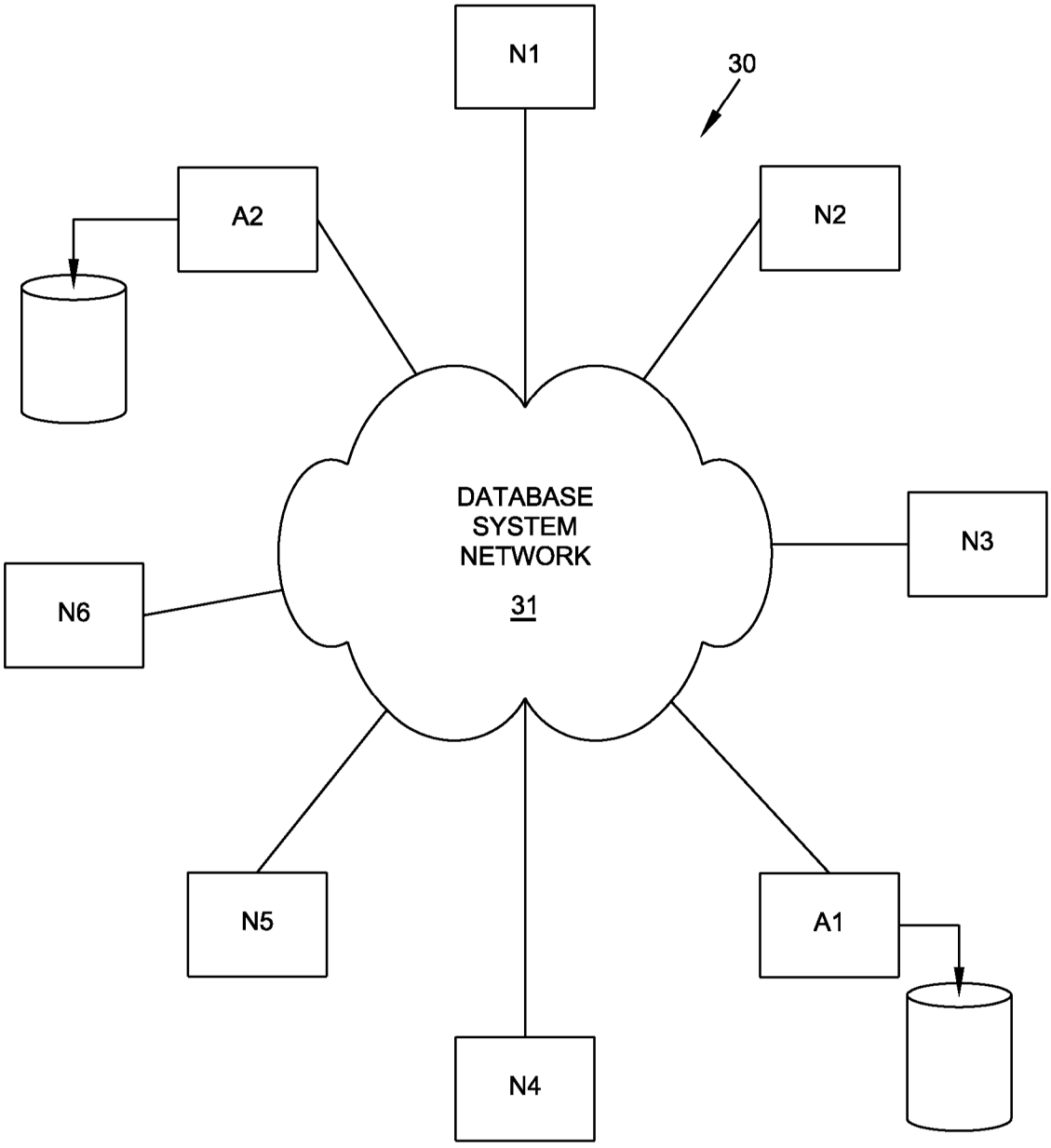

FIG. 1

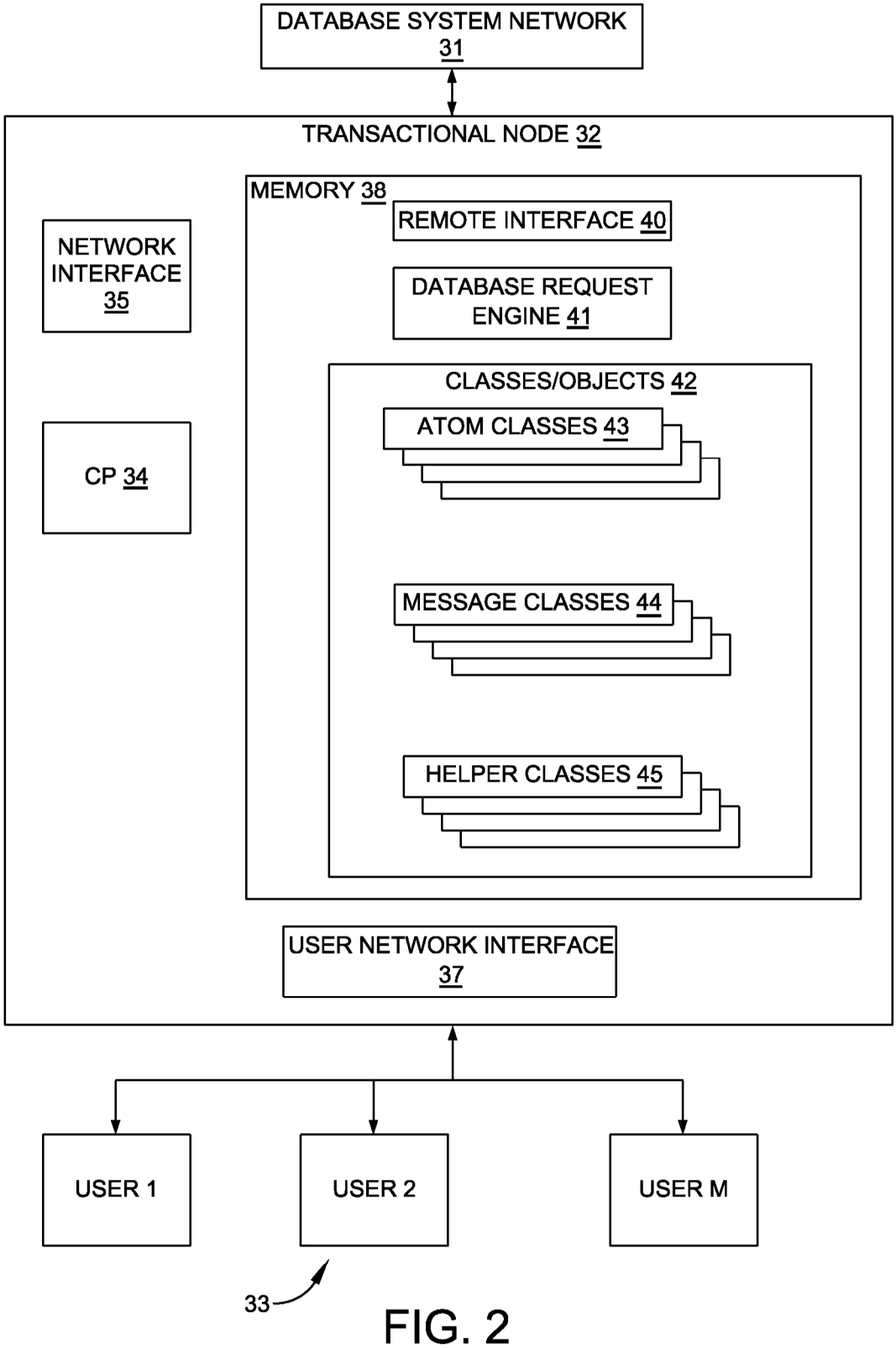
DATABASE SYSTEM NETWORK 31
TRANSACTIONAL NODE 32
MEMORY 38
REMOTE INTERFACE 40
DATABASE REQUEST ENGINE 41
CLASSES/OBJECTS 42
ATOM CLASSES 43
MESSAGE CLASSES 44
HELPER CLASSES 45
NETWORK INTERFACE 35
CP 34
USER NETWORK INTERFACE 37
USER 1
USER 2
USER M
33

