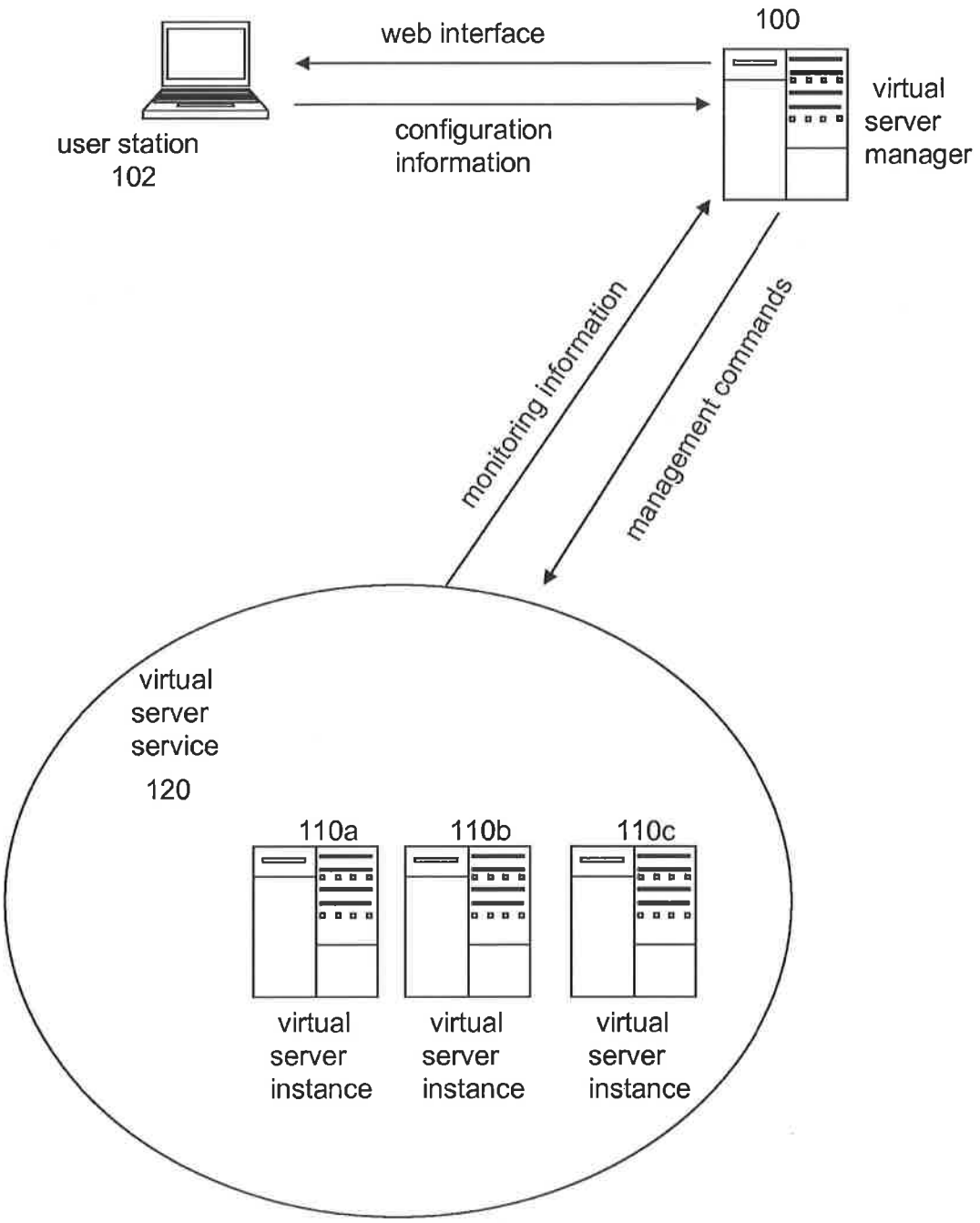

100
web interface
configuration information
user station 102
virtual server manager
monitoring information
management commands
virtual server service 120
110a
110b
110c
virtual server instance
virtual server instance
virtual server instance

Fig. 1A

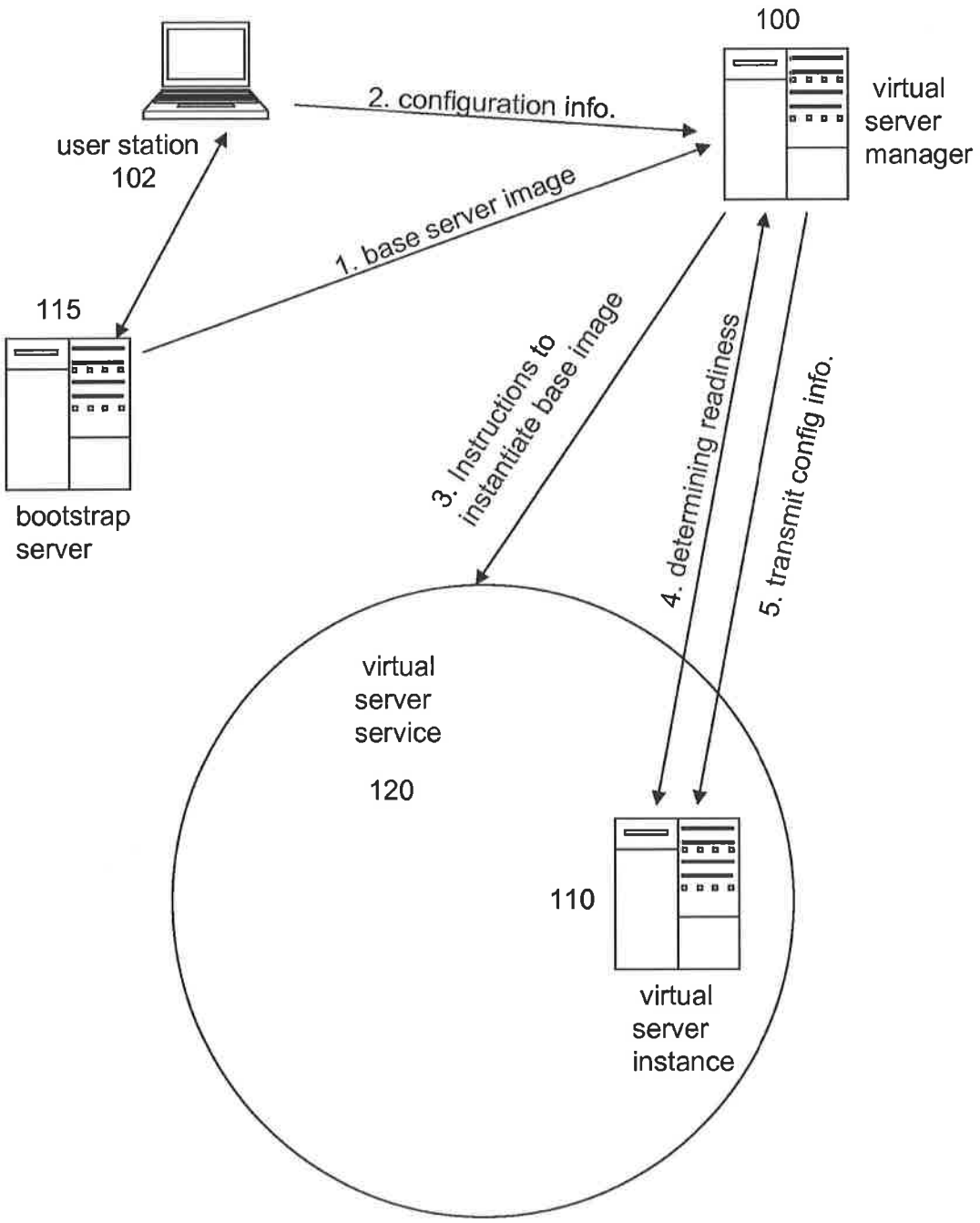

100
virtual server manager
2. configuration info.
user station 102
1. base server image
115
bootstrap server
3. Instructions to instantiate base image
4. determining readiness
5. transmit config info.
virtual server service
120
110
virtual server instance

