

1/14

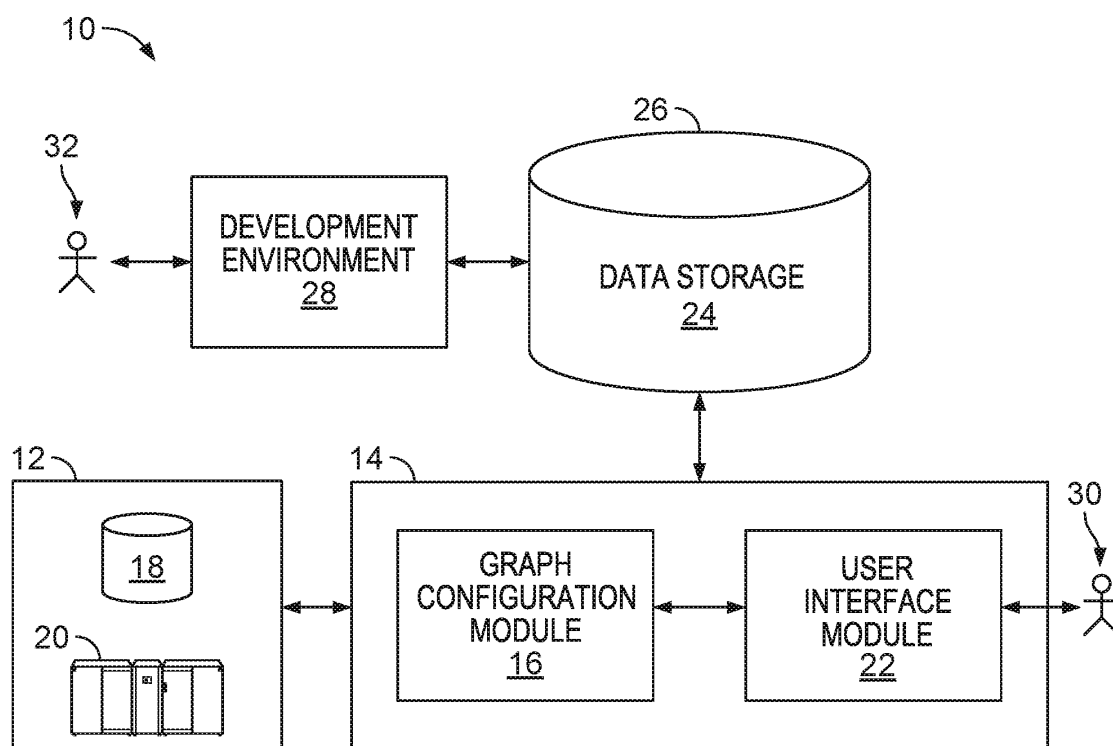


FIG. 1

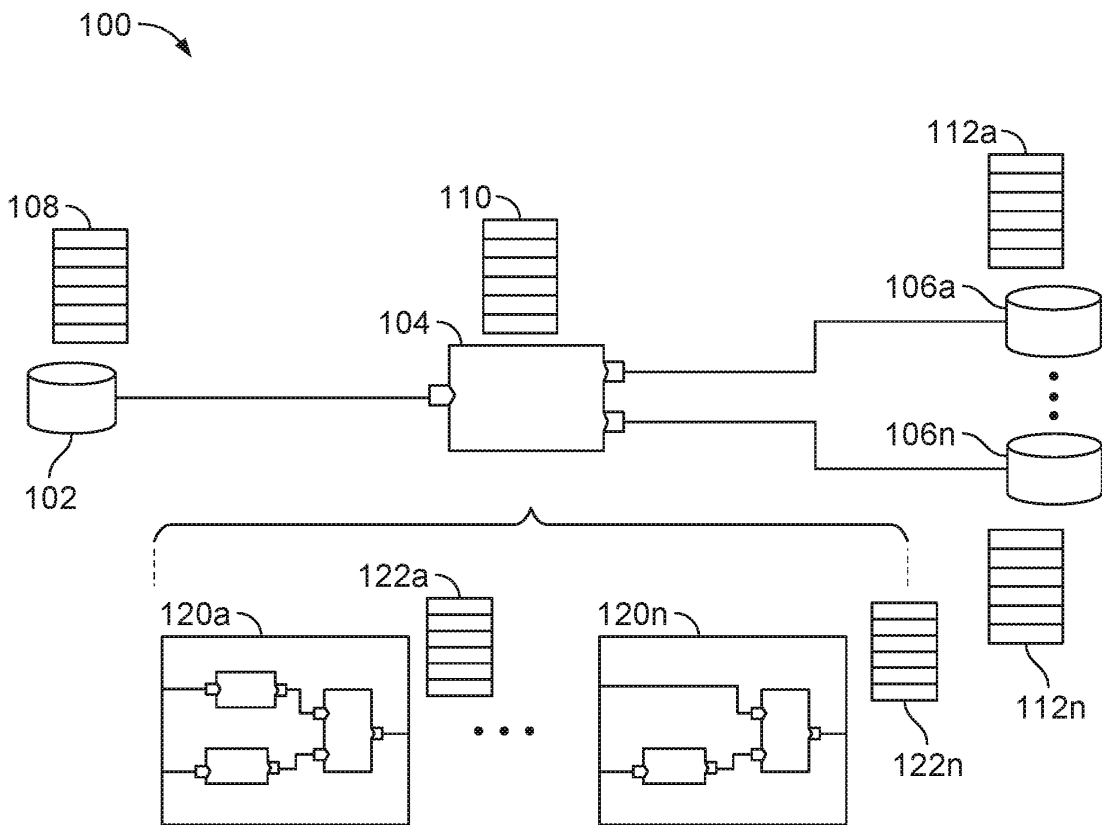


FIG. 2

300 

