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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 September 2018 (27.09.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/174974 A1

(51) International Patent Classification:
H04L 29/06 (2006.01) *G06F 15/163* (2006.01)
H04L 29/08 (2006.01) *G06F 15/173* (2006.01)

(21) International Application Number:
PCT/US2017/065430

(22) International Filing Date:
08 December 2017 (08.12.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
15/466,735 22 March 2017 (22.03.2017) US

(71) Applicant: **QADIUM, INC.** [US/US]; 300 Brannan Street, Suite 601, San Francisco, CA 94107 (US).

(72) Inventors: **KRANING, Matthew**; 300 Brannan Street, Suite 601, San Francisco, CA 94107 (US). **ANDERSON, Matthew**; 300 Brannan Street, Suite 601, San Francisco, CA 94107 (US). **DICKINSON, Peter**; 300 Brannan Street, Suite 601, San Francisco, CA 94107 (US). **FREDERICKS, Corey**; 300 Brannan Street, Suite 601, San Francisco, CA 94107 (US). **HOLLIMAN, John**; 300 Brannan Street, Suite 601, San Francisco, CA 94107 (US). **SEIDEL, Andrew**; 300 Brannan Street, Suite 601, San Francisco, CA 94107 (US).

(74) Agent: **WONG, Terrence, L.**; Van Pelt, Yi & James LLP, 10050 N. Foothill Blvd., Suite 200, Cupertino, CA 95014 (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, 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.

(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,

(54) Title: DISTRIBUTED SCANNING

(57) Abstract: A system for scanning a network includes an interface and a processor. The interface is configured to receive an indication to scan a set of network addresses. The processor is configured to determine a set of available scanning nodes and determine a job plan for scanning the set of network addresses using the set of available scanning nodes. The job plan includes one or more job portions. The processor is configured to, for a job portion of the one or more job portions, select a scanning node of the set of available scanning nodes and provide the job portion to the scanning node.

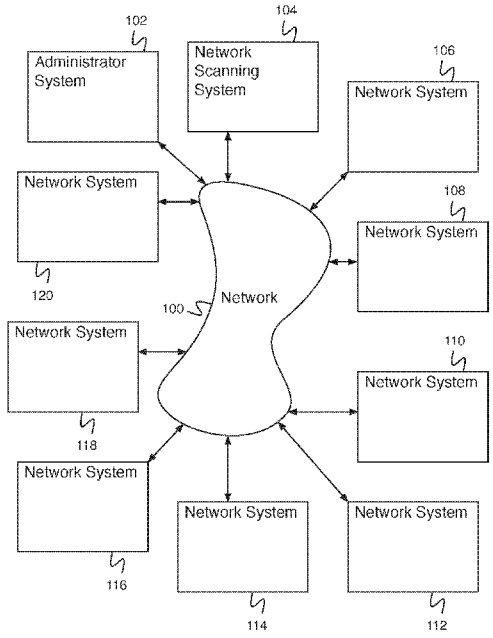

Fig. 1

WO 2018/174974 A1

[Continued on next page]

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))*

DISTRIBUTED SCANNING

BACKGROUND OF THE INVENTION

[0001] Internet connected assets (e.g., computers, mobile devices, server systems, client systems, internet-of-things devices, etc.) comprise computing systems in communication with the Internet. Internet connected assets commonly include one or more publicly addressable communication ports, allowing any Internet connected device to query the asset. Some devices allow a range of connection types (e.g., HTTP connections HTTPS connections, FTP connections, FTPS connections, telnet connections, SSH connections, etc.) over the one or more publicly accessible ports. Internet connected assets can be a wide range of different types of hardware devices running a wide range of software including a wide range of configuration options, creating a myriad of possibilities for security vulnerabilities. A typical systems administrator may not be aware of every detail of every system under his or her watch, creating a problem where system vulnerabilities may go undetected and unfixed. Scanning a wide range of network addresses can provide understanding that is not available in any other way; however, it can require a large amount of computing power and communications bandwidth to perform the scan. A single computer may not be able to perform the scan adequately.

BRIEF DESCRIPTION OF THE DRAWINGS

[0002] Various embodiments of the invention are disclosed in the following detailed description and the accompanying drawings.

[0003] Figure 1 is a block diagram illustrating an embodiment of a network system.

[0004] Figure 2 is a block diagram illustrating an embodiment of a network scanning system.

[0005] Figure 3 is a block diagram illustrating an embodiment of a network scanning system.

[0006] Figure 4 is a block diagram illustrating an embodiment of a network system.

[0007] Figure 5 is a block diagram illustrating an embodiment of a network scanning controller.

[0008] Figure 6 is a block diagram illustrating an embodiment of a scanning node.

[0009] Figure 7 is a flow diagram illustrating an embodiment of a process for scanning a network.

[0010] Figure 8 is a flow diagram illustrating an embodiment of a process for determining a job plan.

[0011] Figure 9 is a flow diagram illustrating an embodiment of a process for modifying a job plan based on a job portion result, if necessary.

DETAILED DESCRIPTION

[0012] The invention can be implemented in numerous ways, including as a process; an apparatus; a system; a composition of matter; a computer program product embodied on a computer readable storage medium; and/or a processor, such as a processor configured to execute instructions stored on and/or provided by a memory coupled to the processor. In this specification, these implementations, or any other form that the invention may take, may be referred to as techniques. In general, the order of the steps of disclosed processes may be altered within the scope of the invention. Unless stated otherwise, a component such as a processor or a memory described as being configured to perform a task may be implemented as a general component that is temporarily configured to perform the task at a given time or a specific component that is manufactured to perform the task. As used herein, the term 'processor' refers to one or more devices, circuits, and/or processing cores configured to process data, such as computer program instructions.

[0013] A detailed description of one or more embodiments of the invention is provided below along with accompanying figures that illustrate the principles of the invention. The invention is described in connection with such embodiments, but the invention is not limited to any embodiment. The scope of the invention is limited only by the claims and the invention encompasses numerous alternatives, modifications and equivalents. Numerous specific details are set forth in the following description in order to provide a thorough understanding of the invention. These details are provided for the purpose of example and the invention may be practiced according to the claims without some or all of these specific details. For the purpose of clarity, technical material that is known in the technical fields related to the invention has not been described in detail so that the invention is not unnecessarily obscured.

[0014] A system for scanning a network comprises an interface to receive an indication to scan a set of network addresses, and a processor to: determine a set of available scanning nodes, determine a job plan for scanning the set of network addresses using the set of scanning nodes, wherein the job plan comprises one or more job portions, and for each job portion of the one or more job portions: select a scanning node of the set of scanning nodes, provide the job portion to the scanning node, and receive a scanning result from the scanning node. In some embodiments, the system additionally comprises a memory coupled to the processor and configured to provide the processor with instructions.

[0015] In some embodiments, a system for network scanning comprises a system for determining information about a network by scanning some or all of the computers on it. In some embodiments, a system for network scanning comprises a system for improving network security by determining a complete outside view of the network. In various embodiments, a system for network scanning scans for devices associated with an organization, for open ports on a device, for services running on a device, for information indicating multiple devices are associated, for device configuration information, for device encryption information, for device network settings, or for any other appropriate information. In some embodiments, the system for network scanning scans a portion of the Internet. In some embodiments, the system for network scanning scans the entire Internet. In some embodiments, when a system for network scanning receives a request to scan a large network (e.g., the entire Internet), the scanning is performed using a distributed system. In some embodiments, the distributed system comprises a master system (e.g., a network scanning controller) and one or more slave systems (e.g., scanning nodes). In some embodiments, the master system determines a job plan based on a requested scan and a set of available scanning nodes, provides portions of the job plan to the available scanning nodes in order to accomplish the requested scan, and combines the results received from the scanning nodes to create the complete scan result.