FIG. 2

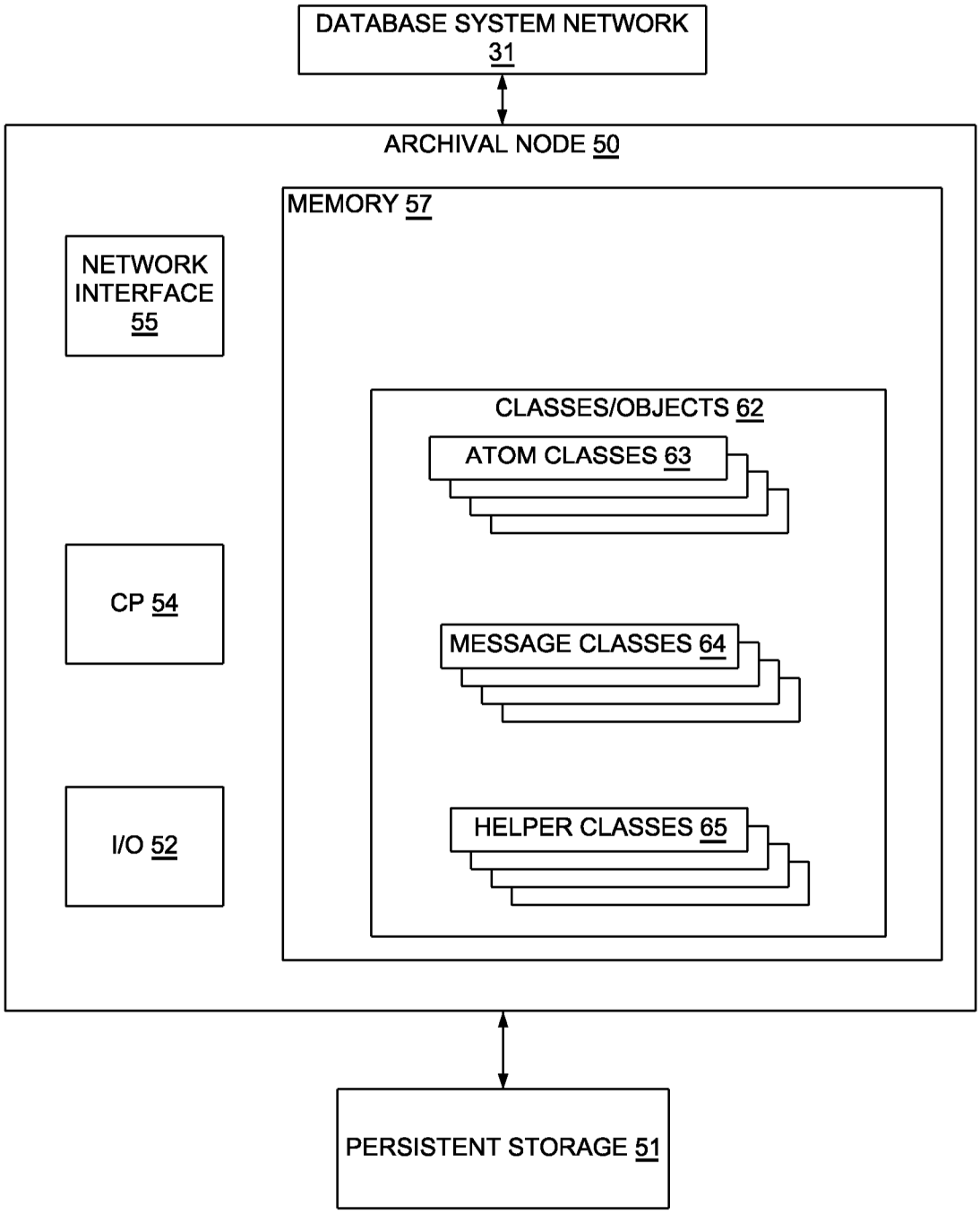
DATABASE SYSTEM NETWORK 31
ARCHIVAL NODE 50
MEMORY 57
NETWORK INTERFACE 55
CLASSES/OBJECTS 62
ATOM CLASSES 63
CP 54
MESSAGE CLASSES 64
I/O 52
HELPER CLASSES 65
PERSISTENT STORAGE 51

FIG. 3

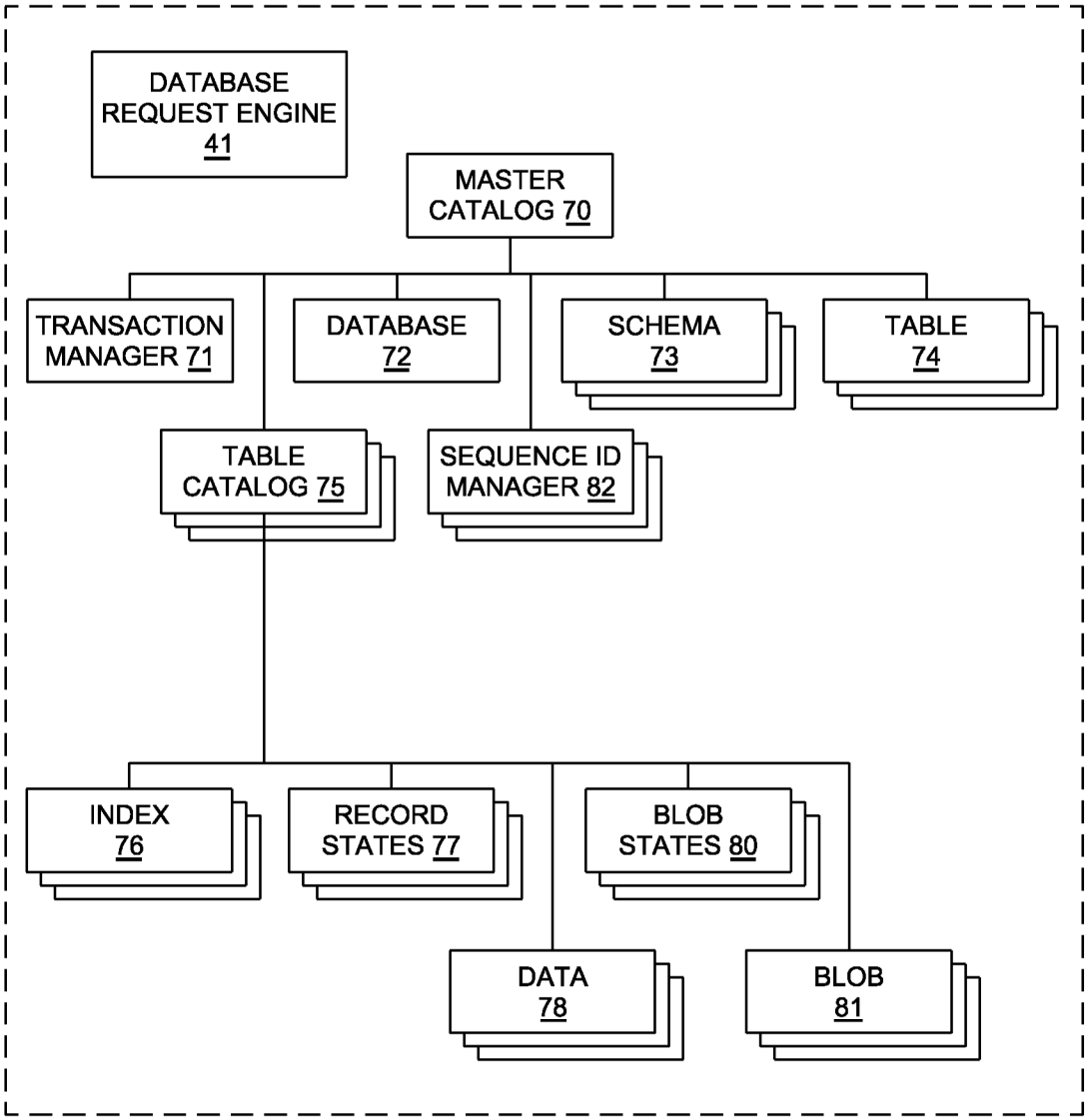
DATABASE REQUEST ENGINE 41
MASTER CATALOG 70
TRANSACTION MANAGER 71
DATABASE 72
SCHEMA 73
TABLE 74
TABLE CATALOG 75
SEQUENCE ID MANAGER 82
INDEX 76
RECORD STATES 77
BLOB STATES 80
DATA 78
BLOB 81

FIG. 4A

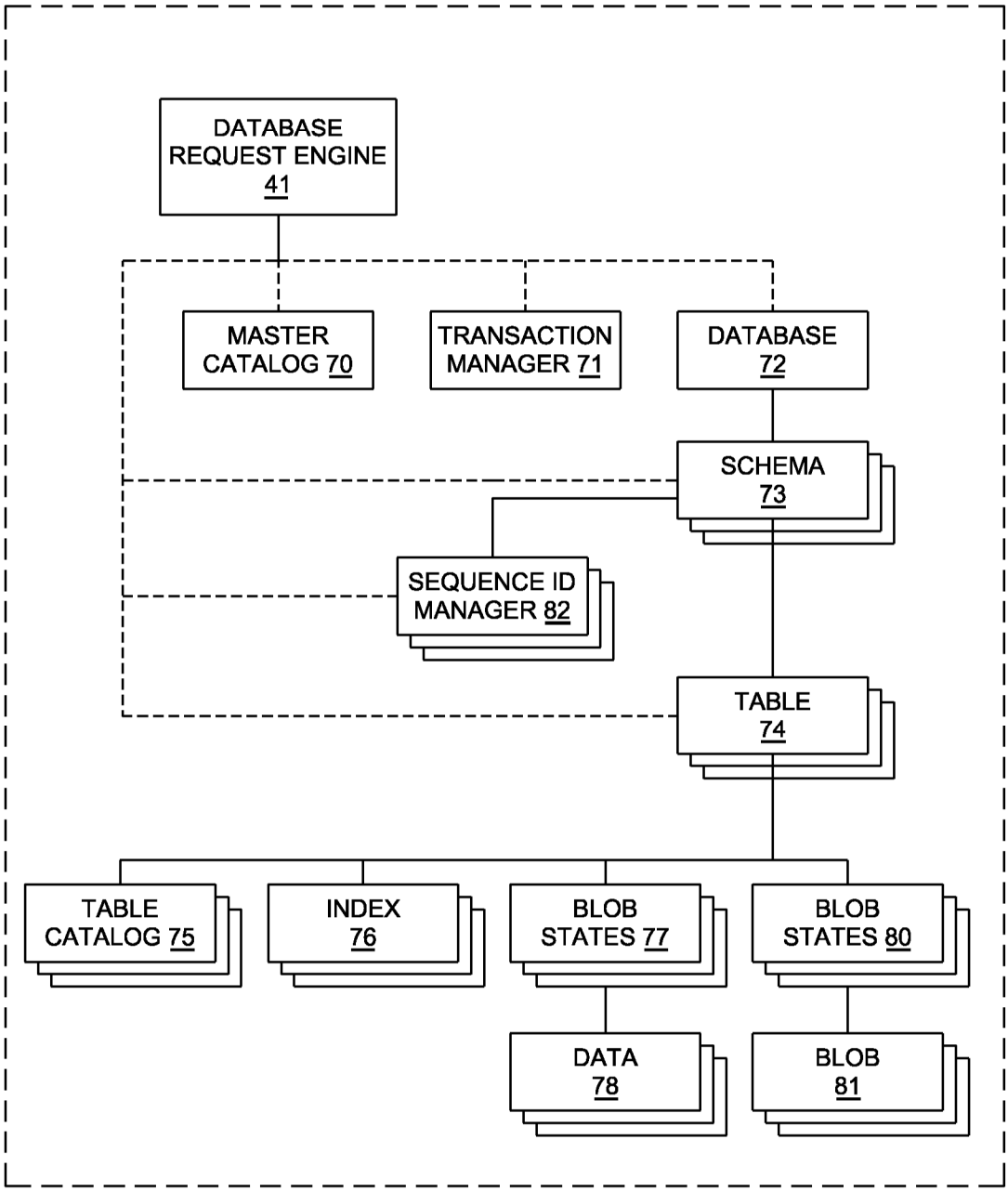
DATABASE REQUEST ENGINE 41
MASTER CATALOG 70
TRANSACTION MANAGER 71
DATABASE 72
SCHEMA 73
SEQUENCE ID MANAGER 82
TABLE 74
TABLE CATALOG 75
INDEX 76
BLOB STATES 77
BLOB STATES 80
DATA 78
BLOB 81