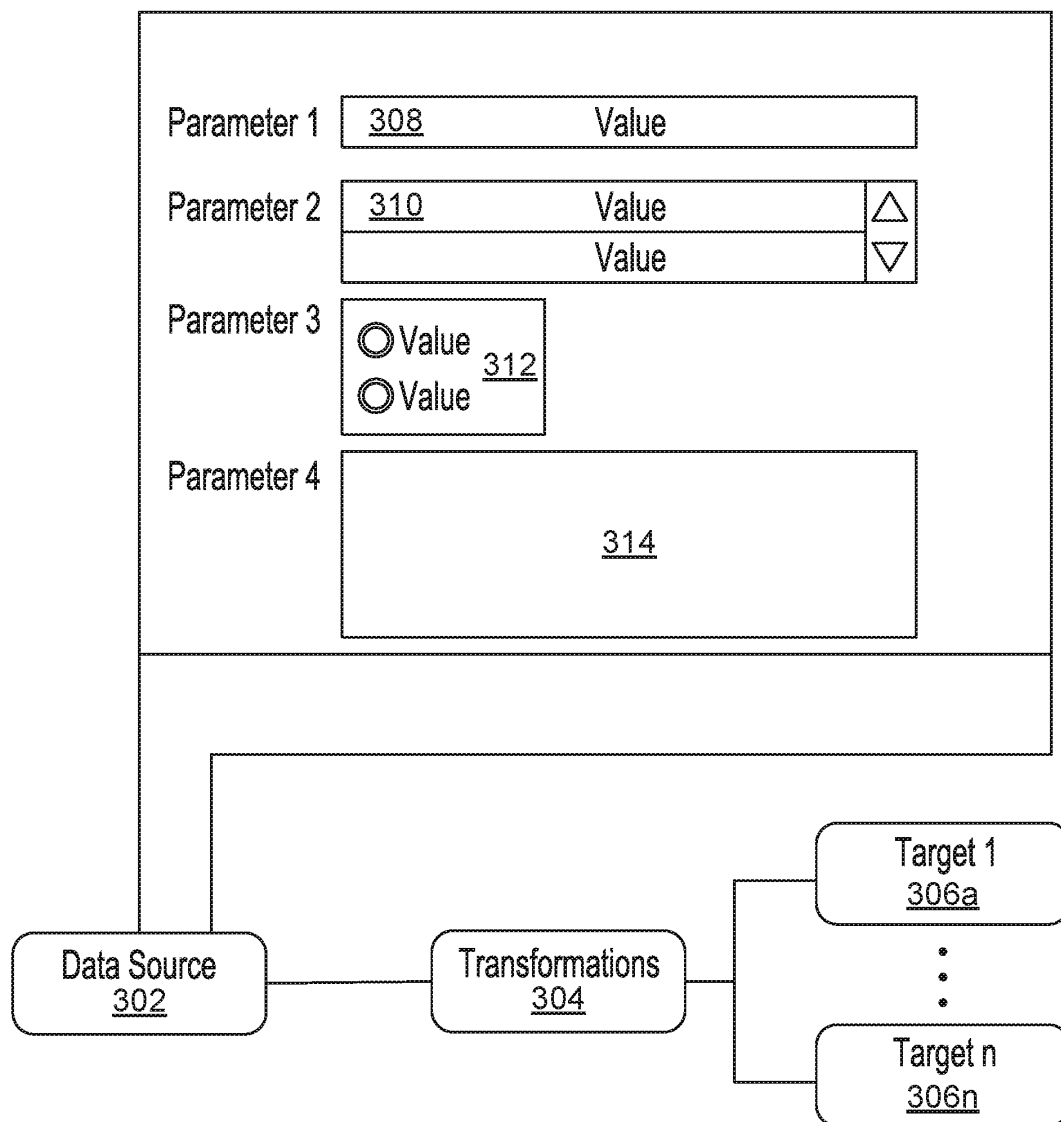


FIG. 3

4/14

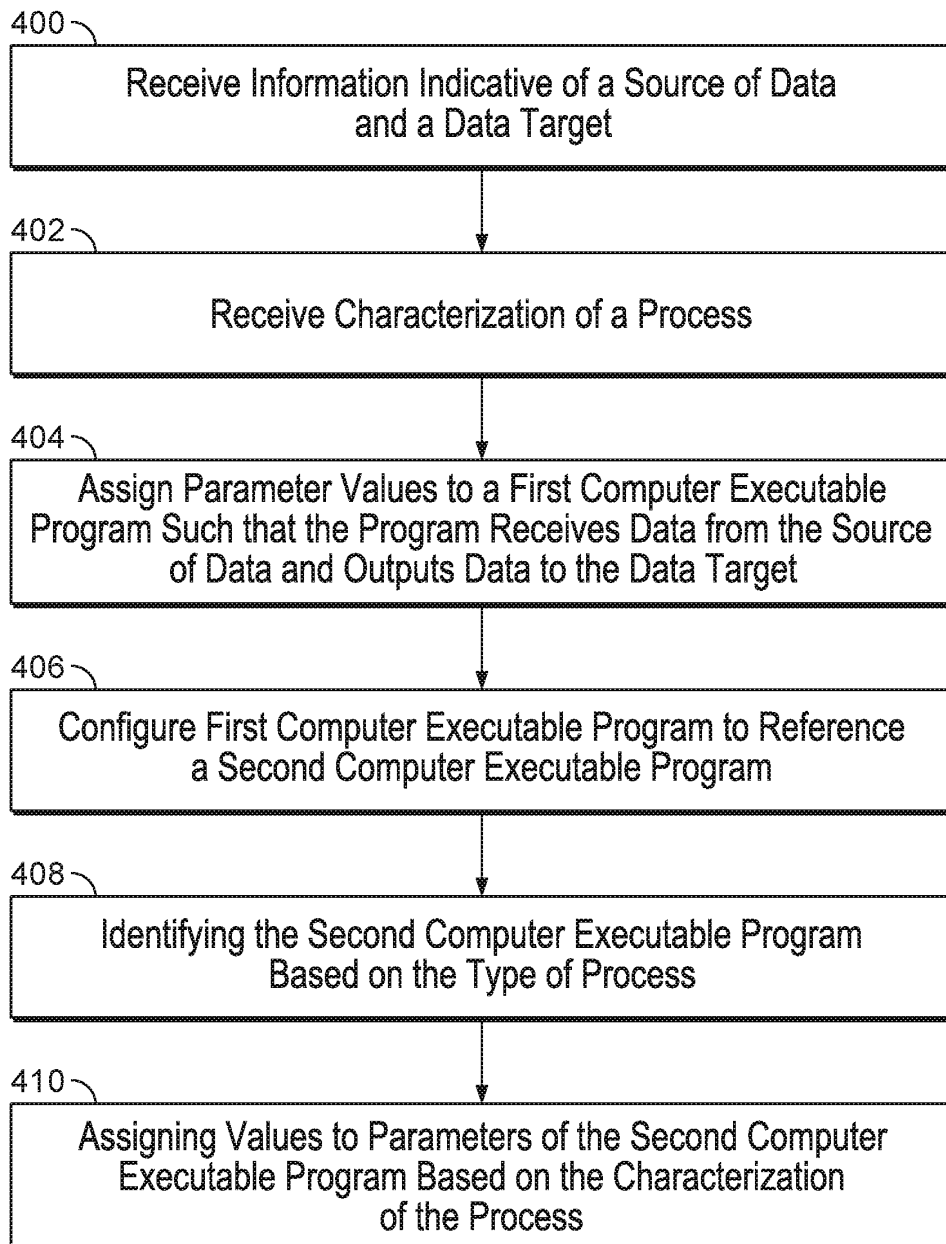


FIG. 4

5/14

