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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 August 2013 (22.08.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/122597 A1

(51) International Patent Classification:
*G06F 9/46* (2006.01)

(21) International Application Number:
PCT/US2012/025577

(22) International Filing Date:
17 February 2012 (17.02.2012)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **AFFIRMED NETWORKS, INC.** [US/US]; 35 Nagog Park. 1st Floor, Acton, MA 01720 (US).

(72) Inventors: **LOGAN, James, J.**; 2 Adams Circle, Southborough, MA 01772 (US). **GARG, Deepak**; 56 Stillwater Drive, Nashua, NH 03062 (US). **MORTSOLF, Tim**; 239 Stanley Street, Amherst, MA 01002 (US).

(74) Agent: **HOBGOOD, John V.**; Wilmer Cutler Pickering Hale And Dorr LLP, 60 State Street, Boston, MA 02109 (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, 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, QA, RO, RS, RU, RW, 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, 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,

*[Continued on next page]*

(54) Title: VIRTUALIZED OPEN WIRELESS SERVICES SOFTWARE ARCHITECTURE

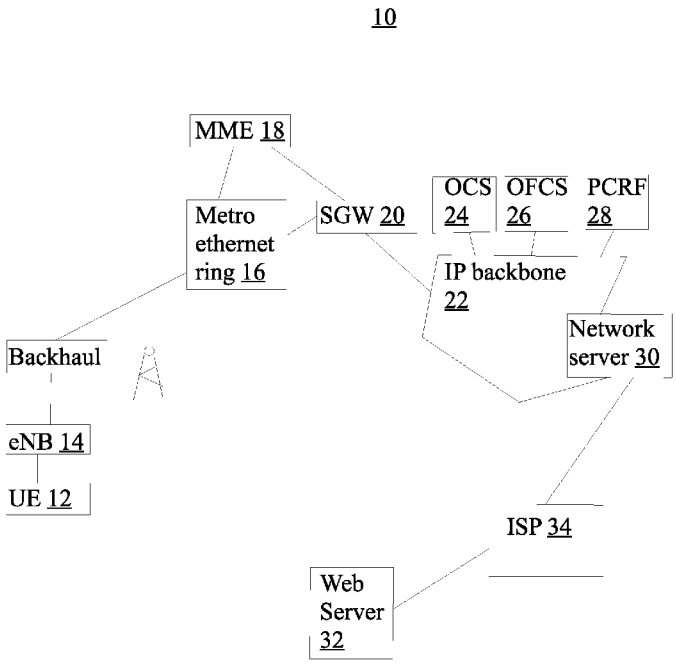



FIG. 1

(57) Abstract: Methods and apparatus, including computer program products, for virtualized open wireless services software architecture. A method includes, in a mobile network, monitoring hardware resources for utilization, receiving a request to instantiate a service, selecting a hardware resource that is underutilized, and mapping the service to the underutilized hardware resource.

WO 2013/122597 A1

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, ML, MR, NE, SN, TD, TG).

**Published**:

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

# VIRTUALIZED OPEN WIRELESS SERVICES SOFTWARE ARCHITECTURE

## BACKGROUND OF THE INVENTION

[001] The invention generally relates to wireless networks, and more specifically to a virtualized open wireless services software architecture.

[002] Wireless networks include many different functions that are typically mapped to different devices. For example, a 3G/4G network includes base stations, Radio Network Controllers (RNCs), Serving GPRS Support Nodes (SGSNs), Gateway GPRS Support Nodes (GGSNs), Serving Gateways (SGWs), Packet Data Network Gateways (PGWs), Mobile Management Entities (MMEs), Digital Phone Interface (DPI) probes, traffic analyzers, firewalls, Policy and Charging Rules Functions (PCRFs), Authentication, Authorization, and Accounting (AAA) servers, Lightweight Directory Access Protocol (LDAP) databases, charging servers, Hypertext Transfer Protocol (HTTP) proxies, traffic optimizers, video optimizers, and so forth. As these wireless networks have evolved, each additional function that is added has required the introduction of a new device that hosts that function or multiple hardware processors in the same device hosting different functions.

## SUMMARY OF THE INVENTION

[003] The following presents a simplified summary of the innovation in order to provide a basic understanding of some aspects of the invention. This summary is not an extensive overview of the invention. It is intended to neither identify key or critical elements of the invention nor delineate the scope of the invention. Its sole purpose is to present some concepts of the invention in a simplified form as a prelude to the more detailed description that is presented later.

[004] The present invention provides methods and apparatus, including computer program products, for a virtualized open wireless services software architecture.

[005] In general, in one aspect, the invention features a method including, in a mobile network, monitoring hardware resources for utilization, receiving a request to instantiate a

service, selecting a hardware resource that is underutilized, and mapping the service to the underutilized hardware resource.

[006] On another aspect, the invention features a server including central processing units, and a memory, the memory including an operating system and grid resource management process, the grid resource management process including monitoring hardware resources in a mobile network for utilization, receiving a request to instantiate a service, selecting a hardware resource that is underutilized, and mapping the service to the underutilized hardware resource.

[007] Other features and advantages of the invention are apparent from the following description, and from the claims.

## BRIEF DESCRIPTION OF THE DRAWINGS

[008] The invention will be more fully understood by reference to the detailed description, in conjunction with the following figures, wherein:

[009] FIG. 1 is a block diagram.

[0010] FIG. 2 is a block diagram.

[0011] FIG. 3 is a flow diagram.

## DETAILED DESCRIPTION

[0012] The subject innovation is now described with reference to the drawings, wherein like reference numerals are used to refer to like elements throughout. In the following description, for purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the present invention. It may be evident, however, that the present invention may be practiced without these specific details. In other instances, well-known structures and devices are shown in block diagram form in order to facilitate describing the present invention.

[0013] As used in this application, the terms “component,” “system,” “platform,” and the like can refer to a computer-related entity or an entity related to an operational machine with one or more specific functionalities. The entities disclosed herein can be either hardware, a combination of hardware and software, software, or software in execution. For example, a

component may be, but is not limited to being, a process running on a processor, a processor, an object, an executable, a thread of execution, a program, and/or a computer. By way of illustration, both an application running on a server and the server can be a component. One or more components may reside within a process and/or thread of execution and a component may be localized on one computer and/or distributed between two or more computers. Also, these components can execute from various computer readable media having various data structures stored thereon. The components may communicate via local and/or remote processes such as in accordance with a signal having one or more data packets (e.g., data from one component interacting with another component in a local system, distributed system, and/or across a network such as the Internet with other systems via the signal).

[0014] In addition, the term "or" is intended to mean an inclusive "or" rather than an exclusive "or." That is, unless specified otherwise, or clear from context, "X employs A or B" is intended to mean any of the natural inclusive permutations. That is, if X employs A; X employs B; or X employs both A and B, then "X employs A or B" is satisfied under any of the foregoing instances. Moreover, articles "a" and "an" as used in the subject specification and annexed drawings should generally be construed to mean "one or more" unless specified otherwise or clear from context to be directed to a singular form.

[0015] Moreover, terms like "user equipment," "mobile station," "mobile," "subscriber station," "communication device," "access terminal," "terminal," "handset," and similar terminology, refer to a wireless device (e.g., cellular phone, smart phone, computer, personal digital assistant (PDA), set-top box, Internet Protocol Television (IPTV), electronic gaming device, printer, etc.) utilized by a subscriber or user of a wireless communication service to receive or convey data, control, voice, video, sound, gaming, or substantially any data-stream or signaling-stream. The foregoing terms are utilized interchangeably in the subject specification and related drawings. Likewise, the terms "access point," "base station," "Node B," "evolved Node B," "home Node B (HNB)," and the like, are utilized interchangeably in the subject application, and refer to a wireless network component or appliance that serves and receives data, control, voice, video, sound, gaming, or substantially any data-stream or signaling-stream from a set of subscriber stations. Data and signaling streams can be packetized or frame-based flows.

[0016] Furthermore, the terms "user," "subscriber," "customer," and the like are employed interchangeably throughout the subject specification, unless context warrants particular distinction(s) among the terms.

[0017] In a distributed, high availability environment, where each hardware node in a network runs a mix of active and standby applications, and, where the standby applications do not require the same amount of physical resources as the active applications, the requirements for physical resources (e.g. CPU, memory, disk space, bandwidth) is not usually taken into consideration when assigning applications to run on a hardware node. The standby applications are backing up active applications distributed across the collection of hardware, and, at any given time only a subset of the applications may go active. Since this fact is not considered when assigning applications to hardware nodes, there is an inherent underutilization of the resources.

