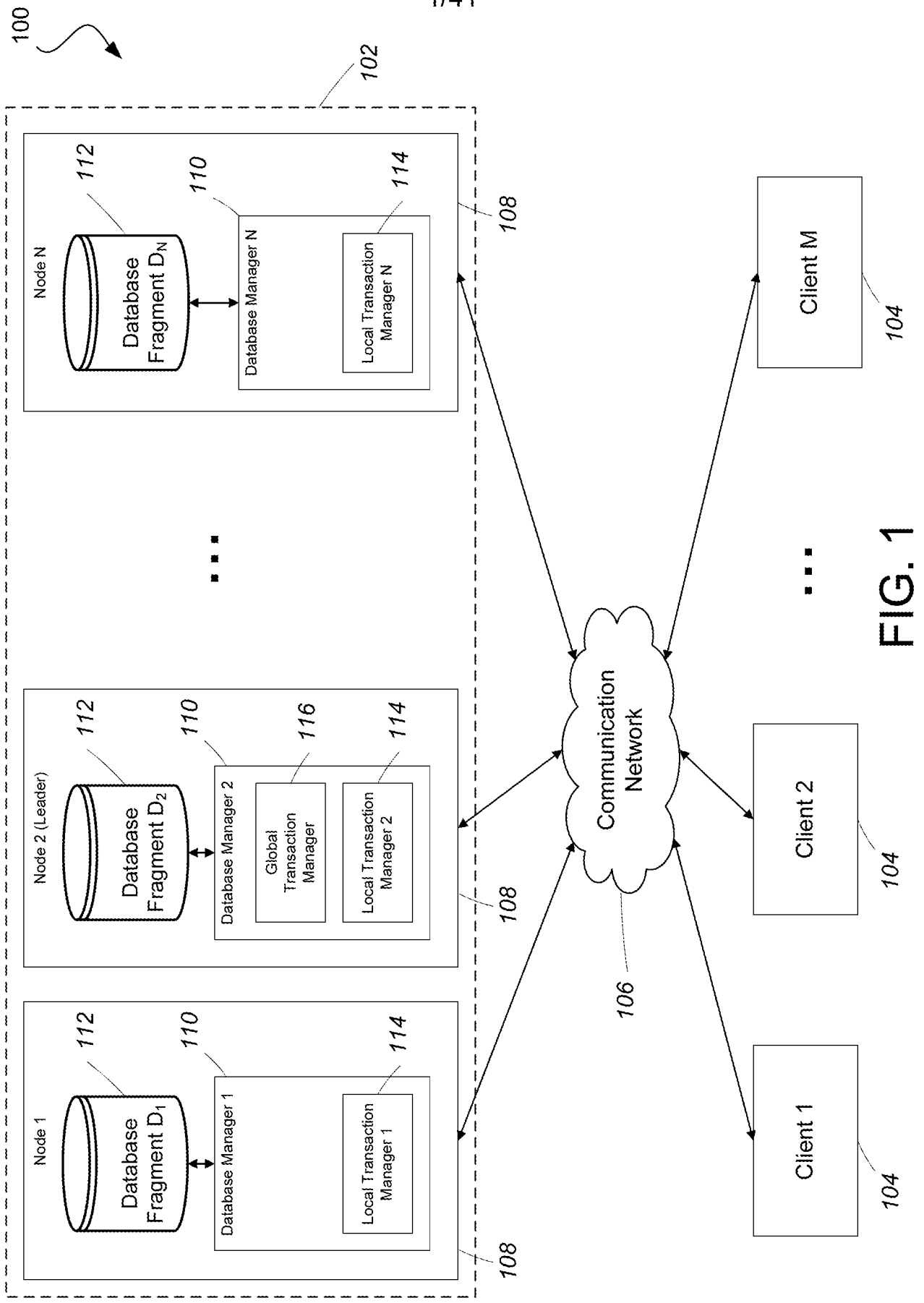
100
102
Node N
112
Database Fragment D_N
110
Database Manager N
114
Local Transaction Manager N
108
Node 2 (Leader)
112
Database Fragment D_2
110
Database Manager 2
116
Global Transaction Manager
114
Local Transaction Manager 2
108
Node 1
112
Database Fragment D_1
110
Database Manager 1
114
Local Transaction Manager 1
108
Communication Network
106
Client M
104
Client 2
104
Client 1
104

FIG. 1

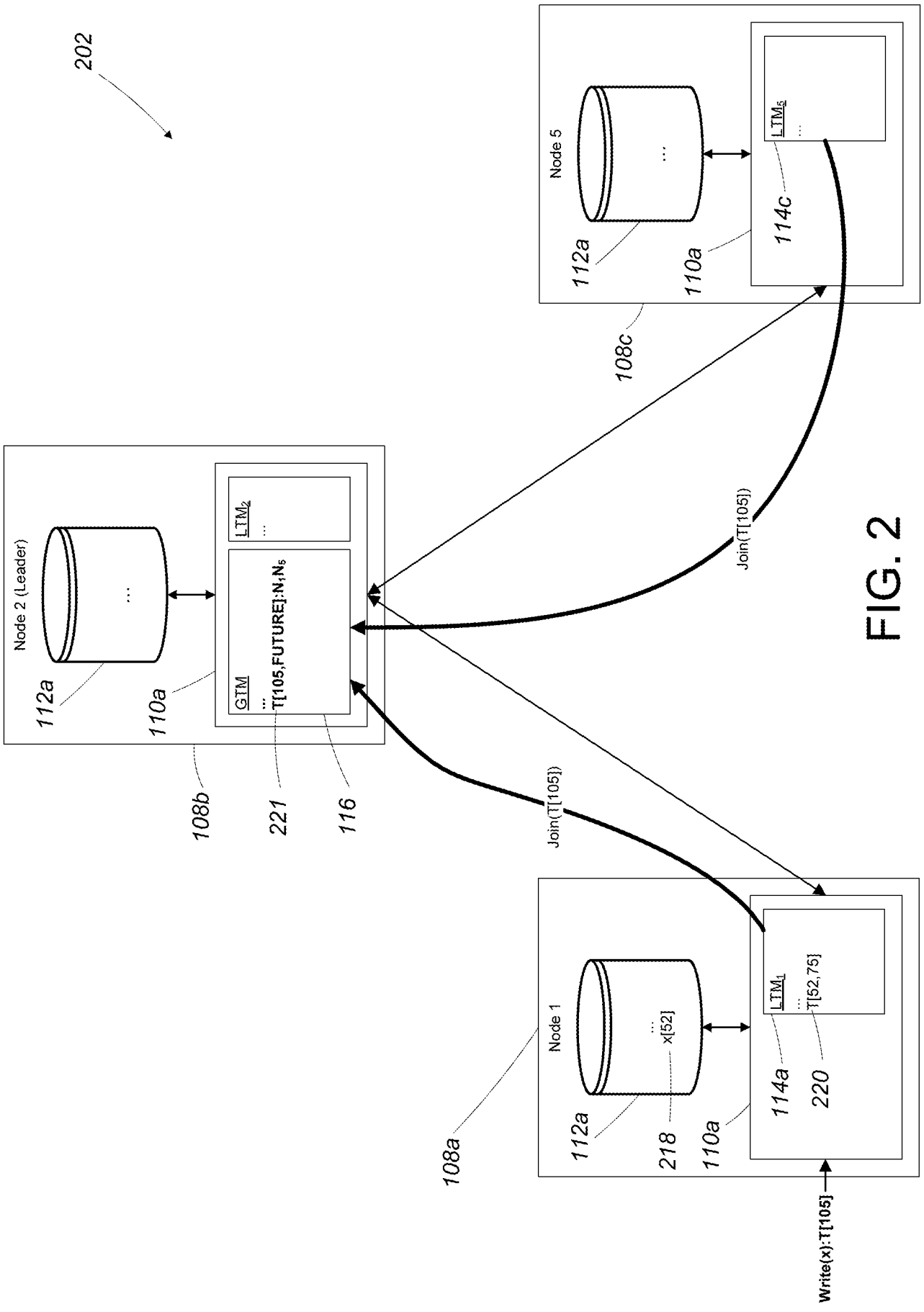
202
Node 2 (Leader)
112a
110a
GTM
T[105,FUTURE]:N1,N6
LTM2
108b
221
116
Node 5
112a
110a
114c
LTM5
108c
Join(T[105])
Join(T[105])
Node 1
112a
218
x[52]
110a
114a
LTM1
T[52,75]
220
108a
Write(x):T[105]

FIG. 2

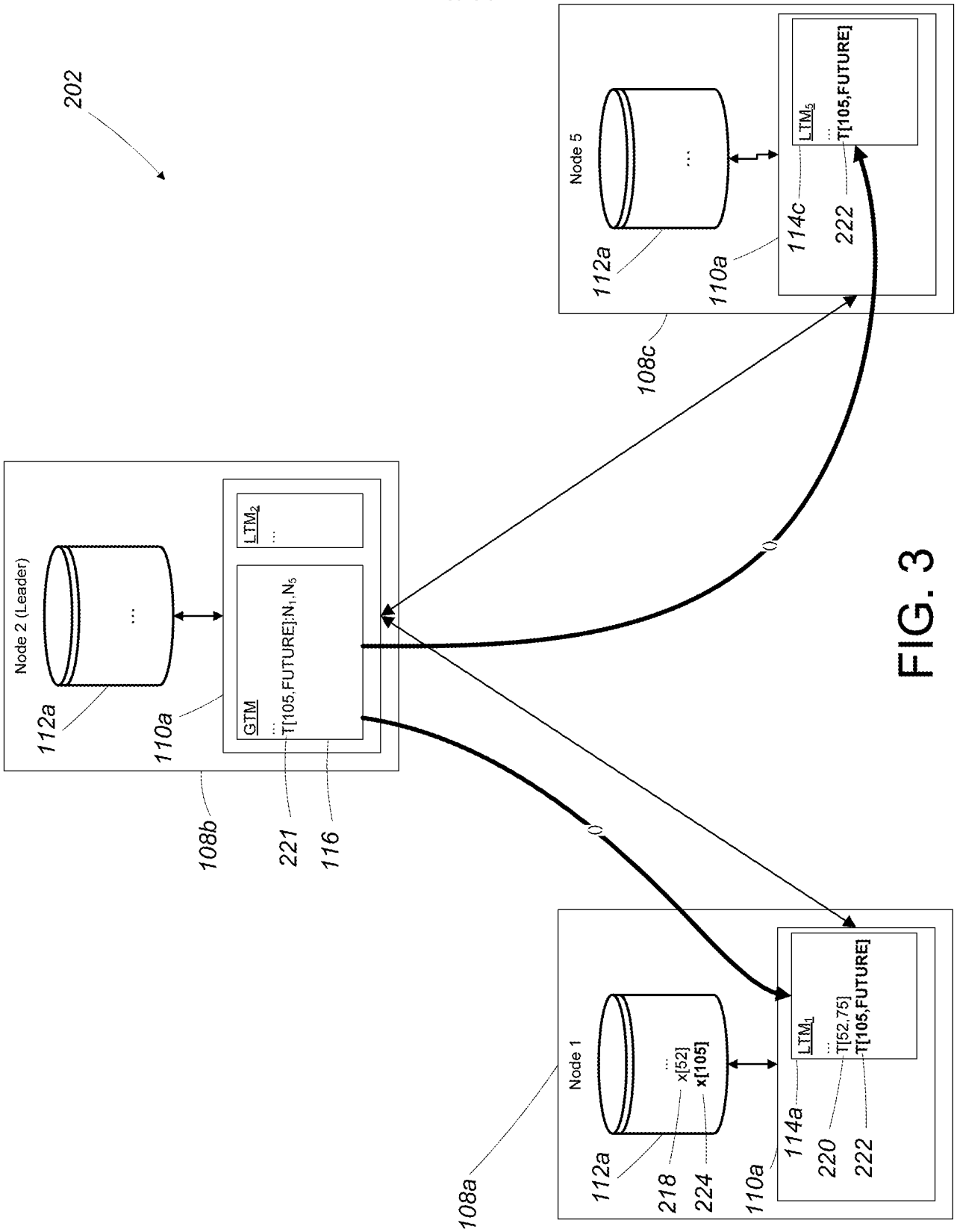
202
Node 2 (Leader)
112a
110a
GTM
T[105,FUTURE]:N1,N5
LTM2
108b
221
116
Node 5
112a
110a
114c
LTM2
T[105,FUTURE]
222
108c
Node 1
112a
218
x[52]
x[105]
224
110a
114a
LTM1
T[52,75]
T[105,FUTURE]
220
222
108a

FIG. 3

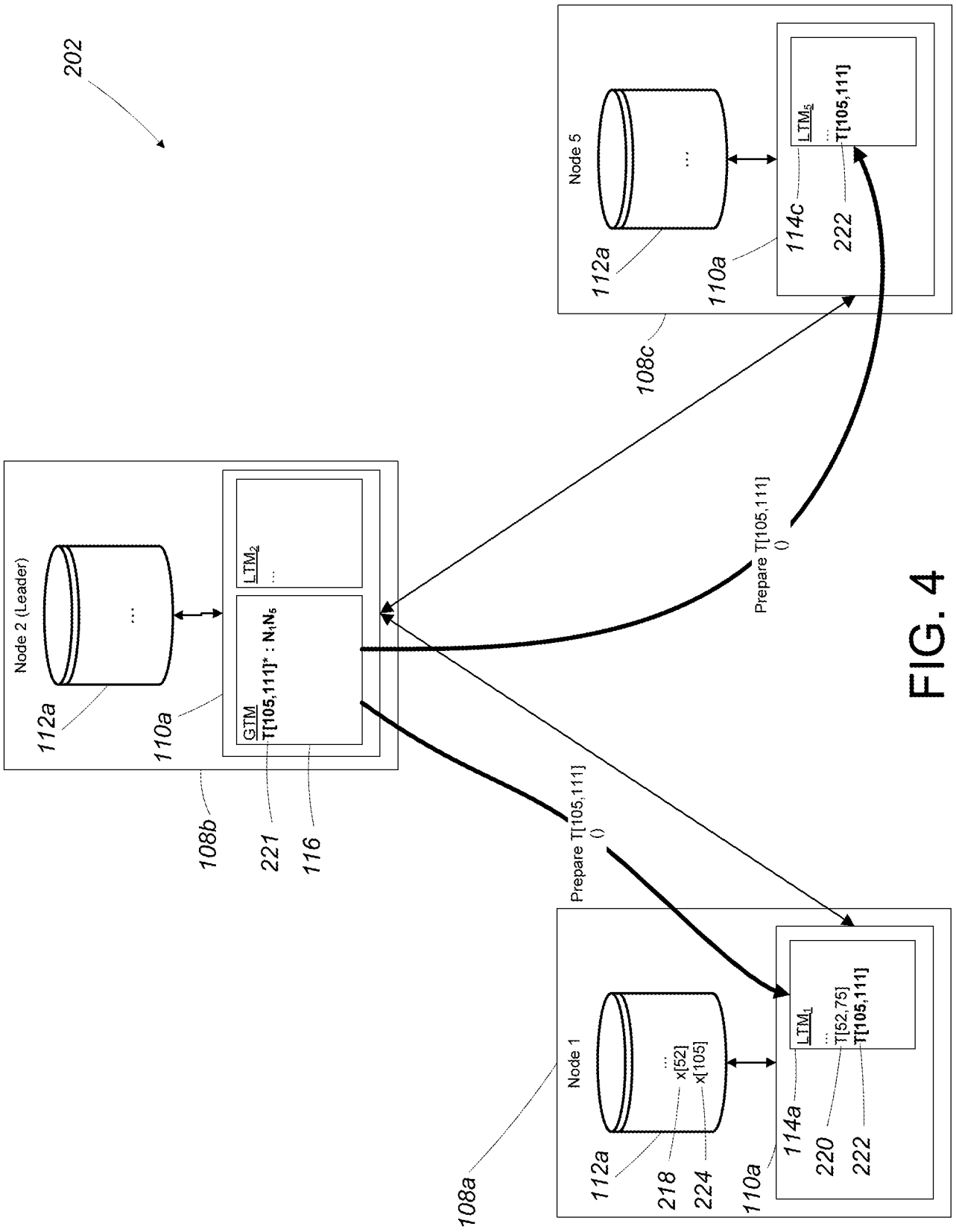
202
Node 2 (Leader)
112a
110a
GTM
T[105,111]* : N1N5
LTM2
108b
221
116
Node 5
112a
110a
114c
LTM5
T[105,111]
222
108c
Prepare T[105,111]
()
Prepare T[105,111]
()
Node 1
112a
218
x[52]
x[105]
224
110a
114a
LTM1
T[62,75]
T[105,111]
220
222
108a

FIG. 4

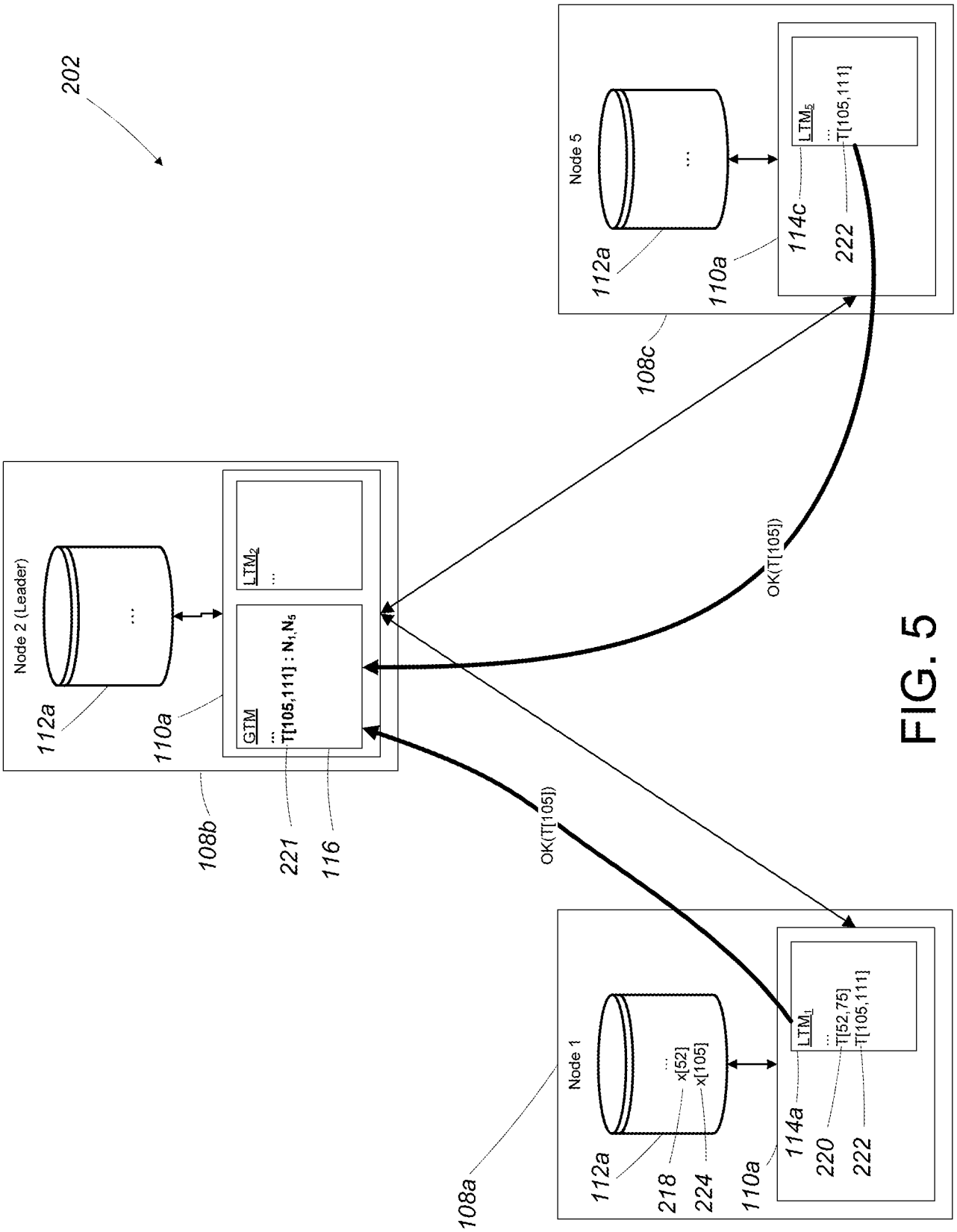
202
Node 2 (Leader)
112a
110a
GTM
T[105,111] : N1,N5
LTM2
108b
221
116
OK(T[105])
Node 5
112a
110a
LTM5
T[105,111]
114c
222
108c
OK(T[105])
Node 1
112a
218
x[52]
x[105]
224
110a
LTM1
T[52,75]
T[105,111]
114a
220
222
108a

