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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 December 2018 (20.12.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/229153 A1

(51) International Patent Classification:
G06F 11/30 (2006.01) *G06F 9/455* (2018.01)

(21) International Application Number:
PCT/EP2018/065708

(22) International Filing Date:
13 June 2018 (13.06.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/521,339 16 June 2017 (16.06.2017) US

(71) Applicant: **AMADEUS S.A.S.** [FR/FR]; 485 route du Pin Montard, Sophia Antipolis, 06410 Biot (FR).

(72) Inventor: **CASTAING, Thibaud Nicolas**; 2805 Stearns Hill Road, Waltham, Massachusetts 02451 (US).

(74) Agent: **SAMSON & PARTNER PATENTANWÄLTE MBB, ASSOCIATION NO. 275**; Widenmayerstrasse 6, 80538 München (DE).

(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, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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.

(54) Title: CROSS-CLUSTER SERVICE PROVISION

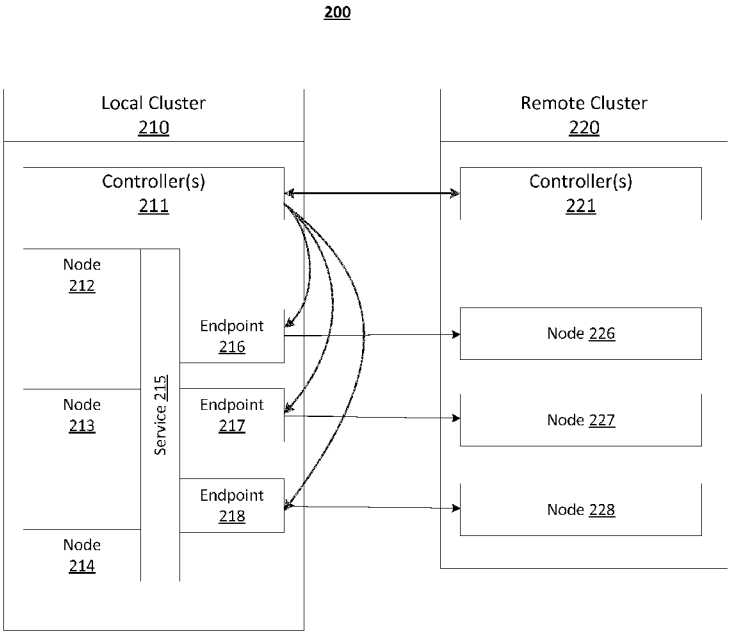


Fig. 2

(57) Abstract: Computer-implemented methods and systems are provided for cross-cluster service provision. Consistent with disclosed embodiments, a system for cross-cluster service provision can include a first computing cluster and a second computing cluster. The first cluster can include service-providing nodes and a first controller that collects status information concerning the service-providing nodes. The status information can indicate at least one service-providing node endpoint. The second computing cluster can include service-consuming nodes, a service that receives requests for the service from the service-consuming nodes and distributes the requests among a set of specified endpoints, and a second controller that performs health checks on the specified endpoints, retrieves the status information from the first controller, and updates the set of specified endpoints based on the status information.

WO 2018/229153 A1

[Continued on next page]

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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, KM, ML, MR, NE, SN, TD, TG).

Published:

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

CROSS-CLUSTER SERVICE PROVISION

TECHNICAL FIELD

[001] The disclosed embodiments concern the use of container-based software deployment and management technologies. More specifically, the disclosed embodiments concern automating deployment, scaling, and management of containerized external services across local and remote clusters.

BACKGROUND

[002] A computing cluster hosting an instance of a cloud service platform can be configured to access an external service provider, such as an SQL database or another computing cluster. The computing cluster may be provisioned with endpoints for accessing services provided by this external service provider. But these endpoints may become unreachable, inhibiting the ability of the computing cluster to access the services provided by the external service provider. Furthermore, configuration changes in the external service provider can affect the endpoints the external service provider exposes, again inhibiting the ability of the computing cluster to access the services provided by the external service provider.

[003] Accordingly, systems and method are needed for automatically provisioning a computer cluster with external service provider information. The specific technical solutions disclosed address at least the problem of accessing the services provided by an external service provider. In this manner, the disclosed systems and methods improve on existing container-based software deployment and management technologies and overcome problems specifically arising in the realm of computer networks.

SUMMARY

[004] The disclosed systems and methods enable automatic provisioning of computer clusters with external service provider information. In some embodiments, a computing cluster can be configured to automatically determine reachable endpoints as such endpoints are added and/or removed from an external service provider. For example, when a service-providing node in the service-providing cluster goes online or offline, the cluster accessing the service-providing cluster may automatically be configured to add or remove an endpoint for this service-providing node. In this manner, the computing cluster accessing the service-providing cluster can be

automatically updated to reflect dynamic changes in the configuration and accessibility of the service-providing cluster.