Fig. 1B

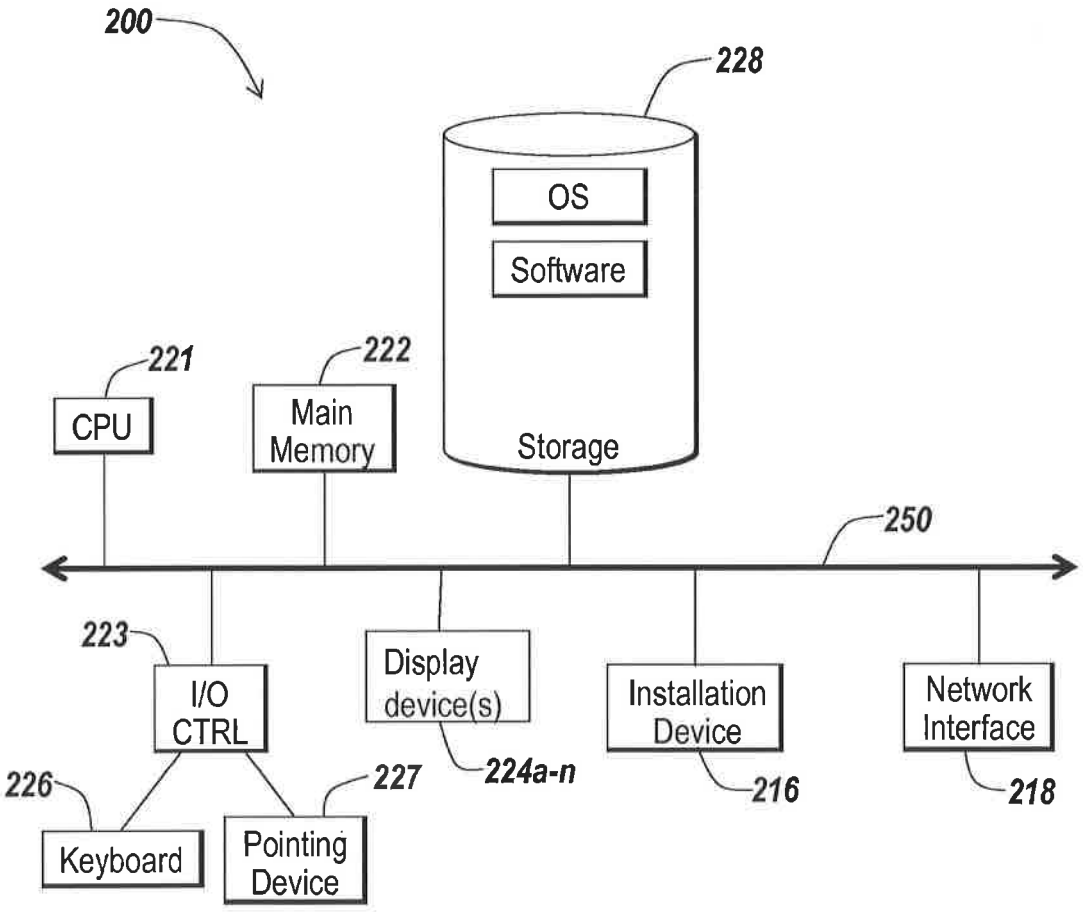

200
228
OS
Software
Storage
221
CPU
222
Main Memory
250
223
I/O CTRL
Display device(s)
224a-n
Installation Device
216
Network Interface
218
226
Keyboard
227
Pointing Device

Fig. 2A

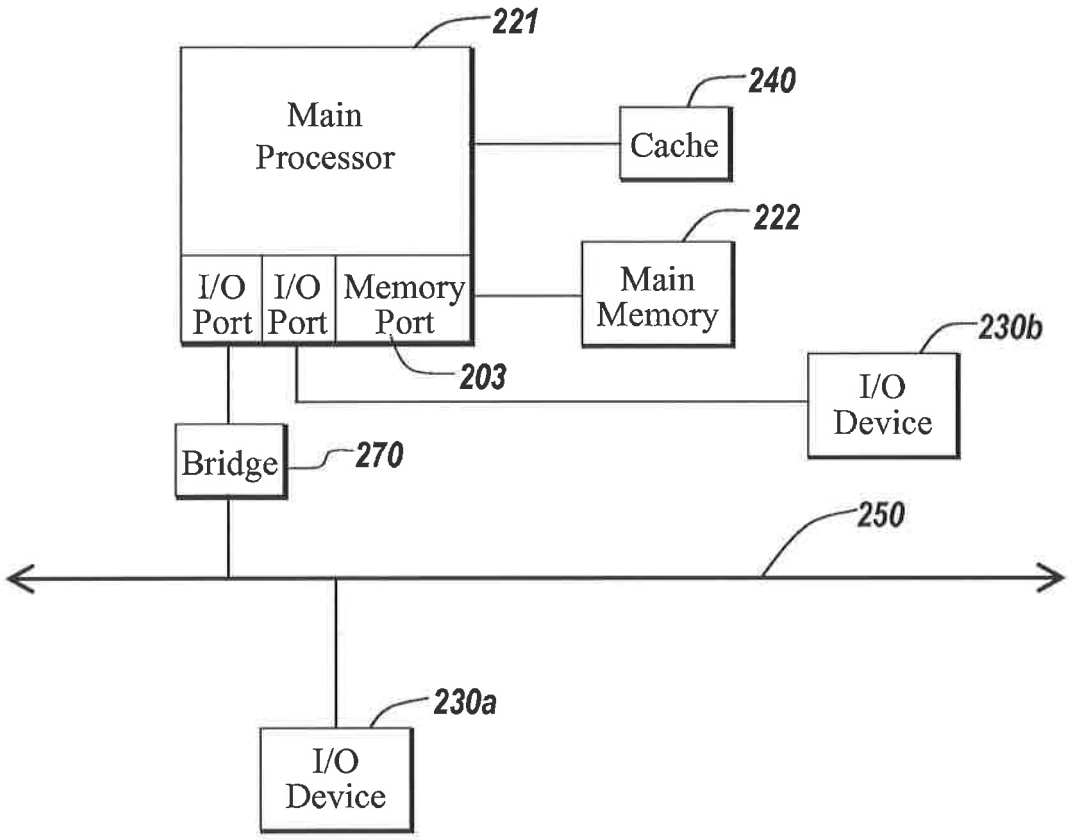

221
Main Processor
240
Cache
222
Main Memory
I/O Port
I/O Port
Memory Port
203
230b
I/O Device
Bridge
270
250
230a
I/O Device