FIG. 4B

MASTER CATALOG ATOM 70

70A	MASTER CATALOG ID=1
70B	POINTER TO MASTER CATALOG
70C	POINTER TO CREATING CATALOG
70D	POINTER TO CHAIRMAN
70E	CHANGE NUMBER
70F	LIST OF RELAYS
70G	CYCLE REFERENCE
70H	ACTIVE NODE LIST
70I	STATUS STATES
70J	GLOBAL NODE IDS
70K	LOCAL NODE IDS
70L	CONNECTION MANAGER
70M	PENDING NODE
70N	DATABASE UUI
70P	CATALOG OF DATABASE, SCHEMA, TABLE AND TABLE CATALOG ATOMS
70Q	PASSWORD
70R	ACTUAL VERSION
70S	SOFTWARE VERSION
70T	POINTER TO TRANSACTION MANAGER ATOM
70U	POINTER TO CONFIGURATION OBJECT
70V	POINTER TO GARBAGE COLLECTION THREAD
70W	POINTER TO PING THREAD

FIG. 5

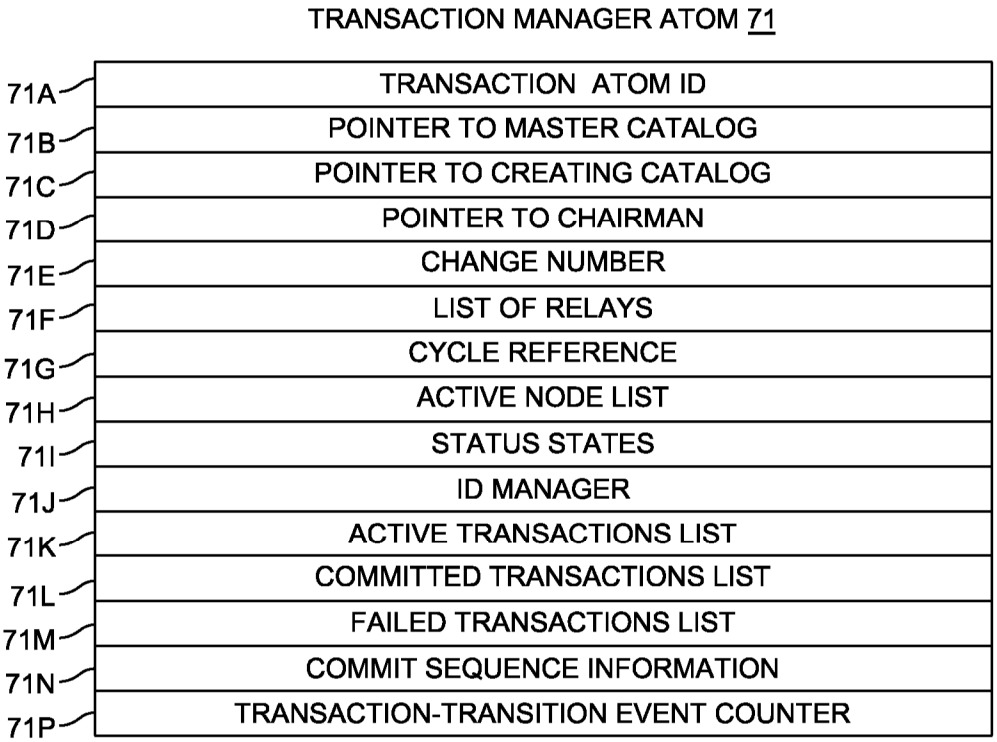
TRANSACTION MANAGER ATOM 71
71A
TRANSACTION ATOM ID
71B
POINTER TO MASTER CATALOG
71C
POINTER TO CREATING CATALOG
71D
POINTER TO CHAIRMAN
71E
CHANGE NUMBER
71F
LIST OF RELAYS
71G
CYCLE REFERENCE
71H
ACTIVE NODE LIST
71I
STATUS STATES
71J
ID MANAGER
71K
ACTIVE TRANSACTIONS LIST
71L
COMMITTED TRANSACTIONS LIST
71M
FAILED TRANSACTIONS LIST
71N
COMMIT SEQUENCE INFORMATION
71P
TRANSACTION-TRANSITION EVENT COUNTER

FIG. 6

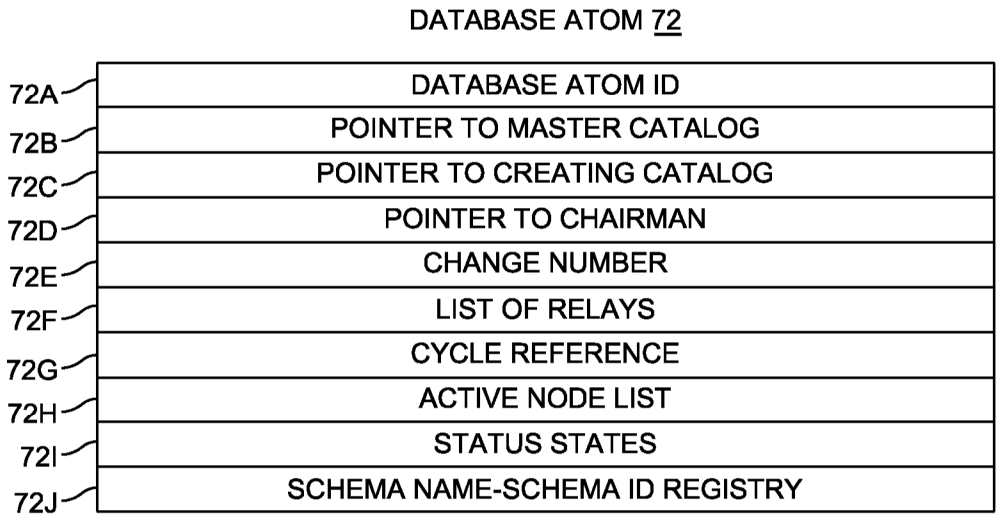
DATABASE ATOM 72
72A
DATABASE ATOM ID
72B
POINTER TO MASTER CATALOG
72C
POINTER TO CREATING CATALOG
72D
POINTER TO CHAIRMAN
72E
CHANGE NUMBER
72F
LIST OF RELAYS
72G
CYCLE REFERENCE
72H
ACTIVE NODE LIST
72I
STATUS STATES
72J
SCHEMA NAME-SCHEMA ID REGISTRY

