

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
16 March 2006 (16.03.2006)

PCT

(10) International Publication Number
WO 2006/027038 A2

(51) International Patent Classification:

G06F 11/34 (2006.01) **G06F 9/50** (2006.01)

(21) International Application Number:

PCT/EP2005/007020

(22) International Filing Date: 29 June 2005 (29.06.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/608,524 9 September 2004 (09.09.2004) US
60/609,544 13 September 2004 (13.09.2004) US

(71) Applicant (*for all designated States except US*): **FUJITSU SIEMENS COMPUTERS, INC.** [US/US]; 1250 East Arques Avenue, Mail Stop 190, Sunnyvale, CA 94085-5401 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **ARMSTRONG, Joseph, W.** [US/US]; 2850 Fresno Street, Santa Cruz, CA 95062 (US). **HSU, Shu-Ching** [US/US]; 38135 Paseo Padre Ct., Fremont, CA 94536 (US). **JOHNSTON, Mark** [US/US]; 545 W. Hacienda Ave #305, Campbell, CA

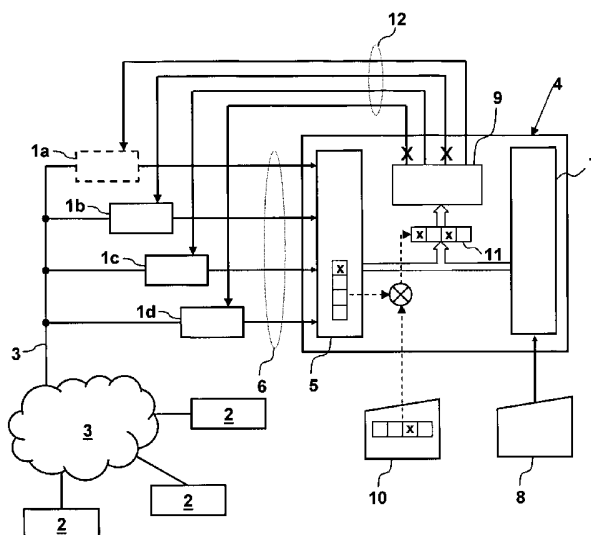
95008 (US). **KELKAR, Rahul** [IN/US]; 4950 Stevenson Blvd., Apt. #49, Fremont, CA 94538 (US). **KING, Judy** [US/US]; 221 Almond St., San Luis Obispo, CA 93405 (US). **KRESS, Brian** [US/US]; 30 Barry Street, Brockport, NY, CA 14420 (US). **PENNEPALLI, Radhika** [IN/US]; 45422 Little Foot Pl., Fremont, CA 94539 (US). **PULIJALA, Kesava** [IN/US]; 3655 Pruneridge Ave., Apt. #64, Santa Clara, CA 95051 (US). **SHEN, Guangji** [CN/US]; 300 Richfield Dr., Apt. #3, San Jose, CA 95129 (US). **SINGH, Pushkar** [IN/US]; 39939 Stevenson Common, Apt. # 2125, Fremont, CA 94538 (US). **STONER, Kevin** [US/US]; 1500 Virginia Place, #104, San Jose, CA 95116 (US). **VISHWANATHAN, Rajendran** [IN/US]; 1403 Petal Way, San Jose, CA 95129 (US).

(74) Agent: **EPPING HERMANN FISCHER PATENTANWALTSGESELLSCHAFT MBH**; Ridlerstr. 55, 80339 München (DE).

(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, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, 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,

[Continued on next page]

(54) Title: COMPUTER ARRANGEMENT FOR PROVIDING SERVICES FOR CLIENTS OVER A NETWORK



(57) Abstract: The invention relates to a computer arrangement for providing services for clients (2) over a network (3), comprising at least one server (1) hosting at least one or more application and control means (4), the control means being set up to start or stop instances of application programs on registered servers (1). A service mode is provided by the control means (4) for each registered server (1). For a server in service mode, status information is not evaluated, no instances of application programs are started or stopped and configuration information of this server is preserved in the registration data structure (7). The service mode allows to easily access a server (1) for maintenance and/or trouble-shooting without any interference from the control means (4). After maintenance, control over the server (1) can be handed back to the control means (4) quickly and without the need to reregister the server (1).

