



(86) Date de dépôt PCT/PCT Filing Date: 2011/06/27
(87) Date publication PCT/PCT Publication Date: 2012/01/12
(45) Date de délivrance/Issue Date: 2017/07/11
(85) Entrée phase nationale/National Entry: 2012/12/20
(86) N° demande PCT/PCT Application No.: US 2011/042050
(87) N° publication PCT/PCT Publication No.: 2012/006034
(30) Priorité/Priority: 2010/06/28 (US12/824,723)

(51) Cl.Int./Int.Cl. *H04L 12/24* (2006.01),
G06Q 10/06 (2012.01), *H04L 12/14* (2006.01)

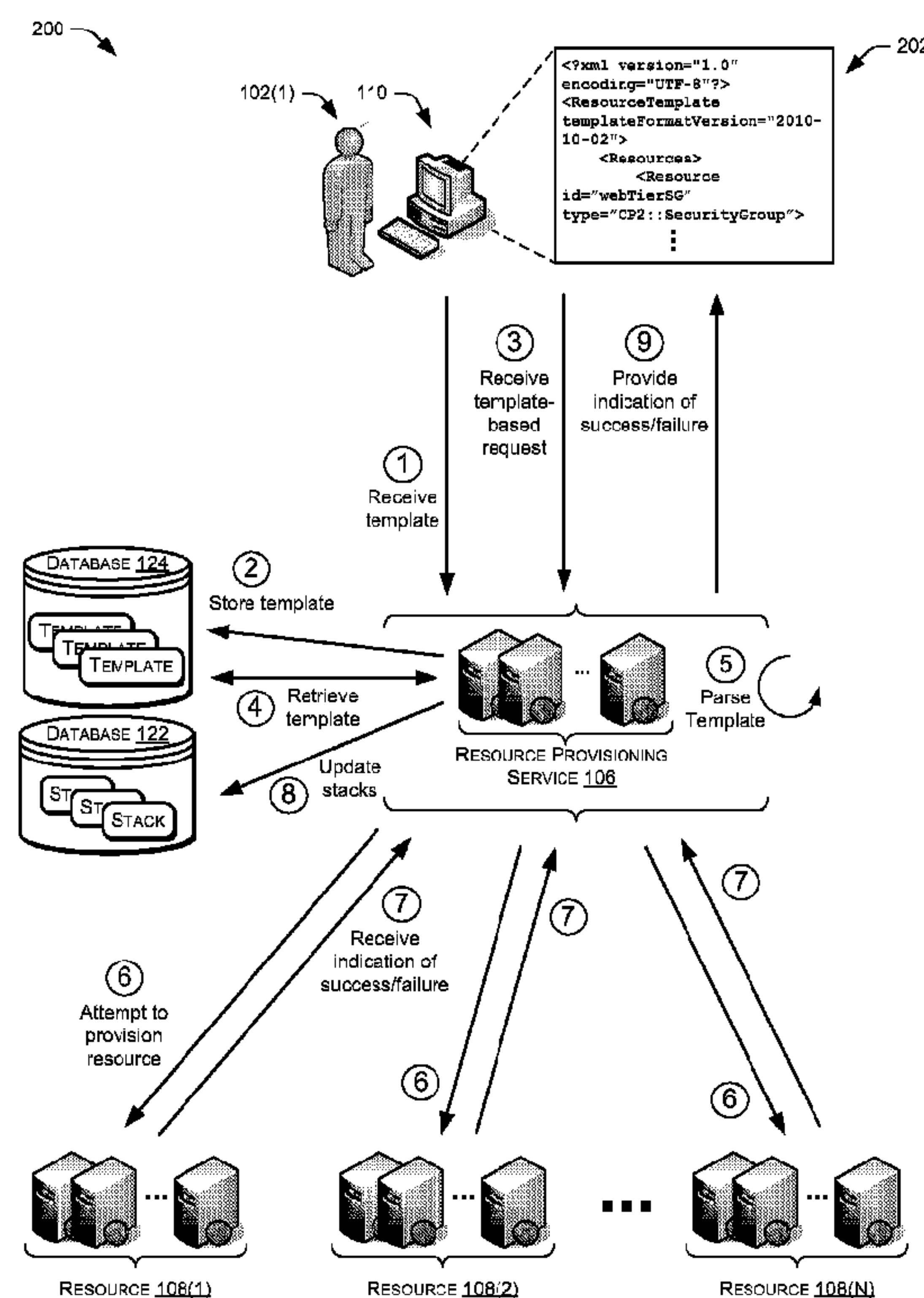
(72) Inventeurs/Inventors:
KRAMER, RETO, US;
OSIECKI, DANIEL LEE, US;
SHANKARAN, NISHANTH, US;
BALAKRISHNAN, VENKATES P., US;
PARE, GEOFFREY SCOTT, US;
MEIKE, BLAKE, US;
WHITAKER, CHRISTOPHER, US

(73) Propriétaire/Owner:
AMAZON TECHNOLOGIES, INC., US

(74) Agent: MARKS & CLERK

(54) Titre : MISE A DISPOSITION DE PLUSIEURS RESSOURCES DE RESEAU

(54) Title: PROVISIONING MULTIPLE NETWORK RESOURCES



(57) Abrégé/Abstract:

A resource provisioning service allows users to provision multiple, different network resources in an atomic manner and with a single call to a resource provisioning service. In some instances, the multiple, different network resources comprise individual types of resources that form a portion of one or more cloud-computing platforms. For instance, one or more entities may host and operate a cloud-computing platform that includes different types of network resources, such a storage service, a load balancing service, a compute service, a security service, or any other similar or different type of network-accessible service.

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
12 January 2012 (12.01.2012)(10) International Publication Number
WO 2012/006034 A1(51) International Patent Classification:
G06Q 10/00 (2006.01)(21) International Application Number:
PCT/US2011/042050(22) International Filing Date:
27 June 2011 (27.06.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
12/824,723 28 June 2010 (28.06.2010) US(71) Applicant (for all designated States except US): **AMAZON TECHNOLOGIES INC.** [US/US]; PO Box 8102, Reno, NV 89507 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **KRAMER, Reto** [CH/US]; c/o 410 Terry Avenue North, Seattle, WA 98109-5210 (US). **OSIECKI, Daniel, Lee** [US/US]; c/o 410 Terry Avenue North, Seattle, WA 98109-5210 (US). **SHANKARAN, Nishanth** [IN/US]; c/o 410 Terry Avenue North, Seattle, WA 98109-5210 (US). **BALAKRISHNAN, Venkates, P.** [IN/US]; c/o 410 Terry AvenueNorth, Seattle, WA 98109-5210 (US). **PARE, Geoffrey, Scott** [CA/US]; c/o 410 Terry Avenue North, Seattle, WA 98109-5210 (US). **MEIKE, Blake** [US/US]; c/o 410 Terry Avenue North, Seattle, WA 98109-5210 (US). **WHITAKER, Christopher** [GB/US]; c/o 410 Terry Avenue North, Seattle, WA 98109-5210 (US).(74) Agents: **HARTMAN, Robert, G.** et al.; Lee & Hayes, PLLC, 601 W. Riverside Ave, Suite 1400, Spokane, WA 99201 (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, RO, RS, RU, 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, SD, SL, SZ, TZ, UG,

[Continued on next page]

(54) Title: PROVISIONING MULTIPLE NETWORK RESOURCES

(57) Abstract: A resource provisioning service allows users to provision multiple, different network resources in an atomic manner and with a single call to a resource provisioning service. In some instances, the multiple, different network resources comprise individual types of resources that form a portion of one or more cloud-computing platforms. For instance, one or more entities may host and operate a cloud-computing platform that includes different types of network resources, such as a storage service, a load balancing service, a compute service, a security service, or any other similar or different type of network-accessible service.

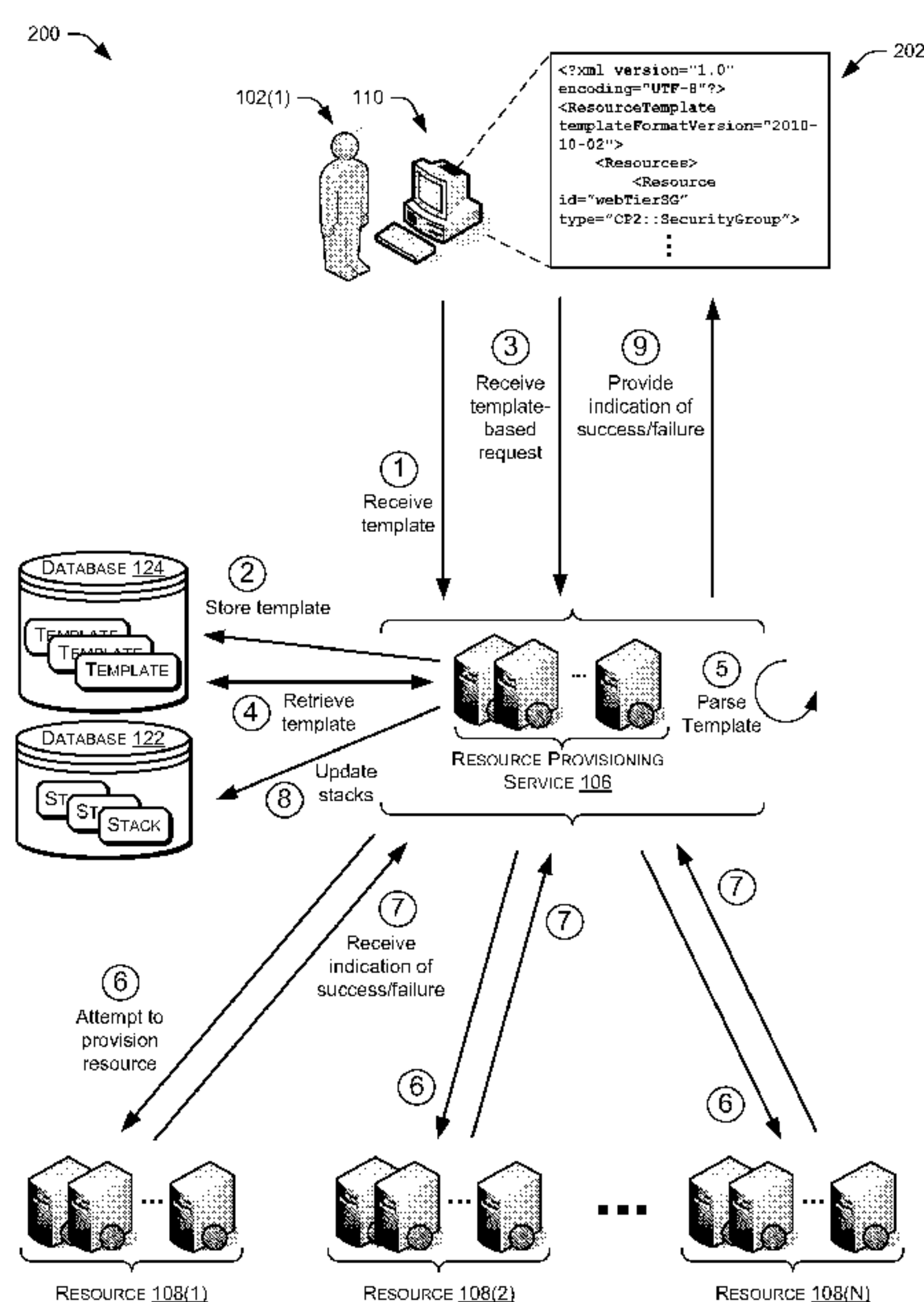


FIG. 2

WO 2012/006034 A1



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, 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))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

PROVISIONING MULTIPLE NETWORK RESOURCES

PRIORITY CLAIM AND RELATED APPLICATIONS

5 [0001]

BACKGROUND

[0002] Large-scale, network-based computing represents a paradigm shift from traditional client-server computing relationships. With large-scale, network-based computing platforms (e.g., data centers), customers are able
10 leverage shared resources on-demand by renting resources that are owned by third parties and that reside “in the cloud.” With these resources, customers of the platform are able to launch and maintain large applications without actually owning or servicing the underlying infrastructure necessary for these applications. As such, network-based computing platforms have expanded the
15 class of individuals and companies able to effectively compete in the realm of computing applications. Despite these advances, these individuals and companies may still experience difficulty in working with hosts of these network-based computing platforms for the purpose of provisioning the desired resources.

20

SUMMARY

According to an aspect, there is provided a method comprising:
under control of one or more computer systems configured with executable instructions,
receiving, from a user, a request to provision a stack of multiple, different
25 types of network resources identified by a template;
parsing the template to identify each of the multiple, different types of network resources requested for provisioning;

attempting to provision each of the multiple, different types of network resources at least partly in response to the parsing; and

subsequent to successfully provisioning each of the multiple, different types of network resources, notifying the user that the stack has been successfully
5 provisioned; or

subsequent to failing to successfully provision one or more of the multiple, different types of network resources, notifying the user that the stack has not been successfully provisioned.