FIG. 7

SCHEMA ATOM 73

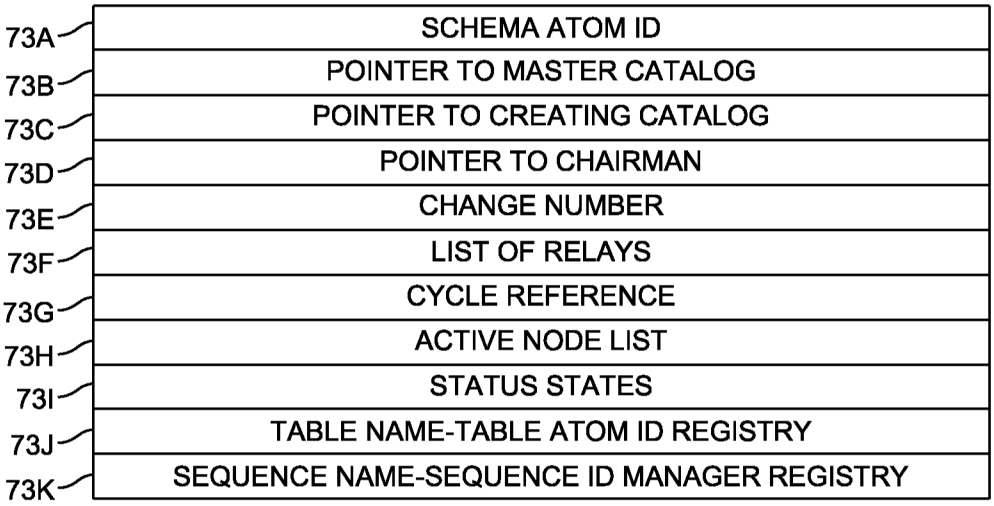

FIG. 8

TABLE ATOM 74

74A	TABLE ATOM ID
74B	POINTER TO MASTER CATALOG
74C	POINTER TO CREATING CATALOG
74D	POINTER TO CHAIRMAN
74E	CHANGE NUMBER
74F	LIST OF RELAYS
74G	CYCLE REFERENCE
74H	ACTIVE NODE LIST
74I	STATUS STATES
74J	POINTER TO CORRESPONDING TABLE CATALOG ATOM
74K	LIST OF FIELDS
74L	ID MANAGER FOR FIELD IDS
74M	ID MANAGER FOR INDEX ATOM IDS
74N	ID MANAGER FOR DATA ATOM IDS
74P	ID MANAGER FOR BLOB ATOM IDS
74Q	ID MANAGER FOR TYPES
74R	SUBTYPE LISTS
74S	ARRAY OF RECORD STATES ATOM IDS
74T	ARRAY OF BLOB STATES ATOM IDS

FIG. 9

TABLE CATALOG ATOM 75

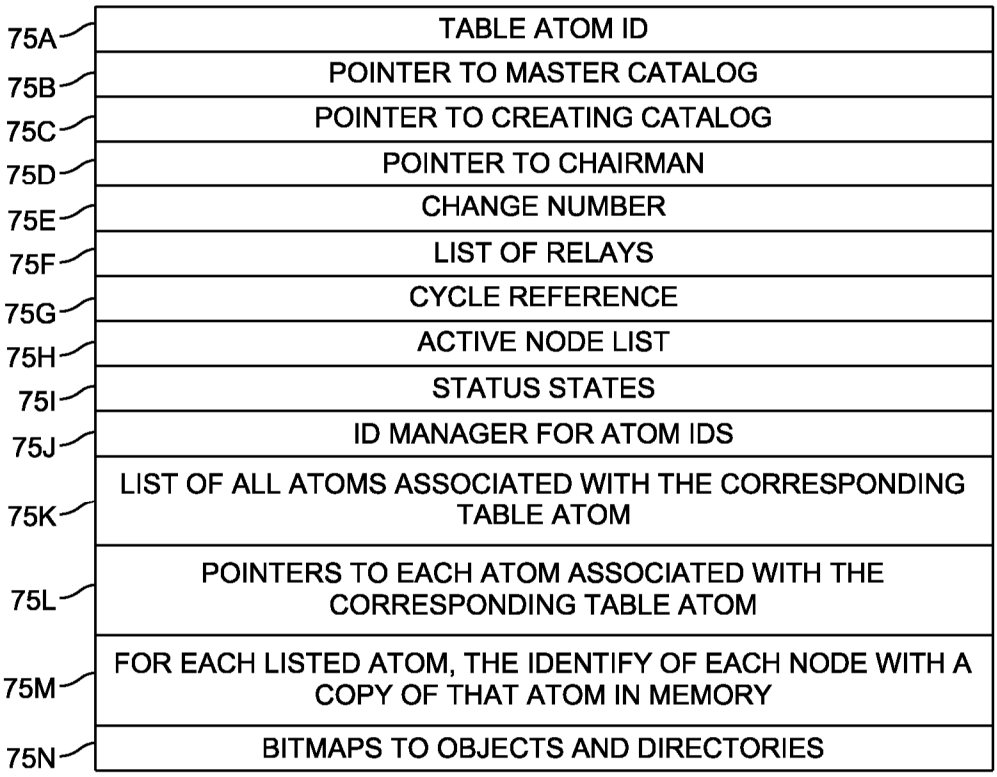

FIG. 10

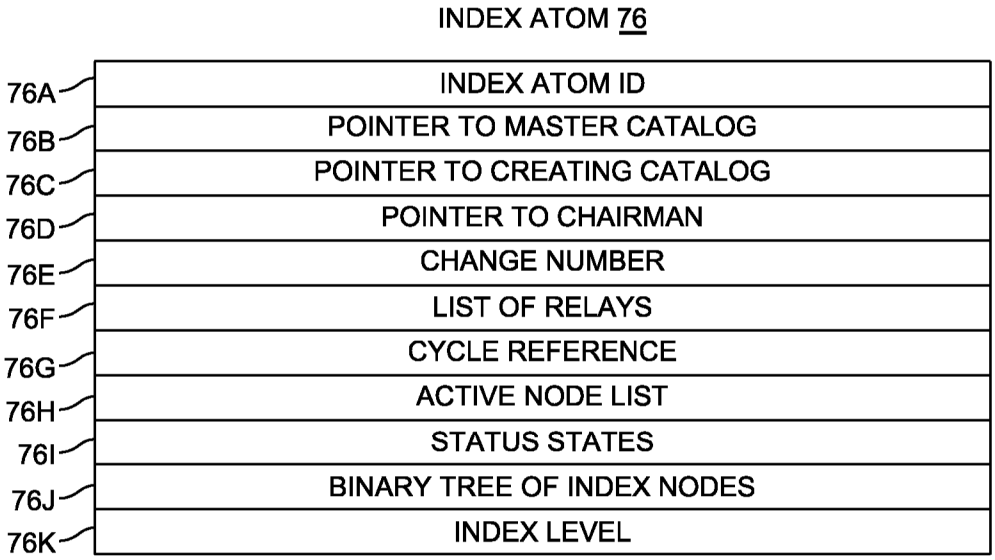
INDEX ATOM 76
76A
INDEX ATOM ID
76B
POINTER TO MASTER CATALOG
76C
POINTER TO CREATING CATALOG
76D
POINTER TO CHAIRMAN
76E
CHANGE NUMBER
76F
LIST OF RELAYS
76G
CYCLE REFERENCE
76H
ACTIVE NODE LIST
76I
STATUS STATES
76J
BINARY TREE OF INDEX NODES
76K
INDEX LEVEL

FIG. 11

RECORD STATES ATOM 77

77A	RECORD STATES ATOM ID
77B	POINTER TO MASTER CATALOG
77C	POINTER TO CREATING CATALOG
77D	POINTER TO CHAIRMAN
77E	CHANGE NUMBER
77F	LIST OF RELAYS
77G	CYCLE REFERENCE
77H	ACTIVE NODE LIST
77I	STATUS STATES
77J	ARRAY OF MANAGED DATA ATOMS
77K	RECORD ID FOR BASE RECORD
77L	POINTER TO TABLE ATOM
77M	BITMAP OF VERSIONED RECORDS
77N	FOR EACH RECORD VERSION
77P	TRANSACTION ID
77Q	FORMAT VERSION
77R	RECORD VERSION SEQUENCE NUMBER
77S	LOCATION OF NEXT OLDER VERSION
77T	INDEX TO DATA ATOM ARRAY
77U	SLOT IN DATA ATOM FOR ACTUAL DATA

FIG. 12

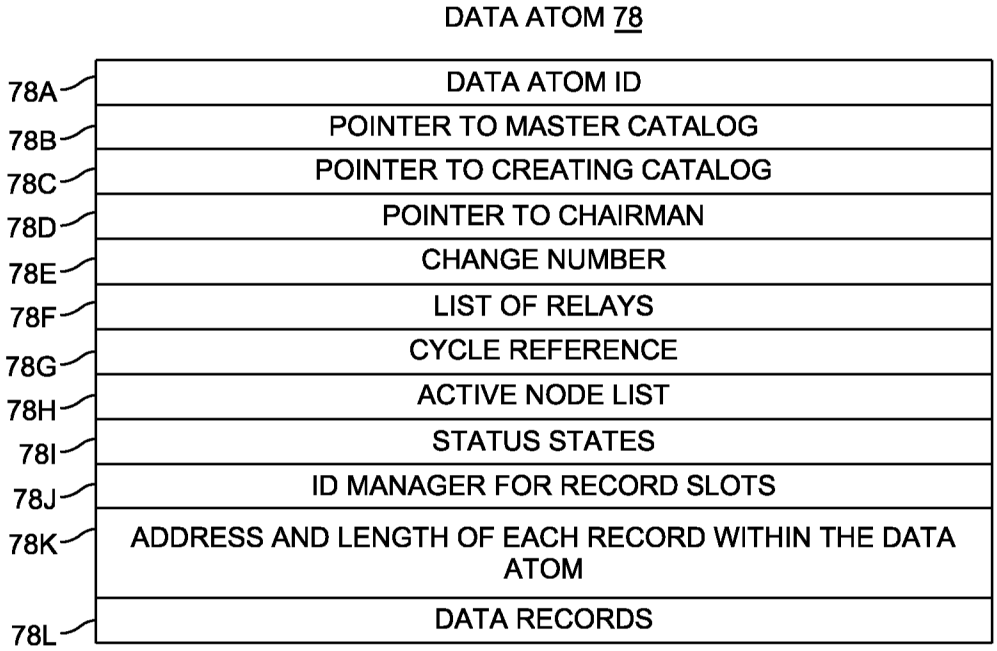
DATA ATOM 78
78A
DATA ATOM ID
78B
POINTER TO MASTER CATALOG
78C
POINTER TO CREATING CATALOG
78D
POINTER TO CHAIRMAN
78E
CHANGE NUMBER
78F
LIST OF RELAYS
78G
CYCLE REFERENCE
78H
ACTIVE NODE LIST
78I
STATUS STATES
78J
ID MANAGER FOR RECORD SLOTS
78K
ADDRESS AND LENGTH OF EACH RECORD WITHIN THE DATA ATOM
78L
DATA RECORDS