Fig. 2B

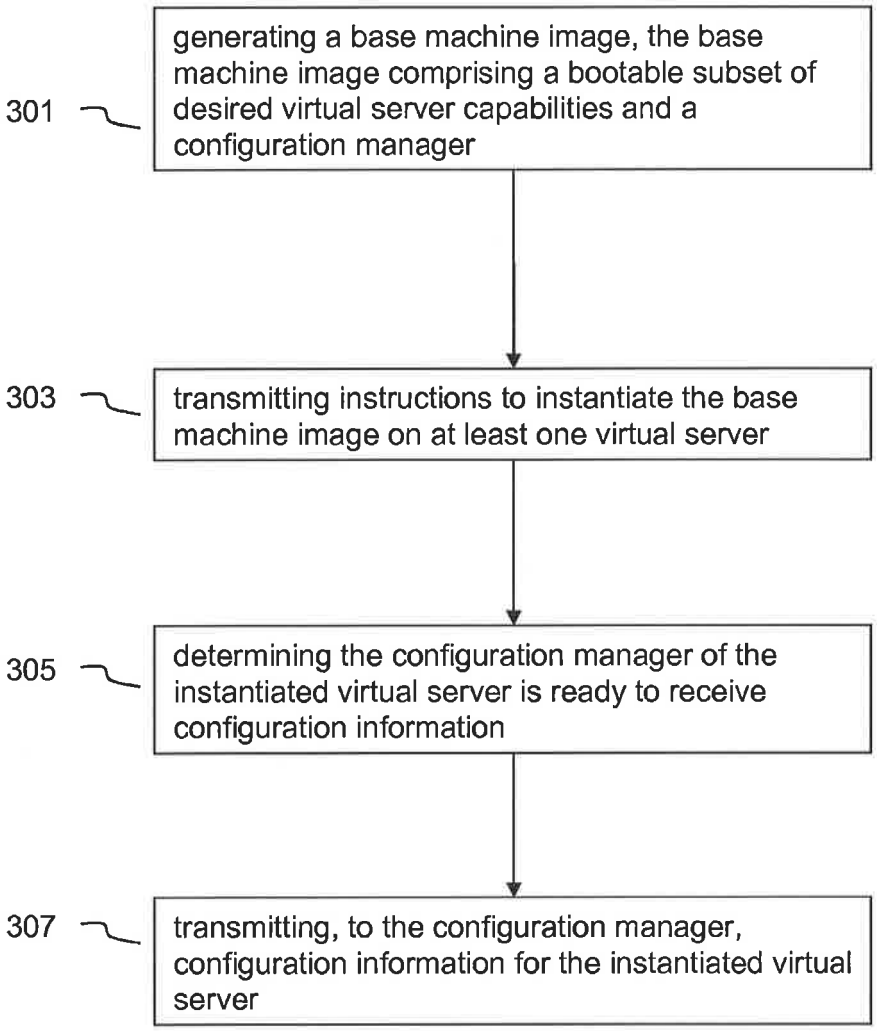

301
generating a base machine image, the base machine image comprising a bootable subset of desired virtual server capabilities and a configuration manager
303
transmitting instructions to instantiate the base machine image on at least one virtual server
305
determining the configuration manager of the instantiated virtual server is ready to receive configuration information
307
transmitting, to the configuration manager, configuration information for the instantiated virtual server

Fig. 3

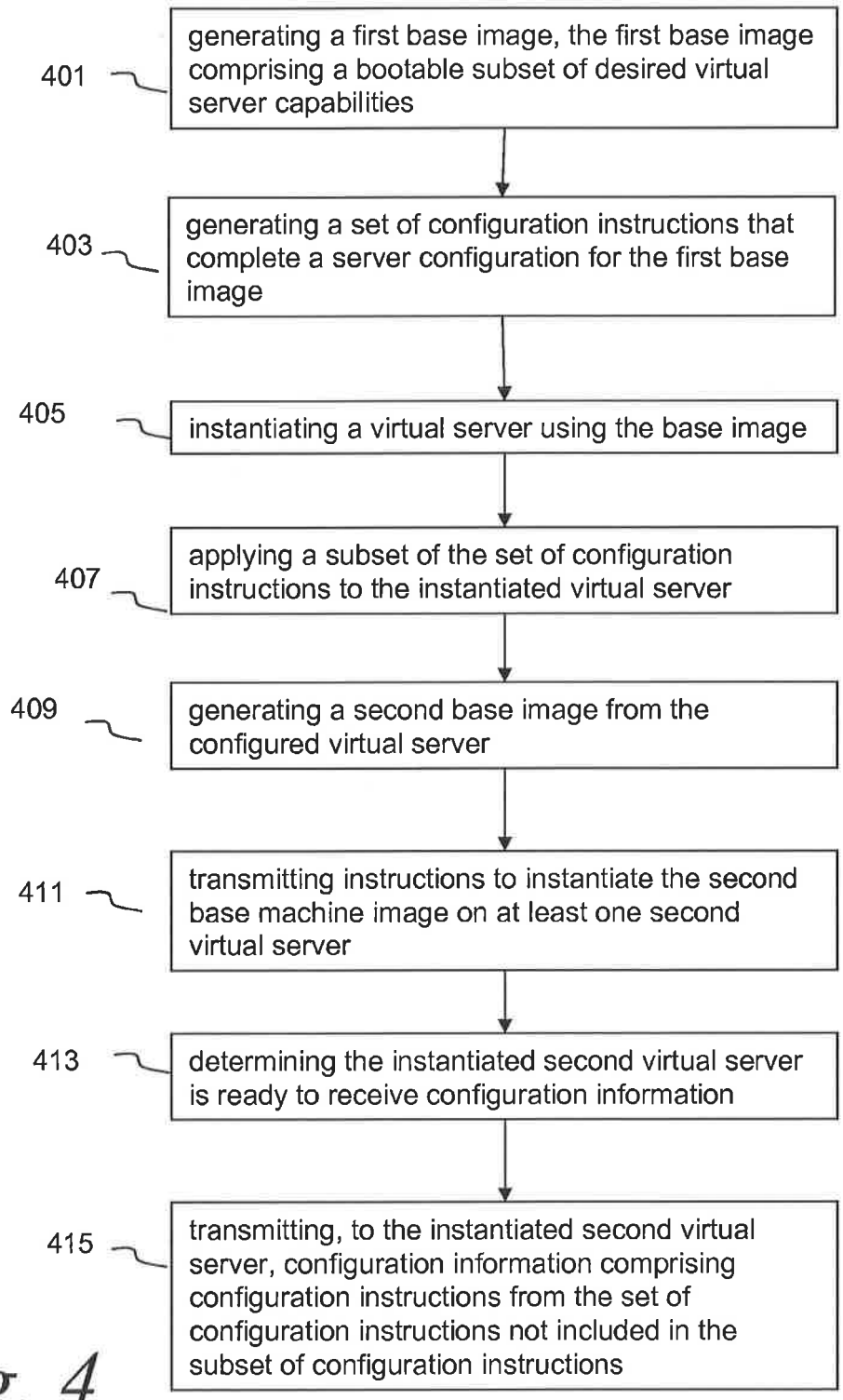
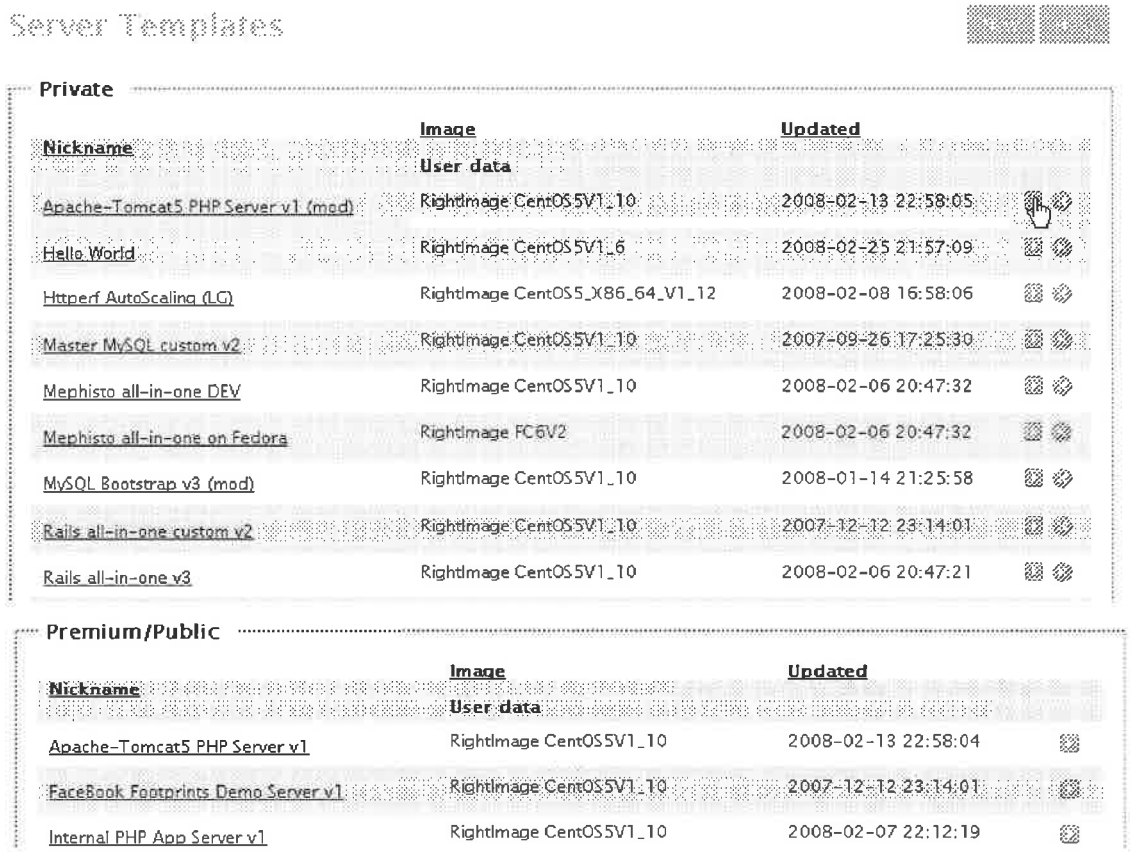

