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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 October 2006 (12.10.2006)

PCT

(10) International Publication Number
WO 2006/106142 A1

(51) International Patent Classification:
G06F 9/50 (2006.01)

(21) International Application Number:
PCT/EP2006/061434

(22) International Filing Date: 7 April 2006 (07.04.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/101,609 7 April 2005 (07.04.2005) US

(71) Applicant *(for all designated States except US)*: **INTERNATIONAL BUSINESS MACHINES CORPORATION** [US/US]; New Orchard Road, Armonk, New York 10504 (US).

(71) Applicant *(for MG only)*: **IBM UNITED KINGDOM LIMITED** [GB/GB]; Po Box 41, Portsmouth Hampshire PO6 3AU (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CHILDRESS, Rhonda** [US/US]; 8515 West Potosi Cove, Austin, Texas 78717 (US). **CRAWFORD, Catherine, Helen** [US/US]; 123 Hemlock Terrace, Carmel, New York 10512 (US). **KUMHYR, David, Bruce** [US/US]; 8934 Appaloosa Run, Austin, Texas 78737 (US). **MAGNONE, Paolo, Franco** [US/US]; 124 Princeton Place, Palisades Park, New Jersey 07650 (US). **PENNELL, Neil** [GB/US]; 106 Colorado Cove, Cedar Creek, Texas 78612 (US).

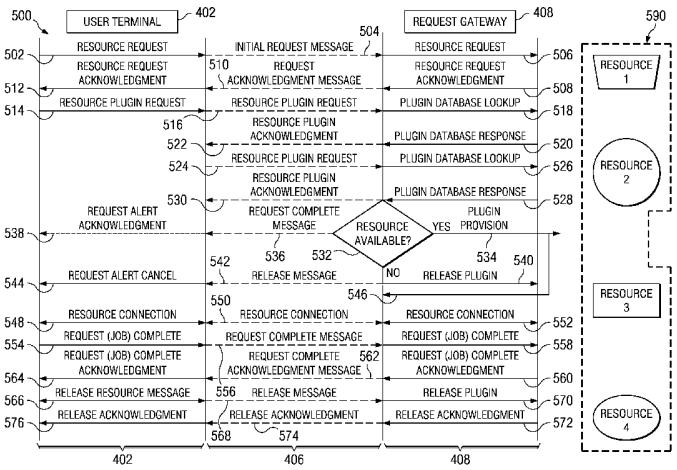
(74) Agent: **WILLIAMS, Julian, David**; IBM United Kingdom Limited, Intellectual Property Law, Hursley Park, Winchester Hampshire SO21 2JN (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: METHOD, SYSTEM AND PROGRAM PRODUCT FOR OUTSOURCING RESOURCES IN A GRID COMPUTING ENVIRONMENT



(57) Abstract: A technique for efficiently processing resource allocation requests in an on-demand environment. When a user logs into a resource allocation request manager, resource requirements entered are immediately sent to a remote resource manager to permit real-time feedback to the user who is entering the resource requirements. The user can thus reconfigure or otherwise modify the resource requests based on this feedback to optimize the overall resource allocation request. This also allows an organization providing services such as computer hosting services to re-allocate resources in advance if a new resource request would exceed a service level agreement outlined in advance with a customer, thereby improving resource planning in a data center environment. In addition, the dynamic feedback on the new resource request that exceeds a service level agreement can be used to develop a new service level agreement or temporary service offer that addresses the unique resource condition.

WO 2006/106142 A1

Published:
— *with international search report*
— *with amended claims*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

METHOD, SYSTEM AND PROGRAM PRODUCT FOR OUTSOURCING RESOURCES IN A GRID COMPUTING ENVIRONMENT

BACKGROUND OF THE INVENTION

Technical Field

The present invention relates to resource management in a data processing system, and more particular relates to real-time analysis of resource availability when generating a comprehensive resource request for remote resource fulfillment.

Description of Related Art

In today's highly competitive, global economic climate, many companies are resorting to outsourcing certain operations to other companies and/or countries in an effort to control costs. A user/customer that desires such outsourcing must work with a supplier (outsourcer) to define and manage the work to be outsourced.

Grids are environments that enable software applications to integrate instruments, displays, computational and information resources that are managed by diverse organizations in widespread locations. Grid computing is all about sharing resources that are located in different places based on different architectures and belonging to different management domains. Computer grids create a powerful pool of computing resources capable of running the most demanding scientific and engineering applications required by researchers and businesses today.

The above described outsourcing environments are evolving to include a Grid computing model where users request remote resources to be allocated dynamically. However, today's technology usually requires significant advanced reservation, e.g. days to weeks. In these scenarios, a maintenance level service agreement may be in place between the customer and the supplier (the outsourcer) whereby the customer will only have access to certain resources to compensate for overload scenarios such as periods in which the local customer resources are over-utilized and additional work must be pushed off-site. However, current service level agreements (SLAs) are typically very static. In order for suppliers to adequately plan and architect a collection of resources, e.g. firewalls, network partitioning, etc., customers ask for a specific set of resources

and are not given access to any resources that do not match that type, or even a certain class of service like security. The customer does not have access to an 'open environment' behind the supplier's firewall. In addition, a resource request in such an environment is often a fairly complex description of requirements (hardware, software, networks, etc.) which must be parsed in its entirety before any decisions can be made about available resource pools, pricing, and time to allocate such resources. In this type of on-demand environment, it would be desirable to process resource allocation requests more efficiently with a dynamic resource request system, where the SLAs can become more generalized and more resource options may become available.

SUMMARY OF THE INVENTION

The present invention provides a method, system and program product for efficiently processing resource allocation requests in an on-demand or real-time environment. Such efficiency is achieved by not just parsing hierarchical levels of resource requests by a remote resource manager, but in addition by also receiving the requests by such remote resource manager as they are entered by a user/requester. When a user logs into a resource allocation request manager, various resource requirements entered are immediately sent to a remote resource manager to allow for dynamic feedback to the user who is entering the resource requirements as to any issues or concerns regarding the requested resource(s) or plug-in element(s), such as lack of availability. The user can thus reconfigure or otherwise modify the various resource requests based on this feedback to optimize the overall resource allocation request. This also allows an organization providing services such as computer hosting services to re-allocate resources in advance if a new resource request would exceed a service level agreement outlined in advance with a customer, thereby improving resource planning in a data center environment. In addition, the dynamic feedback on the new resource request that exceeds an SLA can be used to develop a new SLA or temporary service offer that addresses the unique resource condition.

To achieve the above, a hierarchical description of an atomistic resource model is provided, from which a resource request may be generated. A discrete signal system is provided which generates a resource request, responsive to user interaction with the system, to a remote resource manager. A plug-in is provided to the remote resource manager to handle the dynamic, partially built requests. If a user

deletes or decommits a field, a signal is dynamically sent to the remote resource manager to de-commit resources.

BRIEF DESCRIPTION OF THE DRAWINGS

A preferred embodiment of the present invention will now be described, by way of example only, with reference to the accompanying drawings, wherein:

Figure 1 is a diagram illustrating a distributed data processing system according to a preferred embodiment of the present invention.