[005] The disclosed embodiments include a system for cross-cluster service provision. The system can include a first computing cluster and a second computing cluster. The first computing cluster can include service-providing nodes and a first controller that collects status information concerning the service-providing nodes. The status information can indicate at least one service-providing node endpoint. The second computing cluster can include service-consuming nodes and a service that receives requests for the service from the service-consuming nodes and distributes the requests among a set of specified endpoints. The second computer cluster can further include a second controller that performs health checks on the specified endpoints, retrieves the status information from the first controller, and updates the set of specified endpoints based on the status information.

[006] In some aspects, updating the set of specified endpoints includes at least one of adding another endpoint to the set of specified endpoints or removing an existing endpoint from the set of specified endpoints. The second controller can remove endpoints failing the health checks from the set of specified endpoints.

[007] The second controller can request the status information from the first controller repeatedly, periodically, at scheduled times, or in response to an event. In some aspects, the collected status information can include status information for nodes in the first computing cluster including the service-providing nodes. The first controller can determine a subset of the retrieved status information concerning the service-providing nodes. The first controller can further determine the subset of the status information based on labels included in the status information that categorize the nodes.

[008] In some aspects, the retrieved status information can include status information for nodes in the first computing cluster including the service-providing nodes. The second controller can determine a subset of the retrieved status information concerning the service-providing nodes. The second controller can further determine the subset of the retrieved status information based on labels included in the retrieved status information that categorize the service-providing nodes.

[009] The disclosed embodiments include a method for cross-cluster service provision. The method can include receiving, by a first controller on a first computing cluster that includes

service-providing nodes associated with endpoints, a request for status information for the service-providing nodes. The method can further include providing, by the first controller to a second controller on a second computing cluster that includes a service object and service-consuming nodes, a response to the request including the status information for the service-providing nodes. The method can also include receiving, by the service-providing nodes through the endpoints, health checks from the second controller and service requests from the service object. The method can further include generating, by the first controller, the response based on status information for the service-providing nodes of the first computing cluster and one or more subsetting criteria.

[010] In some aspects, the request for status information can be received repeatedly, periodically, at scheduled times, or in response to an event. The event can include at least one of a determination that a number of reachable endpoints for the service object is less than a predetermined threshold, a determination that additional endpoints are required to satisfy performance criteria for the service, or a configuration request.

[011] The disclosed embodiments include a method for cross-cluster service provision. The method can include providing, by a second controller on a second computing cluster that includes a service object and service-consuming nodes to a first controller on a first computing cluster that includes service-providing nodes associated with endpoints, a request for status information for the service-providing nodes. The method can further include receiving, by the second controller from the first controller in response to the request, status information for the service-providing nodes. The method can also include reconfiguring, by the second controller, the service object based on the status information for the service-providing nodes. The method can further include providing, by the second controller, health checks to the service-providing nodes. The method can also include providing, by the service-consuming nodes, service requests from the service object.

[012] As discussed above, the request for status information can be provided repeatedly, periodically, at scheduled times, or in response to an event at scheduled times, or in response to an event. The event can include a determination that a number of reachable endpoints for the service object is less than a predetermined threshold, a determination that additional endpoints are required to satisfy performance criteria for the service object, and a configuration request.

[013] In some aspects, the method can further include determining, by the second controller, a subset of the status information based on one or more subsetting criteria and using the subset of the status information to reconfigure the service object. The method can also include reconfiguring the service object includes at least one of specifying another endpoint for the service object and deleting a specified endpoint for the service object. For example, the second controller can reconfigure the service object by deleting specified endpoints for the service object that fail the health checks.

[014] It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the disclosed embodiments, as claimed.

BRIEF DESCRIPTION OF THE DRAWINGS

[015] The drawings are not necessarily to scale or exhaustive. Instead, emphasis is generally placed upon illustrating the principles of the inventions described herein. The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate several embodiments consistent with the disclosure and, together with the description, serve to explain the principles of the disclosure. In the drawings:

[016] **Fig. 1** depicts a conventional system for cross-cluster service provision.

[017] **Fig. 2** depicts an exemplary system for cross-cluster service provision.

[018] **Fig. 3** depicts an exemplary flowchart illustrating automatic cross-cluster service provision.

DETAILED DESCRIPTION

[019] Reference will now be made in detail to the disclosed embodiments, examples of which are illustrated in the accompanying drawings. Wherever convenient, the same reference numbers will be used throughout the drawings to refer to the same or like parts.

[020] Cloud-based application management platforms, such as KUBERNETES® or OPENSHIFT®, allow for automated deployment, scaling, and management of containerized applications. These platforms can define groups that include one or more containers that share computing resources. The groups can be configured to provide services (e.g., run applications). A cluster running such a platform can configure nodes to host these groups. Services can be assigned to clusters, which will then automatically handle creation of groups and assignment of groups to nodes to ensure that the service cluster continues to perform the service. For example,

if the cluster attempts to provide the service using a node, and the node cannot perform the service, or ceases to perform the service, then the cluster will provide the service using another node.