401
generating a first base image, the first base image comprising a bootable subset of desired virtual server capabilities
403
generating a set of configuration instructions that complete a server configuration for the first base image
405
instantiating a virtual server using the base image
407
applying a subset of the set of configuration instructions to the instantiated virtual server
409
generating a second base image from the configured virtual server
411
transmitting instructions to instantiate the second base machine image on at least one second virtual server
413
determining the instantiated second virtual server is ready to receive configuration information
415
transmitting, to the instantiated second virtual server, configuration information comprising configuration instructions from the set of configuration instructions not included in the subset of configuration instructions

Fig. 4

Fig. 5A


Server Templates
Private
Nickname
Image
Updated
User data
Apache-Tomcat5 PHP Server v1 (mod) RightImage CentOS5V1_10 2008-02-13 22:58:05
Hello World RightImage CentOS5V1_6 2008-02-25 21:57:09
Httperf AutoScaling (LG) RightImage CentOS5_X86_64_V1_12 2008-02-08 16:58:06
Master MySQL custom v2 RightImage CentOS5V1_10 2007-09-26 17:25:30
Mephisto all-in-one DEV RightImage CentOS5V1_10 2008-02-06 20:47:32
Mephisto all-in-one on Fedora RightImage FC6V2 2008-02-06 20:47:32
MySQL Bootstrap v3 (mod) RightImage CentOS5V1_10 2008-01-14 21:25:58
Rails all-in-one custom v2 RightImage CentOS5V1_10 2007-12-12 23:14:01
Rails all-in-one v3 RightImage CentOS5V1_10 2008-02-06 20:47:21
Premium/Public
Nickname
Image
Updated
User data
Apache-Tomcat5 PHP Server v1 RightImage CentOS5V1_10 2008-02-13 22:58:04
FaceBook Footprints Demo Server v1 RightImage CentOS5V1_10 2007-12-12 23:14:01
Internal PHP App Server v1 RightImage CentOS5V1_10 2008-02-07 22:12:19