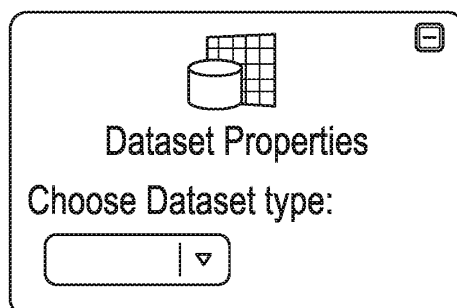


FIG. 5A

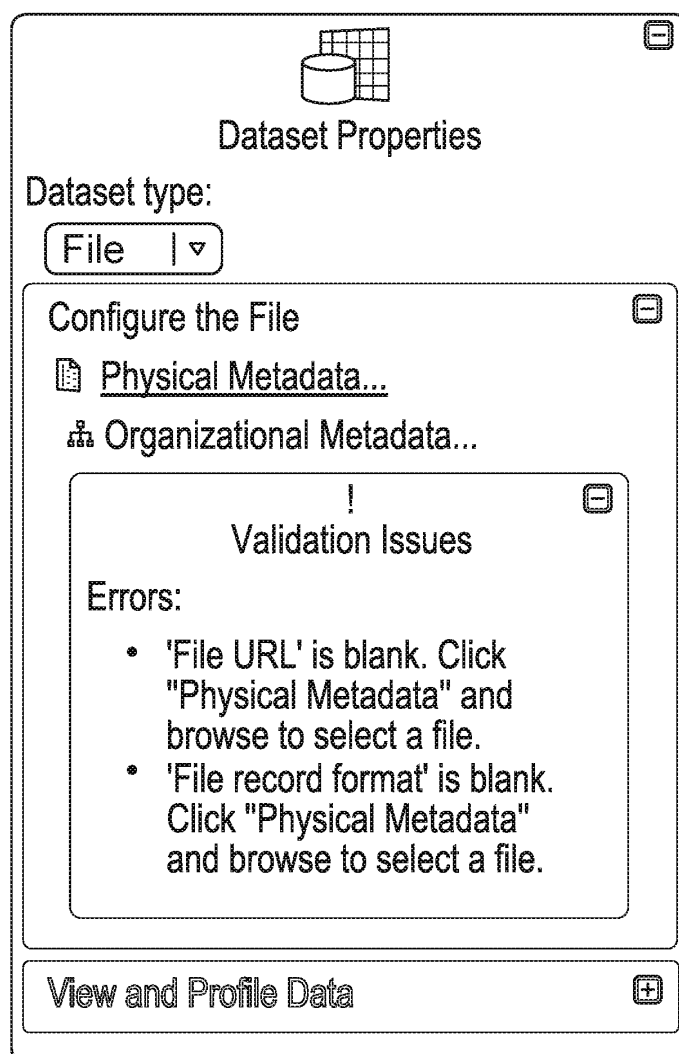
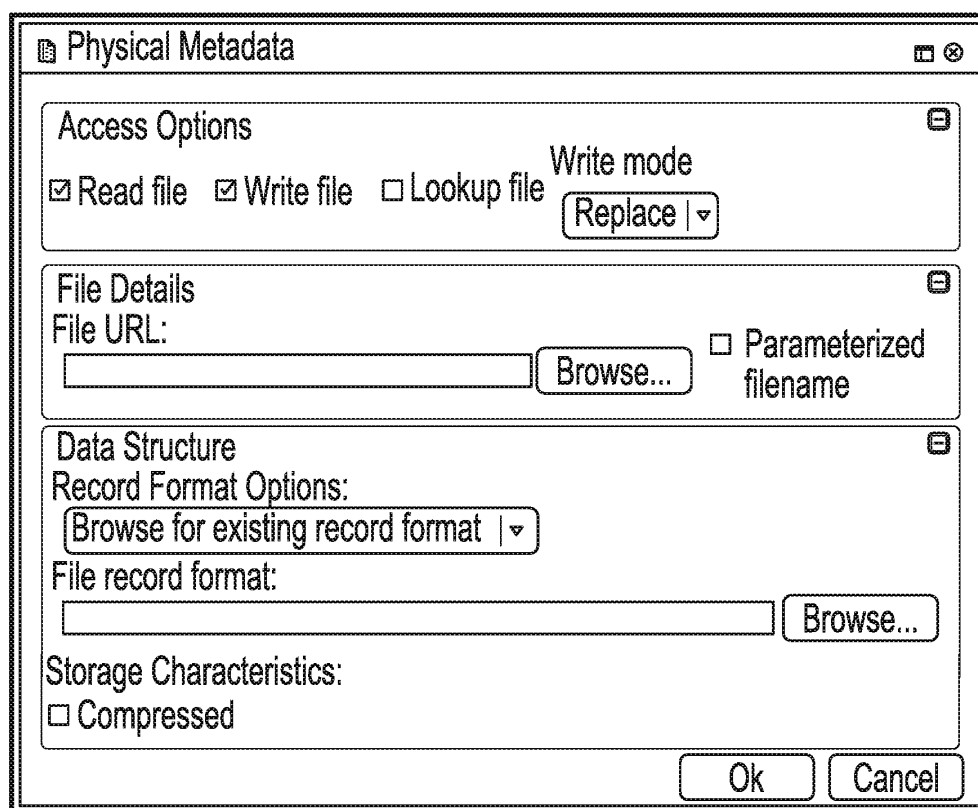