[0018] For example, wireless networks include many different functions which are typically mapped to different devices. A typical 3G/4G network can include base stations (BSs), RNCs, SGSNs, GGSNs, SGWs, PGWs, MMEs, DPI probes, traffic analyzers, firewalls, PCRF, AAA servers, LDAP databases, charging servers, HTTP proxies, traffic optimizers, video optimizers and so forth. As these networks have evolved, each additional function that is added has required the introduction of a new device that hosts that function or multiple hardware processors in the same device hosting different functions. This approach causes some major problems.

[0019] The approach leads to inefficient use of resources. Since the functions are mapped to processing resources in a relatively static manner, this potentially leads to scenarios where some devices are underutilized and some overloaded, with no flexibility to use underutilized processing resources to offload the overload devices.

[0020] The approach leads to scalability issues. As the demand on the network increases, each device has to be scaled up independently to manage the increased load which is sub-optimal from a cost perspective. In addition, as new functions are added, new devices have to be added to the network.

[0021] The approach leads to management issues. The more devices there are, the harder it is to monitor and maintain them. Typically each device has its own management interface,

and integrating them into a single management platform becomes very complex. Alternatively, the operator has to deal with multiple management platforms.

[0022] The approach leads to inflexibility in ordering of functions. Different end-end services may require a different ordering in functions. For example, one service may require charging before the firewall function, and another may require the reverse. Since these functions are on different devices, the ordering is a function of how they are wired up and therefore fairly static. Trying to change the order leads to multiple extra hops for traffic which is highly inefficient as well as increases latency.

[0023] The approach can lead to complexity of service instantiation. Any time a new service is introduced, multiple functions have to be configured to instantiate the end-end service. This is a cumbersome and error prone process since the configuration on multiple devices has to be modified. In addition, an error in configuring one of these devices could impact existing services. Furthermore, the complexity associated with this process means that it takes a lot of effort and a lot of time from when a service is identified to when it is actually executed. As a consequence, most operators only enable services that are very lucrative and long lasting, and miss out on opportunities for services that may be less lucrative and more short term (for example, offering premium access to a sporting event).

[0024] The approach can lead to a closed software environment. The software architecture of the devices that make up the network are designed specifically to support the function they have been built for by the vendors that provide them and are therefore closed in the sense that the operator cannot use the processing resources of these platforms to host nay other functions (related or unrelated to the original function) that they may want to. Even though the operator has bought the processing resources, they are very limited in what these resources can be used for.

[0025] The invention described herein characterizes the amount of physical resources required by applications while in their active and standby roles, and maps the distribution of active/standby processes within the collection of hardware nodes. This results in maximized hardware utilization by strategically assigning applications to run on hardware nodes based on their physical resource requirements and high availability state.

[0026] The invention described herein applies to all wireless networks regardless of access technology, including Global System for Mobile (GSM), Code Division Multiple Access (CDMA), Long Term Evolution (LTE), WiMAX and so forth. For ease of explanation, the invention will is described in a 4G network environment.

[0027] As shown in FIG. 1, an exemplary mobile network 10 includes user equipment (UE) 12, such as a smart phone. Other examples of US 12 include, but are not limited to, cellular phone, computer, personal digital assistant (PDA), set-top box, Internet Protocol Television (IPTV), electronic gaming device, printer, tablet, Wi-Fi Hotspot and so forth. The UE 12 is wirelessly linked to an Evolved Node B (eNB) 14. The eNB 14 is a radio part of a cell site. A single eNB may contain several radio transmitters, receivers, control sections and power supplies. The eNB 14 is backhauled to a metro ethernet ring 16, which includes a Mobility Management Entity (MME) 18 and a Serving Gateway (SGW) 20. Backhaul is a process of transferring packets or communication signals over relatively long distances to a separate location for processing. The SGW 20 routes and forwards user data packets, while also acting as the mobility anchor for a user plane during inter-eNodeB handovers.

[0028] The SGW 20 is linked to an Internet Protocol (IP) backbone 22. The IP backbone 22 includes links to a Online Charging System (OCS) 24, an Offline Charging Subsystem (OFCS) 26 and a Policy Control and Charging Function (PCRF) 28. In general, the OCS 24 is a set of interconnected network elements that enable the identification, rating and posting of charges in real time (or near real time). The OFCS 26 receives charging data in the form of Call Detail Records (CDRs) and Diameter accounting messages from network elements after the subscriber incurs network resource usage.

[0029] The IP backbone 22 includes a network server 30 that implements virtualized open wireless services software architecture for 3G and 4G mobile networks. The network server 30 is linked to a web server 32 through an Internet Service Provider (ISP) 34.

[0030] As shown in FIG. 2, the network server 30 includes one or more a central processing units (CPUs) 50, 52, 54 and a memory 52. Memory 52 includes at least an operating system (OS) 54, such as Linux, and grid resource management process 100. The grid resource management process 100, described below, monitors hardware resources, determines their utilization over time, and maps/re-maps software processes into hardware resources, e.g., the one or more a central processing units (CPUs) 50, 52, 54, according to

their availability and utilizations. In addition, the grid resource management process 100 can monitor utilization of other hardware resources within the exemplary mobile network 10 and map/re-map software processes into these other hardware resources, enabling a virtualized open wireless services software architecture.

[0031] Monitoring the one or more CPUs 50, 52, 54, can be implemented in a number of ways. For example, In the Linux OS, monitoring can be done using "mpstat" and other tools.

[0032] As shown in FIG. 3, the grid resource management process 100 includes monitoring (102) hardware resources residing in a mobile network for utilization.

[0033] The grid resource management process 100 receives (104) a request to instantiate a service.

[0034] Services that the grid resource management service 100 can intelligently map/remap to hardware resources may include a network firewall service, a serving gateway (SGW) service, a packet data network gateway (PGW) service, a HTTP Web proxy service, a video proxy service, a service module, and so forth.

[0035] The SGW service routes and forwards user data packets and acts as an anchor for mobility between LTE and other 3GPP technologies.

[0036] The PGW service provides connectivity from user equipment to external packet data networks by being a point of exit and entry of traffic for the user equipment. The PGW service performs policy enforcement, packet filtering for each user, charging support, lawful interception and packet screening.

[0037] The HTTP Web proxy service acts as an intermediary for requests from clients seeking resources from other servers. The HTTP Web proxy service provides web caching, web translating and web transcoding.

[0038] The video proxy service provides video caching, video translating and video transcoding. In addition, this module provides related services involving content manipulation such as advertisement insertion, splicing content together or rewriting adaptive bit rate manifest files.

[0039] The service module provides a number of services, such as, for example, Radius/Diameter, Policy Personal Eventing Protocol (PEP), Packet Forwarding, Content

Filtering, Session Management, Domain Name System (DNS) Service, Access Control, Packet Inspection, Session Term, IP Transport I/O, a charging function, a policy enforcement function, a traffic steering function, a latency service function, and so forth.

[0040] The charging function refers to an ability to selectively charge mobile subscribers and/or content partners based on a per flow basis by applying differentiated billing plans based on flow characteristics such as volume of packets, amount of time the flow is active, and the application associated with the flow as determined by shallow or deep packet inspection of the packet(s) in the flow or by analyzing the flow heuristics (e.g., signature analysis), or other parameters associated with servicing the flow such as time of day.

[0041] The policy enforcement function refers to an ability to apply subscriber related QoS and gating policies to a subscriber flow on a per-flow basis, where the flow is determined by shallow or deep packet inspection, or heuristic analysis (e.g., signature analysis) of the flow pattern.

[0042] The latency service refers to an ability to measure network round trip latency between any selected subscriber or set of subscribers and the gateway function in the wireless core.

[0043] The grid resource management process 100 selects (106) a hardware resource that is underutilized.

[0044] The grid resource management process 100 maps (108) the service to the underutilized hardware resource.

[0045] The grid resource management process 100 monitors (110) mapped hardware resource.

[0046] The grid resource management process 100 remaps (112) the service to another underutilized hardware resource when the mapped hardware resource is over-utilized.

