1	06/11/2007
		US 7574208 B2	11/08/2009
		US 7644171 B2	05/01/2010
		US 7778260 B2	17/08/2010
		US 7882247 B2	01/02/2011
		US 8060656 B2	15/11/2011
		US 8078727 B2	13/12/2011
		WO 2003-061188 A1	24/07/2003
		WO 2005-125235 A2	29/12/2005
		WO 2005-125235 A3	27/07/2006
KR 10-2013-0142426 A	30/12/2013	KR 10-1386835 B1	21/04/2014
US 2009-0319647 A1	24/12/2009	CA 2728303 A1	23/12/2009
		EP 2307956 A2	13/04/2011
		EP 2307956 A4	19/12/2012
		US 2009-0319247 A1	24/12/2009
		US 2009-0319248 A1	24/12/2009
		US 2009-0319249 A1	24/12/2009
		US 2009-0319906 A1	24/12/2009
		US 2009-0320137 A1	24/12/2009
		US 2014-046644 A1	13/02/2014
		US 8532970 B2	10/09/2013
		US 8751629 B2	10/06/2014

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/US2014/032155

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		WO 2009-155469 A2	23/12/2009
		WO 2009-155469 A3	06/05/2010
US 2006-0045005 A1	02/03/2006	US 8023417 B2	20/09/2011
US 2002-0159410 A1	31/10/2002	BR 0209169 A	23/05/2006
		CN 100340091 C	26/09/2007
		CN 1505889 A	16/06/2004
		EP 1382166 A1	21/01/2004
		HK 1063700 A1	29/02/2008
		JP 2004-536496 A	02/12/2004
		JP 4409176 B2	03/02/2010
		KR 10-0891850 B1	07/04/2009
		KR 10-2004-0015140 A	18/02/2004
		TW 567731 A	21/12/2003
		TW 567731 B	21/12/2003
		US 6625172 B2	23/09/2003
		WO 2002-089432 A1	07/11/2002