[0016] Figure 1 is a block diagram illustrating an embodiment of a network system. In some embodiments, the network system of Figure 1 comprises a system for network scanning. In the example shown, Figure 1 comprises network 100. In various embodiments, network 100 comprises one or more of the following: a local area network, a wide area network, a wired network, a wireless network, the Internet, an intranet, a storage area network, or any other appropriate communication network. Administrator system 102 and network scanning system 104 communicate via network 100. Administrator system 102 comprises a system for an administrator. In various embodiments, administrator system 102 comprises a system for an administrator to

access applications on an application system, to access data on a database system, to indicate to network scanning system 104 to perform a network scanning process, to receive data from network scanning system 104, to configure a network system (e.g., network system 106), to receive data from a network system, or for any other appropriate purpose. In some embodiments, administrator system 102 comprises a processor and a memory. Network scanning system 104 comprises a system for scanning a network. In some embodiments, network scanning system 104 comprises a system for scanning data associated with network systems in response to a command from administrator system 102. In some embodiments, network scanning system 104 comprises a system for scanning data associated with a set of network systems (e.g. network system 106, network system 108, network system 110, network system 112, network system 114, network system 116, network system 118, and network system 120). In some embodiments, scanning data associated with a set of network systems comprises analyzing previously stored data associated with the set of network systems. In some embodiments, scanning data associated with a set of network systems comprises providing a payload to one or more network systems of the set of network systems and analyzing the received response, in the event a response is received. In some embodiments, network scanning system 104 comprises a processor and a memory. Each network system of Figure 1 (e.g., network system 106) comprises an Internet connected system (e.g., a desktop computer, a laptop computer, a smartphone, a tablet computer, a server system, an internet-of-things device, etc.). In various embodiments, the system of Figure 1 comprises 8, 13, 197, 2222, one million, one hundred million, or any other appropriate number of network systems. In some embodiments, each network system of Figure 1 is associated with an Internet address. In some embodiments, an Internet address comprises an IP (e.g., Internet Protocol) address. In the example shown, network scanning system 104 comprises a centralized network scanning system (e.g., the elements of the network scan including determining addresses to scan, scanning, and analyzing the results of the scan are all performed by network scanning system 104).

[0017] Figure 2 is a block diagram illustrating an embodiment of a network scanning system. In some embodiments, the network scanning system of Figure 2 comprises a system for network scanning 104 of Figure 1. In the example shown, network scanning system 200 comprises network scanning controller 202, network scan database 203, network 204, and a plurality of scanning nodes (e.g., scanning node 206, scanning node 208, scanning node 210, scanning node 212, scanning node 214, scanning node 216, and scanning node 218). Network scanning controller 202 comprises a system for receiving an indication to scan a set of network addresses (e.g., from a user or an administrator system). Network scanning controller 202 determines portions of the scan to be performed by a plurality of scanning nodes, directs scanning nodes to each perform a portion

of the requested scan, receives back results from each of the scanning nodes, provides the results to a network scan database (e.g., network scan database 203), determines differences between current results and previous results (e.g., previous results also stored in network scan database 203), and providing the requestor an indication of results, results, or a summary of the results. In some embodiments, network scanning controller 202 comprises a system implemented using cloud computing hardware. In some embodiments, network scanning controller 202 comprises a system to receive an indication of a set of scanning nodes (e.g., scanning node 208 and scanning node 216). In some embodiments, scanning nodes comprise network systems for scanning. Network scanning controller 202 communicates with network scan database 203 for storing network scanning results. Network scanning controller 202 communicates with the plurality of scanning nodes using network 204. In various embodiments, network 204 comprises one or more of the following: a local area network, a wide area network, a wired network, a wireless network, the Internet, an intranet, a storage area network, or any other appropriate communication network. In various embodiments, there are 1, 2, 4, 7, 11, 25, 92, 297, 1001, or any other appropriate number of scanning nodes in communication with network scanning controller 202. In some embodiments, scanning nodes are located physically remotely from one another and from network scanning controller 202. In some embodiments, scanning nodes are implemented using cloud computing hardware. In some embodiments, scanning nodes are implemented using a different system than network scanning controller 202 (e.g., different cloud computing hardware). In some embodiments, scanning nodes do not comprise publicly visible information for identifying their association with network scanning controller 202.

[0018] Figure 3 is a block diagram illustrating an embodiment of a network scanning system. In some embodiments, the network scanning system of Figure 3 comprises a system for network scanning 104 of Figure 1. In the example shown, network scanning system 300 comprises network scanning controller 302, network scan database 303, network 304, anonymizing system 305, and a plurality of scanning nodes (e.g., scanning node 306, scanning node 308, scanning node 310, scanning node 312, scanning node 314, scanning node 316, and scanning node 318). Network scanning controller 302 comprises a system for receiving an indication to scan a set of network addresses (e.g., from a user or an administrator system). Network scanning controller 302 determines portions of the scan to be performed by a plurality of scanning nodes, directs scanning nodes to each perform a portion of the requested scan, receives back results from each of the scanning nodes, provides the results to a network scan database (e.g., network scan database 303), determines differences between current results and previous results (e.g., previous results also stored in network scan database 303), and providing the requestor an indication of results, results,

or a summary of the results. In some embodiments, network scanning controller 302 comprises a system implemented using cloud computing hardware. In some embodiments, network scanning controller 302 comprises a system to receive an indication of a set of scanning nodes (e.g., scanning node 308 and scanning node 316). In some embodiments, scanning nodes comprise network systems for scanning. Network scanning controller 302 communicates with network scan database 303 for storing network scanning results. Network scanning controller 302 communicates with the plurality of scanning nodes via anonymizing system 306. In various embodiments, anonymizing system 306 comprises an anonymizing virtual private network, a tunneling system, or any other appropriate anonymizing system. Anonymizing system 305 makes it difficult for scan target systems to identify the scanning controller behind the plurality of scanning systems. Anonymizing system 305 communicates with a plurality of scanning nodes using network 304. In various embodiments, network 304 comprises one or more of the following: a local area network, a wide area network, a wired network, a wireless network, the Internet, an intranet, a storage area network, or any other appropriate communication network. In various embodiments, there are 1, 2, 4, 7, 11, 25, 92, 297, 1001, or any other appropriate number of scanning nodes in communication with network scanning controller 302. In some embodiments, scanning nodes are located physically remotely from one another and from network scanning controller 302. In some embodiments, scanning nodes are implemented using cloud computing hardware. In some embodiments, scanning nodes are implemented using a different system than network scanning controller 302 (e.g., different cloud computing hardware). In some embodiments, scanning nodes do not comprise publicly visible information for identifying their association with network scanning controller 302.

[0019] Figure 4 is a block diagram illustrating an embodiment of a network system. In some embodiments, network system 400 comprises a network system of Figure 1 (e.g., network system 106). In the example shown, network system 400 comprises processor 402, data storage 404, and network interface 406. In some embodiments, network system 400 comprises an Internet connected asset (e.g., a desktop computer, a laptop computer, a smartphone, a tablet computer, a server system, an Internet-of-things device, or any other appropriate Internet connected asset). In various embodiments, processor 402 comprises a processor for executing instructions, processing data, responding to commands, etc. In various embodiments, processor 402 comprises a general-purpose processor, a microcontroller, a parallel processing system, a cluster of processors, or any other appropriate processor. In various embodiments, data storage 404 comprises a data storage for storing data, for storing instructions for processor 402, for storing configuration information, or for storing any other appropriate information. In various embodiments, data storage 404 comprises one or more of a volatile memory, a non-volatile memory, a magnetic memory, an optical memory, a

phase-change memory, a semiconductor memory, a disc memory, a tape memory, or any other appropriate memory. Network interface 406 comprises a network interface for communicating with a network. In the example shown, network interface 406 comprises network communications information 408 and a plurality of ports (e.g., port 410). In various embodiments, network communications information comprises network communications software, network communications settings, network communications data, or any other appropriate network communications information. The plurality of ports comprises physical ports (e.g., plugs for connecting cables to network system 400) or virtual ports (e.g., virtual communications channels identified by a virtual port number). In some embodiments, network interface 406 comprises a network address (e.g., a network address assigned by an external network addressing authority). In some embodiments, communication with network system 400 is specified by indicating the network address of network 400 along with a port number. In some embodiments, some ports of network interface 406 are configured for communication (e.g., comprising open ports) and some are configured to not respond to communication. In some embodiments, open port configuration information is stored in network communications information 408. In some embodiments, some ports are associated with one or more specific communications services (e.g., hypertext transmission protocol (HTTP), file transfer protocol (FTP), secure shell (SSH), etc.). In some embodiments, configuration information associating services with ports is stored in network communications information 408. In some embodiments, network communications information 408 comprises encryption information (e.g., a public SSH key, a certificate, etc.). In some embodiments, network communications information 408 comprises a network system name or names (e.g., a hostname, a domain name, a set of hostnames, a hostname pattern, etc.). In some embodiments, network communications information comprises text information associated with a service or a set of services (e.g., a welcome text, a connection refused text, a service not supported text, a file not found text, or any other appropriate text information). In some embodiments, network interface 406 comprises a set of network hardware (e.g., a modem) running a set of communications software that has been configured according to a set of communications specifications.