[0047] The grid resource management process 100 provides an open and modular software architecture that maps wireless network functions to processing resources in a virtualized environment that can be run on any standard processor device like a datacenter solution. The grid resource management process 100 maps functions to available processor resources dynamically. Since the functions are run in a virtual environment, they can be

stitched together in a different order for different services. In addition to implementing standard wireless functions as virtual software modules, the architecture allows the operator to plug in their own modules both for protocol analyzers and for additional functions that they want to host on the platform.

[0048] This architecture addresses the problems raised by outdated methods. Since the functions are virtual and can be mapped to processor resources at a granular level (on a per process basis), the processor is allocated to these functions in small increments of processor capacity. This maximizes the efficiency of processor usage since only the amount needed for a function is allocated.

[0049] Since the allocations of functions to resources are done dynamically in small increments, each function can be scaled as required according to the needs of a particular deployment and dynamically changing network conditions, applications and user behavior. In addition, since the processing resources are generic, it is easy to add additional resources to the platform when needed and allocate these resources to any function based on scaling needs.

[0050] Since multiple functions can be mapped to a single scalable platform with generic resources, a number of functions can be hosted on the same platform. It is much easier to manage a single platform running multiple functions in an integrated manner than multiple platforms running single functions.

[0051] Since the functions are running in a virtualized environment as software process on the same platform, they can be invoked in any order for any flow.

[0052] There can be a single management interface for multiple functions that allows these functions to be stitched together in any order with ease, end to end service instantiation becomes easy. Service instantiation involves a planning exercise, and easy configuration through a service aware configuration interface that allows the operator to specify protocol analyzers for flow classification and a service workflow that is series of functions that are instantiated and stitched together in the desired order. This allows the operator to quickly turn services on and off and take advantage of short term service opportunities.

[0053] Creating a modular and virtualized environment enables a “service platform” where the operator is empowered to plug in their own protocol analyzers and software

functions, and use these plugged in modules to instantiate services using the service workflow mechanism.

[0054] Various implementations of the systems and techniques described here can be realized in digital electronic circuitry, integrated circuitry, specially designed ASICs (application specific integrated circuits), computer hardware, firmware, software, and/or combinations thereof. These various implementations can include implementation in one or more computer programs that are executable and/or interpretable on a programmable system including at least one programmable processor, which may be special or general purpose, coupled to receive data and instructions from, and to transmit data and instructions to, a storage system, at least one input device, and at least one output device.

[0055] These computer programs (also known as programs, software, software applications or code) include machine instructions for a programmable processor, and can be implemented in a high-level procedural and/or object-oriented programming language, and/or in assembly/machine language. As used herein, the terms "machine-readable medium" "computer-readable medium" refers to any computer program product, apparatus and/or device (e.g., magnetic discs, optical disks, memory, Programmable Logic Devices (PLDs)) used to provide machine instructions and/or data to a programmable processor, including a machine-readable medium that receives machine instructions as a machine-readable signal. The term "machine-readable signal" refers to any signal used to provide machine instructions and/or data to a programmable processor.

[0056] To provide for interaction with a user, the systems and techniques described here can be implemented on a computer having a display device (e.g., a CRT (cathode ray tube) or LCD (liquid crystal display) monitor) for displaying information to the user and a keyboard and a pointing device (e.g., a mouse or a trackball) by which the user can provide input to the computer. Other kinds of devices can be used to provide for interaction with a user as well; for example, feedback provided to the user can be any form of sensory feedback (e.g., visual feedback, auditory feedback, or tactile feedback), and input from the user can be received in any form, including acoustic, speech, or tactile input.

[0057] The systems and techniques described here can be implemented in a computing system that includes a back-end component (e.g., a data server), or that includes a middleware component (e.g., an application server), or that includes a front-end component

(e.g., a client computer having a graphical user interface or a web browser through which a user can interact with an implementation of the systems and techniques described here), or any combination of such back-end, middleware, or front-end components. The components of the system can be interconnected by any form or medium of digital data communication (e.g., a communication network). Examples of communication networks include a local area network ("LAN"), a wide area network ("WAN"), and the Internet.

[0058] The computing system can include clients and servers. A client and server are generally remote from each other and typically interact through a communication network. The relationship of client and server arises by virtue of computer programs running on the respective computers and having a client-server relationship to each other.

[0059] The foregoing description does not represent an exhaustive list of all possible implementations consistent with this disclosure or of all possible variations of the implementations described. A number of implementations have been described. Nevertheless, it will be understood that various modifications may be made without departing from the spirit and scope of the systems, devices, methods and techniques described here. For example, various forms of the flows shown above may be used, with steps re-ordered, added, or removed. Accordingly, other implementations are within the scope of the following claims.

[0060] What is claimed is:

1. A method comprising:

in a mobile network, monitoring hardware resources for utilization;

receiving a request to instantiate a service;

selecting a hardware resource that is underutilized; and

mapping the service to the underutilized hardware resource.

2. The method of claim 1 further comprising:

monitoring the mapped hardware resource; and

remapping the service to another underutilized hardware resource when the mapped hardware resource is over-utilized.

3. The method of claim 1 wherein the service is selected from the group consisting of a network firewall service, a serving gateway (SGW) service, a packet data network gateway (PGW) service, a HTTP Web proxy service, a video proxy service, a service module and a custom service.

4. The method of claim 3 wherein the SGW service routes and forwards user data packets and acts as an anchor for mobility between LTE and other 3GPP technologies.

5. The method of claim 3 wherein the PGW service provides connectivity from user equipment to external packet data networks by being a point of exit and entry of traffic for the user equipment.

6. The method of claim 3 wherein the HTTP Web proxy service acts as an intermediary for requests from clients seeking resources from other servers.

7. The method of claim 3 wherein the video proxy service provides video caching, video translating and video transcoding.

8. The method of claim 3 wherein the service module is selected from the group consisting of a Radius/Diameter service, a Policy Personal Eventing Protocol (PEP) service, a Packet Forwarding service, a Content Filtering service, a Session Management service, a Domain

Name System (DNS) service, an Access Control service, a Packet Inspection service, a Session Term service, an IP Transport I/O service, a charging function service, a policy enforcement function service, a traffic steering function service, and a latency service function service.

9. The method of claim 1 wherein the hardware resources are selected from the group consisting of central processor units (CPUs), memory, disk space, and bandwidth.

10. A server comprising:

central processing units; and

a memory, the memory comprising an operating system and grid resource management process, the grid resource management process comprising:

monitoring hardware resources in a mobile network for utilization;

receiving a request to instantiate a service;

selecting a hardware resource that is underutilized; and

mapping the service to the underutilized hardware resource.

11. The server of claim 10 wherein the grid resource management process further comprises:

monitoring the mapped hardware resource; and

remapping the service to another underutilized hardware resource when the mapped hardware resource is over-utilized.

12. The server of claim 10 wherein the service is selected from the group consisting of a network firewall service, a serving gateway (SGW) service, a packet data network gateway (PGW) service, a HTTP Web proxy service, a video proxy service, a service module and a custom service.

13. The server of claim 12 wherein the SGW service routes and forwards user data packets and acts as an anchor for mobility between LTE and other 3GPP technologies.

14. The server of claim 12 wherein the PGW service provides connectivity from user equipment to external packet data networks by being a point of exit and entry of traffic for the user equipment.

15. The server of claim 12 wherein the HTTP Web proxy service acts as an intermediary for requests from clients seeking resources from other servers.

16. The server of claim 12 wherein the video proxy service provides video caching, video translating and video transcoding.

17. The server of claim 12 wherein the service module is selected from the group consisting of a Radius/Diameter service, a Policy Personal Eventing Protocol (PEP) service, a Packet Forwarding service, a Content Filtering service, a Session Management service, a Domain Name System (DNS) service, an Access Control service, a Packet Inspection service, a Session Term service, an IP Transport I/O service, a charging function service, a policy enforcement function service, a traffic steering function service, and a latency service function service.

18. The server of claim 10 wherein the hardware resources are selected from the group consisting of central processor units (CPUs), memory, disk space, and bandwidth.

19. The server of claim 10 wherein the service is selected from standby applications backing up active applications distributed across a collection of hardware.

