US 20180103117A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2018/0103117 A1**

Mosko (43) **Pub. Date: Apr. 12, 2018**

(54) **DISTRIBUTED CONSENSUS IN A CONTENT CENTRIC NETWORK**

(71) Applicant: **CISCO TECHNOLOGY, INC.**, SAN JOSE, CA (US)

(72) Inventor: **Marc E. Mosko**, Santa Cruz, CA (US)

(73) Assignee: **CISCO TECHNOLOGY, INC.**, SAN JOSE, CA (US)

(21) Appl. No.: **15/289,739**

(22) Filed: **Oct. 10, 2016**

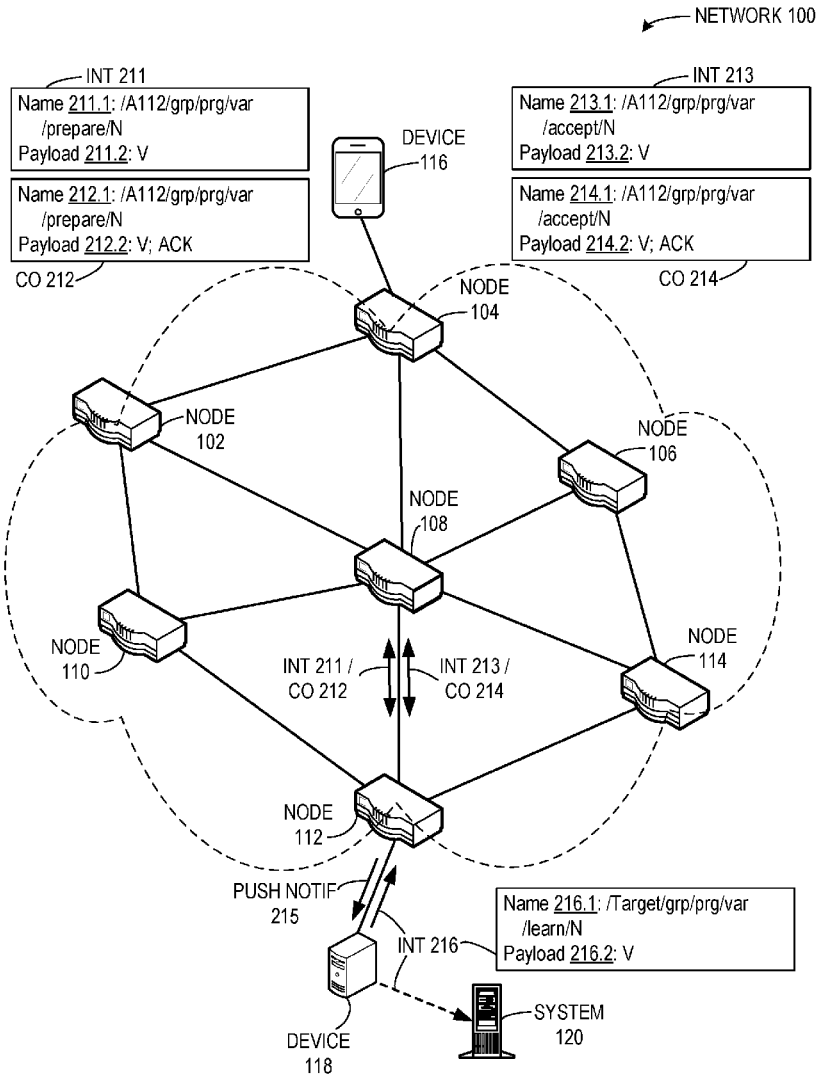
Publication Classification

(51) **Int. Cl.**
H04L 29/08 (2006.01)

(52) **U.S. Cl.**
CPC ***H04L 67/32*** (2013.01); ***H04L 67/10*** (2013.01)

(57) **ABSTRACT**

One embodiment provides a system that facilitates distributed consensus in a content centric network. During operation, the system generates, by a coordinator for a plurality of nodes, a first interest that indicates a request to approve a proposed value for a variable. In response to receiving a positive acknowledgment of the first interest from a majority of the nodes, the coordinator generates a second interest that indicates a request to accept the proposed value. A name for the first interest and a name for the second interest include an identifier of the variable and a round number. A payload of the first interest and a payload of the second interest include the proposed value. In response to receiving a positive acknowledgement of the second interest from the majority of the nodes, the system generates a notification indicating that an agreed-upon value for the variable is the proposed value.



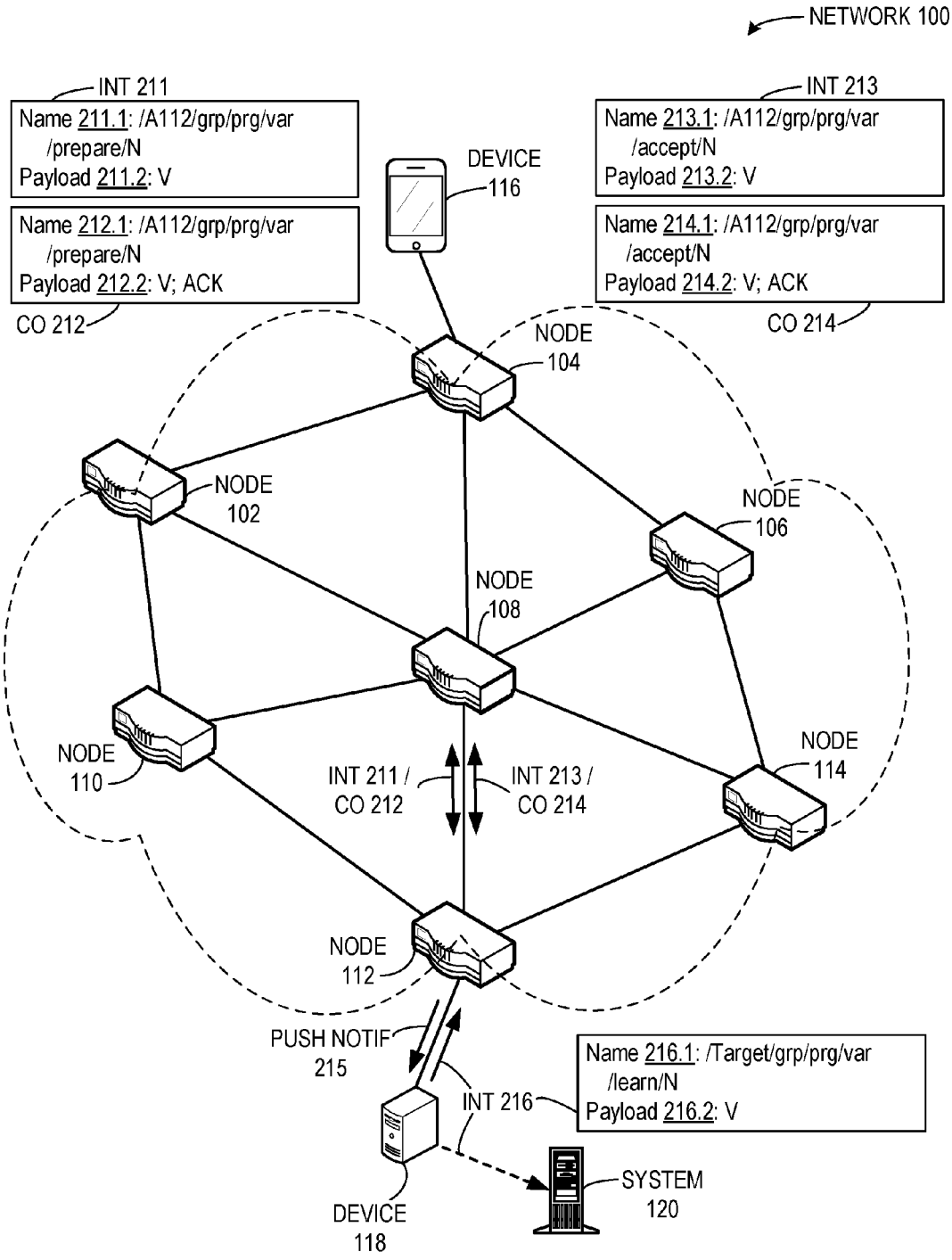

NETWORK 100
INT 211
Name 211.1: /A112/grp/prg/var /prepare/N
Payload 211.2: V
Name 212.1: /A112/grp/prg/var /prepare/N
Payload 212.2: V; ACK
CO 212
DEVICE 116
INT 213
Name 213.1: /A112/grp/prg/var /accept/N
Payload 213.2: V
Name 214.1: /A112/grp/prg/var /accept/N
Payload 214.2: V; ACK
CO 214
NODE 104
NODE 102
NODE 106
NODE 108
NODE 110
NODE 112
NODE 114
INT 211 / CO 212
INT 213 / CO 214
PUSH NOTIF 215
INT 216
Name 216.1: /Target/grp/prg/var /learn/N
Payload 216.2: V
DEVICE 118
SYSTEM 120