[0020] Figure 5 is a block diagram illustrating an embodiment of a network scanning controller. In some embodiments, network scanning controller 500 comprises network scanning controller 202 of Figure 2. In some embodiments, network scanning controller 500 comprises a server system. In the example shown, network scanning controller 500 comprises processor 502, data storage 504, and network interface 506. In various embodiments, processor 502 comprises a processor for executing instructions, processing data, responding to commands, etc. In various

embodiments, processor 502 comprises a general-purpose processor, a microcontroller, a parallel processing system, a cluster of processors, or any other appropriate processor. In some embodiments, processor 502 comprises network scanner 508. In various embodiments, network scanner 508 comprises software and/or hardware implementing hierarchical scanning system functionality. In some embodiments, network scanner 508 comprises a network scanner for determining a job plan. In some embodiments, a job plan comprises a set of scanning instructions. In some embodiments, a job plan comprises a job plan for scanning a set of network addresses using a set of available scanning nodes. In some embodiments, a job plan comprises a set of network addresses and ports to scan. In some embodiments, a job plan comprises one or more payloads and follow-up probes for interacting with a network device. In some embodiments, a job plan comprises one or more job portions. In some embodiments, job portions comprise divisible and independently executable job portions (e.g., job portions that can be run in parallel and do not depend on one another). In various embodiments, data storage 504 comprises a data storage for storing data, for storing instructions for processor 502, for storing configuration information, or for storing any other appropriate information. In various embodiments, data storage 504 comprises one or more of a volatile memory, a non-volatile memory, a magnetic memory, an optical memory, a phase-change memory, a semiconductor memory, a disc memory, a tape memory, or any other appropriate memory. In the example shown, data storage 504 comprises payload database 510 for storing payloads for providing to network devices. In some embodiments, a payload comprises a small data packet for probing a network device in order to elicit a response. Data storage 504 additionally comprises follow-up probe database 512 for storing follow-up probes for interacting with network devices. In some embodiments, a follow-up probe comprises software for interacting with a network device in order to determine information about the network device. In some embodiments, follow-up probe database 512 comprises a set of follow-up probes, each designed to interact with a network device in a specific way to retrieve data about the network device (e.g., establish a secure HTTP (HTTPS) connection and download an encrypted web page). In some embodiments, a follow-up probe is used to interact with a network device once it is determined that the follow-up probe is likely to succeed in receiving data from the network device. Data storage 504 additionally comprises network information database 514 for storing network information received as a result of interacting with network devices (e.g., using a payload or a follow-up probe). In some embodiments, network information is stored remotely (e.g., on a storage server, on a different hierarchical scanning system, on cloud storage, etc.). In the example shown, network interface 506 comprises a network interface for interacting with remote systems via a network. In various embodiments, network interface 506 comprises a network interface for providing a

payload, for executing communications for a follow-up probe, for receiving network information, or for any other appropriate purpose. In some embodiments, network interface 506 comprises a network interface configured for high bandwidth communication. In some embodiments, network interface 506 comprises a network interface for providing a job portion of a job plan to a scanning node. In some embodiments, network interface 506 comprises a network interface for receiving a job portion result from a scanning node.

[0021] Figure 6 is a block diagram illustrating an embodiment of a scanning node. In some embodiments, scanning node 600 comprises a scanning node of Figure 2 (e.g., scanning node 208). In some embodiments, scanning node 600 comprises a client system. In the example shown, scanning node 600 comprises processor 602, data storage 604, and network interface 606. In various embodiments, processor 602 comprises a processor for executing instructions, processing data, responding to commands, etc. In various embodiments, processor 602 comprises a general-purpose processor, a microcontroller, a parallel processing system, a cluster of processors, or any other appropriate processor. In some embodiments, processor 602 comprises network scanner 608. In various embodiments, network scanner 608 comprises software and/or hardware implementing hierarchical scanning system functionality. In various embodiments, data storage 604 comprises a data storage for storing data, for storing instructions for processor 602, for storing configuration information, or for storing any other appropriate information. In various embodiments, data storage 604 comprises one or more of a volatile memory, a non-volatile memory, a magnetic memory, an optical memory, a phase-change memory, a semiconductor memory, a disc memory, a tape memory, or any other appropriate memory. In the example shown, data storage 604 comprises payload database 610 for storing payloads for providing to network devices. In some embodiments, a payload comprises a small data packet for probing a network device in order to elicit a response. Data storage 604 additionally comprises follow-up probe database 612 for storing follow-up probes for interacting with network devices. In some embodiments, a follow-up probe comprises software for interacting with a network device in order to determine information about the network device. In some embodiments, follow-up probe database 612 comprises a set of follow-up probes, each designed to interact with a network device in a specific way to retrieve data about the network device (e.g., establish a secure HTTP (HTTPS) connection and download an encrypted web page). In some embodiments, a follow-up probe is used to interact with a network device once it is determined that the follow-up probe is likely to succeed in receiving data from the network device. Data storage 604 additionally comprises network information database 614 for storing network information received as a result of interacting with network devices (e.g., using a payload or a follow-up probe). In some embodiments, network information is stored remotely (e.g., on a storage

server, on a different hierarchical scanning system, on cloud storage, etc.). In the example shown, network interface 606 comprises a network interface for interacting with remote systems via a network. In various embodiments, network interface 606 comprises a network interface for providing a payload, for executing communications for a follow-up probe, for receiving network information, or for any other appropriate purpose. In some embodiments, network interface 606 comprises a network interface configured for high bandwidth communication.

[0022] In some embodiments, network scanner 608 executes a network scan according to instructions provided to it by a network scanning controller (e.g., network scanning controller 204 of Figure 2 or network scanning controller 302 of Figure 3). In some embodiments, instructions comprise a job portion (e.g., a job portion of a scanning job). In various embodiments, instructions comprise one or more Internet addresses, one or more ports, one or more payloads, one or more follow-up probes, or any other appropriate scan instructions.

[0023] Figure 7 is a flow diagram illustrating an embodiment of a process for scanning a network. In some embodiments, the process of Figure 7 is executed by network scanning controller 204 of Figure 2. In the example shown, in 700, an indication is received to scan a set of network addresses. In 702, a set of available scanning nodes is determined. In 704, a job plan is determined for scanning the set of network addresses using the set of available scanning nodes, wherein the job plan comprises one or more job portions. In 706, a job portion is selected. In 708, a scanning node is selected. In some embodiments, the scanning node is selected based at least in part on the job portion. In various embodiments, the scanning node comprises a randomly chosen scanning node, a pseudorandomly chosen scanning node, the next scanning node in a list of scanning nodes, the scanning node with the least processor load, the scanning node has not been recently selected, the scanning node that has least recently scanned addresses of the job portion, the next scanning node to scan the addresses of the job portion according to a predetermined list, a scanning node chosen based at least in part on its physical location, or a scanning node chosen in any other appropriate way. In 710, the job portion is provided to the scanning node. In 712, a job portion result is received from the scanning node. In 714, the job portion result is scanned. In some embodiments, the job portion result is stored by the network scanning controller. In some embodiments, the job portion result is provided to a network scan database for storage. In 716, the job plan is modified based on the job portion result, if necessary. In various embodiments, the job plan is modified to scan additional address, to scan additional ports, to execute additional scan types, to execute additional follow-up probes, or in any other appropriate way. In 718, it is determined whether there are more job portions (e.g., of the job plan). In the event it is determined that there are more job

portions, control passes to 706. In the event it is determined that there are not more job portions, the process ends.

[0024] Figure 8 is a flow diagram illustrating an embodiment of a process for determining a job plan. In some embodiments, the process of Figure 8 implements 704 of Figure 7. In the example shown, in 800, a set of scanning commands for the scan is determined. In various embodiments, scanning commands comprise shell commands, payloads, follow-up probes, or any other appropriate scanning commands. In some embodiments, scanning commands comprise part of a received indication to scan (e.g., the indication to scan indicates scanning commands for the scan). In some embodiments, scanning commands are determined based at least in part on indicated commands to scan (e.g., additional scanning commands are determined beyond indicated commands to scan, commands to scan are modified, etc.). In 802, a set of ports for the scan is determined. In some embodiments, ports to scan comprise part of a received indication to scan (e.g., the indication to scan indicates ports for the scan). In some embodiments, ports are determined based at least in part on indicated commands to scan (e.g., additional ports are determined beyond indicated commands to scan, port are modified, etc.). In 804, a job portion size is determined. In various embodiments, a job portion size is determined based at least in part on one or more of: a predetermined job portion size, a number of available scanning nodes, a number of addresses to scan, a distribution of addresses to scan, physical locations associated with addresses to scan, historical scanning data, current and historical organizational network footprints, or any other appropriate factor. In 806, a portion of the set of network addresses to scan of the job portion size is selected. In various embodiments, a contiguous set of network addresses is selected, a discontiguous set of network addresses is selected, a random set of network addresses is selected, a pseudorandom set of network addresses is selected, a set of network addresses with common organizational ownership or common network management is selected, a set of network addresses with multiple distinct gateway devices and/or traffic concentrators is selected, or a set of network addresses is selected in any other appropriate way. In 808, a job portion of the job plan is created comprising instructions to scan the portion of the set of ports using the set of network commands. In 810, it is determined whether there are more addresses (e.g., more addresses of the set of network addresses to scan that have not been allocated to a job portion). In the event it is determined that there are more addresses, control passes to 806. In the event it is determined that there are not more addresses, the process ends.