FIG. 5

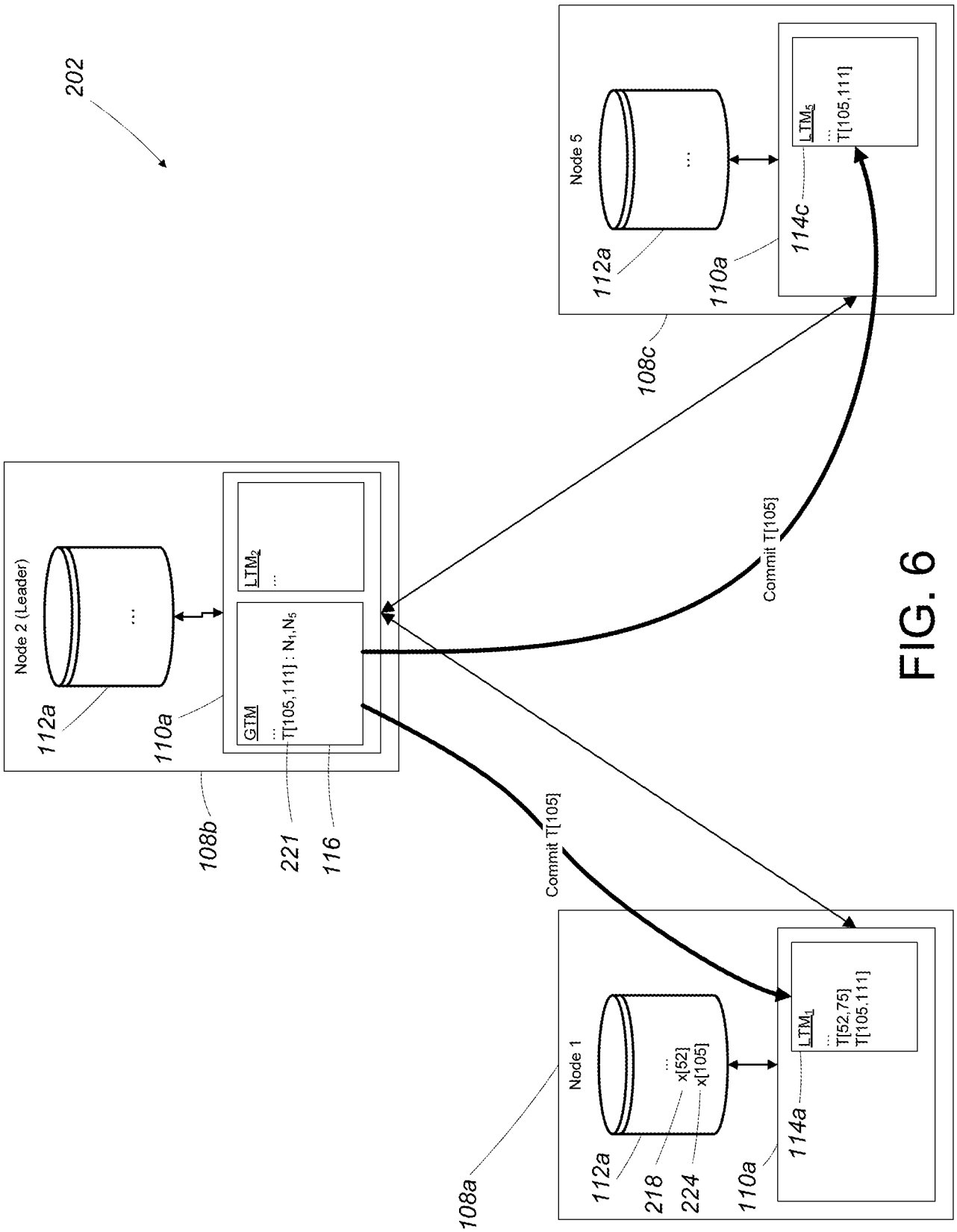
202
Node 2 (Leader)
112a
110a
GTM
T[105,111] : N1,N5
LTM2
108b
221
116
Commit T[105]
Node 5
112a
110a
114c
LTM5
T[105,111]
108c
Commit T[105]
Node 1
112a
218
224
x[52]
x[105]
110a
114a
LTM1
T[52,75]
T[105,111]
108a

FIG. 6

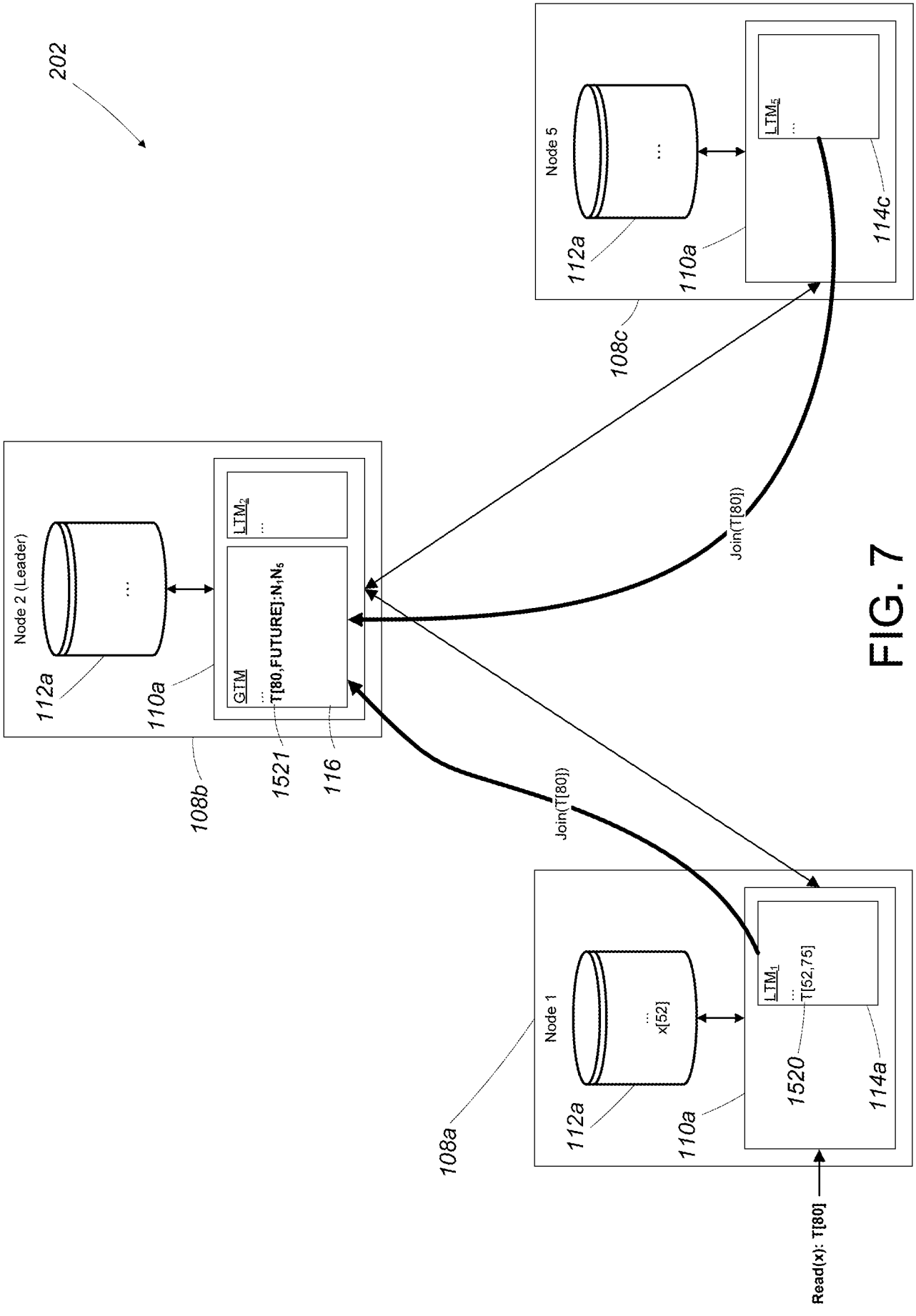
202
Node 2 (Leader)
112a
110a
GTM
...
T[80,FUTURE]:N1,N5
LTM2
...
108b
1521
116
Node 5
112a
110a
LTM5
...
114c
108c
Join(T[80])
Join(T[80])
Node 1
112a
110a
...
x[52]
LTM1
...
T[52,75]
1520
114a
108a
Read(x): T[80]

FIG. 7

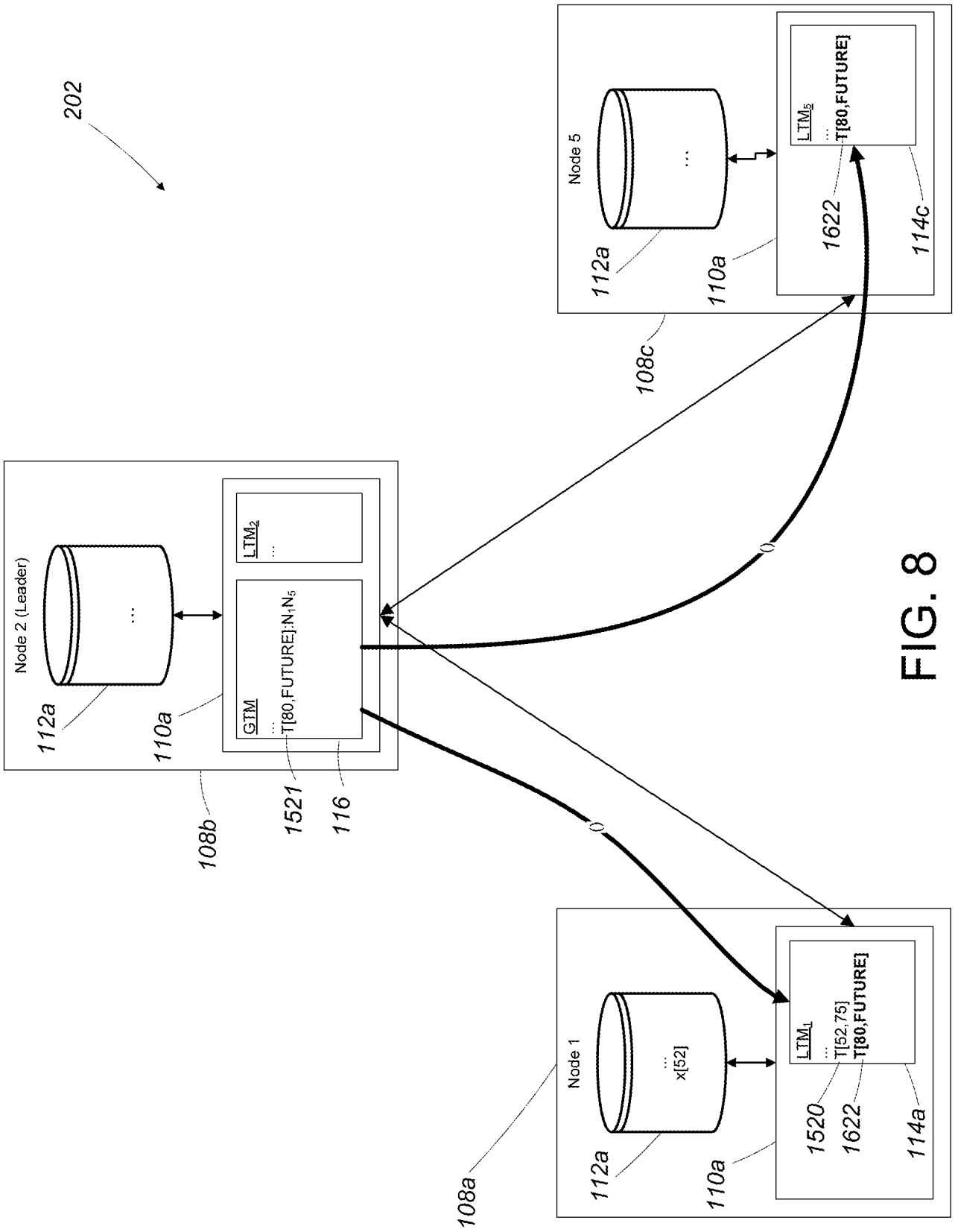
202
Node 2 (Leader)
112a
110a
GTM
T[80,FUTURE]:N1N5
LTM2
108b
1521
116
Node 5
112a
110a
LTM3
T[80,FUTURE]
1622
114c
108c
Node 1
112a
110a
x[52]
LTM1
T[52,75]
T[80,FUTURE]
1520
1622
114a
108a

FIG. 8

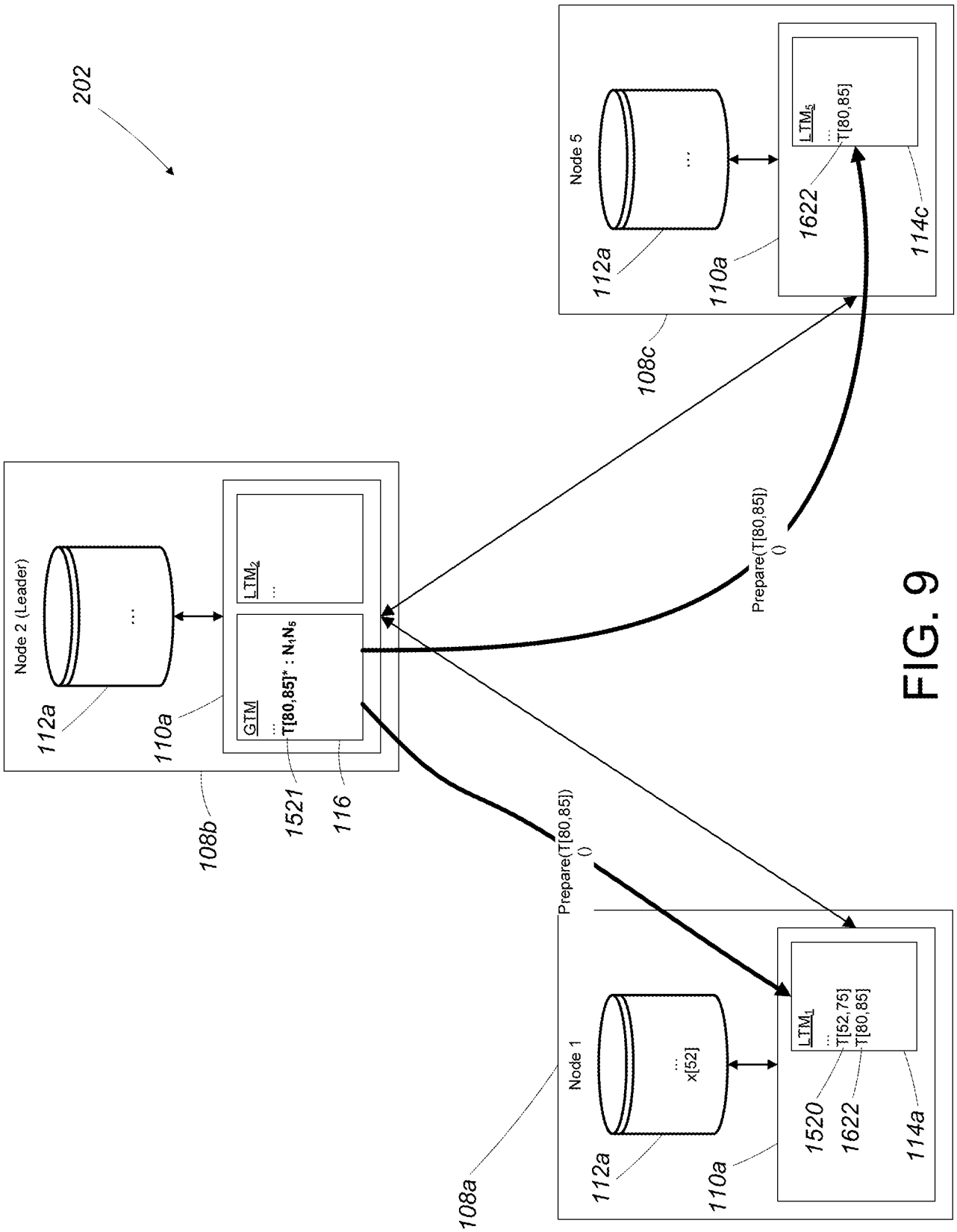
202
Node 2 (Leader)
112a
110a
GTM
T[80,85]* : N1N5
LTM2
108b
1521
116
Node 5
112a
110a
LTM5
T[80,85]
1622
114c
108c
Prepare(T[80,85])
()
Prepare(T[80,85])
()
Node 1
112a
110a
x[52]
LTM1
T[52,75]
T[80,85]
1520
1622
114a
108a

FIG. 9

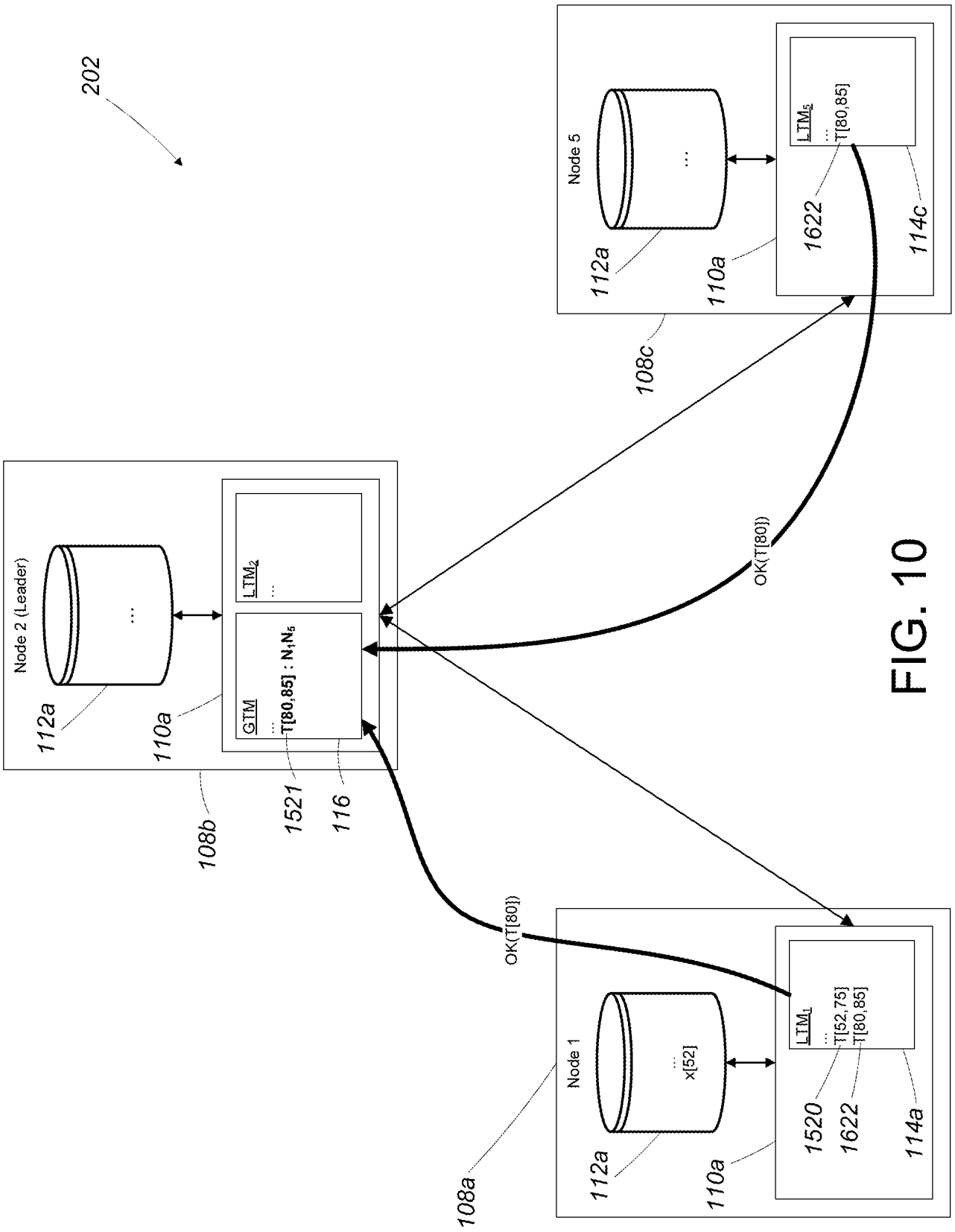
202
Node 2 (Leader)
112a
110a
GTM
T[80,85] : N1N5
LTM2
108b
1521
116
Node 5
112a
110a
LTM2
T[80,85]
1622
114c
108c
Node 1
112a
110a
x[52]
LTM1
T[52,75]
T[80,85]
1520
1622
114a
108a
OK(T[80])
OK(T[80])

FIG. 10

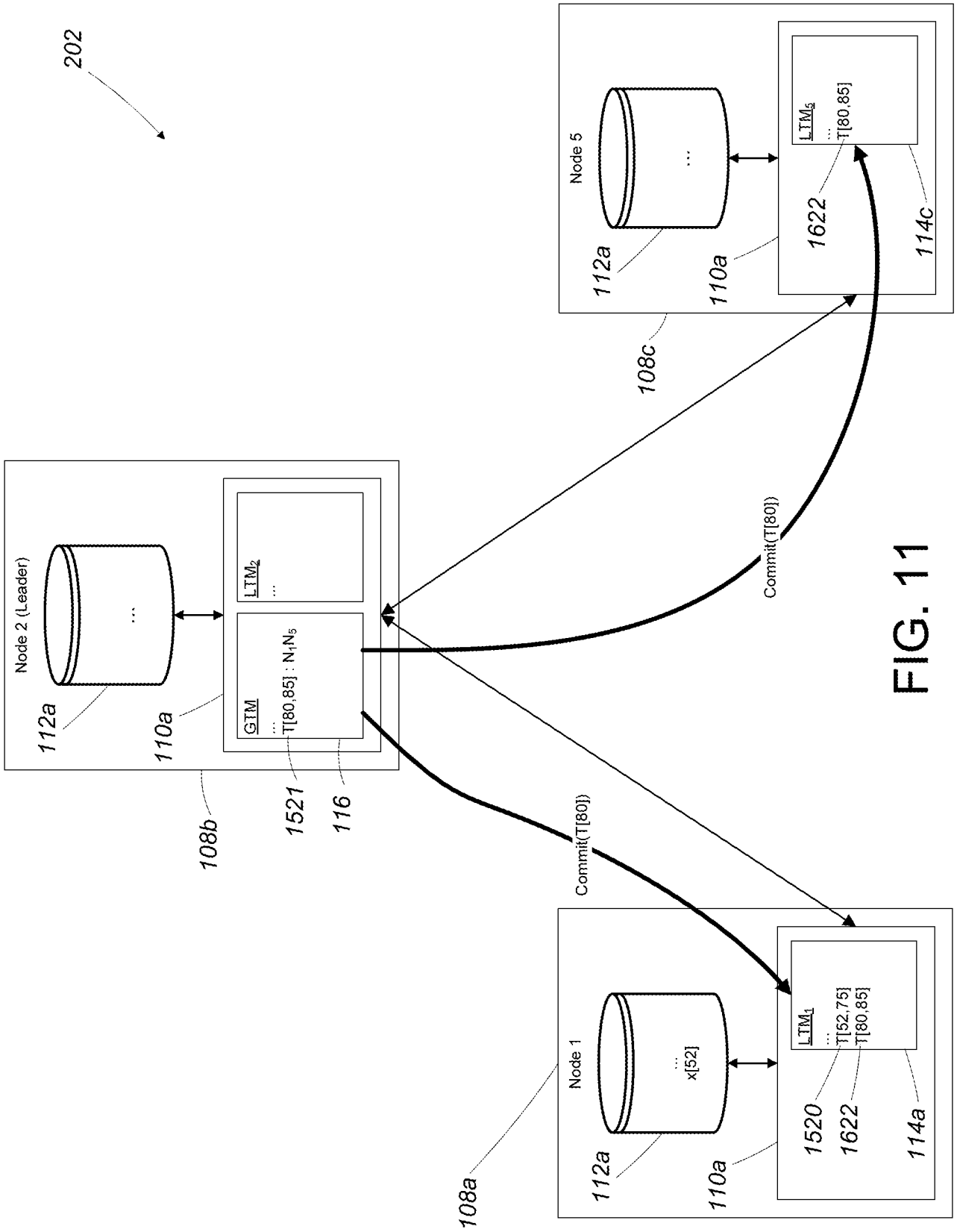
202
Node 2 (Leader)
112a
110a
GTM
T[80,85] : N1N5
LTM2
108b
1521
116
Node 5
112a
110a
LTM5
T[80,85]
1622
114c
108c
Commit(T[80])
Commit(T[80])
Node 1
112a
110a
x[52]
LTM1
T[52,75]
T[80,85]
1520
1622
114a
108a