[021] External services can also be represented in the above-described framework. For example, a service object can be created that represents the external service. During configuration of the service, endpoints can be listed at which the external service can be accessed. The platform can then perform load balancing between the listed endpoints. For example, the external service can be assigned a virtual endpoint (the combination of a virtual Internet Protocol (IP) address and port) and can distribute requests to this virtual endpoint to the listed endpoints provided during the configuration of the service.

[022] A cluster consuming external services may lack the ability to monitor the status of these external services or the nodes providing these external services. In some instances, when the external services are accessible through endpoints, the cluster may not have access to the status of these endpoints. For example, a first cluster may consume services provided by a second cluster. As discussed above, the first cluster can represent the second cluster using a service object with a list of endpoints statically defined upon creation of the service. But should the status of the second cluster change (e.g., a node becomes unavailable, a specific service/port becomes unreachable on the node, etc.), traffic can be sent to an unreachable endpoint. Individual cluster-level controllers configured to monitor node functioning within each cluster do not completely address this problem, as endpoints may become unreachable from the first cluster while the nodes associated with the endpoints continue to function. Monitoring the status of endpoints or the nodes associated with the endpoints using a controller in the first cluster also does not completely address the problem. The controller in the first cluster may continue to monitor the statically defined endpoints or the nodes associated with these endpoints regardless of dynamic changes in the configuration of the second cluster (an increase, decrease, or reassignment of endpoints or nodes in the second cluster).

[023] Though described herein with respect to a first service-consuming cluster and a second service-providing cluster, the envisioned systems and methods are not limited to this configuration. In some embodiments, the first service-consuming cluster can access services provided on multiple service-providing clusters. These service-providing clusters can be represented by one or more service objects listing endpoints. The first service-consuming cluster

can be configured with one or more controllers for monitoring such endpoints. For example, a controller can monitor endpoints listed by multiple service objects, each representing a service-providing cluster. As an additional example, a controller can monitor endpoints listed by a service object representing multiple service-providing clusters. As a further example, a controller can correspond to a service-providing cluster and be configured to monitor endpoints exposed by that service-providing cluster. In various embodiments, a cluster may act as both a service-providing cluster and a service-consuming cluster. For example, the first cluster described above may both consume services provided by the second cluster and provide services for consumption by the second cluster (or another service-consuming cluster). In some aspects, controllers operating on the first and second clusters can be configured to provide status information regarding local service-providing nodes and receive status information regarding remote service-providing nodes.

[024] Fig. 1 depicts a conventional system (system 100) for cross-cluster service provision. This conventional approach introduces a load balancer between the service-consuming cluster and the service-providing cluster. For example, system 100 includes local cluster 110, remote cluster 120, and load balancer 130. Local cluster 110 can include service-consuming nodes (e.g., node 111, node 113, and node 115). Local cluster 110 also includes service 117, which may be a software object representing an external service provided by remote cluster 120. The nodes of local cluster 110 can send requests to service 117.

[025] Remote cluster 120 can include service providing nodes (e.g., node 121, node 123, and node 125). These nodes can be configured to expose endpoints for service requests. Load balancer 130 can be configured to maintain a pool of valid endpoints (e.g., endpoint pool 131) and can periodically confirm that endpoints in endpoint pool 131 are reachable (i.e., load balancer 130 will perform “health checks” of the endpoints). System 100 can be configured to require remote cluster 120 to register endpoints with load balancer 130. These registered endpoints can become part of endpoint pool 131 and can be used by load balancer 130 when distributing requests received from local cluster 110.

[026] Fig. 2 depicts an exemplary system (system 200) for cross-cluster service provision consistent with disclosed embodiments. System 200 improves upon system 100 of Fig. 1 by enabling reliable service provision across clusters without requiring external load balancer 130, thereby removing a point of failure and reducing the expense and the complexity of such cross-

cluster service provision. In some embodiments, system 200 can instead use load balancing functionality built into the cloud computing platform. Because this exemplary system leverages existing services to manage load balancing, it can be lightweight and can consume fewer computing resources than a system that requires additional software or hardware.

[027] System 200 of Fig. 2 can include local cluster 210 and remote cluster 220. Local cluster 210 can include one or more controllers (e.g., controller(s) 211), service-consuming nodes (e.g., node 212, node 213, and node 214), service object 215, and endpoints (e.g., endpoint 216, endpoint 217, and endpoint 218). Remote cluster 220 can include controller(s) 221 and service-providing nodes (e.g., node 226, node 227, and node 228). In some embodiments, controller(s) 221 can be configured to communicate status information regarding the service-providing nodes of remote cluster 220 to controller(s) 211. This status information can indicate, describe, or concern endpoints exposed by the service providing nodes.

[028] Controller(s) 211 can be configured to use the status information of the service-providing nodes to generate a list of endpoints (e.g., endpoint 216, endpoint 217, and endpoint 218) for service object 215. In some aspects, controller(s) 211 can be configured to monitor the status of these endpoints. As endpoints become unreachable, controller(s) 211 can be configured to remove them from the list of endpoints maintained by service object 215. Controller(s) 211 can be configured to update the listed endpoints upon receiving updated status information indicating the addition or removal of endpoints in remote cluster 220. For example, controller(s) 211 can be configured to receive status information from controller(s) 221 indicating that remote cluster 220 added a new endpoint associated with a new service-providing node. Controller(s) 211 can be configured to associate this new endpoint with services 215 and begin monitoring this endpoint.