FIG. 1

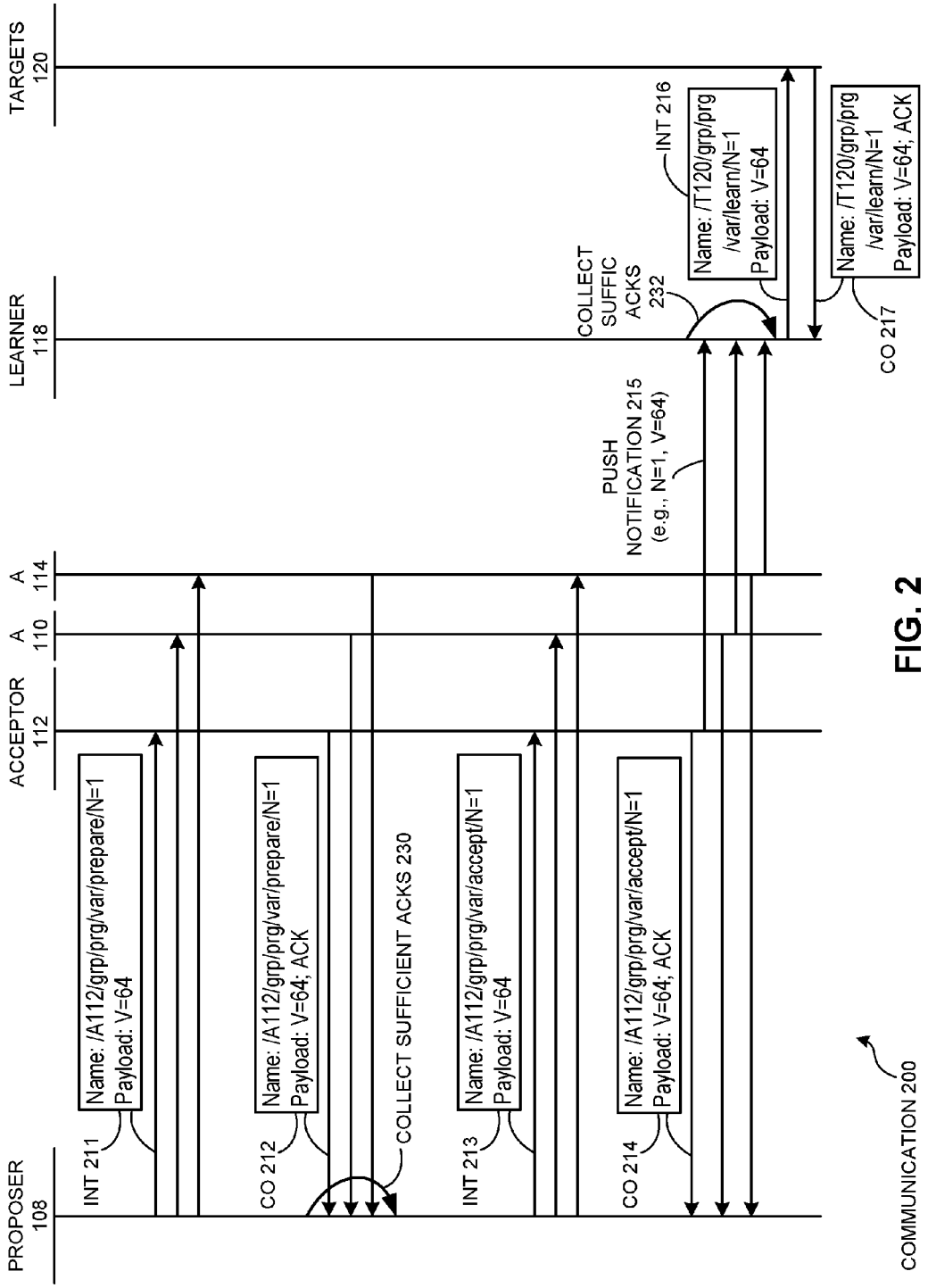

PROPOSER 108
ACCEPTOR A 112
A 110
A 114
LEARNER 118
TARGETS 120
INT 211
Name: /A112/grp/prg/var/prepare/N=1
Payload: V=64
CO 212
Name: /A112/grp/prg/var/prepare/N=1
Payload: V=64; ACK
COLLECT SUFFICIENT ACKS 230
INT 213
Name: /A112/grp/prg/var/accept/N=1
Payload: V=64
CO 214
Name: /A112/grp/prg/var/accept/N=1
Payload: V=64; ACK
PUSH NOTIFICATION 215 (e.g., N=1, V=64)
COLLECT SUFFIC ACKS 232
INT 216
Name: /T120/grp/prg/var/learn/N=1
Payload: V=64
CO 217
Name: /T120/grp/prg/var/learn/N=1
Payload: V=64; ACK
COMMUNICATION 200

FIG. 2

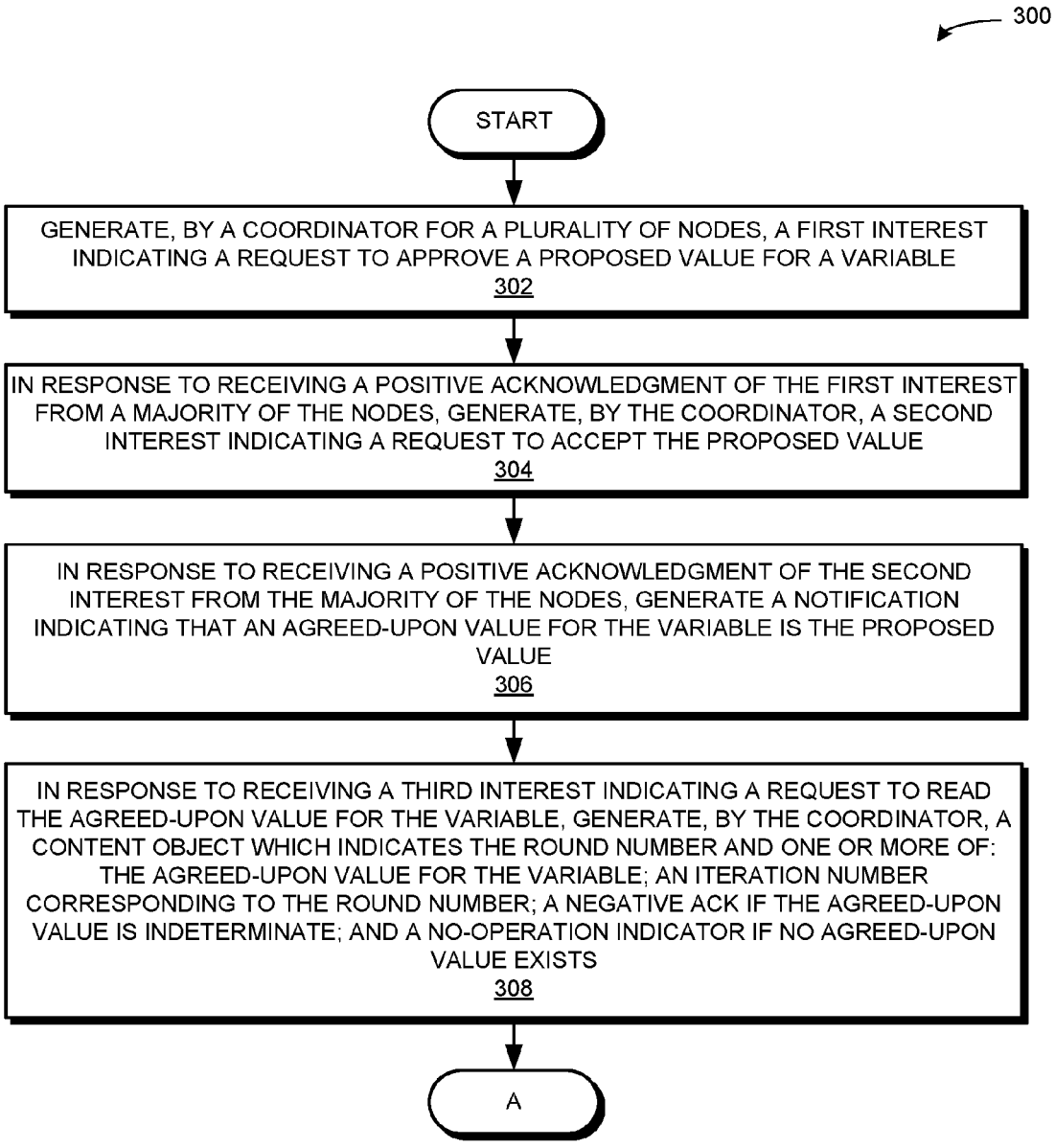

300
START
GENERATE, BY A COORDINATOR FOR A PLURALITY OF NODES, A FIRST INTEREST INDICATING A REQUEST TO APPROVE A PROPOSED VALUE FOR A VARIABLE
302
IN RESPONSE TO RECEIVING A POSITIVE ACKNOWLEDGMENT OF THE FIRST INTEREST FROM A MAJORITY OF THE NODES, GENERATE, BY THE COORDINATOR, A SECOND INTEREST INDICATING A REQUEST TO ACCEPT THE PROPOSED VALUE
304
IN RESPONSE TO RECEIVING A POSITIVE ACKNOWLEDGMENT OF THE SECOND INTEREST FROM THE MAJORITY OF THE NODES, GENERATE A NOTIFICATION INDICATING THAT AN AGREED-UPON VALUE FOR THE VARIABLE IS THE PROPOSED VALUE
306
IN RESPONSE TO RECEIVING A THIRD INTEREST INDICATING A REQUEST TO READ THE AGREED-UPON VALUE FOR THE VARIABLE, GENERATE, BY THE COORDINATOR, A CONTENT OBJECT WHICH INDICATES THE ROUND NUMBER AND ONE OR MORE OF: THE AGREED-UPON VALUE FOR THE VARIABLE; AN ITERATION NUMBER CORRESPONDING TO THE ROUND NUMBER; A NEGATIVE ACK IF THE AGREED-UPON VALUE IS INDETERMINATE; AND A NO-OPERATION INDICATOR IF NO AGREED-UPON VALUE EXISTS
308
A

FIG. 3A

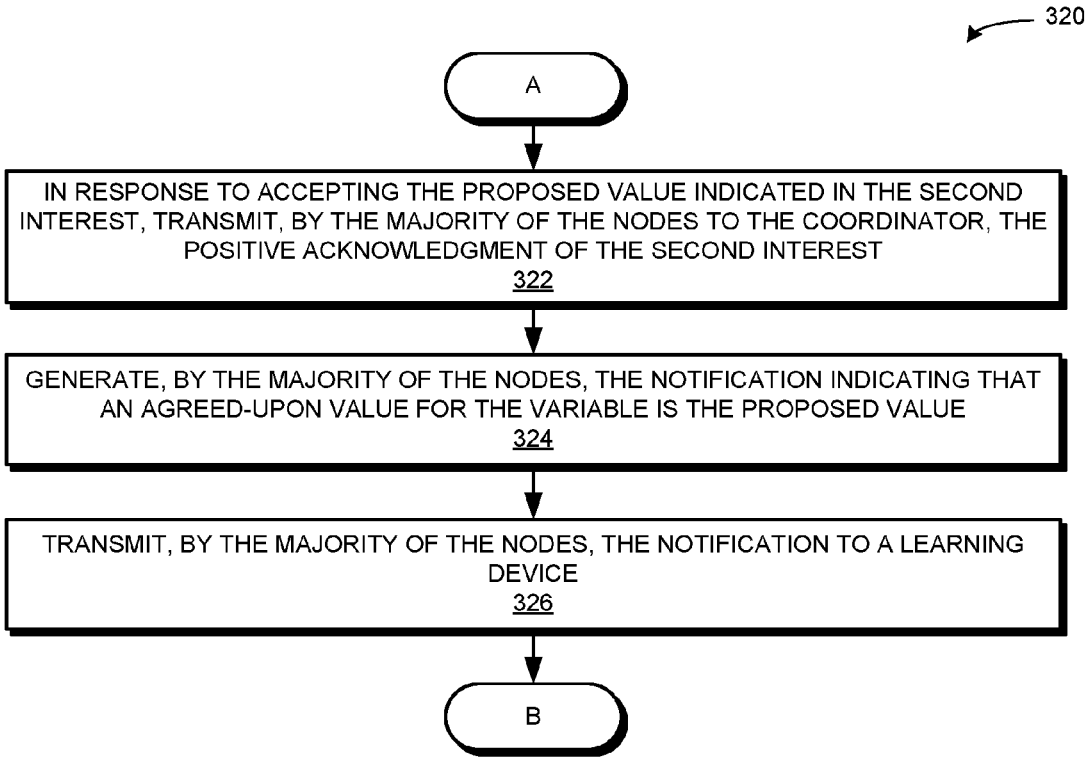

320
A
IN RESPONSE TO ACCEPTING THE PROPOSED VALUE INDICATED IN THE SECOND INTEREST, TRANSMIT, BY THE MAJORITY OF THE NODES TO THE COORDINATOR, THE POSITIVE ACKNOWLEDGMENT OF THE SECOND INTEREST
322
GENERATE, BY THE MAJORITY OF THE NODES, THE NOTIFICATION INDICATING THAT AN AGREED-UPON VALUE FOR THE VARIABLE IS THE PROPOSED VALUE
324
TRANSMIT, BY THE MAJORITY OF THE NODES, THE NOTIFICATION TO A LEARNING DEVICE
326
B