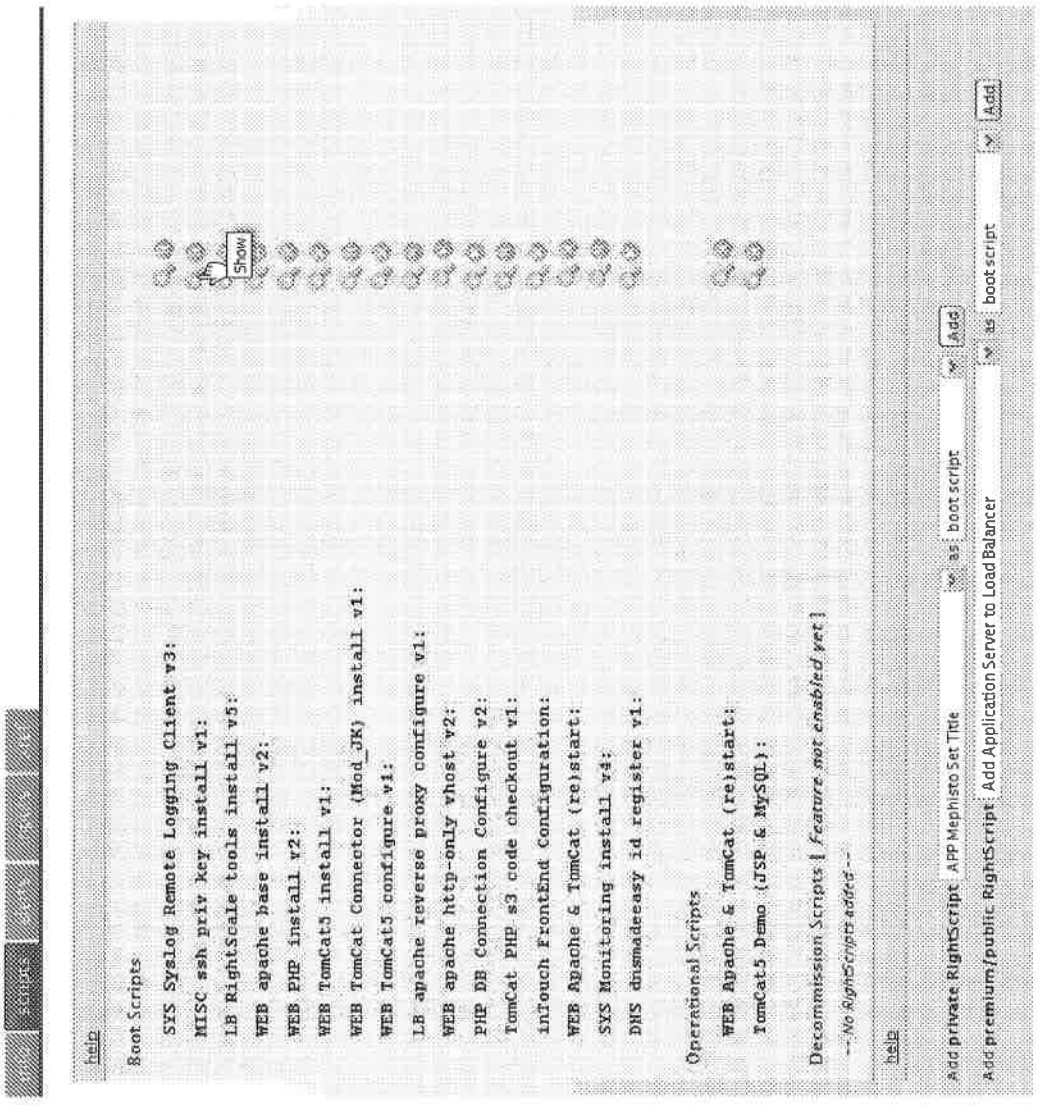

help
Boot Scripts
SYS Syslog Remote Logging Client v3:
MISC ssh priv key install v1:
LB RightScale tools install v5:
WEB apache base install v2:
WEB PHP install v2:
WEB TomCat5 install v1:
WEB TomCat Connector (Mod_JK) install v1:
WEB TomCat5 configure v1:
LB apache reverse proxy configure v1:
WEB apache http-only vhost v2:
PHP DB Connection Configure v2:
TomCat PHP s3 code checkout v1:
inTouch FrontEnd Configuration:
WEB Apache & TomCat (re)start:
SYS Monitoring install v4:
DNS dnsmadeeasy id register v1:
Show
Operational Scripts
WEB Apache & TomCat (re)start:
TomCat5 Demo (JSP & MySQL):
Decommission Scripts [feature not enabled yet]
-- No RightScripts added --
help
Add private RightScript APP Mephisto Set Title as boot script Add
Add premium/public RightScript Add Application Server to Load Balancer as boot script Add

Fig. 5B

Fig. 5C

Alerts
Alert Spec
Condition
Derived from
Escalation
Enabled
Age
Actions
scale-up at 50% busy
if cpu-0/cpu-idle.value < '50' for 3 min then escalate to 'scale-up'.
-
yes
-
scale-down at 20% busy
if cpu-0/cpu-idle.value > '80' for 10 min then escalate to 'scale-down'.
-
scale-down
yes
5m
stranded
if EC2/server-failure.state == 'stranded' for 1 min then escalate to 'critical'.
-
critical
yes
-
rs instance terminated
if EC2/server.state == 'terminated' for 1 min then escalate to 'critical'.
Rails FrontEnd v3
critical
yes
-
rs instance stranded
if EC2/server-failure.state == 'stranded' for 1 min then escalate to 'warning'.
Rails FrontEnd v3
warning
yes
-
rs instance not responding
if cpu-0/cpu-idle.value == 'NaN' for 1440 min then escalate to 'warning'.
Rails FrontEnd v3
warning
yes
-
rs cpu busy
if cpu-0/cpu-idle.value < '15' for 60 min then escalate to 'warning'.
Rails FrontEnd v3
warning
yes
-

Fig. 5D

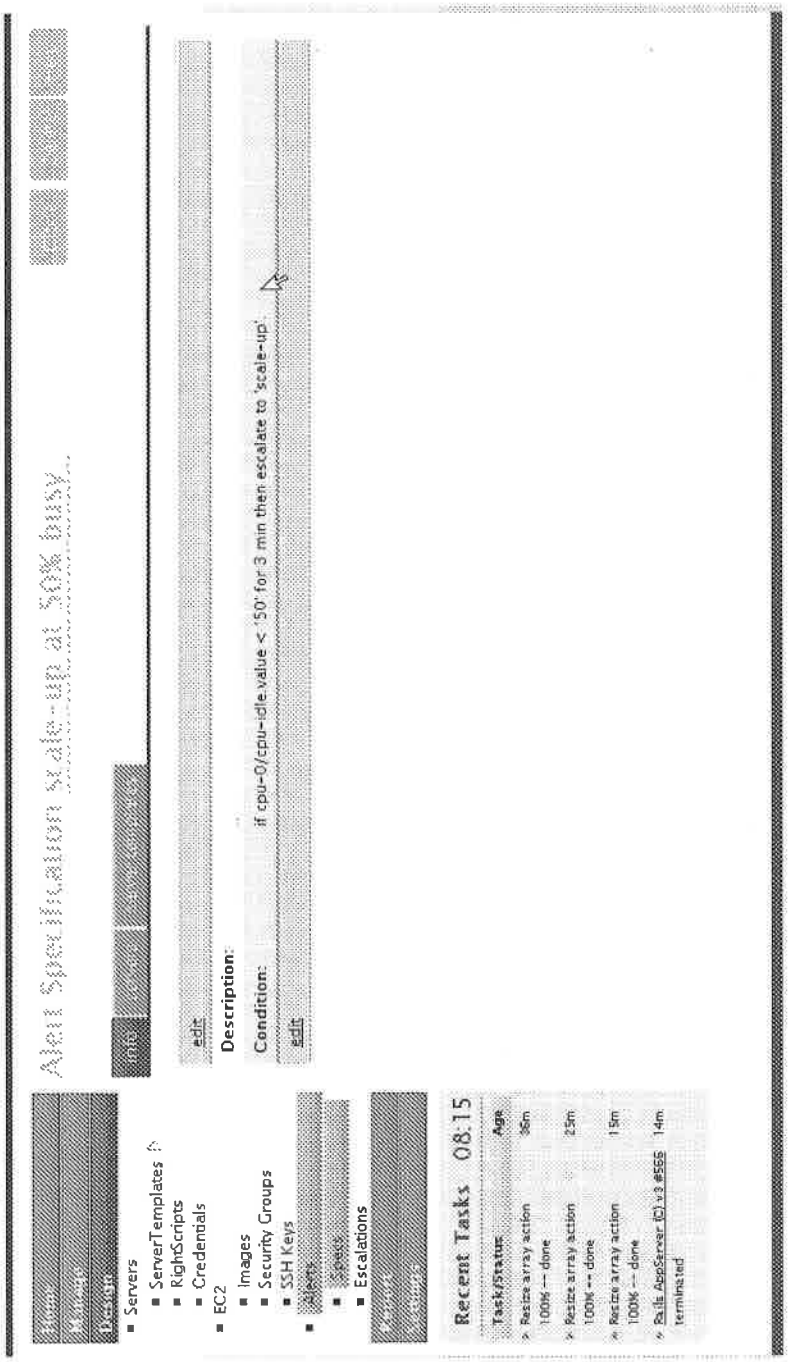

Alert Specification scale-up at 50% busy
info
edit
Description:
Condition:
if cpu-0/cpu-idle.value < 50 for 3 min then escalate to 'scale-up'
edit
Servers
ServerTemplates
RightScripts
Credentials
EC2
Images
Security Groups
SSH Keys
Alerts
Spots
Escalations
Recent Tasks 08:15
Task/Status
Age
Resize array action
100% -- done
35m
Resize array action
100% -- done
25m
Resize array action
100% -- done
15m
Rails AppServer (0).v3 #565 terminated
14m