According to another aspect, there is provided an apparatus comprising:
10 one or more processors;

memory, accessible by the one or more processors; and

a template-based stack builder stored in memory and executable on the one or more processors to:

receive, from a user, a request to create a stack of different
15 types of distinct network resources based at least in part on a template, the template comprising pre-populated instructions for creating the stack that includes the different types of distinct network resources;

parse the template to identify each of the different types of distinct network resources; and

20 attempt to provision each of the different types of distinct network resources on behalf of the user.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] The detailed description is set forth with reference to the accompanying figures. In the figures, the left-most digit(s) of a reference
25 number identifies the figure in which the reference number first appears. The

use of the same reference numbers in different figures indicates similar or identical items or features.

[0004] FIG. 1 illustrates an example environment in which users of a user community may interact with a resource provisioning service for the purpose of provisioning multiple, different network resources. Within this environment, the users may provision the multiple resources with a single request to the provisioning service in some implementations.

[0005] FIG. 2 illustrates an example process that the architecture of FIG. 1 may implement for the purpose of provisioning multiple, different network resources.

[0006] FIG. 3 is a flow diagram of an example process that the resource provisioning service may implement for provisioning multiple different network resources. Here, the process may comprise an atomic process in that the service either successfully provisions each requested resource or none of the requested resources.

[0007] FIG. 4 illustrates one example of a template-based stack builder of the resource provisioning service of FIG. 1 in greater detail.

[0008] FIG. 5 illustrates an example stack management workflow of the resource provisioning service of FIG. 1.

[0009] FIG. 6 illustrates an example process of creating a stack with use of a template within the architecture of FIGS. 1, 4, and 5.

[0010] FIG. 7 illustrates an example process of deleting a stack within the architecture of FIGS. 1, 4, and 5.

[0011] FIG. 8 illustrates an example process of adopting a stack within the architecture of FIGS. 1, 4, and 5.

[0012] FIG. 9 illustrates an example transition of internal stack state during requests to create or delete a stack.

[0013] FIG. 10 illustrates an example transition of external stack state during requests to create or delete a stack.

5

DETAILED DESCRIPTION

[0014] This disclosure describes an architecture and techniques for allowing users to provision multiple, different network resources in an atomic manner and with a single call to a resource provisioning service. In some instances, the multiple, different network resources comprise individual types of resources that form a portion of one or more cloud-computing platforms. For instance, one or more entities may host and operate a cloud-computing platform that includes different types of network resources, such a storage service, a load balancing service, a compute service, a security service, or any other similar or different type of network-accessible service.

[0015] In one example, a user may desire to launch an application (e.g., a web application) with use of multiple network resources offered by one or more third-party entities. The user may desire a load-balancing service to route workload amongst multiple, different resources, a compute service to host the application and to provide computing capability, an online storage service to store client data, and a security service to provide security to users of the application. While a single entity may offer and operate each of these network resources, multiple, different entities may offer and operate one or more of these resources in other implementations.

[0016] In either event, a resource provisioning service may communicate with the user to provision these multiple, different network resources. For

instance, the user may make an initial request to the provisioning service to provision a stack of multiple network resources. In some instances, the request may be based on a template that includes pre-populated instructions for provisioning the multiple resources. This template provides the user with a declarative language to author a stack-creation request, allowing the user to focus on which resources to provision and the interconnection of these resources, rather than on how to build these stacks and which application programming interface (API) calls to invoke. For instance, the template may instruct (e.g., in extensible markup language (XML) format, JavaScript Object Notation (JSON), etc.) the provisioning service to provision one or more different network resources.

[0017] In some instances, the user may then alter the instructions in the template to specify particular parameters associated with the resources and may save this template with the changes as a new template. For instance, the user may specify a particular desired size of a storage service, a number of desired computers of the compute service, and the like to create a new template. Furthermore, the user may assign a logical name to one or more of the network resources or components of the network resources when creating the new template. The user may also alter the original template in any other way by, for example, deleting portions of the instructions or adding new instructions to provision even more network resources than the template previously called out to create the new template. In another example, the user may also make a call, within the template, to another template. As such, templates may become building blocks for use in creating stacks of multiple, different network resources.

[0018] When the user is satisfied with the resulting instructions (with or without alterations to the template), the user provides instructions back to the provisioning service to store the instructions as a new template. For instance, the user may provide a text file comprising the new template. This template
5 may comprise instructions in XML or JSON format, or in any other suitable format.

[0019] After the provisioning service stores the new template, the user may then request to create a stack of resources with use of that template (or with use of another template). In response, the provisioning service may retrieve and
10 parse the template to identify the stack of desired network resources. In one example, for instance, the provisioning service may identify that the user would like to employ a load balancing service of a particular size, a particular number of servers of a compute service, a storage service of a particular size, and a particular type of security service. In response, the provisioning service may
15 attempt to provision each of the requested network resources on behalf of the user. That is, the provisioning service may attempt to secure the desired services of the load balancing service, the compute service, and the like.

[0020] Furthermore, the provisioning service may attempt to provision at least a portion of the network resources in parallel with one another. For
20 instance, the provisioning service may determine dependencies amongst the multiple network resources, and may attempt to provision the resources according to these dependencies. The provisioning service may identify implicit dependencies from the text file, or the provisioning service may receive explicit dependencies specified by the user.

25 [0021] In some instances, the provisioning service attempts to provision atomically the stack of the network resources. That is, the provisioning service

will consider a stack successfully created when the provisioning service is able to successfully provision each of the requested network resources. In response to successfully provisioning the stack, the provisioning service may notify the user of this success. In instances where the provisioning service fails to provision one or more of the requested resources (e.g., after repeated tries), the provisioning service may deem the provisioning of the stack a failure and may roll back or tear down those requested resources that the provisioning service has successfully provisioned. In addition, the provisioning service may notify the user of this failure. As such, the user knows that at the time of submitting the request to provision the stack of multiple network resources that either all of the requested resources will be provisioned, or none of them will.

[0022] The discussion begins with a section entitled “Example Architecture” that describes an example architecture that may implement the techniques introduced above. Next, a section entitled “Example Processes” follows, and describes example processes for provisioning multiple network resources using the described example architecture of FIG. 1. A section entitled “Example Resource Provisioning Service and Processes” follows, describing specific examples of the resource provisioning service described in FIG.1 and processes that the example service may implement for creating, deleting, and adopting stacks. This section also describes example transitions of internal and external stack state during these processes. Finally, the discussion ends with a brief conclusion.

[0023] This brief introduction, including section titles and corresponding summaries, is provided for the reader’s convenience and is not intended to limit the scope of the claims, nor the proceeding sections. Furthermore, the techniques described above and below may be implemented in a number of

ways and in a number of contexts. Several example implementations and contexts are provided with reference to the following figures, as described below in more detail. However, the following implementations and contexts are but a few of many.

5

Example Architecture

[0024] FIG. 1 illustrates an example architecture 100 that may implement the described techniques. The architecture 100 includes a user community 102 that communicates over a network 104 with a resource provisioning service 106 for provisioning multiple network resources 108(1), 108(2), ..., 108(N) (or simply, “resources”). For instance, users of the user community 102 may provision services of one or more cloud-computing platforms, potentially hosted in one or more large-scale data centers. As discussed above, these network resources may include a storage service, a load balancing service, a compute service, a security service, a domain name system (DNS) service, or any other similar or different type of network-accessible service. These resources 108(1)-(N) may each be associated with a common entity or owner, with separate entities or owners, or a combination thereof.

[0025] The network 104, meanwhile, is representative of any one or combination of multiple different types of networks, such as the Internet, cable networks, cellular networks, wireless networks, and wired networks. The network 104 may be a public or a private network. Furthermore, while FIG. 1 illustrates several arrows, each of the illustrated services may communicate directly or indirectly with one another via the network 104 or via a different network.

[0026] In the illustrated example, a representative user 102(1) uses a computing device 110 to access the resource provisioning service 106 to request the provisioning of one or more of the network resources 108(1)-(N). For instance, the user 102(1) may desire to launch a web application using cloud-based resources. As such, the user 102(1) may send a request to provision multiple resources, such as load balancers from a load balancing servers, databases from a storage service, and the like.

[0027] In response to accessing the resource provisioning service 106 and making such a request, the resource provisioning service 106 may serve a user interface (UI) 112 to the computing device 110 of the user 102(1). FIG. 1 illustrates one example of such a UI. As illustrated, the UI 112 includes a first area 112(1) that allows the user to select a template to use for making a request to provision a stack of a specific set of the resources 108(1)-(N). The UI 112 also includes a second area 112(1) that allows the user 102(1) to create a name for the stack that is to be created. Finally, the example UI 112 includes an icon 112(3) that, when selected, sends the request to the resource provisioning service 106. In response to receiving this request, the resource provisioning service 106 may attempt to create the stack of resources specified in the template, as described in detail below. While FIG. 1 illustrates one example UI, other implementations may employ any other similar or different UI.

[0028] As illustrated, each of the resources 108(1)-(N) and the resource provisioning service 106 is embodied as one or more servers that collectively have processing and storage capabilities to receive and send requests. These servers may be embodied in any number of ways, including as a single server, a cluster of servers, a server farm or data center, and so forth, although other server architectures (e.g., mainframe) may also be used.

[0029] The resource provisioning service 106 comprises one or more processors 114 and memory 116, which stores an interface 118 and a template-based stack builder 120. The service 106 also stores or has access to databases 122, 124, and 126. The interface 118 comprises a network-accessible interface that allows users of the user community to access and communicate with the resource provisioning service 106. For instance, users of the user community 102 may make calls into the interface via certain public APIs. In these and other instances, users of the user community 102 access the interface 118 for the purpose of requesting to provision one or more of the network resources 108(1)-(N).

[0030] The template-based stack builder 120, meanwhile, functions to receive requests from users to create a stack of network resources (as described above) and to provide an appropriate template to the users based on the requests. For instance, in the illustrated example the builder 120 may receive a request to provision a named stack of resources with use of a particular template selected by the user 102(1). In response, the builder 120 may locate the selected template from the database 124, which may store or otherwise have access to previously created templates. The resource provisioning service 106, users of the user community 102, and/or other entities may have previously created some or all of the templates within the database 124 in some instances.

[0031] After locating the selected template, the builder 120 provides this template to the computing device 110 of the user 102(1). This template may comprise pre-populated instructions for provisioning a particular set of the network resources 108(1)-(N). The user 102(1) may then use the template in whole or in part to form the actual request for resources. For instance, the user

may alter certain parameters of the pre-populated instructions to specify a size of a requested resource, a number of components (e.g., computers, databases, etc.) of a requested resource, a particular version of a resource, or any other parameter. In some instances, the user 102(1) may specify logical names
5 within the template for one or more components of the resource. These logical names comprise names chosen by the user for the convenience of the user when later managing these resources by making status or request calls to the resource provisioning service 106.

[0032] Additionally or alternatively, the user 102(1) may add instructions to
10 the template to call yet another resource(s) or may remove instructions effective to remove a resource(s) from being provisioned. After modifying (or refraining from modifying) the template provided by the service 106, the user 102(1) submits the altered template for storage by the resource provisioning service 106 as a new template. The user 102(1) may then send a request to
15 provision the resources as specified by the new template. In some instances, the template may form a text file that includes instructions in extensible markup language (XML) format, JavaScript Object Notation (JSON) format, or the like.

[0033] Upon receiving a request to provision resources based on a template,
20 the builder 120 may retrieve and parse the template (e.g., which may be in the form of a text file) to determine the identity and parameters (e.g., size, etc.) of the desired resources, as well as the order in which to provision these resources. The builder 120 may determine the order based on implicit dependencies gleaned from the actual instructions of the text file, and/or based
25 on explicit dependencies provided by the user 102(1).

[0034] After parsing the file, the builder 120 may then attempt to provision each of the requested resources on behalf of the user, potentially in an atomic manner. In these instances, the resource provisioning service 106 attempts to provision each of the requested ones of the network resources 108(1)-(N) prior to notifying the user. If the resource provisioning service 106 successfully provisions each of the requested resources, then the resource provisioning service 106 notifies the user 102(1) that the service 106 has successfully created the requested stack. The service 106 may also update the database 122 to reflect the newly-provisioned stack. Finally, the service 106 may charge costs associated with these provisioned resources to an account of the user 102(1), maintained by the database 126.

[0035] In instances where the builder 120 of the resource provisioning service 106 is unable to successfully provision each requested resource (e.g., after multiple retries), the service 106 may notify the user 102(1) that stack creation has failed. In addition, the service 106 may roll back or tear down those ones of the requested network resources that the service 106 has successfully provisioned. The service 106 may or may not charge a user account for these successfully provisioned resources.

[0036] By operating in this atomic manner, the resource provisioning service 106 provides comfort to the user 102(1), who knows that the service 106 will either fully succeed in provisioning the requested stack of resources or will return to the state of the system prior to receiving the user's request. That is, even in the latter instances the user 102(1) will not be left with a half-completed and non-functioning stack, as the service 106 will roll back each provisioned network resource. In addition to providing this comfort, the architecture 100 and the associated techniques allow the user to make a single

request to the resource provisioning service 106 rather than multiple requests corresponding to each requested resource. The service 106 further eases this single call by providing a template to the user 102(1) on which to base the request.