[0025] Figure 9 is a flow diagram illustrating an embodiment of a process for modifying a job plan based on a job portion result, if necessary. In some embodiments, the process of Figure 9

implements 716 of Figure 7. In the example shown, in 900, it is determined whether the job portion comprises all previously seen results. In some embodiments, previously seen results comprise previously seen devices or services, or previously seen addresses or ports that produce no response. In the event it is determined that the job portion result comprises all previously seen results, the process ends (e.g., no modification of the job plan is made). In the event it is determined that the job portion result does not comprise all previously seen results, control passes to 902. In 902, it is determined whether the job portion result indicates a previously online device or service is offline. In the event it is determined that no previously online device or service is offline, control passes to 908. In the event it is determined that a previously seen device or service is offline, control passes to 904. In 904, a reattempt of the scan using one or more different scanning nodes is added to the job. For example, a scan is added to a different job portion that is assigned to a different selected scanning node than the scan had been previously assigned to so that the resulting scan does not appear to originate from the same source. In various embodiments, the reattempt is conducted using a randomly chosen scanning node, a scanning node that has scanned the previously online device or service least recently, a reserve scanning node only used in the case of determining whether a previously online device has gone offline, or any other appropriate scanning node. In 906, one or more nearby scan locations are added to the job. In some embodiments, scanning one or more nearby scan locations can find a device or service that has been moved. In some embodiments, nearby scan locations comprise nearby addresses. In some embodiments, nearby scan locations comprise nearby ports. In 908, it is determined whether the job portion result indicates a new device or service. In the event it is determined that the job portion result does not indicate a new device or service, the process ends. In the event it is determined that the job portion result indicates a new device or service, control passes to 910. In 910, one or more follow-up scans are added to the job. In some embodiments, follow-up scans comprise scans to collect further information about a newly discovered device or service. In some embodiments, follow-up scans comprise follow-up probes.

[0026] Although the foregoing embodiments have been described in some detail for purposes of clarity of understanding, the invention is not limited to the details provided. There are many alternative ways of implementing the invention. The disclosed embodiments are illustrative and not restrictive.

CLAIMS

1. A system for scanning a network, comprising:

an interface configured to:

receive an indication to scan a set of network addresses; and

a processor configured to:

determine a set of available scanning nodes;

determine a job plan for scanning the set of network addresses using the set of available scanning nodes, wherein the job plan comprises one or more job portions; and

for a job portion of the one or more job portions:

select a scanning node of the set of available scanning nodes; and

provide the job portion to the scanning node.

2. The system of claim 1, wherein the processor communicates with the scanning node via an anonymizing system.

3. The system of claim 1, wherein a job portion of the one or more job portions comprises instructions to scan a portion of the set of network addresses.

4. The system of claim 1, wherein a job portion of the one or more job portions comprises a set of ports.

5. The system of claim 1, wherein a job portion of the one or more job portions comprises a set of scanning commands.

6. The system of claim 1, wherein the scanning node is selected based at least in part on a job portion of the one or more job portions.

7. The system of claim 1, wherein the scanning node is selected based at least in part on a physical location of the scanning node.

8. The system of claim 1, wherein the scanning node is selected based at least in part on the scanning node having not been recently selected.

9. The system of claim 1, wherein the scanning node is selected randomly or pseudorandomly.

10. The system of claim 1, wherein the processor is further configured to provide the job plan result to a database for storage.

11. The system of claim 1, wherein the processor is further configured to receive a scanning result from the scanning node.

12. The system of claim 11, wherein the processor is further configured to modify the job plan based at least in part on a scanning result.

13. The system of claim 12, wherein modifying the job plan comprises adding a reattempt of a scan of a job portion of the one or more job portions using one or more different scanning nodes.

14. The system of claim 12, wherein modifying the job plan comprises adding one or more nearby scan locations to the job portion.

15. The system of claim 12, wherein modifying the job plan comprises adding one or more follow-up scans to the job portion.

16. The system of claim 1, wherein the set of available scanning nodes are implemented using cloud computing hardware.

17. The system of claim 1, wherein the set of available scanning nodes are implemented using a different system than the system for scanning the network.

18. The system of claim 1, wherein the set of available scanning nodes do not comprise publicly visible information for identifying their association with the system for scanning the network.

19. A method for scanning a network, comprising:

receiving an indication to scan a set of network addresses;

determining, using a processor, a set of available scanning nodes;

determining a job plan for scanning the set of network addresses using the set of available scanning nodes, wherein the job plan comprises one or more job portions; and

for a job portion of the one or more job portions:

selecting a scanning node of the set of available scanning nodes; and

providing the job portion to the scanning node

.

20. A computer program product for scanning a network, the computer program product being embodied in a non-transitory computer readable storage medium and comprising computer instructions for:

receiving an indication to scan a set of network addresses;

determining a set of available scanning nodes;

determining a job plan for scanning the set of network addresses using the set of available scanning nodes, wherein the job plan comprises one or more job portions; and

for a job portion of the one or more job portions:

selecting a scanning node of the set of available scanning nodes;

providing the job portion to the scanning node.

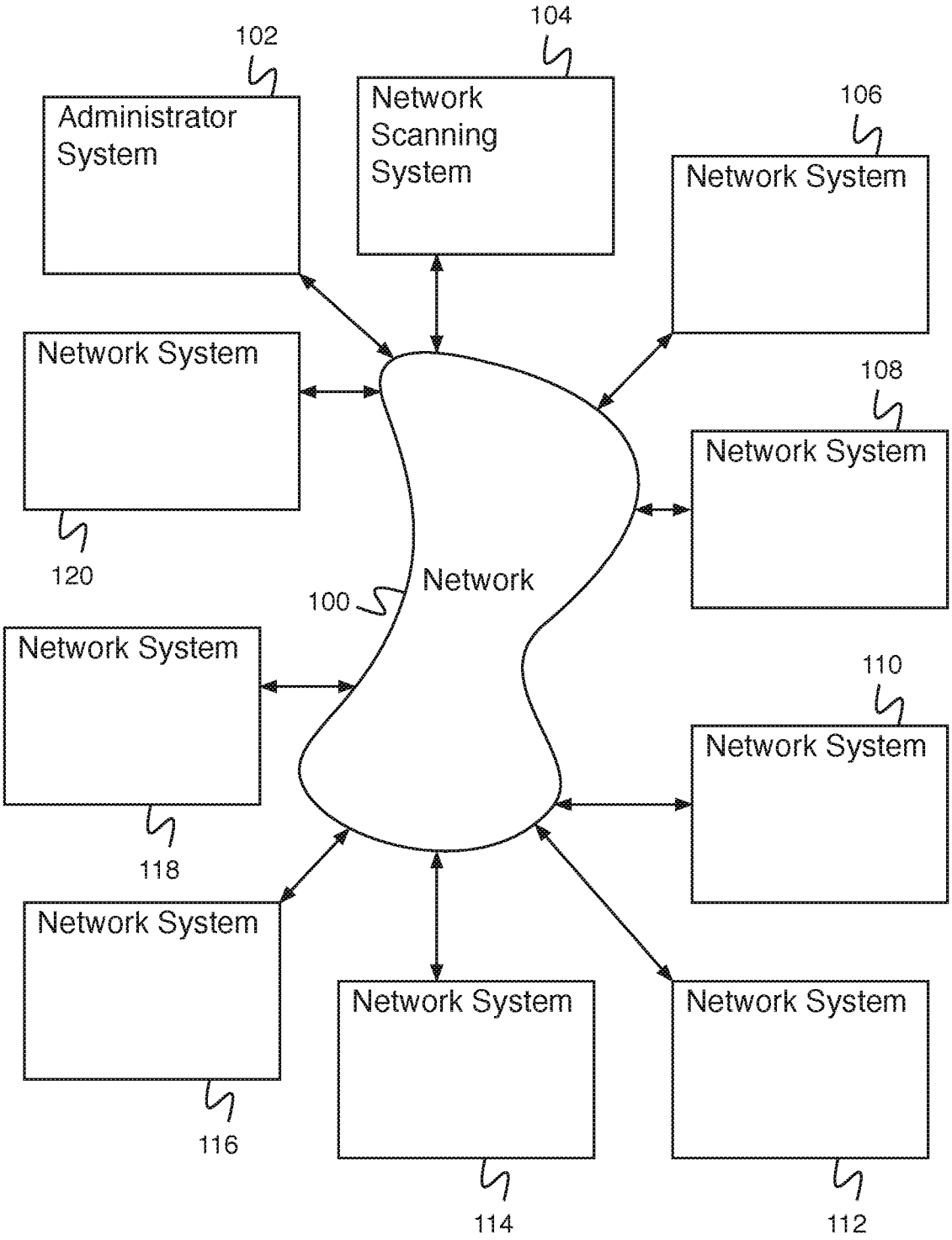
102
Administrator System
104
Network Scanning System
106
Network System
Network System
108
Network System
120
100
Network
Network System
110
Network System
118
Network System
Network System
Network System
116
114
112