FIG. 11

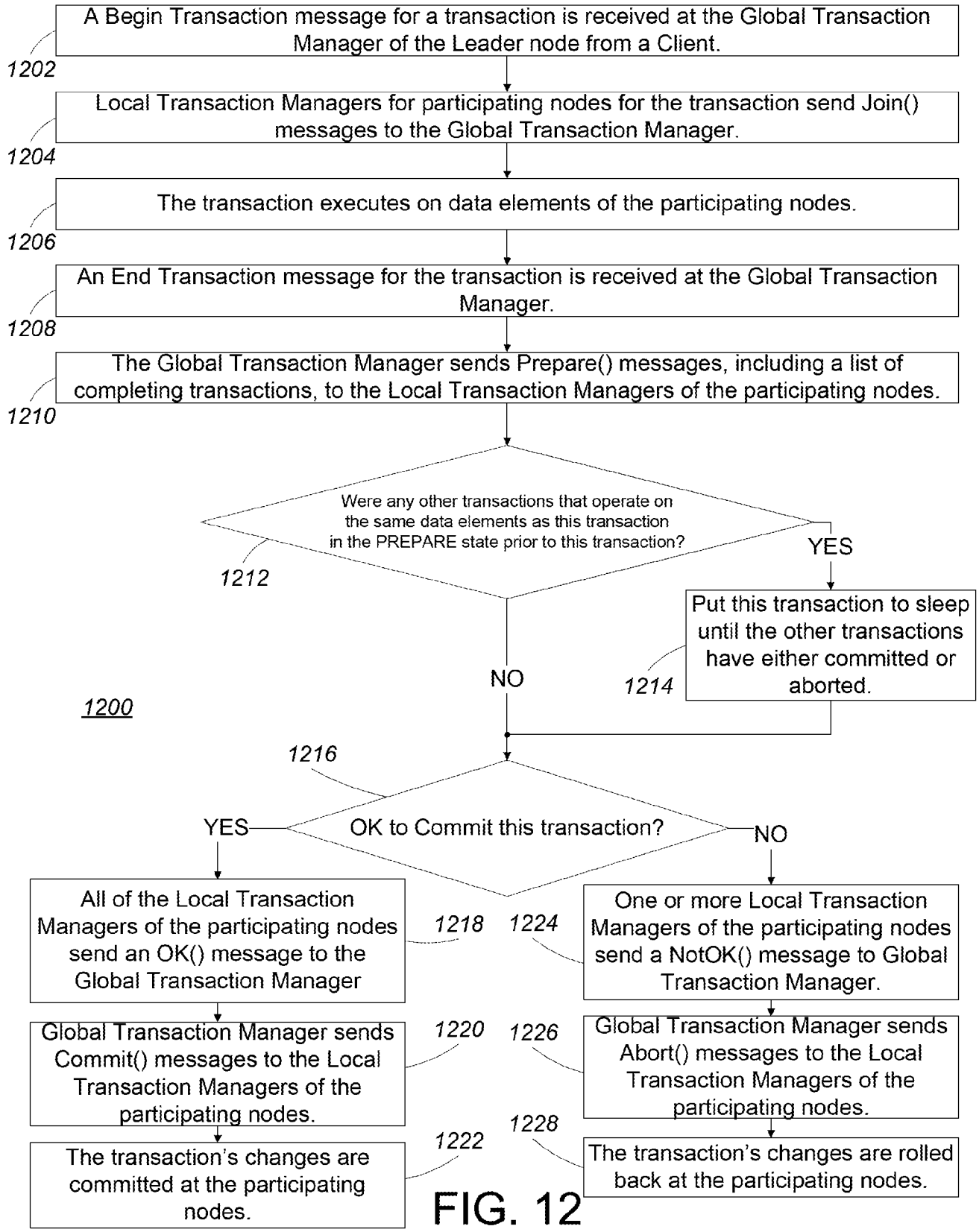
A Begin Transaction message for a transaction is received at the Global Transaction Manager of the Leader node from a Client.
1202
Local Transaction Managers for participating nodes for the transaction send Join() messages to the Global Transaction Manager.
1204
The transaction executes on data elements of the participating nodes.
1206
An End Transaction message for the transaction is received at the Global Transaction Manager.
1208
The Global Transaction Manager sends Prepare() messages, including a list of completing transactions, to the Local Transaction Managers of the participating nodes.
1210
Were any other transactions that operate on the same data elements as this transaction in the PREPARE state prior to this transaction?
1212
YES
Put this transaction to sleep until the other transactions have either committed or aborted.
1214
NO
1200
1216
OK to Commit this transaction?
YES
NO
All of the Local Transaction Managers of the participating nodes send an OK() message to the Global Transaction Manager
1218
Global Transaction Manager sends Commit() messages to the Local Transaction Managers of the participating nodes.
1220
The transaction's changes are committed at the participating nodes.
1222
One or more Local Transaction Managers of the participating nodes send a NotOK() message to Global Transaction Manager.
1224
Global Transaction Manager sends Abort() messages to the Local Transaction Managers of the participating nodes.
1226
The transaction's changes are rolled back at the participating nodes.
1228

FIG. 12

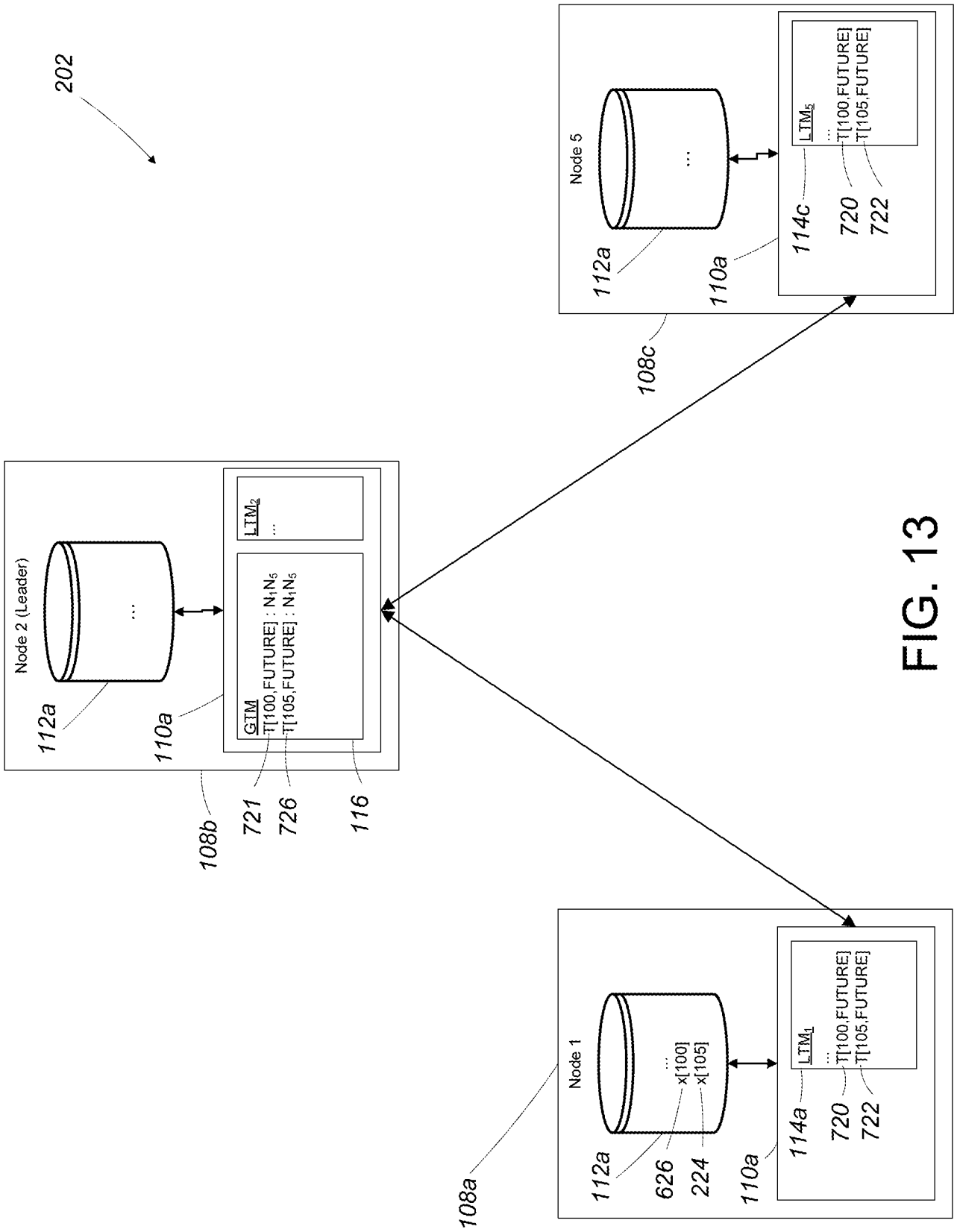
202
Node 2 (Leader)
112a
110a
GTM
T[100,FUTURE] : N1N5
T[105,FUTURE] : N1N5
LTM2
108b
721
726
116
Node 5
112a
110a
LTM5
T[100,FUTURE]
T[105,FUTURE]
114c
720
722
108c
Node 1
112a
626
224
110a
x[100]
x[105]
LTM1
T[100,FUTURE]
T[105,FUTURE]
114a
720
722
108a

FIG. 13

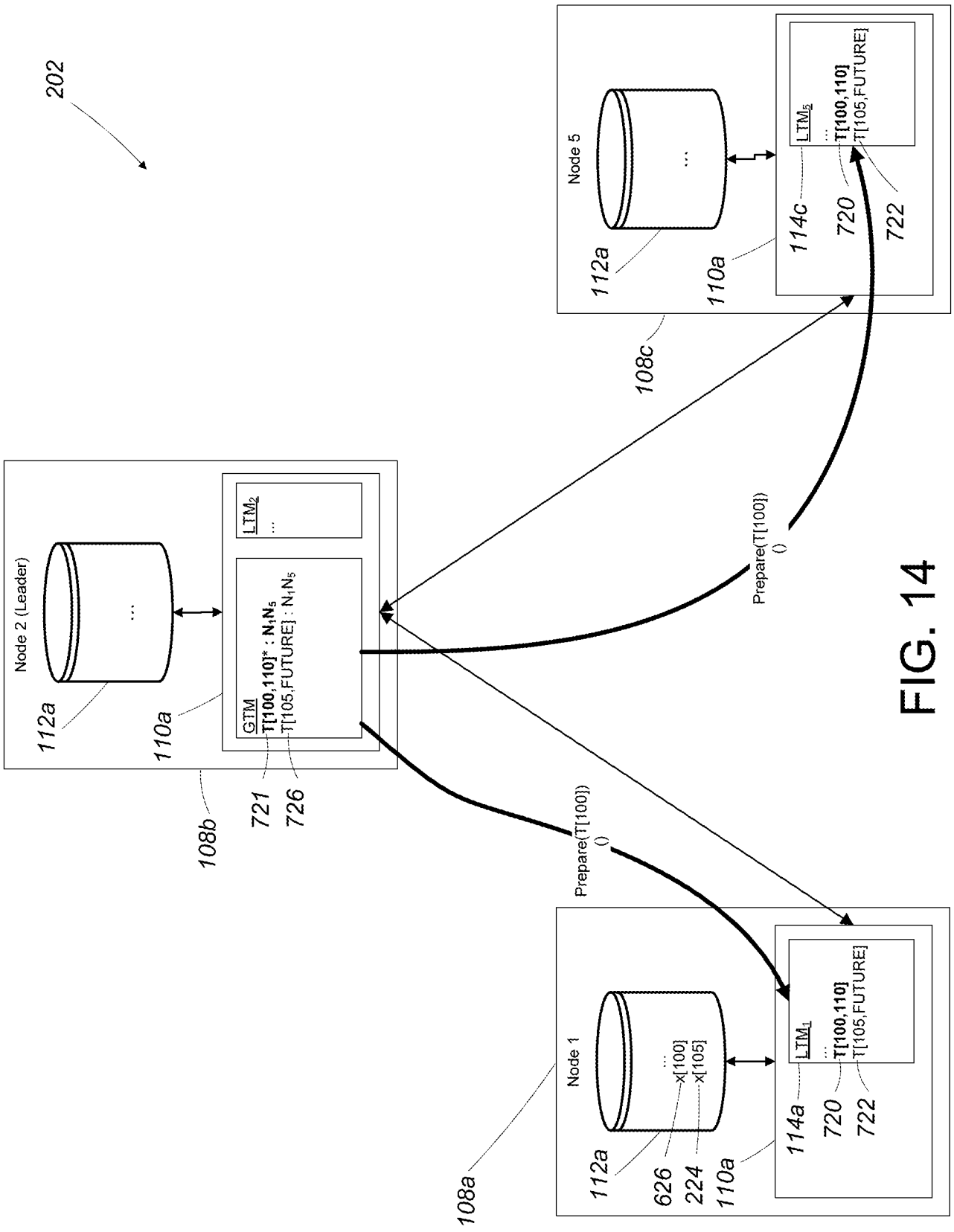
202
Node 2 (Leader)
112a
110a
GTM
T[100,110]* : N1N5
T[105,FUTURE] : N1N5
LTM2
...
108b
721
726
Node 5
112a
110a
114c
LTM2
T[100,110]
T[105,FUTURE]
720
722
108c
Node 1
112a
626
224
110a
114a
LTM1
x[100]
x[105]
T[100,110]
T[105,FUTURE]
720
722
108a
Prepare(T[100])
()
Prepare(T[100])
()

FIG. 14

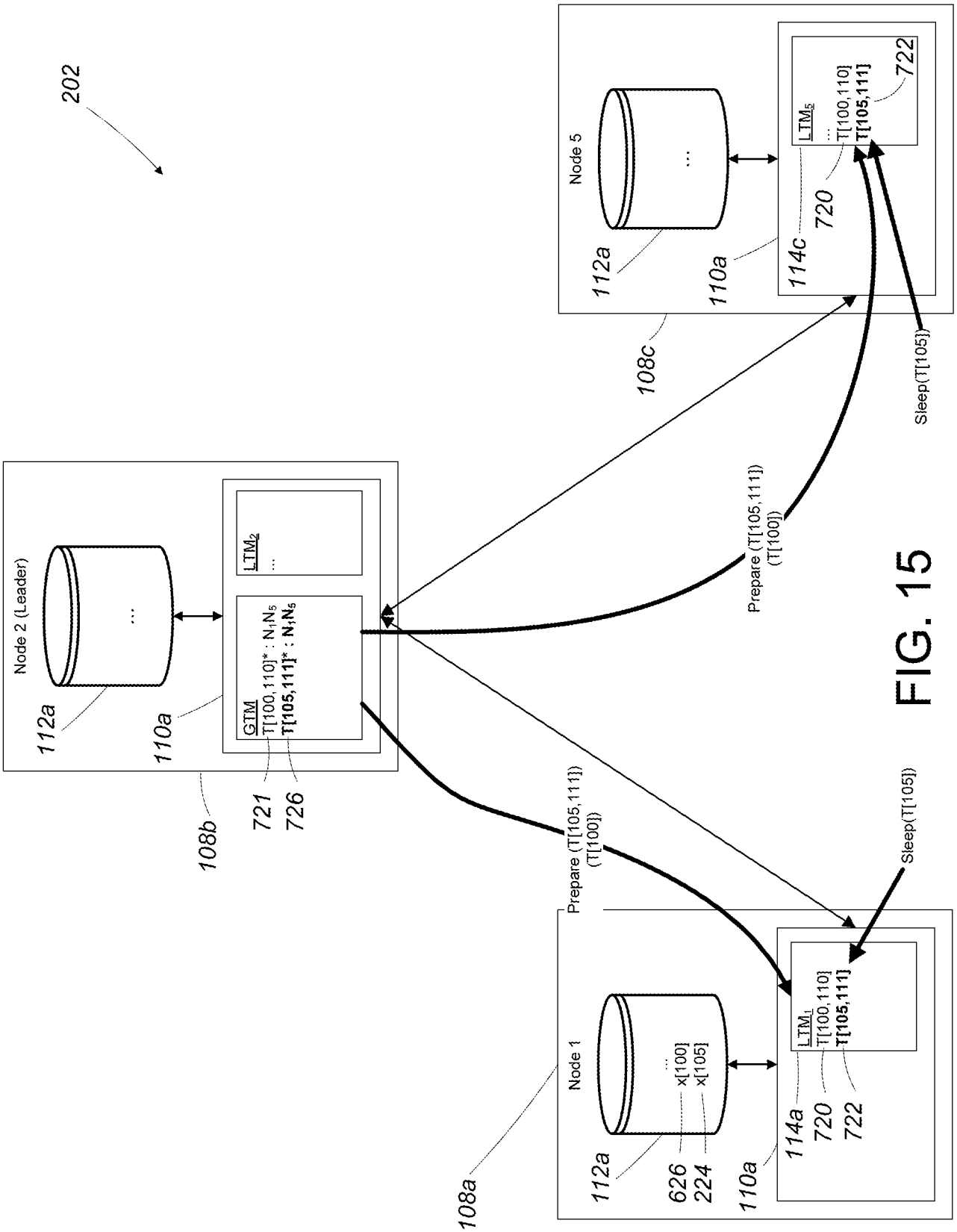
202
Node 2 (Leader)
112a
110a
GTM
T[100,110]* : N1N5
T[105,111]* : N1N5
LTM2
...
108b
721
726
Node 5
112a
110a
114c
LTM5
...
T[100,110]
T[105,111]
720
722
108c
Node 1
112a
626
224
x[100]
x[105]
110a
114a
LTM1
T[100,110]
T[105,111]
720
722
108a
Prepare (T[105,111])
(T[100])
Prepare (T[105,111])
(T[100])
Sleep(T[105])
Sleep(T[105])