Figure 2 is an exemplary block diagram of a server according to a preferred embodiment of the present invention.

Figure 3 is an exemplary block diagram of a client according to a preferred embodiment of the present invention.

Figure 4 is an exemplary overall architecture of a resource request and fulfillment system.

Figure 5 is an exemplary message interchange between a resource requestor application running on a user/client device and a remote resource manager running on a server.

Figure 6 is an exemplary hierarchical resource model.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

With reference now to the figures, **Figure 1** depicts a pictorial representation of a network of data processing systems in which a preferred embodiment of the present invention may be implemented. Network data processing system **100** is a grid network of computers in which a preferred embodiment of the present invention may be implemented. Network data processing system **100** contains a network **102**, which is the medium used to provide communications links between various devices and computers connected together within network data processing system **100**. Network **102** may include connections, such as wire, wireless communication links, or fiber optic cables.

In the depicted example, server **104** is connected to network **102** along with storage unit **106**. In addition, clients **108**, **110**, and **112** are connected to network **102**. These clients **108**, **110**, and **112** may be, for example, personal computers or network computers. In the depicted example, server **104** provides resources, such as operating system images, applications or access to hardware resources, to clients **108-112**. Clients **108**, **110**, and **112** are clients to server **104**. Network data processing system **100** may include additional servers, clients, and other devices not shown. In the depicted example, network data processing system **100** is the Internet with network **102** representing a worldwide collection of networks and gateways that use the Transmission Control Protocol/Internet Protocol (TCP/IP) suite of protocols to communicate with one another. At the heart of the Internet is a backbone of high-speed data communication lines between major nodes or host computers, consisting of thousands of commercial, government, educational and other computer systems that route data and messages. Of course, network data processing system **100** also may be implemented as a number of different types of networks, such as for example, an intranet, a local area network (LAN), or a wide area network (WAN). **Figure 1** is intended as an example, and not as an architectural limitation for the present invention.

Referring to **Figure 2**, a block diagram of a data processing system that may be implemented as a server, such as server **104** in **Figure 1**, is depicted in accordance with a preferred embodiment of the present invention. Data processing system **200** may be a symmetric multiprocessor (SMP) system including a plurality of processors **202** and **204** connected to system bus **206**. Alternatively, a single processor system may be employed. Also connected to system bus **206** is memory controller/cache **208**, which provides an interface to local memory **209**. I/O bus bridge **210** is connected to system bus **206** and provides an interface to I/O bus **212**. Memory controller/cache **208** and I/O bus bridge **210** may be integrated as depicted.

Peripheral component interconnect (PCI) bus bridge **214** connected to I/O bus **212** provides an interface to PCI local bus **216**. A number of modems may be connected to PCI local bus **216**. Typical PCI bus implementations will support four PCI expansion slots or add-in connectors. Communications links to clients **108-112** in **Figure 1** may be provided through modem **218** and network adapter **220** connected to PCI local bus **216** through add-in boards.

Additional PCI bus bridges **222** and **224** provide interfaces for additional PCI local buses **226** and **228**, from which additional modems or network adapters may be supported. In this manner, data processing system **200** allows connections to multiple network computers. A memory-mapped graphics adapter **230** and hard disk **232** may also be connected to I/O bus **212** as depicted, either directly or indirectly.

Those of ordinary skill in the art will appreciate that the hardware depicted in **Figure 2** may vary. For example, other peripheral devices, such as optical disk drives and the like, also may be used in addition to or in place of the hardware depicted. The depicted example is not meant to imply architectural limitations with respect to the present invention.

The data processing system depicted in **Figure 2** may be, for example, an IBM(R) eServer(TM) pSeries(TM) system, a product of International Business Machines Corporation in Armonk, New York, running the Advanced Interactive Executive (AIX) operating system or LINUX(TM) operating system.

With reference now to **Figure 3**, a block diagram illustrating a data processing system is depicted in which the present invention may be implemented. Data processing system **300** is an example of a client computer. Data processing system **300** employs a peripheral component interconnect (PCI) local bus architecture. Although the depicted example employs a PCI bus, other bus architectures such as Accelerated Graphics Port (AGP) and Industry Standard Architecture (ISA) may be used. Processor **302** and main memory **304** are connected to PCI local bus **306** through PCI bridge **308**. PCI bridge **308** also may include an integrated memory controller and cache memory for processor **302**. Additional connections to PCI local bus **306** may be made through direct component interconnection or through add-in boards. In the depicted example, local area network (LAN) adapter **310**, SCSI host bus adapter **312**, and expansion bus interface **314** are connected to PCI local bus **306** by direct component connection. In contrast, audio adapter **316**, graphics adapter **318**, and audio/video adapter **319** are connected to PCI local bus **306** by add-in boards inserted into expansion slots. Expansion bus interface **314** provides a connection for a keyboard and mouse adapter **320**, modem **322**, and additional memory **324**. Small computer system interface (SCSI) host bus adapter **312** provides a connection for hard disk drive **326**, tape drive **328**, and CD-ROM drive **330**.

Typical PCI local bus implementations will support three or four PCI expansion slots or add-in connectors.

An operating system runs on processor **302** and is used to coordinate and provide control of various components within data processing system **300** in **Figure 3**. The operating system may be a commercially available operating system, such as Windows XP(TM), which is available from Microsoft(R) Corporation. An object oriented programming system such as Java may run in conjunction with the operating system and provide calls to the operating system from Java(TM) programs or applications executing on data processing system **300**. "Java" is a trademark of Sun Microsystems, Inc. Instructions for the operating system, the object-oriented operating system, and applications or programs are located on storage devices, such as hard disk drive **326**, and may be loaded into main memory **304** for execution by processor **302**.

Those of ordinary skill in the art will appreciate that the hardware in **Figure 3** may vary depending on the implementation. Other internal hardware or peripheral devices, such as flash read-only memory (ROM), equivalent nonvolatile memory, or optical disk drives and the like, may be used in addition to or in place of the hardware depicted in **Figure 3**. Also, the processes of the present invention may be applied to a multiprocessor data processing system.

As another example, data processing system **300** may be a stand-alone system configured to be bootable without relying on some type of network communication interfaces As a further example, data processing system **300** may be a personal digital assistant (PDA) device, which is configured with ROM and/or flash ROM in order to provide non-volatile memory for storing operating system files and/or user-generated data.

The depicted example in **Figure 3** and above-described examples are not meant to imply architectural limitations. For example, data processing system **300** also may be a notebook computer or hand held computer in addition to taking the form of a PDA. Data processing system **300** also may be a kiosk or a Web appliance.