FIG. 3B

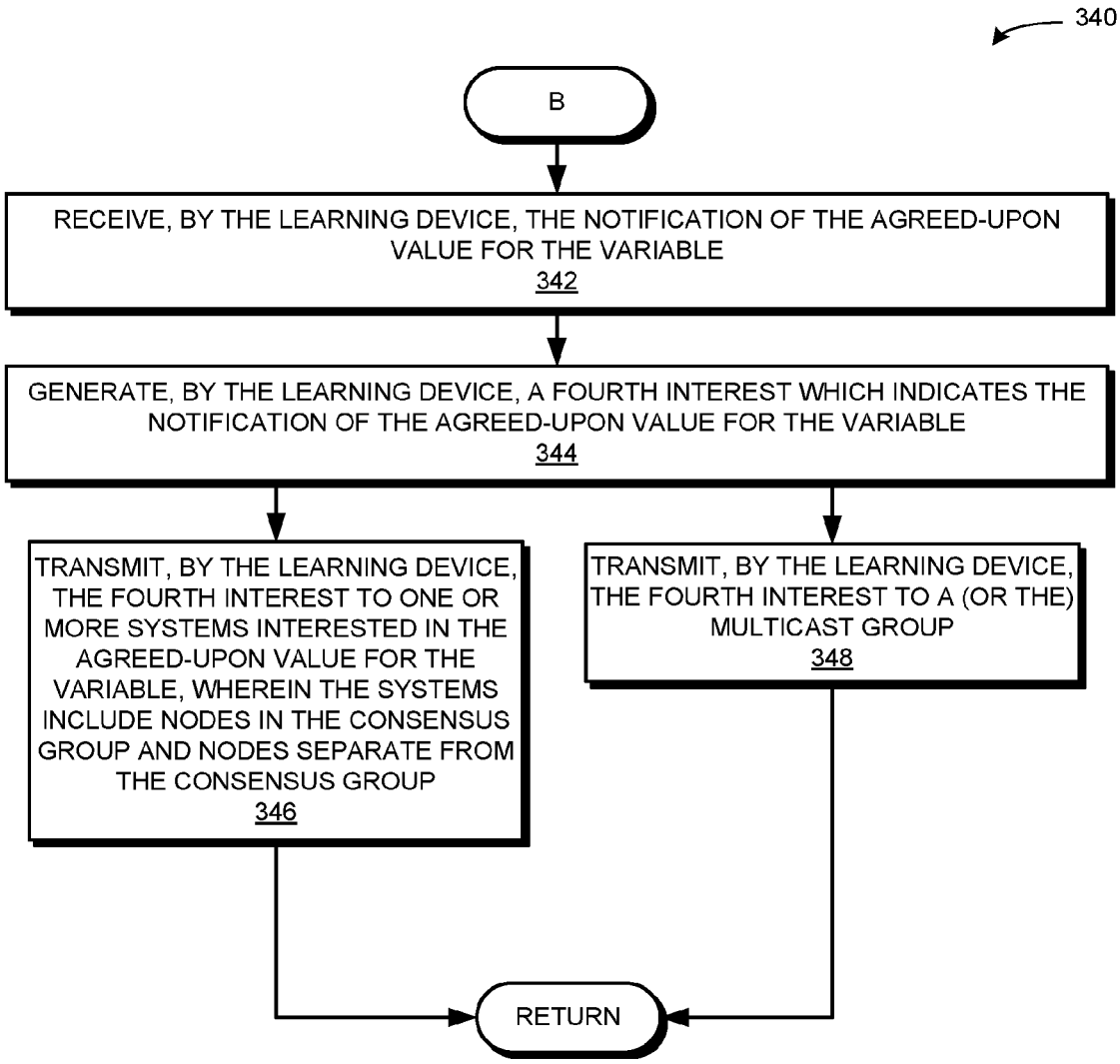

340
B
RECEIVE, BY THE LEARNING DEVICE, THE NOTIFICATION OF THE AGREED-UPON VALUE FOR THE VARIABLE
342
GENERATE, BY THE LEARNING DEVICE, A FOURTH INTEREST WHICH INDICATES THE NOTIFICATION OF THE AGREED-UPON VALUE FOR THE VARIABLE
344
TRANSMIT, BY THE LEARNING DEVICE, THE FOURTH INTEREST TO ONE OR MORE SYSTEMS INTERESTED IN THE AGREED-UPON VALUE FOR THE VARIABLE, WHEREIN THE SYSTEMS INCLUDE NODES IN THE CONSENSUS GROUP AND NODES SEPARATE FROM THE CONSENSUS GROUP
346
TRANSMIT, BY THE LEARNING DEVICE, THE FOURTH INTEREST TO A (OR THE) MULTICAST GROUP
348
RETURN

FIG. 3C

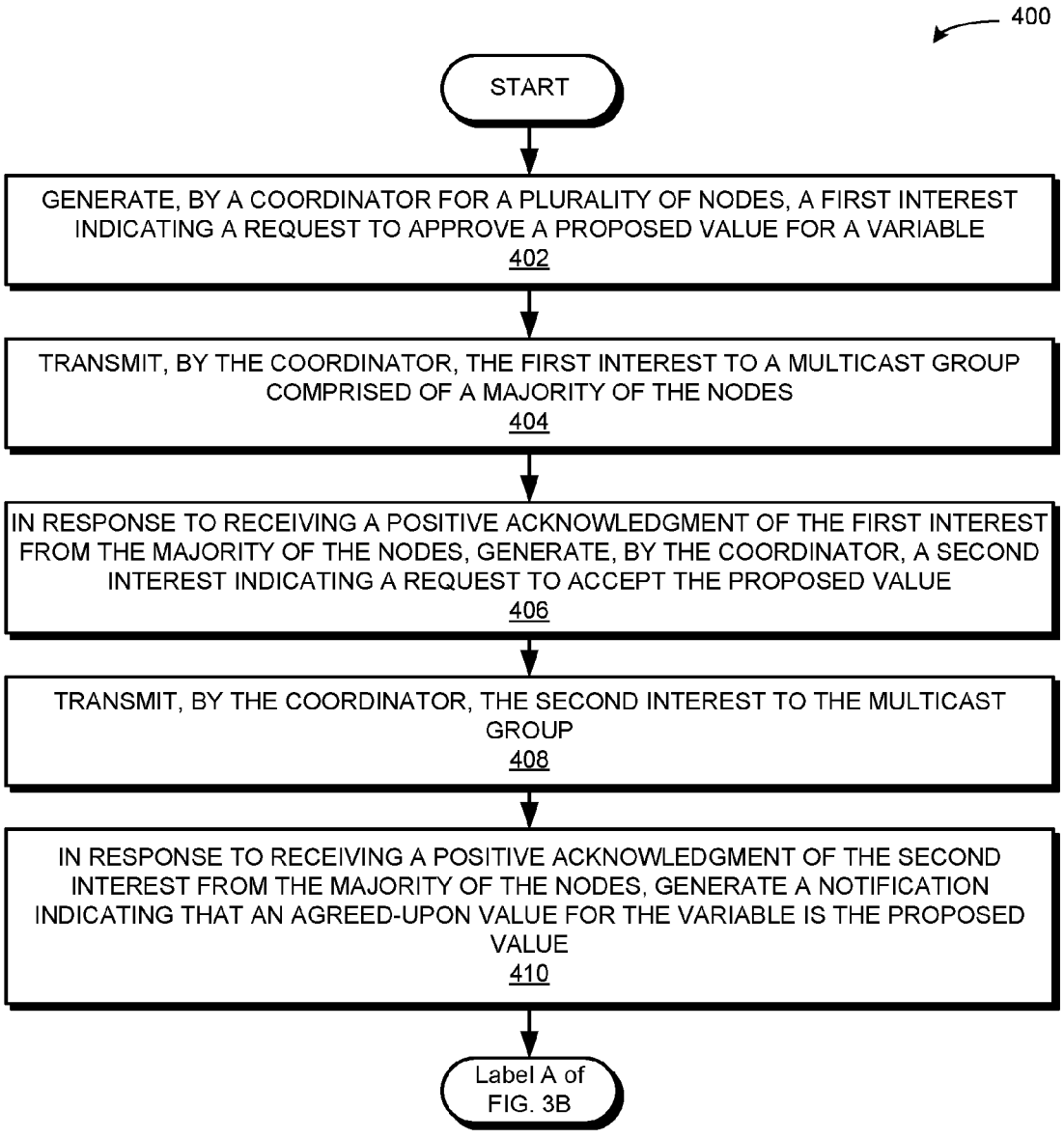

400
START
GENERATE, BY A COORDINATOR FOR A PLURALITY OF NODES, A FIRST INTEREST INDICATING A REQUEST TO APPROVE A PROPOSED VALUE FOR A VARIABLE
402
TRANSMIT, BY THE COORDINATOR, THE FIRST INTEREST TO A MULTICAST GROUP COMPRISED OF A MAJORITY OF THE NODES
404
IN RESPONSE TO RECEIVING A POSITIVE ACKNOWLEDGMENT OF THE FIRST INTEREST FROM THE MAJORITY OF THE NODES, GENERATE, BY THE COORDINATOR, A SECOND INTEREST INDICATING A REQUEST TO ACCEPT THE PROPOSED VALUE
406
TRANSMIT, BY THE COORDINATOR, THE SECOND INTEREST TO THE MULTICAST GROUP
408
IN RESPONSE TO RECEIVING A POSITIVE ACKNOWLEDGMENT OF THE SECOND INTEREST FROM THE MAJORITY OF THE NODES, GENERATE A NOTIFICATION INDICATING THAT AN AGREED-UPON VALUE FOR THE VARIABLE IS THE PROPOSED VALUE
410
Label A of FIG. 3B