5

Example Processes

[0037] FIG. 2 is an example process 200 that the architecture 100 of FIG. 1 may implement. This process (as well as other processes described throughout) is illustrated as a logical flow graph, each operation of which represents a sequence of operations that can be implemented in hardware, software, or a combination thereof. In the context of software, the operations represent computer-executable instructions stored on one or more computer-readable storage media that, when executed by one or more processors, perform the recited operations. Generally, computer-executable instructions include routines, programs, objects, components, data structures, and the like that perform particular functions or implement particular abstract data types. The order in which the operations are described is not intended to be construed as a limitation, and any number of the described operations can be combined in any order and/or in parallel to implement the process.

20 [0038] For discussion purposes, the process 200 (and the other processes) is described with reference to the architecture 100 of FIG. 1, although other architectures may implement this process.

[0039] In this example, at “1,” the resource provisioning service 106 receives, from the computing device 110 of the user 102(1), a new template from a user that is effective to attempt to provision a stack of multiple, different network resources on behalf of the user. For instance, the user 102(1) may have submitted this template via a UI served by the resource provisioning

25

service 106. Further, this template may have been based on a template previously stored by the service. That is, the user 102(1) may have requested and received a template via the UI 112 of Fig. 1, modified this template to form a new template, and then sent the new template for storage by the service 106.

5 In some instances, the user 102(1) may request storage of the template via a “Register Template” call to the resource provisioning service 106. At “2,” the resource provisioning service 106 stores the received template in the database 124.

[0040] At “3,” the resource provisioning service 106 receives a request from
10 the user 102(1) to provision a stack of resources on behalf of the user. In this example, the user 102(1) requests that the service 106 provision the resources in the manner specified by the template stored at “2.” That is, at “3,” the user 102(1) may call the template previously created by the user 102(1) stored by the service 106. Of course, in other instances, the user 102(1) may call a
15 template created by another user of the user community 102, by the resource provisioning service 106, and/or by any other entity.

[0041] FIG. 2 illustrates an example of a template 202, which comprises instructions in XML format for provisioning a set of the network resources 108(1)-(N). While the example template 202 illustrates one example format of
20 these instructions, the provided template may comprise other text-based instructions, selectable icons, radio buttons, or any other interface (graphical or otherwise) with which the user may request to provision multiple ones of the network resources 108(1)-(N).

[0042] At “4,” in this example the resource provisioning service 106
25 retrieves the template on which the user 102(1) based his or her request. In this example, the user 102(1) specifies the recently stored template created by the

user and, hence, the service 106 retrieves this template from the database 124. As discussed above, the user 102(1) may have customized this template. For instance, the user 102(1) may have specified values of certain parameters, such as particular versions of the resources, sizes of the resources, logical names of the resources, and the like.

[0043] At “5,” the resource provisioning service 106 parses the template to identify the requested ones of the network resources 108(1)-(N) and to identify an order in which to attempt to provision these requested resources. The service 106 may base this order in whole or in part upon implicit and/or explicit dependencies. The service 106 may identify the implicit dependencies by analyzing characteristics of the resources that the user 102(1) requests. The service 106 may identify explicit dependencies in response to receiving these dependencies explicitly from the user 102(1). For instance, the user 102(1) may state a desired order in which to attempt to provision the resources, in whole or in part.

[0044] At “6,” the resource provisioning service 106 attempts to provision the requested network resources, potentially in the determined order. In this example, the user 102(1) has requested to provision the network resources 108(1), 108(2), and 108(N). In some instances, the resource provisioning service 106 attempts to provision at least some of these requested resources in parallel.

[0045] At “7,” the resource provisioning service 106 receives an indication from each of the resources indicating whether or not the service 106 successfully provisioned the respective resource (after one or more multiple tries). In this example, each of the resources provides an indication of success and, as such, the resource provisioning service 106 considers the creation of the

stack a success. As such, the resource provisioning service 106 updates the database 122 to reflect the new stack at “8.” The resource provisioning service 106 may also update the database 126 to associate this stack with the account of the user 102(1) and to charge the account of the user 102(1) for the created
5 resources.

[0046] Finally, at “9,” the resource provisioning service 106 provides an indication of success or failure to the user 102(1). For instance, the resource provisioning service 106 may provide this indication to the computing device 110 or may communicate this indication via another communication channel.
10 In this example, the resource provisioning service 106 provides an indication that the service 106 has successfully provisioned each of the requested network resources 108(1), 108(2), and 108(N) on behalf of the user 102(1). In some instances, the user 102(1) may send a query regarding the status of the requested stack creation prior to receiving the indication at “9.” The service
15 106 may in turn respond to this request, informing the user 102(1) as to whether or not the stack has yet to be successfully created.

[0047] FIG. 3 is a flow diagram of an example process 300 that the resource provisioning service 106 may implement for provisioning multiple different network resources. Here, the process 300 may comprise an atomic process in
20 that the service either successfully provisions each requested resource or none of the requested resources.

[0048] The process 300 includes, at 302, the resource provisioning service 106 receiving a template from a client computing device and, in response, the service storing the received template. This template may facilitate provisioning
25 of a stack of network resources. At 304, the resource provisioning service 106 receives a request to provision a stack of network resources with use of the

template. In some instances, the template and the request may be received simultaneously (e.g., the template may comprises the request), although they need not. At 306, the resource provisioning service 106 parses the template to identify the resources and, at 308, the resource provisioning service 106
5 attempts to provision each of the identified resources.

[0049] At 310, the resource provisioning service 106 determines whether each provisioning attempt has proven successful. If not, then the resource provisioning service 106 rolls back each successfully provisioned network resources at 312. In some instances, the service ceases attempting to provision
10 resources as soon as a single resource call fails. In either instance, the resource provisioning service 106 also proceeds to notify the requesting user of the failure at 314. In some instances, the resource provisioning service 106 may also charge a user account of the requesting user for the successfully provisioned resources (e.g., as soon as the service 106 successfully provisions
15 these resources and without regard as to whether or not the creation of the entire stack succeeds).

[0050] If, however, the resource provisioning service 106 determines that each provisioning attempt has proven successful, then the resource provisioning service 106 notifies the requesting user of this success at 316. At
20 318, the resource provisioning service 106 may charge a user account of the requesting user for the successfully provisioned resources.

Example Resource Provisioning Service and Processes

[0051] FIG. 4 illustrates one example of the resource provisioning service
25 106 of FIG. 1 in great detail. As illustrated, the resource provisioning service 106 includes the interface 118 and the template-based stack builder 120, as discussed above. In one example, the template-based stack builder 120 may

include a stack builder service 402, a stack builder daemon 404, and one or more stack management workflows 406.

[0052] As discussed above, the interface 118 is a network-facing component that allows users of the user community 102 to call through public APIs. One function of this layer is to accept, validate, authenticate, authorize, and route requests to the network resources 108(1)-(N). Other functions include throttling, request logging, and delegation token exchange. The interface 118 may proxy one or more of the following APIs for the network resources 108(1)-(N): Register Template, Describe Templates, Delete Template, Create Stack, Describe Stack, Delete Stack, Describe Stack Resources, Describe Stack Events, and Adopt Stack.

[0053] The stack builder service 402, meanwhile, provides concrete implementations of various APIs fronted by the interface 118. This service 402 may also provide one or more internal APIs. APIs provided by this service can be classified as template management APIs, stack management APIs, and audit trail APIs in some instances.

[0054] Template management APIs are responsible for the storage and processing of the templates (e.g., in the database 124 of FIG. 1). Raw, user-provided templates are accepted and validated before being transformed into a canonical/reduced form for storage alongside the raw original. The stack builder service 402 provides the ability to look up these uploaded templates by name, and to mark them as deprecated. Example template management APIs may include Register Template, Describe Template, and/or Delete Template.

[0055] Stack management APIs, meanwhile, are the entry point for initiating stack creation and deletion and obtaining stack description. The stack creation and deletion calls may be handled in an asynchronous manner. That

is, the stack builder service 402 validates and records user requests for stack creations and deletions, while the stack builder daemon 404 asynchronously handles the actual processing of the users' intent. Stack management APIs also enable the reading and writing of stack resource data into underlying storage.

5 Users of the user community 102 can request to inspect the logical-physical bindings for their stack's resource via these APIs. Similarly, when the stack management workflows 406 allocate or de-allocate a stack's resources, the stack management workflows 406 use the stack management APIs to record/update the physical-logical mappings. The stack management APIs may
10 include Create Stack, Describe Stack, Delete Stack, Adopt Stack, Describe Stack Resources, Add Stack Resource (Internal APIs), and/or Delete Stack Resource (Internal APIs).

[0056] The stack builder service 402 may also maintain audit trail APIs, which enable logging of resource creation/deletion actions performed on behalf
15 of a user. In some instances, each action is logged for the purpose of maintaining a log of costs charged to the each user account in the database 126. In addition, each user of the user community 102 having an account within the database 126 may retrieve the user's log by invoking the Describe Stack Events API. The audit trail APIs may include Describe Stack Event and/or Add Stack
20 Event.

[0057] The stack builder daemon 404, meanwhile, is responsible for integrating the stack management workflows 406 with the job state storage described above. This daemon 404 may periodically check the stack request table for new/un-started requests and may spawn stack management workflows
25 406 to handle these requests. The stack builder daemon 404 may also periodically check up on existing stack management workflows 406 and may

mark some of these workflows as complete in the job storage records. This may effectively decouple stack management workflows 406 calls from any front-end services.

[0058] Finally, the stack management workflows 406 are responsible for
5 executing various steps for the creation and configuration of network resources 108(1)-(N) during stack creation, as well as during graceful shutdown and termination of network resources 108(1)-(N) during stack deletion.

[0059] FIG. 5 illustrates an example stack management workflow 406 of the resource provisioning service 106 of FIGS. 1 and 4. As illustrated, the example
10 stack management workflow 406 is built atop a Simple Workflow Service (SWF) 502. The stack management workflow 406 comprises two workflow types: a Create Stack Workflow Type and A Delete Stack Workflow Type. The Create Stack Workflow Type is associated with a create stack decider 504, while the Delete Stack Workflow Type is associated with a delete stack decider
15 506. In addition, these workflows are associated with corresponding create activities 508, describe activities 510, and delete activities 512, which together enable the creation, description, and deletion of network resources 108(1)-(N) in the context of a stack.

[0060] FIG. 6 illustrates an example process 600 of creating a stack with use
20 of a template within the architecture of FIGS. 1, 4, and 5. At “1,” the example user 102(1) calls the CreateStack API, passing in the template name and stack name to the interface 118. At “2,” an authentication service 602 authenticates the request.

[0061] At “3,” the interface 118 delegates the request to the stack builder
25 service 402 and, particularly, to an API server 604 of the stack builder service 402. The stack builder service 402 first validates the stack name to ensure that

the chosen name is not a duplicate of another stack by referencing stack resources 606. The stack builder service 402 then verifies whether the template name is valid by referencing template metadata 608. If both the stack name and the template name are valid, the stack builder service 402 persists stack metadata (Namespace, StackName, TemplateName, Status=CREATE-PENDING) in a stack metadata (SDB) table 610 and a successful reply is sent back to the user 102(1).

[0062] At “4,” The stack builder daemon 404 picks up stacks with Status = CREATE_PENDING and starts a CreateStack workflow, passing the Namespace, TemplateID, StackID to the created workflow. The daemon 404 then updates the status of those stacks to CREATE_IN_PROGRESS. The daemon 404 also enters an audit trail entry into an audit trail table 612 indicating that the stack creation is in progress.

[0063] At “5,” create stack deciders 504 of the CreateStack workflow type periodically poll the SWF 502 for decision events. When one or more of the CreateStack deciders 504 receives the start workflow request from SWF 502, the respective decider 504 starts processing the create workflow request.

[0064] At “6,” upon receiving the create stack request, the CreateStack decider 504 enters an entry into the audit trail table 612 regarding the stack creation.

[0065] At “7,” the create stack decider 504 contacts the template service (e.g., template metadata 608) to obtain the internal representation of the template for the given template ID. Upon obtaining the template, the decider 504 computes the order in which to create the resources. This may be based on implicit dependencies from the template or text file or on explicit dependencies specified by the user 102(1). Once the decider 504 determines the order of

resource creation, the decider 504 schedules the first set of CreateResource activities 508 corresponding to the network resources 108(1)-(N) in the computed list (along with configuration metadata captured in the template) to successfully create and configure one of the network resources 108(1)-(N).

5 [0066] At “8,” the CreateResource activities 508 periodically poll the SWF 502 for work. At “9,” upon receiving work to create and configure a specific resource from the SWF 502, the CreateResource activity 508 contacts the corresponding network resource to create and configure the resource. Once the create-resource call to the resource returns successfully, the CreateResource
10 activity 508 indicates to the SWF 502, via the successful completion of the activity, that the resource has successfully been created. The physical identifier of the resource that was successfully created by the activity may accompany the activity completion indication.