FIG. 5B

6/14



The 'Physical Metadata' dialog box contains three main sections. The 'Access Options' section has checkboxes for 'Read file', 'Write file', and 'Lookup file', along with a 'Write mode' dropdown set to 'Replace'. The 'File Details' section includes a 'File URL' text field with a 'Browse...' button and a 'Parameterized filename' checkbox. The 'Data Structure' section features a 'Record Format Options' dropdown set to 'Browse for existing record format', a 'File record format' text field with a 'Browse...' button, and a 'Storage Characteristics' section with a 'Compressed' checkbox. 'Ok' and 'Cancel' buttons are at the bottom right.

Physical Metadata

Access Options

☒ Read file ☒ Write file ☐ Lookup file Write mode

Replace

File Details

File URL: Browse... ☐ Parameterized filename

Data Structure

Record Format Options:

Browse for existing record format

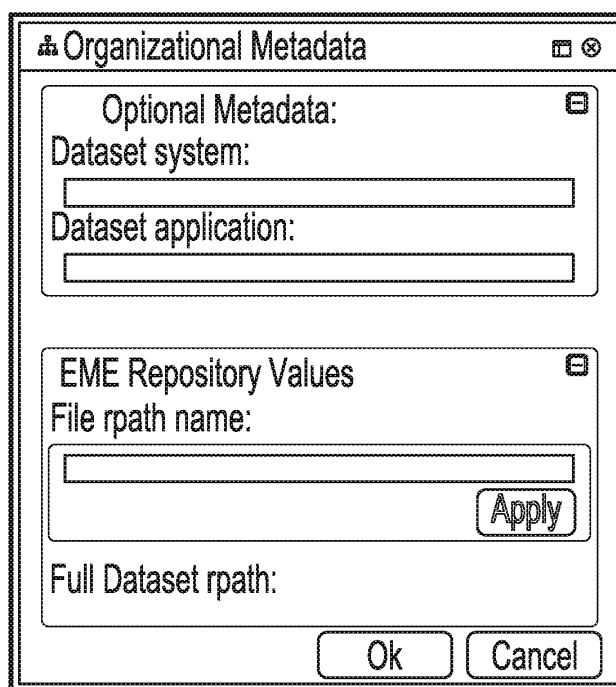
File record format: Browse...

Storage Characteristics:

☐ Compressed

Ok Cancel

FIG. 5C



The 'Organizational Metadata' dialog box contains two main sections. The 'Optional Metadata' section has text fields for 'Dataset system' and 'Dataset application'. The 'EME Repository Values' section has a 'File rpath name' text field with an 'Apply' button and a 'Full Dataset rpath' text field. 'Ok' and 'Cancel' buttons are at the bottom right.

Organizational Metadata

Optional Metadata:

Dataset system:

Dataset application:

EME Repository Values

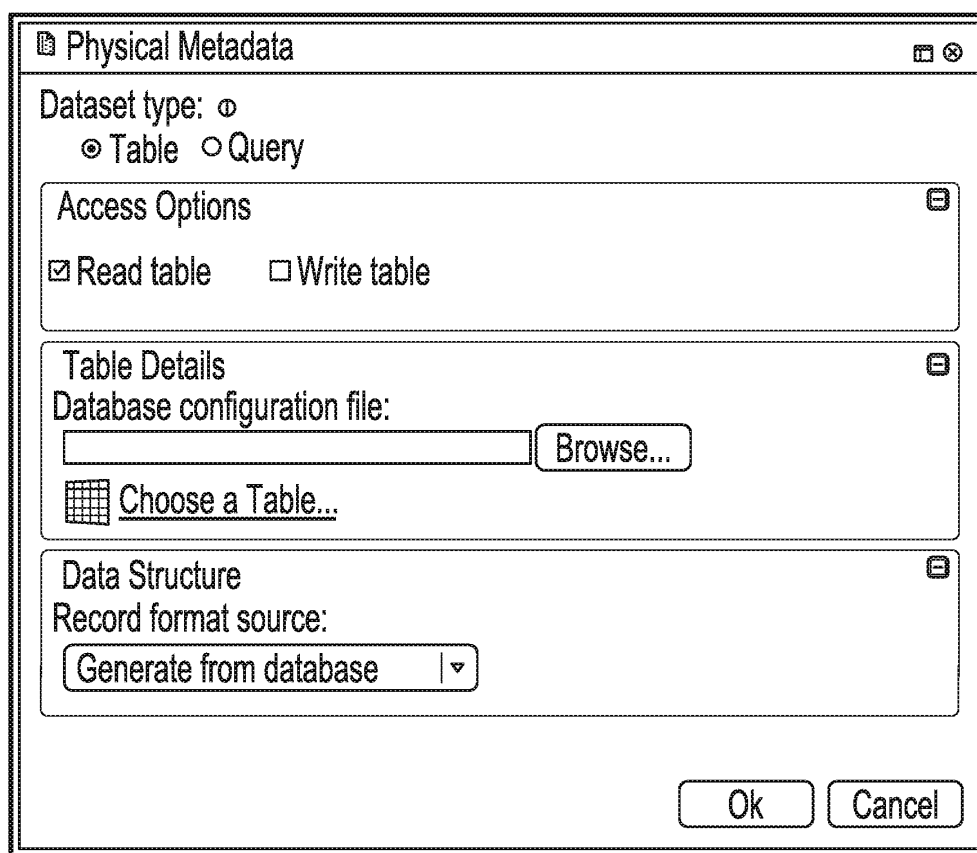
File rpath name: Apply

Full Dataset rpath:

Ok Cancel

FIG. 5D

7/14



The dialog box is titled "Physical Metadata" and contains three main sections: "Dataset type", "Access Options", "Table Details", and "Data Structure".