FIG. 4

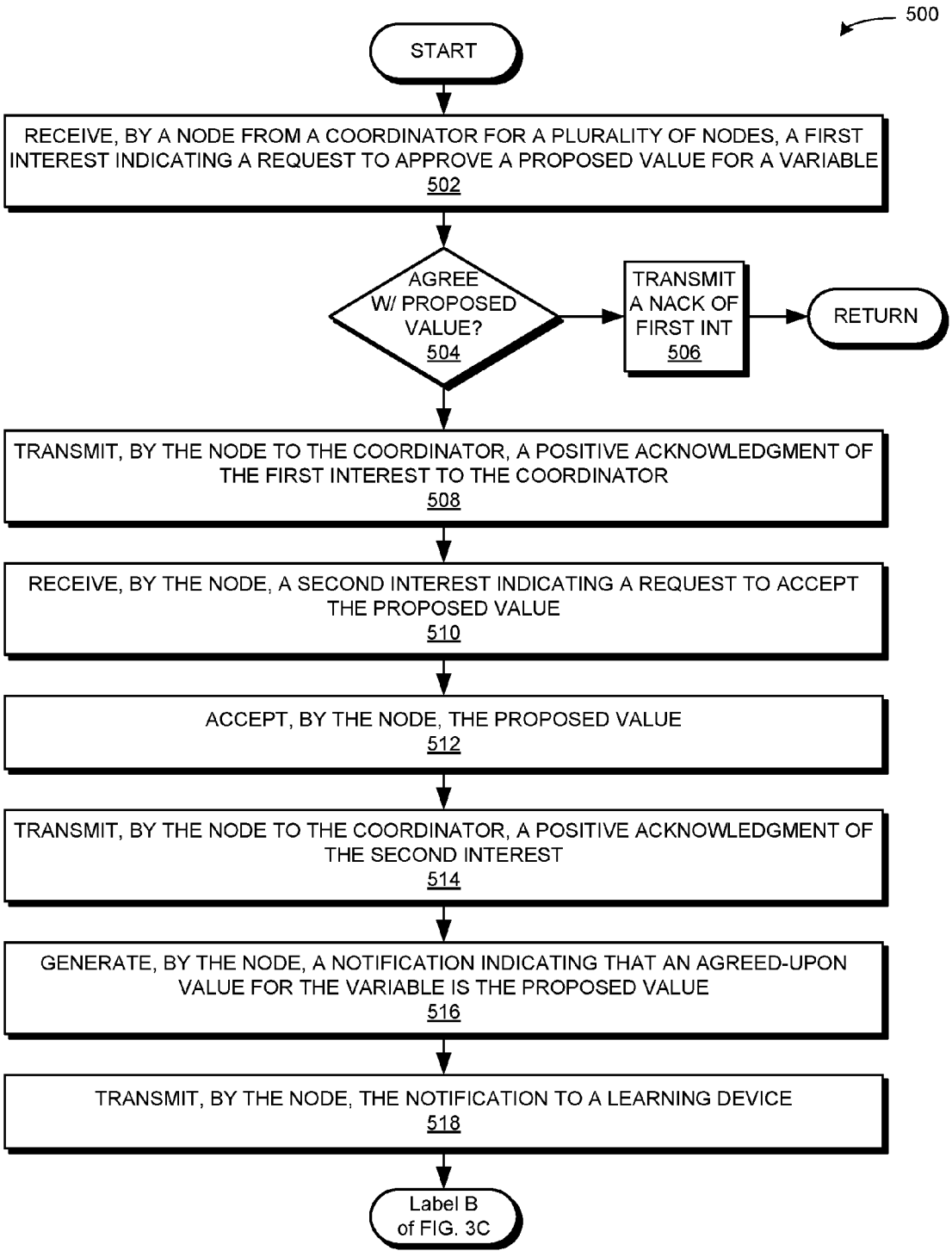

500
START
RECEIVE, BY A NODE FROM A COORDINATOR FOR A PLURALITY OF NODES, A FIRST INTEREST INDICATING A REQUEST TO APPROVE A PROPOSED VALUE FOR A VARIABLE
502
AGREE W/ PROPOSED VALUE?
504
TRANSMIT A NACK OF FIRST INT
506
RETURN
TRANSMIT, BY THE NODE TO THE COORDINATOR, A POSITIVE ACKNOWLEDGMENT OF THE FIRST INTEREST TO THE COORDINATOR
508
RECEIVE, BY THE NODE, A SECOND INTEREST INDICATING A REQUEST TO ACCEPT THE PROPOSED VALUE
510
ACCEPT, BY THE NODE, THE PROPOSED VALUE
512
TRANSMIT, BY THE NODE TO THE COORDINATOR, A POSITIVE ACKNOWLEDGMENT OF THE SECOND INTEREST
514
GENERATE, BY THE NODE, A NOTIFICATION INDICATING THAT AN AGREED-UPON VALUE FOR THE VARIABLE IS THE PROPOSED VALUE
516
TRANSMIT, BY THE NODE, THE NOTIFICATION TO A LEARNING DEVICE
518
Label B of FIG. 3C

FIG. 5

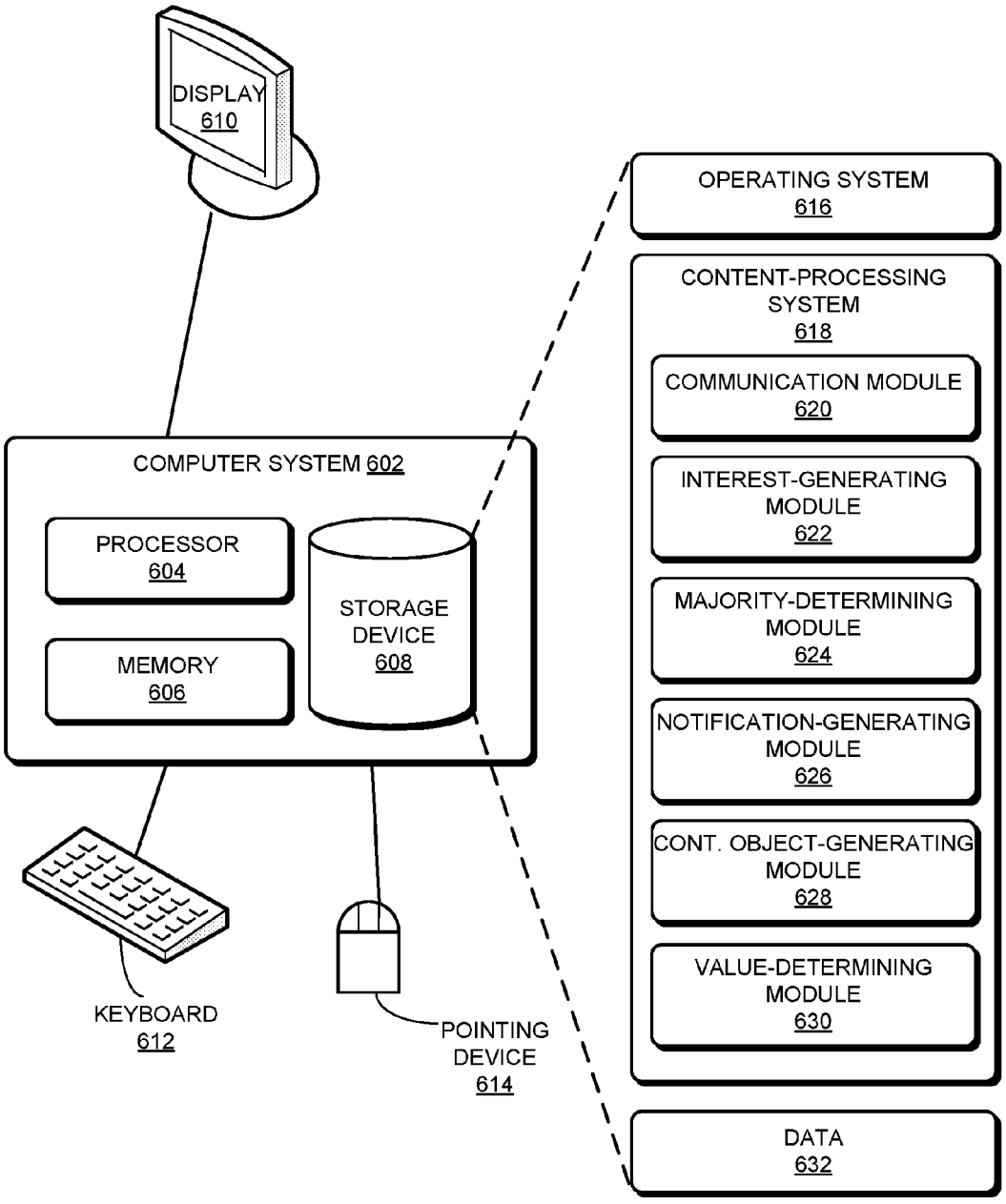

DISPLAY
610
COMPUTER SYSTEM 602
PROCESSOR
604
MEMORY
606
STORAGE
DEVICE
608
KEYBOARD
612
POINTING
DEVICE
614
OPERATING SYSTEM
616
CONTENT-PROCESSING
SYSTEM
618
COMMUNICATION MODULE
620
INTEREST-GENERATING
MODULE
622
MAJORITY-DETERMINING
MODULE
624
NOTIFICATION-GENERATING
MODULE
626
CONT. OBJECT-GENERATING
MODULE
628
VALUE-DETERMINING
MODULE
630
DATA
632

FIG. 6

DISTRIBUTED CONSENSUS IN A CONTENT CENTRIC NETWORK

RELATED APPLICATION

[0001] The subject matter of this application is related to the subject matter in the following applications:

[0002] U.S. patent application Ser. No. 13/847,814 (Attorney Docket No. PARC-20120537-US-NP), entitled "ORDERED-ELEMENT NAMING FOR NAME-BASED PACKET FORWARDING," by inventor Ignacio Solis, filed 20 Mar. 2013 (hereinafter "U.S. patent application Ser. No. 13/847,814");

[0003] U.S. patent application Ser. No. 12/338,175 (Attorney Docket No. PARC-20080626-US-NP), entitled "CONTROLLING THE SPREAD OF INTERESTS AND CONTENT IN A CONTENT CENTRIC NETWORK," by inventors Van L. Jacobson and Diana K. Smetters, filed 18 Dec. 2008 (hereinafter "U.S. patent application Ser. No. 12/338,175"); and

[0004] U.S. patent application Ser. No. 14/231,515 (Attorney Docket No. PARC-20140190US01), entitled "AGGREGATE SIGNING OF DATA IN CONTENT CENTRIC NETWORKING," by inventors Ersin Uzun, Marc E. Mosko, Michael F. Plass, and Glenn C. Scott, filed 31 Mar. 2014 (hereinafter "U.S. patent application Ser. No. 14/231,515");

the disclosures of which are herein incorporated by reference in their entirety.

BACKGROUND

Field

[0005] This disclosure is generally related to distribution of digital content. More specifically, this disclosure is related to a system for facilitating distributed consensus in a content centric network based on a Paxos algorithm.

Related Art

[0006] The proliferation of the Internet and e-commerce continues to create a vast amount of digital content. Content centric network (CCN) architectures have been designed to facilitate accessing and processing such digital content. A CCN includes entities, or nodes, such as network clients, forwarders (e.g., routers), and content producers, which communicate with each other by sending interest packets for various content items and receiving content-object packets in return. CCN interests and content objects are identified by their unique names, which are typically hierarchically structured variable length identifiers (HSVLI). An HSVLI can include contiguous name components ordered from a most general level to a most specific level.