[0067] If, however, the CreateResource activity 508 receives an error/failure
15 from the resource, the activity 508 reports a failure to the SWF 502 via the activity failed API. The error/failure message obtained from the resource by the activity 508 may accompany the activity-failed call to the SWF 502.

[0068] At “10,” The CreateStack deciders 504 periodically poll the SWF 502 for work. At “11,” if one of the CreateStack deciders 504 receives the
20 information indicating that the create resource activity 508 completed successfully, then the CreateStack decider 504 updates the binding service with the information regarding the resource that was created. This includes the mapping of logical resource identifier to physical resource identifier and a creation timestamp. The decider 504 also enters an entry into the audit trail
25 table 612 regarding the successful creation of the resource. The decider 504

then schedules a describe resource activity 510 to describe the status of the created resource.

[0069] If the decider 504 receives the information that the create resource activity timed out, then the decider 504 reschedules the create resource activity 508 for specified number of times with back off. Steps 8–11 may repeat until the activity 508 completes successfully or until the maximum number of retries is reached. If the maximum number of retries for the create resource activity is reached, the decider 504 enters a corresponding entry into the audit trail table 612 and initiates the closure of the workflow. If the decider 504 receives the information that the create resource activity 508 failed, the decider 504 examines the error code returned by the activity 508. Based on the error code, the decider 504 may handle the error in a fashion similar to an activity timeout and may retry the activity.

[0070] Conversely, based on the error code, the decider 504 may decide not to retry the activity and may instead update the binding service with the information regarding the failure in resource creation. This update may include logging the error messages received by the activity 508 while trying to create the resource, entering a corresponding entry into the audit trail table 612, and initiating the closure of the workflow.

[0071] When the decider 504 initiates the closure of the workflow, the following sequence of steps may be executed. First, the decider 504 annotates the workflow (using the state field of the workflow record) with the CREATE_FAILED tag. Once a workflow has been annotated with the CREATE_FAILED tag, no decider 504 will schedule any more activities for that workflow. Instead, once the previously scheduled create resource activities 508 for the workflow have been accounted for (e.g., either completed

or failed), or a maximum number of retries has been reached, the decider 504 closes the workflow (via the CloseWorkflow SWF API) with an error message indicating that the stack could not be created successfully.

[0072] At “12,” The DescribeResource activities 510 periodically poll the
5 SWF 502 for work. At “13,” upon receiving work to describe a specific network resource from the SWF 502, the DescribeResource activity 510 contacts the corresponding network resource to check the status of the resource. If the resource is in “active” state, the activity 510 completes successfully. Otherwise, the activity 510 reports the failure to the SWF 502.

10 [0073] At “14,” The CreateStack decider 504 periodically polls the SWF 502 for work. If the CreateStack decider 504 receives the information indicating that the describe resource activity 510 completed successfully, the decider 504 enters a corresponding entry into the audit trail table 612 and once again contacts the template service (e.g., template metadata 608) to obtain the
15 internal representation of the template. The decider 504 compares the resources described in the template to the resources that have been created (obtained from the workflow stack history) for the given stack and determines the resources that are yet to be created. If the set of resources to be created is empty, which indicates that the stack has been successfully created, the decider
20 504 closes the workflow. If, however, the resources are yet to be created, the decider 504 schedules the next batch of create source activities 508. The process 600 may repeat steps 8–14 until the entire stack is built.

[0074] At “15,” if the decider 504 receives the information indicating that the describe resource activity 510 has timed out or failed, the decider 504
25 reschedules the create resource activity 510 for a specified number of times with back off. The process 600 may repeat steps 12–14 until the activity

completes successfully or the maximum number of retries is reached. If the maximum number of retries for the create resource activity 508 is reached, the decider 504 enters a corresponding entry into the audit trail table 612 and initiates the closure of the workflow. The process by which the decider closes
5 the workflow may be the same as the detailed earlier in step 10.

[0075] At “16,” the stack builder daemon 404 periodically polls the SWF 502 to check if the workflow is complete for stacks with status = CREATE-IN-PROGRESS. Once the stack builder daemon 404 finds that the stack creation has been successfully completed, the stack builder daemon 404 marks the
10 stacks status as CREATE-COMPLETE. However, if the stack builder daemon 404 observes that the stack creation has failed, via the CREATE-FIELD tag in the workflow status, the daemon 404 marks the stack status as CREATE-FAILED. Details of clean up of a failed stack creation are discussed below in the delete stack section. Once the stack creation workflow closes, the stack
15 builder daemon 404 enters an entry into the audit trail table 612 regarding the success/failure of the stack creation.

[0076] FIG. 7 illustrates an example process 700 of deleting a stack within the architecture of FIGS. 1, 4, and 5. At “1,” the example user 102(1) calls the DeleteStack API passing in a stack name to the interface 118. At “2,” the
20 authentication service 602 authenticates the request.

[0077] At “3,” interface 118 delegates the request to the stack builder service 402 where the stack name is validated. The status of the corresponding stack is marked as DELETE_PENDING.

[0078] At “4,” the Stack Builder daemon 404 looks for stacks in
25 DELETE_PENDING status and starts a workflow of type DeleteStack by contacting the SWF 502. Inputs to the workflow include the stackId and the

template ID used by the stack. The stacks status in the stack metadata 610 is now updated to DELETE_IN_PROGRESS. The Stack Builder daemon 404 writes an entry into the audit trail table 612 indicating that the start deletion has started.

5 [0079] At “5,” the deciders 506 of the DeleteStack workflow type periodically poll the SWF 502 for decision events. When one or more of the DeleteStack deciders 506 receives the start workflow request from the SWF 502, the decider 506 begins processing the delete stack request.

[0080] At “6,” upon receiving the delete stack request, the decider 506
10 enters a corresponding entry into the audit trail table 612. At “7,” the decider 506 contacts the template service (e.g., the template metadata 608) to obtain the template corresponding to the stack that is being deleted. Upon receiving the template, the decider 506 computes the order in which the resources are to be deleted. This may be done based on dependencies that are implicit in the stack
15 and/or the template, and/or based on explicit dependencies that the user 102(1) provides.

[0081] At “8,” based on the resource order computed in step 7, the decider 506 contacts the binding service to obtain the physical resource identifier and the resource type of the resources that are not in DELETED state. Upon
20 obtaining the resource type and identifier, the decider 506 schedules the appropriate delete resource activity 512 to delete the resource that is part of the given stack. If the decider 506 is unable to find any resource that is not in DELETED state, the stack deletion is complete. The decider 506 then closes the workflow.

25 [0082] At “9,” The DeleteResource activities 512 periodically poll the SWF 502 for work. At “10,” upon receiving work to delete a network resource from

the SWF 502, the DeleteResource activity 512 contacts the corresponding network resource to delete the resource. Input to the specific DeleteResource activity 512 includes the physical identifier of the resource to be deleted. Upon submitting a delete request to the resource, the DeleteResource activity 512
5 periodically polls the corresponding resource to check the status of the resource. Once the resource declares that the requested resource has been successfully deleted, the activity 512 indicates to the SWF 502, via the successful completion of the activity 512, that the deletion has successfully occurred. If the network returns a particular error type, the activity 512 reports
10 a failure and accompanies the error message returned by the resource to the decider 506 along with the ActivityFailed API.

[0083] At “11,” deciders of 506 of the DeleteStack workflow type periodically poll the SWF 502 for decision events. At “12,” upon receiving the information regarding the activity completion, the binding service is updated to
15 reflect the deletion of a resource within the stack. The decider 506 enters a corresponding entry into the audit trail table 612 regarding the successful resource deletion. The process 700 may repeat steps 7–12 until the entire stack is deleted. If the decider 506 receives information that the delete resource activity 512 has timed out or failed, the decider 506 reschedules the delete
20 resource activity 506 a specified number of times. If the maximum number of retry attempts has been exceeded, the decider 506 updates the status and reason filed of the resources in the binding service to DELETE_FAILED and “max delete attempt reached.” The decider 506 also enters a corresponding entry in the audit trail table 612 and annotates the workflow with the
25 DELETE_FAILED tag using the status filed in the workflow metadata.

[0084] At “13,” the stack builder daemon 404 periodically polls the SWF 502 to check if the workflow is closed for stacks with status=DELETE_IN_PROGRESS. Once it finds that the stack deletion has been successfully completed, the stack builder daemon 404 marks the stacks status to DELETE_COMPLETE. If the daemon 404 discovers that the workflow has been annotated with the DELETE_FAILED tag, it updates the stack status to DELETE_FAILED.

[0085] In some instances, the example user 102(1) may initiate a deletion of a stack while the creation of a stack is in progress and before the stack is fully created. In this instance, the resource provisioning service 106 may interpret this as a request to terminate the stack creation. Under this scenario, the stack builder service 402 marks the stack status as CREATE_CANCEL_PENDING. The stack builder daemon 404 scans for stacks in CREATE_CANCEL_PENDING status and notes into the corresponding create stack workflow requesting the workflow to terminate. The stack builder daemon 404 then updates the stack status to CREATE_CANCEL_IN_PROGRESS.

[0086] Upon receiving the note event, the CreateStack decider 504 initiates closure of the workflow by waiting for outstanding activities to be accounted for. The procedure for closing the workflow may be the same as the procedure described above in the create stack scenario.

[0087] The stack builder daemon 404 periodically polls the SWF 502 to check if the create stack workflow is closed for stacks with status CREATE_CANCEL_IN_PROGRESS. Once the stack builder daemon 404 discovers that the stack creation has been terminated, the stack builder daemon

404 marks the stacks status to DELETE_PENDING. The process of stack deletion cleans up any resources that were created during stack creation.

[0088] FIG. 8 illustrates an example process of adopting a stack within the architecture of FIGS. 1, 4, and 5. At “1,” the example user 102(1) calls the AdoptStack API, passing in the template name, stack name, and a map that captures the stack binding. At “2,” The authentication service 602 authenticates the request.

[0089] At “3,” the interface 118 delegates the request to the stack builder service 402 where the stack name is validated (e.g., checked for duplicates etc). If the stack name is invalid, a fault is returned to the user. The Stack builder service 402 attempts to validate the template name, and if the template name is valid, the service 402 then validates the contents of the stack binding map to ensure that the template identified by the template name and the binding map are consistent with one another. If the template name is invalid, or the template identified by the name and the binding map are inconsistent, a fault is returned to the user.

[0090] At “4,” the stack builder service 402 then invokes the DescribeResource API describing the resources listed in the template to verify that the resource listed in the binding map exists and is owned by the user 102(1). If any of the resources listed in the binding map do not exist, a fault is returned to the user. The stack builder service 402 creates a unique identifier for the stack and inserts a new entry into the binding service for the newly generated stack identifier with the binding information present in the binding map. The stack builder service 402 also inserts/updates the stack metadata table 610 to add/update the association between the user-provided stack name

and the generated stack identifier. The stack builder service 402 also inserts an entry into the audit trail table 612 indicating the adoption of the stack.

[0091] FIG. 9 illustrates an example transition of internal stack state during requests to create or delete a stack. In this example, the internal state of the example stack state is initially set to “Create Stack Pending,” as shown at 902. At 904, the resource provisioning service 106 starts a create stack workflow. As shown at 906, this alters the internal stack state to “Create in Progress.” If, however, the service 106 receives a request to delete the stack as shown at 908, the internal stack state becomes “Create Cancel Pending,” as shown at 910.

10 [0092] The service 106 may also receive a request to delete the stack at 912 while the state is in “Create in Progress” (as shown at 906). In response, the internal stack state will also transition to “Create Cancel Pending,” as shown at 910.

[0093] When a create stack workflow is in progress and no requests to delete the stack are received, the service 106 determines whether or not the create stack workflow is successful at 914. If so, then the internal stack state transition to “Create Complete” at 916. If not, then the internal stack state transitions to “Create Failed” at 918. In the former instances, the service 106 may receive a request to delete the stack at 920. In this instance, the internal stack state transitions to “Delete Pending,” as shown at 922.

[0094] In instances where the internal stack state becomes “Create Failed,” meanwhile, the transition 900 initiates a stack clean up at 924. Thereafter, the internal stack state transitions to “Delete Pending,” as shown at 922.

[0095] Returning to the internal stack state of “Create Cancel Pending” at 910, the transition 900 notes into the create stack workflow at 926, as discussed above. The internal stack state then transitions to “Create Cancel in Progress”

at 928. Thereafter, the create stack workflow completes at 930 and the stack state is updated to “Delete Pending,” as shown at 922.

[0096] When the internal stack state is “Delete Pending,” a delete stack workflow begins at 932. As shown at 934, the internal stack state then transitions to “Delete in Progress.” At 936, the service 106 determines whether or not the delete stack workflow has successfully completed. If so, then the internal stack state transitions to “Delete Complete” at 938. If not, then the internal stack state transitions to “Delete Failed” at 940. Thereafter, the service 106 may clean up the workflow and may again receive a request to delete the stack at 942. In response, the internal stack state transitions to “Delete Pending” as shown at 922.