FIG. 15

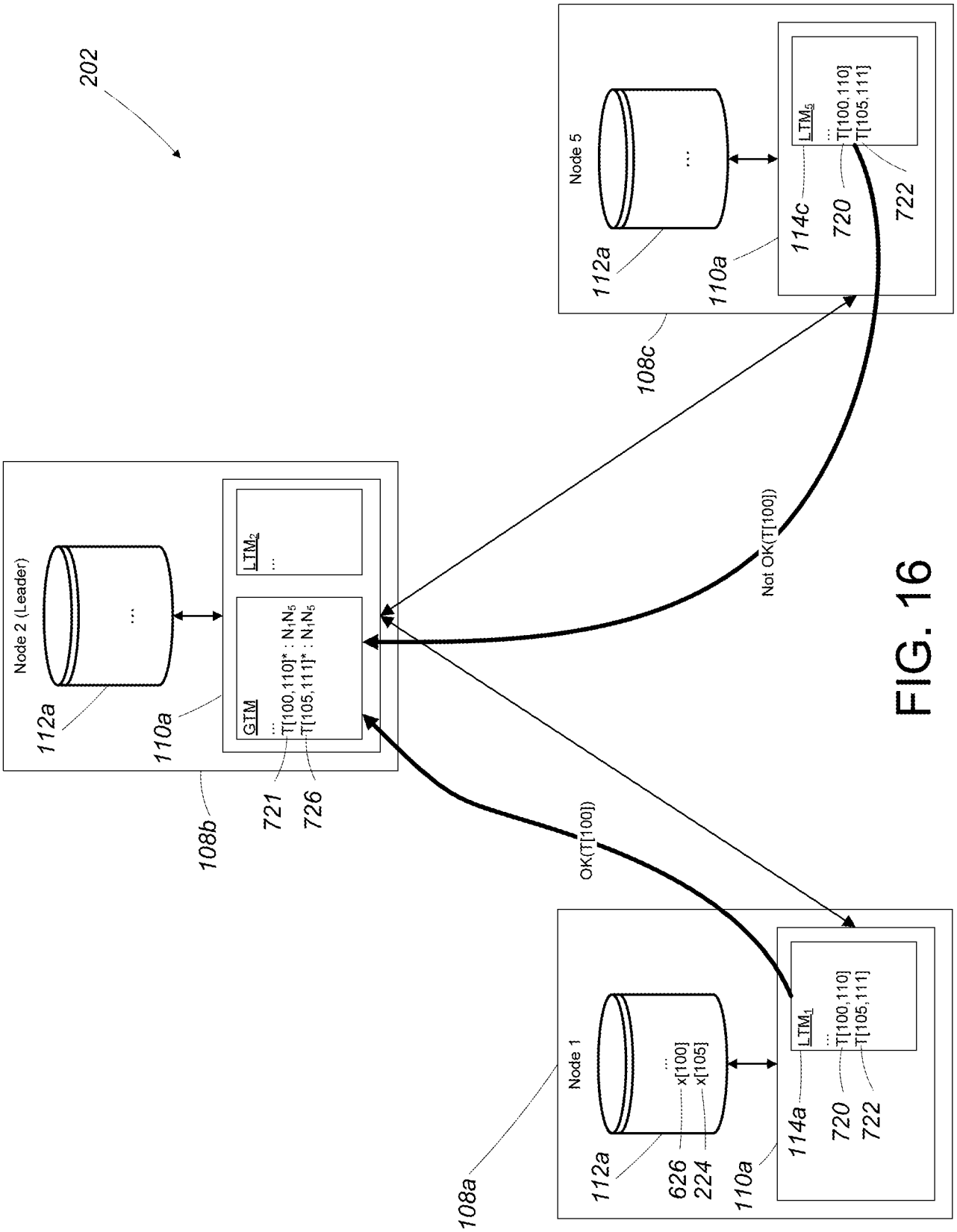
202
Node 2 (Leader)
112a
110a
GTM
LTM2
T[100,110]* : N1N5
T[105,111]* : N1N5
108b
721
726
Node 5
112a
110a
LTM5
T[100,110]
T[105,111]
114c
720
722
108c
Not OK(T[100])
OK(T[100])
Node 1
112a
x[100]
x[105]
626
224
110a
LTM1
T[100,110]
T[105,111]
114a
720
722
108a

FIG. 16

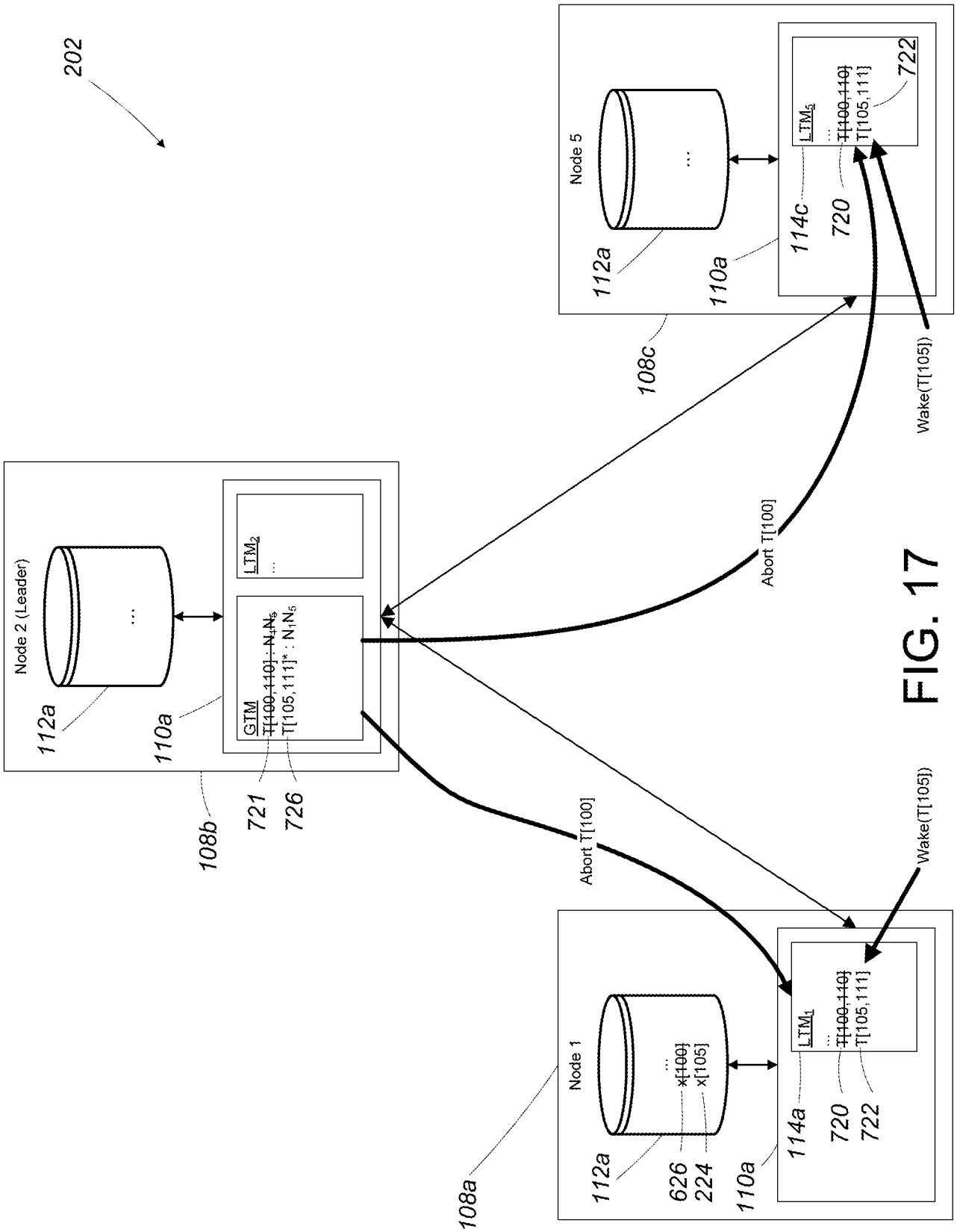
202
Node 2 (Leader)
112a
110a
GTM
T[100,110] : N1N5
T[105,111]* : N1N5
LTM2
108b
721
726
Node 5
112a
110a
LTM5
T[100,110]
T[105,111]
722
720
114c
108c
Abort T[100]
Wake(T[105])
Node 1
112a
626
224
110a
x[100]
x[105]
LTM1
T[100,110]
T[105,111]
114a
720
722
108a
Abort T[100]
Wake(T[105])

FIG. 17

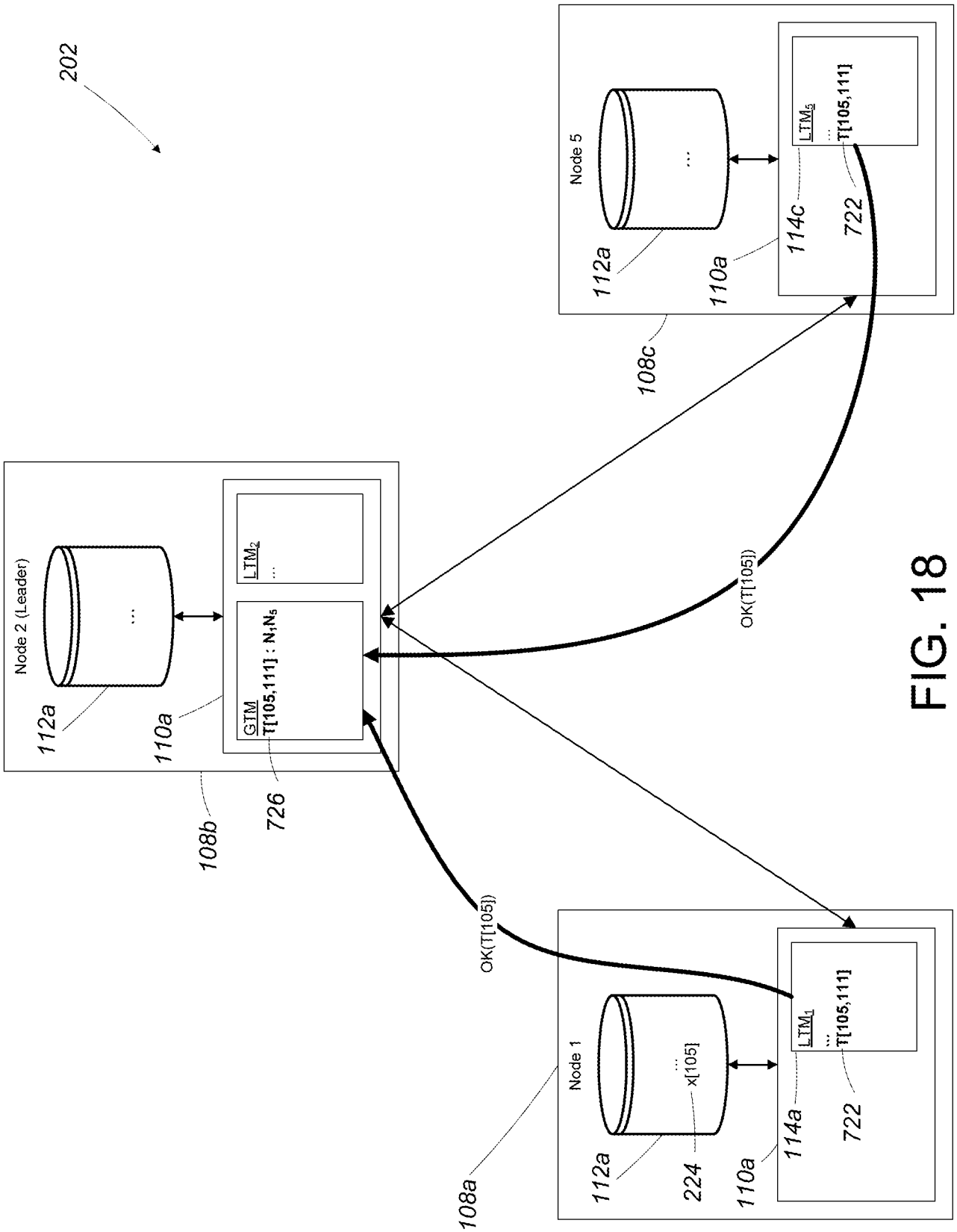
202
Node 2 (Leader)
112a
110a
GTM
T[105,111] : N1,N5
LTM2
726
108b
Node 5
112a
110a
114c
LTM3
T[105,111]
722
108c
Node 1
112a
224
x[105]
110a
114a
LTM1
T[105,111]
722
108a
OK(T[105])
OK(T[105])

FIG. 18

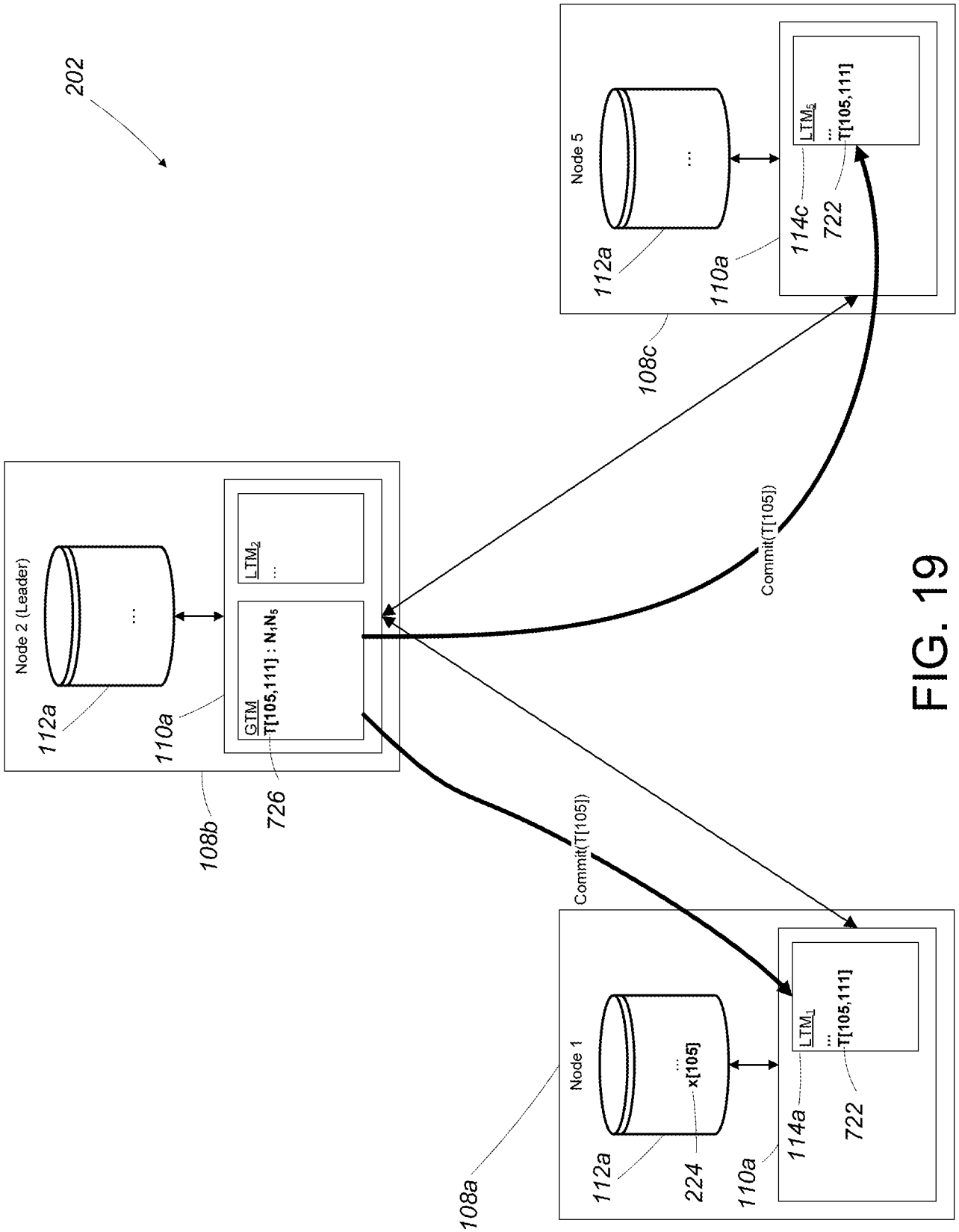
202
Node 2 (Leader)
112a
110a
GTM
T[105,111] : N1N5
LTM2
...
726
108b
Node 5
112a
110a
114c
LTM2
...
T[105,111]
722
108c
Commit(T[105])
Commit(T[105])
Node 1
112a
224
x[105]
110a
114a
LTM1
...
T[105,111]
722
108a

FIG. 19

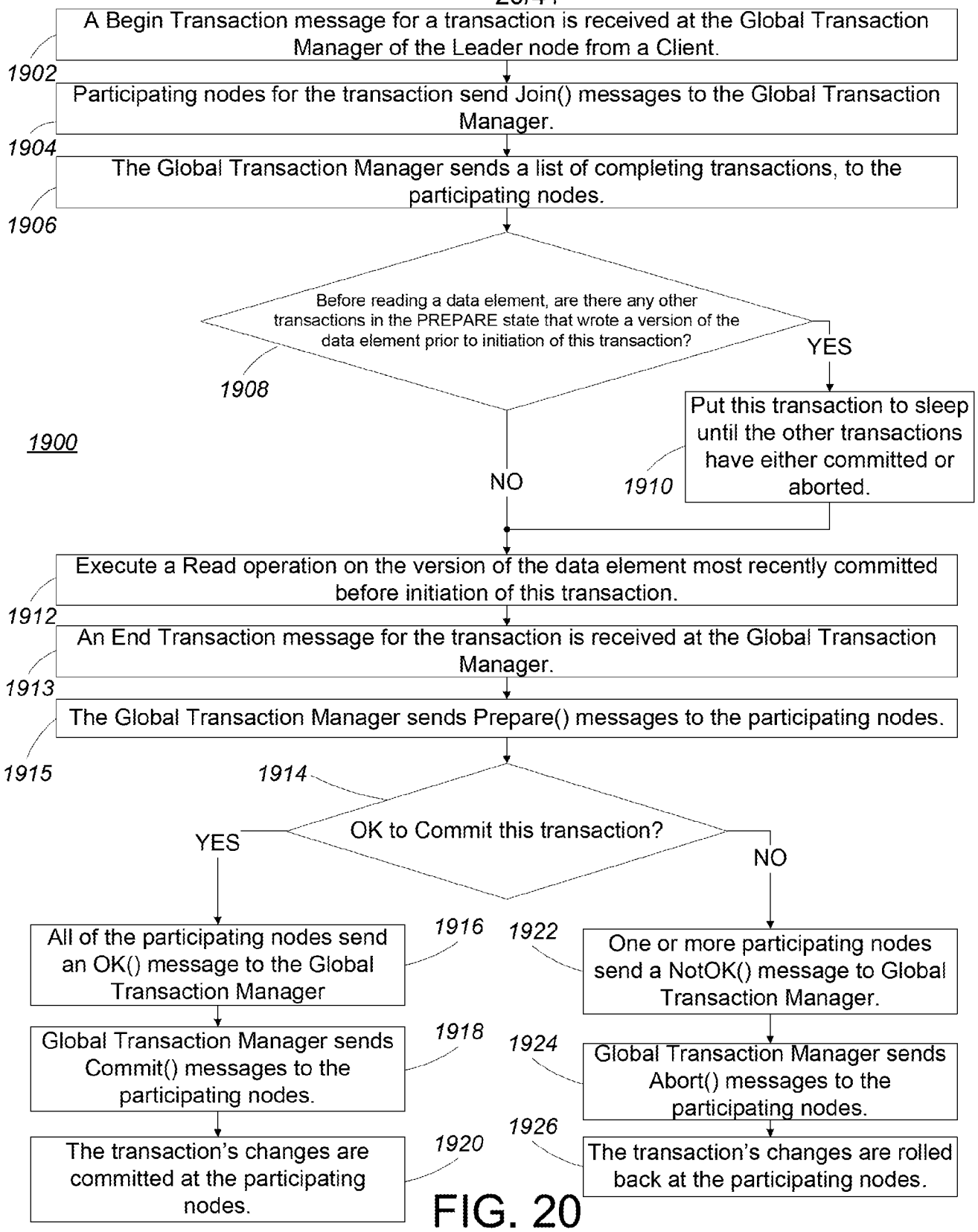
A Begin Transaction message for a transaction is received at the Global Transaction Manager of the Leader node from a Client.
1902
Participating nodes for the transaction send Join() messages to the Global Transaction Manager.
1904
The Global Transaction Manager sends a list of completing transactions, to the participating nodes.
1906
Before reading a data element, are there any other transactions in the PREPARE state that wrote a version of the data element prior to initiation of this transaction?
1908
YES
Put this transaction to sleep until the other transactions have either committed or aborted.
1910
1900
NO
Execute a Read operation on the version of the data element most recently committed before initiation of this transaction.
1912
An End Transaction message for the transaction is received at the Global Transaction Manager.
1913
The Global Transaction Manager sends Prepare() messages to the participating nodes.
1915
1914
OK to Commit this transaction?
YES
NO
All of the participating nodes send an OK() message to the Global Transaction Manager
1916
Global Transaction Manager sends Commit() messages to the participating nodes.
1918
The transaction's changes are committed at the participating nodes.
1920
1922
One or more participating nodes send a NotOK() message to Global Transaction Manager.
1924
Global Transaction Manager sends Abort() messages to the participating nodes.
1926
The transaction's changes are rolled back at the participating nodes.