Fig. 5E

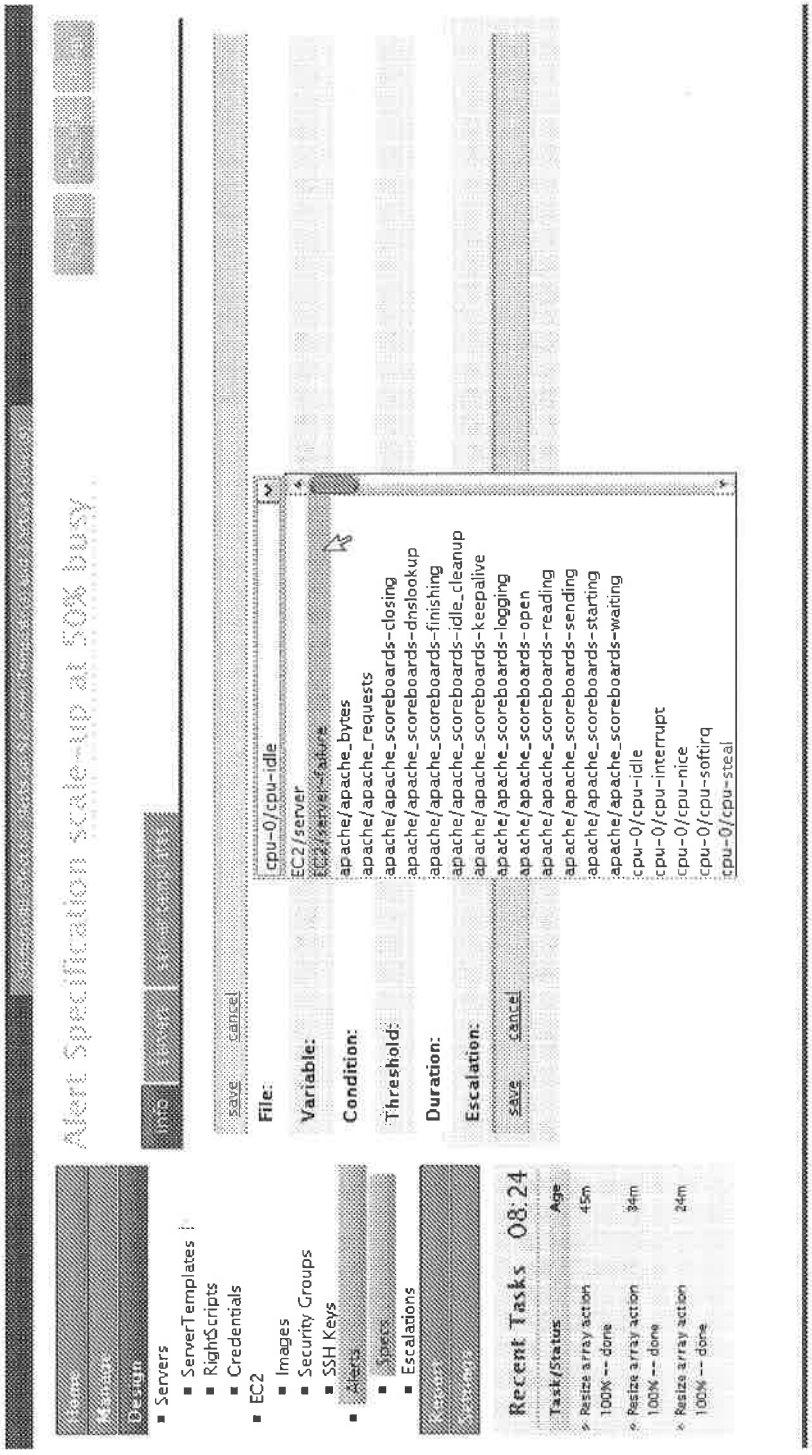

Alert Specification scale-up at 50% busy
Recent Tasks 08:24
Task/Status
Age

Fig. 5F

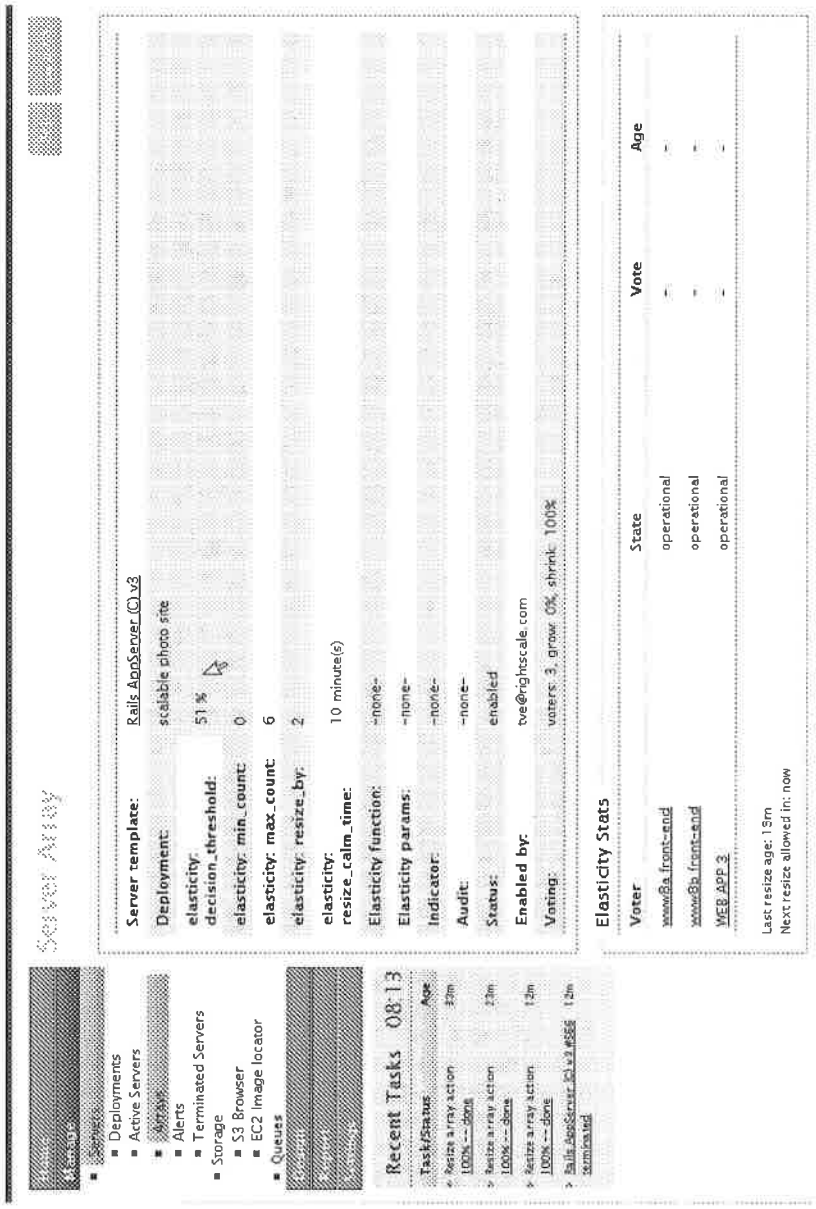
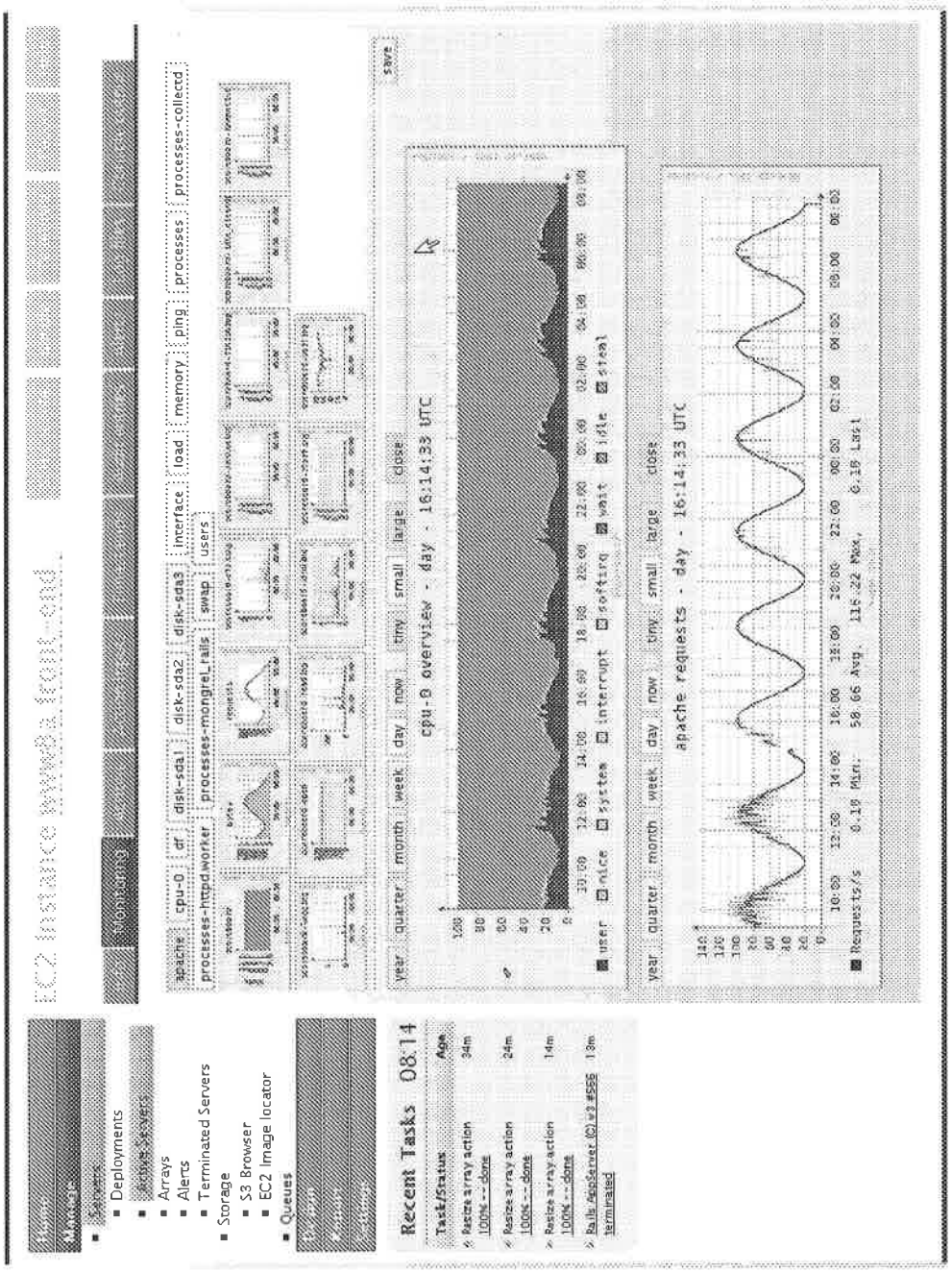