Referring now to **Figure 4**, there is shown at **400** a system for requesting resources in a distributed computing environment in accordance with the present invention. A user request is input by a user using terminal **402**, which may part of or attached to a client device such as

client **108** shown in **Figure 1.** The user request may be any of a plurality of different types of requests, such as those shown in block **404**, and may include a login request, a system-level request for systems such as servers or storage, a system software request for software such as an operating system (O/S), microcode (ucode) or device/adapter driver(s), a system hardware request for hardware such as a processor, network or storage, an application middleware request, a geography request, a security request, and a capacity/performance request. The user request is transmitted across path **406**, which may be implemented as a network such as network **102** shown in **Figure 1**, to a request gateway **408**, which may be implemented by a server such as server **104** shown in **Figure 1.** At the request gateway, a thread is created from thread pool **410**, and this thread creates a finite state machine (FSM) **412** in the request gateway **408** to handle subsequent signals and errors such as request signals from the user and allowed/disallowed resource allocation attempts. The finite state machine **412** dynamically creates a plug-in workflow **414** which manages, in conjunction with the state table **416**, different states of the request, including error conditions. The finite state machine uses a resource database **418** to determine if requested resources are available, and to temporarily commit resources until all signals in the request have been received. Plug-ins to translate requirements and coordinate with the provisioning engine **422** are dynamically executed within a runtime engine **424** in the finite state machine. These plug-ins are shown at **420** and are a part of a plug-in library. These plug-ins may provide functions such as Login, System Request (e.g. server, storage), System Software Request (e.g. operating system, microcode, drivers), System Hardware Request (e.g. processor, network, storage), Application Middleware Request, Geography, Security and Capacity/Performance, as shown at **420.** An error or unavailable signal can be generated at any point based upon the state of the user request.

Turning now to **Figure 5**, a representative series of signals or transactions that flow between a user terminal such as **402** and a request gateway such as **408** are shown at **500.** In this scenario, a user is generating a conglomerate resource request file or model, and inputs individual resource requests which are received at user terminal **402** and merged into a conglomerate resource request template (not shown). Responsive to receiving a first of many requests, a first resource request is signaled at **502.** This signal causes an initial request message to be generated and transmitted to the request gateway at **504.** This initial

request message is received by the Request Gateway **408** at **506**, and an acknowledgement is signaled by the request gateway at **508**. This signal causes a request acknowledgement message to be transmitted across the network at **510**, where it is received by the user terminal at **512**. This initial resource request could be a login command, for example. The specific *n* plug-in resources are then sequentially transmitted to the request gateway, as initiated at **514**. The sequential transmission of individual resources is done even when using a conglomerate template. An example of this interplay between transmission of individual resources and a client conglomerate template is provided by the IBM Websphere(R) Application Server, available from International Business Machines Corporation, Armonk, N.Y. Business Gateway. The first of these plug-in requests is transmitted at **516**, and received by the request gateway where the plug-in database/library, such as that shown at **420** in **Figure 4**, is queried or looked-up to determine availability of such plug-in resource at **518**. A status response is sent by the request gateway at **520** and transmitted across network **406** at **522**. Responsive to receiving this response, a gateway client (not shown, but further described below) then sequentially signals a request for the next 2 to *n* plug-in requests at **524** in similar fashion to the first plug-in request, and the request gateway **408** processes these in the same manner as the first request at **526** and **528**, with resulting acknowledgements being signaled back to the requestor at **530**. The gateway client is an executable which is acting on behalf of the user/requestor to integrate into the dynamic request system, aggregates the received responses and presents the appropriate information to the user. Such client executable does not have to reside or execute on the requestor's machine, but rather has to act on behalf of the requestor as it processes responses and presents information to the requestor. An example of such a client would be a Java applet. Representative plug-in resources are shown at **590** in **Figure 5**.

As resource plug-in requests are received by request gateway **408**, and responsive to a determination being made at **532** on whether the requested resource is available, the plug-in is either provided at **534** if available, or the plug-in reservation request is released at **540** if not available. If available, a signal is generated at **546** (for inclusion in the response message that is subsequently sent across resource connection **550** and **548**). If the resource is not available and has been released, a release message is sent to the requestor at **542**, resulting in an alert message at the user terminal indicating that the resource request has been

cancelled at **544**. This alert message advantageously provides for dynamic, real-time feedback of resource status from a remote resource manager, enabling the user/requestor to adjust their overall resource request package to account for such resource plug-in unavailability.

Once all resource requests have been processed, the resource connection is provided by **548**, **550** and **552** to provide the requested sources, which are remotely provided in the grid computer environment in the preferred embodiment, to the requester/customer who is outsourcing such requested resources. Once the outsourced service or job has completed, the user terminal signals a job complete request at **554**, resulting in a request complete message being sent at **556**, and received by the request gateway at **558**. The gateway **408** signals an acknowledgement at **560**, resulting in an acknowledgement message being sent to the user/requester terminal at **562**, and received at **564**. The resource allocation request manager application running at user terminal **402**, responsive to the job complete request acknowledgement signal being received at **564**, signals a release resource message at **566**, which is sent across the network at **568** and results in the request gateway releasing the 1-*n* plug-ins relating to this completed job at **570**. The request gateway acknowledges the release at **572**, which is sent to the user terminal at **574** and received at **576** to thereby end the current outsourced job.

This dynamic resource allocation scheme advantageously allows an organization providing services such as computer hosting services to re-allocate resources in advance if a service level agreement with a customer would otherwise be in violation without such re-allocation, thereby improving resource planning in a data center environment.

To facilitate the above data flow, a hierarchical description of an atomistic resource model is provided, where several computer systems have a lowest common building block. For instance, as shown in **Figure 6**, a cluster **602** is a system built using software **604** and hardware **606**. The software **604** is built using an application environment **608**, system management **610**, data management **612** and workload management **614**. Data management **612** is built using one or more databases **616** and one or more file systems **618**. The hardware is built using servers, disks, and a network, such as is shown by server(s) **620**, storage **622** and network **624**. It is probably difficult to define disks beyond a certain capacity, so storage **622** is considered to be an atomistic resource. Alternatively, if

a multitude of heterogeneous storage devices were provided, such as disk storage devices and tape storage devices, the storage could be further modeled at a lower level to include both disk and tape storage devices, which themselves could be considered to be an atomistic resource. Server **620** and network **624** can be described with even further detail. A server **620** may be described using operating system **626**, I/O **628**, on board memory **630**, and processor **632**. A network may be described by a switch **634** and the type of connectivity **636**. In each of these instances, a hierarchy of atomistic (e.g. processor **632**, storage **622**, database **616**, system management **610**) and then compound resources (e.g. software **604** and hardware **606**) are used to define the cluster. A plurality of such clusters can be further organized into a multi-cluster implementation (not shown), for use in a grid computing environment.

Furthermore, this hierarchical model is expandable to cover different events or responses that may be required in a requested resource/system. For instance, a single event such as "More Capacity is required to execute Application 'A'" can be broken down into several smaller events: (i) "More boxes of type E required", (ii) "More network bandwidth on switch plane 3 required", (iii) "Application J on server 9.2.9.32 must be suspended", which would correspond to a similar set of responses, (iv) "System manager suspending application J on server 9.3.9.32", (v) "Switch configuration updated", (vi) "Linux kernel rebuilt on server 9.3.9.32", which would in total correspond to the aggregated response of "More Application A systems have been allocated and built". Complex responses for resource requests are thus built from fundamental or atomistic responses to the basic building blocks in the system.