[0097] FIG. 10 illustrates an example transition 1000 of external stack state during requests to create or delete a stack. In this example, the external state of the example stack state is initially set to “Create in Progress,” as shown at 1002. At 1004, the resource provisioning service 106 determines whether or not the creation is successful. If so, then the external stack state transitions to “Create Complete” at 1006. If not, then the external stack state transitions to “Create Failed” at 1008.

[0098] At 1010, the service 106 receives a request to delete the stack. As such, the external stack state transitions to “Delete in Progress” at 1012. At 1014, the service 106 determines whether or not the deletion is successful. If so, then the external stack state becomes “Delete Complete” at 1016. If not, then the external stack state becomes “Delete Failed” at 1018. Thereafter, the service 106 may again receive a request to delete the stack, as shown at 1010.

[0099] Clause 1. One or more computer-readable media storing computer-executable instructions that, when executed, cause one or more processors to perform acts comprising:

receiving, from a computing device of a user, a template comprising
5 pre-populated instructions effective at least in part to provision a stack of multiple, different network resources according to one or more parameters associated with each of the multiple, different network resources;

storing the template at least partly in response to the receiving;

receiving, from the computing device, a request to provision the stack of
10 multiple, different network resources with use of the template;

retrieving and parsing the template to identify each of the multiple, different network resources requested for provisioning and each of the associated one or more parameters; and

at least partly in response to the parsing, attempting to provision each of
15 the multiple, different network resources according to the associated one or more parameters on behalf of the user.

[00100] Clause 2. One or more computer-readable media as recited in clause 1, wherein the multiple, different network resources comprise a storage service, a load balancing service, a compute service, or a security service.

20 [00101] Clause 3. One or more computer-readable media as recited in clause 1, wherein the multiple, different network resources form at least a portion of a cloud-computing platform.

[00102] Clause 4. One or more computer-readable media as recited in clause 1, wherein the one or parameters associated with a respective one of the
25 multiple, different network resources specifies a requested size, capability, or characteristic of the respective network resource.

[00103] Clause 5. One or more computer-readable media as recited in clause 1, wherein the template identifies at least one of the multiple, different network resources by a logical name specified by the user.

5 [00104] Clause 6. One or more computer-readable media as recited in clause 1, wherein the template specifies a particular version of multiple different versions for at least one of the multiple, different network resources.

[00105] Clause 7. One or more computer-readable media as recited in clause 1, wherein the attempting to provision each of the multiple, different network
10 resources comprises attempting to provision at least two of the multiple, different network resources in parallel.

[00106] Clause 8. One or more computer-readable media as recited in clause 1, wherein the attempting to provision each of the multiple, different network resources comprises attempting to provision each of the multiple, different
15 network resources in an order based at least in part on specified dependencies between a portion of the multiple, different network resources.

[00107] Clause 9. One or more computer-readable media as recited in clause 8, wherein the specified dependencies comprise implicit dependencies, and wherein the parsing of the template is effective to identify the implicit
20 dependencies between the portion of the multiple, different network resources.

[00108] Clause 10. One or more computer-readable media as recited in clause 8, wherein the specified dependencies comprise explicit dependencies specified by the user.

[00109] Clause 11. One or more computer-readable media as recited in clause
25 1, further storing computer-executable instructions that, when executed, cause the one or more processors to perform an act comprising notifying the user that

the stack has been provisioned at least partly in response to successfully provisioning each of the multiple, different network resources.

[00110] Clause 12. One or more computer-readable media as recited in clause 1, further storing computer-executable instructions that, when executed, cause
5 the one or more processors to perform acts comprising:

receiving an indication that provisioning of one of the multiple, different network resources has failed; and

rolling back each successfully provisioned one of the multiple, different network resources at least partly in response to the receiving of the indication.

10 [00111] Clause 13. One or more computer-readable media as recited in clause 12, further storing computer-executable instructions that, when executed, cause the one or more processors to perform an act comprising notifying the user that provisioning of the stack has failed at least at least partly in response to the receiving of the indication.

15 [00112] Clause 14. A method comprising
under control of one or more computer systems configured with executable instructions,

receiving, from a user, a request to provision a stack of multiple, different network resources identified by a template;

20 parsing the template to identify each of the multiple, different network resources requested for provisioning;

attempting to provision each of the multiple, different network resources at least partly in response to the parsing;

when each of the multiple, different network resources are successfully
25 provisioned, notifying the user that the stack has been successfully provisioned;
and

when one or more of the multiple, different network resources is not successfully provisioned, notifying the user that the stack has not been successfully provisioned.

[00113] Clause 15.A method as recited in clause 14, wherein each of the
5 multiple, different network resources reside on a cloud-computing platform.

[00114] Clause 16.A method as recited in clause 14, further comprising maintaining a user account associated with the user, and wherein the attempting to provision the multiple, different network resources comprises attempting to provision the multiple, different network resources on behalf of the user such
10 that a cost associated with a successful provisioning of a network resource is charged to the user account associated with the user.

[00115] Clause 17.A method as recited in clause 14, wherein the template includes pre-populated instructions for requesting to provision at least two of the multiple different resources.

15 [00116] Clause 18.A method as recited in clause 14, wherein the template comprises a text file comprising text-based instructions that include calls to each of the multiple, different network resources.

[00117] Clause 19.A method as recited in clause 14, wherein the template specifies a particular version of multiple different versions for at least one of
20 the multiple, different network resources.

[00118] Clause 20.A method as recited in clause 14, wherein the template specifies a logical name for a component of one of the multiple, different network resources that has been selected by the user.

[00119] Clause 21.A method as recited in clause 14, wherein the parsing of
25 the template comprises parsing the template to identify implicit dependencies amongst the multiple, different network resources, and wherein the attempting

to provision the multiple, different network resources comprises attempting to provision the multiple, different network resources in an order based at least in part on the implicit dependencies.

[00120] Clause 22.A method as recited in clause 14, wherein the request
5 explicitly specifies a dependency between at least two of the multiple, different network resources, and wherein the attempting to provision the multiple, different network resources comprises attempting to provision the multiple, different network resources in an order based at least in part on the dependency.

10 [00121] Clause 23.A method as recited in clause 14, wherein the attempting to provision the multiple, different network resources comprises attempting to provision at least some of the multiple, different network resources in parallel.

[00122] Clause 24.A method as recited in clause 14, further comprising
tearing down those ones of the multiple, different network resources that have
15 been successfully provisioned when one or more of the multiple, different network resources is not successfully provisioned.

[00123] Clause 25.An apparatus comprising:

one or more processors;

memory, accessible by the one or more processors; and

20 a template-based stack builder stored in memory and executable on the one or more processors to:

receive, from a user, a request to create a stack of distinct
network resources based at least in part on a template, the template comprising
pre-populated instructions for creating the stack that includes the distinct
25 network resources;

parse the template to identify each of the distinct network resources; and

attempt to provision each of the distinct network resources on behalf of the user.

5 [00124] Clause 26. An apparatus as recited in clause 25, wherein the template-based stack builder is further executable on the one or more processors to:

notify the user that the stack has been successfully provisioned when the template-based stack builder has successfully provisioned each of the
10 distinct network resources on behalf of the user;

notify the user that the stack has not been successfully provisioned when the template-based stack builder fails to successfully provision at least one of the distinct network resources on behalf of the user.

[00125] Clause 27. An apparatus as recited in clause 26, wherein the
15 template-based stack builder is further executable on the one or more processors to:

roll back each of the distinct network resources that the template-based stack builder has successfully provisioned when the template-based stack builder fails to successfully provision the at least one of the distinct network
20 resources on behalf of the user.

[00126] Clause 28. An apparatus as recited in clause 25, wherein the request comprises a request to create the stack for a web application, and the distinct network resources comprise a cloud-based load balancing service, a cloud-based compute service, and a cloud-based storage service.

25

[00127] Clause 29. An apparatus as recited in clause 25, wherein the distinct network resources form at least a portion of a cloud-computing platform managed by a single entity.

[00128] Clause 30. An apparatus as recited in clause 25, wherein one of the
5 distinct network resources forms at least a portion of a cloud-computing platform managed by a first entity and another of the distinct network resources forms at least a portion of a cloud-computing platform managed by a second entity that is independent of the first entity.

[00129] Clause 31. An apparatus as recited in clause 25, further comprising a
10 database storing or having access to multiple different templates, each of the multiple different templates comprising pre-populated instructions for creating a stack that includes distinct network resources.

[00130] Clause 32. An apparatus as recited in clause 31, wherein at least a
15 portion of the multiple different templates have been provided by a community of independent users.

[00131] Clause 33. One or more computer-readable media storing computer-executable instructions that, when executed, cause one or more processors to render a graphical user interface (GUI), the GUI comprising:

a first area that facilitates selection of one of multiple, different
20 templates, each of the multiple, different templates comprising pre-populated instructions for creating a stack that includes a set of distinct network resources;

a second area to name a stack to be created with use of the selected template; and

25 a third area to request creation of the stack.

[00132] Clause 34. One or more computer-readable media as recited in clause 33, wherein each of the multiple different templates comprise instructions in extensible markup language (XML) format or JavaScript Object Notation (JSON) format.

- 5 [00133] Clause 35. One or more computer-readable media as recited in clause 33, wherein at least a portion of the multiple different templates have been provided by different users of a community of users.

Conclusion

- 10 [00134] Although the subject matter has been described in language specific to structural features and/or methodological acts, it is to be understood that the subject matter defined in the appended claims is not necessarily limited to the specific features or acts described. Rather, the specific features and acts are disclosed as exemplary forms of implementing the claims.

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

1. A method comprising:
 - under control of one or more computer systems configured with executable instructions,
 - receiving, from a user, a request to provision a stack of multiple, different types of network resources identified by a template;
 - parsing the template to identify each of the multiple, different types of network resources requested for provisioning;
 - attempting to provision each of the multiple, different types of network resources at least partly in response to the parsing; and
 - subsequent to successfully provisioning each of the multiple, different types of network resources, notifying the user that the stack has been successfully provisioned; or
 - subsequent to failing to successfully provision one or more of the multiple, different types of network resources, notifying the user that the stack has not been successfully provisioned.
2. A method as recited in claim 1, further comprising maintaining a user account associated with the user, and wherein the attempting to provision the multiple, different types of network resources comprises attempting to provision the multiple, different types of network resources on behalf of the user such that a cost associated with a successful provisioning of a network resource is charged to the user account associated with the user.
3. A method as recited in claim 1 or 2, wherein the template includes pre-populated instructions for requesting to provision at least two of the multiple different types of resources.

4. A method as recited in any one of claims 1 to 3, wherein the template comprises a text file comprising text-based instructions that include calls to each of the multiple, different types of network resources.
5. A method as recited in any one of claims 1 to 4, wherein the template specifies a particular version of multiple different versions for at least one of the multiple, different types of network resources.
6. A method as recited in any one of claims 1 to 5, wherein the template specifies a logical name for a component of one of the multiple, different types of network resources that has been selected by the user.
7. A method as recited in any one of claims 1 to 6, wherein the parsing of the template comprises parsing the template to identify implicit dependencies amongst the multiple, different types of network resources, and wherein the attempting to provision the multiple, different types of network resources comprises attempting to provision the multiple, different types of network resources in an order based at least in part on the implicit dependencies.
8. A method as recited in any one of claims 1 to 6, wherein the request explicitly specifies a dependency between at least two of the multiple, different types of network resources, and wherein the attempting to provision the multiple, different types of network resources comprises attempting to provision the multiple, different types of network resources in an order based at least in part on the dependency.
9. A method as recited in any one of claims 1 to 8, further comprising tearing down those ones of the multiple, different types of network resources that have been successfully provisioned when one or more of the multiple, different types of network resources is not successfully provisioned.
10. An apparatus comprising:

one or more processors;
 memory, accessible by the one or more processors; and
 a template-based stack builder stored in memory and executable on the one or more processors to:

receive, from a user, a request to create a stack of different types of distinct network resources based at least in part on a template, the template comprising pre-populated instructions for creating the stack that includes the different types of distinct network resources;

parse the template to identify each of the different types of distinct network resources; and

attempt to provision each of the different types of distinct network resources on behalf of the user.

11. An apparatus as recited in claim 10, wherein the template-based stack builder is further executable on the one or more processors to:

notify the user that the stack has been successfully provisioned subsequent to the template-based stack builder successfully provisioning each of the different types of distinct network resources on behalf of the user; and

notify the user that the stack has not been successfully provisioned subsequent to the template-based stack builder failing to successfully provision at least one of the different types of distinct network resources on behalf of the user.

12. An apparatus as recited in claim 11, wherein the template-based stack builder is further executable on the one or more processors to:

roll back each of the different types of distinct network resources that the template-based stack builder has successfully provisioned when the template-based stack builder fails to successfully provision the at least one of the different types of distinct network resources on behalf of the user.

13. An apparatus as recited in any one of claims 10 to 12, wherein the request comprises a request to create the stack for a web application, and the different types of

distinct network resources comprise a cloud-based load balancing service, a cloud-based compute service, and a cloud-based storage service.