[0007] Distributed consensus is crucial in today's network to provide fast, reliable, and lively services. Paxos is a family of protocols for solving consensus in a network of unreliable processors. In distributed consensus, the process of agreeing on one result among a group of participants presents challenges when the participants or their communication medium may experience failures. Paxos protocols typically involve three types of entities: proposers; acceptors; and learners. In Basic Paxos, a proposer sends a "prepare" request with a counter "N" and a value "V" to at least a majority of acceptors. Upon receiving the prepare request, an acceptor can respond with an ACK that N is the current maximum, and can also include any previously accepted value for V. When the proposer has received ACKs from the majority of the acceptors, the proposer sends an "accept" request for (N, V) (i.e., the consensus value) to the acceptors. Upon receiving the accept request, an acceptor can both respond with an ACK and notify the learner of the consensus value. The learner can inform other interested systems (including the involved proposer and acceptors) of the consensus value.

[0008] A CCN is a distributed system where consensus among nodes is an important feature (e.g., agreeing on a single value that is the outcome of an election or an environmental observation). Consensus is necessary, for example, if multiple writers wish to agree on the current version number of a CCNx name or if multiple distributed systems wish to elect a leader for fast transaction processing. Though CCN brings many desirable features to a network, some issues remain unsolved for achieving distributed consensus.

SUMMARY

[0009] One embodiment provides a system that facilitates distributed consensus in a content centric network. During operation, the system generates, by a coordinator for a plurality of nodes, a first interest that indicates a request to approve a proposed value for a variable. In response to receiving a positive acknowledgment of the first interest from a majority of the nodes, the coordinator generates a second interest that indicates a request to accept the proposed value. A name for the first interest and a name for the second interest include an identifier of the variable and a round number. A payload of the first interest and a payload of the second interest include the proposed value. In response to receiving a positive acknowledgement of the second interest from the majority of the nodes, the system generates a notification indicating that an agreed-upon value for the variable is the proposed value.

[0010] In some embodiments, in response to receiving a third interest indicating a request to read the agreed-upon value for the variable, wherein a name for the third interest includes the variable identifier and the round number, the system generates, by the coordinator, a content object which indicates the round number and one or more of: the agreed-upon value for the variable; an iteration number corresponding to the round number; a negative acknowledgment if the agreed-upon value for the variable is indeterminate; and a no-operation indicator if no agreed-upon value for the variable exists.

[0011] In some embodiments, the positive acknowledgement of the second interest is transmitted to the coordinator by the majority of the nodes. The notification is generated by the majority of the nodes and further transmitted to a learning device, which transmits the notification as a fourth interest to one or more systems interested in the agreed-upon value for the variable.

[0012] In some embodiments, the name for the first interest, the name for the second interest, the name for the third interest, and the name for the fourth interest further include one or more of: a routable prefix for one of the majority of the nodes; an identifier for a consensus group to which the one of the majority of the nodes belongs, wherein the plurality of nodes belong to the consensus group; an indi-

cator of a logical program associated with the variable identifier; and an iteration number corresponding to the round number.

[0013] In some embodiments, the name for the first interest indicates the request to approve the proposed value for the variable, the name for the second interest indicates the request to accept the proposed value, the name for the third interest indicates the request to read the agreed-upon value, and the name for the fourth interest indicates the notification to allow a receiving device to learn the agreed-upon value.

[0014] In some embodiments, the system transmits, by the coordinator, the first interest to a multicast group comprised of the majority of the nodes. The coordinator transmits the second interest to the multicast group. The name for the first interest and the name for the second interest further include one or more of: an identifier for a consensus group to which the one of the majority of the nodes belongs, wherein the plurality of nodes belong to the consensus group, wherein the consensus group identifier is the most general level name component; and an indicator of a group version to which the majority of the nodes belongs. The payload of the first interest further includes a routable prefix of the coordinator to be used by a node in response to the first interest, and the payload of the second interest further includes a routable prefix of a target to be used by a node in response to the second interest.

[0015] In some embodiments, the proposed value is one or more of: a link to a piece of content which describes a current state of an algorithm; a link to a manifest, which is a content object indicating a collection of other content objects; and the manifest embedded in the proposed value.

[0016] In some embodiments, a response by one of the majority of the nodes to an interest is a content object with a same name as the name for the interest, and the content object has a lifetime set to a small or a zero value.

[0017] One embodiment provides a system that facilitates distributed consensus in a content centric network. During operation, the system receives, from a coordinator for a plurality of nodes by a node, a first interest indicating a request to approve a proposed value for a variable. In response to agreeing with the proposed value, the node transmits a positive acknowledgement of the first interest to the coordinator. The node receives a second interest indicating a request to accept the proposed value. A name for the first interest and a name for the second interest include an identifier of the variable and a round number, and a payload of the first interest and a payload of the second interest include the proposed value. In response to accepting the proposed value, the node transmits a positive acknowledgement of the second interest to the coordinator.

[0018] In some embodiments, in response to accepting the proposed value, the node transmits a notification indicating that an agreed-upon value for the variable is the proposed value to a learning device, which transmits the notification as a third interest to one or more systems interested in the agreed-upon value for the variable.

[0019] In some embodiments, in response to not agreeing with the proposed value, the node transmits a negative acknowledgment of the first interest to the coordinator. The negative acknowledgment includes a previous value for the variable corresponding to a previous round number.

BRIEF DESCRIPTION OF THE FIGURES

[0020] FIG. **1** illustrates an exemplary network facilitating distributed consensus in a content centric network, in accordance with an embodiment of the present invention.

[0021] FIG. **2** presents exemplary communication between a proposer, acceptors, a learner, and targets, in accordance with an embodiment of the present invention.

[0022] FIG. **3**A presents a flow chart illustrating a method by a coordinator for facilitating distributed consensus in a content centric network, in accordance with an embodiment of the present invention.

[0023] FIG. **3**B presents a flow chart illustrating a method by acceptor nodes for facilitating distributed consensus in a content centric network, in accordance with an embodiment of the present invention.

[0024] FIG. **3**C presents a flow chart illustrating a method by a learning device for facilitating distributed consensus in a content centric network, in accordance with an embodiment of the present invention.

[0025] FIG. **4** presents a flow chart illustrating a method by a coordinator for facilitating distributed consensus in a content centric network, based on a multicast group, in accordance with an embodiment of the present invention.

[0026] FIG. **5** presents a flow chart illustrating a method by an acceptor node for facilitating distributed consensus in a content centric network, in accordance with an embodiment of the present invention.

[0027] FIG. **6** illustrates an exemplary computer system that facilitates distributed consensus in a content centric network, in accordance with an embodiment of the present invention.

[0028] In the figures, like reference numerals refer to the same figure elements.

DETAILED DESCRIPTION

[0029] The following description is presented to enable any person skilled in the art to make and use the embodiments, and is provided in the context of a particular application and its requirements. Various modifications to the disclosed embodiments will be readily apparent to those skilled in the art, and the general principles defined herein may be applied to other embodiments and applications without departing from the spirit and scope of the present disclosure. Thus, the present invention is not limited to the embodiments shown, but is to be accorded the widest scope consistent with the principles and features disclosed herein.

Overview

[0030] Embodiments of the present invention provide a system which facilitates distributed consensus in a CCN based on the Paxos algorithm. Distributed consensus is crucial in today's network to provide fast, reliable, and lively services. Paxos is a family of protocols for solving consensus in a network of unreliable processors. In distributed consensus, the process of agreeing on one result among a group of participants presents challenges when the participants or their communication medium may experience failures. Paxos protocols typically involve three types of entities: proposers; acceptors; and learners. In Basic Paxos, a proposer sends a "prepare" request with a counter N and a value V to at least a majority of acceptors. Upon receiving the prepare request, an acceptor can respond with an acknowledgment (ACK) that N is the current maximum, and

can also include any previously accepted value for V. When the proposer has received ACKs from the majority of the acceptors, the proposer sends an "accept" request for (N, V) (i.e., the consensus value) to the acceptors. Upon receiving the accept request, an acceptor can both respond with an ACK, and notify the learner of the consensus value. The learner can then inform other interested systems (including the involved proposer and acceptors) of the consensus value.

[0031] In Multi-Paxos, which is based on a series of iterations of Basic Paxos, a consensus value V can evolve over time, as $\{V_0, \ldots, V_i\}$. A single master proposer may be selected using Basic Paxos, and after the master proposer has succeeded in Phase 1 (e.g., round N_0 corresponding to V_0), the master proposer can submit as many values as it wishes in subsequent phases, by submitting pairs $\{i, V_i\}$. The distinction between the three types of entities in Paxos (proposer, acceptor, and learner) is not exclusive. That is, each entity can be a potential proposer, acceptor, or learner. Each entity can contend for the proposer role and all entities can act as an acceptor or a learner.

[0032] A CCN is a distributed system where consensus among nodes is an important feature (e.g., agreeing on a single value that is the outcome of an election or an environmental observation). Consensus is necessary, for example, if multiple writers wish to agree on the current version number of a CCN name or if multiple distributed systems wish to elect a leader for fast transaction processing. Embodiments of the present invention provide a system that facilitates distributed consensus in a CCN based on the Paxos algorithm. One form is based on the standard CCN exchange of an interest (i.e., request) and a content object (i.e., response), as described below in relation to FIGS. **1**, **2**, **3**A-**3**C. Another form is based on an exchange of a Push request and a response, as described below in relation to FIG. **4**.

[0033] Thus, the present system provides improvements to the distribution of digital content, where the improvements are fundamentally technological. Embodiments of the present invention provide a technological solution (e.g., determining distributed consensus on values for system-related variables based on the Paxos algorithm using interest/content object exchanges and/or push request/response exchanges) to the technological problem of efficiently achieving consensus in a distributed system such as a CCN.

[0034] The following terms describe elements of a CCN architecture:

[0035] Content Object or "content object": A single piece of named data, which is bound to a unique name. Content Objects are "persistent," which means that a Content Object can move around within a computing device, or across different computing devices, but does not change. If any component of the Content Object changes, the entity that made the change creates a new Content Object that includes the updated content, and binds the new Content Object to a new unique name.

[0036] Unique Names: A name in a CCN is typically location independent and uniquely identifies a Content Object. A data-forwarding device can use the name or name prefix to forward a packet toward a network node that generates or stores the Content Object, regardless of a network address or physical location for the Content Object. In some embodiments, the name may be a hierarchically structured variable-length identifier (HSVLI). The HSVLI can be divided into several hierarchical components, which can be structured in various ways. For example, the individual name components parc, home, ccn, and test.txt can be structured in a left-oriented prefix-major fashion to form the name "/parc/home/ccn/test.txt." Thus, the name "/parc/home/ccn" can be a "parent" or "prefix" of "/parc/home/ccn/test.txt." Additional components can be used to distinguish between different versions of the content item, such as a collaborative document. The HSVLI can comprise contiguous name components ordered from a most general level to a most specific level.

[0037] In some embodiments, the name can include a non-hierarchical identifier, such as a hash value that is derived from the Content Object's data (e.g., a checksum value) and/or from elements of the Content Object's name. A description of a hash-based name is described in U.S. patent application Ser. No. 13/847,814, which is hereby incorporated by reference. A name can also be a flat label. Hereinafter, "name" is used to refer to any name for a piece of data in a name-data network, such as a hierarchical name or name prefix, a flat name, a fixed-length name, an arbitrary-length name, or a label (e.g., a Multiprotocol Label Switching (MPLS) label).