[029] In some embodiments, system 200 can be configured to maintain multiple instances of controller(s) 211 and controller(s) 221. Maintaining multiple instances of these controllers can increase system robustness and reliability. For example, each of local cluster 210 and remote cluster 220 can be configured to include one primary controller and at least one backup controller. Should the primary controller cease functioning, the backup controller can assume responsibility for provisioning the cluster. As a non-limiting example, local cluster 210 and/or remote cluster 220 can be configured with three or more instances of a controller.

[030] Fig. 3 depicts a flowchart illustrating automatic cross-cluster service provision consistent with disclosed embodiments. In step 301, consistent with disclosed embodiments, controller(s) 211 can be configured to provide a request for status information to controller(s) 221. The request protocol is not intended to be limiting. As a non-limiting example, the request can comprise a Simple Object Access Protocol (SOAP) object, a JavaScript Object Notation (JSON) object, an Extensible Markup Language (XML) message, a remote procedure call, a Hypertext Transfer Protocol (HTTP) request message, a Simple Mail Transfer Protocol (SMTP) message, or a message in another similar protocol.

[031] In step 303, controller 221 can be configured to request status information for remote cluster 220. In some aspects, controller(s) 221 can be configured to maintain a repository (e.g., a data structure, file, or database) of status information, and can update the repository by requesting status information for remote cluster 220 repeatedly, periodically, at scheduled times, or in response to an event. In various aspects, controller(s) 221 can be configured to request status information for remote cluster 220 in response to a request from controller(s) 211. Controller(s) 221 can request this status information from remote cluster 220 (e.g., a management or configuration program running on remote cluster 220), from the nodes running on cluster 220, or from some combination of these sources.

[032] In step 305, controller(s) 221 can receive the status information for remote cluster 220. In some aspects, the status information can include a label categorizing the node. These labels or categories can differentiate the nodes based on function, exposed endpoints, current load or state (e.g., running or halted), applications running within the node, or similar criteria. The status information can describe all nodes on remote cluster 120, or a subset of the nodes on remote cluster 120. The subset of the nodes can include nodes matching one or more subsetting criteria provided by controller(s) 221. In various aspects, the status information can indicate accessible endpoints for the nodes, or a subset of the nodes, on remote cluster 220.

[033] In some embodiments, controller(s) 221 can be configured to select a subset of the endpoints exposed by the nodes running on the service-providing cluster (e.g., remote cluster 220). For example, controller(s) 221 can generate a subset of endpoints including only the endpoints corresponding to nodes satisfying one or more subsetting criteria. Non-limiting exemplary subsetting criteria include whether a node provides a particular service, whether a node is currently running, whether a node has an IP address falling within a particular range of IP

addresses, whether a node has a particular key/value pair (e.g., tier/frontend, importance/mission-critical, location/Germany, stores_customer_data/false, etc.) attached to the node, whether a node is experiencing a low memory condition, and/or whether a node is experiencing a high CPU utilization condition, or like subsetting criteria. The one or more subsetting criteria can be predetermined, or provided by controller(s) 211 in the request. In some embodiments, a node may satisfy the one or more subsetting criteria when one or more labels for the node satisfy the one or more subsetting criteria.

[034] In step 307, controller(s) 221 can provide a response to controller(s) 211. The response can indicate the subset of endpoints. For example, the response can include a list of endpoints, such as endpoints corresponding to nodes satisfying one or more subsetting criteria. The response protocol is not intended to be limiting. As a non-limiting example, the response can comprise a SOAP object, a JSON object, an XML message, a remote procedure call, an HTTP request message, an SMTP message, or a message in another similar protocol. In this manner, remote cluster 220 can be configured to provide an interface that allows controller(s) 211 to access the relevant topology of the remote cluster 220 to retrieve status information for nodes currently operating on this cluster. Additionally or alternatively, controller(s) 221 can provide status information concerning all the nodes on remote cluster 220. Controller(s) 211 may then determine a subset of endpoints satisfying one or more subsetting criteria based on the status information for all of the nodes of cluster 220.

[035] In some embodiments, controller(s) 221 can be configured to require authentication and/or authorization of controller(s) 211 before providing status information to controller(s) 211. This authentication and/or authorization can be performed by controller(s) 221 or by another authentication system (e.g., a centralized authentication and authorization server). In some embodiments, the authentication system can be a component of remote cluster 220, or of a different system. In some embodiments, local cluster 210, remote cluster 220, and the authentication system can be associated with the same entity. For example, they can all be directed or controlled by the same company or individual. In various embodiments, local cluster 210, remote cluster 220, and the authentication system can be associated with different entities. For example, the authentication system can be directed or controlled by a trusted third party.