14. An apparatus as recited in any one of claims 10 to 12, wherein one of the different types of distinct network resources forms at least a portion of a cloud-computing platform managed by a first entity and another of the different types of distinct network resources forms at least a portion of a cloud-computing platform managed by a second entity that is independent of the first entity.

15. An apparatus as recited in any one of claims 10 to 12, further comprising a database storing or having access to multiple different templates, each of the multiple different templates comprising pre-populated instructions for creating a stack that includes different types of distinct network resources.

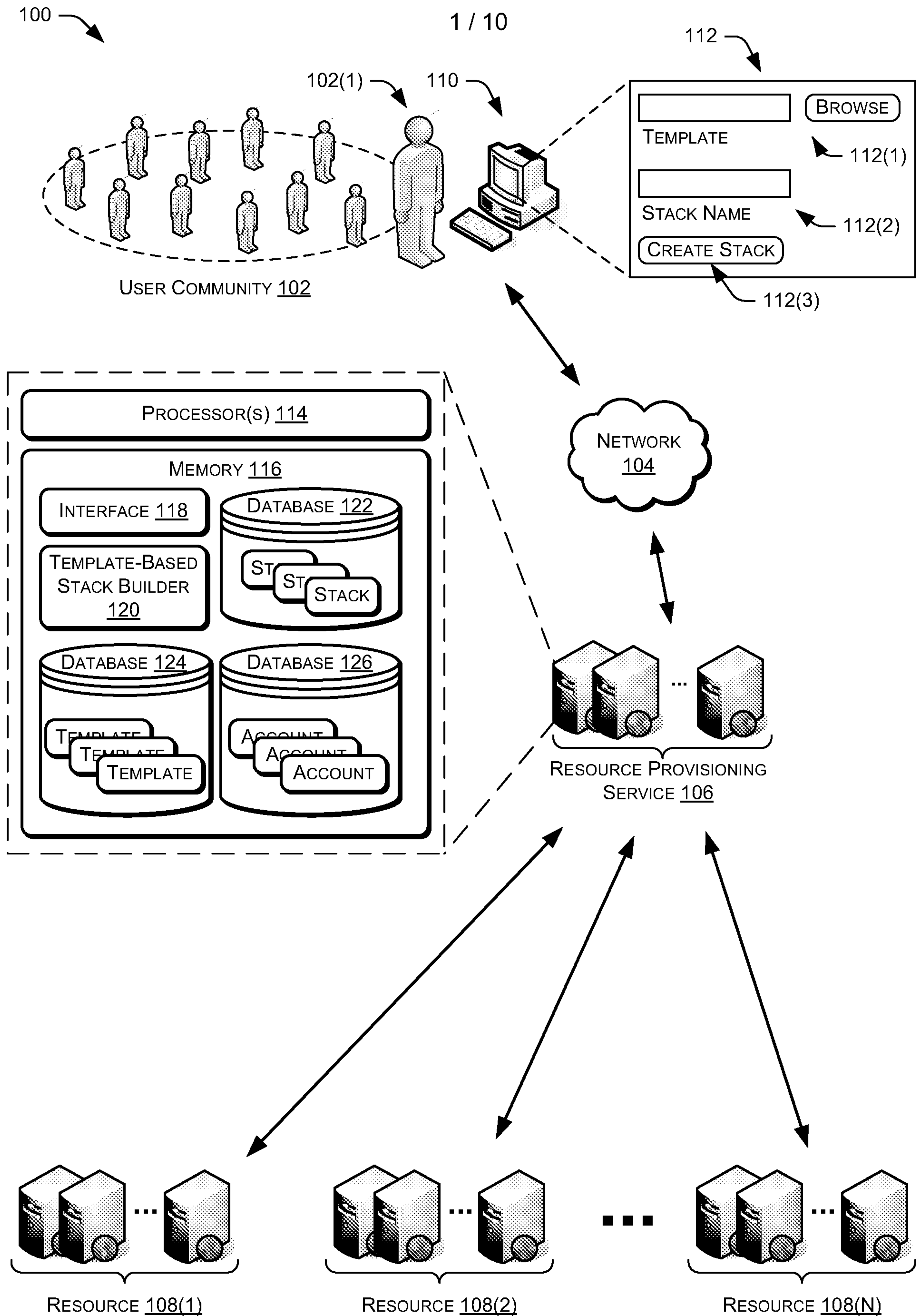


FIG. 1

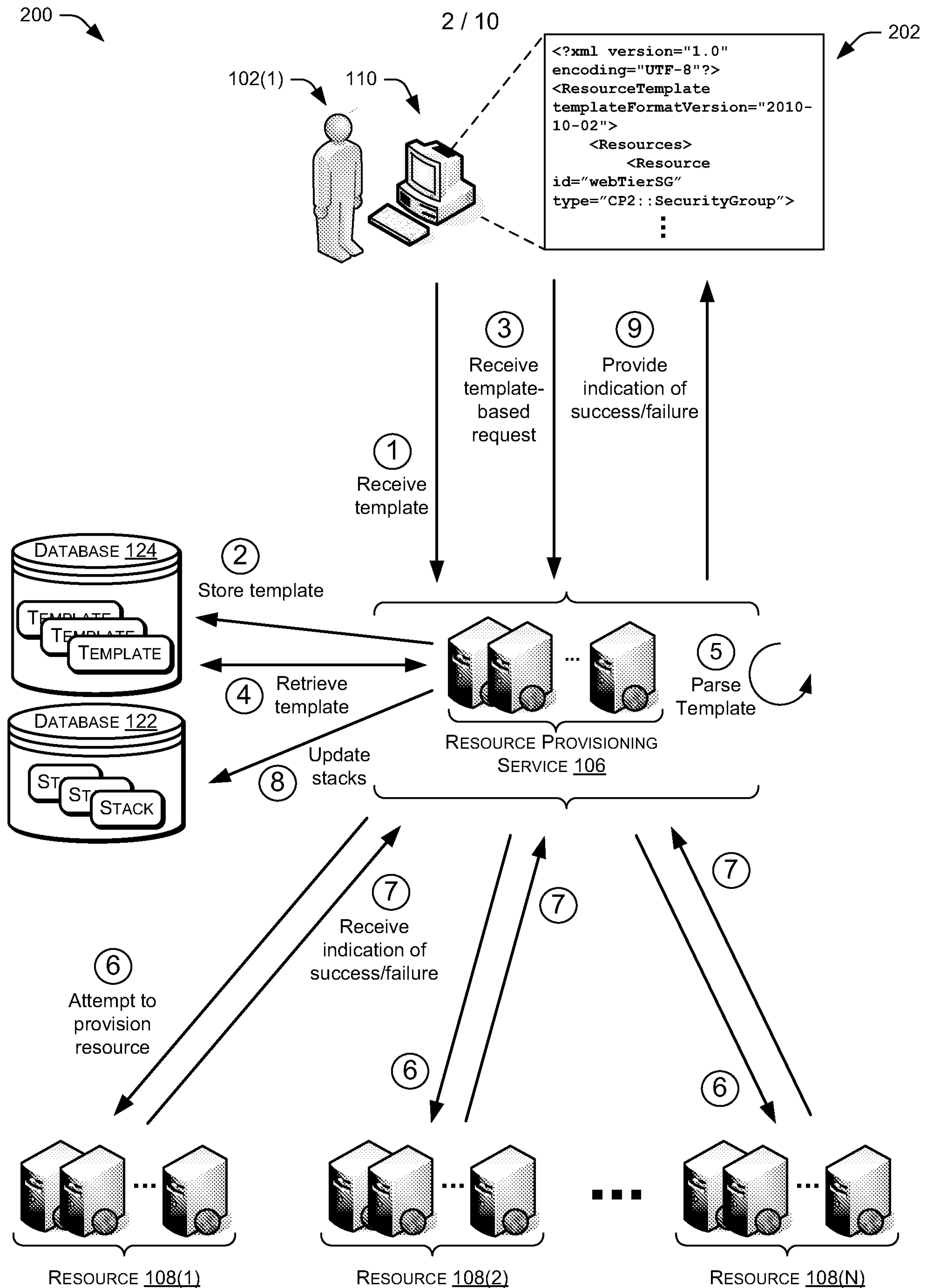
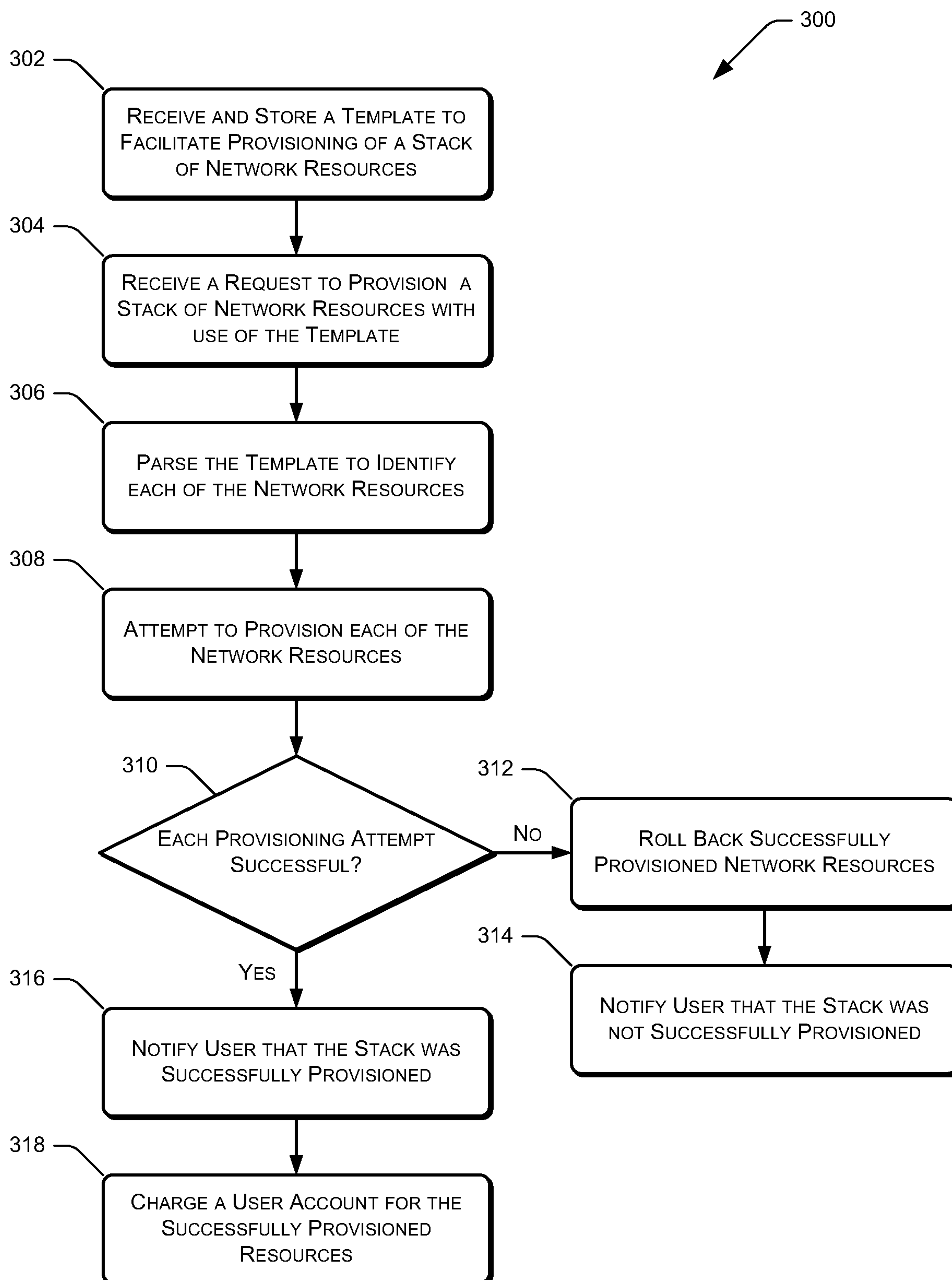
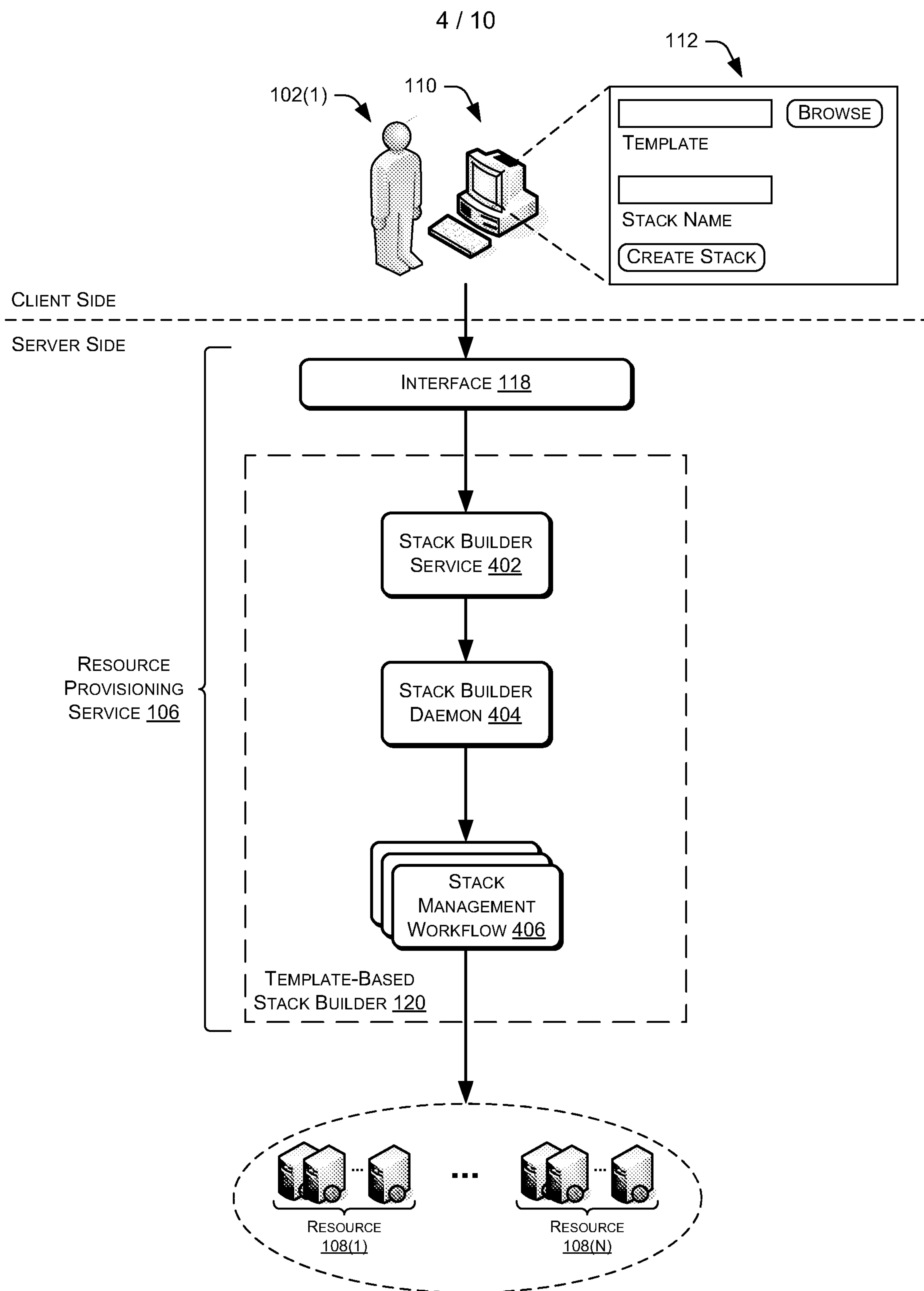
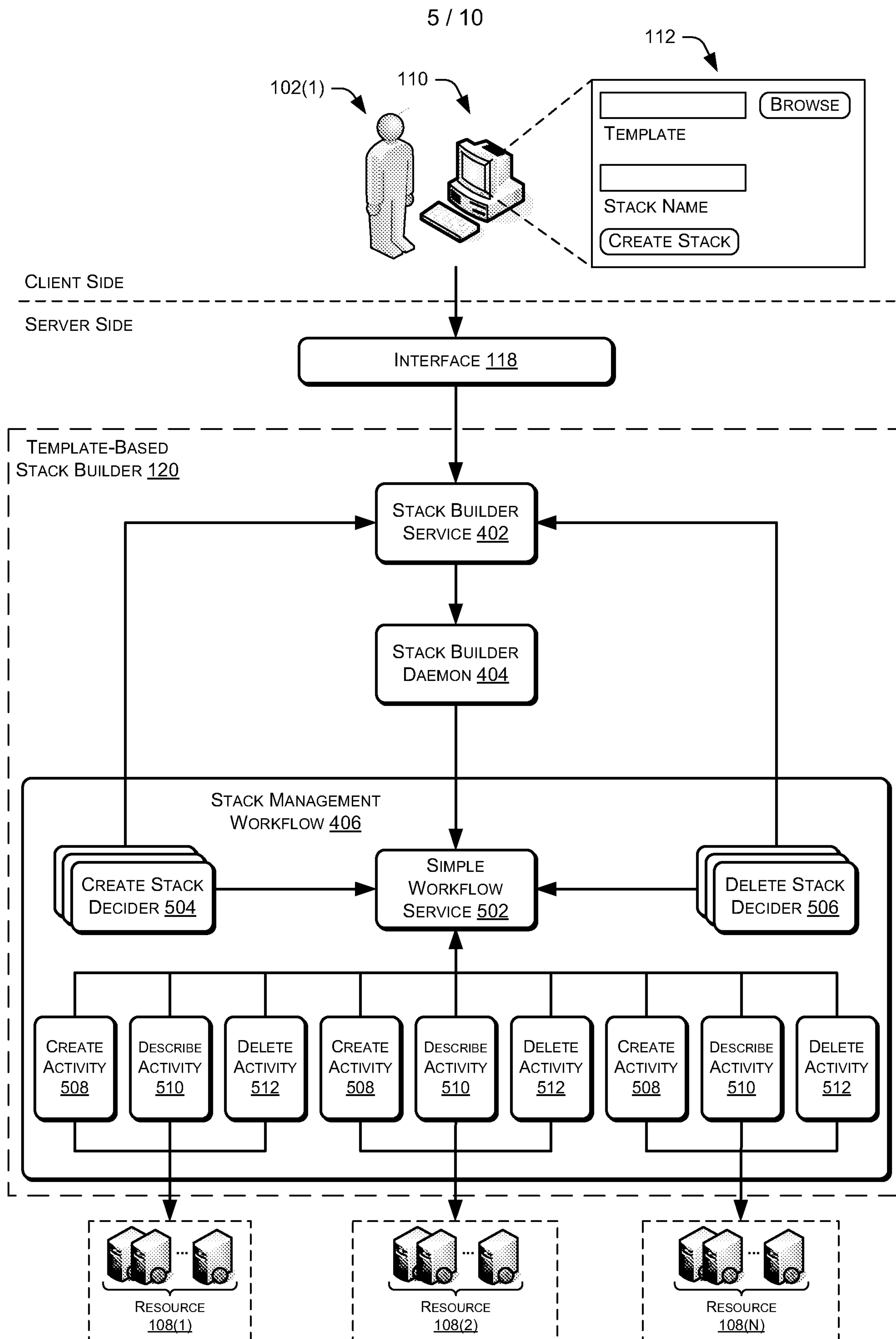