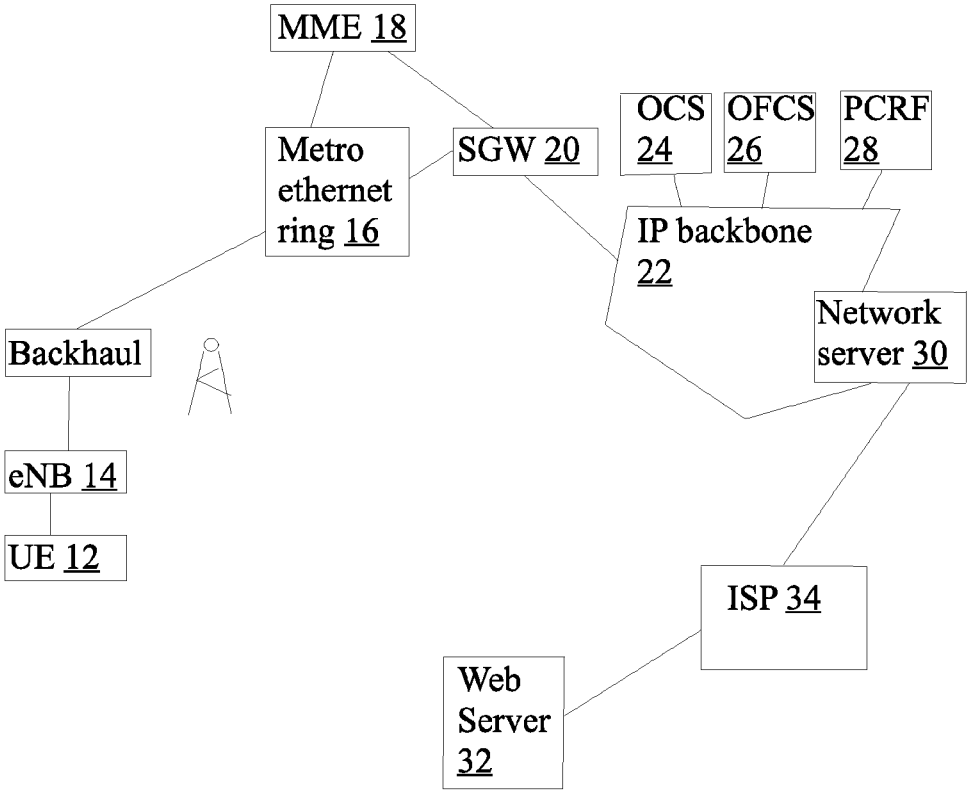

10
MME 18
Metro ethernet ring 16
SGW 20
OCS 24
OFCS 26
PCRF 28
IP backbone 22
Network server 30
Backhaul
eNB 14
UE 12
ISP 34
Web Server 32

FIG. 1

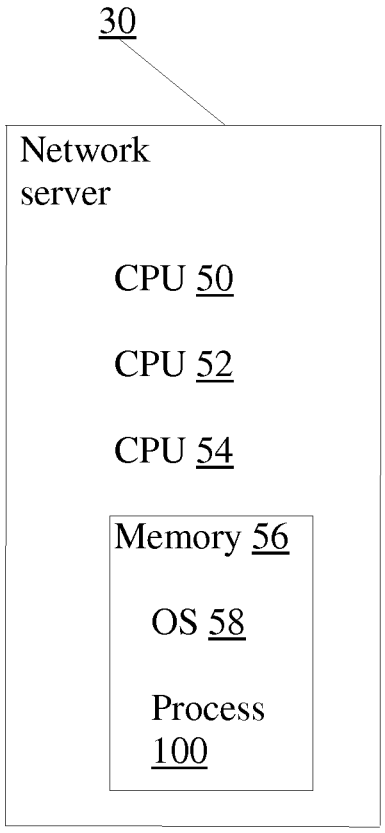

30
Network server
CPU 50
CPU 52
CPU 54
Memory 56
OS 58
Process 100

FIG. 2

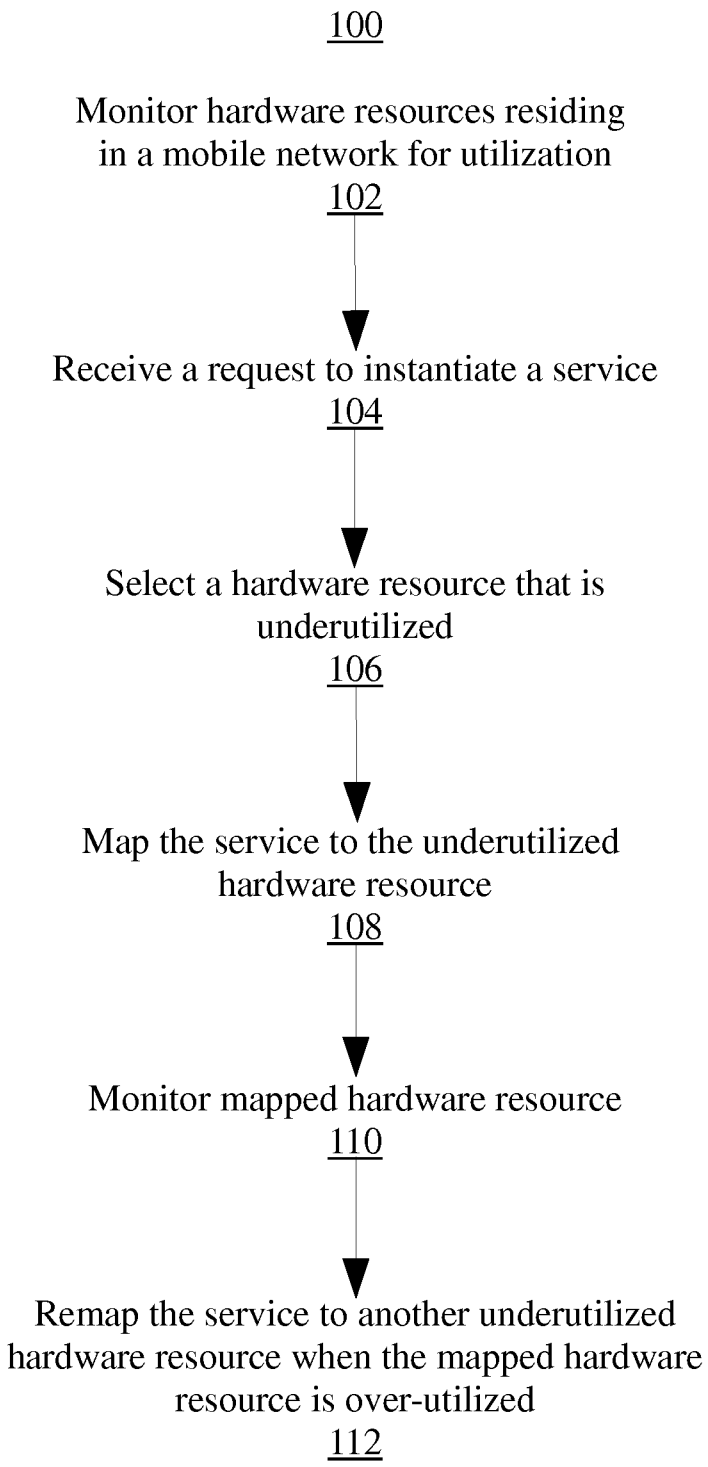

100
Monitor hardware resources residing in a mobile network for utilization
102
Receive a request to instantiate a service
104
Select a hardware resource that is underutilized
106
Map the service to the underutilized hardware resource
108
Monitor mapped hardware resource
110
Remap the service to another underutilized hardware resource when the mapped hardware resource is over-utilized
112

FIG. 3

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 12/25577

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - G06F 9/46 (2012.01)
USPC - 718/105
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)
USPC: 718/105

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC: 455/73; 718/105 (keyword limited - see terms below)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWEST (PGPB, USPT, USOC, EPAB, JPAB); GOOGLE; GoogleScholar
Search Terms: software, hardware, mapping, resource, monitoring, mapping, mobile, wireless, cellular, utilized, under-utilized, grid, packet, virtual, server, proxy, service, packet, HTTW, SGW, PGW, LTE, 3G, session, video, CPU, bandwidth, latency

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| Y | US 2003/0187982 A1 (Petit) 02 October 2003 (02.10.2003), entire document, especially; abstract, para. [0018]-[0020], [0041], [0049], [0078]-[0080] | 1 - 19 |
| Y | US 2006/0294238 A1 (Naik et al.) 28 December 2006 (28.12.2006), entire document, especially; abstract, para. [0010]-[0012], [0056], [0077], [0103]-[0107], [0126] | 1 - 19 |
| A | US 2010/0238840 A1 (Lu et al.) 23 September 2010 (23.09.2010), entire document | 1 - 19 |