WO 2006/027038 A2



SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC,
VN, YU, ZA, ZM, ZW.

SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN,
GQ, GW, ML, MR, NE, SN, TD, TG).

(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, MC, NL, PL, PT, RO,

Published:

— without international search report and to be republished upon receipt of that report

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.

Description of the InventionCOMPUTER ARRANGEMENT FOR PROVIDING SERVICES FOR CLIENTS OVER
A NETWORK

5

This invention relates to a computer arrangement for providing services for clients over a network, comprising at least one server hosting at least one or more applications which provide services to clients over the network and control means, said control means being set up to start or stop instances of said application programs on the servers.

10

Computer networks typically include at least one computer called a server and one or more computers called clients. Clients can request certain services over the network which are provided by the servers.

15

There is a huge variety of different types of services that can be offered by servers. One example is a file service, where clients can request the transmission of a file which is stored on the server or on a storage device associated with the server. Another example is a web service, where web pages are transmitted upon request. Also known are database services, where, in a similar fashion, data sets of a database are transmitted or processed upon request.

20

25

Usually each individual service is provided by an application program running on a server. Depending on requirements or capacity a server can host a single or also several instances of different application programs. Often a network service provider runs a plurality of servers (called a server farm) in order to be able to handle all incoming requests. Considering the growing size of modern server farms which can

30

consist of hundreds of servers there is an increasing demand to automate administrative tasks, for example allocating the available resources to different types of services. The approach to automate administrative tasks is known as

5 autonomous computing (AC). A central element in an AC-environment is a control means to perform administrative tasks which is sometimes referred to as a decision engine (DE). The decision engine is capable of dynamically allocating or freeing servers and of starting new or stopping

10 running instances of different applications on these servers. Usually the control means is hosted on a computer within the network that is neither server nor client but specifically designated for administrative tasks.

15 Decisions of the control means are based on a number of criteria and basic conditions, like actual demands, the capacity that is theoretically available and the actual capacity which might be decreased due to failures. Also, non-technical but economical considerations, for example

20 different priorities of customers, might be taken into account.

For the control means to meet all these demands, it has to be provided with the appropriate information. For one, it is a

25 necessity for the control means to have detailed knowledge of all servers which are to be controlled. Before a server can be controlled by a control means it has to be registered with the control means. In this registration procedure all relevant configuration information of the server is made

30 known to the control means and kept in an appropriate registration data structure of the control means. The control means usually also comprises means to receive or collect status information of all registered servers. From the status

information the control can, for example, extract load data, i.e. information on capacity utilization.

Based on an evaluation of the status information and of the
5 information kept in the registration data structure and
possibly based on data from other sources (for example data
provided by an additional network load balancer of the
computer arrangement) tasks can be distributed between
servers and additional applications can be started or stopped
10 on servers or servers can be assigned to certain jobs or to
certain customers. Also a failure of a server, a "crash",
will become apparent in the status information. To ensure
continued provision of services the control means can be set
up to reboot the crashed server.

15 This results in a computer arrangement which is autonomous
and, to some extent, self-healing. This autonomous approach
breaks down if a server failure occurs on a regular basis due
to an inherent problem which can only be solved by human
20 interference. According to prior art it would in this case be
necessary for an administrator to deregister the problematic
server from the control means to gain manual control over the
server for maintenance and/or trouble-shooting. Having
finished the maintenance, the registration procedure would
25 have to be repeated to incorporate the server again into the
pool of servers controlled by the control means. This method
is time-consuming and error-prone. Also, if a server is de-
registered during maintenance all logging of status
information by the control means is stopped, although logging
30 could be useful for the maintenance or trouble-shooting
process.