Fig. 1

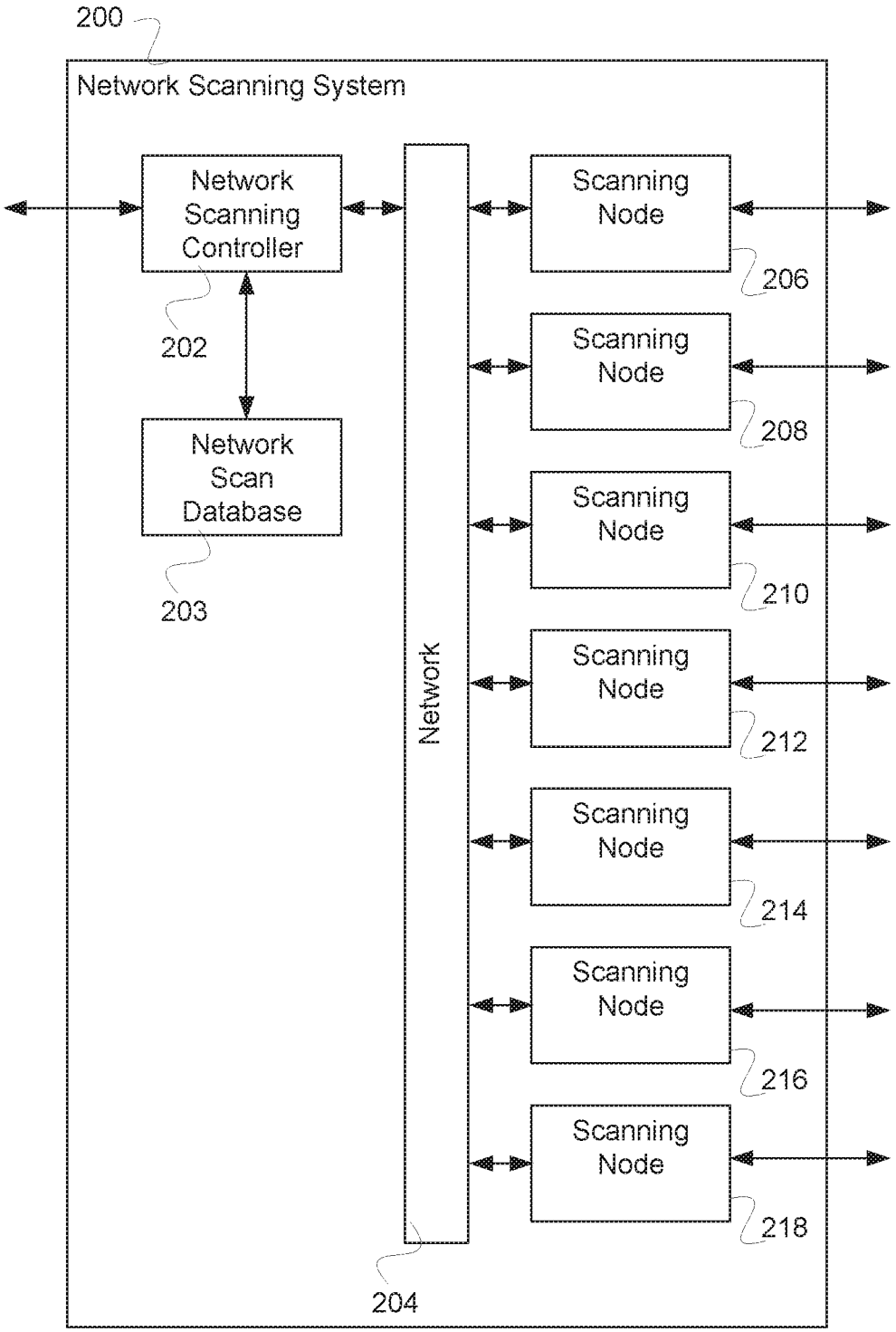
200
Network Scanning System
Network Scanning Controller
202
Network Scan Database
203
Network
204
Scanning Node
206
Scanning Node
208
Scanning Node
210
Scanning Node
212
Scanning Node
214
Scanning Node
216
Scanning Node
218

FIG. 2

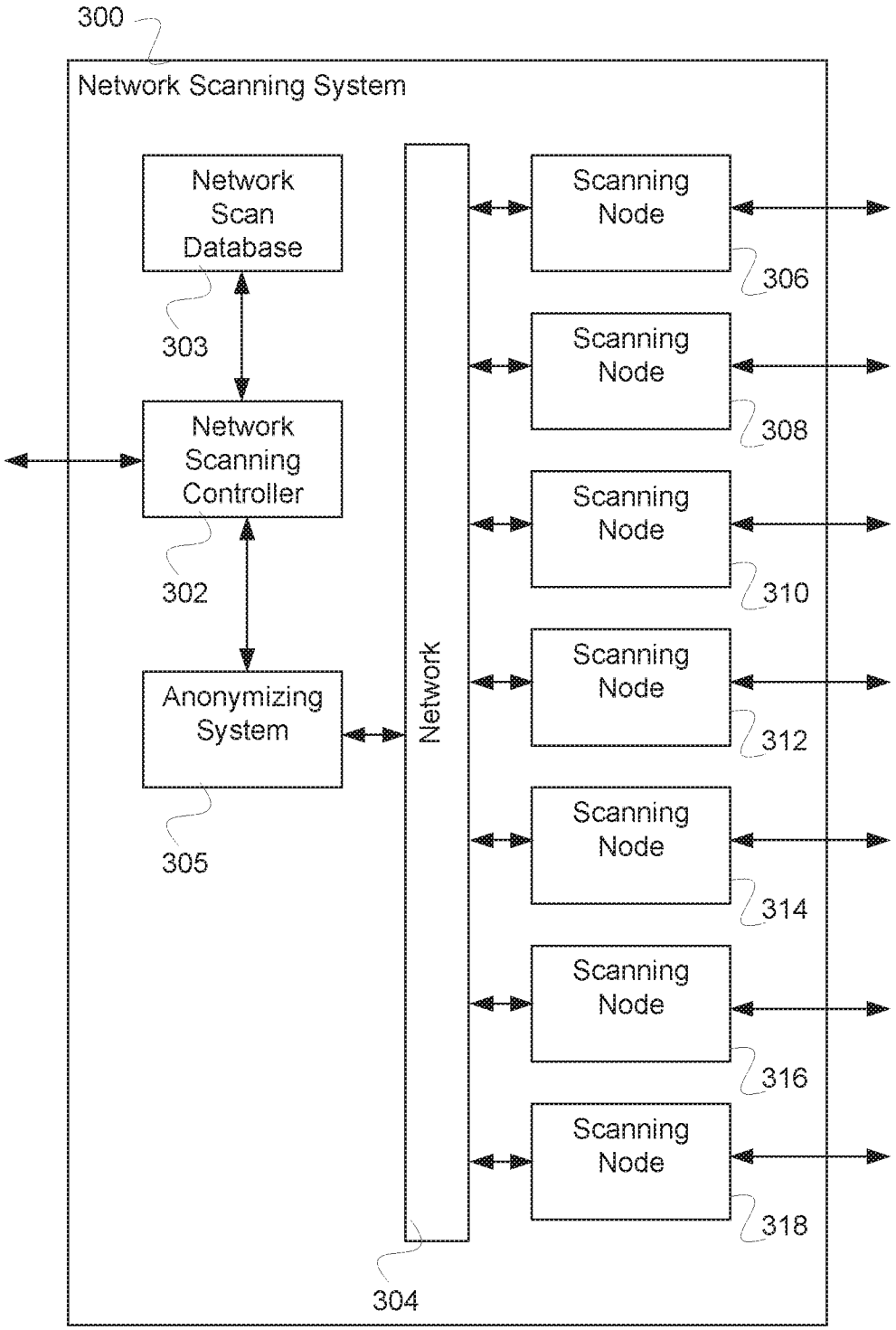
300
Network Scanning System
Network Scan Database
303
Network Scanning Controller
302
Anonymizing System
305
Network
304
Scanning Node
306
Scanning Node
308
Scanning Node
310
Scanning Node
312
Scanning Node
314
Scanning Node
316
Scanning Node
318