☐ Further documents are listed in the continuation of Box C. ☐

* 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 |
|---|---|
| 01 May 2012 (01.05.2012) | 21 MAY 2012 |
| Name and mailing address of the ISA/US<br>Mail Stop PCT, Attn: ISA/US, Commissioner for Patents<br>P.O. Box 1450, Alexandria, Virginia 22313-1450<br>Facsimile No. 571-273-3201 | Authorized officer:<br>Lee W. Young<br>PCT Helpdesk: 571-272-4300<br>PCT OSP: 571-272-7774 |

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

(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 104508633 A
(43)申请公布日 2015.04.08

(21)申请号 201280072000.1

(22)申请日 2012.02.17

(85)PCT国际申请进入国家阶段日
2014.09.28

(86)PCT国际申请的申请数据
PCT/US2012/025577 2012.02.17

(87)PCT国际申请的公布数据
WO2013/122597 EN 2013.08.22

(71)申请人 阿弗梅德网络公司
地址 美国马萨诸塞州

(72)发明人 詹姆斯·J·洛根 迪帕克·加格
蒂姆·莫特索夫

(74)专利代理机构 北京品源专利代理有限公司
11332
代理人 杨生平 钟锦舜

(51)Int.Cl.
*G06F 9/46*(2006.01)

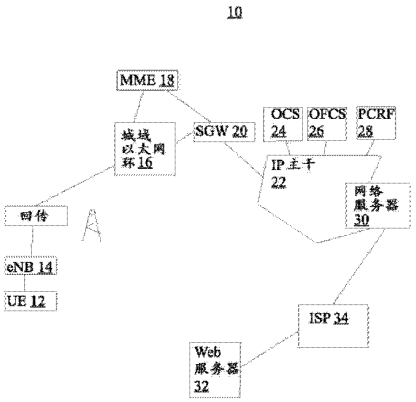
权利要求书2页 说明书6页 附图2页

(54)发明名称

虚拟化开放式无线服务软件体系结构

(57)摘要

用于虚拟化开放式无线服务软件架构的方法和装置，其包括计算机程序产品。一种方法，包括在移动网络中监控供利用的硬件资源、接收实例化服务的请求、选择未充分利用的硬件资源以及将所述服务映射至所述未充分利用的硬件资源。




CN 104508633 A

1. 一种方法，其包括：

在移动网络内监控供利用的硬件资源；

接收实例化服务的请求；

选择未充分利用的硬件资源；以及

将所述服务映射至所述未充分利用的硬件资源。

2. 根据权利要求 1 所述的方法，其还包括：

监控所述被映射的硬件资源；以及

当所述被映射的硬件资源被过度利用时，将所述服务重新映射至另一个未充分利用的硬件资源。

3. 根据权利要求 1 所述的方法，其中所述服务选自由网络防火墙服务、服务网关 (SGW) 服务、分组数据网络网关 (PGW) 服务、HTTP Web 代理服务、视频代理服务、服务模块和定制服务组成的组。

4. 根据权利要求 3 所述的方法，其中所述 SGW 服务路由并转发用户数据分组，并充当 LTE 与其它 3GPP 技术之间的移动锚点。

5. 根据权利要求 3 所述的方法，其中所述 PGW 服务通过充当用户设备的流量出入点来提供从所述用户设备至外部分组数据网络的连接。

6. 根据权利要求 3 所述的方法，其中所述 HTTP Web 代理服务充当来自客户端的从其它服务器查找资源的请求的媒介。

7. 根据权利要求 3 所述的方法，其中所述视频代理服务提供视频高速缓存、视频翻译和视频转码。

8. 根据权利要求 3 所述的方法，其中所述服务模块选自由 Radius/Diameter 服务、策略个人事件协议 (PEP) 服务、分组转发服务、内容过滤服务、会话管理服务、域名系统 (DNS) 服务、接入控制服务、分组检查服务、会话术语服务、IP 传输 I/O 服务、计费功能服务、策略执行功能服务、流量引导功能服务和延时服务功能服务组成的组。

9. 根据权利要求 1 所述的方法，其中所述硬件资源选自由中央处理器 (CPU)、存储器、磁盘空间和带宽组成的组。

10. 一种服务器，其包括：

中央处理器；和

存储器，所述存储器包括操作系统和网格资源管理进程，所述网格资源管理进程包括：

在移动网络中监控供利用的硬件资源；

接收实例化服务的请求；

选择未充分利用的硬件资源；以及

将所述服务映射至所述未充分利用的硬件资源。

11. 根据权利要求 10 所述的服务器，其中所述网格资源管理进程还包括：

监控所述被映射的硬件资源；以及

当所述被映射的硬件资源被过度利用时，将所述服务重新映射至另一个未充分利用的硬件资源。

12. 根据权利要求 10 所述的服务器，其中所述服务选自由网络防火墙服务、服务网关

(SGW)服务、分组数据网络网关(PGW)服务、HTTP Web代理服务、视频代理服务、服务模块和定制服务组成的组。

13.根据权利要求12所述的服务器，其中所述SGW服务路由并转发用户数据分组，并充当LTE与其它3GPP技术之间的移动锚点。

14.根据权利要求12所述的服务器，其中所述PGW服务通过充当用户设备的流量出入点来提供从所述用户设备至外部分组数据网络的连接。

15.根据权利要求12所述的服务器，其中所述HTTP Web代理服务充当来自客户端的从其它服务器查找资源的请求的媒介。

16.根据权利要求12所述的服务器，其中所述视频代理服务提供视频高速缓存、视频翻译和视频转码。

17.根据权利要求12所述的服务器，其中所述服务模块选自由Radius/Diameter服务、策略个人事件协议(PEP)服务、分组转发服务、内容过滤服务、会话管理服务、域名系统(DNS)服务、接入控制服务、分组检查服务、会话术语服务、IP传输I/O服务、计费功能服务、策略执行功能服务、流量引导功能服务和延时服务功能服务组成的组。

18.根据权利要求10所述的服务器，其中所述硬件资源选自由中央处理器(CPU)、存储器、磁盘空间和带宽组成的组。

19.根据权利要求10所述的服务器，其中所述服务选自支持分配给硬件集合的活动应用程序的备用应用程序。

# 虚拟化开放式无线服务软件体系结构

[0001] 发明背景

[0002] 本发明大体涉及无线网络，更具体地涉及虚拟化开放式无线服务软件体系结构。

[0003] 无线网络包括通常映射到不同设备的许多不同功能。例如，3G/4G 网络包括基站、无线电网络控制器 (RNC)、服务 GPRS 支持节点 (SGSN)、网关 GPRS 支持节点 (GGSN)、服务网关 (SGW)、分组数据网络网关 (PGW)、移动管理实体 (MME)、数字电话界面 (DPI) 探针、流量分析器、防火墙、策略和计费规则功能 (PCRF)、认证、授权和计费 (AAA) 服务器、轻量级目录访问协议 (LDAP) 数据库、计费服务器、超文本传输协议 (HTTP) 代理、流量优化器、视频优化器等。随着这些无线网络的发展，每一个增加的附加功能需要引入在托管不同功能的同一设备中托管该功能或多个硬件处理器的新设备。

## 发明概要

[0004] 以下对本发明进行了简单概述以提供对本发明一些方面的基本理解。本概述并非是对本发明的广泛综述。其并非意在确定本发明的关键或重要元素，也并非意在勾画本发明的范围。其唯一目的是以简要形式展示本发明的一些概念，以作为下文所做详细描述的前序。

[0005] 本发明提供了用于虚拟化开放式无线服务软件体系结构的方法和装置，包括计算机程序产品。

[0006] 一般而言，一方面本发明的特征在于一种方法，包括在移动网络中监控供利用的硬件资源、接收实例化服务的请求、选择未充分使用的硬件资源以及将服务映射至未充分使用的硬件资源。

[0007] 另一个方面，本发明的特征在于包括中央处理器和存储器的服务器，存储器包括操作系统和网格资源管理进程，该网格资源管理进程包括在移动网络中监控供利用的硬件资源、接收实例化服务的请求、选择未充分使用的硬件资源以及将服务映射至未充分使用的硬件资源。