It is therefore an objective of the present invention to present a computer arrangement for providing services for clients over a network based on automatically controlled and registered servers which allows for an easy access to the servers for maintenance and/or trouble-shooting purposes. It is another objective to present a method for providing services for clients over a network.

In order to accomplish the above object, a computer arrangement and a method for providing services for clients over a network of the aforementioned type is described where a service mode is provided for each server by the control means, said service mode is characterized in that the status information of a server in service mode is not evaluated, no instances of application programs are started or stopped on a server in service mode and the configuration information of a server in service mode is preserved in the registration data structure.

The service mode allows to easily access a server for maintenance and/or trouble-shooting without any interference from the control means, while all configuration information of the server is preserved in the registration data structure. Control over the server can be handed back to the control means afterwards quickly and without the need to re-register the server and transfer all configuration information for a second time.

According to one embodiment of the invention the service mode is entered upon a user request. This would be the preferred embodiment for user scheduled maintenance purposes.

In another embodiment of the present invention the service mode is entered if an evaluation of the status information of a server reveals that the server does not operate correctly. An advantage of this embodiment is that servers are not automatically re-booted which might override valuable debugging information. This is of particular interest for computer arrangements which are set up for the first time or which had to undergo major configuration changes.

Other features which are considered as characteristic for the invention or which describe advantageous embodiments of the present invention are set forth in the dependent claims.

The above and other objects, features and advantages of the present invention will become apparent from the following description in conjunction with the accompanying drawings.

In the drawings:

Fig. 1 is a conceptual diagram of one embodiment of a computer arrangement according to the invention and

Fig. 2 is a flow-diagram of one embodiment of a method for providing services for clients over a network according to the invention.

Fig. 1 shows several servers 1 (1a, 1b, 1c, 1d) which are set up to communicate to clients 2 over a network 3. Each server 1 could host one or more applications to provide services for clients 2. A control means 4 is provided which comprises receiving means 5, a registration data structure 7 and an evaluation unit 9. The receiving means 5 is connected to the servers 1 by status lines 6 for receiving status information.

Two manual input devices 8, 10 are connected to the control means. In particular, input device 8 is coupled to the registration data structure 7 and input device 10 to the evaluation unit 9. The evaluation unit 9 comprises service mode flags 11 which have an influence on the information evaluated by the evaluation unit 9. The output of the evaluation unit 9 is connected to the servers 1 by control lines 12.

- 10 The configuration shown in Fig. 1 is typical for a server farm where a provider runs a plurality of servers 1. Often, several tens or even hundreds of blade servers, named after their geometric outline, are used in server farms. However, the invention is not restricted to a situation where each
- 15 server 1 is an independent physical unit. The invention could e.g. also be applied to virtual servers. Virtual servers share a common hardware which runs a software that emulates several hardware units.
- 20 Providing services that have a fast response time on the one hand, and that do not waste any server capacity on the other hand, requires that an appropriate number of instances of an application 4 is running and, accordingly, that an appropriate number of servers 1 is dedicated to host these
- 25 instances. Provisioning of servers 1 and/or configuring servers 1 is performed by the control means 4. The control means 4 may be software-implemented and hosted on a separate computer. For reasons of high availability and redundancy there may also be a second control means that could take over
- 30 in case of a failure of the first control means 4.

Different methods to provision, control and configure servers 1 by the control means 4 are known. One way is to start

remotely or stop instances of an application depending on requirements. This bears the danger of collecting remains of all applications that were once started and stopped again (like temporary files) which might lead to an unstable and slow system. Therefore, often another approach, sometimes called "bare-metal provisioning", is chosen, where for each new task (different customer, different applications) a whole boot image is transferred to a server and the server is re-booted with the new boot image. In the following

5 "controlling" a server is meant to comprise all methods that influence applications running on the server.

10

The control lines 12 between the control means 4 and the servers 1 which are used to provision and/or control servers 1 could form an independent network for security and/or performance reasons. In an other embodiment, the same network 3 that connects the servers 1 and the clients 2 could also be used to control the servers 1. They same holds true for status lines 6.