An example of a conglomerate template can be seen in **figure 4**. In this figure, the template is depicted on the client side (at **404**), which has been built by a user constructing various portions of a resource request. On the server or system side **408**, a conglomerate is reconstructed by receiving pieces of this request and then using the appropriate plug-ins as determined by the finite state machine to dynamically assemble the resource request. Although in the preferred embodiment the conglomerate is not required to be assembled in its entirety on the server side before the plug-ins execute, a historical trace of the actions that were taken by the plug-ins can be used to reconstruct the conglomerate on the server side.

Thus, there is provided an improved technique for efficiently processing resource allocation requests in an on-demand environment. Such

efficiency is achieved by not just parsing hierarchical levels of resource requests by a remote resource manager, but in addition by also receiving the requests by such remote resource manager as they are entered by a user/requester. When a user logs into a resource allocation request manager, various resource requirements entered are immediately sent to a remote resource manager to allow for feedback to the user entering the resource requirements as to any issues or concerns regarding the requested resource(s), such as lack of availability. The user can thus reconfigure or otherwise modify the various resource requests based on this feedback to optimize the overall resource allocation request.

It is important to note that while the present invention has been described in the context of a fully functioning data processing system, those of ordinary skill in the art will appreciate that the processes of the present invention are capable of being distributed in the form of a computer readable medium of instructions and a variety of forms and that the present invention applies equally regardless of the particular type of signal bearing media actually used to carry out the distribution. Examples of computer readable media include recordable-type media, such as a floppy disk, a hard disk drive, a RAM, CD-ROMs, DVD-ROMs, and transmission-type media, such as digital and analog communications links, wired or wireless communications links using transmission forms, such as, for example, radio frequency and light wave transmissions. The computer readable media may take the form of coded formats that are decoded for actual use in a particular data processing system.

The description of the present invention has been presented for purposes of illustration and description, and is not intended to be exhaustive or limited to the embodiment of invention in the form disclosed. Many modifications and variations will be apparent to those of ordinary skill in the art. The embodiment was chosen and described in order to best explain the principles of the invention, the practical application, and to enable others of ordinary skill in the art to understand the various embodiments with various modifications as are suited to the particular use contemplated.

CLAIMS

1. A method for accessing on demand a plurality of resources in a data processing environment comprising a resource requestor and a resource provider, comprising the steps of:

generating a conglomerate resource request by receiving user input specifying a plurality of resource requests;

responsive to receiving a first user input for a first resource request, sending a signal to the resource provider identifying the first resource request;

determining, by the resource provider, if a resource is available for fulfilling the first resource request; and

if the resource is available, providing a response from the resource provider to the resource requestor indicating that the resource is available, wherein the response is provided to the resource requestor while the conglomerate resource request is being generated.

2. The method of Claim 1, wherein if the resource is available, reserving the resource by the resource provider for use by the resource requestor.

3. The method of Claim 1, wherein if the resource is not available, releasing the resource by the resource provider, and providing a response from the resource provider to the resource requestor indicating that the resource is not available for use.

4. The method of Claim 1, wherein step (d) further comprises reserving the resource for subsequent use.

5. The method of Claim 1, wherein the step of generating a conglomerate resource request comprises a step of populating a hierarchical resource request template with the specified plurality of resource requests.

6. A system for accessing on demand a plurality of resources in a data processing environment, comprising

a resource requestor that generates a conglomerate resource request by receiving user input specifying a plurality of resource requests and

responsive to receiving a first user input for a first resource request, sends a signal identifying the first resource request; and

a resource provider that, responsive to receiving the signal, determines if a resource is available for fulfilling the first resource request and, if the resource is available, provides a response indicating that the resource is available, wherein the response is provided to the resource requestor while the resource requestor is generating the conglomerate resource request.

7. The system of Claim 6, wherein if the resource is available, the resource is reserved by the resource provider for use by the resource requestor.

8. The system of Claim 6, wherein if the resource is not available, the resource is released by the resource provider, and a response is provided by the resource provider to the resource requestor indicating that the resource is not available for use.

9. The system of Claim 6, wherein the conglomerate resource request comprises a hierarchical resource request template with the specified plurality of resource requests.

10. A computer program comprising computer program code to, when loaded into a computer system and executed thereon, cause said computer system to perform the steps of a method as claimed in any of claims 1 to 5.

AMENDED CLAIMS
received by the International Bureau on 24 August 2006 (24.08.2006)

CLAIMS

1. A method for accessing on demand a plurality of resources in a data processing environment comprising a resource requestor and a resource provider, comprising the steps of:

generating a conglomerate resource request by receiving user input specifying a plurality of resource requests;

responsive to receiving a first user input for a first resource request, sending a signal to the resource provider identifying the first resource request;

determining, by the resource provider, if a resource is available for fulfilling the first resource request; and

if the resource is available, providing a response from the resource provider to the resource requestor indicating that the resource is available, wherein the response is provided to the resource requestor while the conglomerate resource request is being generated; or

if the resource is not available, releasing the resource by the resource provider, and providing a response from the resource provider to the resource requestor indicating that the resource is not available for use.

2. The method of Claim 1, wherein if the resource is available, reserving the resource by the resource provider for use by the resource requestor.

3. The method of Claim 2, wherein the step of reserving further comprises reserving the resource for subsequent use.

4. The method of Claim 1, wherein the step of generating a conglomerate resource request comprises a step of populating a hierarchical resource request template with the specified plurality of resource requests.

5. A system for accessing on demand a plurality of resources in a data processing environment, comprising

a resource requestor that generates a conglomerate resource request by receiving user input specifying a plurality of resource requests and

responsive to receiving a first user input for a first resource request, sends a signal identifying the first resource request; and

a resource provider that, responsive to receiving the signal, determines if a resource is available for fulfilling the first resource request and,

if the resource is available, provides a response indicating that the resource is available, wherein the response is provided to the resource requestor while the resource requestor is generating the conglomerate resource request; or

if the resource is not available, the resource is released by the resource provider, and a response is provided by the resource provider to the resource requestor indicating that the resource is not available for use.

6. The system of Claim 5, wherein if the resource is available, the resource is reserved by the resource provider for use by the resource requestor.

7. The system of Claim 5, wherein the conglomerate resource request comprises a hierarchical resource request template with the specified plurality of resource requests.

8. A computer program comprising computer program code to, when loaded into a computer system and executed thereon, cause said computer system to perform the steps of a method as claimed in any of claims 1 to 4.