FIG. 2

3 / 10

**FIG. 3**

**FIG. 4**

**FIG. 5**

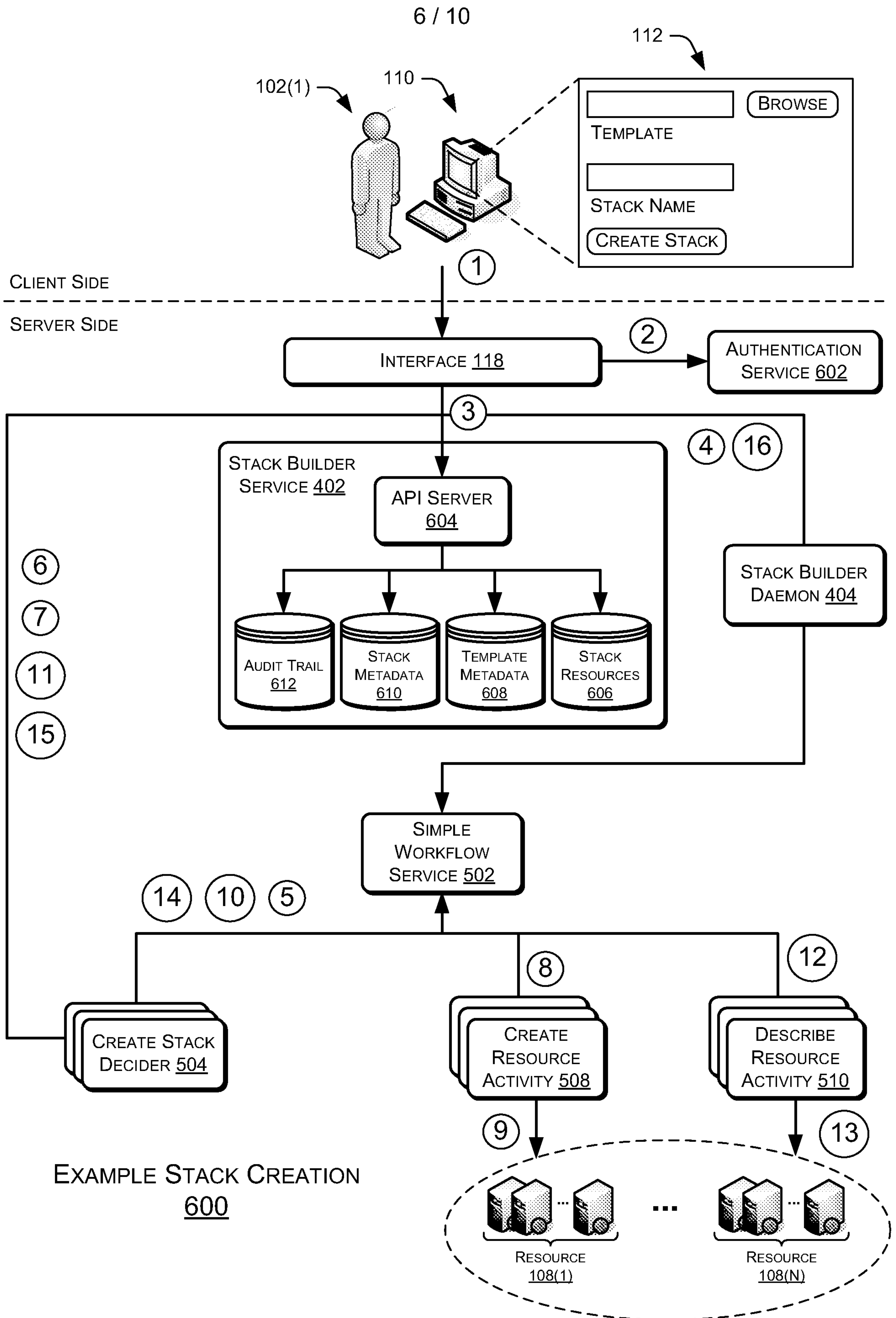


FIG. 6

7 / 10

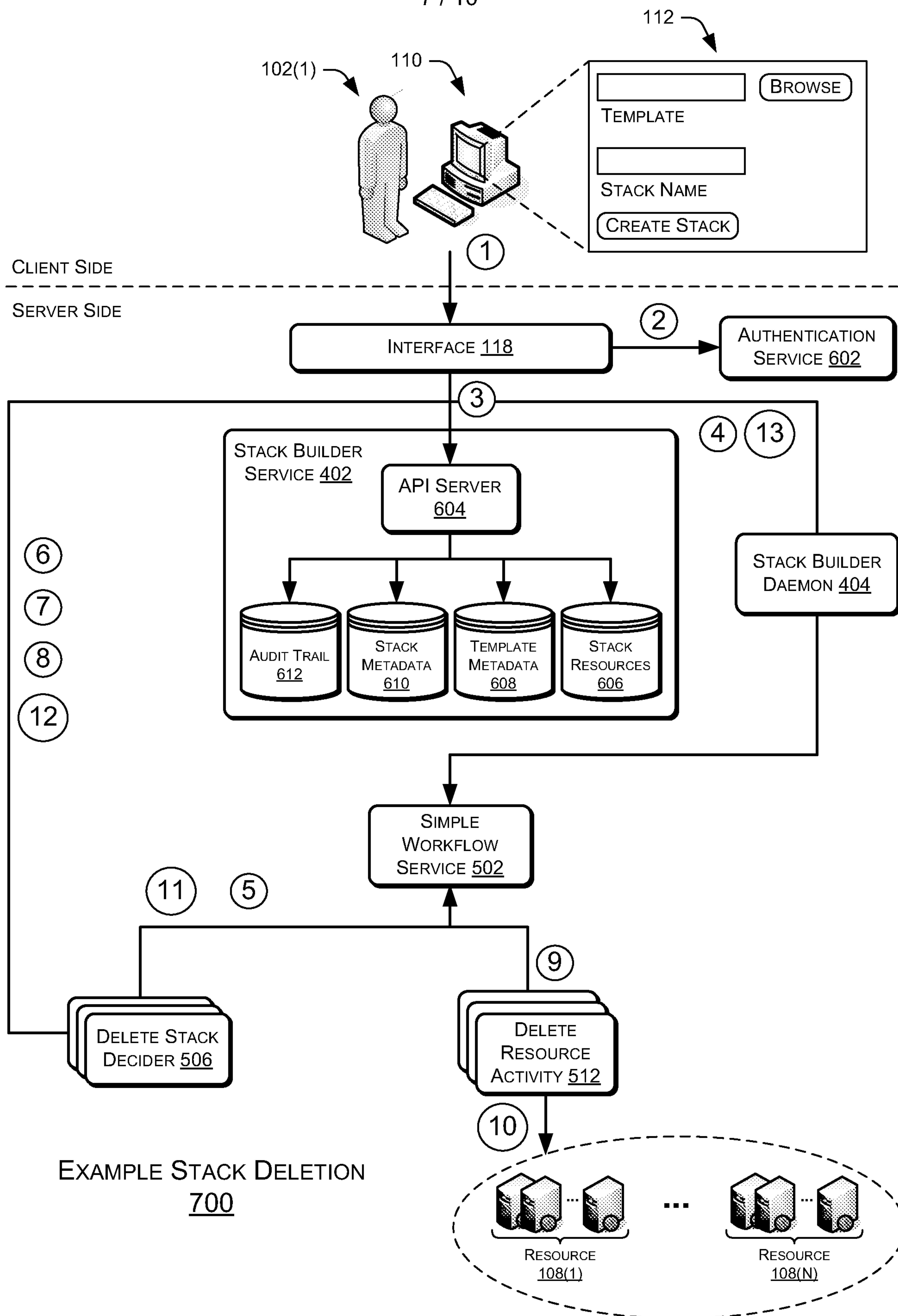
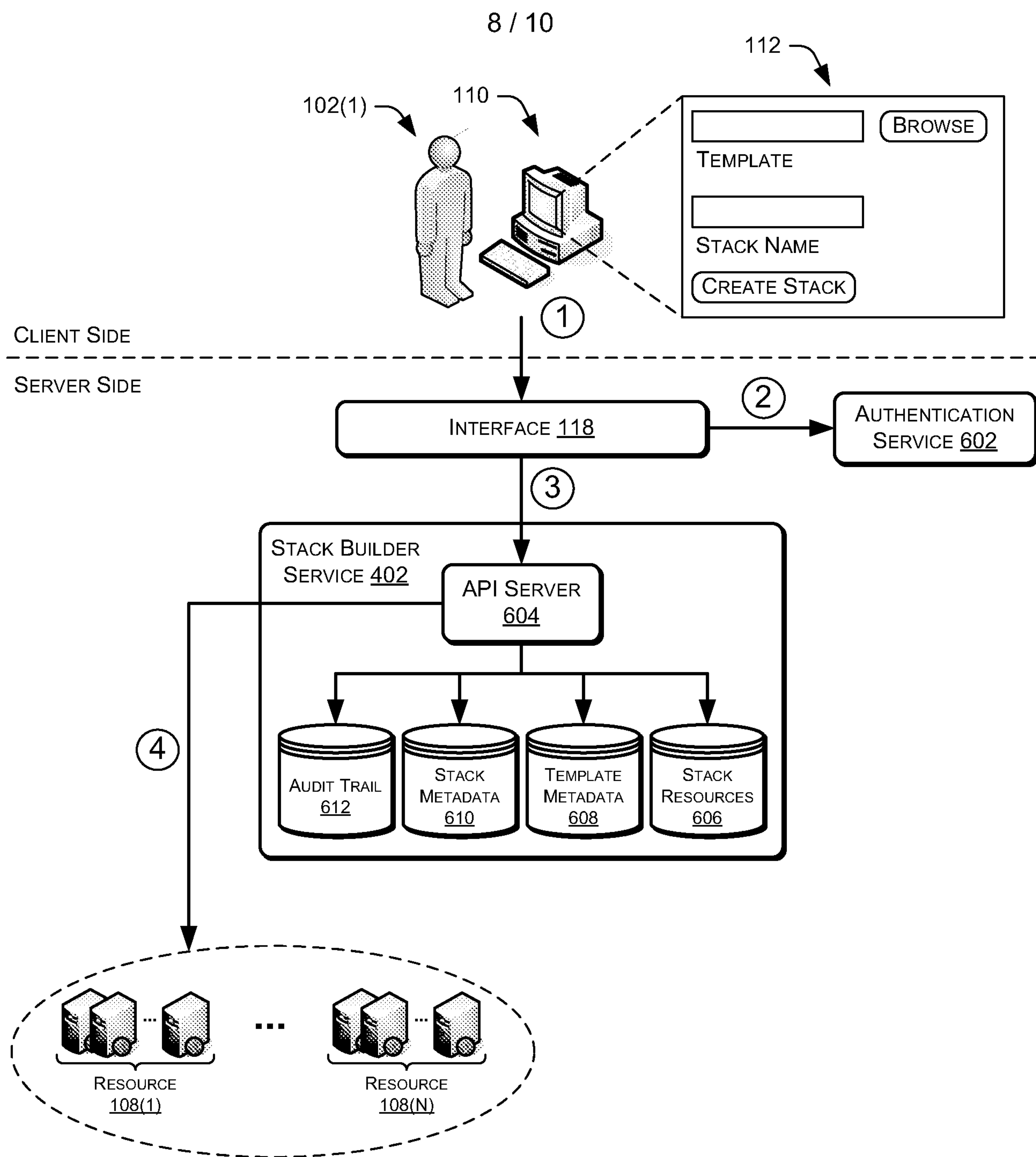
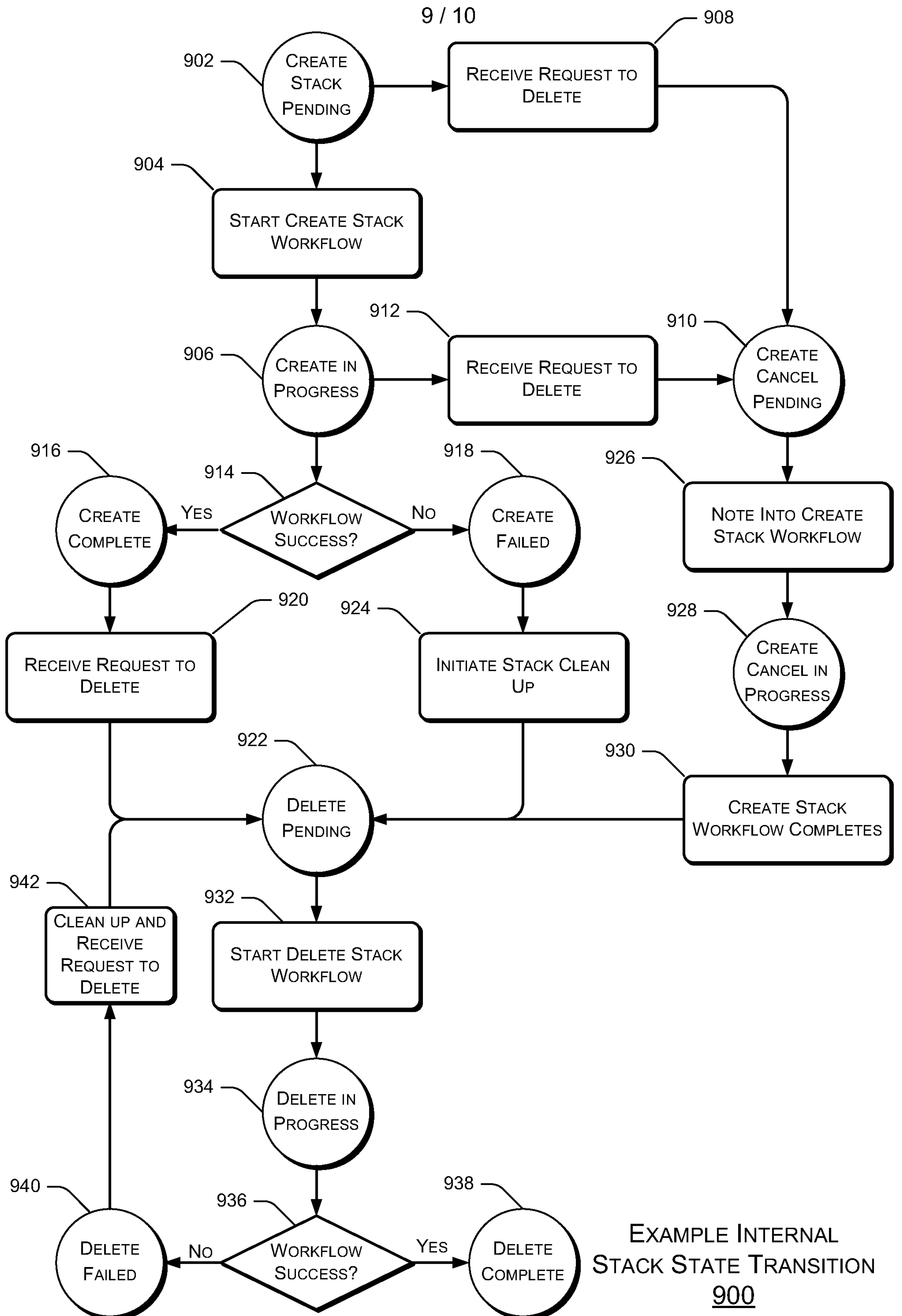


FIG. 7

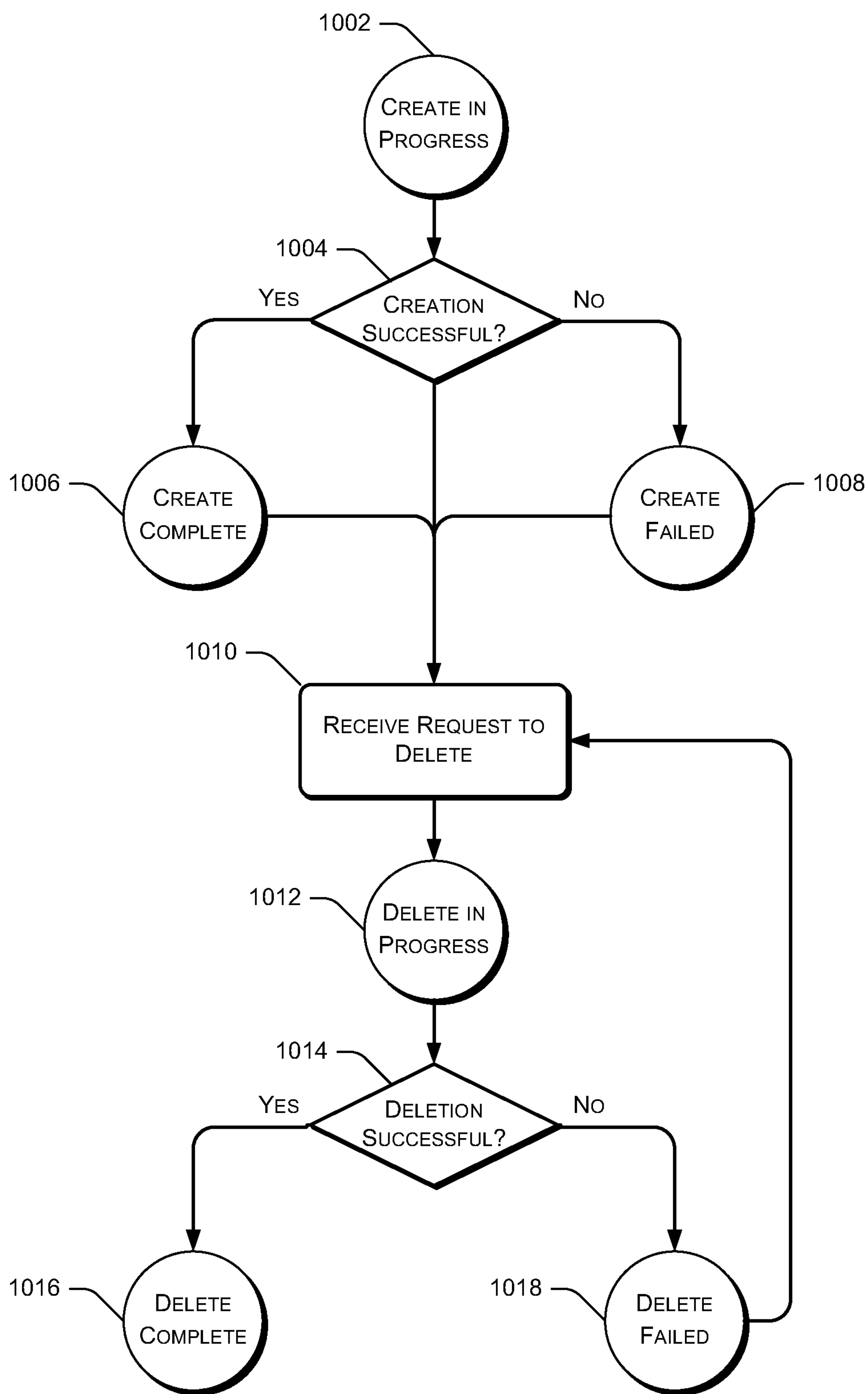


EXAMPLE STACK ADOPTION
800

FIG. 8

**FIG. 9**

10 / 10



EXAMPLE EXTERNAL
STACK STATE TRANSITION
1000

FIG. 10

200

202

102(1)

110

```
<?xml version="1.0"
encoding="UTF-8"?>
<ResourceTemplate
templateFormatVersion="2010-
10-02">
  <Resources>
    <Resource
id="webTierSG"
type="CP2::SecurityGroup">
      ...
    </Resource>
  </Resources>
</ResourceTemplate>
```