FIG. 20

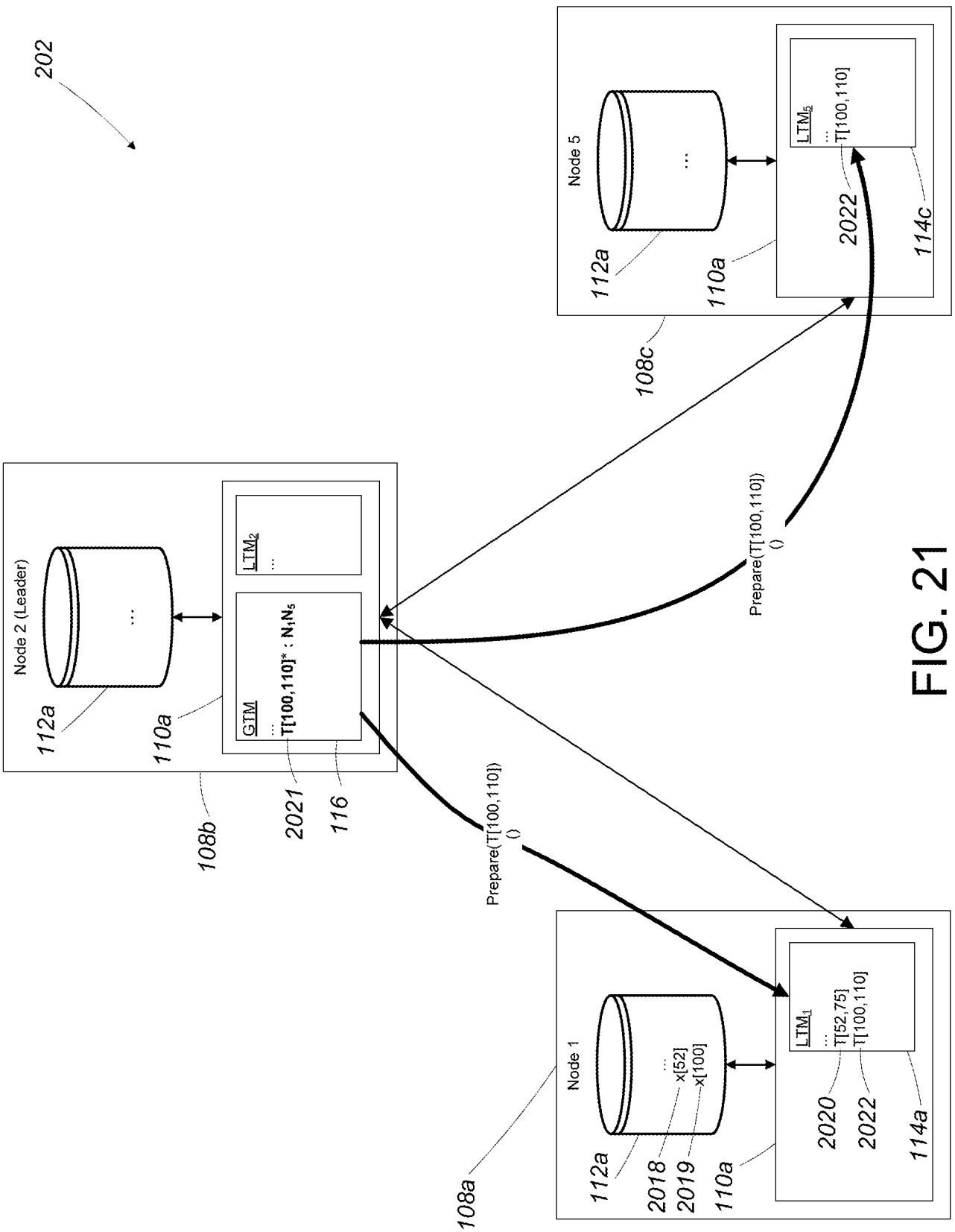
202
Node 2 (Leader)
112a
110a
GTM
T[100,110]* : N1,N5
LTM2
108b
2021
116
Node 5
112a
110a
LTM3
T[100,110]
2022
114c
108c
Prepare(T[100,110])
()
Prepare(T[100,110])
()
Node 1
112a
2018
2019
110a
x[52]
x[100]
LTM1
T[52,75]
T[100,110]
2020
2022
114a
108a

FIG. 21

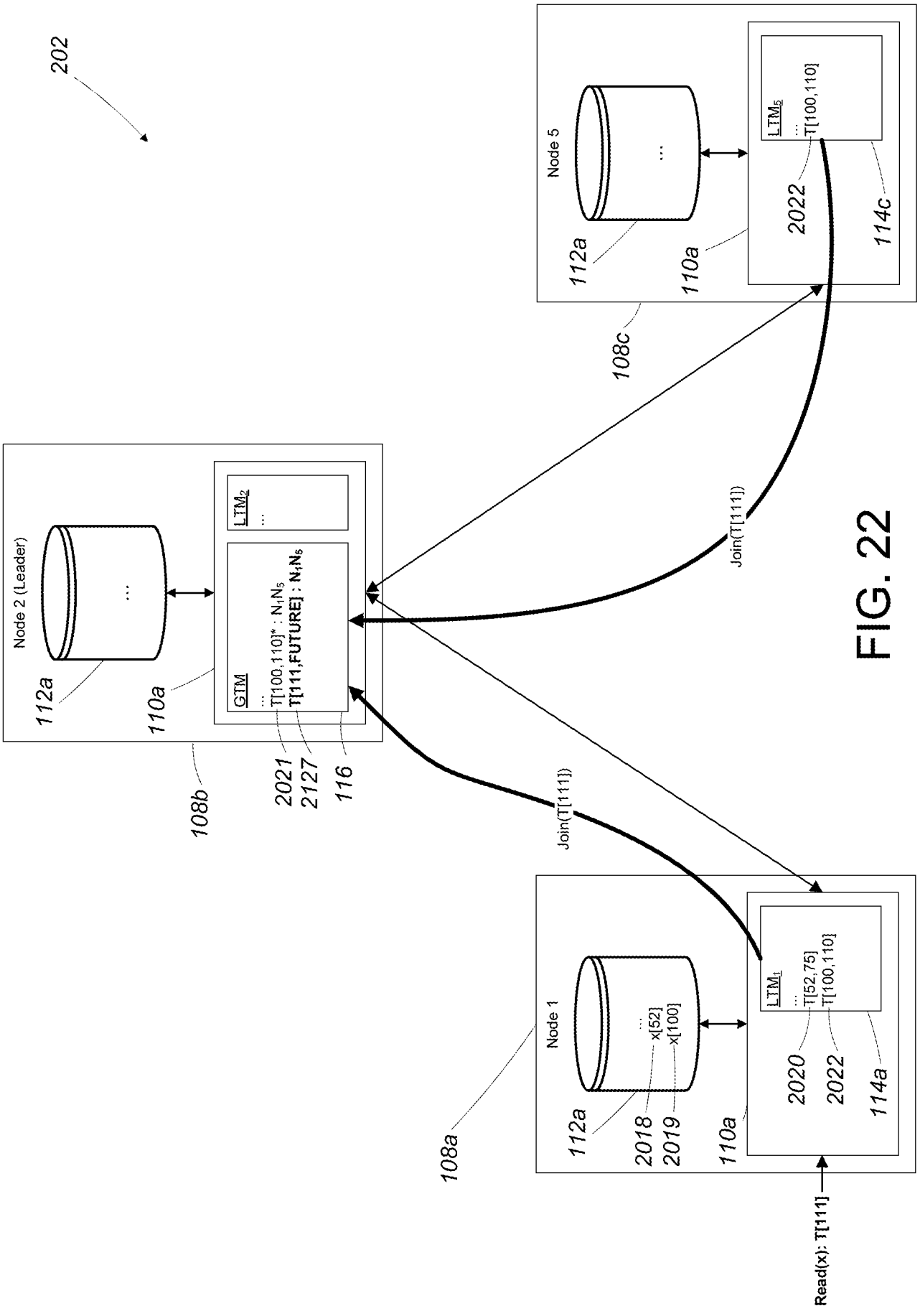
202
Node 2 (Leader)
112a
110a
GTM
...
T[100,110]* : N1N5
T[111,FUTURE] : N1N6
LTM2
...
108b
2021
2127
116
Node 5
112a
110a
LTM5
...
T[100,110]
2022
114c
108c
Join(T[111])
Join(T[111])
Node 1
112a
2018
2019
x[52]
x[100]
110a
LTM1
...
T[52,75]
T[100,110]
2020
2022
114a
108a
Read(x): T[111]

FIG. 22

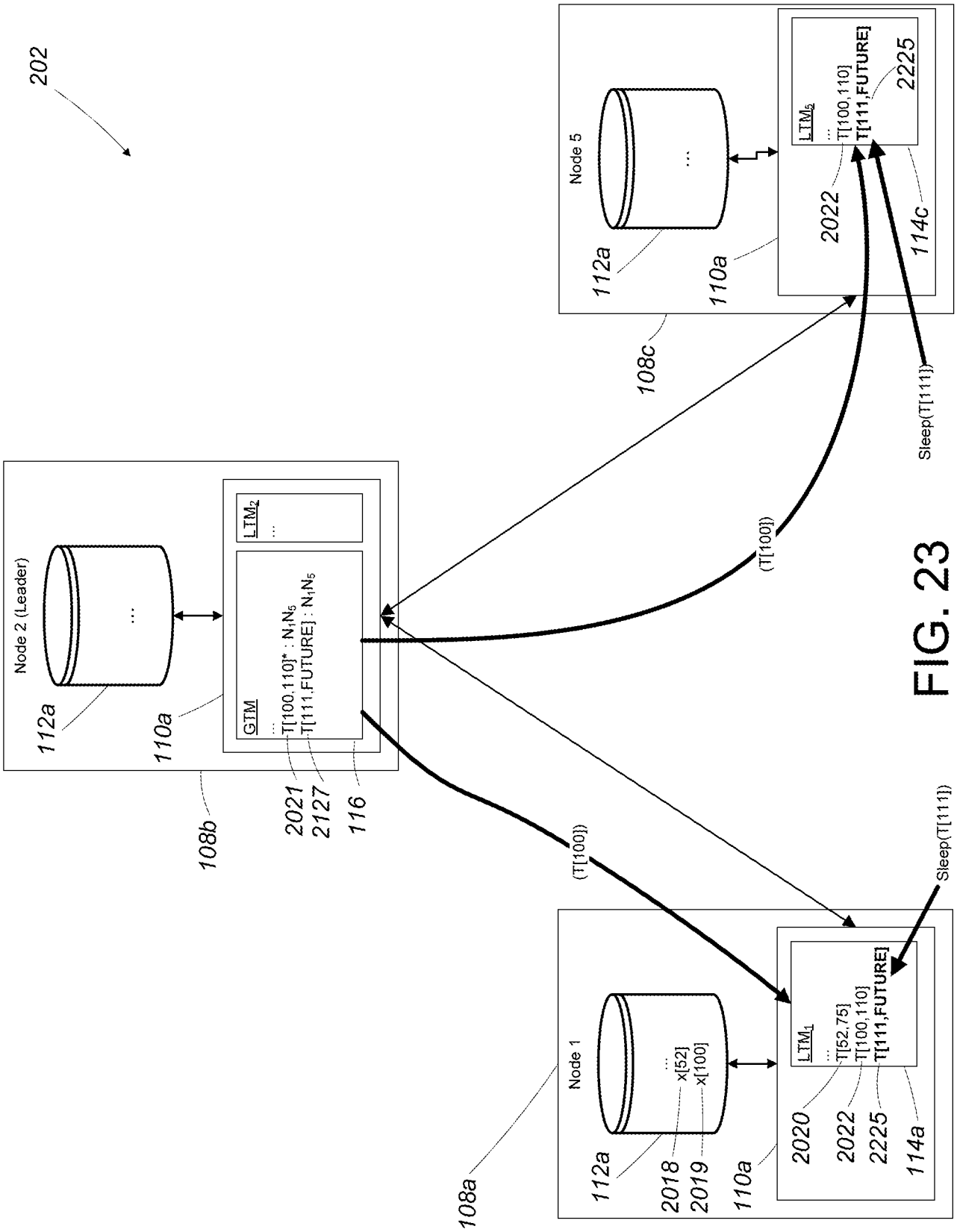
202
Node 2 (Leader)
112a
110a
GTM
LTM2
T[100,110]* : N1N5
T[111,FUTURE] : N1N5
108b
2021
2127
116
(T[100])
Node 5
112a
110a
LTM2
T[100,110]
T[111,FUTURE]
2225
2022
114c
108c
Sleep(T[111])
Node 1
112a
2018
2019
110a
LTM1
T[52,75]
T[100,110]
T[111,FUTURE]
2020
2022
2225
114a
108a
(T[100])
Sleep(T[111])

FIG. 23

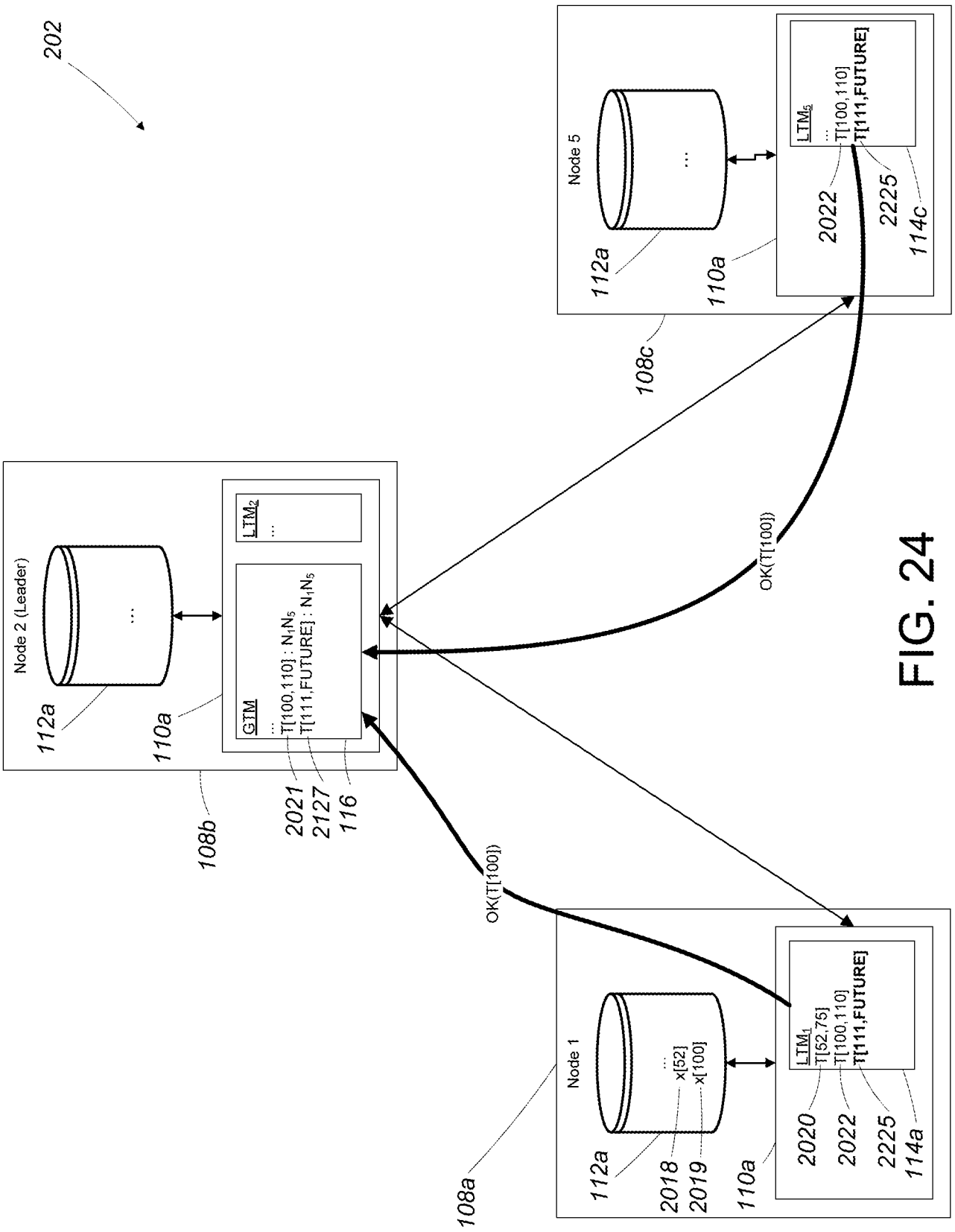
202
Node 2 (Leader)
112a
110a
GTM
T[100,110] : N1N5
T[111,FUTURE] : N1N5
LTM2
108b
2021
2127
116
Node 5
112a
110a
LTM5
T[100,110]
T[111,FUTURE]
2022
2225
114c
108c
Node 1
112a
2018
2019
x[52]
x[100]
110a
LTM1
T[52,75]
T[100,110]
T[111,FUTURE]
2020
2022
2225
114a
108a
OK(T[100])
OK(T[100])

FIG. 24

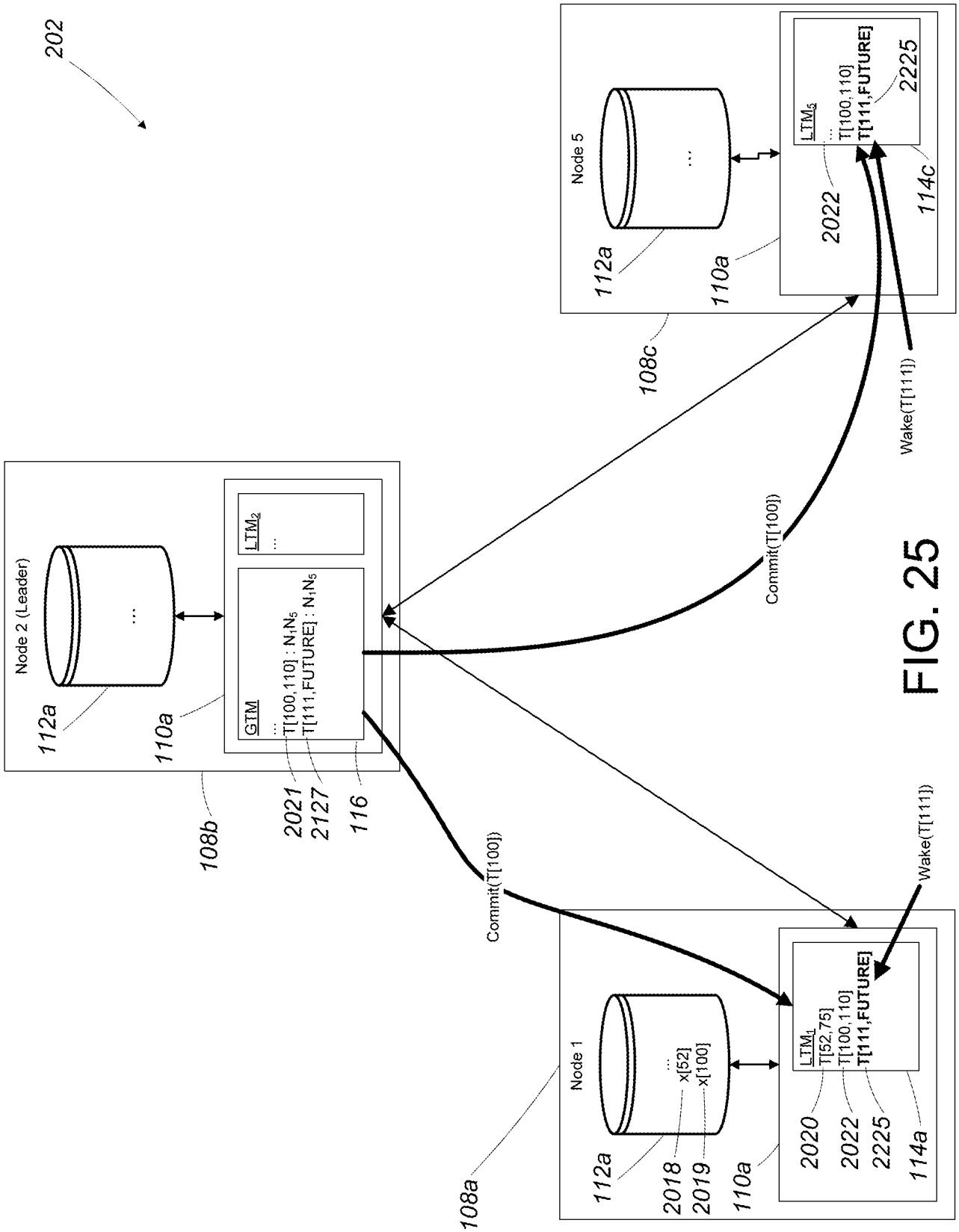
202
Node 2 (Leader)
112a
110a
GTM
LTM2
T[100,110] : N1N5
T[111,FUTURE] : N1N5
108b
2021
2127
116
Node 5
112a
110a
LTM5
T[100,110]
T[111,FUTURE]
2225
2022
114c
108c
Node 1
112a
2018
2019
x[52]
x[100]
110a
LTM1
T[52,75]
T[100,110]
T[111,FUTURE]
2020
2022
2225
114a
108a
Commit(T[100])
Wake(T[111])
Commit(T[100])
Wake(T[111])

FIG. 25

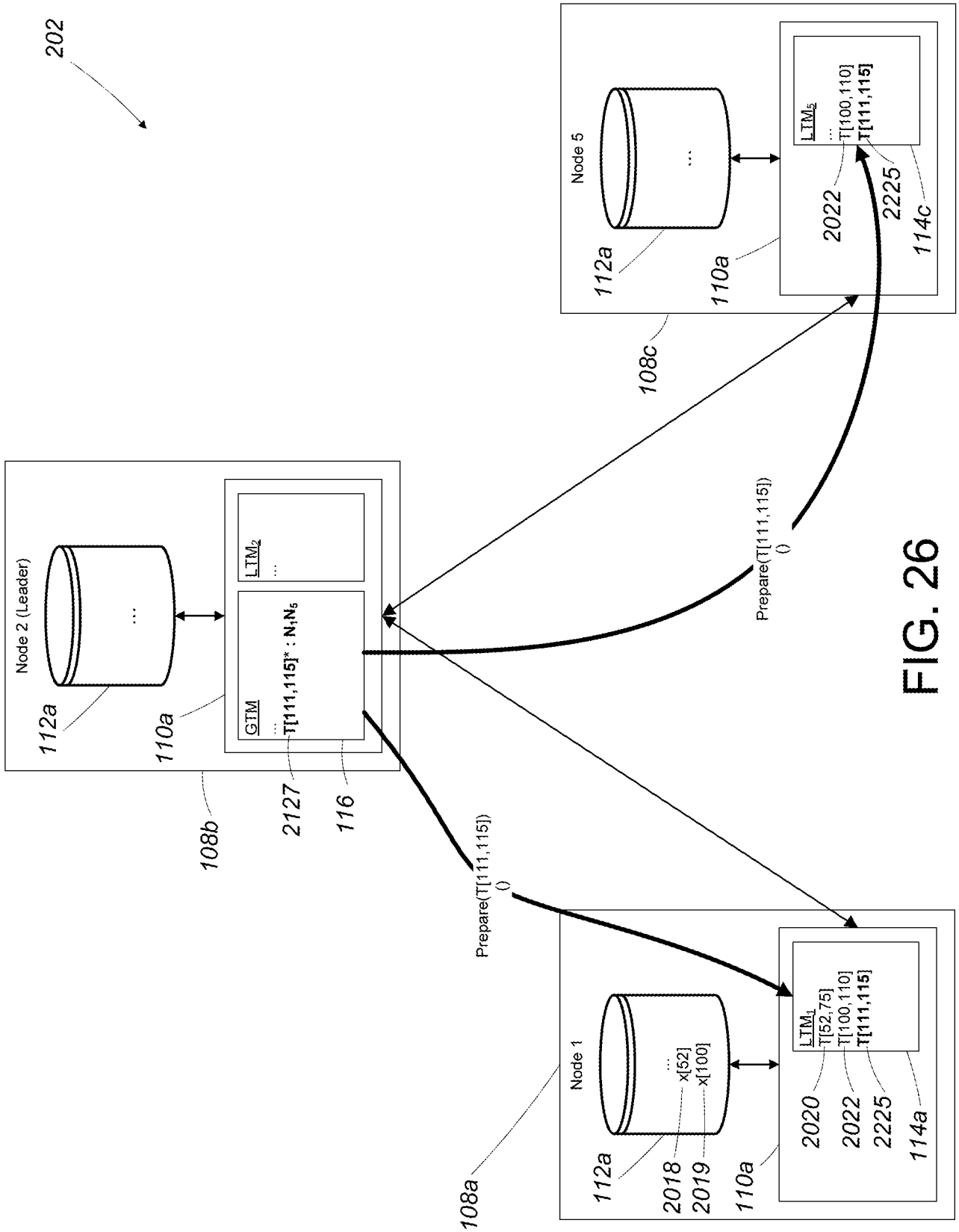
202
Node 2 (Leader)
112a
110a
GTM
T[111,115]* : N1,N5
LTM2
108b
2127
116
Prepare(T[111,115])
()
Node 5
112a
110a
LTM3
T[100,110]
T[111,115]
2022
2225
114c
108c
Prepare(T[111,115])
()
Node 1
112a
2018
2019
x[52]
x[100]
110a
LTM1
T[52,75]
T[100,110]
T[111,115]
2020
2022
2225
114a
108a