FIG. 1

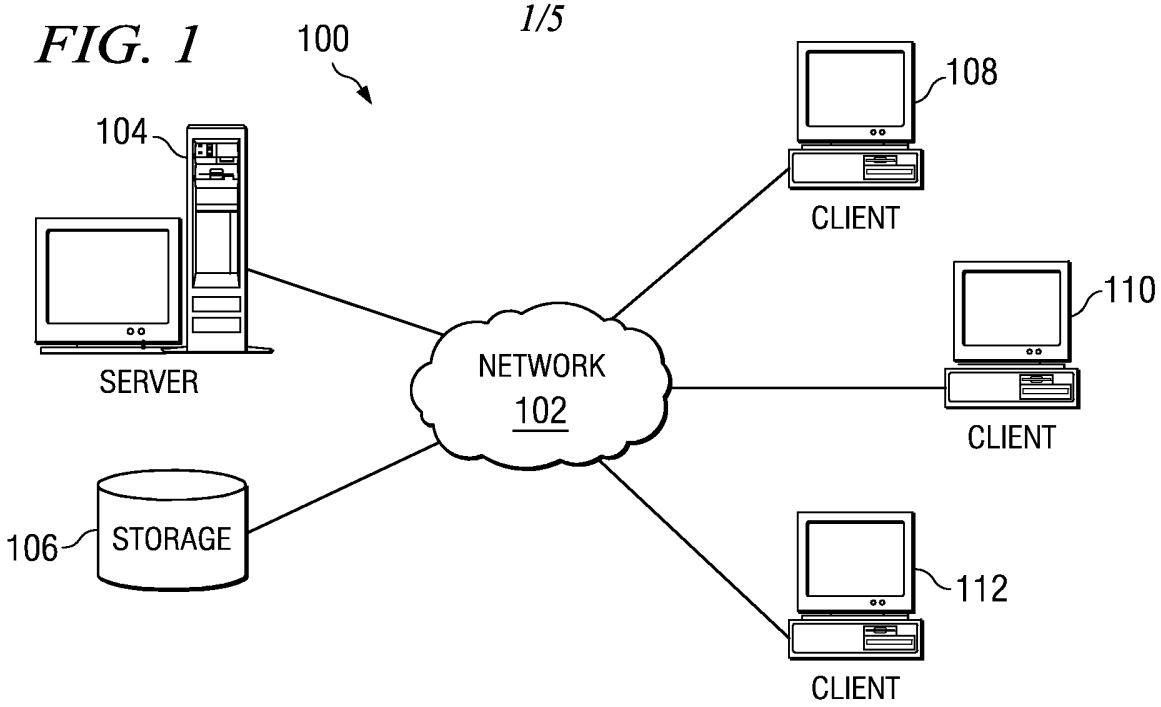

100
104
SERVER
108
CLIENT
110
CLIENT
NETWORK
102
106
STORAGE
112
CLIENT

FIG. 2

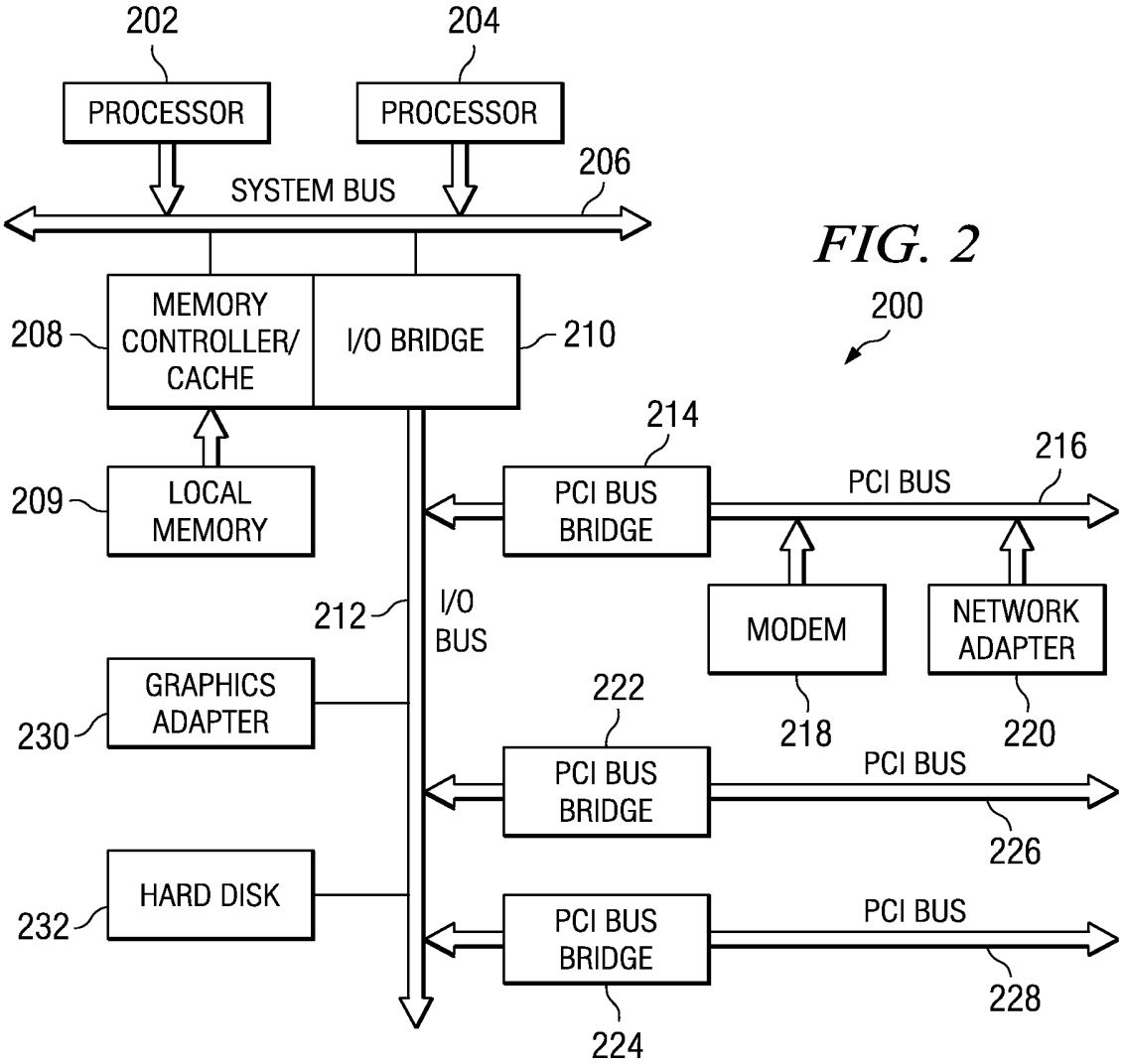
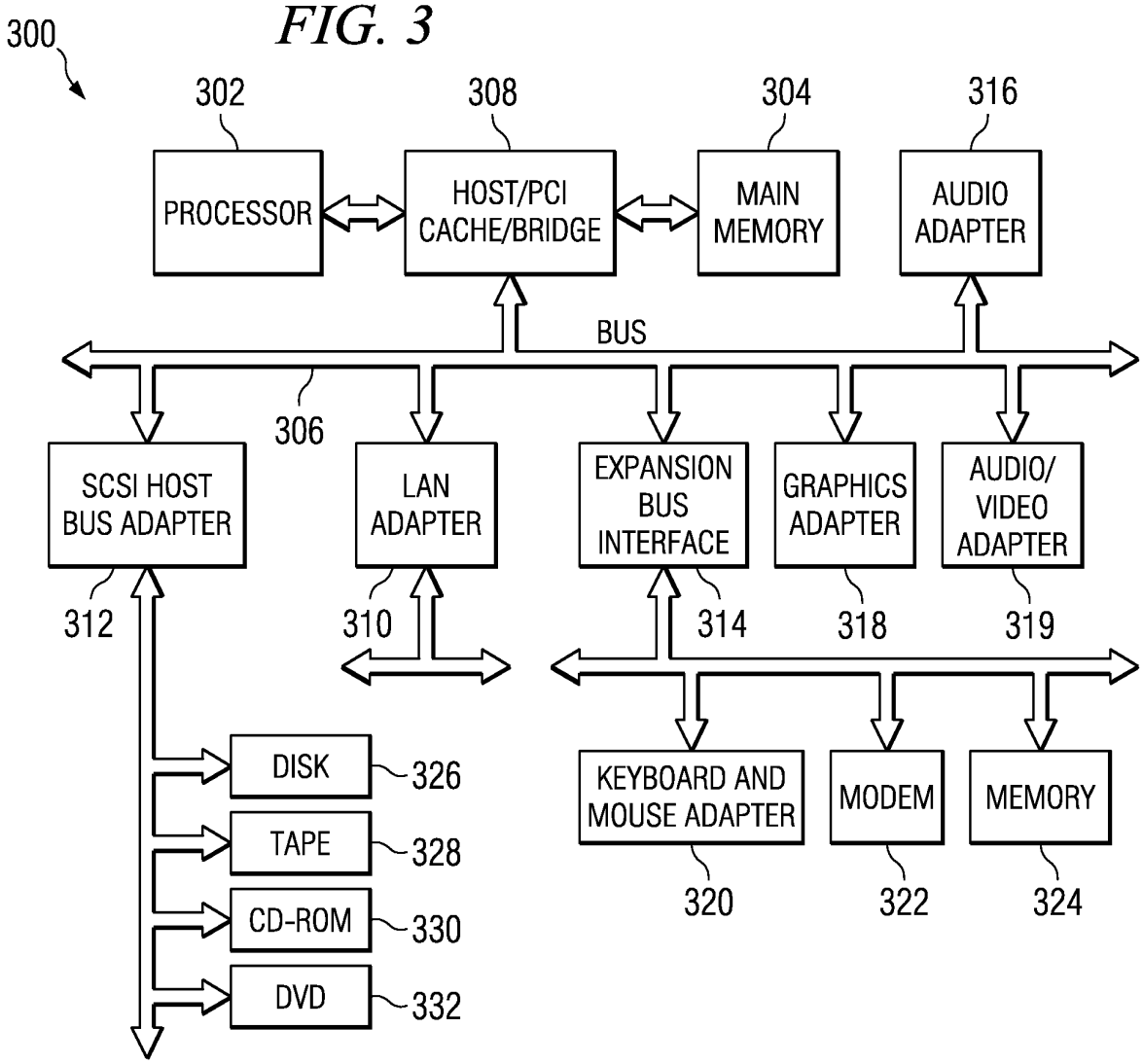