[036] Controller(s) 211 can be authenticated and/or authorized using a password authentication protocol, a challenge-handshake authentication protocol, an extensible

authentication protocol, or another similar authentication protocol. Controller(s) 211 can be configured to respond to a request for status information with a request for authentication information or can be configured to redirect requests for status information to another system for authentication and authorization. In some embodiments, controller(s) 211 can be configured to communicate with controller(s) 221 using a secure communication protocol (e.g., HTTP Secure, HTTPS).

[037] In step 309, controller(s) 211 can configure service 215 based on the status information received from controller(s) 221. For example, when service 215 is a service object (e.g., when the cluster is a KUBERNETES® or OPENSHIFT® cluster or other container-based software deployment cluster), controller(s) 211 can update the service object by adding or removing endpoints specified for that service object based on the status information received from controller(s) 221. For example, the service object can include, as a property or field of the service object, a list of endpoints of the nodes of remote cluster 220 for providing the service. Controller(s) 211 can provide instructions for updating the values of the property or field of the service object. For example, controller(s) 211 can provide instructions to replace the existing list with a new list of endpoints, or instructions to add and/or remove endpoints from the existing list of endpoints. Service 215 can then route requests from the service-consuming nodes in local cluster 210 to service-providing nodes in remote cluster 220 based on the retrieved status information.

[038] In steps 311a, 311b, and 311c, controller(s) 211 can intermittently communicate with the service-providing nodes to ensure that they are reachable from the specified endpoints (e.g., endpoint 216, endpoint 217, and endpoint 218) of service 215. For example, the controller(s) 211 may perform a health check on the endpoint. This health check may include one or more attempts to contact the endpoint using a particular protocol (e.g., Transmission Control Protocol=TCP, Transmission Control Protocol=UDP, Transmission Control Protocol=SCTP, Transmission Control Protocol=SSL, HTTP, HTTPS). In some embodiments, controller(s) 211 can be configured to communicate with the service-providing nodes (e.g., node 226, node 227, and node 228) directly. In various embodiments, controller(s) 211 can be configured to communicate with the nodes using calls to service 215. In this manner, controller(s) 211 can determine an “endpoint readiness” for the service-providing nodes of remote cluster 220. For example, by calling service 215, controller(s) 211 may obtain a list of reachable endpoints of the

service-providing notes of remote cluster 220 that are visible to service 215 running on local cluster 210. The particular protocol used for the communication between controller(s) 211 and the service-providing nodes could be customizable, and could be determined during configuration of controller(s) 211, or changed during operation of controller(s) 211.

[039] When controller(s) 211 determines that an endpoint is no longer reachable, that endpoint can be removed from the list of endpoints providing the service. For example, as shown in Fig. 3, node 228 fails to respond to communication attempts from controller(s) 211 in step 311c. In some embodiments, controller(s) 211 can reconfigure services 215 in response to this failure. For example, controller(s) 211 can contact service 215 as described above with regard to step 309 and remove the endpoint corresponding to node 228 from the list of endpoints for service 215. Because the service 215 performs load balancing using a virtual IP address, such a change to the list of endpoints will not be visible to the service-consuming nodes running on local cluster 210. Thus the service-consuming nodes can benefit from a consistent service address.

[040] As another benefit, additional resources can be dynamically added and/or removed from remote cluster 220 without causing problems with provision of the service in local cluster 210. For example, as service-providing nodes are removed, controller(s) 211 and controller(s) 221 can cooperate to determine that the endpoints corresponding to these nodes are no longer active and to remove them from the list of endpoints. As service-providing nodes are added, controller(s) 211 and controller(s) 221 can cooperate to add endpoints corresponding to these nodes are added to the list of endpoints (or status information) retrieved by controller(s) 211. These new endpoints can then be integrated into service provision.

[041] Consistent with disclosed embodiments, controller(s) 211 can repeatedly retrieve the status information from controller(s) 221. For example, the status information can be retrieved periodically, with some predetermined duration separating the requests for status information. As an additional example, the status information can be retrieved at scheduled times. As a further example, the status information can be retrieved in response to an event. For example, controller(s) 211 can retrieve status information from controller(s) 221 in response to a determination that the number of reachable endpoints for service 215 has fallen below a predetermined threshold. Similarly, controller(s) 211 can retrieve status information from controller(s) 221 in response to a determination that additional endpoints are required to satisfy

performance criteria for service 215. As a further example, controller(s) 211 can retrieve status information from controller(s) 221 in response to a configuration request received from system 200, from a node of local cluster 210, or another system. As would be appreciated by one of skill in the art, controller(s) 221 can be configured to similarly push status information to controller(s) 211. For example, controller(s) 221 can push such information to controller(s) 211 repeatedly, periodically, at scheduled times, or in response to an event.

[042] In some embodiments, controls(s) 221 and controllers(s) 211 can include data structures and instructions stored in a non-transitory computer-readable medium, which when executed, cause one or more processors to perform the steps or functions consistent with the present disclosure. The computer-readable medium may be a storage device having the data structures and instructions stored thereon. Examples of a computer-readable storage medium include, but are not limited to, a phase-change random access memory (PRAM), a static random access memory (SRAM), a dynamic random access memory (DRAM), other types of random access memories (RAMs), a read-only memory (ROM), an electrically erasable programmable read-only memory (EEPROM), a flash memory or other memory technologies, a cache, a register, a compact disc read-only memory (CD-ROM), or any other non-transitory computer-readable medium that may be used to store information capable of being accessed by a computer device.