FIG. 3

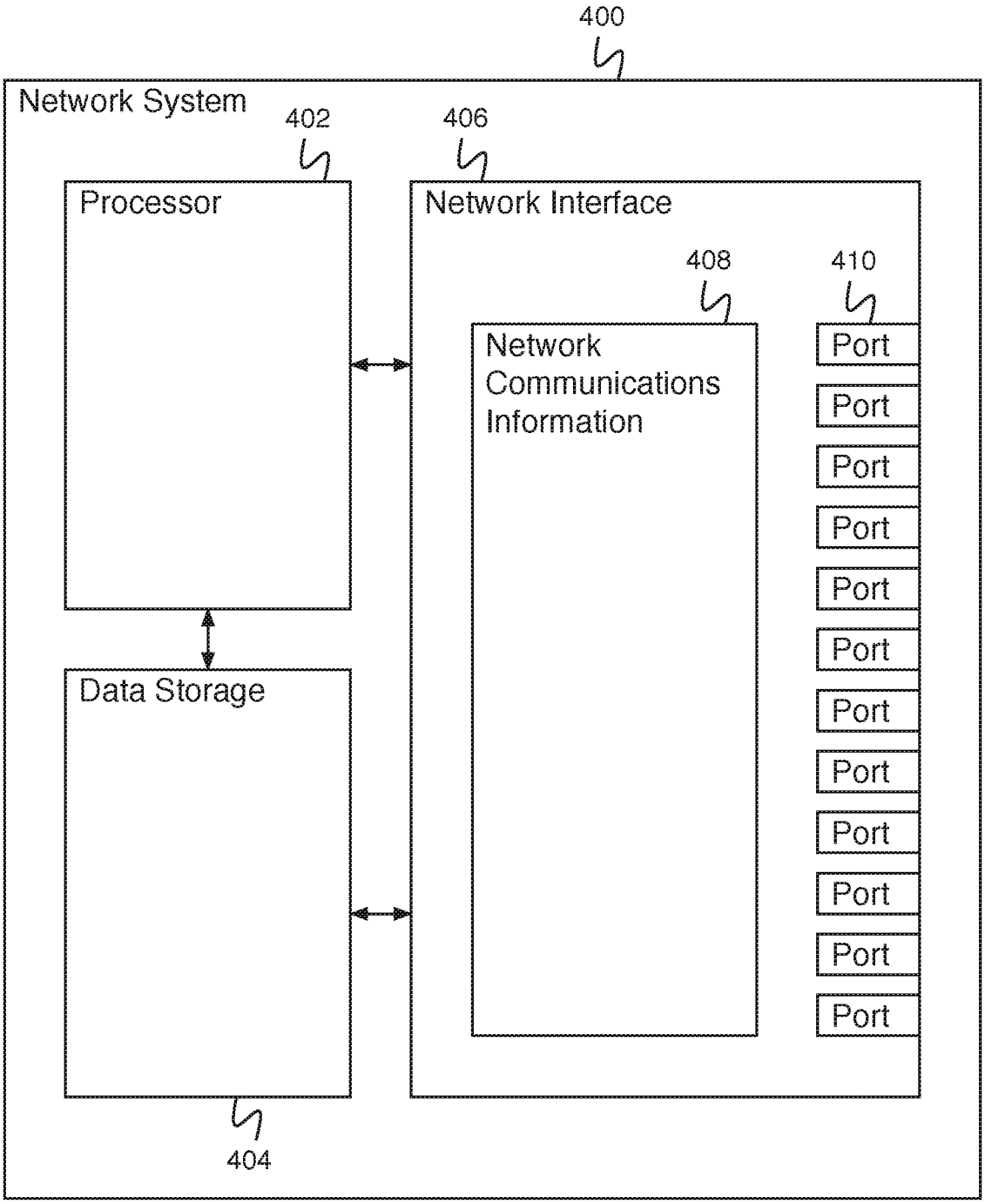
400
Network System
402
406
Processor
Network Interface
408
410
Network Communications Information
Port
Port
Port
Port
Port
Port
Port
Port
Port
Port
Port
Port
Data Storage
404

Fig. 4

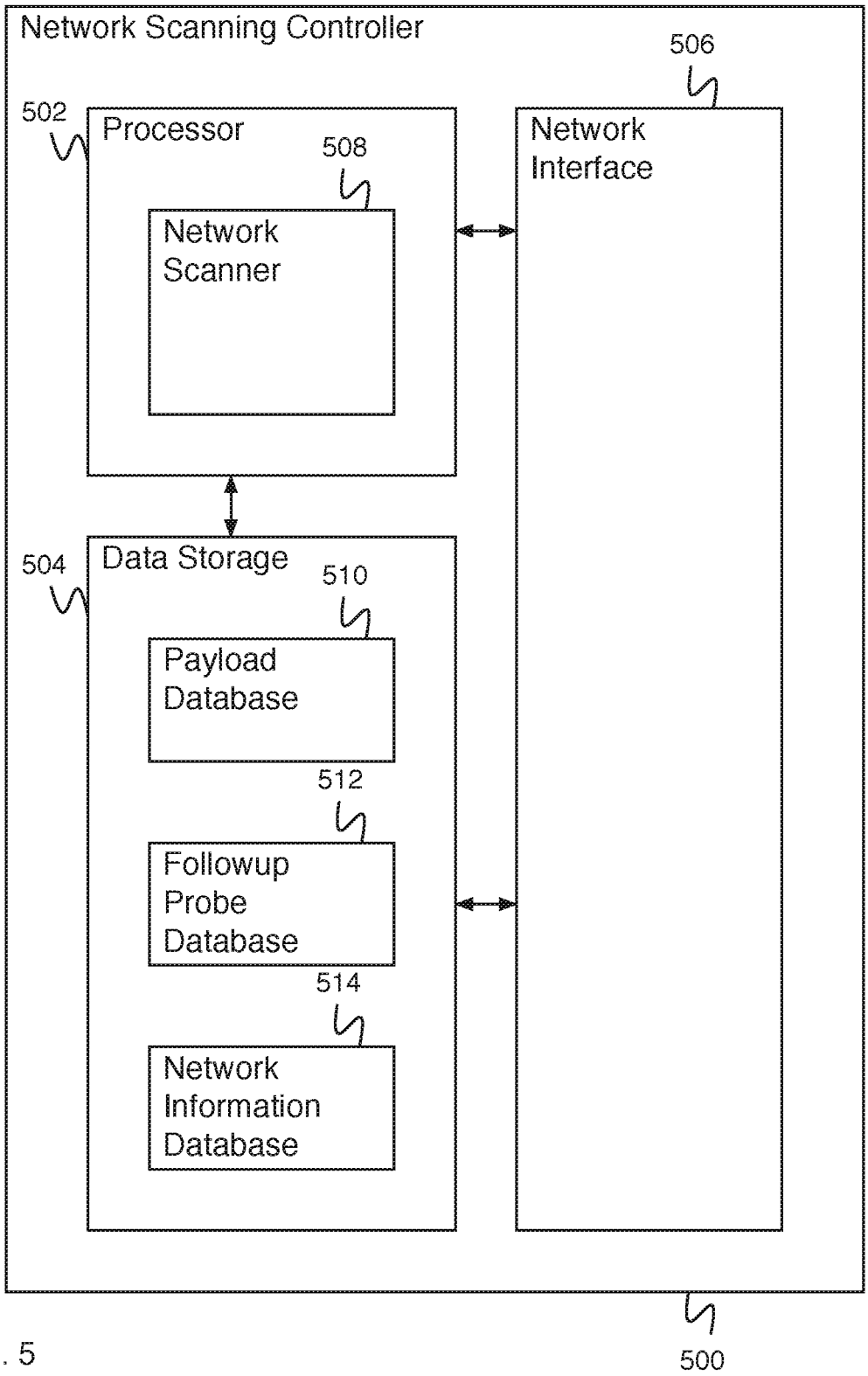
Network Scanning Controller
506
502
Processor
508
Network Scanner
Network Interface
504
Data Storage
510
Payload Database
512
Followup Probe Database
514
Network Information Database
500