202
204
PROCESSOR
PROCESSOR
206
SYSTEM BUS
208
MEMORY CONTROLLER/ CACHE
I/O BRIDGE
210
209
LOCAL MEMORY
200
214
216
PCI BUS BRIDGE
PCI BUS
212
I/O BUS
MODEM
NETWORK ADAPTER
GRAPHICS ADAPTER
230
222
218
220
PCI BUS BRIDGE
PCI BUS
226
HARD DISK
232
PCI BUS BRIDGE
PCI BUS
228
224

FIG. 3


300
302
308
304
316
PROCESSOR
HOST/PCI CACHE/BRIDGE
MAIN MEMORY
AUDIO ADAPTER
BUS
306
SCSI HOST BUS ADAPTER
LAN ADAPTER
EXPANSION BUS INTERFACE
GRAPHICS ADAPTER
AUDIO/ VIDEO ADAPTER
312
310
314
318
319
DISK
326
TAPE
328
CD-ROM
330
DVD
332
KEYBOARD AND MOUSE ADAPTER
MODEM
MEMORY
320
322
324

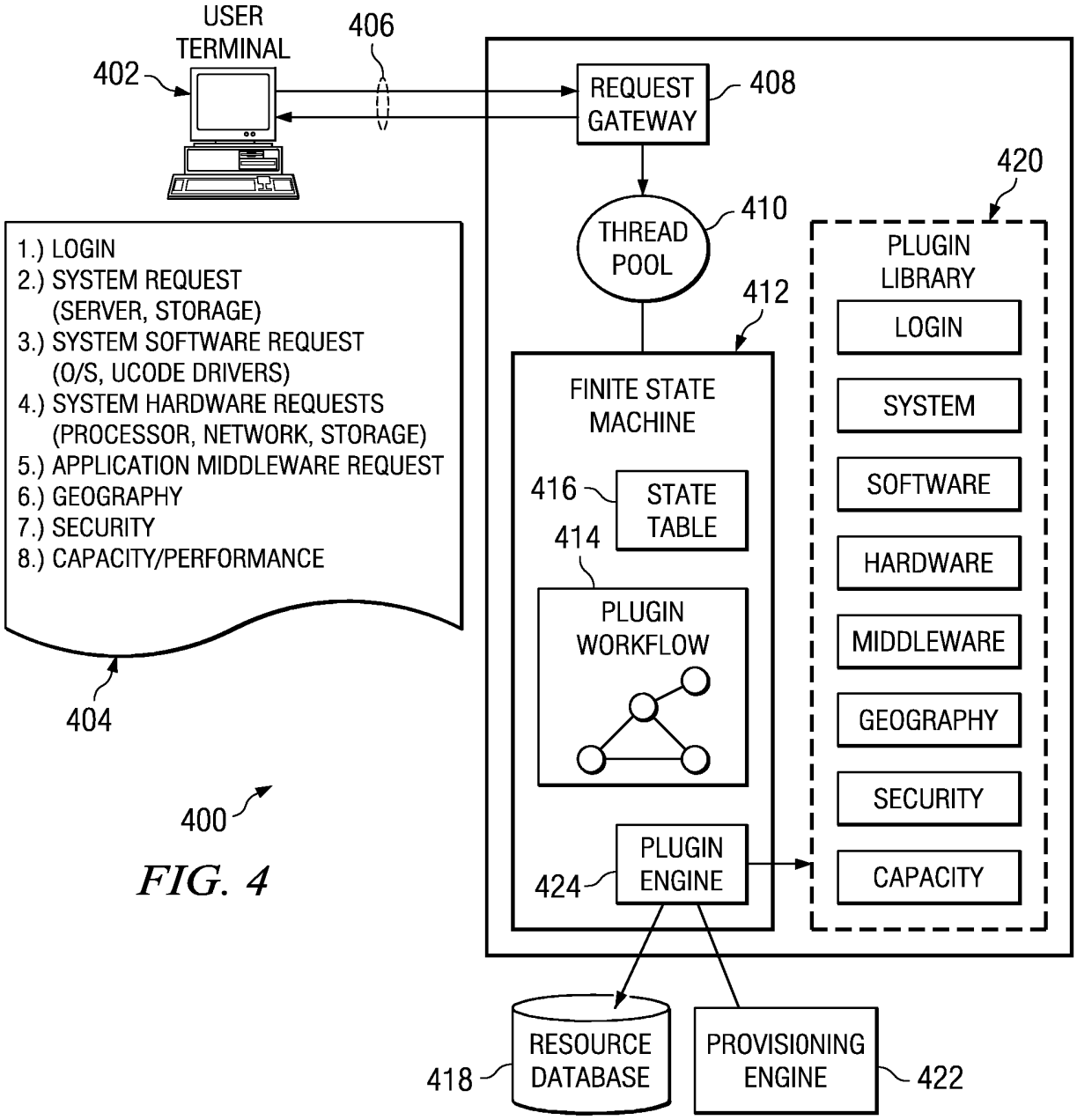