[0008] 根据以下描述和权利要求，本发明的其它特征和优点将变得更明显。

[0009] 附图简述

[0010] 通过参考结合附图的详细描述，将更全面地理解本发明，其中：

[0011] 图 1 是框图。

[0012] 图 2 是框图。

[0013] 图 3 是流程图。

## 具体实施方式

[0014] 现在将结合附图描述本发明，其中相同附图标记在所有附图中用于指示相同元件。在以下的描述中，出于说明目的，对很多具体细节进行了阐述，以提供对本发明的全面理解。然而，明显的是，没有这些具体细节也可实践本发明。在其它情况下，为了便于描述本发明，以框图形式示出公知的结构和设备。

[0015] 本申请中使用的术语“组件”、“系统”、“平台”等可指与计算机相关的实体或与具有一个或多个特定功能的操作机器相关的实体。本文公开的实体可以是硬件、硬件和软件的组合、软件，或执行中的软件。例如，组件可以是（但不限于）处理器上运行的进程、处理器、对象、可执行程序、执行线程、程序和/或计算机。作为说明，在服务器上运行的应用程序和服务器本身都可以是组件。一个或多个组件可驻留在进程和/或执行线程内，且组件可在一台计算机上本地化和/或分布在两台或多台计算机之间。这些组件也可从其上存储有各种数据结构的各种计算机可读介质执行。组件可通过本地和/或远程进程通信，例如根据具有一个或多个数据分组（例如，来自一个与本地系统、分布式系统中的另一组件交互、和/或通过诸如因特网等网络与其它系统交互的组件的数据）的信号进行通信。

[0016] 此外，术语“或”意指包括性“或”而非排他性“或”。即，除非另外指明或从上下文可明确，否则“X 使用 A 或 B”意指任何自然的包括性排列。即，如果 X 使用 A；X 使用 B；或 X 使用 A 和 B，则“X 使用 A 或 B”满足上述任何一种实例。此外，除非另外指明或可从上下文清楚得知要指示单数形式，否则本说明书和附图中使用的冠词“a”和“an”通常应解释为表示“一个或多个/一种或多种”。

[0017] 此外，术语如“用户设备”、“移动站”、“移动”、“订户站”、“通信设备”、“接入终端”、“终端”、“手持设备”以及相似术语指的是无线通信服务的订户或用户用于接收或传送数据、控制、语音、视频、声音、游戏或基本上任何数据流或信令流的无线设备（例如，蜂窝电话、智能电话、计算机、个人数字助理（PDA）、机顶盒、互联网协议电视（IPTV）、电子游戏设备、打印机等）。以上术语在说明书和相关附图中可互换使用。同样地，术语“接入点”、“基站”、“节点 B”、“演进节点 B”、“家庭节点 B(HNB)”等在本说明书中可互换使用，且指的是提供服务并从一组订户站接收数据、控制、语音、视频、声音、游戏或基本上任何数据流或信令流的无线网络组件或设备。数据和信令流可以是分组化或基于帧的流。

[0018] 此外，除非上下文特别指出术语之间的区别，否则术语“用户”、“订户”、“客户”等在整个说明书中可互换使用。

[0019] 在分布式高可用性环境中，其中网络中的每一个硬件节点混合运行活动和备用应用程序，且其中备用应用程序所要求的物理资源量与活动应用程序不同，当分派应用程序在硬件节点上运行时，对物理资源（例如，CPU、存储器、磁盘空间、带宽）的要求通常不考虑在内。备用应用程序支持分配给硬件集合的活动应用程序，且在任何给定时间，仅应用程序的子集活动。由于当将应用程序分派至硬件节点时不考虑该情况，因此存在固有的资源利用不充分的情况。

[0020] 例如，无线网络包括通常映射到不同设备的许多不同功能。典型的 3G/4G 网络可包括基站（BS）、RNC、SGSN、GGSN、SGW、PGW、MME、DPI 探针、流量分析器、防火墙、PCRF、AAA 服务器、LDAP 数据库、计费服务器、HTTP 代理、流量优化器、视频优化器等。随着这些网络的演进，每一个增加的附加功能需要引入在托管不同功能的同一设备中托管该功能或多个硬件处理器的新设备。这种方法会造成一些主要问题。

[0021] 该方法导致资源利用率低。由于功能以相对静态的方式被映射到处理资源，因此这潜在地导致一些设备未充分利用而一些设备过载的情形，而不会灵活地使用未充分利用的处理资源来对过载设备进行卸载。

[0022] 该方法导致可扩展性问题。由于对网络的需求增加，因此每一个设备需独立地扩

展以管理增加的负载，从成本角度而言，这并非最佳的。此外，由于增加了新功能，新设备需添加至网络。

[0023] 该方法导致管理问题。设备越多，对其进行监控与维修越难。通常，每一个设备具有各自的管理界面，因此将设备集成到单个管理平台中变得非常复杂。如若不然，运营商需处理多个管理平台。

[0024] 该方法导致功能排序不灵活。不同的端对端服务可要求不同的功能排序。例如，一种服务可要求防火墙功能在计费功能之前，而另一种服务的要求可相反。由于这些功能在不同设备上，因此排序根据如何连接设备变化，因此相当静态。试图改变顺序导致流量发生多个额外跳跃，这是非常低效的且增加了延时。

[0025] 该方法可导致服务实例化复杂。每当引入新服务时，需对多个功能进行配置以实例化端对端服务。由于需对多个设备上的配置进行修改，因此这是一个麻烦且易出错的过程。此外，对这些设备的其中之一进行配置时出现的错误会影响现有服务。此外，与该过程关联的复杂性意味着从识别服务至实际执行服务需要花费更多努力和时间。其结果是，大多数运营商只启用赚钱多且持久的服务，而错过可能赚钱少且时间短的服务（例如提供对体育赛事的高级访问）。

[0026] 该方法可导致封闭式软件环境。组成网络的设备的软件体系结构被设备供应商具体设计成支持构建设备所为的功能，且因此从运营商无法使用这些平台的处理资源来托管其想要托管的任何其它功能（与初始功能相关或不相关的功能）的意义上说是封闭的。即使运营商已购买了处理资源，其在这些资源的用途方面也非常受限。

[0027] 本文描述的本发明的特征在于应用程序在其活动和备用角色时所需要的物理资源的量，且在硬件节点的集合内映射活动 / 备用进程的分布。这通过基于应用程序的物理资源要求和高可用性状态战略性地分派应用程序在硬件节点上运行，使得硬件的利用率被最大化。

[0028] 不论接入技术如何，本文描述的本发明应用于所有无线网络，包括全球移动通信系统 (GSM)、码分多址 (CDMA)、长期演进 (LTE)、WiMAX 等。为了便于解释，将在 4G 网络环境下描述本发明。

[0029] 如图 1 所示，示例性移动网络 10 包括用户设备 (UE)12，例如智能电话。US 12 的其它实例包括，但不限于蜂窝电话、计算机、个人数字助理 (PDA)、机顶盒、互联网协议电视 (IPTV)、电子游戏设备、打印机、平板电脑、Wi-Fi 热点等。UE 12 无线链接至演进节点 B(eNB)14。eNB 14 是小区基站的无线电部分。单个 eNB 可包含几个无线电发送器、接收器、控制部分和电源。eNB 14 回传至城域以太网环 16，城域以太网环 16 包括移动管理实体 (MME)18 和服务网关 (SGW)20。回传是将分组或通信信号长距离传送至分开的地点进行处理的过程。SGW 20 路由并转发用户数据分组，同时在 eNodeB 间切换期间作为用户平面的移动锚点。

[0030] SGW 20 链接至互联网协议 (IP) 主干 22。IP 主干 22 包括至在线计费系统 (OCS)24、离线计费子系统 (OFCS)26 和政策控制和计费功能 (PCRF)28 的链路。一般而言，OCS 24 为一组相互连接的使得能够实时（或近乎实时）地进行确认、计账以及过账的网络元件。订户使用网络资源后，OFCS 26 从网络元件接收呼叫详细记录 (CDR) 和 Diameter 计费消息形式的计费数据。

[0031] IP 主干 22 包括实施用于 3G 和 4G 移动网络的虚拟化开放式无线服务软件体系结构的网络服务器 30。Web 服务器 32 通过互联网服务提供商 (ISP)34 链接至网络服务器 30。