[0038] Interest or "interest": A packet that indicates a request for a piece of data, and includes a name (or a name prefix) for the piece of data. A data consumer can disseminate a request or Interest across an information-centric network, which CCN routers can propagate toward a storage device (e.g., a cache server) or a data producer that can provide the requested data to satisfy the request or Interest.

[0039] The methods disclosed herein are not limited to CCN networks and are applicable to other architectures as well. A description of a CCN architecture is described in U.S. patent application Ser. No. 12/338,175, which is hereby incorporated by reference. In addition, CCNx 1.0 is a specific protocol suite and implementation of CCN.

Distributed Consensus in a CCN: Interest and Content Object Exchanges

[0040] In a system which facilitates distributed consensus in a CCN based on the Paxos algorithm, three types of entities can exist: a proposer; an acceptor; and a learner. A plurality or group of nodes can participate in a consensus decision on a value for a variable, and a majority of the participating nodes must agree on the value for the variable in order for the consensus value to be accepted. Any node in the group may be chosen to act as the proposer or coordinator in a round related to the consensus decision. Assume that a current proposer (or "master" or "master proposer") of a consensus group is elected using distributed consensus where each contending proposer bids to have its value accepted. The accepted consensus value determines the master proposer. The actual value can be the name of the CCNx content object that describes the proposer. The set of acceptors can be maintained as a consensus value. A new system may enter as an acceptor or be removed, if non-responsive, by the proposer, which can perform such an action based on a protected variable. This allows the proposer to know what constitutes a majority. The identity of the learner is also maintained as a protected value. The acceptors have knowledge of the identity of the current learner, and can inform the learner of the acceptors' accept choices. The learner can use the identities of the acceptor group associated with the given consensus value and notify

all acceptors and proposers. The learner can also notify other systems interested in the given consensus value.

[0041] As described above, one form of achieving distributed consensus is based on the standard CCN exchange of an interest (i.e., request) and a content object (i.e., response). Below are four exemplary formats for names of interest (and corresponding responsive content objects):

/proposer/grp/prg/var/read/[N[/iter]] Format (1)

/acceptor/grp/prg/var/prepare/N[/iter] Format (2)

/acceptor/grp/prg/var/accept/N[/iter] Format (3)

/target/grp/prg/var/learn/N[/iter] Format (4)

[0042] The name components "/proposer," "/acceptor," and "/target" identify, respectively, a proposer, an acceptor, and a target (such as a learner, a proposer, an acceptor, or other system). The identifier can be a routable prefix for the respective entity. The substring of name components "/grp/prg/var" identifies, respectively, the consensus group grp in which the acceptor participates, the logical program prg, and the protected variable var. The name components "/read," "/prepare," "/accept," and "/learn" identify the type of request. The suffix "/N/[iter]" identifies, respectively, the ordering N (which can identify the round number) and the optional iteration iter. When using CCNx 1.0 labeled names, the suffix can take the form of, e.g., "App:prepare=N" and "App:iter=iter."

[0043] The payload of the request carries the state of the request. Specifically, the payload carries the value V. The value can be a CCNx 1.0 link to a piece of content which describes a current state of an algorithm. The value can also be a link to a manifest, which is a content object indicating a collection of other content objects, or can be the actual manifest embedded in the value. Manifests, or secure content catalogs, are described in U.S. patent application Ser. No. 14/231,515.

[0044] Any system may read the current consensus value by sending an interest of Format (1) ("read request"). In response, the proposer returns the current consensus value, which is the tuple (N, iter, V_{iter}). A requesting system may also specify a specific N or (N, iter) pair. If the proposer does not know the consensus value for a read request, the proposer can respond with a NACK if the consensus value is indeterminate, or the proposer can respond with a no-operation indicator if no consensus value exists to satisfy the read request.

[0045] A proposer can send an interest of Format (2) ("prepare request") to a majority of acceptors. Upon receiving the prepare request, an acceptor can return an ACK content object response to the proposer. The acceptor can also return nothing or a NACK content object response to the proposer, and the NACK content object response can include the acceptor's most recent consensus value and round for the variable.

[0046] When the proposer has collected sufficient ACKS (e.g., received an ACK from a majority of the acceptors), the proposer can send an interest of Format (3) ("accept request") to a majority of acceptors. Upon receiving the accept request, an acceptor can return an ACK content object response to the proposer. In addition, the acceptor can send a push request (or notification or interest of Format (4)) to the learner, which indicates the ACK of the accept request for the given variable for the given round and/or iteration. When the learner has collected sufficient ACKs (e.g., received an ACK from a majority of the acceptors), the learner can send an interest of Format (4) ("learn request") to one or more targets. The targets can include the proposer and the acceptors, as well as other systems or devices not involved in the consensus group.

[0047] Note that an acceptor's content object response (to the prepare request and to the accept request) follows the reverse path of the request back to the proposer. The content object response carries in its payload the consensus value or state for the current round or iteration. To avoid excessive caching by a node, the content object response can have a lifetime (e.g., MaxAge) set to a small or a zero value.

Distributed Consensus in a CCN: Interest Multicast

[0048] Another form of achieving distributed consensus in a CCN is based on multicasting an interest. A proposer can transmit a single Push request message to an interest multicast group, and all listening acceptors can receive the single Push request message. Because the proposer requires knowledge of when it has received a majority of responses (e.g., collected a sufficient number of ACKs), the group of acceptors listening to the group name is identified by a specific group version with a known number of acceptors. The payload of the Push message sent by the proposer can include an identifier for the proposer, which allows each acceptor to send an individual Push response message directly back to the proposer. Thus, the Push responses do not necessarily follow the reverse path of the corresponding Push request.

[0049] This multicast method uses similar signaling as the previously described method based on interest and content object exchanges. One difference in the multicast method is that the routable prefix is now the group name "grp" rather than the individual acceptor:

/grp/grpver/prg/var/prepare/N[/iter] Format (5)

/grp/grpver/prg/var/accept/N[/iter] Format (6)

/grp/grpver/prg/var/learn/N[/iter] Format (7)

[0050] The name component "grpver" identifies the version corresponding to the group grp. Another difference in the multicast method is that in some embodiments, the payload of a request carries the target name to use in the response. However, including the target name in the payload of an individual Push request message is not strictly necessary because all systems have knowledge of the current system state as well as the identity of the proposer and the learner from the consensus state and the group version grpver. This is because the proposer is the response target for a prepare or accept request and the learner is the ACK destination for a learn request.

[0051] In the multicast method, a proposer can send a first Push interest message (e.g., a prepare request or message which is an interest of Format (5)) to the interest multicast group grp, which allows all listening acceptors to receive the interest and respond directly to the proposer with a first Push response, based on a name or identifier for the proposer as carried in the payload of the Push interest message. When the proposer has collected sufficient positive responses, the proposer can send a second Push interest message (e.g., an accept request or message which is an interest of Format (6)) to the interest multicast group grp, and all listening acceptors

can receive the second Push interest message and respond directly to the proposer or a learner with a second Push response, again based on the identifier as carried in the payload of the second Push interest message.

[0052] In addition, in the "learning" stage, the learner may use a mixture of interest and content object exchanges and of multicast Push learn messages of Format (7). A node can respond to a Push learn message with a Push acknowledgment message. In the above example, the identifier carried in the payload of the second Push interest message can be an identifier for a learner that is an entity distinct from the proposer. As such, the listening acceptors that receive the second Push interest message will send a response directly to the learner. Upon collecting sufficient second Push responses, the learner can send a third Push interest message (e.g., a learn message which is an interest of Format (7)) to the interest multicast group, and all listening acceptors can respond directly to the learner with a third Push response, which is an ACK of the learn message. The learner can also send a similar third Push interest message to any other interested systems, and can receive a responsive Push ACK from the other systems.

[0053] In some embodiments, the proposer can act as the learner, and the identifier carried in the payload of the second Push interest message can be the proposer identifier. Upon collecting sufficient second Push responses, the proposer can send the third Push interest message (e.g., a learn message which is an interest of Format (7)), and all listening acceptors can again respond directly to the proposer with a third Push response, which is an ACK of the learn message.

Exemplary Network and Communication

[0054] FIG. **1** illustrates an exemplary network **100** facilitating distributed consensus in a content centric network, in accordance with an embodiment of the present invention. Network **100** can include a consumer or content-consuming device **116**, a producer or content-producing device **118**, and a router or other forwarding device at nodes **102**, **104**, **106**, **108**, **110**, **112**, and **114**. A node can be a computer system, an end-point representing users, and/or a device that can generate interests or originate content. A node can also be an edge router (e.g., CCN nodes **102**, **104**, **112**, and **114**) or a core router (e.g., intermediate CCN routers **106-110**).

[0055] Node **108** can be a proposer ("P108"); nodes **104**, **106**, **110**, **112**, and **114** can be acceptors comprising a plurality of nodes involved in a consensus group ("A104," "A106," "A110," "A112", and "A114"); and device **118** can be a learner ("L118"). That is, these five nodes (A104, A106, A110, A112, and A114) are acceptors in a distributed consensus protocol. The majority of these five nodes is thus three nodes (where the majority is more than half of five).

[0056] During operation, P108 sends a first interest (i.e., a prepare request) to a majority of the acceptors, e.g., to three acceptors such as A110, A112, and A114. For example, P108 can send to A112 an interest **211** with a name **211.1** of "/A112/grp/prg/var/prepare/N" and a payload **211.2** with a value of "V." Name **211.1** can also include an iteration number (not shown). Upon receiving interest **211**, A112 can determine to agree with the prepare request by sending a responsive ACK content object **212** with a name **212.1** of "/A112/grp/prg/var/prepare/N" and a payload **212.2** with a value of "V; ACK." The acknowledgment can be included or indicated in content object **212** in other ways.

[0057] When P108 has received ACKs from the majority (e.g., the three nodes A110, A112, and A114), P108 sends a second interest (i.e., an accept request) to the majority of the acceptors (e.g., A110, A112, and A114). For example, P108 can send to A112 an interest **213** with a name **213.1** of "/A112/grp/prg/var/accept/N" and a payload **213.2** with a value of "V." Upon receiving interest **213**, A112 can determine to agree with the accept request by sending a responsive ACK content object **214** with a name **214.1** of "/A112/grp/prg/var/accept/N" and a payload **214.2** with a value of "V; ACK." A112 can also send an interest or push notification **215** to learner **118** ("L118").

[0058] L118 can respond to A112 with a responsive content object ACK or a push response (not shown). When L118 has received push notifications from the majority (e.g., collected sufficient ACKs), L118 can also generate and transmit another interest (i.e., a learn request) to one or more targets, which can include P108, any of acceptors A102, A104, A106, A110, A112, and A114 (including acceptors both involved and not involved in the consensus decision as well as other nodes that are not acceptors), and one or more other targets. For example, L118 can send to system **120** an interest **216** with a name **216.1** of "/Target/grp/prg/var/learn/N" and a payload **216.2** with a value of "V," where "/Target" identifies system **120**. Upon receiving interest **216**, system **120** can send a responsive ACK content object (not shown). As another example, L118 can send to A112 an interest **216** with a name **216.1** of "/Target/grp/prg/var/learn/N" and a payload **216.2** with a value of "V," where "/Target" identifies A112. Upon receiving interest **216**, A112 can send a responsive ACK content object (not shown).