[043] Other embodiments will be apparent to those skilled in the art from consideration of the specification and practice of the disclosed embodiments disclosed herein. It is intended that the specification and examples be considered as exemplary only, with a true scope and spirit of the disclosed embodiments being indicated by the following claims. Furthermore, although aspects of the disclosed embodiments are described as being associated with data stored in memory and other tangible non-transitory computer-readable storage media, one skilled in the art will appreciate that these aspects can also be stored on and executed from many types of tangible computer-readable media, such as secondary storage devices, like hard disks, floppy disks, or CD-ROM, or other forms of RAM or ROM. Accordingly, the disclosed embodiments are not limited to the above-described examples, but instead are defined by the appended claims in light of their full scope of equivalents.

[044] Moreover, while illustrative embodiments have been described herein, the scope includes any and all embodiments having equivalent elements, modifications, omissions, combinations (e.g., of aspects across various embodiments), adaptations or alterations based on

the present disclosure. The elements in the claims are to be interpreted broadly based on the language employed in the claims and not limited to examples described in the present specification or during the prosecution of the application, which examples are to be construed as non-exclusive. Further, the steps of the disclosed methods can be modified in any manner, including by reordering steps or inserting or deleting steps. It is intended, therefore, that the specification and examples be considered as example only, with a true scope and spirit being indicated by the following claims and their full scope of equivalents.

CLAIMS

1. A system for cross-cluster service provision, comprising:

a first computing cluster, including:

service-providing nodes, and

a first controller which is arranged to collect status information concerning the service-providing nodes, the status information indicating at least one service-providing node endpoint; and

a second computing cluster comprising,

service-consuming nodes,

a service module which is arranged to receive requests for the service from the service-consuming nodes and to distribute the requests among a set of specified endpoints, and

a second controller which is arranged to perform health checks on the specified endpoints, to retrieve the status information from the first controller, and to update the set of specified endpoints based on the status information.

2. The system of claim 1, wherein the second controller being arranged to update the set of specified endpoints includes the second controller being arranged to add another endpoint to the set of specified endpoints or remove an existing endpoint from the set of specified endpoints.

3. The system of claim 1 or claim 2, wherein the second controller is arranged to remove endpoints failing the health checks from the set of specified endpoints.

4. The system of any one of claims 1 to 3, wherein the second controller is arranged to request the status information from the first controller repeatedly, periodically, at scheduled times, or in response to an event.

5. The system of any one of claims 1 to 4, wherein the first controller being arranged to collect the status information comprises the first controller being arranged to receive status information for nodes in the first computing cluster including the service-providing nodes and to determine a subset of the received status information concerning the service-providing nodes.

6. The system of claim 5, wherein the first controller is arranged to determine the subset of the received status information based on labels included in the received status information that categorize the nodes.

7. The system of claim 5 or claim 6, wherein the second controller is arranged to determine the subset of the received status information based on labels included in the retrieved status information that categorize the service-providing nodes.

8. The system of any one of claims 1 to 4, wherein the retrieved status information comprises status information for nodes in the first computing cluster including the service-providing nodes and the second controller is arranged to determine a subset of the retrieved status information concerning the service-providing nodes.

9. A method for cross-cluster service provision, comprising:

receiving, by a first controller on a first computing cluster that includes service-providing nodes associated with endpoints, a request for status information for the service-providing nodes;

providing, by the first controller to a second controller on a second computing cluster that includes a service object and service-consuming nodes, a response to the request including the status information for the service-providing nodes; and

receiving, by the service-providing nodes through the endpoints, health checks from the second controller and service requests from the service object.

10. The method of claim 9, wherein the request for status information is received repeatedly, periodically, at scheduled times, or in response to an event.

11. The method of claim 10, wherein the event comprises a determination that a number of reachable endpoints for the service object is less than a predetermined threshold, a determination that additional endpoints are required to satisfy performance criteria for the service, and a configuration request.

12. The method of any one of claims 9 to 11, further comprising generating, by the first controller, the response based on status information for the service-providing nodes of the first computing cluster and one or more subsetting criteria.

13. A method for cross-cluster service provision, comprising:

providing, to a first controller on a first computing cluster that includes service-providing nodes associated with endpoints by a second controller on a second computing cluster that includes a service object and service-consuming nodes, a request for status information for the service-providing nodes;

receiving, from the first controller in response to the request by the second controller, status information for the service-providing nodes;

reconfiguring, by the second controller, the service object based on the status information for the service-providing nodes;

providing, by the second controller, health checks to the service-providing nodes; and

providing, by the service-consuming nodes, service requests from the service object.

14. The method of claim 13, wherein the request for status information is provided repeatedly, periodically, at scheduled times, or in response to an event at scheduled times, or in response to an event.