15

Each server 1 transmits status information over the status lines 6 to the control means 4. The status information is received by the receiving means 5, where it is buffered and where it might be logged, e.g. in the form of logging-files.

20

25 Status information is all information concerning the actual state of operation of a server 1, e.g. load data, performance data or a signal indicating proper functioning (alive-signal). The receiving means 5 could also perform a pre-evaluation of the status information, for example sorting or

30 filtering the data with respect to its importance or for clarity of the logging files.

The status information is then passed on to the evaluation unit 9. Evaluation unit 9 evaluates this status information and controls the servers 1 according to the result of the evaluation. Apart from the status information, additional data describing the actual state of operation could be used in the evaluation process. Such data could for example be provided by a network load balancer which might be used within the system to rout requests of clients 2 to the least busiest server 1. To be able to do so, network load balancers usually keep track of the utilization of the servers 1 and could share their information on this matter with the control means 4.

Despite dynamic data describing the actual state of operation, configuration information of all available servers 1 is needed by the evaluation unit 9 to base its control decisions on. Configuration information comprises all static information on the servers 1, e.g. hardware information like the processor type, memory capacity, network addresses, theoretical performance data. The configuration information is kept in the registration data structure 7. It has to be provided by a system administrator in a registration (or setup) procedure for a server 1 to be used within the computer arrangement and to be controlled automatically by the control means 4. By way of example, the registration procedure is carried out using the manual input 8.

Methods for evaluating status information, taking configuration information into consideration and making decisions for controlling the servers 1 to provide services for clients 2 are known from the art and will therefore not be described in further detail here.

According to the invention, a service mode is provided for each server 1 controlled by the control means 4. The service mode is characterized in that a server 1 which is set to service mode is not controlled by the control means 4, nonetheless the server 1 remains registered in the registration data structure 7 and status information of this server 1 are still received by the receiving means 5 of the control means 4. To achieve this, the control means 4 comprises service flags 11 - a data structure indicating which server 1 is in service mode. The evaluation unit 9 is set up to only use information of servers 1 which are not in service mode. In the embodiment shown, neither status information nor configuration information of servers 1 in service mode are transmitted to the evaluation unit 9 and consequently, no control commands concerning those servers 1 are issued by the control means 4. In the drawing, this is indicated by crossed-out control lines 12.

Two different ways to set a server 1 into service mode are shown in Fig. 1. One way is to manually set the service mode by using the manual input 10 which is connected to the control unit 4. The manual option is e.g. used for a planned or scheduled maintenance session for a server 1, during which any disturbance due to an automatic control by the control unit 4 would be detrimental. The manual inputs 8 and 10 can of course be integrated into one common administration interface.

Another way is that a server 1 is automatically set into service mode under certain circumstances which mainly relate to the failure of servers 1. To achieve this, the receiving means 5 is set up to detect if a server 1 does not operate correctly any more. There are different methods known to

monitor a server 1 via status lines 6 or a network in general. The server 1 could e.g. periodically send a message, called alive-signal. If the alive-signal is not received by the receiving means 5 for a certain period of time, this could be considered as an indication for the server 1 not operating correctly or being "crashed". In the example in Fig. 1, the topmost server 1a is crashed as indicated by its dashed outline. After having detected the crash, the receiving means 5 requests service mode for this server 1a.

10

The requests for service mode of the manual input 10 and of the receiving means 5 are combined to set the service mode flags 11 in such a way, that a server can be set into service mode by means of the manual input (manually) and/or the receiving means (automatically), but that the service mode can be switched off manually only. In the example, server 1a is set to service mode automatically and server 1c is set to service mode manually. Once being set to service mode, the crashed servers 1a and 1c are excluded from any automatic control, they will not be controlled by the control means 4, for example not be re-booted, which would otherwise overwrite valuable debugging information.