FIG. 26

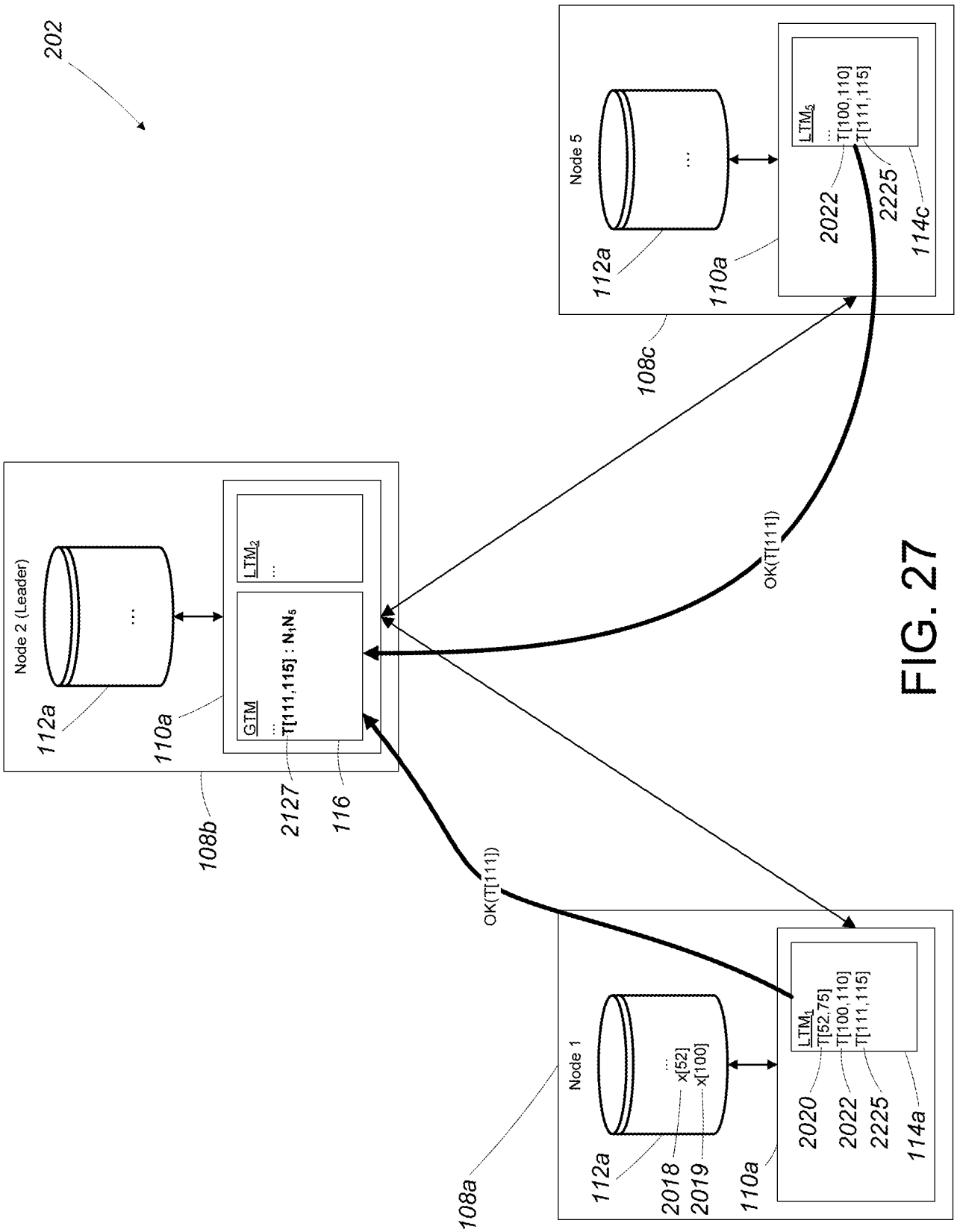
202
Node 2 (Leader)
112a
110a
GTM
T[111,115] : N1N5
LTM2
108b
2127
116
Node 5
112a
110a
LTM2
T[100,110]
T[111,115]
2022
2225
114c
108c
OK(T[111])
OK(T[111])
Node 1
112a
2018
2019
x[52]
x[100]
110a
LTM1
T[52,75]
T[100,110]
T[111,115]
2020
2022
2225
114a
108a

FIG. 27

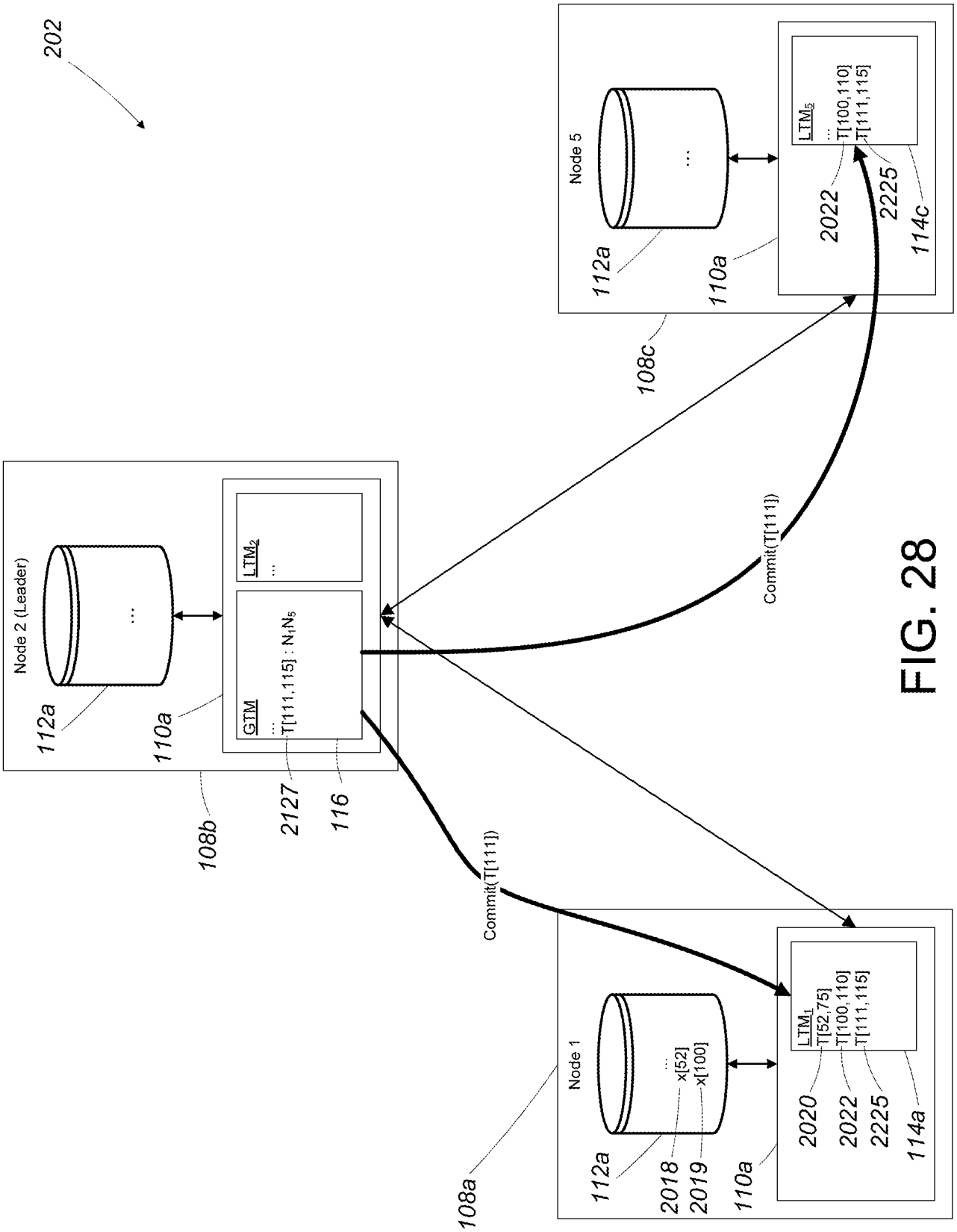
202
Node 2 (Leader)
112a
110a
GTM
T[111,115] : N1N5
LTM2
108b
2127
116
Node 5
112a
110a
LTM5
T[100,110]
T[111,115]
2022
2225
114c
108c
Commit(T[111])
Commit(T[111])
Node 1
112a
2018
2019
x[52]
x[100]
110a
LTM1
T[52,75]
T[100,110]
T[111,115]
2020
2022
2225
114a
108a

FIG. 28

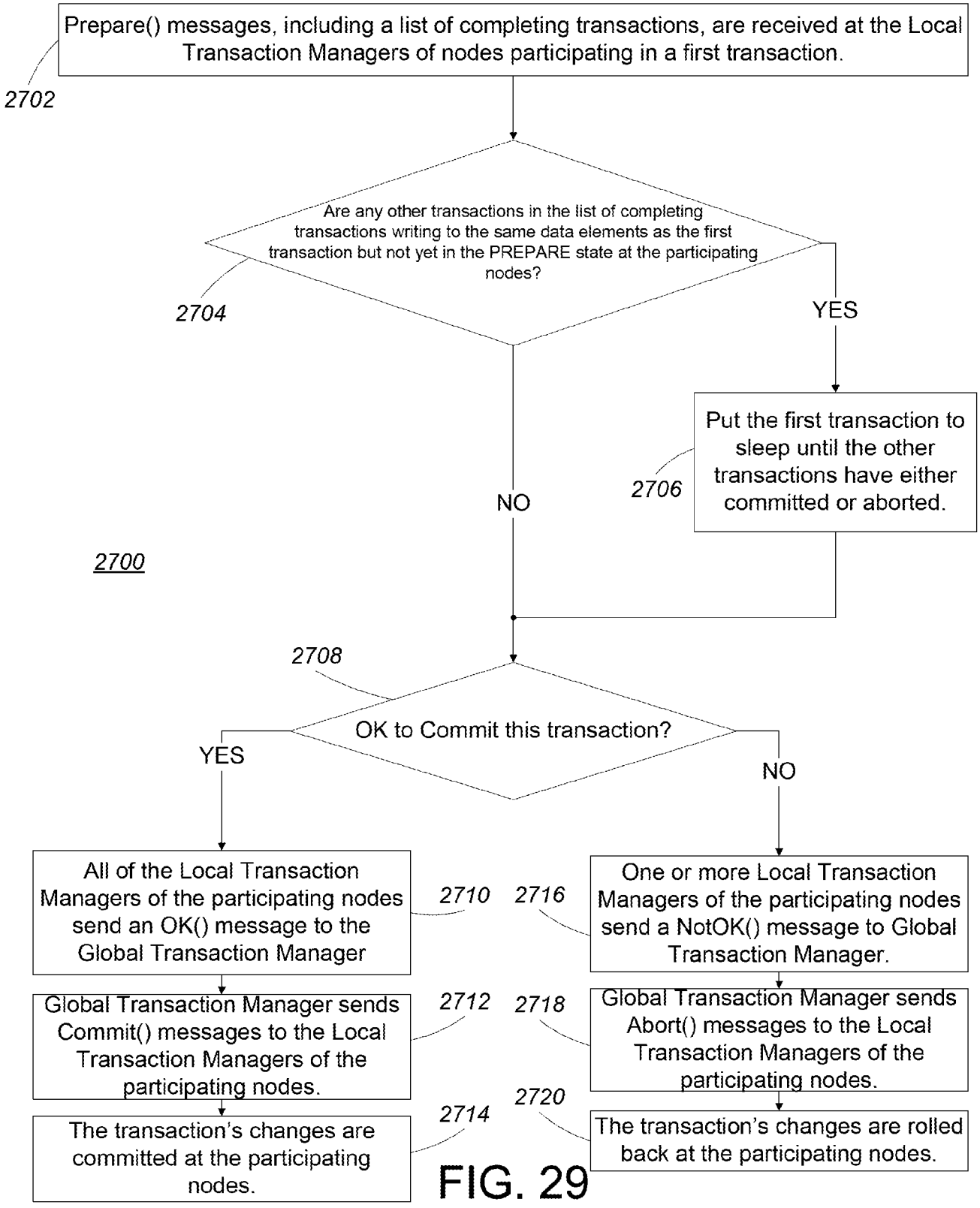
Prepare() messages, including a list of completing transactions, are received at the Local Transaction Managers of nodes participating in a first transaction.
2702
Are any other transactions in the list of completing transactions writing to the same data elements as the first transaction but not yet in the PREPARE state at the participating nodes?
2704
YES
Put the first transaction to sleep until the other transactions have either committed or aborted.
2706
NO
2700
2708
OK to Commit this transaction?
YES
NO
All of the Local Transaction Managers of the participating nodes send an OK() message to the Global Transaction Manager
2710
Global Transaction Manager sends Commit() messages to the Local Transaction Managers of the participating nodes.
2712
The transaction's changes are committed at the participating nodes.
2714
One or more Local Transaction Managers of the participating nodes send a NotOK() message to Global Transaction Manager.
2716
Global Transaction Manager sends Abort() messages to the Local Transaction Managers of the participating nodes.
2718
The transaction's changes are rolled back at the participating nodes.
2720

FIG. 29

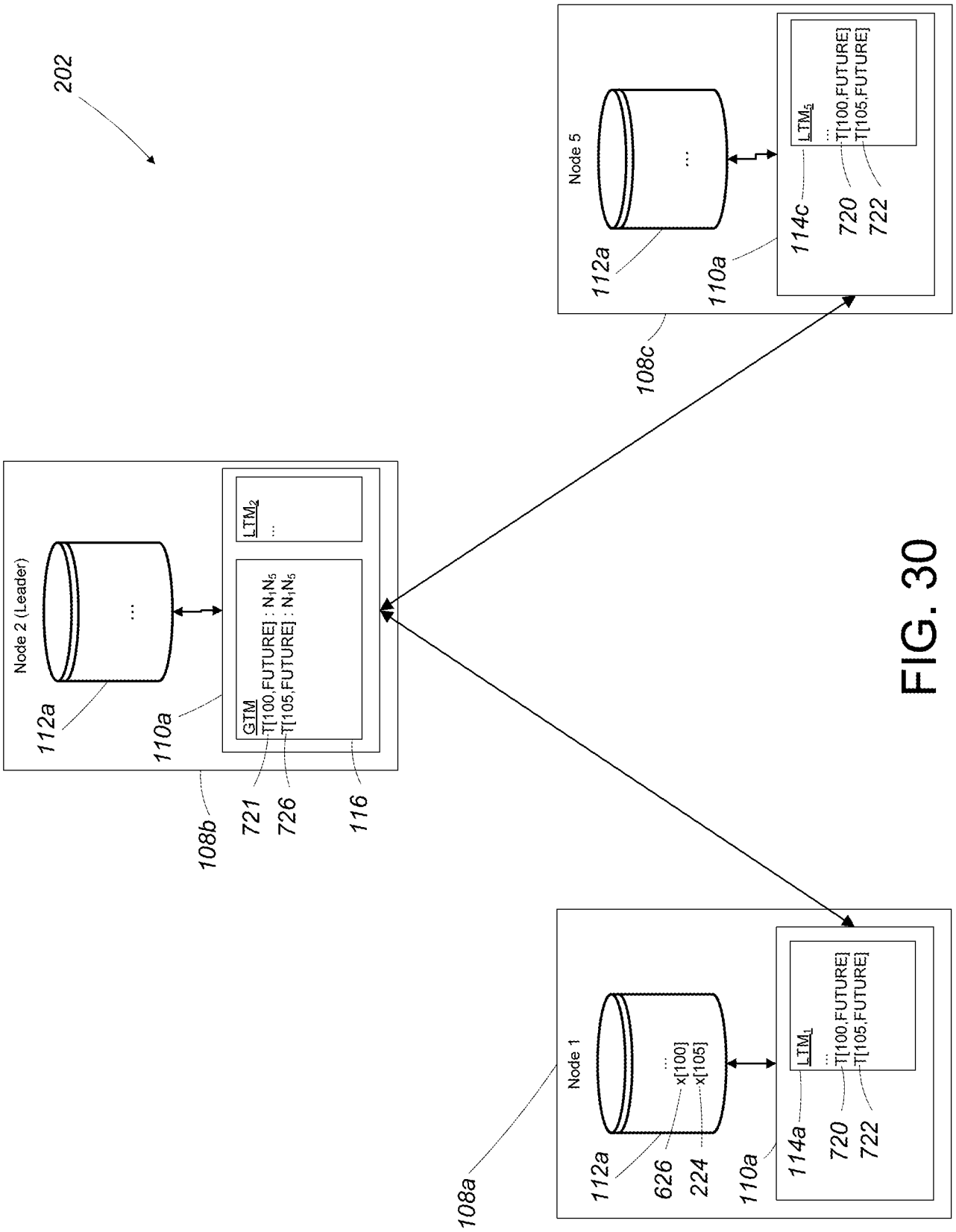
202
Node 2 (Leader)
112a
110a
GTM
T[100,FUTURE] : N1N5
T[105,FUTURE] : N1N5
LTM2
108b
721
726
116
Node 5
112a
110a
LTM5
T[100,FUTURE]
T[105,FUTURE]
114c
720
722
108c
Node 1
112a
x[100]
x[105]
626
224
110a
LTM1
T[100,FUTURE]
T[105,FUTURE]
114a
720
722
108a

FIG. 30

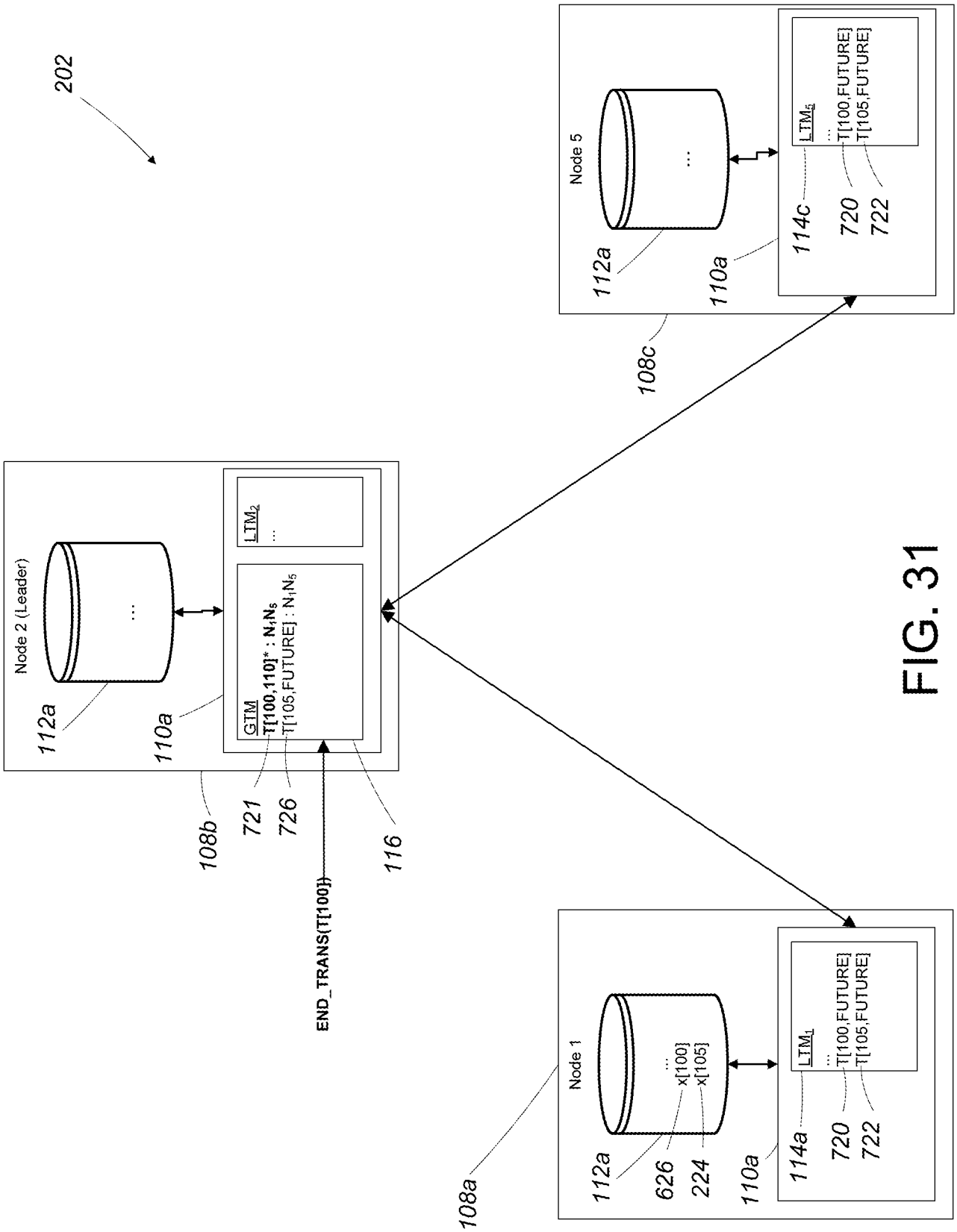
202
Node 2 (Leader)
112a
110a
GTM
T[100,110]* : N1N5
T[105,FUTURE] : N1N5
LTM2
108b
721
726
116
END_TRANS(T[100])
Node 5
112a
110a
LTM5
T[100,FUTURE]
T[105,FUTURE]
114c
720
722
108c
Node 1
112a
626
224
110a
x[100]
x[105]
LTM1
T[100,FUTURE]
T[105,FUTURE]
114a
720
722
108a