15. The method of claim 14, wherein the event comprises at least one of a determination that a number of reachable endpoints for the service object is less than a predetermined threshold, a determination that additional endpoints are required to satisfy performance criteria for the service object, and a configuration request.

16. The method of any one of claims 13 to 15, further comprising determining, by the second controller, a subset of the status information based on one or more subsetting criteria and using the subset of the status information to reconfigure the service object.

17. The method of any one of claims 13 to 16, wherein reconfiguring the service object includes at least one of specifying another endpoint of the service object and deleting a specified endpoint of the service object.

18. The method of any one of claims 13 to 18, wherein the second controller reconfigures the service object by deleting specified endpoints of the service object that fail the health checks.

19. Computer program product comprising program code instructions stored on a computer readable medium to execute the method according to any one of the claims 9 to 18 when the computer program is executed on a computer.

20. Computer program configured to instruct a computer system to perform the method of any one of claims 9 to 18.


100
Local Cluster
110
Service 117
Node
111
Node
113
Node
115
Load Balancer
130
Endpoint
Pool
131
Remote Cluster
120
Node
121
Node
123
Node
125

Fig. 1

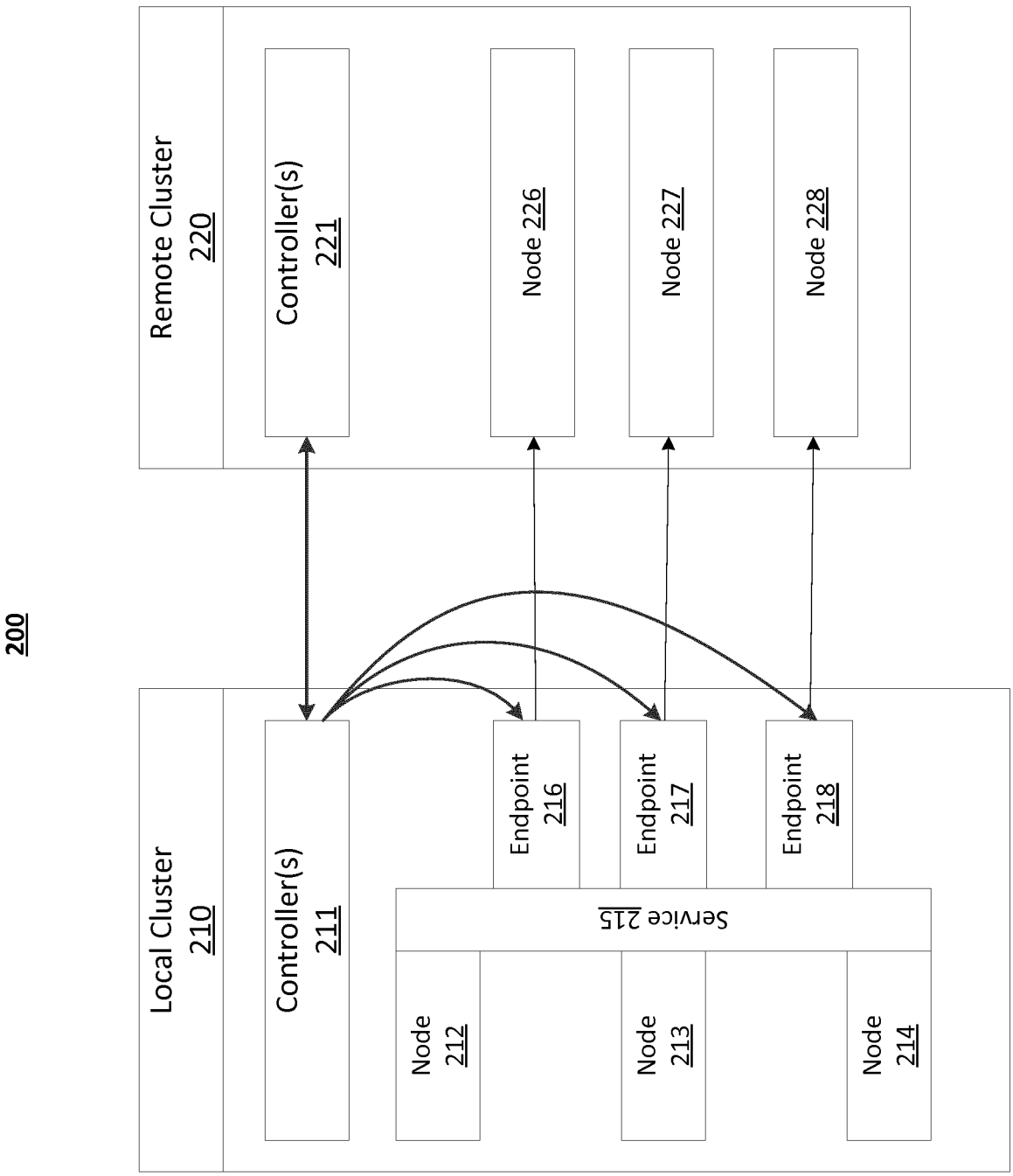

200
Remote Cluster 220
Controller(s) 221
Node 226
Node 227
Node 228
Local Cluster 210
Controller(s) 211
Endpoint 216
Endpoint 217
Endpoint 218
Service 215
Node 212
Node 213
Node 214