[0059] Thus, the system facilitates distributed consensus in a CCN based on the Paxos algorithm by mapping Paxos communications to interest and content object exchanges. The other embodiment (multicasting interests), is described below in relation to FIG. **4**.

Detailed Description of Exemplary Communication

[0060] FIG. **2** presents exemplary communication **200** between a proposer **108** (P108), acceptors **110**, **112**, and **114** (A110, **112**, and A114), a learner **118** (L118), and targets **120** (T120), in accordance with an embodiment of the present invention. Communication **200** of FIG. **2** corresponds to the communication described above in relation to FIG. **1**. Note that the values of N and V in messages **211**, **212**, **213**, **214**, **215**, **216**, and **217** are illustrated as "N=1" and "V=64" for exemplary purposes only.

[0061] During operation, P108 sends a prepare request to a majority of the acceptors (e.g., the three nodes A110, A112, and A114). For example, P108 can send interest **211** to A112. Upon receiving interest **211**, A112 can determine to agree with the prepare request by sending a responsive ACK content object **212**. When P108 has received ACKs from the majority (collect sufficient ACKs function **230**), P108 sends an accept request to the majority of the acceptors (e.g., A110, A112, and A114). For example, P108 can send interest **213** to A112. Upon receiving interest **213**, A112 can determine to agree with the accept request by sending a responsive ACK content object **214**. A112 (and each of majority of acceptors) can also send an interest or push notification **215** to L118.

[0062] L118 can respond to A112 with a responsive content object ACK or a push response (not shown). When L118 has received push notifications from the majority (collect

sufficient ACKs **232**), L118 can generate and transmit another interest (i.e., a learn request) to one or more targets. For example, L118 can send interest **216** to system **120**. Upon receiving interest **216**, system **120** can send a responsive ACK content object **217**. Recall that L118 can also send a learn request to any proposer, acceptor, or other system (not shown).

[0063] Alternatively, instead of the acceptors notifying L118 directly and L118 performing function **232**, P108 can collect sufficient ACKs from acceptors of the accept request (not shown) and P108 can subsequently send a learn request to L118 (not shown), which can then send interest **216** as depicted above.

Method for Facilitating Distributed Consensus in a CCN: Interest and Content Object Exchanges (Proposer; Acceptors; Learner)

[0064] FIG. **3**A presents a flow chart **300** illustrating a method by a coordinator for facilitating distributed consensus in a content centric network, in accordance with an embodiment of the present invention. During operation, the system generates, by a coordinator for a plurality of nodes, a first interest indicating a request to approve a proposed value for a variable (operation **302**). The coordinator can be a proposer. In response to receiving a positive acknowledgment of the first interest from a majority of the nodes, the coordinator generates a second interest indicating a request to accept the proposed value (operation **304**). A name for the first interest and a name for the second interest include an identifier of the variable and a round number, and a payload of the first interest and a payload of the second interest include the proposed value. In response to receiving a positive acknowledgment of the second interest from the majority of the nodes, the system generates a notification indicating that an agreed-upon value for the variable is the proposed value (operation **306**). In response to receiving a third interest indicating a request to read the agreed-upon value for the variable, the coordinator generates a content object which indicates the round number and one or more of: the agreed-upon value for the variable; an iteration number corresponding to the round number; a negative ACK if the agreed-upon value is indeterminate; and a no-operation indicator if no agreed-upon value exists (operation **308**). Note that operation **308** (i.e., a read request) can occur at any time, e.g., independent of and separate from operations **302**-**306**. The operation continues at Label A of FIG. **3**B.

[0065] FIG. **3**B presents a flow chart **320** illustrating a method by acceptor nodes for facilitating distributed consensus in a content centric network, in accordance with an embodiment of the present invention. During operation, in response to accepting the proposed value indicated in the second interest, the majority of nodes transmits to the coordinator the positive acknowledgment of the second interest (operation **322**). The majority of the nodes generates the notification indicating that an agreed-upon value for the variable is the proposed value (operation **324**). The majority of the nodes transmits the notification to a learning device (operation **326**). The operation continues at Label B of FIG. **3**C.

[0066] FIG. **3**C presents a flow chart **340** illustrating a method by a learning device for facilitating distributed consensus in a content centric network, in accordance with an embodiment of the present invention. During operation, a learning device receives the notification of the agreed-upon value for the variable (operation **342**). The notification can also include the round number and other information as described above in relation to Formats (1)-(4) and FIGS. **1** and **2**. The learning device generates a fourth interest which indicates the notification of the agreed-upon value for the variable (operation **344**). The learning device can transmit the fourth interest to one or more system interested in the agreed-upon value for the variable (operation **346**). The interested systems can include nodes in the consensus group such as the proposer and the acceptors (whether in the majority of the acceptors or not). The interested systems can also include nodes or systems that are separate from and not related to the consensus group. A name for the fourth interest can include a name, identifier, or routable prefix for an interested system. The learning device can also transmit the fourth interest to a multicast group (operation **348**) (or to "the" multicast group if Label B is reached from operation **410** of FIG. **4**).

Method for Facilitating Distributed Consensus in a CCN: Interest Multicast

[0067] FIG. **4** presents a flow chart **400** illustrating a method by a coordinator for facilitating distributed consensus in a content centric network, based on multicast group, in accordance with an embodiment of the present invention. During operation, the system generates, by a coordinator for a plurality of nodes, a first interest indicating a request to approve a proposed value for a variable (operation **402**). The coordinator transmits the first interest to a multicast group comprised of a majority of the nodes (operation **404**). In response to receiving a positive acknowledgment of the first interest from the majority of the nodes, the coordinator generates a second interest indicating a request to accept the proposed value (operation **406**). A name for the first interest and a name for the second interest include an identifier of the variable and a round number, and a payload of the first interest and a payload of the second interest include the proposed value. The coordinator transmits the second interest to the multicast group (operation **408**). In response to receiving a positive acknowledgment of the second interest from the majority of the nodes, the system generates a notification indicating that an agreed-upon value for the variable is the proposed value (operation **410**). The operation continues at Label A of FIG. **3**B.

Role of Acceptor in Facilitating Distributed Consensus in a CCN: Interest and Content Object Exchanges

[0068] FIG. **5** presents a flow chart **500** illustrating a method by an acceptor node for facilitating distributed consensus in a content centric network, in accordance with an embodiment of the present invention. During operation, the system receives, by a node from a coordinator for a plurality of nodes, a first interest indicating a request to approve a proposed value for a variable (operation **502**). If the node does not agree with the proposed value (decision **504**), the node transmits a negative acknowledgment of the first interest (operation **506**) and the operation returns. The node can also include a previous value for the variable corresponding to a previous round number or iteration, along with the previous round number or iteration. Alternatively, the node can transmit nothing back to the coordinator, which can result in a timeout and is handled by the coordinator like a failure or a NACK.

[0069] If the node agrees with the proposed value (decision **504**), the node transmits to the coordinator a positive acknowledgment of the first interest (operation **508**). The node receives a second interest indicating a request to accept the proposed value (operation **510**). The node will typically only receive the second interest in response to operation **508** (and not in response to operation **506**). A name for the first interest and a name for the second interest include an identifier of the variable and a round number, and a payload of the first interest and a payload of the second interest include the proposed value.

[0070] The node accepts the proposed value (operation **512**). The node can also determine to not accept the proposed value, and can send a negative acknowledgment to the coordinator (not shown). Upon accepting the proposed value, the node transmits to the coordinator a positive acknowledgment of the second interest (operation **514**). The node generates a notification indicating that an agreed-upon value for the variable is the proposed value (operation **516**), and the node transmits the notification to a learning device (operation **518**). The operation continues at Label B of FIG. **3**C.

Exemplary Computer System

[0071] FIG. **6** illustrates an exemplary computer system that facilitates distributed consensus in a content centric network, in accordance with an embodiment of the present invention. Computer system **602** includes a processor **604**, a memory **606**, and a storage device **608**. Memory **606** can include a volatile memory (e.g., RAM) that serves as a managed memory, and can be used to store one or more memory pools. Furthermore, computer system **602** can be coupled to a display device **610**, a keyboard **612**, and a pointing device **614**. Storage device **608** can store an operating system **616**, a content-processing system **618**, and data **630**.

[0072] Content-processing system **618** can include instructions, which when executed by computer system **602**, can cause computer system **602** to perform methods and/or processes described in this disclosure. Specifically, content-processing system **618** may include instructions for sending and/or receiving data packets to/from other network nodes across a computer network, such as a content centric network, where a data packet can correspond to an interest or a content object with a name and a payload, and can also correspond to a push interest, request, response, or notification message (communication module **620**). Content-processing system **618** may include instructions for generating a first interest indicating a request to approve a proposed value for a variable (interest-generating module **622**). Content-processing system **618** can include instructions for, in response to receiving a positive ACK of the first interest from a majority of the nodes (majority-determining module **624**), generating a second interest indicating a request to accept the proposed value (interest-generating module **622**). Content-processing system **618** can include instructions for, in response to receiving a positive ACK of the second interest from a majority of the nodes (majority-determining module **624**), generating a notification indicating that an agreed-upon value for the variable is the proposed value (notification-generating module **626**).

[0073] Content-processing system **618** can include instructions for, in response to receiving a third interest indicating a request to read the agreed-upon value for the variable (communication module **620**), generating a content object which indicates the round number and one or more of: the agreed-upon value for the variable; an iteration number; a NACK; and a no-operation indicator (content object-generating module **628**). Content-processing system **618** can include instructions for transmitting the first interest and the second interest to a multicast group (communication module **620**).

[0074] Content-processing system **618** can further include instructions for receiving a first interest indicating a request to approve a proposed value for a variable (communication module **620**). Content-processing system **618** can include instructions for, in response to agreeing with the proposed value (value-determining module **630**), transmitting a positive ACK of the first interest to the coordinator (communication module **620**; content object-generating module **628**). Content-processing system **618** can include instructions for receiving a second interest indicating a request to accept the proposed value (communication module **620**). Content-processing system **618** can include instructions for, in response to accepting the proposed value (value-determining module **630**), transmitting a positive ACK of the second interest to the coordinator (communication module **620**; content object-generating module **628**), and transmitting a notification indicating that an agreed-upon value for the variable is the proposed value to a learning device (notification-generating module **626**). Content-processing system **618** can include instructions for, in response to not agreeing with the proposed value (value-determining module **630**), transmitting a negative ACK of the first interest to the coordinator (communication module **620**).