FIG. 31

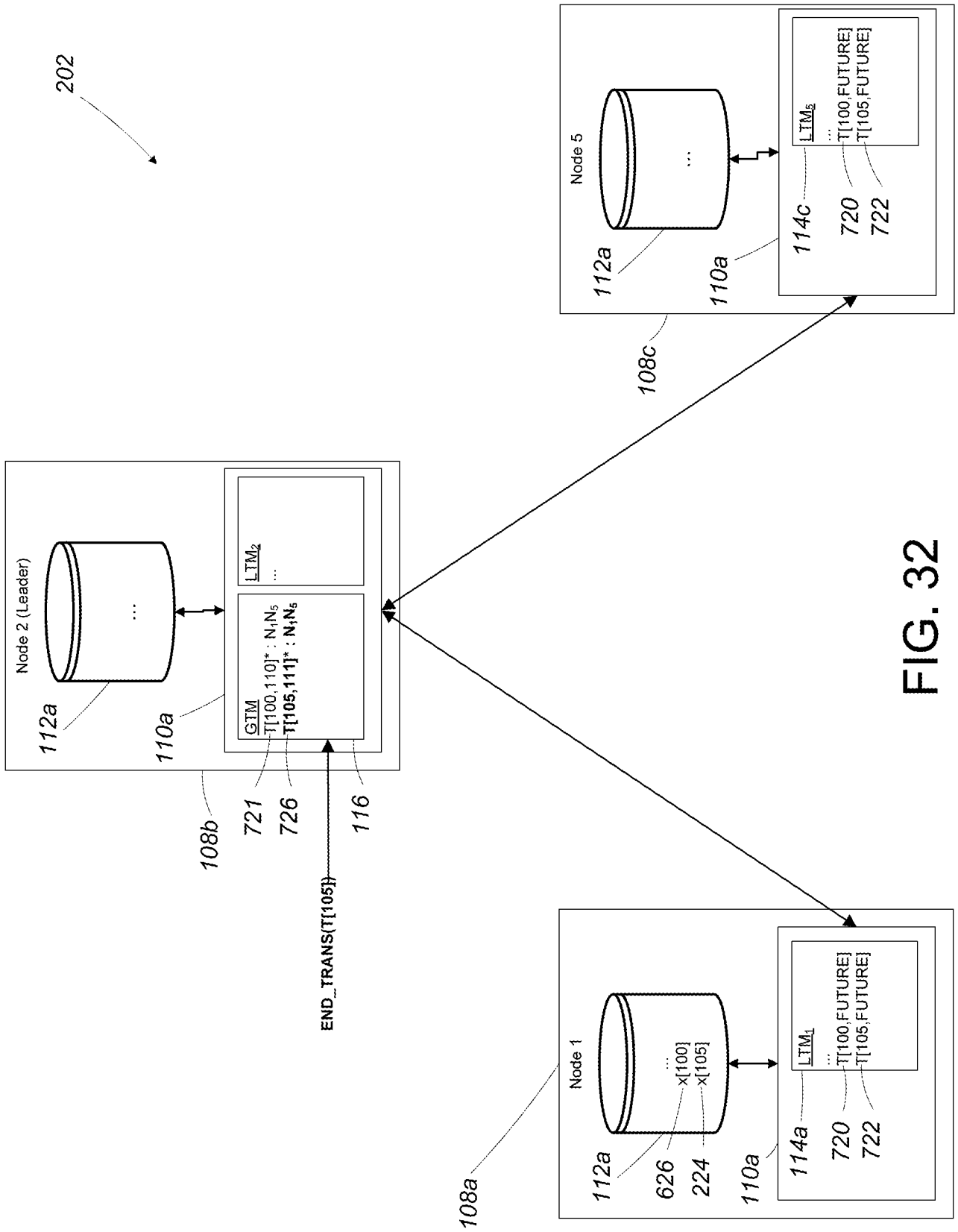
202
Node 2 (Leader)
112a
110a
GTM
T[100,110]* : N1N5
T[105,111]* : N1N5
LTM2
108b
721
726
116
END_TRANS(T[105])
Node 5
112a
110a
LTM3
T[100,FUTURE]
T[105,FUTURE]
114c
720
722
108c
Node 1
112a
x[100]
x[105]
626
224
110a
LTM1
T[100,FUTURE]
T[105,FUTURE]
114a
720
722
108a

FIG. 32

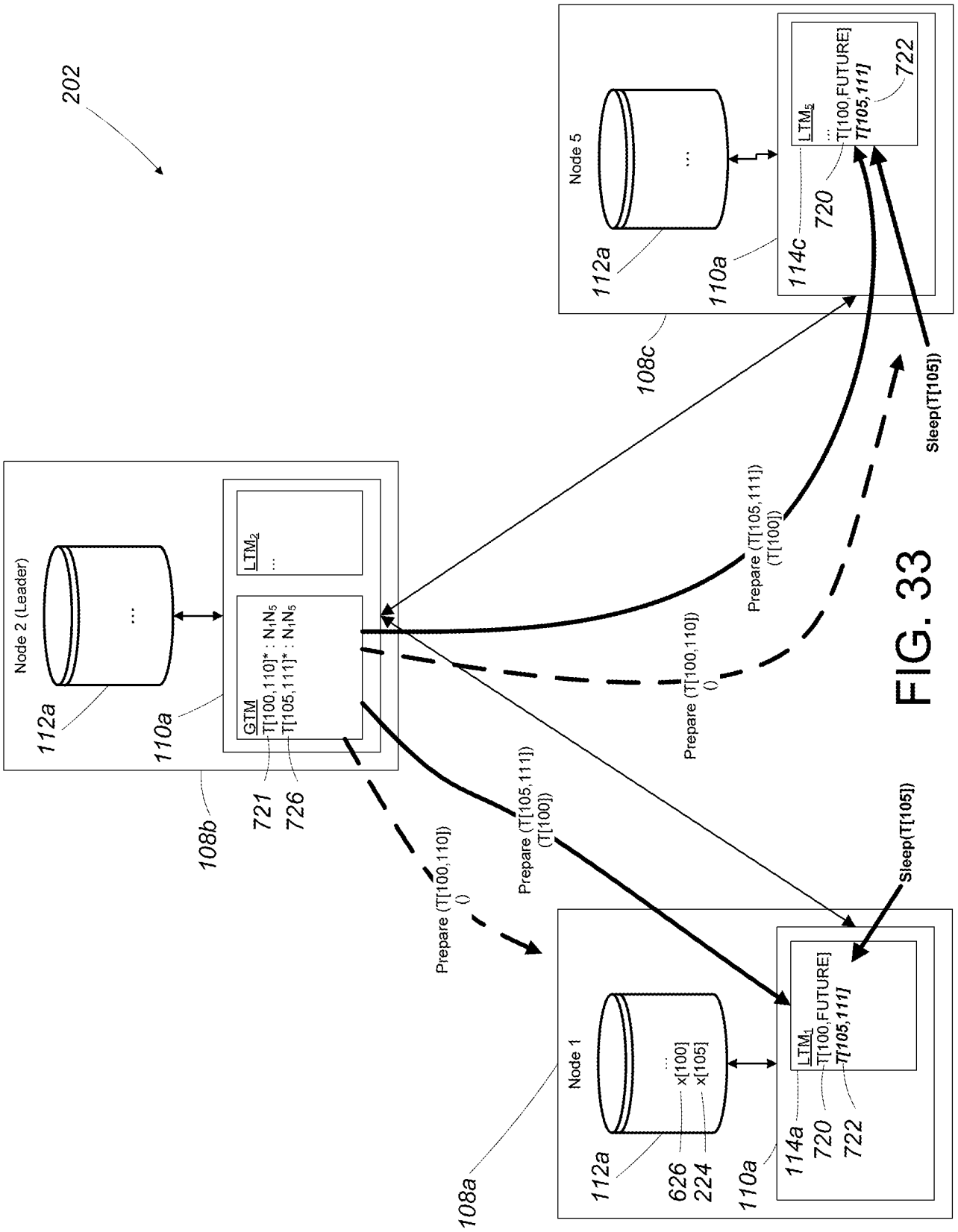
202
Node 2 (Leader)
112a
110a
GTM
T[100,110]* : N1N5
T[105,111]* : N1N5
LTM2
108b
721
726
Node 5
112a
110a
114c
720
LTM5
T[100,FUTURE]
T[105,111]
722
108c
Node 1
112a
626
224
x[100]
x[105]
110a
114a
720
722
LTM1
T[100,FUTURE]
T[105,111]
108a
Prepare (T[100,110])
()
Prepare (T[105,111])
(T[100])
Prepare (T[100,110])
()
Prepare (T[105,111])
(T[100])
Sleep(T[105])
Sleep(T[105])

FIG. 33

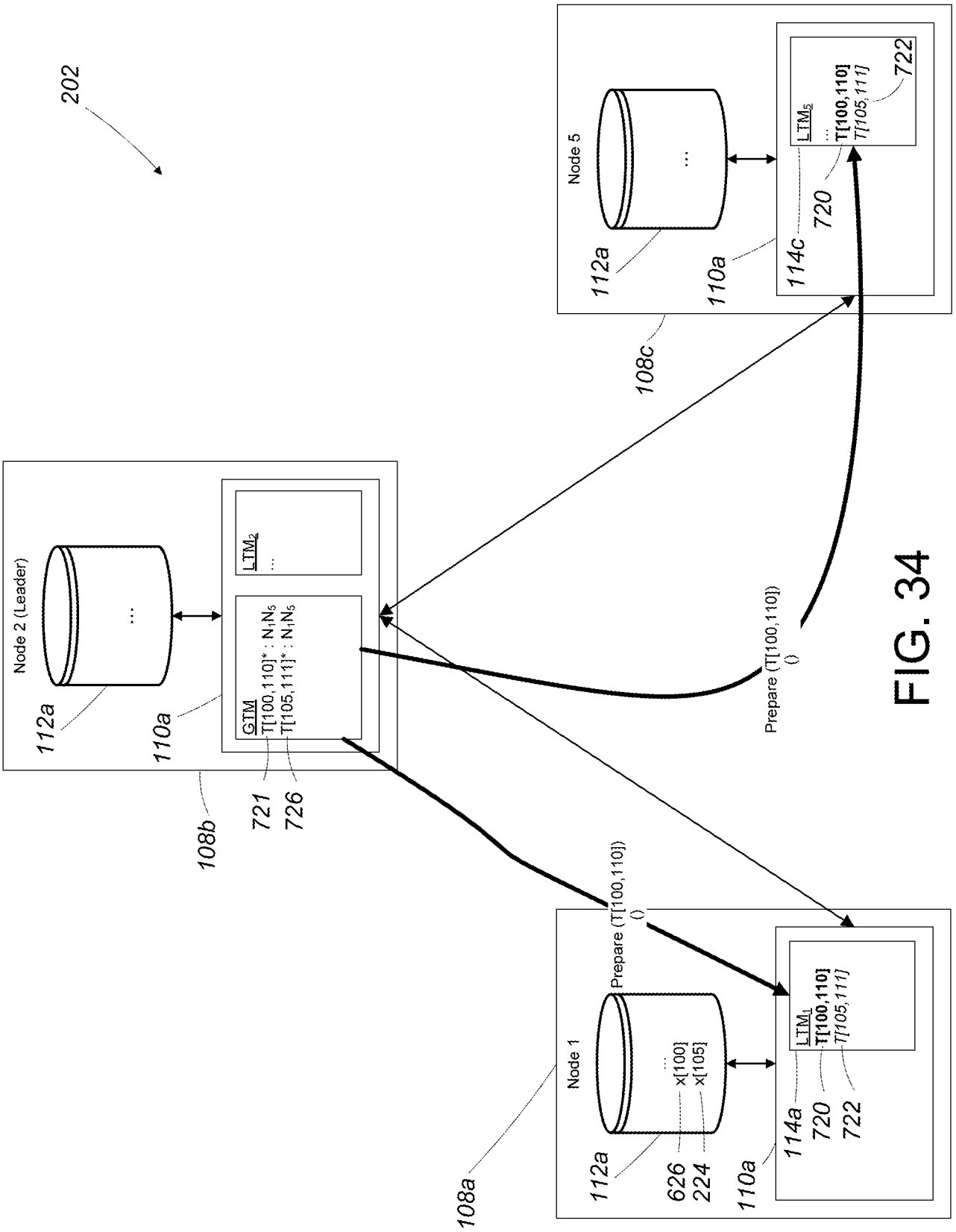
202
Node 2 (Leader)
112a
110a
GTM
T[100,110]* : N1N5
T[105,111]* : N1N5
LTM2
108b
721
726
Prepare (T[100,110])
()
Node 5
112a
110a
114c
LTM5
T[100,110]
T[105,111]
720
722
108c
Prepare (T[100,110])
()
Node 1
112a
626
224
x[100]
x[105]
110a
114a
LTM1
T[100,110]
T[105,111]
720
722
108a

FIG. 34

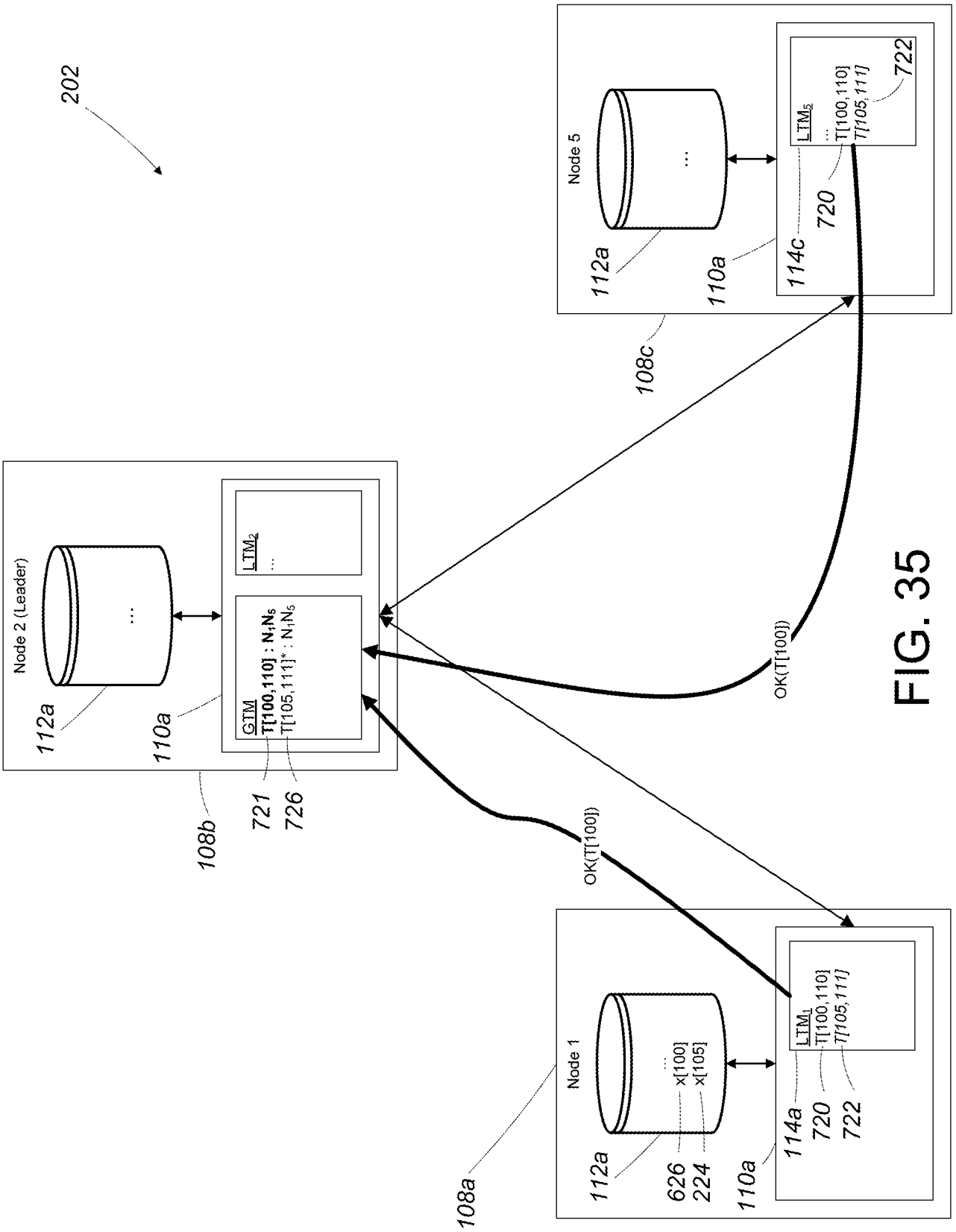
202
Node 2 (Leader)
112a
110a
GTM
T[100,110] : N1N5
T[105,111]* : N1N5
LTM2
108b
721
726
Node 5
112a
110a
114c
LTM5
T[100,110]
T[105,111]
720
722
108c
OK(T[100])
OK(T[100])
Node 1
112a
626
224
x[100]
x[105]
110a
114a
LTM1
T[100,110]
T[105,111]
720
722
108a

FIG. 35

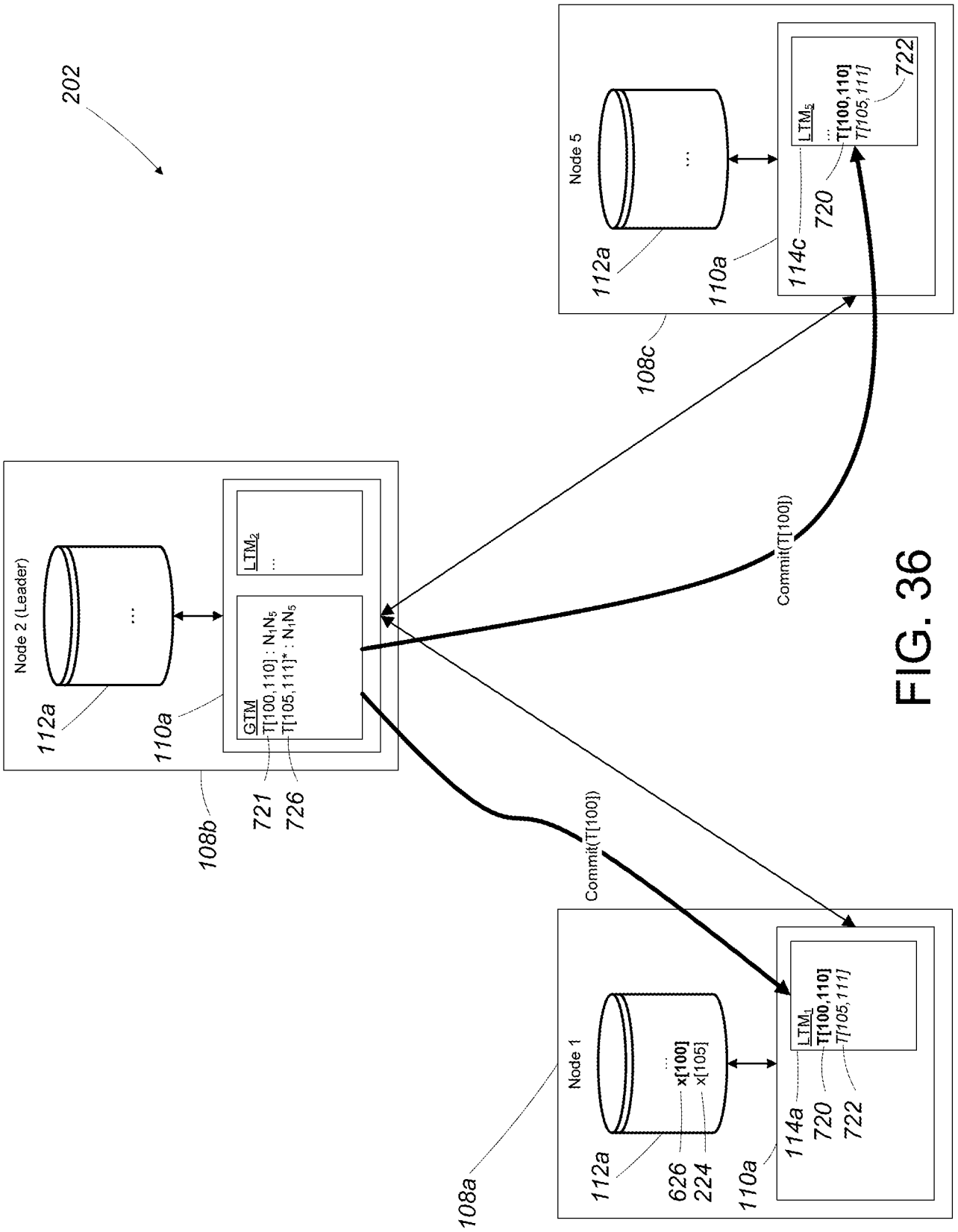
202
Node 2 (Leader)
112a
110a
GTM
T[100,110] : N1N5
T[105,111]* : N1N5
LTM2
108b
721
726
Node 5
112a
110a
114c
720
LTM5
T[100,110]
T[105,111]
722
108c
Commit(T[100])
Commit(T[100])
Node 1
112a
626
224
x[100]
x[105]
110a
114a
720
722
LTM1
T[100,110]
T[105,111]
108a