FIG. 13

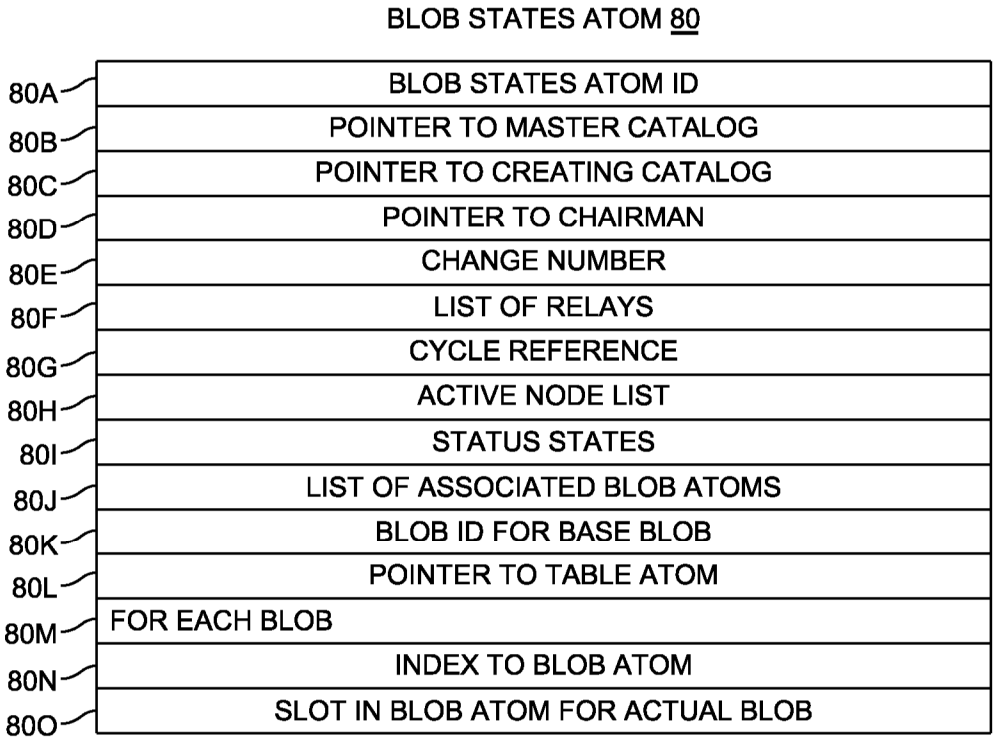
BLOB STATES ATOM 80
80A
BLOB STATES ATOM ID
80B
POINTER TO MASTER CATALOG
80C
POINTER TO CREATING CATALOG
80D
POINTER TO CHAIRMAN
80E
CHANGE NUMBER
80F
LIST OF RELAYS
80G
CYCLE REFERENCE
80H
ACTIVE NODE LIST
80I
STATUS STATES
80J
LIST OF ASSOCIATED BLOB ATOMS
80K
BLOB ID FOR BASE BLOB
80L
POINTER TO TABLE ATOM
80M
FOR EACH BLOB
80N
INDEX TO BLOB ATOM
80O
SLOT IN BLOB ATOM FOR ACTUAL BLOB

FIG. 14

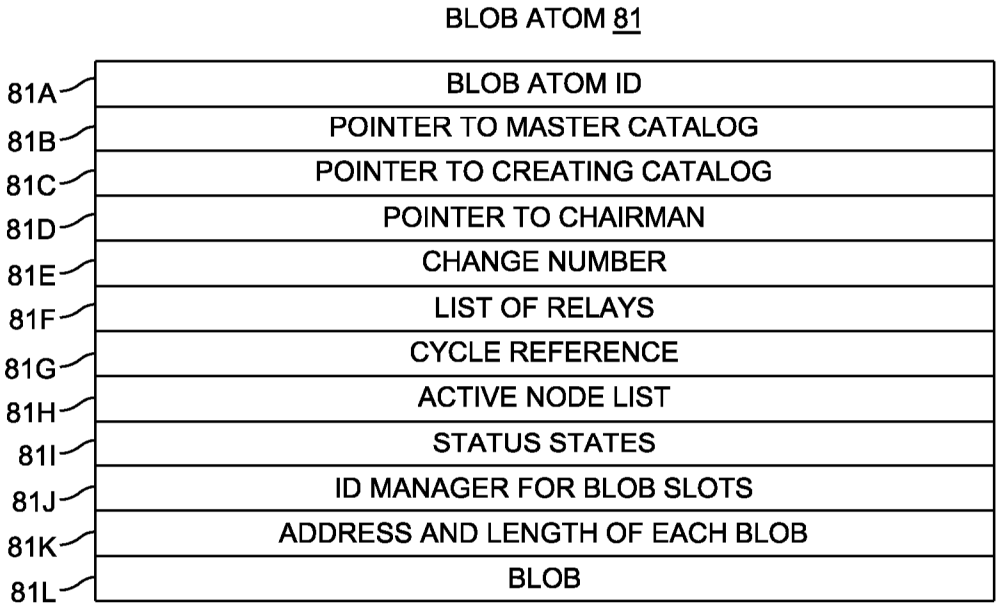
BLOB ATOM 81
81A
BLOB ATOM ID
81B
POINTER TO MASTER CATALOG
81C
POINTER TO CREATING CATALOG
81D
POINTER TO CHAIRMAN
81E
CHANGE NUMBER
81F
LIST OF RELAYS
81G
CYCLE REFERENCE
81H
ACTIVE NODE LIST
81I
STATUS STATES
81J
ID MANAGER FOR BLOB SLOTS
81K
ADDRESS AND LENGTH OF EACH BLOB
81L
BLOB

FIG. 15

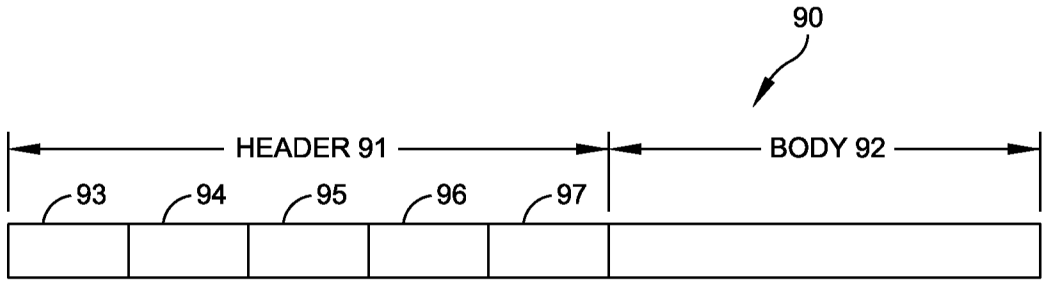

FIG. 16

MESSAGES

110	PING
111	PING ACKNOWLEDGE
112	CONNECT
113	WELCOME
114	NEW NODE
115	NODE STATE
116	OBJECT REQUEST
117	OBJECT
118	OBJECT AVAILABLE
119	OBJECT ACKNOWLEDGED
120	OBJECT COMPLETE
121	OBJECT UNAVAILABLE
122	BOUNCE OBJECT
123	REGISTER OBJECT
124	UNREGISTER OBJECT
125	OBJECT DELETE
126	OBJECT WRITTEN
127	ID REQUEST
128	ID DELEGATION
129	TABLE FIELD ADDED
130	TABLE FORMAT
131	TABLE REQUEST RECORDS
132	TABLE RECORDS OBJECT
133	TABLE RECORD
134	TABLE INDEX ADDED
135	TABLE BLOBS OBJECT
136	TABLE BLOB
137	TABLE TYPE
138	TABLE RECORD EXPUNGE
139	TABLE GARBAGE COLLECT
140	RECORDS UPDATE REQUEST
141	RECORDS UPDATE RESPONSE
142	RECORDS UPDATE
143	RECORDS DATA OBJECT
144	RECORDS RECORD
145	RECORDS PRUNE
146	BACKOUT RECORD
147	INDEX SPLIT REQUEST
148	INDEX SPLIT
149	INDEX PRUNE
150	INDEX NODE ADDED
151	TABLE INDEX READY
152	BLOBS BLOB
153	DATA RECORD
154	TRANSACTION STATE
155	TRANSACTION BLOCKAGE
156	TRANSACTION DEADLOCK
157	REQUEST WRITE TIMES
158	WRITE TIMES