Dataset type: ☒ Table ☐ Query

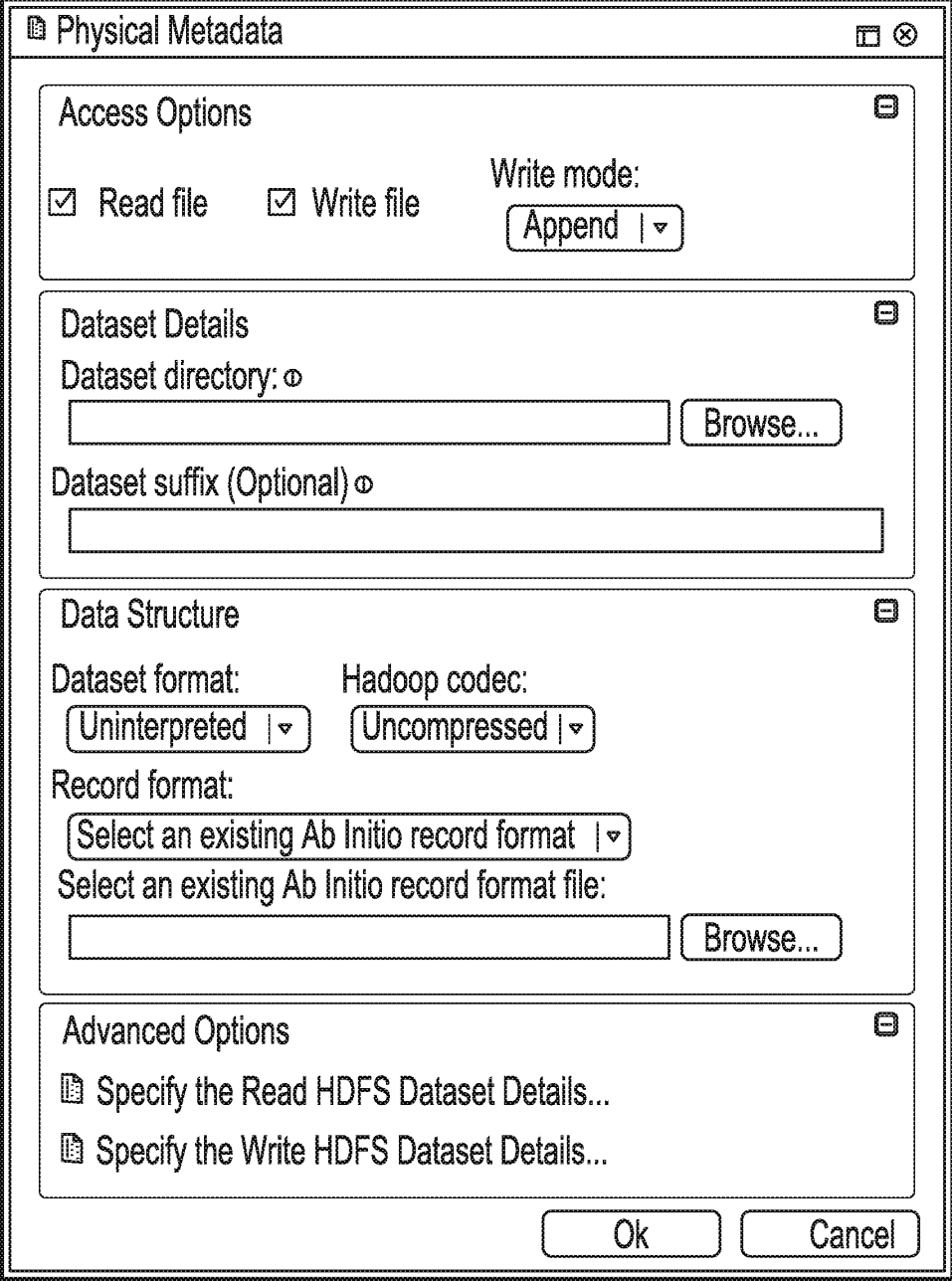
Access Options: ☒ Read table ☐ Write table

Table Details: Database configuration file:

Data Structure: Record format source:

FIG. 5E

8/14



The figure shows a 'Physical Metadata' dialog box with four main sections: Access Options, Dataset Details, Data Structure, and Advanced Options. The Access Options section has checkboxes for 'Read file' and 'Write file', and a 'Write mode' dropdown set to 'Append'. The Dataset Details section has a 'Dataset directory' field with a 'Browse...' button and a 'Dataset suffix (Optional)' field. The Data Structure section has dropdowns for 'Dataset format' (set to 'Uninterpreted') and 'Hadoop codec' (set to 'Uncompressed'), a 'Record format' dropdown set to 'Select an existing Ab Initio record format', and a 'Select an existing Ab Initio record format file' field with a 'Browse...' button. The Advanced Options section has two expandable items: 'Specify the Read HDFS Dataset Details...' and 'Specify the Write HDFS Dataset Details...'. At the bottom are 'Ok' and 'Cancel' buttons.

Physical Metadata

Access Options

☒ Read file ☒ Write file Write mode: Append | ▾

Dataset Details

Dataset directory: ○ Browse...

Dataset suffix (Optional) ○

Data Structure

Dataset format: Uninterpreted | ▾ Hadoop codec: Uncompressed | ▾

Record format: Select an existing Ab Initio record format | ▾

Select an existing Ab Initio record format file: Browse...