Fig. 5

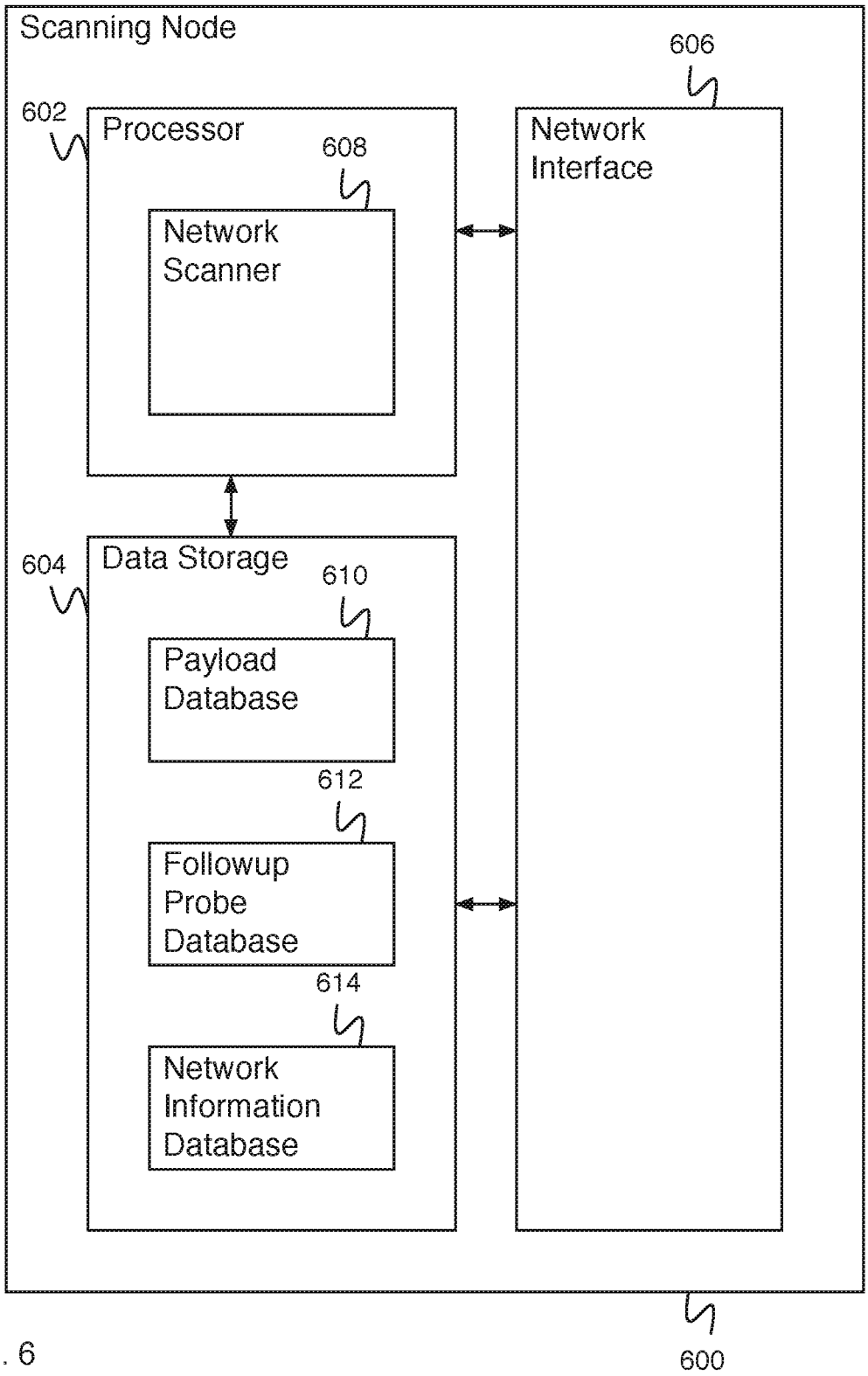
Scanning Node
606
602
Processor
608
Network Scanner
Network Interface
604
Data Storage
610
Payload Database
612
Followup Probe Database
614
Network Information Database
600

Fig. 6

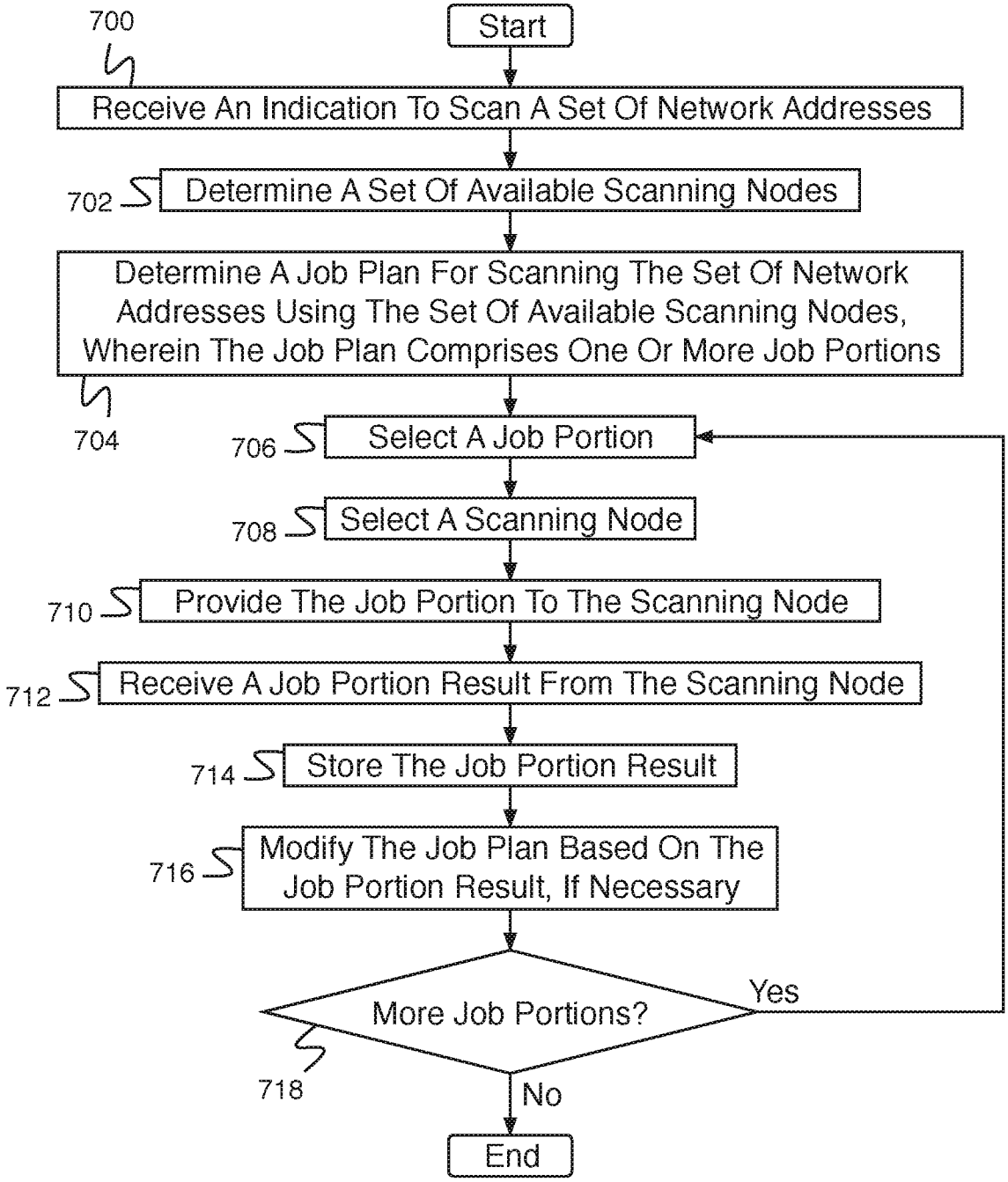
Start
700
Receive An Indication To Scan A Set Of Network Addresses
702
Determine A Set Of Available Scanning Nodes
Determine A Job Plan For Scanning The Set Of Network Addresses Using The Set Of Available Scanning Nodes, Wherein The Job Plan Comprises One Or More Job Portions
704
706
Select A Job Portion
708
Select A Scanning Node
710
Provide The Job Portion To The Scanning Node
712
Receive A Job Portion Result From The Scanning Node
714
Store The Job Portion Result
716
Modify The Job Plan Based On The Job Portion Result, If Necessary
More Job Portions?
718
Yes
No
End