20

Different strategies when to set a crashed server into service mode are possible and useful. If a computer arrangement is set up for the first time or had been subject to major configuration changes, it is sensible to set a server into service mode every time a crash is detected. If, instead, a computer arrangement is running more or less stable and servers do merely crash on the odd occasion, it might be useful to set a server into service mode only if it crashes twice within a given time span. A single crash once in a while would not disturb the system performance

25

30

significantly and an automatic re-boot or a restart of applications initiated by the control means 4 would keep the system operable without any human interference.

- 5 Fig. 2 illustrates the method for providing services for clients over a network according as used in the embodiment described in connection with Fig. 1 in a flowchart representation.
- 10 In step A of the method configuration information of servers 1 are retrieved from the manual input 8 and stored in the registration data structure 7. In step B requests for manually setting a server 1 into service mode are read from the manual input 10, while requests for automatically setting
- 15 a server into service mode are gathered from the receiving unit 5. The requests are combined and written to the service mode flags 11 in step D. The combination is such that the service mode can be set manually and automatically, but reset manually only. In another embodiment an automatic reset
- 20 might be possible if the receiving means determines unambiguously that a server operates correctly again. Then, actual status information is logged for all servers 1 in step E. An automatic control by the control means 4 in step F, in contrast, is restricted to servers 1 which are not set to
- 25 service mode by using the information of the service flags 11. The method then starts again with steps B and C.

List of reference signs:

	1	server
	2	client
5	3	network
	4	control means
	5	receiving means
	6	status lines
	7	registration data structure
10	8	manual input
	9	evaluation unit
	10	manual input
	11	service mode flags
	12	control lines

Claims:

1. A computer arrangement for providing services for clients
(2) over a network (3), comprising
- 5 - at least one server (1) hosting one or more applications
 which provide services to clients (2) over the network
 (3),
- control means (4), comprising
- a registration data structure (7) containing
- 10 configuration information of servers (1), and
- means to receive (5) status information of all
 registered servers (1),
- said control means (4) being set up to control
 registered servers (1) by starting or stopping instances
- 15 of said application programs based on an evaluation of
 the received status information and on the configuration
 information,
- where a service mode is provided for each server (1) by
 the control means (4), said service mode is characterized
- 20 in that
- the status information of a server (1) in service mode is
 not evaluated,
- no instances of application programs are started or
 stopped on a server (1) in service mode and
- 25 - the configuration information of a server (1) in service
 mode is preserved in the registration data structure
 (7).
2. The computer arrangement of claim 1, where
- 30 the service mode for a server (1) is entered upon a user
 request.
3. The computer arrangement of claim 1, where

the service mode for a server (1) is entered if the evaluation of the status information of a server (1) reveals that the server (1) does not operate correctly.

- 5 4. The computer arrangement of claim 1, where the service mode for a server (1) is entered if the evaluation of the status information of a server reveals that the server (1) did not operate correctly for the second time within a predetermined period of time.

10

5. The computer arrangement of one of the claims 1 to 4, where the service mode for a server (1) is exited upon a user request.

- 15 6. The computer arrangement of one of the claims 1 to 5, where the control means (4) comprises means to log received status information of all registered servers (1).

7. A method for providing services for clients (2) over a network (3) with the following steps:
- 20 - providing a computer arrangement comprising at least one server (1) hosting one or more application programs which provide services to clients (2) over the network (3), and a control means (4),
- 25 - register a server (1) with the control means (4) by adding configuration information of the server (1) to a registration data structure (7) provided by the control means (4),
- indicate to the control means (4) that one or more
- 30 server (1) are in a service mode,
- receiving status information of all registered servers (1),

- evaluate the status information for all servers (1) which are not set to service mode,
- control all registered servers (1) which are not in service mode based on the evaluation of the status information.

5

8. The method according to claim 7, where the control unit (4) controls a registered server (1) by starting and stopping instances of application programs.

10

9. The method according to claim 7, where the control unit (4) controls a registered server (1) by booting the server (1).