Server Array
Menu
Servers
Deployments
Active Servers
Arrays
Alerts
Terminated Servers
Storage
S3 Browser
EC2 Image locator
Queues
Recent Tasks 08:13
Task/Status
Age
Resize array action 100% -- done 33m
Resize array action 100% -- done 23m
Resize array action 100% -- done 12m
Rails AppServer (C) v3 #566 terminated 12m
Server template: Rails AppServer (C) v3
Deployment: scalable photo site
elasticity: decision_threshold: 51 %
elasticity: min_count: 0
elasticity: max_count: 6
elasticity: resize_by: 2
elasticity: resize_calm_time: 10 minute(s)
Elasticity function: -none-
Elasticity params: -none-
Indicator: -none-
Audit: -none-
Status: enabled
Enabled by: tve@rightscale.com
Voting: voters: 3, grow: 0%, shrink: 100%
Elasticity Stats
Voter
State
Vote
Age
www@a front-end operational - -
www@b front-end operational - -
WEB APP 3 operational - -
Last resize age: 15m
Next resize allowed in: now

Fig. 5G


EC2 instance myweb.trait.edu
Monitoring
Servers
Deployments
Arrays
Alerts
Terminated Servers
Storage
S3 Browser
EC2 image locator
Queues
Recent Tasks
Task/Status
Age
cpu-0 overview - day - 16:14:33 UTC
apache requests - day - 16:14:33 UTC

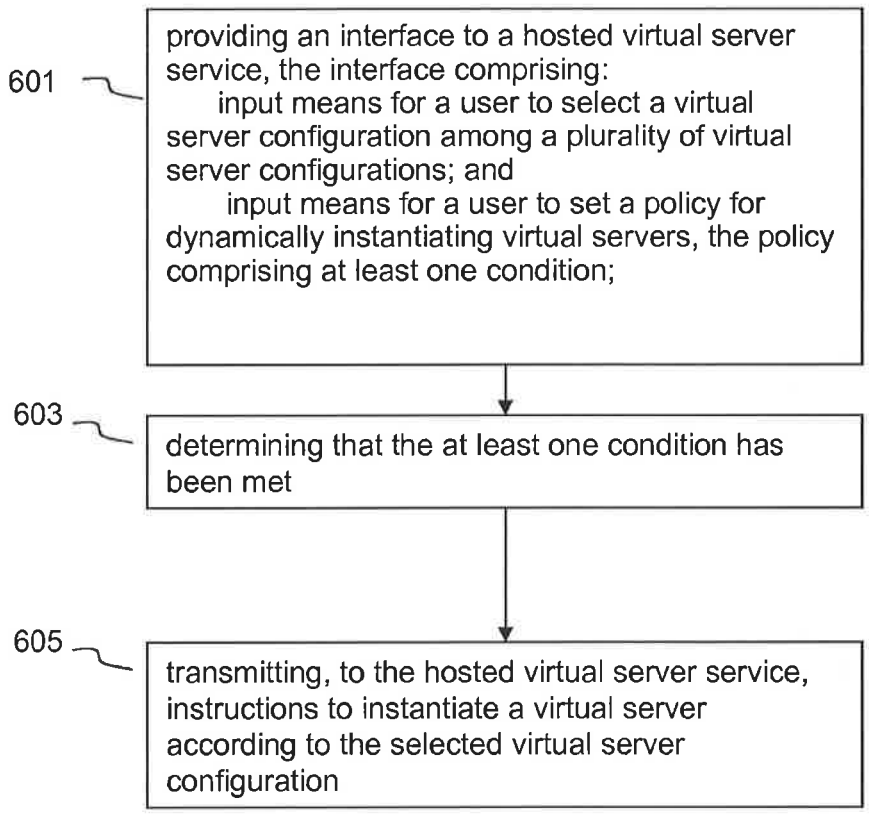

601
providing an interface to a hosted virtual server service, the interface comprising:
input means for a user to select a virtual server configuration among a plurality of virtual server configurations; and
input means for a user to set a policy for dynamically instantiating virtual servers, the policy comprising at least one condition;
603
determining that the at least one condition has been met
605
transmitting, to the hosted virtual server service, instructions to instantiate a virtual server according to the selected virtual server configuration