Advanced Options

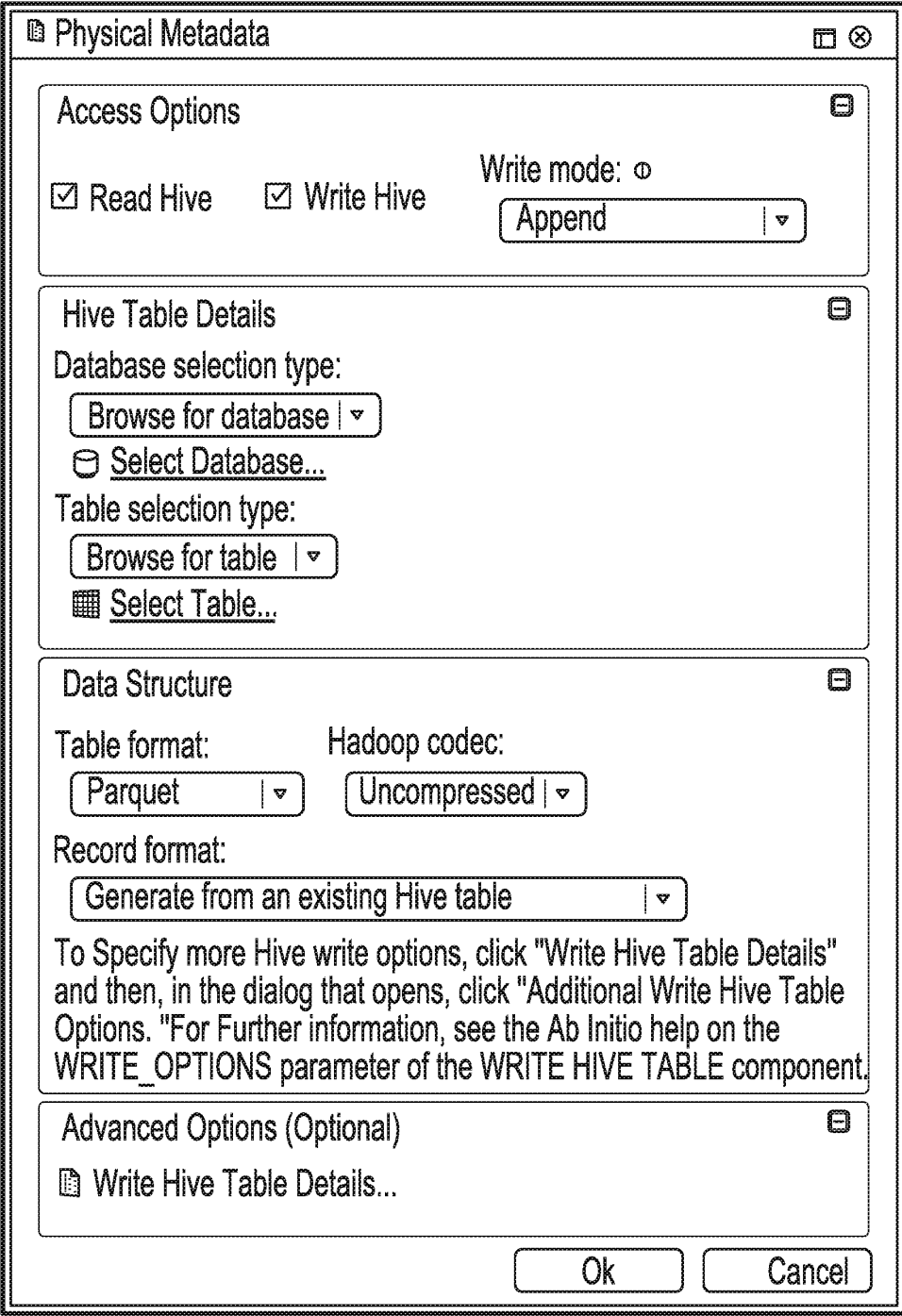
☐ Specify the Read HDFS Dataset Details...

☐ Specify the Write HDFS Dataset Details...

Ok Cancel

FIG. 5F

9/14



The image shows a software dialog box titled "Physical Metadata". It contains four main sections: "Access Options", "Hive Table Details", "Data Structure", and "Advanced Options (Optional)".

Access Options

- ☒ Read Hive
- ☒ Write Hive
- Write mode: $\oplus$
Append

Hive Table Details

Database selection type:

- Browse for database
- ☐ Select Database...

Table selection type:

- Browse for table
- ☒ Select Table...

Data Structure

Table format: Parquet Hadoop codec: Uncompressed

Record format: Generate from an existing Hive table

To Specify more Hive write options, click "Write Hive Table Details" and then, in the dialog that opens, click "Additional Write Hive Table Options." For Further information, see the Ab Initio help on the WRITE_OPTIONS parameter of the WRITE HIVE TABLE component.

Advanced Options (Optional)

- ☒ Write Hive Table Details...

At the bottom right are "Ok" and "Cancel" buttons.