15 10. The method according to one of the claims 7 to 9, where a user indicates that a server (1) is in service mode.

11. The method according to one of the claims 7 to 10, where the control means (4) sets a server (1) into service mode if the evaluation of the status information of the server (1) reveals that the server (1) does not operate correctly.

20

12. The method according to one of the claims 7 to 10, where the control means (4) sets a server (1) into service mode if the evaluation of the status information of the server (1) reveals that the server (1) did not operate correctly for the second time within a predetermined period of time.

25

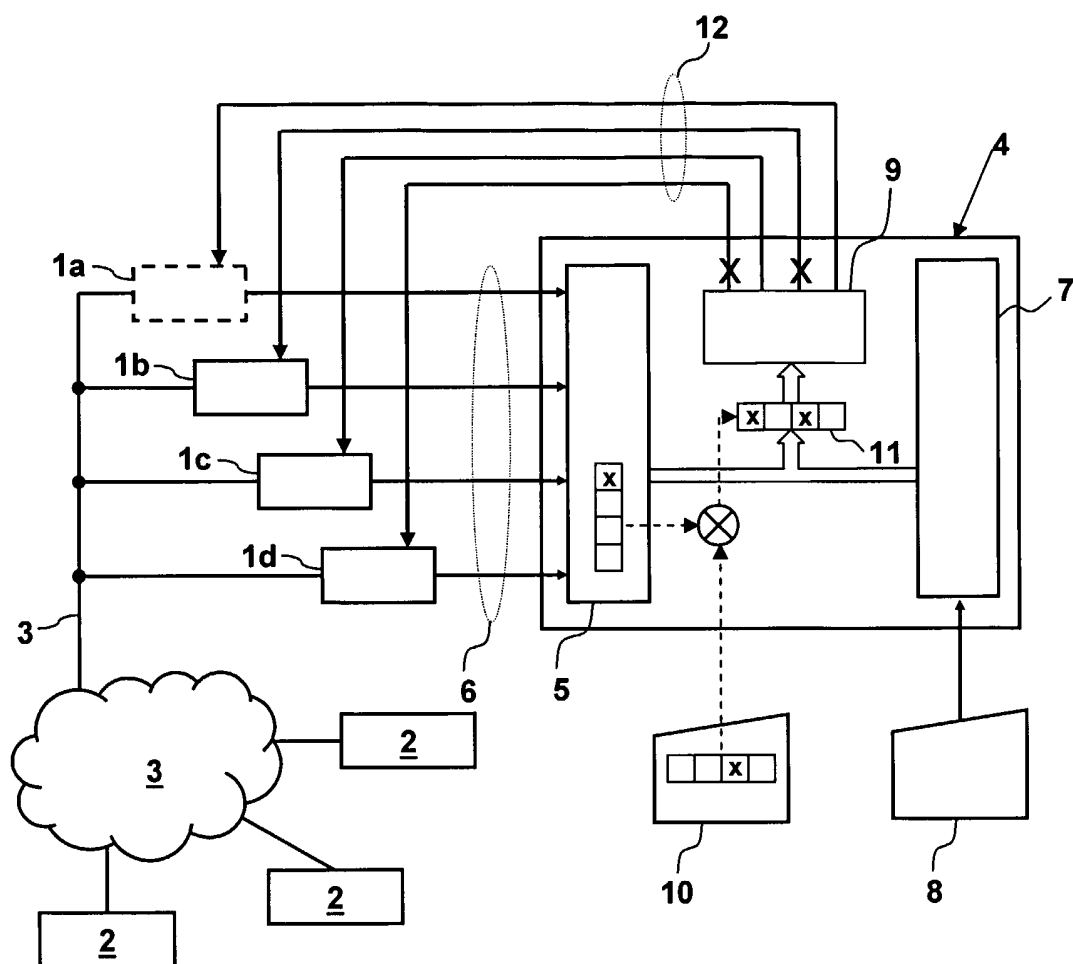
Fig. 1

Fig. 2