Fig. 6A

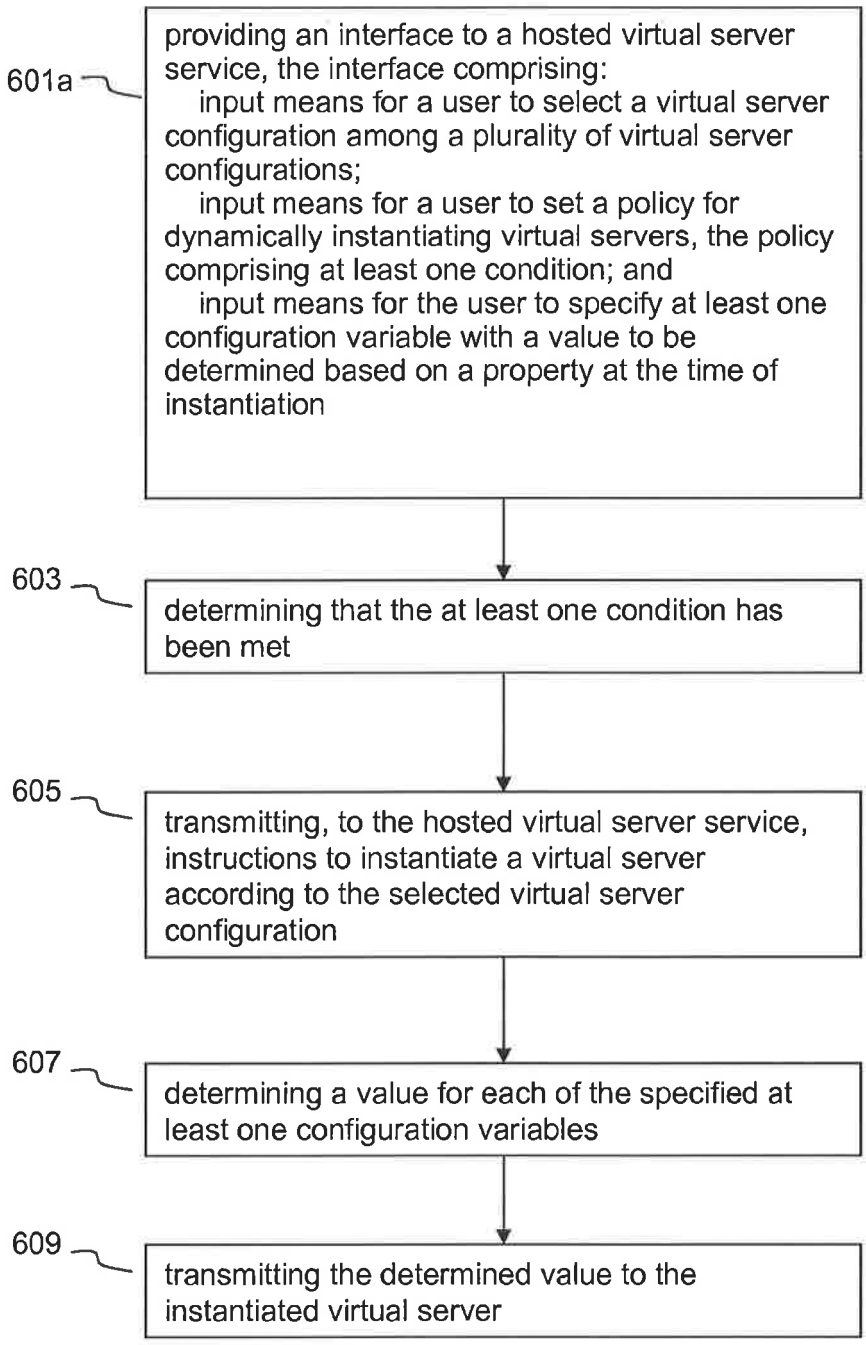

601a
providing an interface to a hosted virtual server service, the interface comprising:
input means for a user to select a virtual server configuration among a plurality of virtual server configurations;
input means for a user to set a policy for dynamically instantiating virtual servers, the policy comprising at least one condition; and
input means for the user to specify at least one configuration variable with a value to be determined based on a property at the time of instantiation
603
determining that the at least one condition has been met
605
transmitting, to the hosted virtual server service, instructions to instantiate a virtual server according to the selected virtual server configuration
607
determining a value for each of the specified at least one configuration variables
609
transmitting the determined value to the instantiated virtual server

Fig. 6B

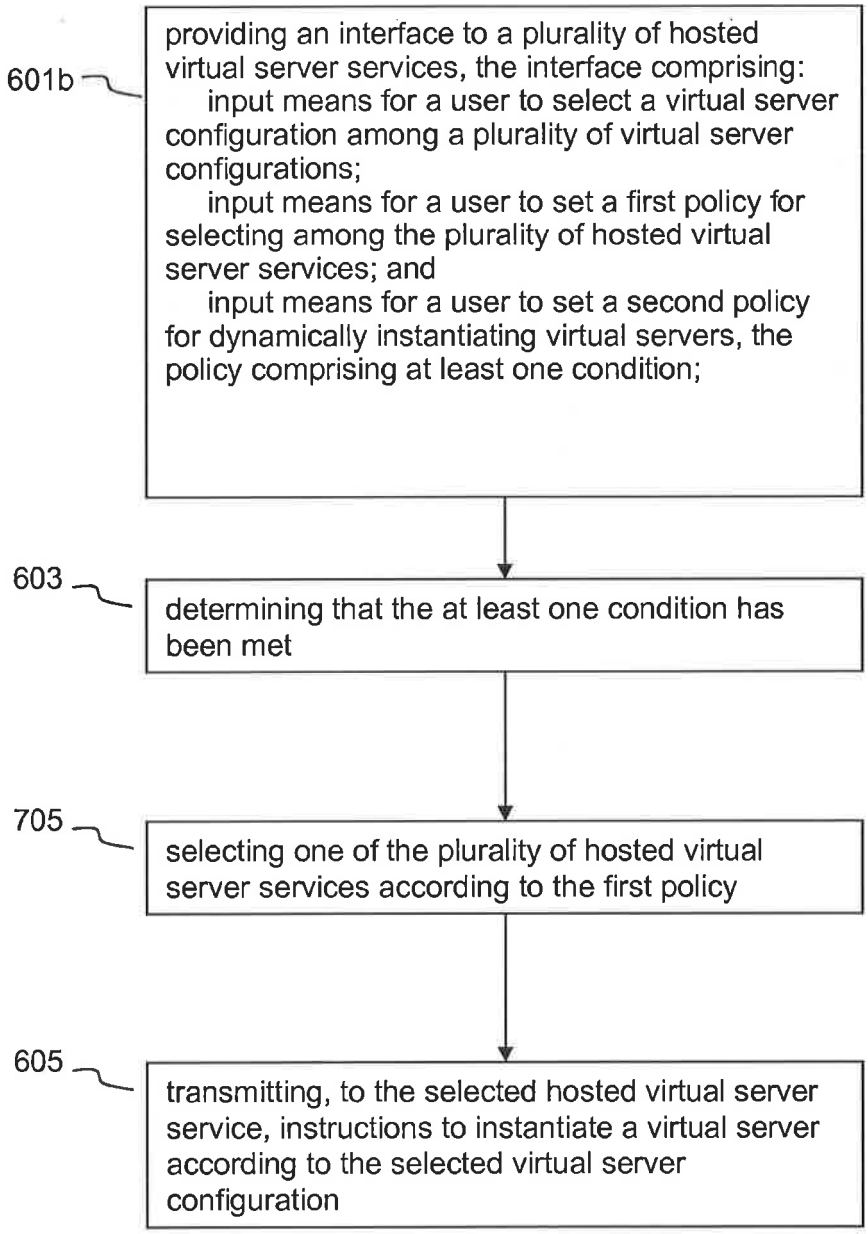

601b
providing an interface to a plurality of hosted virtual server services, the interface comprising:
input means for a user to select a virtual server configuration among a plurality of virtual server configurations;
input means for a user to set a first policy for selecting among the plurality of hosted virtual server services; and
input means for a user to set a second policy for dynamically instantiating virtual servers, the policy comprising at least one condition;
603
determining that the at least one condition has been met
705
selecting one of the plurality of hosted virtual server services according to the first policy
605
transmitting, to the selected hosted virtual server service, instructions to instantiate a virtual server according to the selected virtual server configuration

Fig. 7