FIG. 5G

10/14

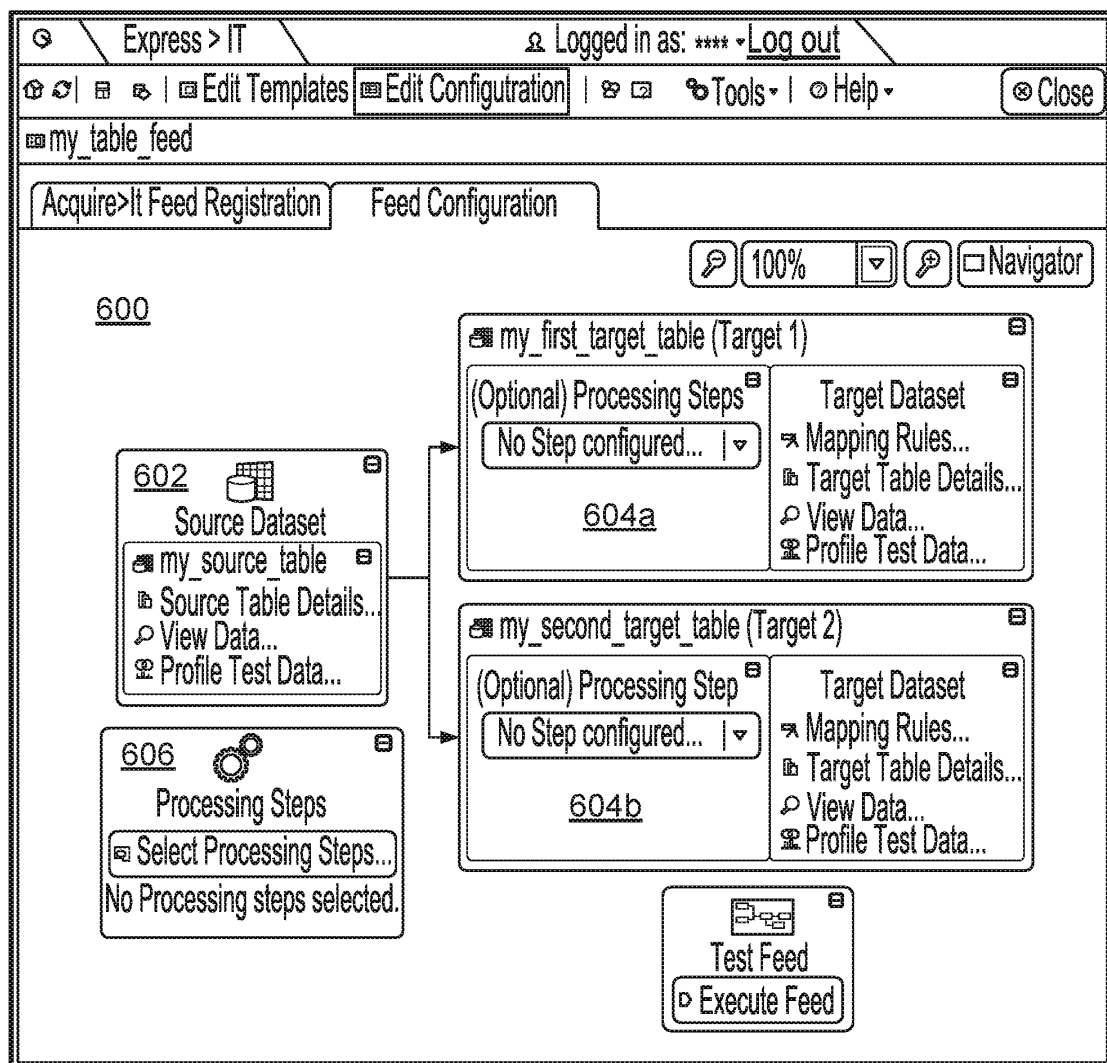


FIG. 6

11/14

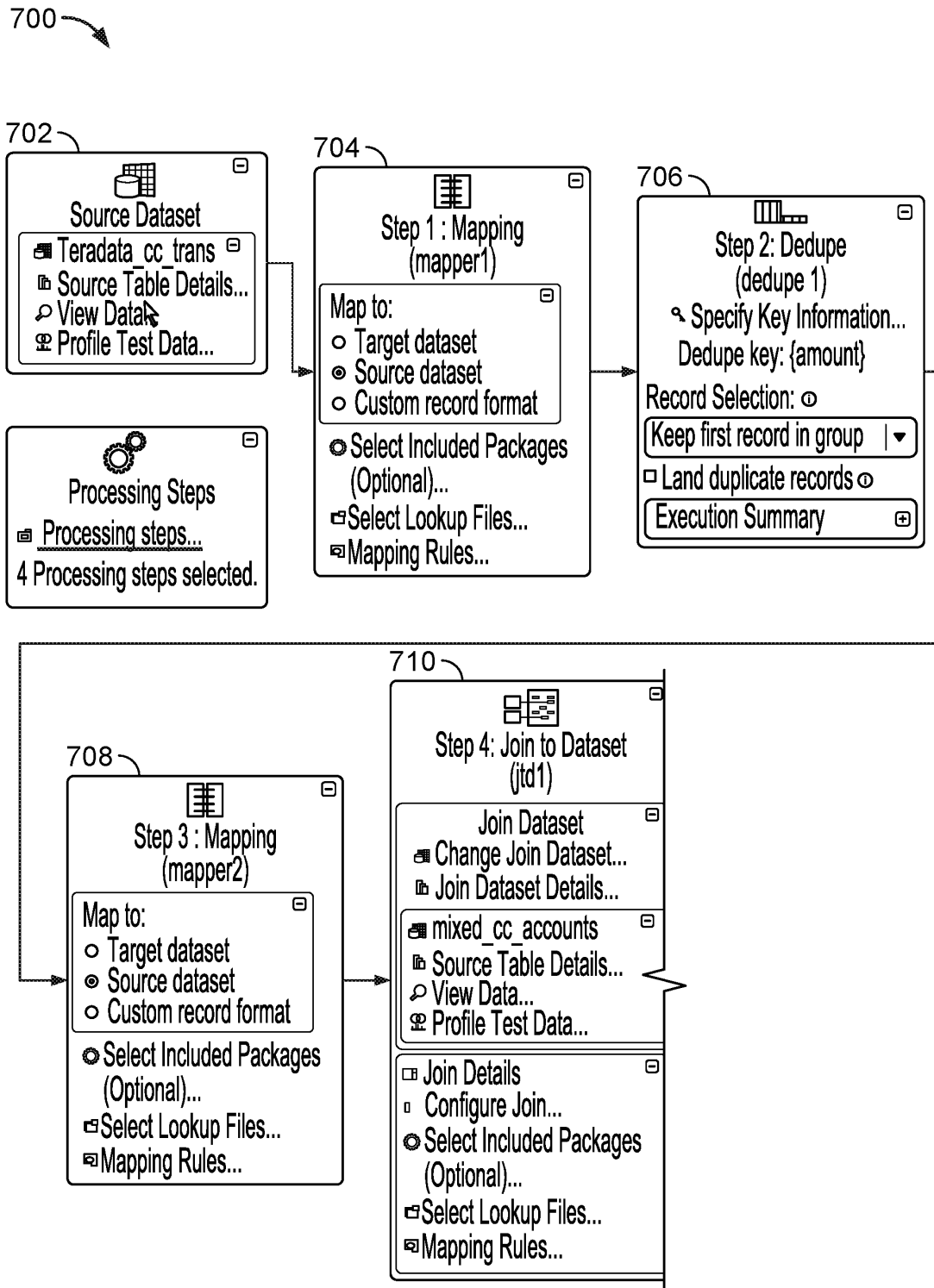


FIG. 7

12/14



Step 1: Dedupe
(Dedupe)



Dedupe Type:

Sorted (Default) | ▼



Specify Key Information...