Fig. 7

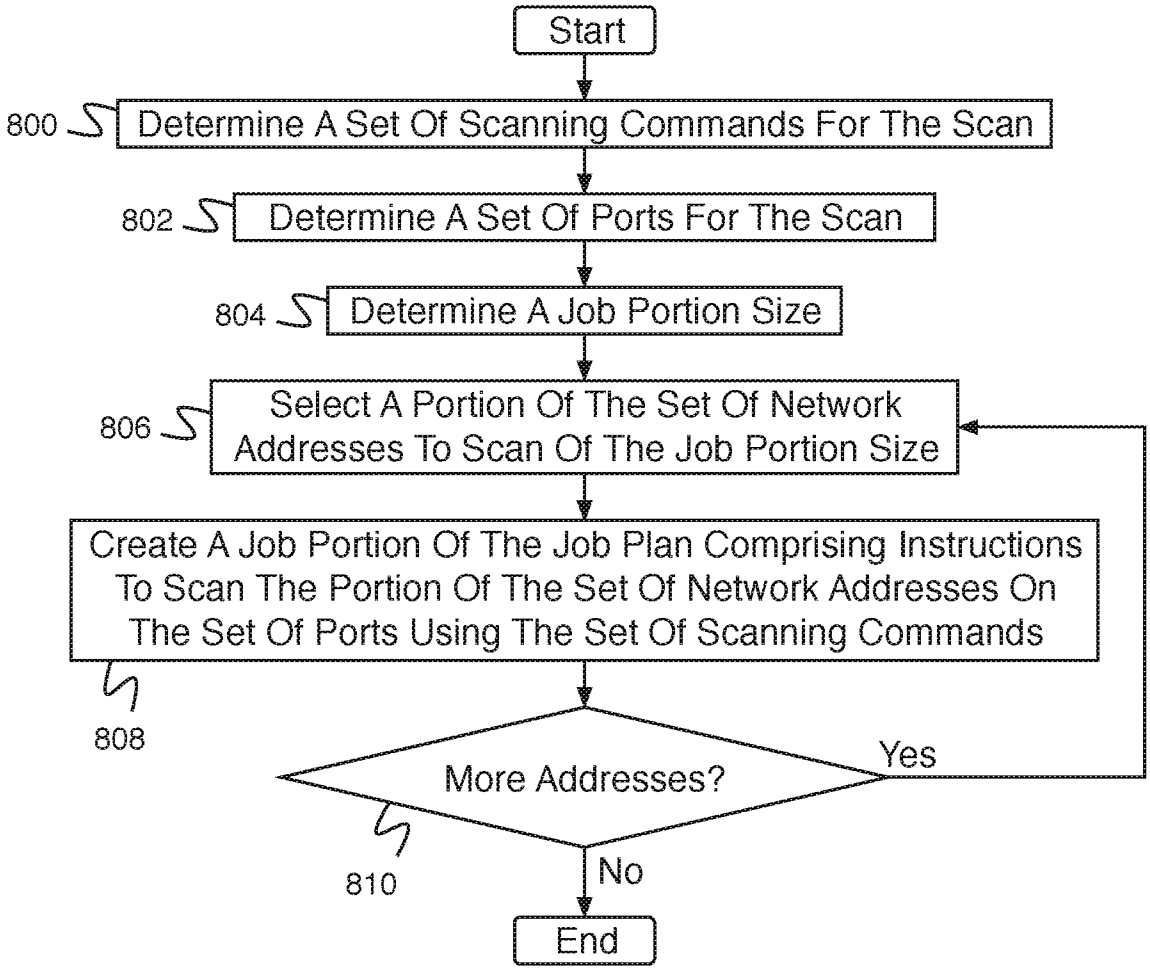
Start
800
Determine A Set Of Scanning Commands For The Scan
802
Determine A Set Of Ports For The Scan
804
Determine A Job Portion Size
806
Select A Portion Of The Set Of Network Addresses To Scan Of The Job Portion Size
Create A Job Portion Of The Job Plan Comprising Instructions To Scan The Portion Of The Set Of Network Addresses On The Set Of Ports Using The Set Of Scanning Commands
808
More Addresses?
Yes
810
No
End

Fig. 8

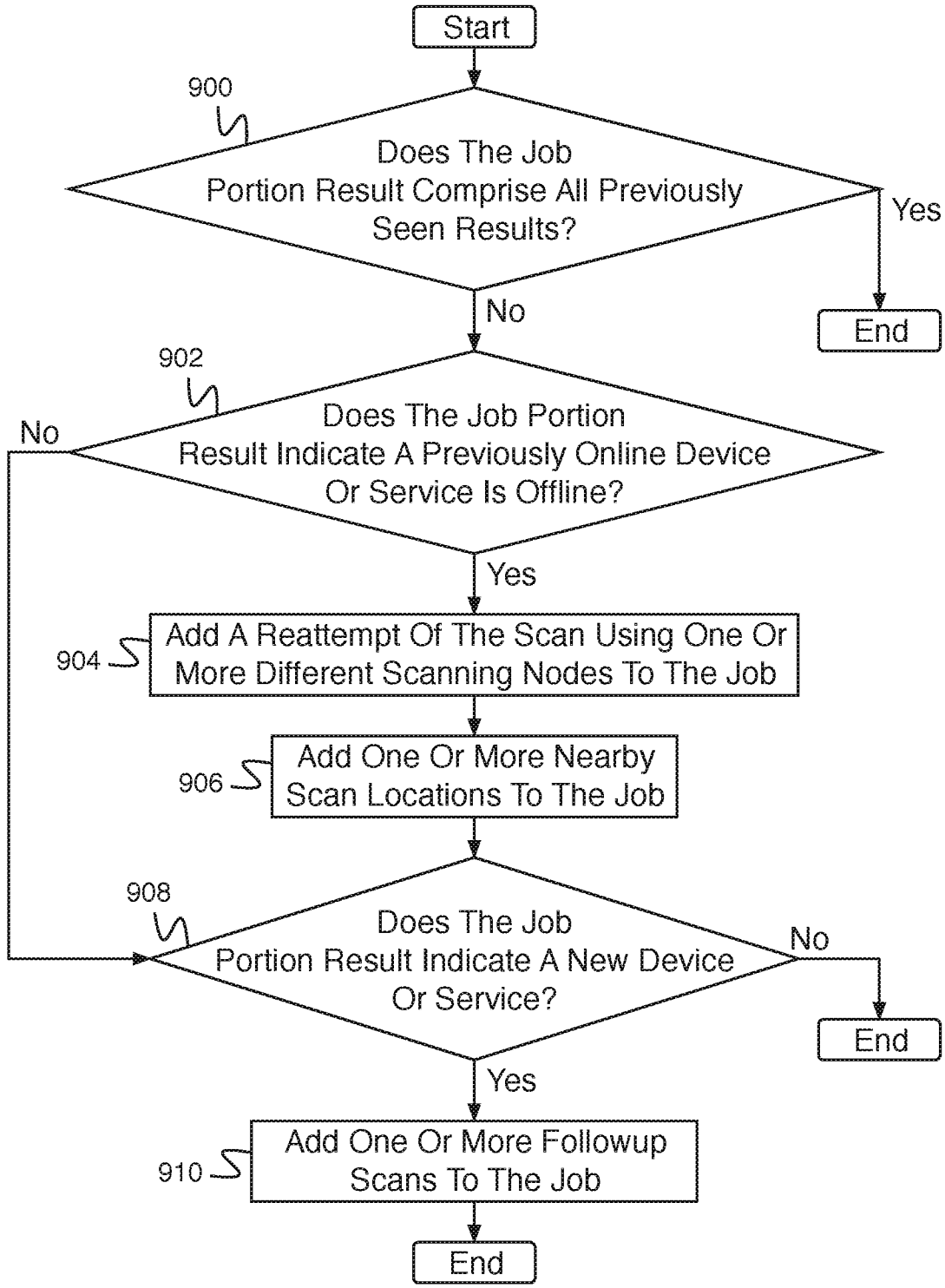
Start
900
Does The Job Portion Result Comprise All Previously Seen Results?
Yes
End
No
902
Does The Job Portion Result Indicate A Previously Online Device Or Service Is Offline?
No
Yes
904
Add A Reattempt Of The Scan Using One Or More Different Scanning Nodes To The Job
906
Add One Or More Nearby Scan Locations To The Job
908
Does The Job Portion Result Indicate A New Device Or Service?
No
End
Yes
910
Add One Or More Followup Scans To The Job
End

Fig. 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 17/65430

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - H04L 29/06, H04L 29/08, G06F 15/163, G06F 15/173 (2018.01)
CPC - H04L 61/103, H04L 41/12, H04L 61/1541, H04L 29/08648, H04L 29/12113, H04L 61/2015, H04L 63/1416

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)
See Search History Document

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
See Search History Document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 2007/0005738 A1 (Alexion-Tiernan et al.) 04 January 2007 (04.01.2007) entire document, especially: para [0004], [0009], [0012], [0014], [0016], [0032], [0035], [0041]	1, 3, 5, 10-12, 15, 19, 20 --- 2, 4, 6-9, 13, 14, 16-18
Y	US 9,544,327 B1 (International Business Machines Corporation) 10 January 2017 (10.01.2017) entire document, especially: col 2, ln 7-10; col 22, ln 49-55; col 24, ln 45-56	2, 16, 18
Y	US 2014/0373161 A1 (FoxGuardSolutions, Inc.) 18 December 2014 (18.12.2014) entire document, especially: Abstract, para [0049], [0056], [0059]	4, 13
Y	US 2012/0192276 A1 (Andrews et al.) 26 July 2012 (26.07.2012) entire document, especially: para [0002], [0017], [0028], [0029], [0030]	6, 8, 14, 17
Y	US 2016/0314307 A1 (LARC Networks, Inc) 27 October 2016 (27.10.2016) entire document, especially: Abstract, para [0032], [0035]	7, 9
A	CN106161450 A (SHANGHAI CTRIP BUSINESS CO LTD) 23 November 2016 (23.11.2016) entire document	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
01 February 2018	06 MAR 2018
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774