FIG. 17

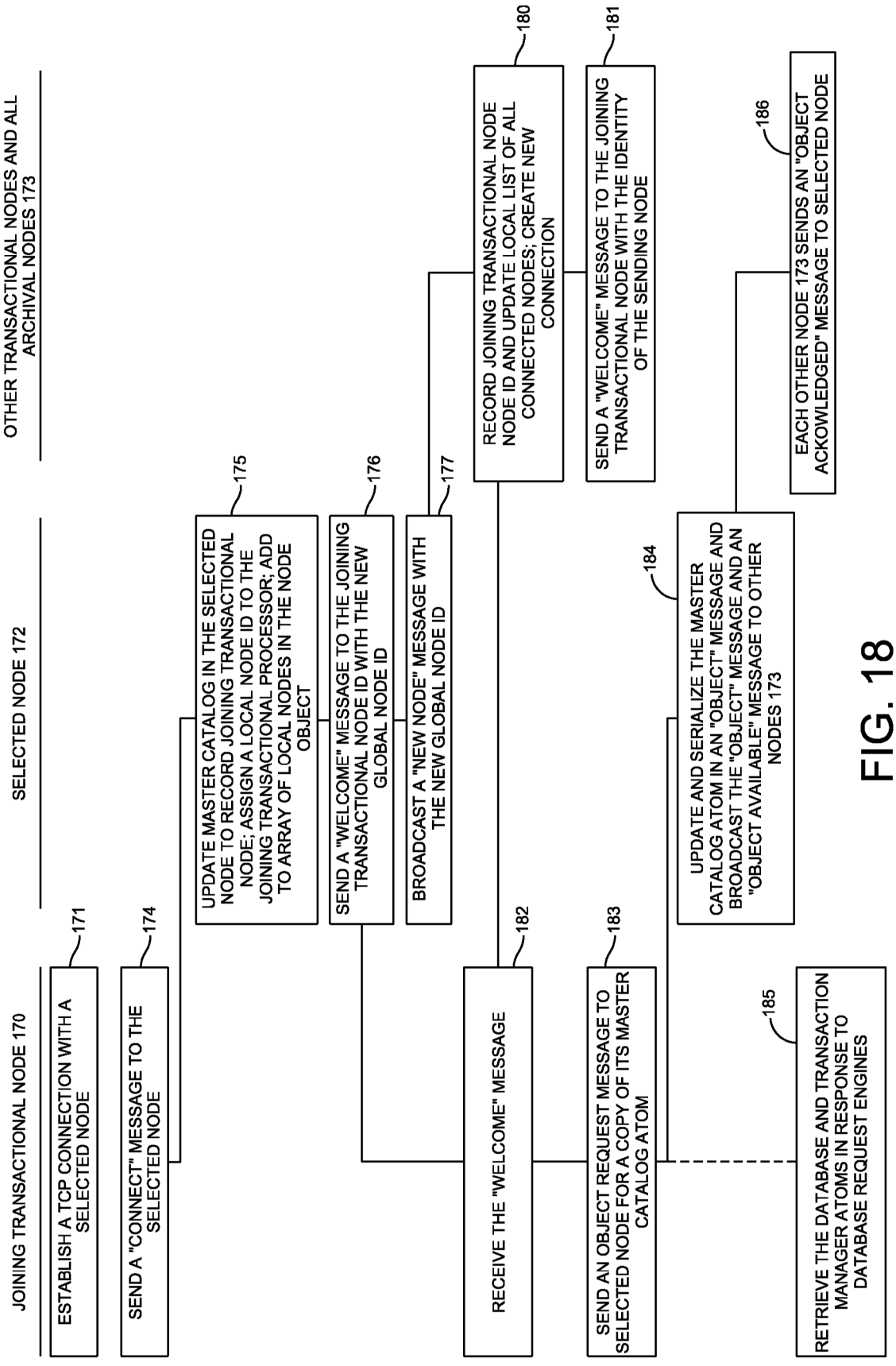
JOINING TRANSACTIONAL NODE 170
SELECTED NODE 172
OTHER TRANSACTIONAL NODES AND ALL ARCHIVAL NODES 173
ESTABLISH A TCP CONNECTION WITH A SELECTED NODE
171
SEND A "CONNECT" MESSAGE TO THE SELECTED NODE
174
UPDATE MASTER CATALOG IN THE SELECTED NODE TO RECORD JOINING TRANSACTIONAL NODE; ASSIGN A LOCAL NODE ID TO THE JOINING TRANSACTIONAL PROCESSOR; ADD TO ARRAY OF LOCAL NODES IN THE NODE OBJECT
175
SEND A "WELCOME" MESSAGE TO THE JOINING TRANSACTIONAL NODE ID WITH THE NEW GLOBAL NODE ID
176
BROADCAST A "NEW NODE" MESSAGE WITH THE NEW GLOBAL NODE ID
177
RECORD JOINING TRANSACTIONAL NODE NODE ID AND UPDATE LOCAL LIST OF ALL CONNECTED NODES; CREATE NEW CONNECTION
180
SEND A "WELCOME" MESSAGE TO THE JOINING TRANSACTIONAL NODE WITH THE IDENTITY OF THE SENDING NODE
181
EACH OTHER NODE 173 SENDS AN "OBJECT ACKOWLEDGED" MESSAGE TO SELECTED NODE
186
RECEIVE THE "WELCOME" MESSAGE
182
SEND AN OBJECT REQUEST MESSAGE TO SELECTED NODE FOR A COPY OF ITS MASTER CATALOG ATOM
183
UPDATE AND SERIALIZE THE MASTER CATALOG ATOM IN AN "OBJECT" MESSAGE AND BROADCAST THE "OBJECT" MESSAGE AND AN "OBJECT AVAILABLE" MESSAGE TO OTHER NODES 173
184
RETRIEVE THE DATABASE AND TRANSACTION MANAGER ATOMS IN RESPONSE TO DATABASE REQUEST ENGINES
185

FIG. 18

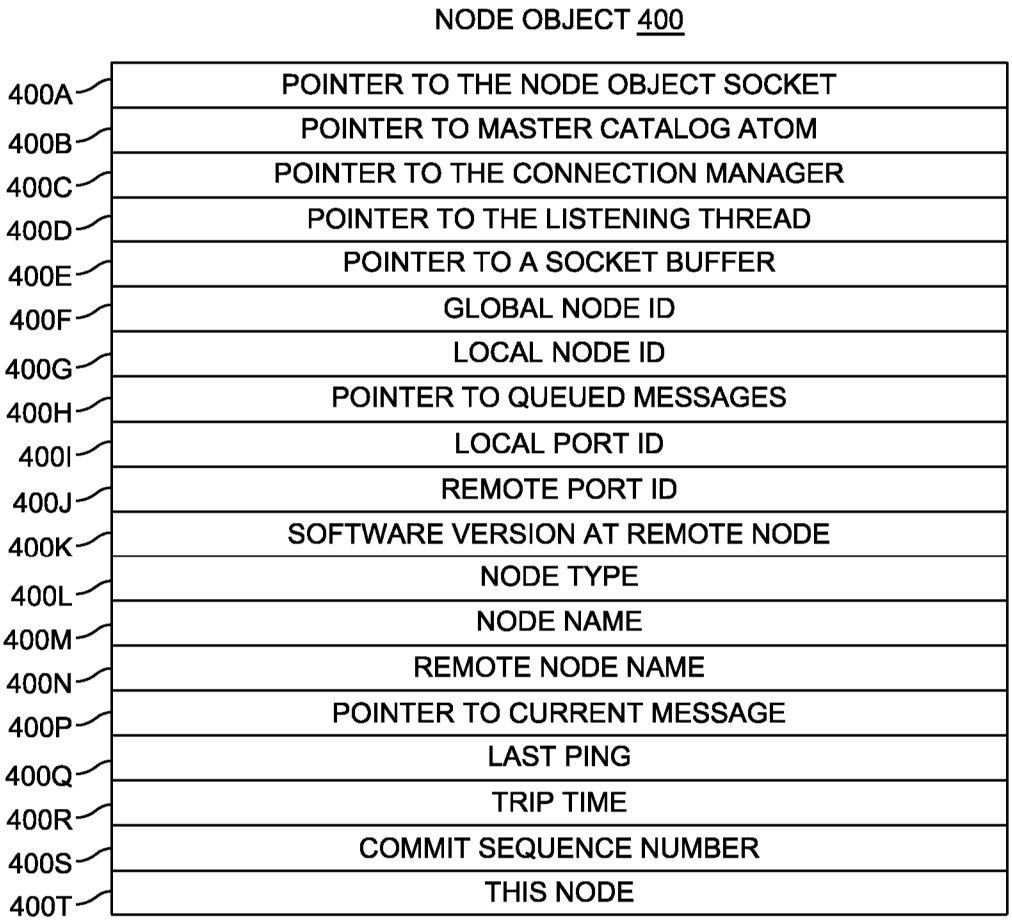
NODE OBJECT 400
400A
POINTER TO THE NODE OBJECT SOCKET
400B
POINTER TO MASTER CATALOG ATOM
400C
POINTER TO THE CONNECTION MANAGER
400D
POINTER TO THE LISTENING THREAD
400E
POINTER TO A SOCKET BUFFER
400F
GLOBAL NODE ID
400G
LOCAL NODE ID
400H
POINTER TO QUEUED MESSAGES
400I
LOCAL PORT ID
400J
REMOTE PORT ID
400K
SOFTWARE VERSION AT REMOTE NODE
400L
NODE TYPE
400M
NODE NAME
400N
REMOTE NODE NAME
400P
POINTER TO CURRENT MESSAGE
400Q
LAST PING
400R
TRIP TIME
400S
COMMIT SEQUENCE NUMBER
400T
THIS NODE