USER TERMINAL
402
406
REQUEST GATEWAY
408
420
THREAD POOL
410
PLUGIN LIBRARY
412
LOGIN
SYSTEM
SOFTWARE
HARDWARE
MIDDLEWARE
GEOGRAPHY
SECURITY
CAPACITY
FINITE STATE MACHINE
416
STATE TABLE
414
PLUGIN WORKFLOW
PLUGIN ENGINE
424
1.) LOGIN
2.) SYSTEM REQUEST (SERVER, STORAGE)
3.) SYSTEM SOFTWARE REQUEST (O/S, UCODE DRIVERS)
4.) SYSTEM HARDWARE REQUESTS (PROCESSOR, NETWORK, STORAGE)
5.) APPLICATION MIDDLEWARE REQUEST
6.) GEOGRAPHY
7.) SECURITY
8.) CAPACITY/PERFORMANCE
404
400
RESOURCE DATABASE
418
PROVISIONING ENGINE
422

FIG. 4

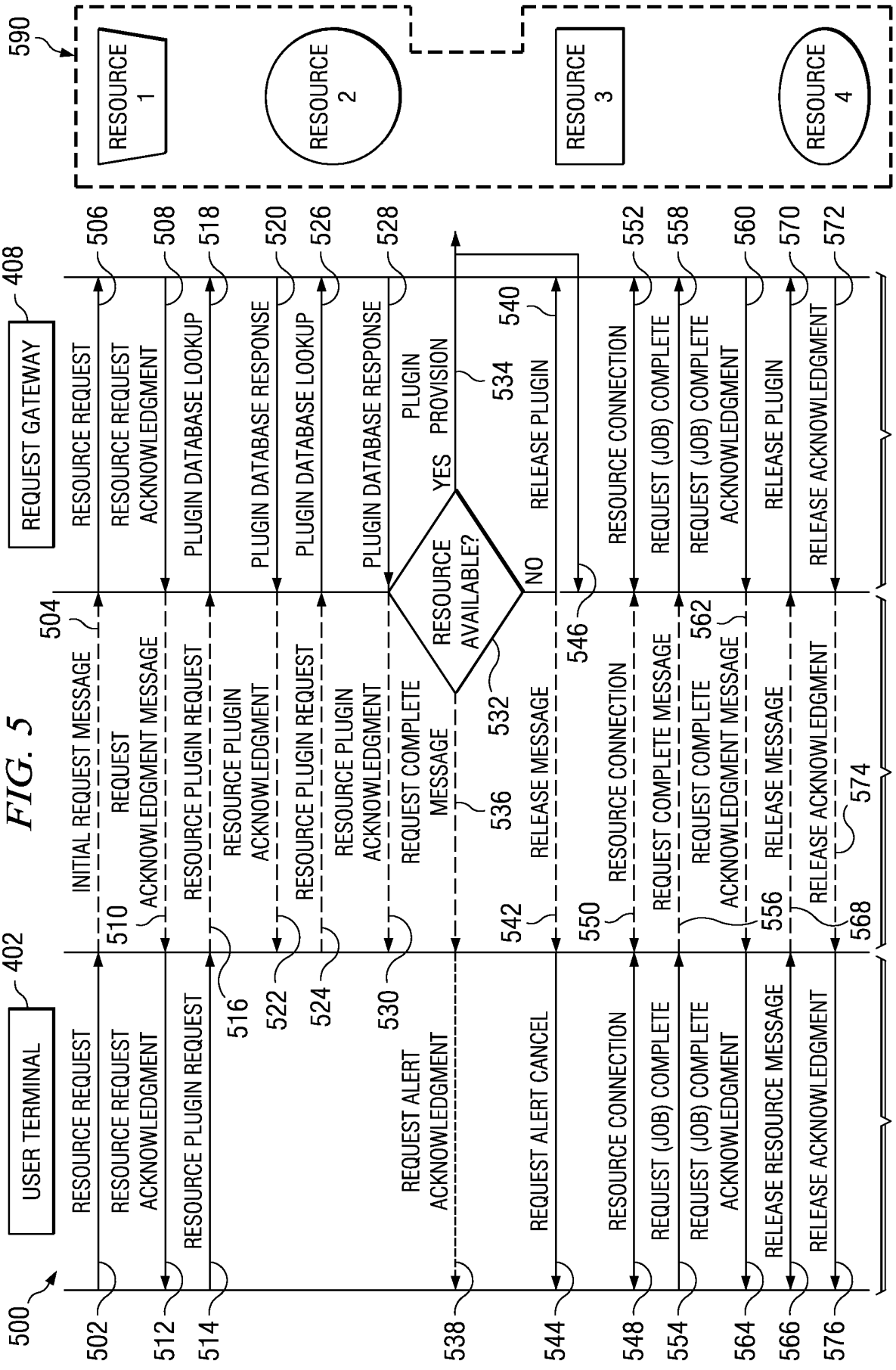
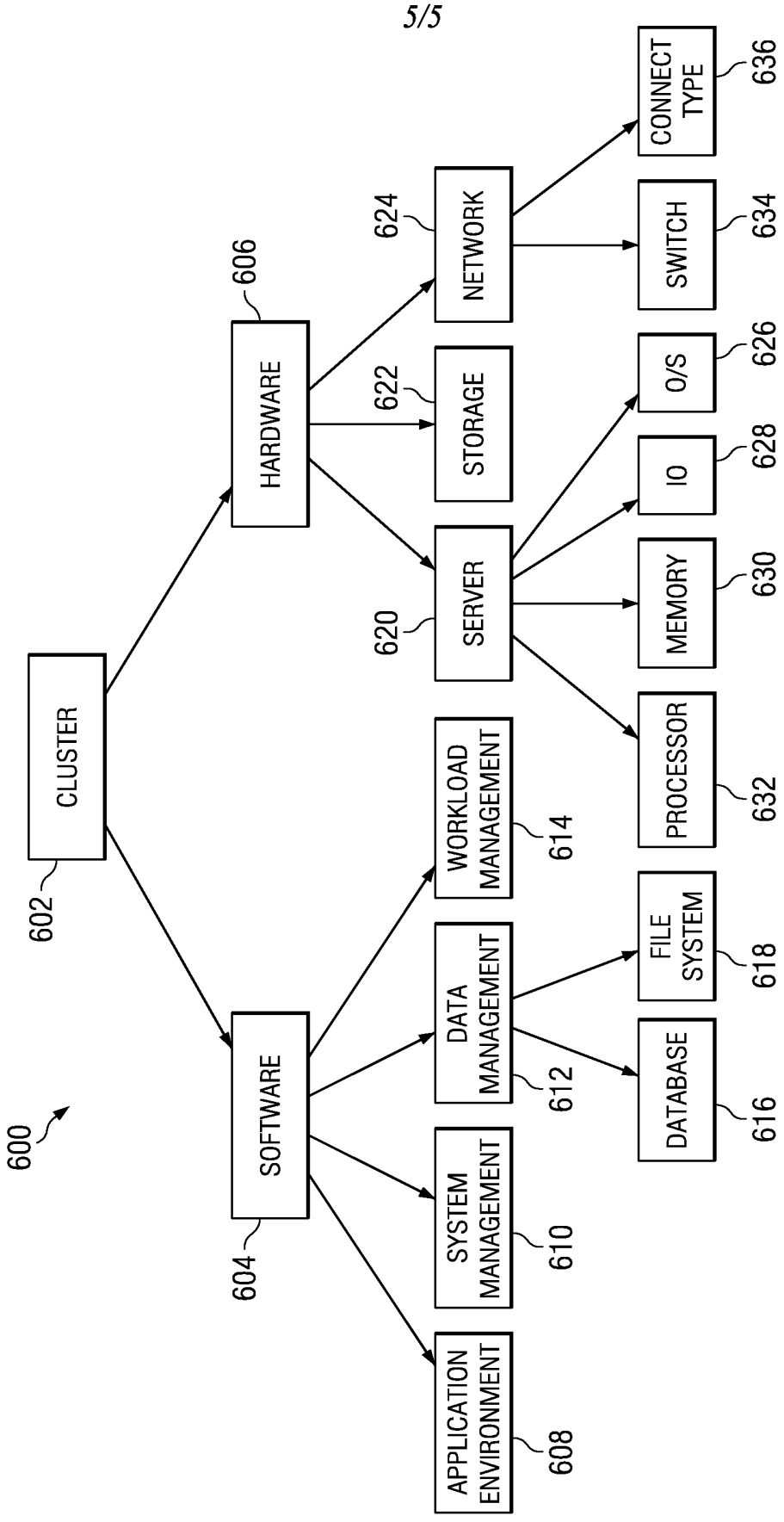