FIG. 36

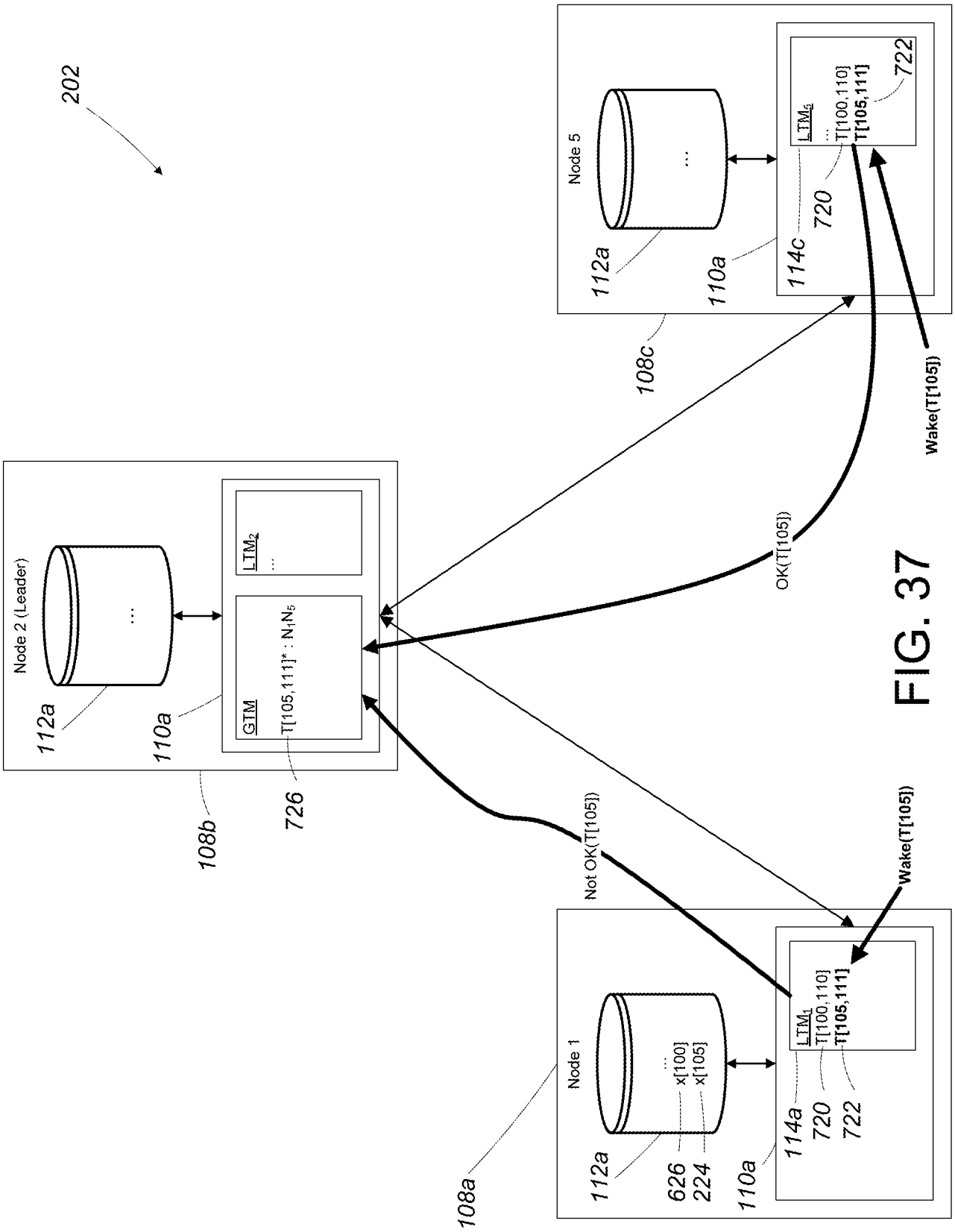
202
Node 2 (Leader)
112a
110a
GTM
T[105,111]* : N1N5
LTM2
...
726
108b
Node 5
112a
110a
114c
LTM5
T[100,110]
T[105,111]
720
722
108c
Node 1
112a
626
224
x[100]
x[105]
110a
114a
LTM1
T[100,110]
T[105,111]
720
722
108a
Not OK(T[105])
OK(T[105])
Wake(T[105])
Wake(T[105])

FIG. 37

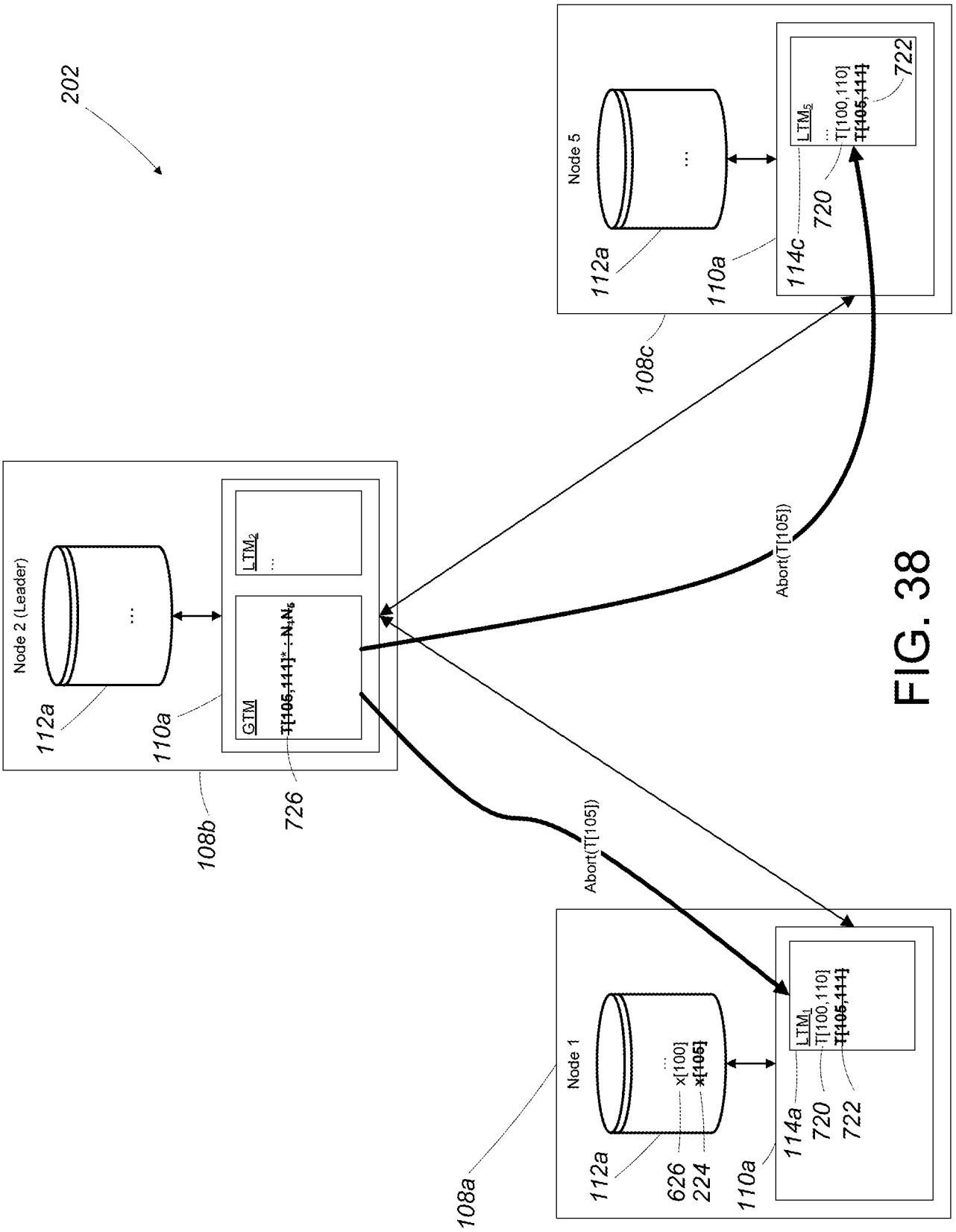
202
Node 2 (Leader)
112a
110a
GTM
LTM2
T[105,111]* : N1,N6
726
108b
Node 5
112a
110a
114c
LTM5
T[100,110]
T[105,111]
720
722
108c
Abort(T[105])
Abort(T[105])
Node 1
112a
626
224
x[100]
x[105]
110a
114a
LTM1
T[100,110]
T[105,111]
720
722
108a

FIG. 38

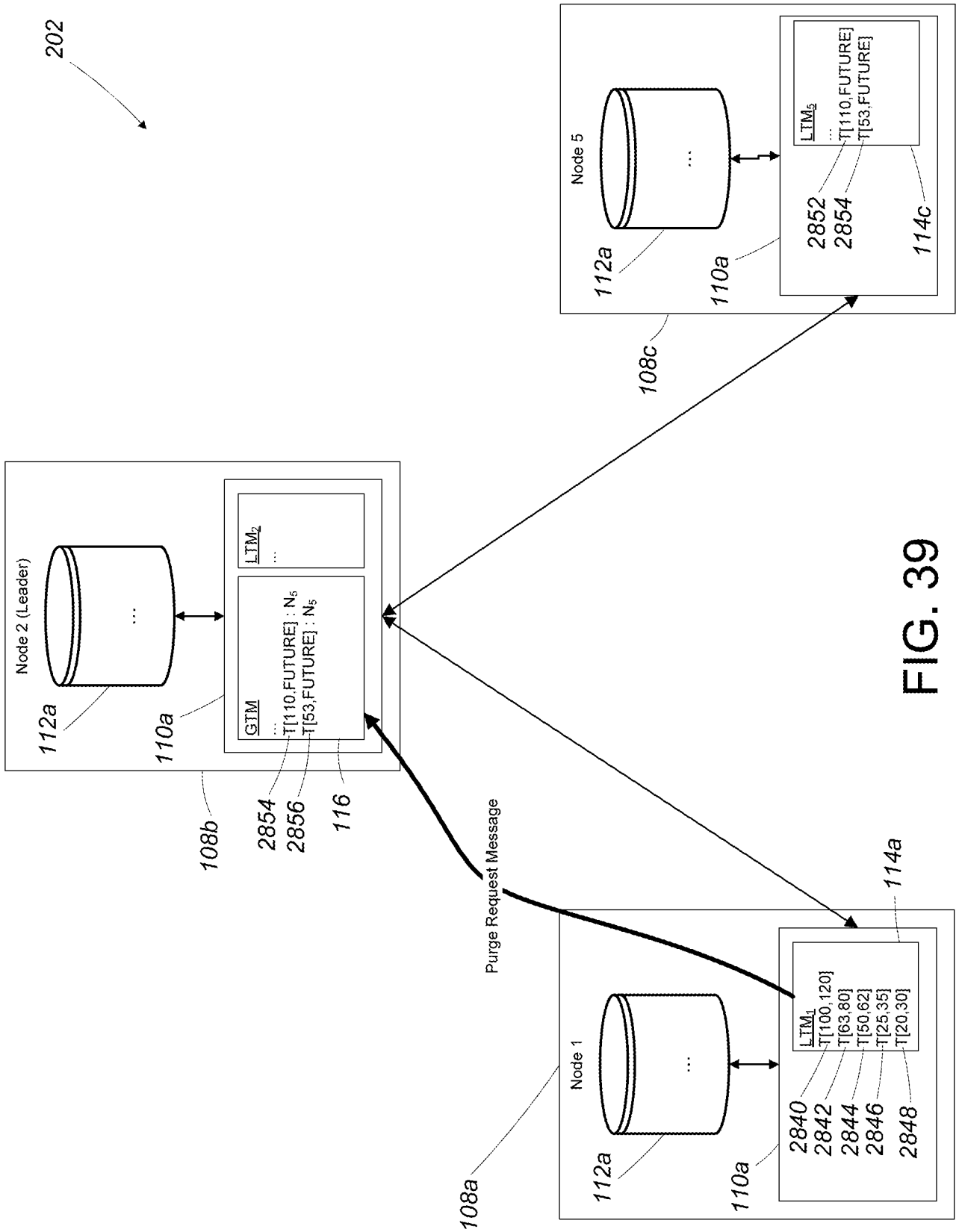
202
Node 2 (Leader)
112a
110a
GTM
...
T[110,FUTURE] : N5
T[53,FUTURE] : N5
LTM2
...
108b
2854
2856
116
Purge Request Message
Node 5
112a
110a
LTM5
...
T[110,FUTURE]
T[53,FUTURE]
2852
2854
114c
108c
Node 1
112a
110a
LTM1
T[100,120]
T[63,80]
T[50,62]
T[25,35]
T[20,30]
2840
2842
2844
2846
2848
114a
108a

FIG. 39

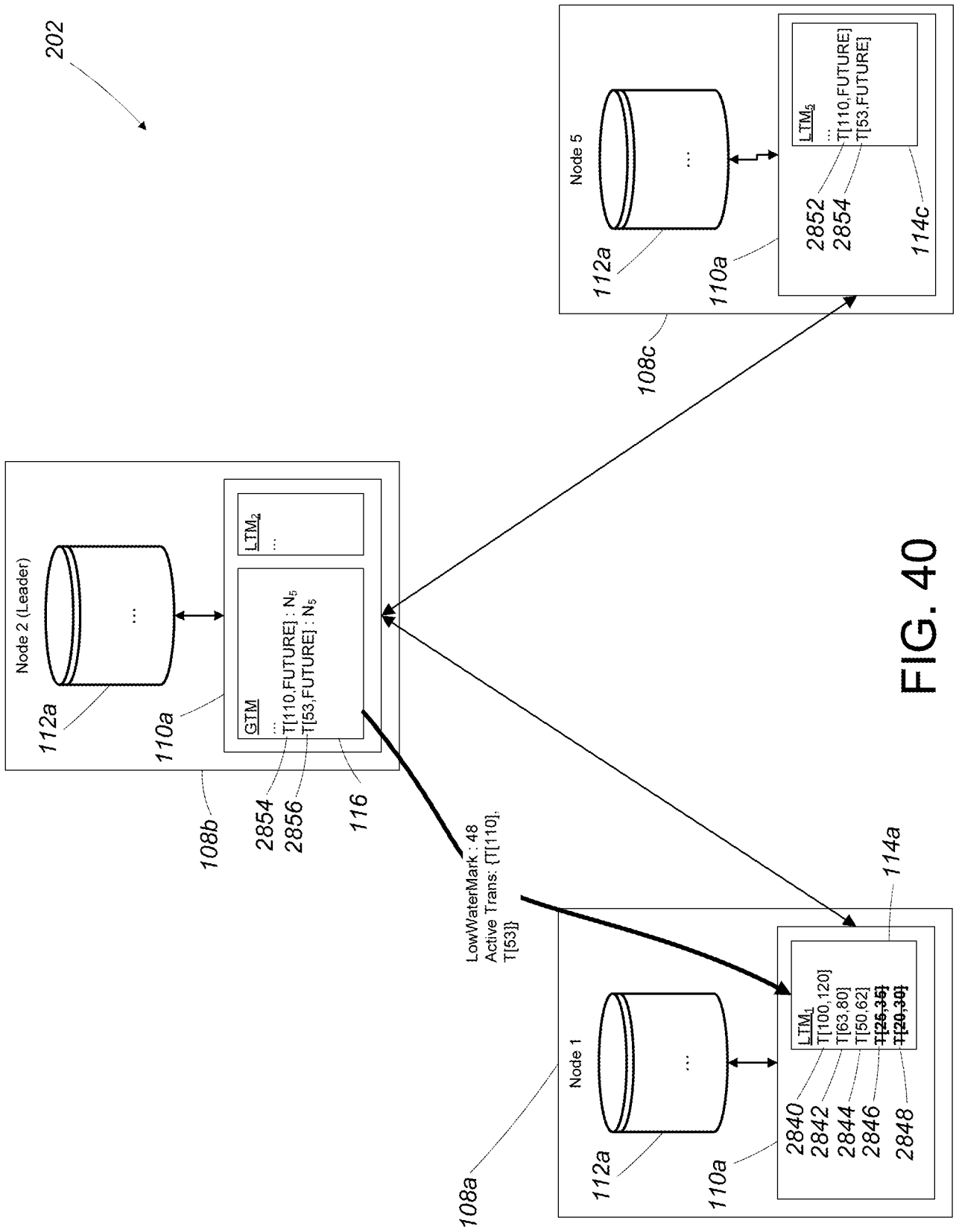
202
Node 2 (Leader)
112a
110a
GTM
T[110,FUTURE] : N5
T[53,FUTURE] : N5
LTM2
108b
2854
2856
116
Node 5
112a
110a
LTM5
T[110,FUTURE]
T[53,FUTURE]
2852
2854
114c
108c
LowWaterMark : 48
Active Trans: {T[110], T[53]}
Node 1
112a
110a
LTM1
T[100,120]
T[63,80]
T[50,62]
T[25,35]
T[20,30]
2840
2842
2844
2846
2848
114a
108a

FIG. 40

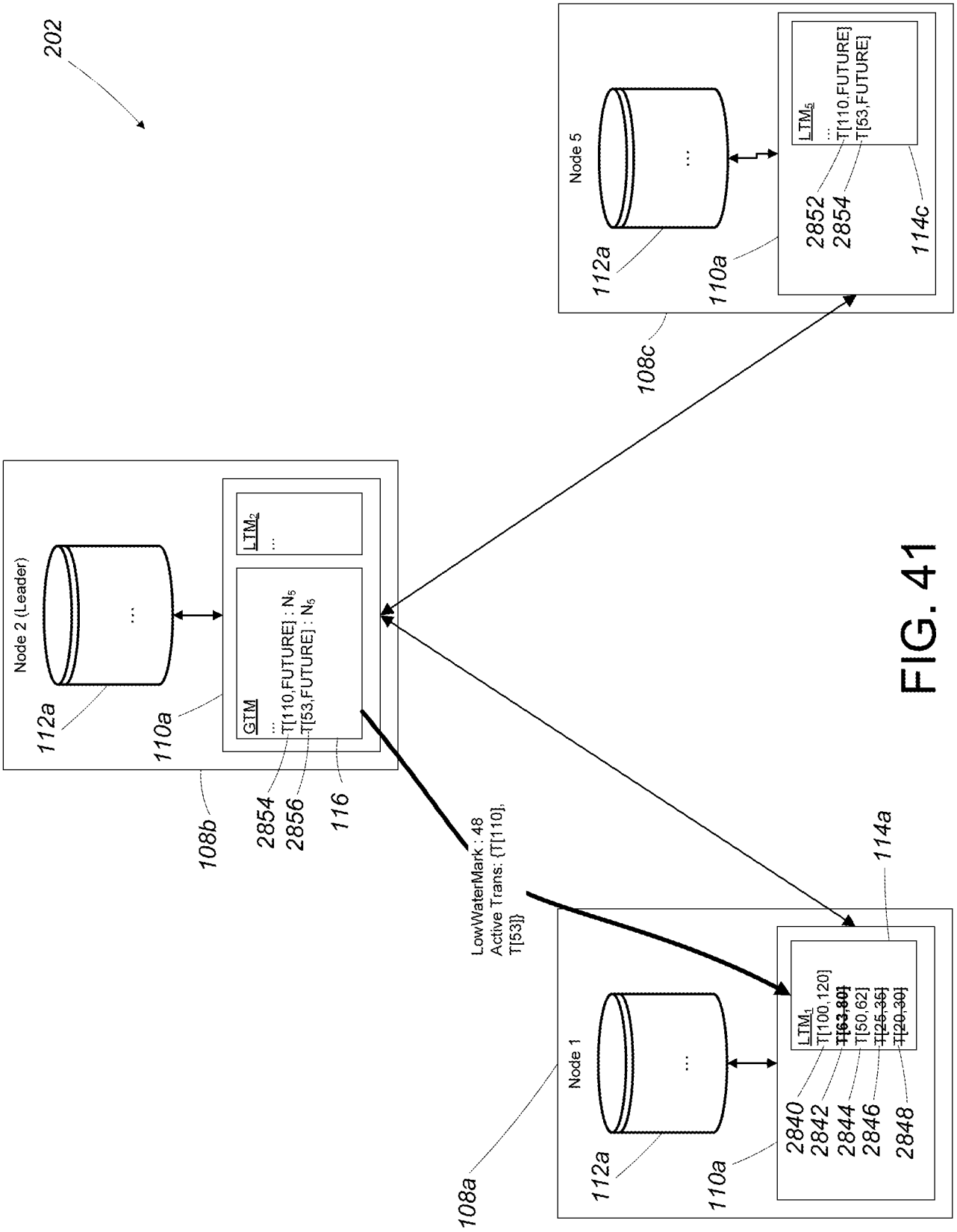
202
Node 2 (Leader)
112a
110a
GTM
...
T[110,FUTURE] : N5
T[53,FUTURE] : N6
LTM2
...
108b
2854
2856
116
LowWaterMark : 48
Active Trans: {T[110], T[53]}
Node 5
112a
110a
LTM5
...
T[110,FUTURE]
T[53,FUTURE]
2852
2854
114c
108c
Node 1
112a
110a
LTM1
T[100,120]
T[63,80]
T[50,62]
T[25,35]
T[20,30]
2840
2842
2844
2846
2848
114a
108a

FIG. 41