FIG. 19

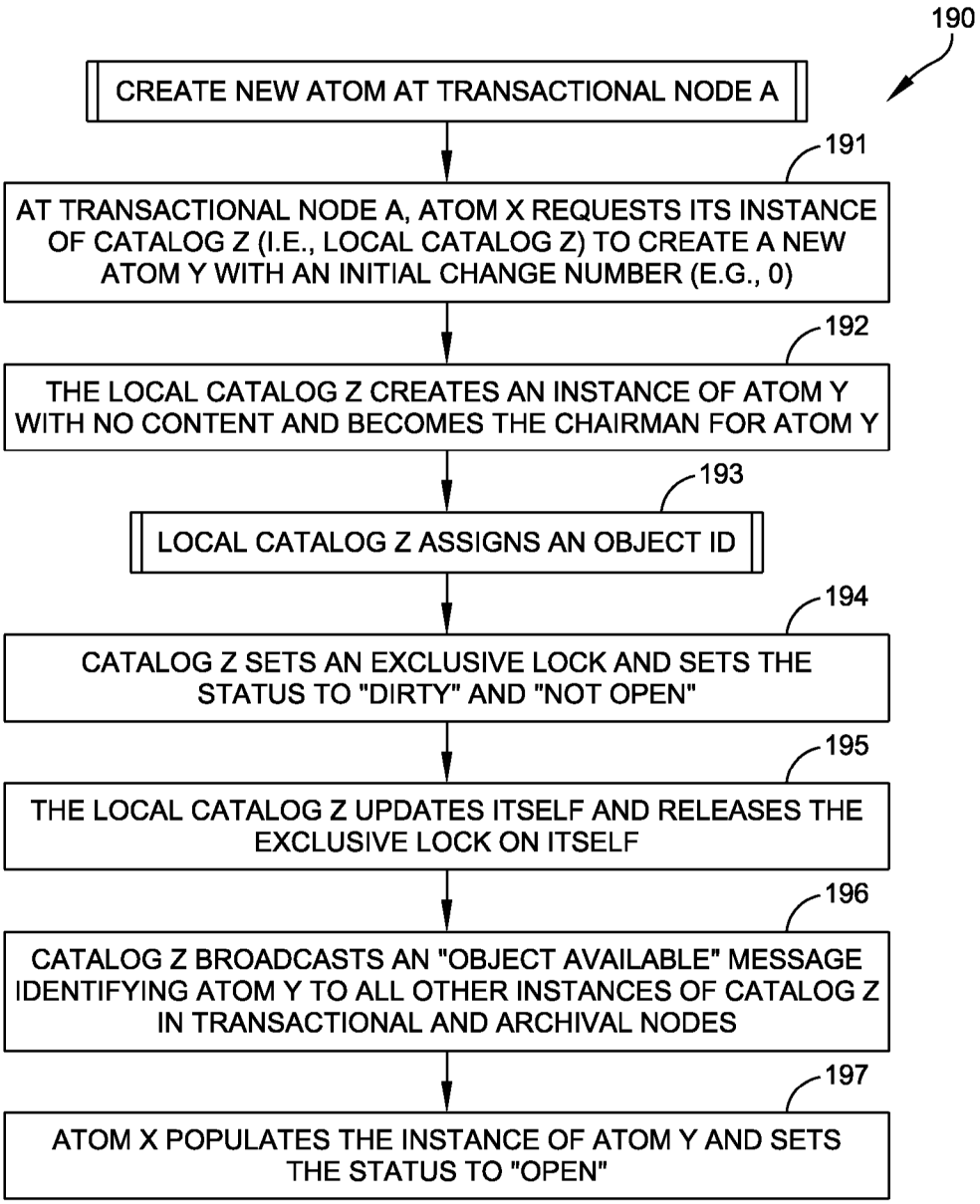
190
CREATE NEW ATOM AT TRANSACTIONAL NODE A
191
AT TRANSACTIONAL NODE A, ATOM X REQUESTS ITS INSTANCE OF CATALOG Z (I.E., LOCAL CATALOG Z) TO CREATE A NEW ATOM Y WITH AN INITIAL CHANGE NUMBER (E.G., 0)
192
THE LOCAL CATALOG Z CREATES AN INSTANCE OF ATOM Y WITH NO CONTENT AND BECOMES THE CHAIRMAN FOR ATOM Y
193
LOCAL CATALOG Z ASSIGNS AN OBJECT ID
194
CATALOG Z SETS AN EXCLUSIVE LOCK AND SETS THE STATUS TO "DIRTY" AND "NOT OPEN"
195
THE LOCAL CATALOG Z UPDATES ITSELF AND RELEASES THE EXCLUSIVE LOCK ON ITSELF
196
CATALOG Z BROADCASTS AN "OBJECT AVAILABLE" MESSAGE IDENTIFYING ATOM Y TO ALL OTHER INSTANCES OF CATALOG Z IN TRANSACTIONAL AND ARCHIVAL NODES
197
ATOM X POPULATES THE INSTANCE OF ATOM Y AND SETS THE STATUS TO "OPEN"

FIG. 20

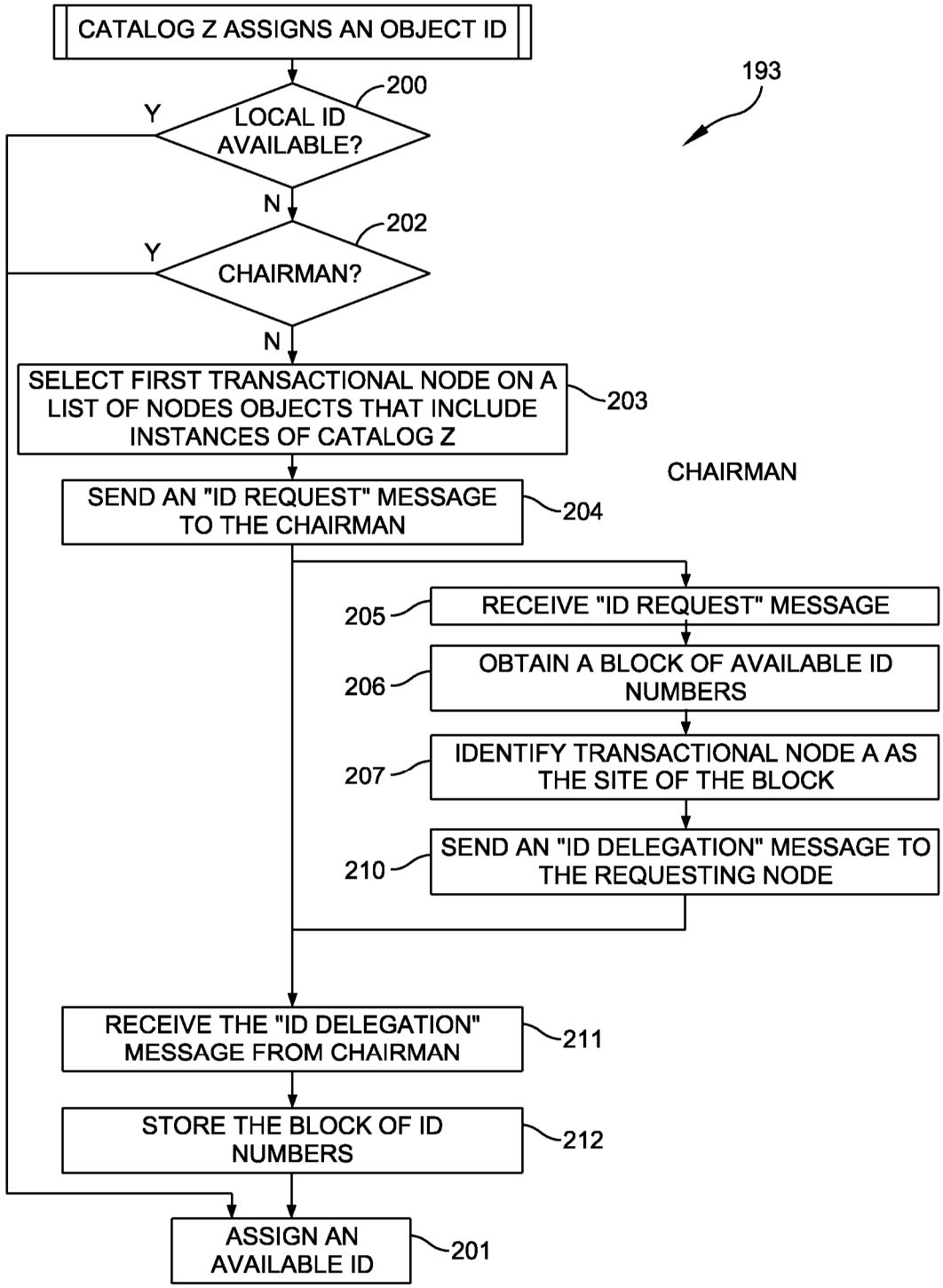
CATALOG Z ASSIGNS AN OBJECT ID
193
200
LOCAL ID AVAILABLE?
Y
N
202
CHAIRMAN?
Y
N
SELECT FIRST TRANSACTIONAL NODE ON A LIST OF NODES OBJECTS THAT INCLUDE INSTANCES OF CATALOG Z
203
CHAIRMAN
SEND AN "ID REQUEST" MESSAGE TO THE CHAIRMAN
204
205
RECEIVE "ID REQUEST" MESSAGE
206
OBTAIN A BLOCK OF AVAILABLE ID NUMBERS
207
IDENTIFY TRANSACTIONAL NODE A AS THE SITE OF THE BLOCK
210
SEND AN "ID DELEGATION" MESSAGE TO THE REQUESTING NODE
RECEIVE THE "ID DELEGATION" MESSAGE FROM CHAIRMAN
211
STORE THE BLOCK OF ID NUMBERS
212
ASSIGN AN AVAILABLE ID
201