[0075] Data **632** can include any data that is required as input or that is generated as output by the methods and/or processes described in this disclosure. Specifically, data **632** can store at least: a packet or message that is an interest, a content object, a push interest or request, a push response, or a notification; a value for a variable; a proposed value for the variable; an agreed-upon value for the variable; an identifier of the variable; a round number; a payload of an interest or a content object; a name for an interest or a content object; a name that is a hierarchically structured variable length identifier (HSVLI) comprised of contiguous name components ordered from a most general level to a most specific level; an indicator or identifier of a majority of nodes, a coordinator for a plurality of nodes, a node, an acceptor, or a learner; an iteration number corresponding to the round number; a negative acknowledgment if the agreed-upon value for the variable is indeterminate; a no-operation indicator if no agreed-upon value for the variable exists; a positive acknowledgment (ACK); a negative acknowledgment (NACK); a content object that includes a responsive ACK or NACK; a routable prefix for one of the majority of the nodes; an identifier for a consensus group to which the one of the majority of the nodes belongs, wherein the plurality of nodes belong to the consensus group; an indicator of a logical program associated with the variable identifier; a name that indicates a request to approve the proposed value for the variable, a request to accept the proposed value, a request to read the agreed-upon value, or a notification to allow a receiving device to learn the agreed-upon value; an indicator of a multicast group; an indicator of a group version to which the majority of the nodes belongs; a link to a piece of content which describes a current state of an algorithm; a link to a manifest, which

is a content object indicating a collection of other content objects; the manifest embedded in the proposed value; and a content object with a lifetime set to a small or a zero value;

[0076] The data structures and code described in this detailed description are typically stored on a computer-readable storage medium, which may be any device or medium that can store code and/or data for use by a computer system. The computer-readable storage medium includes, but is not limited to, volatile memory, non-volatile memory, magnetic and optical storage devices such as disk drives, magnetic tape, CDs (compact discs), DVDs (digital versatile discs or digital video discs), or other media capable of storing computer-readable media now known or later developed.

[0077] The methods and processes described in the detailed description section can be embodied as code and/or data, which can be stored in a computer-readable storage medium as described above. When a computer system reads and executes the code and/or data stored on the computer-readable storage medium, the computer system performs the methods and processes embodied as data structures and code and stored within the computer-readable storage medium.

[0078] Furthermore, the methods and processes described above can be included in hardware modules. For example, the hardware modules can include, but are not limited to, application-specific integrated circuit (ASIC) chips, field-programmable gate arrays (FPGAs), and other programmable-logic devices now known or later developed. When the hardware modules are activated, the hardware modules perform the methods and processes included within the hardware modules.

[0079] The foregoing descriptions of embodiments of the present invention have been presented for purposes of illustration and description only. They are not intended to be exhaustive or to limit the present invention to the forms disclosed. Accordingly, many modifications and variations will be apparent to practitioners skilled in the art. Additionally, the above disclosure is not intended to limit the present invention. The scope of the present invention is defined by the appended claims.

What is claimed is:

1. A computer system for facilitating distributed consensus in a content centric network, the system comprising:

a processor; and

a storage device storing instructions that when executed by the processor cause the processor to perform a method, the method comprising:

generating a first interest indicating a request to approve a proposed value for a variable, wherein the computer system is a coordinator for a plurality of nodes;

in response to receiving a positive acknowledgment of the first interest from a majority of the nodes, generating a second interest indicating a request to accept the proposed value, wherein a name for the first interest and a name for the second interest include an identifier of the variable and a round number, and wherein a payload of the first interest and a payload of the second interest include the proposed value; and

in response to receiving a positive acknowledgement of the second interest from the majority of the nodes, generating a notification indicating that an agreed-upon value for the variable is the proposed value.

2. The computer system of claim **1**, wherein the method further comprises:

in response to receiving a third interest indicating a request to read the agreed-upon value for the variable, wherein a name for the third interest includes the variable identifier and the round number, generating a content object which indicates the round number and one or more of:

the agreed-upon value for the variable;

an iteration number corresponding to the round number;

a negative acknowledgment if the agreed-upon value for the variable is indeterminate; and

a no-operation indicator if no agreed-upon value for the variable exists.

3. The computer system of claim **2**, wherein the positive acknowledgement of the second interest is transmitted to the coordinator by the majority of the nodes,

wherein the notification is generated by the majority of the nodes and further transmitted to a learning device, which transmits the notification as a fourth interest to one or more systems interested in the agreed-upon value for the variable.

4. The computer system of claim **3**, wherein the name for the first interest, the name for the second interest, the name for the third interest, and the name for the fourth interest further include one or more of:

a routable prefix for one of the majority of the nodes;

an identifier for a consensus group to which the one of the majority of the nodes belongs, wherein the plurality of nodes belong to the consensus group;

an indicator of a logical program associated with the variable identifier; and

an iteration number corresponding to the round number.

5. The computer system of claim **3**, wherein the name for the first interest indicates the request to approve the proposed value for the variable,

wherein the name for the second interest indicates the request to accept the proposed value,

wherein the name for the third interest indicates the request to read the agreed-upon value, and

wherein the name for the fourth interest indicates the notification to allow a receiving device to learn the agreed-upon value.

6. The computer system of claim **1**, wherein the method further comprises:

transmitting, by the coordinator, the first interest to a multicast group comprised of the majority of the nodes; and

transmitting, by the coordinator, the second interest to the multicast group,

wherein the name for the first interest and the name for the second interest further include one or more of:

an identifier for a consensus group to which the one of the majority of the nodes belongs, wherein the plurality of nodes belong to the consensus group, wherein the consensus group identifier is the most general level name component; and

an indicator of a group version to which the majority of the nodes belongs,

wherein the payload of the first interest further includes a routable prefix of the coordinator to be used by a node in response to the first interest, and

wherein the payload of the second interest further includes a routable prefix of a target to be used by a node in response to the second interest.

7. The computer system of claim **1**, wherein the proposed value is one or more of:

a link to a piece of content which describes a current state of an algorithm;

a link to a manifest, which is a content object indicating a collection of other content objects; and

the manifest embedded in the proposed value.

8. The computer system of claim **1**, wherein a response by one of the majority of the nodes to an interest is a content object with a same name as the name for the interest, wherein the content object has a lifetime set to a small or a zero value.

9. A computer system for facilitating distributed consensus in a content centric network, the system comprising:

a processor; and

a storage device storing instructions that when executed by the processor cause the processor to perform a method, the method comprising:

receiving a first interest indicating a request to approve a proposed value for a variable, wherein the first interest is received from a coordinator for a plurality of nodes by a node;

in response to agreeing with the proposed value, transmitting a positive acknowledgement of the first interest to the coordinator;

receiving a second interest indicating a request to accept the proposed value, wherein a name for the first interest and a name for the second interest include an identifier of the variable and a round number, and wherein a payload of the first interest and a payload of the second interest include the proposed value; and

in response to accepting the proposed value, transmitting a positive acknowledgement of the second interest to the coordinator.

10. The computer system of claim **9**, wherein in response to accepting the proposed value, the method further comprises:

transmitting a notification indicating that an agreed-upon value for the variable is the proposed value to a learning device, which transmits the notification as a third interest to one or more systems interested in the agreed-upon value for the variable.

11. The computer system of claim **9**, wherein the method further comprises:

in response to not agreeing with the proposed value, transmitting a negative acknowledgment of the first interest to the coordinator, wherein the negative acknowledgment includes a previous value for the variable corresponding to a previous round number.

12. The computer system of claim **9**, wherein the name for the first interest, the name for the second interest, and the name for the third interest further include one or more of:

a routable prefix for one of the majority of the nodes;

an identifier for a consensus group to which the one of the majority of the nodes belongs, wherein the plurality of nodes belong to the consensus group;

an indicator of a logical program associated with the variable identifier; and

an iteration number corresponding to the round number; and

wherein the name for the first interest indicates the request to approve the proposed value for the variable,

wherein the name for the second interest indicates the request to accept the proposed value, and

wherein the name for the third interest indicates the notification to allow a receiving device to learn the agreed-upon value.

13. A computer-implemented method for facilitating distributed consensus in a content centric network, the method comprising:

generating a first interest indicating a request to approve a proposed value for a variable, wherein the computer system is a coordinator for a plurality of nodes;

in response to receiving a positive acknowledgment of the first interest from a majority of the nodes, generating a second interest indicating a request to accept the proposed value, wherein a name for the first interest and a name for the second interest include an identifier of the variable and a round number, and wherein a payload of the first interest and a payload of the second interest include the proposed value; and

in response to receiving a positive acknowledgement of the second interest from the majority of the nodes, generating a notification indicating that an agreed-upon value for the variable is the proposed value.

14. The method of claim **13**, further comprising:

in response to receiving a third interest indicating a request to read the agreed-upon value for the variable, wherein a name for the third interest includes the variable identifier and the round number, generating a content object which indicates the round number and one or more of:

the agreed-upon value for the variable;

an iteration number corresponding to the round number;

a negative acknowledgment if the agreed-upon value for the variable is indeterminate; and

a no-operation indicator if no agreed-upon value for the variable exists.

15. The method of claim **14**, wherein the positive acknowledgement of the second interest is transmitted to the coordinator by the majority of the nodes,

wherein the notification is generated by the majority of the nodes and further transmitted to a learning device, which transmits the notification as a fourth interest to one or more systems interested in the agreed-upon value for the variable.

16. The method of claim **15**, wherein the name for the first interest, the name for the second interest, the name for the third interest, and the name for the fourth interest further include one or more of:

a routable prefix for one of the majority of the nodes;

an identifier for a consensus group to which the one of the majority of the nodes belongs, wherein the plurality of nodes belong to the consensus group;

an indicator of a logical program associated with the variable identifier; and

an iteration number corresponding to the round number.

17. The method of claim **16**, wherein the name for the first interest indicates the request to approve the proposed value for the variable,

wherein the name for the second interest indicates the request to accept the proposed value,

wherein the name for the third interest indicates the request to read the agreed-upon value, and

wherein the name for the fourth interest indicates the notification to allow a receiving device to learn the agreed-upon value.

18. The method of claim **13**, further comprising:

transmitting, by the coordinator, the first interest to a multicast group comprised of the majority of the nodes; and

transmitting, by the coordinator, the second interest to the multicast group,

wherein the name for the first interest and the name for the second interest further include one or more of:

an identifier for a consensus group to which the one of the majority of the nodes belongs, wherein the plurality of nodes belong to the consensus group, wherein the consensus group identifier is the most general level name component; and

an indicator of a group version to which the majority of the nodes belongs,

wherein the payload of the first interest further includes a routable prefix of the coordinator to be used by a node in response to the first interest, and

wherein the payload of the second interest further includes a routable prefix of a target to be used by a node in response to the second interest.

19. The method of claim **13**, wherein the proposed value is one or more of:

a link to a piece of content which describes a current state of an algorithm;

a link to a manifest, which is a content object indicating a collection of other content objects; and

the manifest embedded in the proposed value.

20. The method of claim **1**, wherein a response by one of the majority of the nodes to an interest is a content object with a same name as the name for the interest, wherein the content object has a lifetime set to a small or a zero value.

* * * * *