[0032] 如图 2 所示，网络服务器 30 包括一个或多个中央处理器 (CPU)50、52、54 以及存储器 52。存储器 52 包括至少一个操作系统 (OS)54，例如 Linux，以及网格资源管理进程 100。以下描述的网格资源管理进程 100 监控硬件资源、确定其随着时间的利用情况，并根据其可用性和利用情况将软件进程映射 / 重新映射到硬件资源上，例如一个或多个中央处理器 (CPU)50、52、54。此外，网格资源管理进程 100 可监控其它硬件资源在示例性移动网络 10 内的利用情况，并将软件进程映射 / 重新映射到这些其它硬件资源上，从而实现虚拟化开放式无线服务软件体系结构。

[0033] 可以多种方式对一个或多个 CPU 50、52、54 进行监控。例如，在 Linux OS 中，可使用 "mpstat" 和其它工具进行监控。

[0034] 如图 3 所示，网格资源管理进程 100 包括监控驻留在移动网络中以供利用的硬件资源 (102)。

[0035] 网格资源管理进程 100 接收请求以实例化服务 (104)。

[0036] 网格资源管理服务 100 可智能地映射 / 重新映射至硬件资源的服务可包括网络防火墙服务、服务网关 (SGW) 服务、分组数据网络网关 (PGW) 服务、HTTP Web 代理服务、视频代理服务、服务模块等。

[0037] SGW 服务路由并转发用户数据分组，且充当 LTE 与其它 3GPP 技术之间的移动锚点。

[0038] PGW 服务通过充当用户设备的流量出入点来提供从用户设备至外部分组数据网络的连接。PGW 服务执行策略实施，针对每一个用户的分组过滤，计费支持，合法拦截和分组筛选。

[0039] HTTP Web 代理服务充当来自客户端的从其它服务器查找资源的请求的媒介。在 HTTP Web 代理服务提供了 web 缓存，web 翻译和 web 转码。

[0040] 视频代理服务提供视频高速缓存、视频翻译和视频转码。此外，这种模块提供涉及内容操纵（例如，广告插入、拼接内容或重写自适应比特率清单文件）的相关服务。

[0041] 服务模块提供多种服务，诸如例如 Radius/Diameter、策略个人事件协议 (PEP)、分组转发、内容过滤、会话管理、域名系统 (DNS) 服务、接入控制、分组检查、会话术语、IP 传输 I/O、计费功能、策略执行功能、流量引导功能、延时服务功能等。

[0042] 计费功能指的是基于每个流通过采用基于流特性（例如，分组容量、流活动的时间量、以及通过在流中对分组进行浅度或深度分组检查或通过分析流试探（例如，签名分析）确定的与流关联的应用，或其它与服务流关联的参数（例如，一天中的时间））的有区别的计费模式选择性地对移动订户和 / 或内容合作方进行收费的能力。

[0043] 策略执行功能指的是基于每个流对订户流应用与订户相关的 QoS 和门控策略，其中通过对流模式进行浅度或深度分组检查或试探性分析（例如，签名分析）而确定流。

[0044] 延时服务指的是测量任何选择的订户或一组订户与无线核心中的网关功能之间的网络往返延时的能力。

[0045] 网格资源管理进程 100 选择未充分利用的硬件资源 (106)。

[0046] 网格资源管理进程 100 将服务映射至未充分利用的硬件资源 (108)。

[0047] 网格资源管理进程 100 监控被映射的硬件资源 (110)。

[0048] 当被映射的硬件资源被过度利用时，网格资源管理进程 100 将服务重新映射至另一个未充分利用的硬件资源 (112)。

[0049] 网格资源管理进程 100 提供开放式模块化软件体系结构，其在虚拟化环境下将可在任何标准处理器设备（如数据中心解决方案）上运行的无线网络功能映射至处理资源。网格资源管理进程 100 将功能动态地映射至可用的处理器资源。由于功能在虚拟环境下运行，因此可以不同顺序将功能拼接起来以提供不同服务。除了将标准无线功能作为虚拟软件模块实施外，体系结构还允许运营商插入其各自的模块用于协议分析器以及其想要在平台上托管的附加功能。

[0050] 这种体系结构解决了过时的方法所引起的问题。由于功能是虚拟的且可以细化级别（以每一个进程为基础）被映射至处理器资源，因此处理器以很小的处理器容量增量分配给这些功能。由于仅分配了功能所需的量，因此这使得处理器的使用效率最大化。

[0051] 由于以很小的增量将功能动态地分配给资源，因此可根据特别部署以及动态地改变的网络状况、应用和用户行为的要求，按需要扩缩每一种功能。此外，由于处理资源是通用的，因此很容易在需要时将附加资源添加至平台，并基于扩缩要求将这些资源分配给任何功能。

[0052] 由于多个功能可映射至具有通用资源的单个可扩缩平台，因此许多功能可托管在同一平台上。管理以集成方式运行多个功能的单个平台比管理多个运行单个功能的平台要容易得多。

[0053] 由于功能与软件进程一样在虚拟化环境中的同一平台上运行，因此对于任何流可以任何顺序调用功能。

[0054] 可存在用于多个功能的单个管理界面，该管理界面允许很容易地以任何顺序将这些功能拼接起来，端对端服务实例化变得很容易。服务实例化涉及通过服务感知配置界面进行计划练习和简易配置，服务感知配置界面允许运营商明确用于流分类和服务工作流的协议分析器，所述流分类和服务工作流是被实例化且以所需的顺序被拼接在一起的一系列功能。这允许运营商快速开启并关掉服务，并利用短期服务机会。

[0055] 创建模块化虚拟化的环境使得能够实现“服务平台”，其中平台运营商被允许插入其各自的协议分析器和软件功能，并使用这些插入的模块来实例化使用服务工作流机制的服务。

[0056] 文中描述的系统和技术的各种实现可在数字电子电路、集成电路、特别设计的 ASIC（专用集成电路）、计算机硬件、固件、软件和 / 或其组合中实现。这些不同的实现可包括在可编程系统上可执行和 / 或可解释的一个或多个计算机程序中的实现，该可编程系统包括至少一个可编程处理器，该至少一个可编程处理器可以是专用或通用的，并且连接至存储系统、至少一个输入设备和至少一个输出设备，以便从其接收数据和指令，并向其发送数据和指令。

[0057] 这些计算机程序（也称作程序、软件、软件应用程序或代码）包括用于可编程处理器的计算机指令，且可以高级程序和 / 或面向对象的编程语言、和 / 或汇编 / 机器语言来实现。如文中所用，术语“机器可读介质”、“计算机可读介质”是指用于向可编程处理器提供机器指令和 / 或数据的任何计算机程序产品、装置和 / 或设备（例如，磁盘、光盘、存储器、可编程逻辑设备 (PLD)），包括接收机器指令作为机器可读信号的机器可读介质。术语“机

器可读信号”是指任何用于向可编程处理器提供机器指令和 / 或数据的信号。

[0058] 为了提供与用户的交互，文中描述的系统和技术可在具有显示设备（例如，CRT（阴极射线管）监视器或 LCD（液晶显示器）监视器）和键盘以及定点设备（例如，鼠标或跟踪球）的计算机上实现，其中显示器用于向用户显示信息，通过键盘和定点设备用户可向计算机输入。其它种类的设备也可用于提供与用户的交互；例如，提供给用户的反馈可以是任何形式的感知反馈（例如，视觉反馈、听觉反馈或触觉反馈），并且可以任何形式接收来自用户的输入，包括声音、语音或触觉输入。

[0059] 文中描述的系统和技术可在以下计算系统中实现，该计算系统包括后端组件（例如，数据服务器），或包括中间件组件（例如，应用程序服务器），或包括前端组件（例如，具有图形用户界面和网络浏览器的客户端计算机，通过该图形用户界面和网络浏览器用户可与文中描述的系统和技术的实现进行交互），或这种后端、中间件或前端组件的任意组合。系统组件可以采用任何形式或介质的数字数据通信（例如，通信网络）来实现互连。通信网络的实例包括局域网（"LAN"）、广域网（"WAN"）和互联网。

[0060] 计算系统可包括客户端和服务器。客户端与服务器一般彼此远离，且通常通过通信网络交互。客户端与服务器的关系依靠在各个计算机上运行且彼此具有客户端服务器关系的计算机程序而产生。