590
RESOURCE 1
RESOURCE 2
RESOURCE 3
RESOURCE 4
USER TERMINAL
402
REQUEST GATEWAY
408
504
406
500
RESOURCE REQUEST
RESOURCE REQUEST ACKNOWLEDGMENT
RESOURCE PLUGIN REQUEST
RESOURCE PLUGIN ACKNOWLEDGMENT
RESOURCE PLUGIN REQUEST
RESOURCE PLUGIN ACKNOWLEDGMENT
REQUEST COMPLETE MESSAGE
REQUEST ALERT ACKNOWLEDGMENT
REQUEST ALERT CANCEL
RESOURCE CONNECTION
REQUEST (JOB) COMPLETE
REQUEST (JOB) COMPLETE ACKNOWLEDGMENT
RELEASE RESOURCE MESSAGE
RELEASE ACKNOWLEDGMENT
INITIAL REQUEST MESSAGE
REQUEST ACKNOWLEDGMENT MESSAGE
RESOURCE PLUGIN REQUEST
RESOURCE PLUGIN ACKNOWLEDGMENT
RESOURCE PLUGIN REQUEST
RESOURCE PLUGIN ACKNOWLEDGMENT
REQUEST COMPLETE MESSAGE
RELEASE MESSAGE
RESOURCE CONNECTION
REQUEST COMPLETE MESSAGE
REQUEST COMPLETE ACKNOWLEDGMENT MESSAGE
RELEASE MESSAGE
RELEASE ACKNOWLEDGMENT
RESOURCE REQUEST
RESOURCE REQUEST ACKNOWLEDGMENT
PLUGIN DATABASE LOOKUP
PLUGIN DATABASE RESPONSE
PLUGIN DATABASE LOOKUP
PLUGIN DATABASE RESPONSE
RESOURCE AVAILABLE?
YES
PLUGIN PROVISION
NO
RELEASE PLUGIN
RESOURCE CONNECTION
REQUEST (JOB) COMPLETE
REQUEST (JOB) COMPLETE ACKNOWLEDGMENT
RELEASE PLUGIN
RELEASE ACKNOWLEDGMENT
502
510
512
514
516
522
524
530
532
534
536
538
540
542
544
546
548
550
552
554
556
558
560
562
564
566
568
570
572
574
576

FIG. 5

FIG. 6


600
CLUSTER
602
SOFTWARE
604
HARDWARE
606
APPLICATION ENVIRONMENT
608
SYSTEM MANAGEMENT
610
DATA MANAGEMENT
612
WORKLOAD MANAGEMENT
614
DATABASE
616
FILE SYSTEM
618
SERVER
620
STORAGE
622
NETWORK
624
PROCESSOR
632
MEMORY
630
IO
628
O/S
626
SWITCH
634
CONNECT TYPE
636

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2006/061434

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F9/50
ADD. H04L29/06

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)
G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, IBM-TDB, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2001/029519 A1 (HALLINAN MICHAEL ET AL) 11 October 2001 (2001-10-11) paragraphs [0042], [0059]; figure.4	1-10
A	US 6 009 275 A (DEKONING ET AL) 28 December 1999 (1999-12-28) column 5, line 60 - column 7, line 50	1-10
A	WO 02/31672 A (ZUCOTTO WIRELESS INC; COMEAU, GUILLAUME; REBEIRO, SARAH; NOWAK, CLIFTO) 18 April 2002 (2002-04-18) page 65, line 16 - page 67, line 12; figure 17	1,5,6,9
A	US 6 785 675 B1 (GRAVES JOHN ET AL) 31 August 2004 (2004-08-31) the whole document	1,5,6,9

-/--

[X] Further documents are listed in the continuation of Box C.

[X] 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 document 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
26 July 2006	04/08/2006

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Michel, T

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2006/061434

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2003/191677 A1 (AKKIRAJU RAMA K ET AL) 9 October 2003 (2003-10-09) paragraph [0036] _____	1,5,6,9

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2006/061434

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2001029519 A1	11-10-2001	NONE	
US 6009275 A	28-12-1999	DE 69521549 D1	09-08-2001
		DE 69521549 T2	25-10-2001
		EP 0676699 A2	11-10-1995
		JP 8044681 A	16-02-1996
WO 0231672 A	18-04-2002	AU 9533401 A	22-04-2002
		US 2002091826 A1	11-07-2002
US 6785675 B1	31-08-2004	US 2004230578 A1	18-11-2004
US 2003191677 A1	09-10-2003	NONE	