Fig. 2

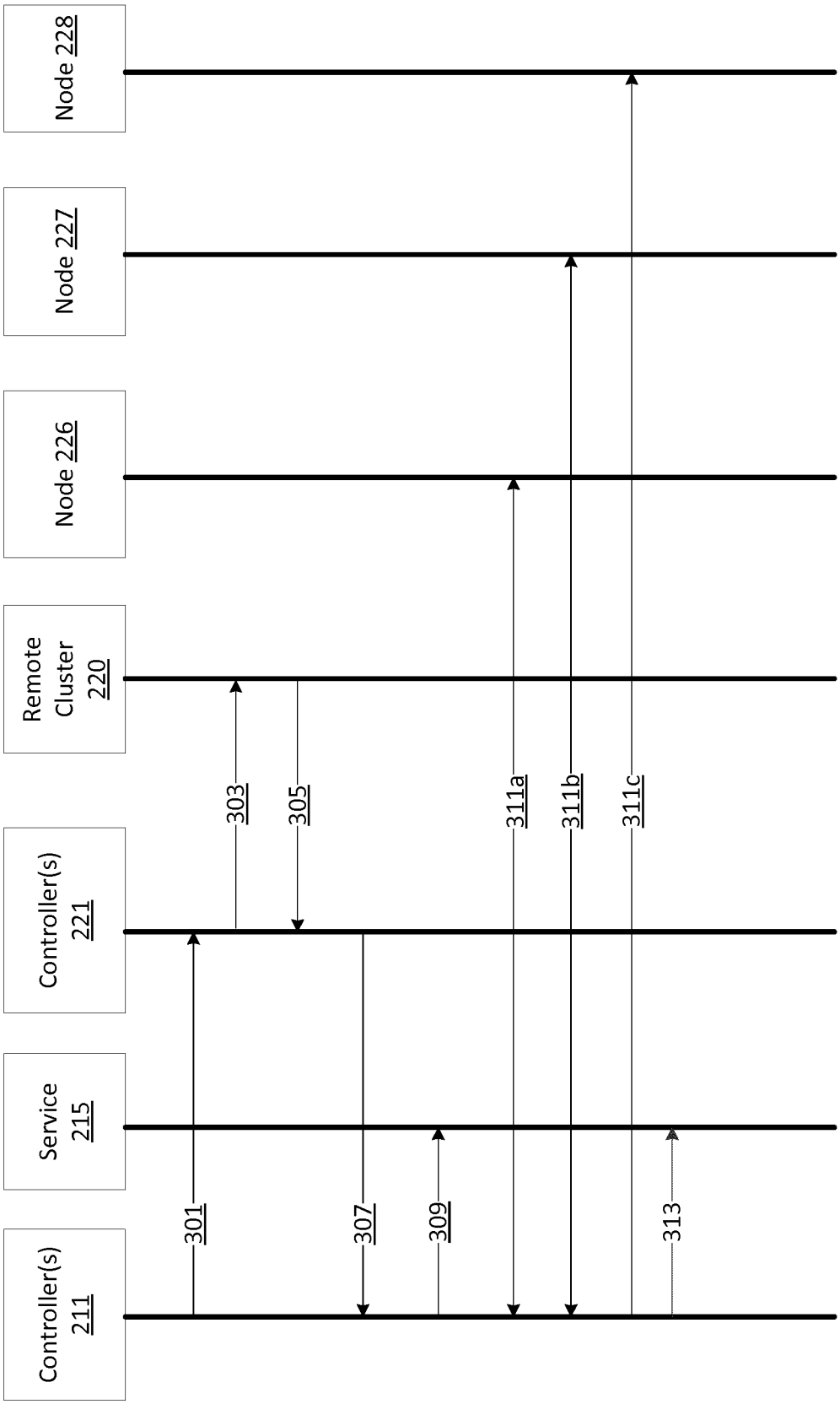

Node 228
Node 227
Node 226
Remote Cluster 220
Controller(s) 221
Service 215
Controller(s) 211
303
305
311a
311b
311c
301
307
309
313

Fig. 3

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2018/065708

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F11/30 G06F9/455
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2013/042139 A1 (BHALERAO ANAND [IN] ET AL) 14 February 2013 (2013-02-14) paragraphs [0001], [0005], [0023], [0031], [0033], [0036], [0037], [0039], [0052] figure 2	1-20
X	US 2017/123874 A1 (LEWIS LEITA BACH [US]) 4 May 2017 (2017-05-04) paragraphs [0024], [0026], [0034], [0041], [0057] figures 1, 3	1-20

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

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 September 2018	19/09/2018

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

2

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2018/065708

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2013042139 A1	14-02-2013	NONE	
US 2017123874 A1	04-05-2017	NONE	