[0061] 以上描述并不代表与本公开相一致的所有可能的实现或描述的实现的所有可能变化的详尽列表。已描述了许多实现。然而，应理解，在不脱离文中描述的系统、设备、方法和技术的精神和范围的情况下可进行各种修改。例如，可使用以上示出的各种形态的流程，可对其步骤进行重新排序，增加或去除。相应地，其它实现在以下权利要求书的范围内。

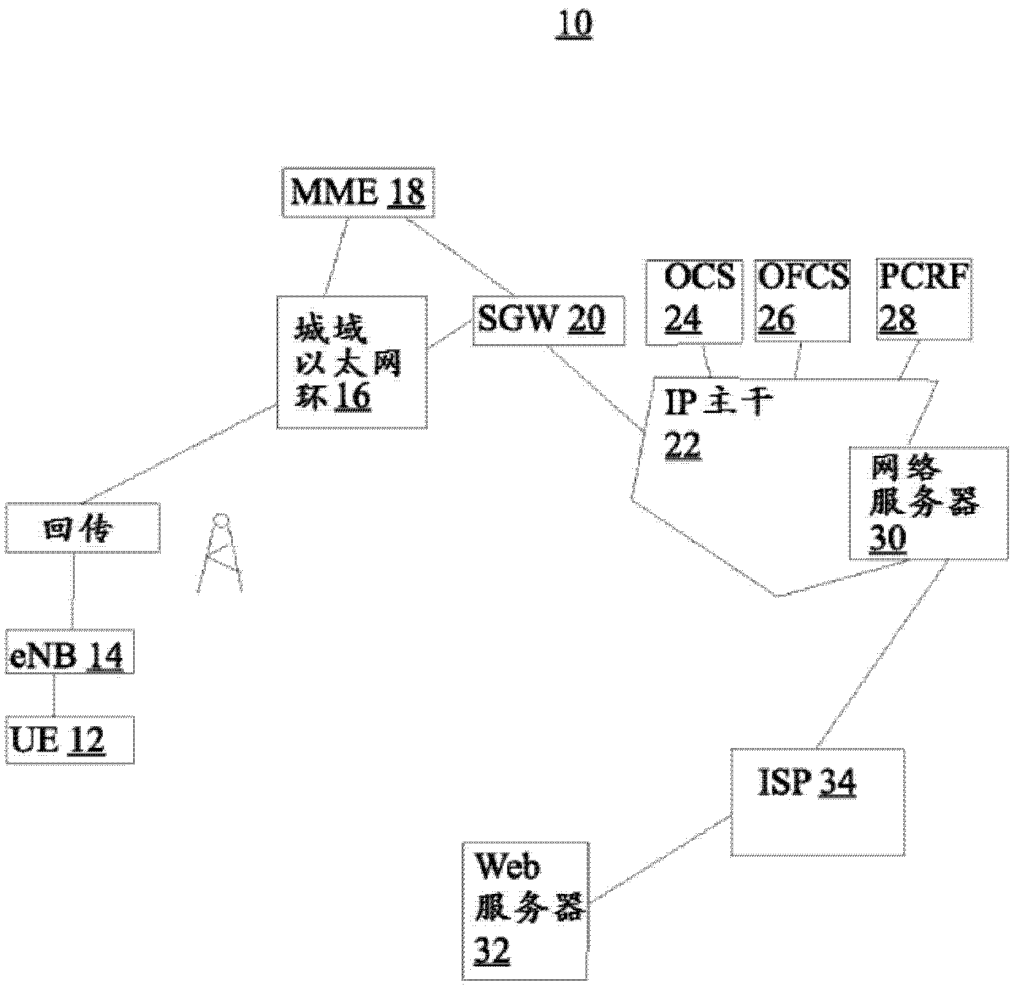

10
MME 18
城域以太网环 16
SGW 20
OCS 24
OFCS 26
PCRF 28
IP 主干 22
网络服务器 30
回传
eNB 14
UE 12
ISP 34
Web 服务器 32

图 1

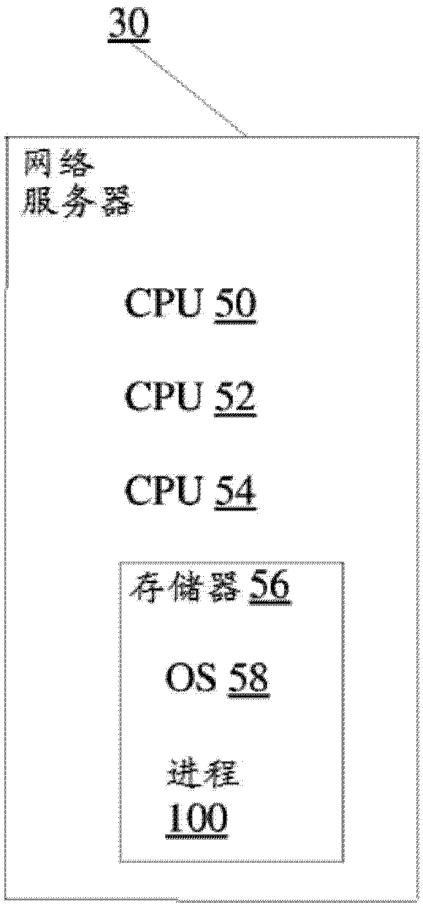

30
网络
服务器
CPU 50
CPU 52
CPU 54
存储器 56
OS 58
进程
100

图 2

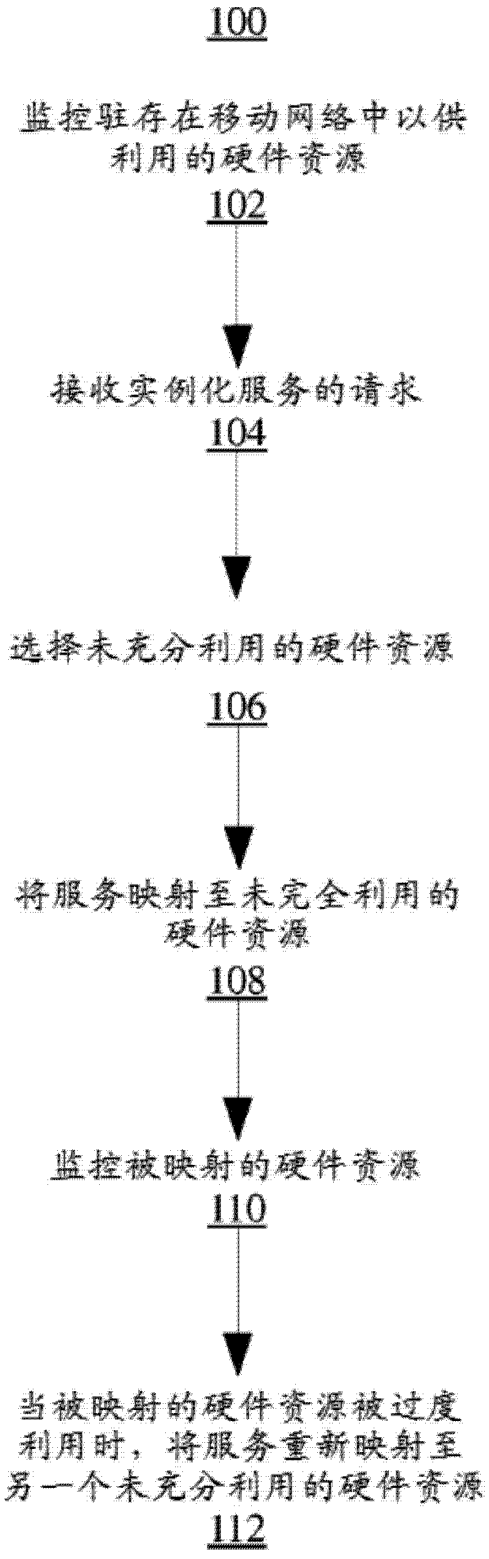

100
监控驻存在移动网络中以供利用的硬件资源
102
接收实例化服务的请求
104
选择未充分利用的硬件资源
106
将服务映射至未完全利用的硬件资源
108
监控被映射的硬件资源
110
当被映射的硬件资源被过度利用时，将服务重新映射至另一个未充分利用的硬件资源
112

图 3

# Abstract

Methods and apparatus, including computer program products, for virtualized open wireless services software architecture. A method includes, in a mobile network, monitoring hardware resources for utilization, receiving a request to instantiate a service, selecting a hardware resource that is underutilized, and mapping the service to the underutilized hardware resource.