FIG. 21

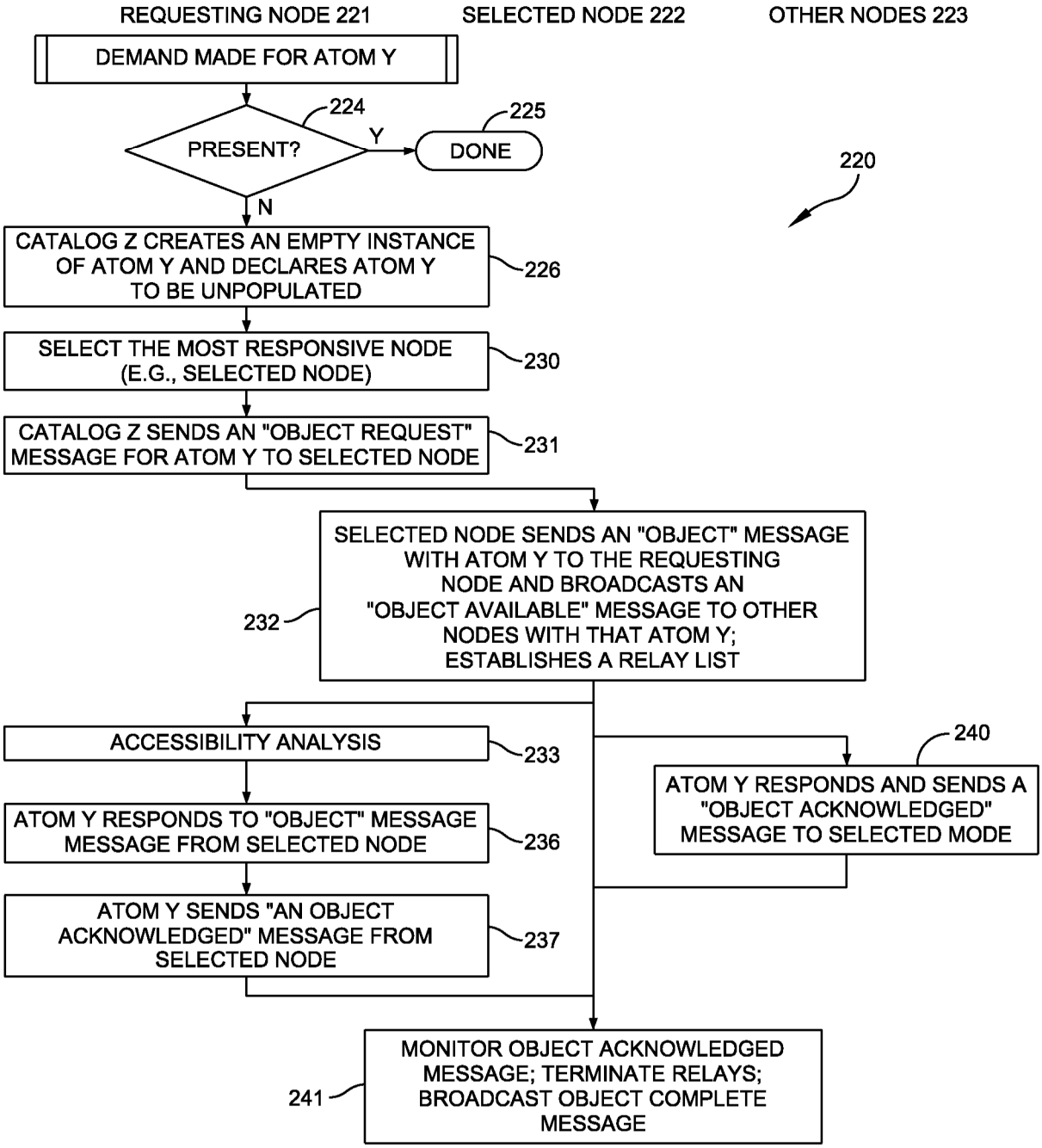
REQUESTING NODE 221
SELECTED NODE 222
OTHER NODES 223
DEMAND MADE FOR ATOM Y
224
PRESENT?
Y
225
DONE
N
220
CATALOG Z CREATES AN EMPTY INSTANCE OF ATOM Y AND DECLARES ATOM Y TO BE UNPOPULATED
226
SELECT THE MOST RESPONSIVE NODE (E.G., SELECTED NODE)
230
CATALOG Z SENDS AN "OBJECT REQUEST" MESSAGE FOR ATOM Y TO SELECTED NODE
231
SELECTED NODE SENDS AN "OBJECT" MESSAGE WITH ATOM Y TO THE REQUESTING NODE AND BROADCASTS AN "OBJECT AVAILABLE" MESSAGE TO OTHER NODES WITH THAT ATOM Y; ESTABLISHES A RELAY LIST
232
ACCESSIBILITY ANALYSIS
233
240
ATOM Y RESPONDS AND SENDS A "OBJECT ACKNOWLEDGED" MESSAGE TO SELECTED MODE
ATOM Y RESPONDS TO "OBJECT" MESSAGE MESSAGE FROM SELECTED NODE
236
ATOM Y SENDS "AN OBJECT ACKNOWLEDGED" MESSAGE FROM SELECTED NODE
237
MONITOR OBJECT ACKNOWLEDGED MESSAGE; TERMINATE RELAYS; BROADCAST OBJECT COMPLETE MESSAGE
241

FIG. 22

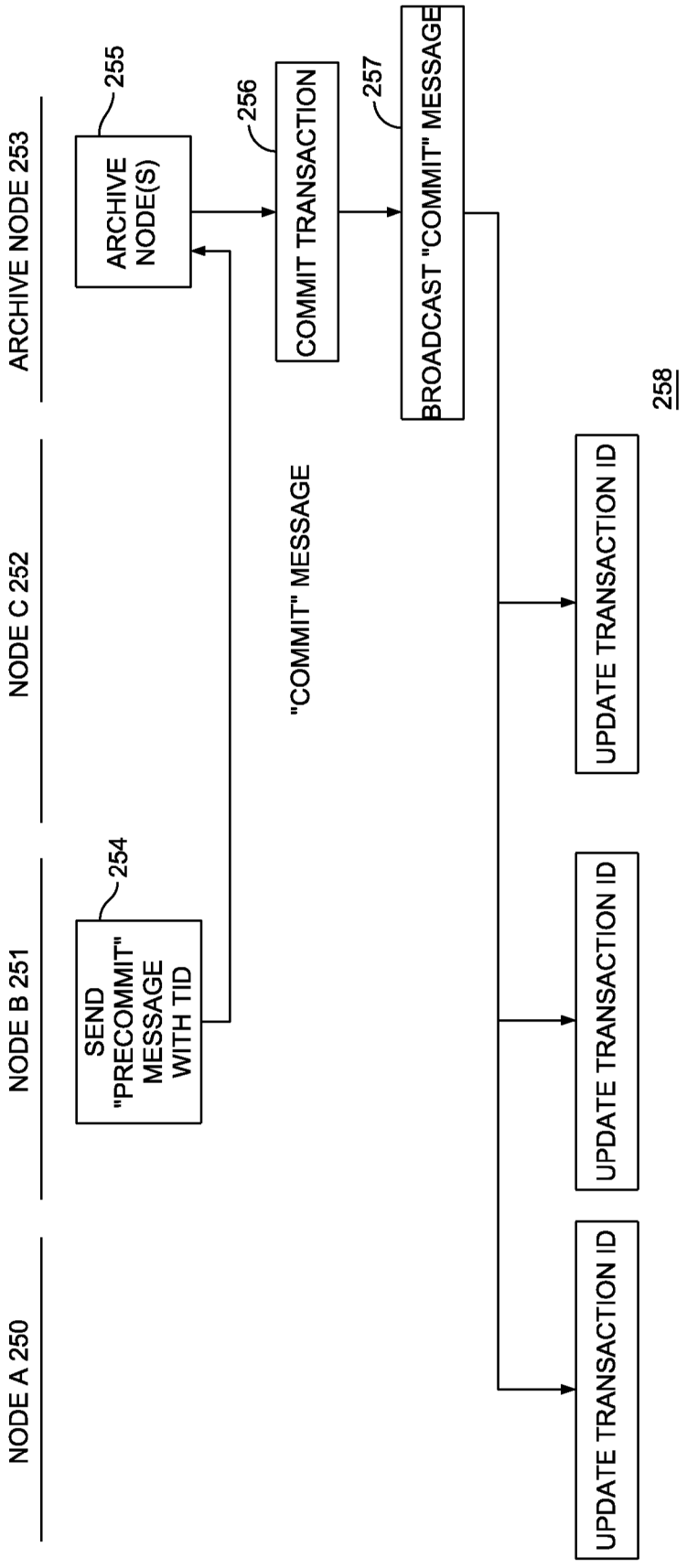
NODE A 250
NODE B 251
NODE C 252
ARCHIVE NODE 253
SEND "PRECOMMIT" MESSAGE WITH TID
254
ARCHIVE NODE(S)
255
COMMIT TRANSACTION
256
BROADCAST "COMMIT" MESSAGE
257
"COMMIT" MESSAGE
UPDATE TRANSACTION ID
UPDATE TRANSACTION ID
UPDATE TRANSACTION ID
258

FIG. 23