Dedupe key: {startmarketingdate}

Record Selection: ⓘ

Keep first record in group | ▼

☐ Land duplicate records ⓘ

Execution Summary



Records Read: 115,911

Records Written: 7,720

Duplicate Records: 108,191

FIG. 8

13/14

The interface for 'my_first_target_file (Target 1)' is divided into two main panels. The left panel, titled '(Optional) Processing Step', contains a dropdown menu set to 'Mapping'. Below this is a 'Map to:' section with three radio button options: 'Target dataset' (selected), 'Source dataset', and 'Custom record format'. Underneath is a 'Select Target Dataset:' dropdown menu showing 'my_first_target_file'. At the bottom of this panel are two links: 'Select Included Packages (Optional)...' and 'Mapping Rules...'. The right panel, titled 'Target Dataset', contains four links: 'Mapping Rules...', 'Target File Details...', 'View Data...', and 'Profile Test Data...'.

FIG. 9A

The interface for 'my_second_target_file (Target 2)' is divided into two main panels. The left panel, titled '(Optional) Processing Step', contains a dropdown menu set to 'Filter'. Below this are two links: 'Select Included Packages (Optional)...' and 'Define Filter...'. The right panel, titled 'Target Dataset', contains four links: 'Mapping Rules...', 'Target File Details...', 'View Data...', and 'Profile Test Data...'.

FIG. 9B

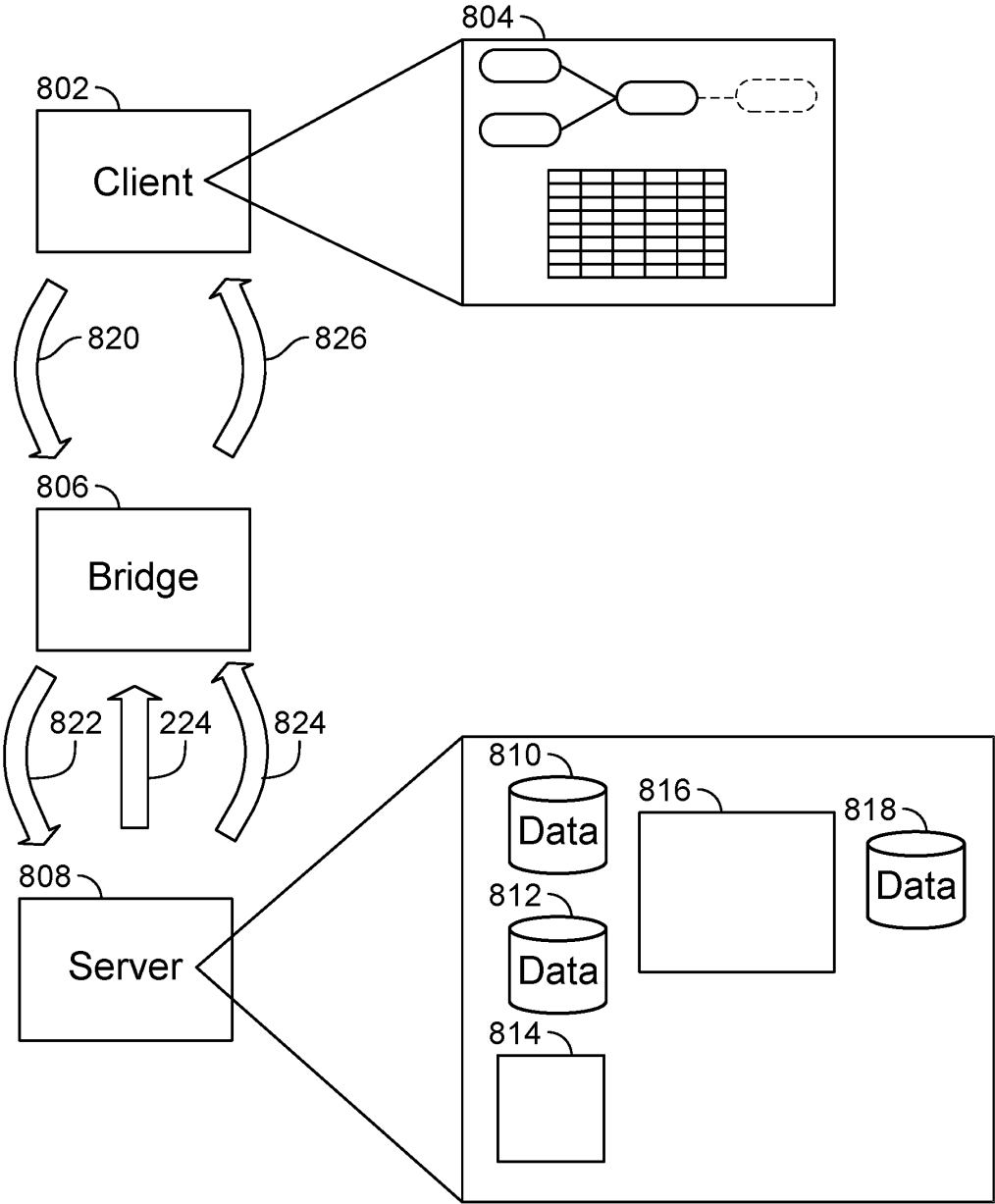


FIG. 10